Visual Aesthetics and Human Preference

Stephen E. Palmer, Karen B. Schloss, and Jonathan Sammartino

Department of Psychology, University of California, Berkeley, California 94720; email: sepalmer@gmail.com, kschloss@berkeley.edu, jonathansammartino@gmail.com

Keywords
visual perception, color, harmony, shape, spatial composition, art

Abstract
Human aesthetic preference in the visual domain is reviewed from definitional, methodological, empirical, and theoretical perspectives. Aesthetic science is distinguished from the perception of art and from philosophical treatments of aesthetics. The strengths and weaknesses of important behavioral techniques are presented and discussed, including two-alternative forced-choice, rank order, subjective rating, production/adjustment, indirect, and other tasks. Major findings are reviewed about preferences for colors (single colors, color combinations, and color harmony), spatial structure (low-level spatial properties, shape properties, and spatial composition within a frame), and individual differences in both color and spatial structure. Major theoretical accounts of aesthetic response are outlined and evaluated, including explanations in terms of mere exposure effects, arousal dynamics, categorical prototypes, ecological factors, perceptual and conceptual fluency, and the interaction of multiple components. The results of the review support the conclusion that aesthetic response can be studied rigorously and meaningfully within the framework of scientific psychology.
1. INTRODUCTION

We make judgments and decisions every day based on our internal aesthetic responses to aspects of the world around us. We decide to wear this sweater rather than that one because we prefer its color, we buy this poster rather than that one because we like its graphic composition, or we simply choose to sit facing this direction rather than that one in the park because we find the view more pleasurable. Such mundane but ubiquitous aesthetic considerations are so deeply woven into the fabric of our mental lives that we seldom reflect on what our preferences are or why we have them, but these are questions of scientific interest and import. Wearing this or that sweater, buying this or that poster, or facing this or that direction in the park may not seem to have much impact on our material lives, and yet, if we consider the alternative—a world in which we have no such preferences or could make no such choices—what a drab, dull, wearisome world it would be!

Given the pervasive influence of aesthetic responses on our mental lives, it is surprising how little is known about them. The study of aesthetic preference is actually one of the oldest topics in psychology, having been pioneered by Gustav Fechner (1876), one of the founders of modern scientific psychology. Interest in aesthetics has waxed and waned greatly since then, but a revival is now in progress as part of a movement to create an interdisciplinary aesthetic science (Shimamura & Palmer 2012). Our goal in this article is to review recent contributions to this field. Because there are so many, we have concentrated on behavioral research about aesthetic preferences in the visual domain. There is thus no coverage of many important topics in empirical aesthetics, including studies of nonvisual modalities (e.g., music cognition), facial attractiveness (Rhodes 2006), neuroaesthetics (Chatterjee 2011), or aesthetic processes in the creator (Turner 2006), all of which are active fields of current research.

2. FOUNDATIONAL ISSUES

2.1. Aesthetics Versus Art

Because art is so closely associated with visual aesthetics in most people’s minds, many readers are likely to assume that the present article

---

17.2 Palmer • Schloss • Sammartino
reviews the psychological study of visual art. It does not. To understand why, we begin by making a principled distinction between art and aesthetics, which are often taken to be essentially equivalent.

Perhaps the most straightforward difference is that significant aesthetic experiences can (and do) occur anywhere in response to seeing any sort of object, scene, or event, whereas art is limited to the subset of human artifacts intended to be viewed as art, whether in a museum, a gallery, or one’s own living room. Who among us has not experienced the ineffable delight produced by the sight of a delicate dew-kissed rose, a majestic snow-capped mountain, a lone seagull gliding silently over a sunset sea, or some other scene of comparable natural beauty? These objects and scenes are not construed as art—unless and until someone paints or photographs them—and yet who can deny that seeing them produces aesthetic experiences?

A second difference is that art has traditionally been identified with positive aesthetic experiences, whereas aesthetic response, in general, spans the range from very positive experiences (consider again viewing the rose, mountain, and seagull described above) to very negative ones (consider instead the sight of a decaying carcass oozing with maggots or a bloody, festering wound). Thus, although art can and (sometimes) does produce significant positive aesthetic experiences in beholders, there is much more to aesthetic experience than just people’s reactions to art. By the same token, however, there is much more to art than just the aesthetic responses it evokes in viewers. Art is a social, cultural, institutional, and commercial enterprise of enormous proportions, involving museums, galleries, curators, auctions, critics, collectors, historians, books, royalties, reproductions, and a host of other aspects whose relation to personal aesthetic response to works of art is often tangential, at best. Although aesthetics and art are indeed related, they are conceptually distinct and should be treated as such. The present article reviews the psychology of visual aesthetics, not the psychology of art.

2.2. Defining Aesthetics

Ideally, we would now provide a proper, closed-form definition of aesthetics, something like “The study of human minds and emotions in relation to the sense of beauty.” This seems helpful because it identifies minds and emotions as central concepts, but its problem lies in presupposing that the reader already knows what “the sense of beauty” is. For those who do not, this kind of definition is either disturbingly vacuous or implicitly circular. For present purposes, we take aesthetics to be the study of those mental processes that underlie disinterested evaluative experiences that are anchored at the positive end by feelings that would accompany verbal expressions such as “Oh wow! That’s wonderful! I love it!” and at the negative end by “Oh yuck! That’s awful! I hate it!” Such a definition is nakedly and unapologetically subjective, yet grounded scientifically in certain kinds of behaviors, much like defining pain as that dimension of human experience that leads one to say “Ouch!” and to hold or rub the injured body part. This definition will do nothing to explain the nature of aesthetic experiences to someone who has never had any, but for those who have, it allows them to be identified within our mental lives.

One might try to anchor such a definition objectively in the physical environment by proposing prototypical exemplars that would tie aesthetic experiences to external objects that regularly and reliably elicit them. In dealing with more objective experiential dimensions, such as “red,” prototypes are indeed useful: “Red” is that visual experience common to the appearance of ripe strawberries, fresh blood, classic fire engines, etc., when viewed under standard daylight conditions. Unfortunately, this strategy is not available for aesthetics simply because people so seldom agree. If one were to claim that aesthetic response refers to the kind of positive experiences elicited by viewing, say, Van Gogh’s *Starry Night*, Michelangelo’s *David*, Frank Lloyd Wright’s *Fallingwater*, and Ansel Adams’s *Moon Over Half Dome*, at least two important problems arise.
One problem is that different people's aesthetic responses differ so greatly that there may be some individuals who have strongly negative aesthetic experiences to all of the supposedly positive prototypes. This potential lack of agreement about aesthetic response even to such prototypes effectively derails the prototype strategy. The other problem is that aesthetic response, at least in our view, does not refer just to positive experiences or even just to extreme experiences. We assume that virtually everyone has some aesthetic response to virtually everything they see (e.g., Palmer et al. 2012b, Reber 2012). In most cases it may linger only fleetingly, if at all, in the “fringe” of human consciousness but it can come into focal awareness under appropriate circumstances, such as when the aesthetic response is extreme (seeing something so wonderful or so awful that the aesthetic response spontaneously bursts into consciousness), when one’s attention is directed to aesthetic response by context (in viewing paintings in a museum or gallery or even in shopping for home furnishings), or when one is given explicit instructions to do so (in a laboratory experiment). The fact that aesthetic responses are not always conscious, however, does not mean that people don’t have them all the time.

Another possible strategy for defining aesthetic response objectively is to “neurologize” it, as advocated in the emerging field of neuroaesthetics (e.g., Ramachandran & Hirstein 1999, Zeki 1999). The substantive claim is that aesthetics can be defined by some specified type(s) of activity in some specified set(s) of neurons in some specified area(s) of the brain. The hope is to identify the neural activity that actually produces aesthetic experiences, whatever that might be. Mary may love Van Gogh’s *Starry Night* and hate Da Vinci’s *Mona Lisa*, whereas Bill may have the opposite reactions, but if positive (and negative) aesthetic experiences arise from correlated amounts and/or types of neural activity in the same or functionally similar brain areas, perhaps aesthetic response can be objectively identified with that neural activity. The difficulty is that such neurological criteria cannot logically supplant the behavioral manifestations of subjective, experiential criteria within individuals. The reason is that, even though neural activity may be ontologically prior to subjective experience (i.e., the specified brain activity causes the aesthetic experience), it is epistemologically secondary and derivative. The logic of this claim is that, to find the kind of brain activity that causes people’s aesthetic responses, one must first identify behavioral manifestations of the target subjective aesthetic experiences and then determine what brain activity is correlated with those kinds of behavioral events. Without the behavioral measures, the neural correlates project can never get off the ground.

2.3. Philosophical Foundations

There is a long history of philosophical inquiry into aesthetics that predates all of the psychological studies we review below. As with many philosophical issues, its history begins in antiquity with treatments by Plato and Aristotle, and it continues, largely unresolved, to the present day. We do not attempt to review this extensive literature in any detail, as it has been reviewed elsewhere (e.g., Dickie 1997). We do want to comment, however, on a few central issues that are specifically relevant to the content of this review.

Much of the philosophical literature on aesthetics is explicitly concerned with art. Plato and Aristotle actually took quite different views on this subject. Plato denounced art as mere imitation that imparts no true knowledge of reality and even damages the soul. Aristotle exulted it precisely because people are indeed able to learn from (and even delight in) imitation. They agreed, however, on the importance of unity, harmony, and integration—see Plato’s discussion of a good argument in the *Phaedrus* and Aristotle’s analysis of tragic plot structure in the *Poetics*—a theme that was elaborated by later writers, including Kant and Dewey.

Kant’s (1892/1951) views on philosophical aesthetics in his *Critique of Judgment* have been particularly influential. His approach was decidedly psychological because he identified
beauty explicitly with the viewer’s mental experiences rather than the object’s physical properties. Aesthetic judgments, he claimed, were characterized by three key features: their subjectivity, their disinterested nature, and their claim of universality. Aesthetic judgments are subjective because they rest on personal experiences (e.g., liking/disliking), which have no objective empirical content with respect to the object itself, but only with respect to the relation between the object and the viewer. Unlike many other such liking/disliking judgments, however, aesthetic judgments are “disinterested” in the sense that they do not involve desire. Preferring a larger to a smaller piece of cake would not count as an aesthetic judgment in Kant’s framework, because such a judgment is (presumably) about one’s desire to consume the larger one.

Kant argued further that aesthetic judgments “claim universal validity.” This assertion can be confusing. Clearly, aesthetic judgments are not universally valid because not everyone agrees about them. Nevertheless, Kant claims that when someone makes an aesthetic judgment, he or she believes that others ought to share that judgment, as if beauty were an objective property of things rather than of our experiences on viewing them. Kant thus argued that aesthetic judgments involve more complex cognitive investments than judgments of “mere agreeability” or “individual preference,” which are merely matters of taste for which there appears to be no basis for discussion. Here Kant refers to the importance of the “harmonious free play of the imagination” in aesthetic judgments. Although it may not be entirely clear what this phrase means, it is evident that it refers to some relatively complex mental processes that are likely absent in the appreciation of “mere preferences.”

We agree with Kant that aesthetic judgments are subjective and disinterested, but not that they necessarily involve claims of universal validity and/or free play of the imagination. Although the latter may be a hallmark of many people’s views about works of art, they are hardly required in characterizing people’s preferences for colors, either singly or in combination (see section 4) or in their preferences for the shapes of rectangles or the spatial composition of simple pictures (see section 5). Rather, we see the aesthetic preferences reviewed below to be basic aspects of what become quite complex aesthetic responses that people have to artworks or scenes of natural splendor. Our belief is that eventually enough will be understood about how such simple preferences combine and interact with cognitive and emotional factors to explain more complex evaluative aesthetic claims that are often couched in universal terms and rely on Kantian free play of the imagination. We freely admit that whether this belief is true and how this understanding might unfold in an aesthetic science are as yet unclear.

2.4. On the Possibility of a Science of Aesthetics

Although many writers and researchers find it obvious that aesthetic experiences can be studied scientifically (e.g., Arnheim 1974, Berlyne 1971, Fechner 1876, Jacobson 2006, Shimamura & Palmer 2012), others do not (e.g., Dickie 1962). Indeed, some would surely claim that a science of aesthetics is not only impossible, but also oxymoronic (because science is objective and lawful, whereas aesthetic response is subjective and whimsical). Aesthetic responses are surely subjective (see above), but just as surely, that does not preclude their being studied objectively through behavioral methods (see below). For example, people’s experiences of color are not only subjective, but also can differ substantially over individuals owing to color blindness and/or color weakness, and yet there is nevertheless a well-established and technically sophisticated science of color vision (e.g., Kaiser & Boynton 1996, Koenderink 2010). There is no logical certainty that a scientific approach to aesthetics will succeed in identifying regularities in people’s aesthetic responses to visual displays, but this is clearly an empirical issue that can only be settled by trying. Below we summarize numerous results that provide significant insight into aesthetic preferences using rigorous scientific methods.
methods, providing strong empirical support for the viability of a science of visual aesthetics.

It is important to note that the factual basis of a science of aesthetics is not to settle whether some image or object is “objectively beautiful”—we agree with Kant that this is impossible—but rather to determine whether (or to what degree) some representative set of individuals judge or experience it as beautiful (or ugly). A science of aesthetics thus concerns accurately describing people’s aesthetic judgments and discovering the causes and/or reasons for those judgments. We now turn to the question of how one might do so.

3. METHODOLOGICAL ISSUES

Typically, the goal of a descriptive science of visual aesthetics is to determine average relative aesthetic preferences for some set of visual displays among a particular population, given a specific task to judge some aesthetic quality or qualities. The population, the displays, the task, and the aesthetic qualities of interest must all be specified prior to collecting data. The displays can be virtually any visible object, event, or image. The task is typically for the observer to attend to his/her experience of the designated aesthetic quality of the displays (e.g., their color or shape or overall attractiveness) and to indicate his/her evaluation. A variety of behavioral responses are possible—ratings, rankings, descriptions of likes or dislikes—but whatever they are, they must be measured behaviorally. Additional physiological measurements can also be made, such as galvanic skin response, functional magnetic resonance imaging, or event-related electrical potentials. Our review focuses on behavioral measures of aesthetic experiences, however.

Many of the methodological choices will be dictated by the specific hypothesis being tested: e.g., that American observers (the population) will tend to prefer color combinations of similar hue (the aesthetic quality) for a smaller square within a larger square (the displays). The sample of observers is important because people in different cultures, economic groups, and social groups may have systematically different aesthetic experiences to the same visual displays. The large interobserver variability in aesthetic preferences within such groups dictates using larger rather than smaller samples of randomly selected individuals.

If the goal is direct measurement (i.e., scaling) of average preferences among a full set of \( N \) visual displays, there are three primary behavioral tasks: two-alternative forced choice (2AFC), rank ordering, and preference ratings. Other methods can provide useful information but suffer from problems that limit their utility.

3.1. Two-Alternative Forced Choice

In most respects the optimal task is 2AFC, in which observers indicate which of two simultaneously presented visual displays they “like better” (prefer aesthetically) for all possible pairs. The average probability of choosing each display over all others is then taken as a global measure of its relative preference. 2AFC paradigms have the advantages of simple responses to indicate choice, trials containing only two displays, essentially no memory load, and minimal response bias effects. Its primary drawback is requiring \( N(N-1) = N^2 - N \) trials to measure preferences for \( N \) stimuli (e.g., 90 trials for 10 stimuli, but 9,900 trials for 100 stimuli). A more efficient alternative that has recently been developed is a Markov Chain Monte Carlo procedure that can be used when the display set is embedded within a dimensional space (Sanborn et al. 2010).

3.2. Rank Ordering

At the other end of the methodological spectrum is a rank-ordering task in which all displays of interest are presented simultaneously, and the observer is required to order them from most to least preferred. The average rank order for each display is then taken as a measure of relative preference. Although individual rank orders are, by definition, merely ordinal, the average rank over observers provides more quantitative information. The primary advantage is
that rank ordering requires only a single trial. Its disadvantages derive from the fact that this one trial can be inordinately complex, requiring simultaneous presentation of all \( N \) displays (difficult or impossible on a computer screen if the displays are large and/or spatially complex) and a potentially long and complex series of judgments to arrive at the required single, coherent ordering.

### 3.3. Rating

Ratings of aesthetic preference are an attractive alternative paradigm, especially when the number of displays to be measured (\( N \)) is large and/or the visual displays require significant space. The ratings can be made on discrete response scales of a given resolution (e.g., a seven-point Likert scale) or continuous scales (e.g., a line-mark rating), where the latter are often better if memory for specific prior ratings is a concern. Observers are shown a single display on each trial and asked to rate how much they like it. Average ratings for displays are taken as relative preference measurements. The responses are relatively simple, and only \( N \) trials are required to measure preferences for \( N \) displays. Because it can be difficult for observers to make consistent ratings across trials, especially at the beginning, it is useful to display the entire set (or a representative sample) of displays together with instructions to indicate the most and least liked alternatives to anchor the response scale before the experimental trials begin.

### 3.4. Production and Adjustment Tasks

Perhaps the most important alternative tasks for assessing aesthetic response are production and adjustment tasks, in which observers are required to produce the most aesthetically preferred possibility given various constraints. Such tasks are particularly useful when the experimental requirements imply a combinatorial explosion of possible displays. For example, suppose that one wanted to study aesthetic preference for all color pairs using a set of 100 possible colors. Because there are 100 \( \times \) 99 distinct permutations of the two colors, 9,900 trials would be required for a single observation per display, even using an efficient rating task. An adjustment task, such as fixing one color on each trial and asking the observer to adjust the second color to produce the most aesthetically pleasing combination, would be a much more efficient procedure, requiring only 100 trials. The dependent variable would be the probability of each adjusted color being chosen for each fixed color. This efficiency comes at the considerable cost of producing only a very complex and incomplete partial ordering, with many ties for the many combinations that are seldom, if ever, produced.

### 3.5. Indirect Measures

Various physiological measurements (e.g., galvanic skin response, functional magnetic resonance imaging, event-related electrical potentials) can theoretically be taken as covert aesthetic responses. For example, observers can be instructed simply to look at each display without requiring any behavioral response, with the average magnitude of the physiological measurement(s) of interest being used to create a scaling of the displays as if they were behavioral ratings. Obviously, such methods will be valid only to the extent that the physiological measurements have been validated by comparing direct behavioral measures (e.g., 2AFC probabilities or aesthetic ratings) with the given covert physiological measure(s).

One can also study indirect behavioral measures, such as the probability of looking first at one of two presented displays or the total time spent fixating each of them. Like physiological measures, they must be validated against direct behavioral measures of preference, which are logically primary. When this is not possible—such as with infants or animals—indirect measures are the only alternative. When combined with direct measures, indirect measures can often reveal nonconscious processes at work that are difficult or impossible to study with direct measures.
3.6. Other Measures

All of the above methods rely to some degree on direct, quantifiable, behavioral measures of aesthetic response in which observers are required to evaluate or create images in terms of their aesthetic reactions to them. There are other kinds of measurements that may prove relevant to understanding aesthetic response within the framework of some theory or hypothesis. For example, one might believe that the aspect ratio of an image’s rectangular shape or the power function of its spatial frequency spectrum contribute to its aesthetic success (see below). Such physical measurements can be combined (e.g., by correlation) with concurrent behavioral measurements, as described above. The behavioral measures are sometimes neglected, however, because an alternative index is taken as a proxy for aesthetic success, such as the fact that images hang in museums, are included in art books, or are included in some other collection of high repute.

There are several pitfalls to this strategy. One is the implicit assumption that images included in high-profile collections are aesthetically superior to other images for the studied population. This might be true, but it should not be taken for granted. A second more serious caveat is that correlation does not imply causation. Even if it were true, say, that people aesthetically prefer pictures that are wider than they are tall to those that are taller than they are wide, it is not necessarily true that the cause of this preference is their aspect ratios. People might actually prefer landscapes to portraits, for example, where landscapes tend to be wider and portraits taller.

The third problem is perhaps the most serious of all: the frequent failure of art theorists and critics to analyze appropriate contrast sets of images. If an author attributes the success of some acknowledged masterpiece to, say, the upward-pointing triangular composition of the principal elements, one should not take the claim as conclusive unless corresponding samples of other paintings with such upward-triangle compositions are judged aesthetically and contrasted with similar judgments of paintings with, say, “downward-triangle” or “quadrilateral” compositions. Few art theorists or critics consider such contrast sets, however.

4. COLOR PREFERENCE

4.1. Preference for Single Colors

Early researchers claimed that adult preferences for single colors were not systematic enough to warrant further investigation (e.g., Allesch 1924, Chandler 1928, Cohn 1884). Eysenck (1941) rejected this conclusion, arguing that previous failures arose from unstandardized colors and inadequate statistical analyses. He found reliable effects in analyses of both his own data and von Allesch’s previously reported “chaotic” data. Modern studies using standardized colors and sophisticated statistical techniques have clearly established that, despite large individual differences, group color preferences show systematic and reliable patterns as a function of the three primary dimensions of color: hue (basic color), saturation (vividness, purity, or chroma), and lightness (brightness or value) (e.g., Granger 1955b, Guilford & Smith 1959, Hurlbert & Ling 2007, McManus et al. 1981, Ou et al. 2004, Palmer & Schloss 2010).

Hue. Hue preferences among American and British adults follow a relatively smooth curvilinear function in which cool colors (green, cyan, blue) are generally preferred to warm colors (red, orange, yellow), with a maximum at blue and a minimum around yellow to yellow-green (see Figure 1). The majority of the variance is due to differences along the violet to yellow-green axis of cone-contrast color space [i.e., S–(L+M), where S, M, and L refer to the output of the short-, medium-, and long-wavelength cones] and/or the blue-yellow dimension of higher-level color-appearance space, with only minor differences along the red to blue-green dimension of cone-contrast (L–M) and/or the red-green dimension of color appearance (Hurlbert & Ling 2007, Ling & Hurlbert 2009, Palmer & Schloss 2010).
Saturation. Western adults generally prefer colors of higher saturation to those of lower saturation for context-free patches of color (McManus et al. 1981, Ou et al. 2004, Palmer & Schloss 2010). Some findings suggest that preferences decrease for colors of very high saturation, which were reportedly “too vivid” (Granger 1955b). Preference for high-saturation colors varies as a function of gender, culture, and object-context, however (see below).

Lightness. Western adults tend to prefer colors of increasing lightness, at least to some point (Guilford & Smith 1959, McManus et al. 1981), although this effect is not always evident (Palmer & Schloss 2010) (see Figure 1). Lightness is confounded with saturation across hues, however, and because people generally prefer highly saturated colors, different hues have their peak preference at different lightness levels: e.g., yellow at high lightness levels, red and green at medium lightness levels, and blue and purple at low lightness levels (Guilford & Smith 1959).

Hue x lightness interactions. Of considerable theoretical importance is the frequent finding that dark colors show a different hue preference function than light and medium-lightness colors of equal chroma, at least for warm colors (see Figure 1). In particular, dark shades of orange (browns) and yellow (olives) are strongly disliked relative to lighter, equally saturated oranges and yellows and relative to dark red (Guilford & Smith 1959, Palmer & Schloss 2010). Such effects are theoretically important because they are difficult to explain within classic psychophysical models of color appearance. They are consistent with ecological explanations of color preference, however (see below). The overall pattern of preferences for single colors is thus complex, but there are clear and repeatable regularities. For additional reviews of previous work on single color preferences, see Ball (1965) and Whitfield & Wiltshire (1990).

4.1.1. Theories of single color preference. The foregoing discussion specifies which colors people like, but not why they like the ones they do or even why they have color preferences at all. Several types of explanations have been suggested, including physiological, psychophysical, emotional, and ecological hypotheses. Because all of these models have been fit to the results shown in Figure 1 (see Palmer & Schloss 2010), we use them as a benchmark.

The most physiologically based theory suggests that people like colors to an extent that depends on a weighted average of cone contrasts relative to the background color, as computed very early in visual processing: \( L-M \) and \( S-(L+M) \), where \( S, M, \) and \( L \) represent the outputs of short-, medium-, and long-wavelength cones, respectively (Hurlbert & Ling 2007). This model fits Hurlbert & Ling’s (2007) own data well, but even an augmented model with two additional psychophysical variables—luminance and saturation (Ling & Hurlbert
2009)—accounted for only 37% of the variance in the data of Figure 1, presumably because their color sample did not include the highly saturated and easily named colors of Palmer & Schloss’s (2010) sample.

A similar, but purely psychophysical, hypothesis is that color preferences are based on conscious color appearances. Palmer & Schloss (2010) tested this possibility using a weighted average of observer-rated redness-greenness, blueness-yellowness, saturation, and lightness of each color, roughly analogous to their coordinates in Natural Color System color space (Härd & Sivik 1981). This model accounted for 60% of the variance in Figure 1. The fact that this is considerably better than the cone-contrast model suggests that later, conscious representations of color predict color preferences better than earlier, nonconscious ones.

A third explanation can be constructed in terms of the emotional associations of colors. Perhaps people like colors to the extent that they like the emotions that are evoked by or associated with those colors. Ou et al. (2004) measured color-emotions through subjective ratings and used those ratings to predict color preferences. Their results showed that three factor-analytic dimensions predicted color preferences: active-passive, light-heavy, and cool-warm, with active, light, and cool colors being preferred to passive, heavy, and warm ones. Palmer & Schloss (2010) fit observers’ subjective ratings of those three dimensions to the data in Figure 1 and found that it accounted for 55% of the variance. Ou et al. (2004) did not discuss how color-emotions arise or why some color-emotions predict preferences better than others. It is surprising, for example, that happy-sad was not included because it is such an important evaluative dimension. The difficulty is that the more desirable emotion (happy) is associated with less preferred hues (yellow), whereas the less desirable emotion (sad) is associated with more desirable hues (blue) (Terwogt & Hoeksma 1995).

An ecological explanation, called the “ecological valence theory” (EVT), was formulated and tested by Palmer & Schloss (2010). They proposed that people like/dislike a specific color to the degree that they like/dislike all of the environmental objects that are associated with that color. The ecological rationale is that it will be adaptive for organisms to approach objects whose colors they like and avoid objects whose colors they dislike to the extent that their color preferences are correlated with objects that are beneficial versus harmful to them (cf. Humphrey 1976). Palmer & Schloss (2010) reported strong support for the EVT through empirical measurements of what they call weighted affective valence estimates (WAVEs) for the 32 chromatic colors in Figure 1. The WAVE for each color measures the extent to which people like the set of objects that are associated with that color. For example, they like blues and cyans, at least in part, because they like clear sky and clean water, and they dislike browns and olive-colors, at least in part, because they dislike feces and rotting food. Average valence (liking/disliking) ratings were measured for all objects named as associates of each color, with each object valence rating being weighted by the similarity between the given object’s color and the color with which it was associated. The weighted average valence over all associates for each color explained 80% of the variance in the data shown in Figure 1 with no estimated parameters.

Several further predictions of the EVT have since been confirmed. One is that if preferences for colored objects causally influence preferences for corresponding colors, then color preferences could be changed by exposure to affectively biased samples of colored objects. Strauss and colleagues (2012) found that when one group saw positive red images (e.g., strawberries and cherries) and negative green images (e.g., mold and snot) and another group saw positive green images (e.g., kiwi fruit and leafy trees) and negative red images (e.g., blood and lesions), the two groups showed the predicted cross-over interaction in preference changes for the corresponding colors from ratings made before versus after seeing the affectively biased colored images. Seeing positive red images selectively increased preference for red,
for example, and seeing positive green things selectively increased preference for green.

The EVT also predicts changes in color preference for people who have highly positive (or negative) emotional investments in a social institution that has strong color associations, such as athletic teams, gangs, religious orders, and universities. People should come to like the colors associated with the institution correspondingly more or less, depending on whether their feelings about it are positive or negative. Indeed, students at two rival universities—the University of California, Berkeley and Stanford University—preferred their own university’s colors more than their rival’s colors, and the magnitude of these differences was correlated with their self-reports of school spirit for their own university (Schloss et al. 2011). These results support the EVT’s claim that color preference is influenced by preferences for the institution because it is highly unlikely that people choose their university and cultivate their level of school spirit based on pre-existing color preferences.

4.1.2. Infant color preferences. Infant color preferences are studied by measuring infants’ looking biases when presented with side-by-side color pairs. Longer (e.g., Bornstein 1975) and/or earlier fixations (Teller 1979) are taken as indices of relative preference. Infant color preference tends to be investigated from around 3–4 months, by which time infants’ color vision is fully trichromatic (e.g., Knoblauch et al. 2001, Teller 1998).

When infant looking biases are measured for highly saturated colors, the hue preference function tends to have the same general shape as the corresponding adult hue preference function (Figure 1, open circles), with a maximum around blue and a minimum around yellow or yellow-green (Bornstein 1975, Franklin et al. 2008, Teller et al. 2004, Zemach et al. 2007). Other patterns have sometimes been reported (Adams 1987), perhaps due to differences in luminance and/or saturation. The similarity of the infant and adult hue preference functions for saturated colors has led to the general view that color preference might be innate. More recent findings undercut this possibility, however. Franklin et al. (2009) measured looking biases in 4- to 6-month-old infants using the same colors Hurlbert & Ling (2007) used for adults and found that, unlike those of adults (see above), infant preferences varied primarily on a red to blue-green dimension (L-M cone contrast), with redder colors being more preferred, and were not strongly related to the violet to yellow-green dimension [S-(L+M) cone contrast]. More recently, Taylor et al. (2012b) have discovered that 4- to 6-month-old infants exhibit a hue x lightness interaction that is clearly different from the corresponding interaction in adults. Indeed, babies show a bias toward looking at dark-yellow and light-red and a bias against looking at light-blue and dark-green, nearly the opposite of corresponding looking biases in adults ($r = -0.46$). These findings clearly demonstrate that color preferences change dramatically during an individual’s lifetime.

4.1.3. Gender differences. Gender differences have been reported among Western adults in their relative preferences for saturated versus desaturated colors. In particular, men tend to prefer saturated colors more than women do, and these differences are strongly correlated ($r = +0.73$) with observers’ judgments of how active/passive colors are, with males generally preferring more active colors and females more passive ones (Palmer & Schloss 2011). Gender differences in saturation develop with age, being absent for young children (~6–9 years), beginning to appear during adolescence (12–13 years), and being clearly apparent by adulthood (17–18 years) (Child et al. 1968). Gender differences have also been reported in hue preference at certain ages. Specifically, girls ages 3–12 tend to prefer pink and purple, whereas boys of that age tend to prefer red and blue (Chiu et al. 2006, Iijima et al. 2001, Picariello et al. 1990). Such gender differences in color preference may develop from exposure to gender-specific toys that are stereotypically colored: pink and purple for girls’ toys and red,
blue, and black for boys’ toys (Jadva et al. 2010, LoBue & DeLoache 2011, Pennell 1994). This explanation fits well with an ecological theory (Palmer & Schloss 2010), according to which color preferences are determined by preferences for correspondingly colored objects.

Further evidence about ecological effects in gender differences has come from fitting WAVE data (see section 4.1.1) collected separately for adult males and females in England. Taylor & Franklin (2012) found a pattern consistent with EVT predictions: male WAVEs predicted male preferences \( (r = 0.86) \) better than they predicted female preferences \( (r = 0.58) \), whereas female WAVEs showed the opposite trend (female-female \( r = 0.67 \); female-male \( r = 0.58 \)).

4.1.4. Cross-cultural differences. Although some writers have claimed that color preferences are universal across cultures (e.g., Birren 1961, Eysenck 1941), modern empirical research reveals that both similarities and differences exist. The strongest case for a universal preference is for bluish colors (Adams & Osgood 1973, Hurlbert & Ling 2007, Ou et al. 2004, Saito 1996). There are exceptions, however, such as blue being ranked among the least preferred colors in Kuwait (Choungourian 1968). Blue was apparently also disliked in Ancient Rome, as evidenced by its relative disuse in historical artifacts (Pastoureau 2001), which may have been due to ecological factors. Because blue was greatly liked by Rome’s archenemies, the Celts, the Romans may have disliked blue by the same logic as Berkeley students dislike of Stanford’s colors and vice versa (see section 4.1.1). A second contender for universality is the robust dislike of dark-yellow (olive), which has been reported for Chinese, British (Ou et al. 2004), Japanese (Yokosawa et al. 2012), and American (Palmer & Schloss 2010) observers. Note that even if such universal color preferences exist, they may reflect universal features of human ecology rather than innate preferences: e.g., clear sky and clean water are universally blue, whereas biological wastes and rotting food are universally dark yellow.

According to the ecological valence theory, WAVEs for a given culture should predict that culture’s color preferences better than another culture’s color preferences, and vice versa. This prediction is based on the idea that different cultures frequently have different color-object associations and/or different valences for the same objects, both of which can affect color preferences. Preliminary results support this prediction for American versus Japanese cultures (see Palmer & Schloss 2010). Japanese WAVEs accounted for Japanese preferences \( (r = 0.66) \) better than for American preferences \( (r = 0.55) \), whereas American WAVEs accounted for American preferences \( (r = 0.89) \) better than for Japanese preferences \( (r = 0.74) \).

It must be noted, however, that a radically different pattern of preferences emerged in the nonindustrial Himba people of rural Namibia, who greatly preferred highly saturated colors, largely independent of hue (Taylor et al. 2012a). This large saturation effect is most consistent with a psychophysical account of color preference because saturation is a potent variable in high-level models of color appearance. It may also be due to novelty, as the Himba’s chromatic environment is considerably more natural than that of industrialized societies, with far fewer instances of highly saturated colors.

Many researchers have appealed to color symbolism to explain cultural differences. For example, there appears to be a stronger preference for white and whitish colors in Japan, Korea, and Taiwan than in other countries (e.g., the United States, Germany, Australia, Papua New Guinea, and South Africa), perhaps because white symbolizes cleanliness, purity, and the sun, which are more highly valued in these Asian cultures than in the other cultures studied (Saito 1996). Similarly, Chinese observers preferred red more than British participants did, perhaps due to its role as a symbol of good luck in China (Hurlbert & Ling 2007). The idea that color symbolism influences color preferences is conceptually consistent with ecological accounts of color preference (Palmer & Schloss 2010), according to which color preferences are determined by people’s preferences for the
“things” that are associated with those colors, provided that the “things” include abstractions such as purity and good luck as well as concrete objects and social institutions. Solid empirical evidence for the role of color symbolism in color preference is largely lacking, however.

4.1.5. Object-based differences. A question of considerable applied interest is whether preferences for context-free colored patches (see above) generalize to preferences for colored objects. The answer depends importantly on the degree to which the objects in question have prototypical or characteristic colors. It is highly improbable, for example, that even a resolute blue-lover and yellow-hater would prefer blue bananas to yellow ones. The more sensible question is whether context-free color preferences generalize to artifacts that could conceivably be any color.

Schloss et al. (2012) studied adult preferences for visually presented colors of many such artifacts: interior room walls, room trim, couches, throw pillows, dress shirts/blouses, ties/scarves, and T-shirts. The overall pattern of hue preferences was similar to that for context-free colored squares (Figure 1), with cool hues (especially blues) being liked better than warm ones. The primary exception was that people tended not to like large things (such as walls and couches) to be red. In contrast, the pattern of preferences for different lightness levels were markedly different across objects, often depending on practical considerations, such as liking lighter wall colors to make rooms appear brighter and/or larger. Interestingly, context-free squares were the only case in which highly saturated colors were preferred to less saturated ones (Saito 1983, Schloss et al. 2012).

Object color preferences are most likely to deviate from context-free color preferences when there are color conventions for that type of object (Taft 1997). Such conventionality effects can be present even within a basic-level object category, such as luxury sedans being most preferred in achromatic colors (black, grays, or white), consistent with their conventional formality as serious, sophisticated cars, whereas a VW “Bugs” were preferred in brighter, more saturated colors, including yellow, consistent with their conventional informality as fun, sporty cars (Schloss et al. 2012). These effects may be considered weaker cases of the prototypical banana example described above, but ones in which sociocultural conventions, rather than naturalness, drive color preferences.

4.2. Preference for Color Combinations

Historically, theories about the aesthetics of color combinations have been predicated on several untested assumptions that have led to a confusing literature. For example, the terms “preference” and “harmony” are often used interchangeably, and preference for a combination as a whole is frequently confused with preference for a figural color against a background color (e.g., Chevreul 1839; Granger 1955a,c; Itten 1973). Schloss & Palmer (2011) used a clearer empirical framework for assessing the aesthetics of color combinations by distinguishing among three types of judgments for figure-ground color pairs: (a) pair preference (how much the two colors are liked together), (b) pair harmony (how well the colors go together, regardless of preference), and (c) figural preference (how much the foreground color is liked when viewed against a colored background). They explain the distinction between preference and harmony in terms of a musical analogy: Nearly everyone would agree that Mozart’s music is more harmonious than Stravinsky’s, and yet some people prefer Stravinsky’s music to Mozart’s whereas others prefer Mozart’s to Stravinsky’s (see also Albers 1971, Ou et al. 2004, Whitfield & Wiltshire 1990). Although figural preference involves a judgment about a single color, it is relevant to color combinations because the same color can look strikingly different on different background colors due to the classic perceptual phenomenon of simultaneous color contrast (Chevreul 1839, da Vinci 1942/1956, Helmholtz 1866/1925, Walraven 1976).
4.2.1. Theories of pair preference and/or harmony. Chevreul (1839) presented one of the most influential theories of color harmony and preference, terms that he used interchangeably. He proposed two distinct types of harmony: harmony of analogous colors and harmony of contrast. Other theories include Itten’s (1973) claim that any combination of colors is harmonious if the colors produce neutral gray when mixed together as paints, Munsell’s (1921) and Ostwald’s (1932) theories that colors are harmonious when they have certain relations in color space (e.g., constant in hue and saturation but varying in lightness), and other theories proposed by Goethe (1810/1970), and Moon & Spencer (1944a,b). Note that none of these theories were formulated on the basis of aesthetic measurements, although some have since been tested empirically. (For a more thorough review, see Westland et al. 2007.)

4.2.2. Empirical research on pair preference and/or harmony. Studies of aesthetic preferences for color combinations have produced conflicting claims, often because researchers have not distinguished among the three types of judgments described above. For example, Granger (1955a,b) reported that harmony and preference (he also used the terms interchangeably) increased with hue contrast, but several subsequent studies show that both harmony and preference decrease with hue contrast (Ou et al. 2004, Ou & Luo 2006, Schloss & Palmer 2011, Szabó et al. 2010). Granger’s finding that preference increases with hue contrast is consistent, however, with subsequent findings on figural color preferences, which increase with hue contrast (Helson & Lansford 1970, Schloss & Palmer 2011).

Schloss & Palmer (2011) measured both pair preference and pair harmony for the same figure-ground color pairs (a small square within a larger square) with the same observers given different instructions (see above). Both measures increased with hue similarity (cf. Chuang & Ou 2001, Ou & Luo 2006, Ou et al. 2011, Szabó et al. 2010), consistent with Chevreul’s harmony of analogous colors. The two measures differed in relation to lightness, however, in that pair preference increased with lightness contrast, whereas pair harmony did not. Harmony also tends to be greater for lighter pairs and for pairs that are more similar in saturation (Ou & Luo 2006, Schloss & Palmer 2011, Szabó et al. 2010). Nevertheless, the overall correlation between ratings of pair preference and pair harmony is highly positive ($r = +0.79$) in Schloss & Palmer (2011). This fact explains why these concepts have so often been equated: people tend, on average, to like harmonious color combinations (see also Ou et al. 2004). Observers agreed more about their ratings of pair harmony (average between-observer $r = +0.51$) than about their ratings of pair preference (average between-observer $r = +0.36$), primarily because pair harmony ratings were almost completely independent of pair preferences for the two component colors. Strikingly, more than 80% of the variance in people’s pair preference ratings can be explained by a linear combination of the pair’s harmony rating, the preference ratings for the two colors alone, and the lightness difference between figure and ground colors (multiple-$r = 0.89$).

Ratings of figural preference for colors against colored backgrounds are measurably different from both pair preference and pair harmony ratings (Schloss & Palmer 2011). Average figural preferences are highly correlated with preference for the corresponding single figural color against a neutral background ($r = 0.87$) and also with average preference for pairs of colors containing the figural color as figure ($r = 0.74$). After the influences of these two preference factors were removed, however, preferences for hue contrast and lightness contrast were evident (cf. Helson & Lansford 1970). It is likely that these effects are responsible for Chevreul’s claims about harmonies of contrast. Indeed, given the general preference for saturated colors (see Figure 1), it is not surprising that observers would prefer figural colors against highly contrastive background hues; simultaneous color contrast effects would
tend to increase the perceived saturation of the figural color.

5. PREFERENCE FOR SPATIAL STRUCTURE

Virtually every two-dimensional visual display—from painted portraits to photographed landscapes to abstract graphic designs—is composed of elements arranged in space, usually within a rectangular frame. It is therefore important to consider how the spatial properties and arrangements of those elements influence people’s aesthetic responses to such displays (e.g., Arnheim 1988).

5.1. Aesthetics of Low-Level Visual Properties

A common theme in aesthetic judgments of low-level visual properties is that images are preferred when their structure mirrors that of natural scenes. Such findings suggest that people prefer images that have the statistical structure to which the human visual system has adapted, whether evolutionarily or ontogenetically.

5.1.1. Spatial frequency. By the time visual information reaches the cortex in area V1, spatial analysis is dominated by local spatial frequency filters that perform something like a piecewise Fourier analysis of the input (cf. De Valois & De Valois 1990). Spectral analyses of paintings and natural scenes reveal that they have similar amplitude spectra, being highest for low spatial frequencies and decreasing approximately linearly with the logarithm of spatial frequency in the classic spectrum of 1/F noise (Graham & Field 2007, Graham & Redies 2010). Not only do more “natural”-looking paintings have the same 1/F power spectrum, but increases within two octaves of three cycles/degree are correlated with viewers’ feelings of visual discomfort (Fernandez & Wilkins 2008). It is not clear, however, whether 1/F power spectra are either necessary or sufficient for producing positive aesthetic responses to images.

5.1.2. Line orientation. Viewers generally prefer horizontal and vertical lines to oblique ones in Mondrian-like images (e.g., Latto et al. 2000), and, generally speaking, more paintings contain horizontal and vertical lines than oblique ones (Latto & Russell-Duff 2002). Given that the visual stimulation to which people are exposed also tends to have more horizontal and vertical lines than oblique lines, at least at higher spatial frequencies (Switkes et al. 1978), these results provide further evidence that people prefer images that mirror the statistical properties of low-level spatial structure in their seen environment.

5.2. Aesthetics of Object Shape

5.2.1. The golden ratio. One of the oldest and most frequently studied questions about the aesthetics of shape concerns the so-called golden ratio or golden section. It is obtained by dividing a line into two parts such that the proportion of the entire line to the longer segment is equal to the proportion of the longer segment to the shorter segment, a ratio of approximately 1.6:1. If the golden ratio characterizes the length-to-width ratio, or aspect ratio, of a rectangle, it defines a shape that has been claimed from antiquity to be particularly pleasing aesthetically (e.g., Green 1995). The golden ratio appears frequently, both in nature (e.g., a person’s height relative to his/her arm span) and in aesthetically acclaimed human artifacts (e.g., the base-to-height ratio of the Great Pyramids and the facial dimensions of Da Vinci’s Mona Lisa) (e.g., Atalay 2004, Konecni 2003, Konecni & Cline 2002, Livio 2002).

Early empirical studies showed that so-called golden rectangles are most preferred, with preference diminishing in both directions as the length-to-width proportions deviate from 1.6:1 (Fechner 1871, 1876). Such claims remain controversial (e.g., Boselie 1984, Green 1995), with some successful replications (e.g., Lalo 1908, Thorndike 1917) but...
many failures (e.g., Angier 1903, Haines & Davies 1904, Thorndike 1917). Interestingly, preferred aspect ratios of rectangles differ for different semantic categories of objects. People preferred invitation cards for “serious” events (e.g., a classical piano recital) to have aspect ratios close to the golden ratio (peaking at 1.4:1), but the preference functions of those for more casual “fun” events (e.g., a child’s birthday party) were virtually flat from 1:1 to 1.5:1 (Raghubir & Greenleaf 2006). A reasonable summary of this field of research is that although many people prefer shapes whose dimensions are in the general neighborhood of the golden ratio, such preferences can and do vary relatively widely across both observers and contexts (McManus 1980).

5.2.2. Complexity and symmetry. In an influential theory of preference for abstract polygon shapes, Birkhoff (1933) proposed that aesthetic preference (M) should vary directly with the number of elements (O) and inversely with the complexity (C, expressed as the number of noncollinear sides) according to the equation $M = \frac{O}{C}$. Experimental tests of this claim have yielded disappointingly low correlations, however, with various other authors giving their own formulations (e.g., Boselie & Leeuwenberg 1985, Eysenck 1941, Eysenck & Castle 1971). One problem with this equation is that it predicts monotonic increases in preference with complexity. In contrast, intermediate complexity of about 10 sides is generally preferred by both adults (Martindale et al. 1988, Munsinger & Kessen 1964b) and children (Munsinger & Kessen 1964a). These results are consistent with Berlyne’s (1971) arousal theory of aesthetics (see below). Interestingly, preference effects of complexity show strong contrast effects with massive amounts of familiarization: People familiarized with simple stimuli later tended to prefer more complex stimuli, whereas those familiarized with complex stimuli tended later to prefer simpler stimuli (Tonio & Leder 2009).

More recent research on shape preferences has focused on symmetry structure while holding the number of elements constant, following prior results showing that more symmetrical dot configurations were more easily processed perceptually and better remembered (Garner & Clement 1963). In general, people tend to prefer shapes that are more symmetrical, although there are large and relatively stable individual differences in such effects (Jacobson & Höfel 2002, Palmer & Griscom 2012), as is discussed below.

5.2.3. Contour curvature. A recently discovered phenomenon of shape preference is that people tend to like objects with curved contours more than similar objects with sharp contours (Bar & Neta 2006). Further research has shown this to be true for abstract shapes as well as recognizable objects (Silvia & Barona 2009) and that the preference for curved objects holds for objects with neutral or positive valences, but not for ones with negative valences (Leder et al. 2011). These findings have been interpreted to mean that sharp contours are more threatening than curved ones, given the tendency for objects with sharp contours to be harmful.

5.2.4. Categorical Prototypes. People also tend to prefer object shapes to the extent that they conform to categorical prototypes (Rosch 1975). There is a large, well-known literature on facial attractiveness, showing that people tend to like symmetrical, average faces (for a review, see Rhodes 2006). We do not count this as an aesthetic preference because it is not indifferent in the Kantian sense, presumably being driven by sexual attraction. Nevertheless, the idea that people tend to like prototypical exemplars has found support in other domains, such as color (Martindale & Moore 1988), furniture (Whitfield & Slatter 1979), surrealist paintings (Farkas 2002), and exemplars of other semantic categories (Martindale et al. 1988).

5.3. Aesthetics of Spatial Composition

Whatever the shape of the objects within an image, an important further aspect of aesthetic consideration is spatial composition: the way
in which the objects are positioned relative to each other and to the surrounding frame. A primary focus of studies of spatial composition is the balance of the image in terms of how the elements are distributed around the frame’s center. Arnheim (1988) and Alexander (2002) both view the understanding and analysis of centers as indispensable aspects of art and design. They typically provide many illustrations of their ideas, but no experimental evidence to justify them. Below we review some of the scientific findings related to frame structure, balance, and centers, as well as more complex aspects of spatial composition.

5.3.1. The structure of a rectangular frame. Many of the factors related to spatial composition depend critically on the relation between the elements and their surrounding rectangular frame. Arnheim (1974) speculated that a frame has a “structural skeleton” defined by dynamic forces propagated from its sides. Recent empirical measurements have supported Arnheim’s intuitions remarkably well (Palmer 1991, Palmer & Griscom 2012, Palmer & Guidi 2011). When observers rated the “goodness of fit” for the placement of a single probe circle at 35 locations within a rectangular frame, the center of the frame always had the highest average goodness rating, which decreased monotonically with distance from the center. The next highest ratings were along the vertical and horizontal axes of symmetry, with smaller increases along the angle bisectors at each corner of the frame, but notably not along the diagonal that runs through the center. Indeed, the central axis of many shapes holds a special salience for other perceptual measures, such as contrast sensitivity (Kovacs & Julesz 1994). Subsequent research revealed that people generally prefer images in which the dot “fits well” within the frame, with average preference ratings being very highly correlated with average goodness-of-fit ratings ($r = 0.95$), although large individual differences were evident (Palmer & Griscom 2012) (see section 6).

5.3.2. Balance and centers. Some of the earliest empirical studies of balance investigated images containing a fixed line of some particular length, width, and color plus a test line of some particular length, width, and color. Observers were asked to adjust the position of the test line to make the entire display appeared balanced (Pierce 1894, Puffer 1903). All else being equal, test lines were placed farther from the center than the fixed line if they were shorter and/or thinner than the fixed line, but they were placed closer to the center if they were longer and/or thicker. Although these results are consistent with physical concepts of balance, darker colors were treated as if they were heavier (Bullough 1907), thus defying explanation by a purely mechanical analogy, since colors have no physical weight.

Other studies have evaluated perceived balance for works of art and compositions of abstract shapes. When asked to place a fulcrum below the balance point of an artwork image, observers generally indicated pictures were balanced near the center of the frame, with more images being balanced slightly left of center (McManus et al. 1985). For abstract displays of one or two colored squares (red, green, or blue), position strongly influenced perceived balance, but there were also significant interactions with color, with red affecting balance most and blue least. Locher et al. (2005) found similar effects of color when comparing perceived balance in Piet Mondrian’s original abstract paintings versus variants with interchanged colors. Observers judged the originals to be balanced near the center and to be more balanced than the color variants. Based on judgments of “balance center,” the perceived weight of a region varied as a function of its size and color, with red being perceived as heaviest, blue intermediate, and yellow lightest. In another study, when artwork was modified in ways that influenced balance and composition (e.g., by deleting elements), observers perceived compositionally

---

1By “frame” we mean to include cases in which there is no actual frame around the image, but simply the rectangular borders of the image.
balanced images to be more visually “right” (Locher 2003). Finally, when participants arranged paper cutout shapes in an empty frame to produce displays that were “both interesting and pleasant,” about half of the designs displayed symmetry at some point in the creation process, and the physical weight was equally distributed around the center in the majority of final designs (Locher et al. 1998).

A quite different claim for the aesthetic importance of centers comes from Tyler’s (1998) finding that one of the two eyes in nonprofile portraits of human faces almost always lies at or very near the frame’s vertical midline. This strong center bias was much more pronounced for the eye than for the face as a whole, the mouth, or even the single eye in profile portraits. McManus & Thomas (2007) challenged Tyler’s claim, finding that people showed no preference for portraits in which one eye was centered compared with otherwise identical images in which neither eye was centered. Controversy continues over reasons for the centered-eye phenomenon, with McManus & Thomas (2007) claiming that it arises from geometric constraints on positioning a single head within a pictorial frame, and Tyler (2007) rebutting their objections.

5.3.3. Compositional biases in pictures of meaningful objects. Palmer, Gardner/Sammartino, and colleagues have attempted to understand aesthetic preferences for spatial composition of pictures containing just one or two meaningful objects within a rectangular frame. Using several different measures, Palmer and colleagues (2008) found strong, systematic tendencies for symmetrical objects that faced forward to be preferred at the center of the frame (the center bias) and for objects that faced rightward or leftward to be located off-center to the left or right of center, respectively (the inward bias). An inward bias is also evident in artists’ paintings and drawings of animals (Bertamini et al. 2011).

Sammartino & Palmer (2012a) found similar biases in preferences for vertical position of single object pictures. For example, when the image of an object was symmetrical about a horizontal axis (e.g., an eagle seen from directly above or below), they found a center bias. When objects “faced” upward (e.g., a bowl) or downward (e.g., a light fixture), they found an inward bias: People preferred the bowl below the frame’s center and the light fixture above it. Ecological biases were also evident, in that people preferred the position of the object within the frame to be consistent with its typical location relative to the observer: Eagles were preferred higher in the frame, whereas stingrays were preferred lower. Complex patterns of preferences arose from the combined influences of these different biases.

Other ecological effects are also evident in people’s aesthetic responses to the depicted perspective and size of a focal object in an image. Previous research on canonical perspective had demonstrated that people can recognize everyday objects better from the perspectives from which they are more likely to be seen (e.g., Palmer et al. 1981). Preferences for different perspective views follow a similar pattern (Palmer et al. 1981, Sammartino & Palmer 2012b). Analogously, there is a canonical size effect in aesthetic judgments. People tend to prefer small objects (such as a butterfly) to be smaller within the frame and for larger objects (such as an elephant) to be larger within the frame (Konkle & Oliva 2011), even if each observer sees only different-sized pictures of a single object (Linsen et al. 2011).

Further studies have examined semantic effects in the composition of pictures containing two objects that were either closely related in everyday use (e.g., a champagne bottle and a cake or a liquid detergent bottle and a sponge) or quite distantly related (e.g., a champagne bottle and a sponge or a liquid detergent bottle and a cake) (Leyssen et al. 2012). The results consistently showed that people prefer related objects to be much closer together and unrelated objects to be much farther apart within the frame, even if placing two related objects nearby produced an unbalanced composition. Indeed, when 2AFC comparisons were performed to look explicitly for balance versus semantic
relatedness effects, the semantic bias was clearly evident, but there was no bias toward balance. Although this finding appears to contradict previous research on balance (see above), prior research has almost exclusively used abstract geometrical shapes that do not have meaningful ecological relations to one another.

McManus et al. (2011) recently explored various factors that affect the way in which different complex photographic images were cropped within a fixed-size rectangular frame. Participants adjusted the position of the rectangular frame within a larger image so that they achieved the most aesthetically pleasing image that contained a specified location. Some low-level, image-based features, such as color, had little effect on cropping choices, whereas others, such as detail, had more substantial effects, but higher-level, meaning-based features seemed to dominate.

The bulk of these results on spatial composition of pictures of meaningful objects support a general conclusion that prior knowledge plays a crucial role. Even the inward bias depends on observers knowing which side is an object’s front, and all the other forms of ecological bias rest on observers knowing the ecological statistics of meaningful objects: the perspective from which they are seen (canonical perspective), how big they are (canonical size), and where they tend to be positioned vertically relative to eye level (which might be termed “canonical elevation”). Although it has never been demonstrated in aesthetic judgments, it seems clear that there would also be large aesthetic differences due to an object’s orientation relative to the frame (canonical orientation). All of these effects are consistent with perceptual fluency theories of aesthetic preference (see below): People like pictures of objects to the degree that they are easily perceived (Reber et al. 2004, Winkielman et al. 2006).

5.3.4. Higher-level spatial composition: Effects of meaning, titles, and context. Higher-level factors, such as meaning, emotion, and/or the interpretation of images in different contexts, are often thought to be important in aesthetic preference, especially in philosophical discussions, such as Kant’s (1892/1951) writings about the importance of “free play of the imagination.” Only recently have such factors been investigated experimentally.

Millis (2001) explored the effect of different kinds of titles for illustrations and photographs on four qualities of the aesthetic experience: interest, thoughts, emotions, and aesthetic appreciation. Metaphorical titles, which gave additional information not immediately evident from the image itself (e.g., titling an image of a woman picking flowers One Day at a Time) increased aesthetic appreciation over purely descriptive titles (e.g., Woman Picking Flowers) or no titles at all. Russell (2003) examined the effects of providing viewers with the artist’s names and actual titles and/or brief descriptions of abstract or semi-abstract images on both the viewers’ judgments of “meaningfulness” and their aesthetic judgments. He expected meaningfulness to increase aesthetic appeal, but found this to be true only in a within-subjects design. Leder et al. (2006) extended this kind of investigation to both abstract and representational artworks and examined the effects of exposure duration. Titles did not influence liking or content understanding overall, but they did interact with durations. Descriptive titles were more effective at brief 1-second exposures, and elaborative titles were more effective at longer 10-second exposures.

Extending their prior research on biases in object placement within a frame (Palmer et al. 2008), Sammartino & Palmer (2012b) studied systematic effects of different titles on preference for the horizontal placement and perspective views of objects. When the title was purely descriptive (e.g., Racehorse and Man Walking), they replicated their previous results, with observers preferring the object near the center of the frame and facing inward. When the title elaborated the context in a way that implied unseen objects or portions of an event (e.g., Front Runner versus Dead Last for the running horse or Journey’s End versus Starting Out for the man walking), they found that the most preferred composition was dramatically
different: Compositions that were least preferred with neutral titles were most preferred with elaborative titles that fit the composition. They found analogous effects of elaborative titles on different perspective views of objects.

6. INDIVIDUAL DIFFERENCES

It is widely acknowledged that people differ enormously in their aesthetic preferences for all kinds of different modalities and domains (e.g., McManus 1980, McManus et al. 1981). This fact, more than any other, underlies the well-known adages, “Beauty is in the mind of the beholder” and “There is no accounting for taste.” In spatial composition, for example, McManus & Weatherby (1997) found that average positional preferences were close to the golden section value in horizontal placement, but individuals differences (IDs) were so large that few, if any, participants showed preference functions that looked much like the group averages.

In color preferences, Ling & Hurlbert (2009) used their extended, four-parameter cone contrast model (see section 4.1.1) to fit individuals’ preference data. It accounted well for IDs (average multiple-$r = 0.71$), indicating that observers differ in the polarity and importance of these four dimensions for single colors. From an ecological perspective, Palmer et al. (2012a) found that, as predicted by the EVT, WAVE measurements for individual observers were more highly correlated with their own color preferences (average $r = +0.55$) than with those of other observers (average $r = +0.40$).

In preferences for two-color combinations, Schloss & Palmer (2011) found that even though average preference ratings correlated very strongly with average harmony ratings (+0.79; see above), the same correlations for individuals ranged widely from about zero to +0.75. Furthermore, these IDs varied with the amount of formal color training participants had completed, following an inverted-U function, with a maximum for intermediate amounts of training and lower correlations for both untrained and highly trained individuals. This pattern is consistent with Berlyne’s inverted-U prediction (see below).

Eysenck (1940) studied people’s aesthetic rankings of sets of spatial images of 18 different types (including portraits, landscapes, and photographs of medieval clocks) as well as to colors, odors, and polygons. After correlating the rankings and performing a factor analysis, he identified a single “general objective factor of aesthetic appreciation” (which he called “t” for “good taste”) that varied with the extent to which an individual’s rankings agreed with the average rankings of the entire group. It turned out to be relatively constant for individuals across domains. Although Eysenck’s interpretation of the t-factor as an objective measure of good taste seems misguided (see section 2.3), this measure might be an important variable in characterizing systematic IDs in aesthetic preferences.

Palmer & Griscom (2012) have recently proposed that “preference for harmony” may be the primary aesthetic ID that underlies Eysenck’s t-factor. Following Schloss & Palmer’s (2011) findings of IDs in preference for color pairs (see above), they studied the extent to which people’s judgments of aesthetic preference among stimuli in different domains (color pairs, configural shapes, spatial compositions, and music) correlated with their own judgments for the same stimuli in their degree of harmony, where harmony is a dimension characterized by simplicity, regularity, and parts that fit well together, regardless of preference. They found that the correlations between average preference and average harmony ratings over all individuals were quite high in all domains (ranging from +0.97 for music to +0.47 for configural shape) but that the same correlations for individuals were extremely variable: Some people like harmonious stimuli and others dislike them, with IDs in preference for harmony in music ranging from −0.73 to +0.97. Most importantly, the correlations between a difference-score measure of preference for harmony were systematically above chance for all pairs of different domains, ranging from...
+0.60 for color pairs and music to +0.32 for spatial composition and music. Preference for harmony can explain Eysenck’s t-factor because the preference judgments of individuals who prefer harmonious stimuli will necessarily be more highly correlated with average preferences to the degree that people generally prefer harmonious stimuli in that domain, as it was for all four of Palmer and Griscom’s domains. Furthermore, they found that specific training in an aesthetic domain (art or music) systematically lowered people’s preference for harmony in that domain, consistent with Berlyne’s (1971) analysis of the effects of experience.

7. THEORIES OF AESTHETIC RESPONSE
The foregoing review summarizes many of the most important findings about human aesthetic preference in visual domains as indexed by behavioral responses. Although the preferences discussed are fairly basic, the hope is that they constitute a foundation on which a better understanding of more complex aesthetic responses can be built, perhaps eventually including people’s powerful aesthetic reactions to works of art. It is by no means obvious that this can be accomplished, and some writers have explicitly claimed that it cannot (Markovic 2010). We now turn to ask why people like the things they do. Several accounts have been suggested.

7.1. Mere Exposure
One possible explanation of why people like some things more than others is the “mere exposure” effect: People tend to like objects and images more as the frequency of seeing them increases (Zajonc 1968). One might try to explain aesthetic preference for, say, photographs of inward-facing objects over outward-facing objects (Palmer et al. 2008) by appealing to the fact that people simply see more pictures of inward-facing objects. The problem is that one must also explain the latter fact. It is tempting to say that it is because photographers like inward-facing pictures more, but then one still needs to explain this preference in some way other than invoking the mere exposure effect again—i.e., that photographers have seen more pictures of inward-facing objects. The root problem is thus that explanations based purely on exposure frequency lead to infinite regress.

We do not deny that exposure frequency influences people’s aesthetic judgments; certainly it does. Cutting, for example, has provided a compelling historical analysis of its influence on aesthetic judgments of certain paintings that are included in the Impressionist canon (Cutting 2003, 2006). The ultimate explanation of their inclusion does not come from mere exposure, however, but rather from differences in the availability of certain paintings at an important historical juncture when the canon was defined: The pictures that were available to the public were included rather than those only available to private owners. The important conclusion for present purposes is that, although mere exposure can explain the perpetuation and even the amplification of a bias across time, it cannot satisfactorily explain its ultimate cause.

7.2. Arousal Dynamics
Daniel Berlyne developed a complex and influential theory of aesthetic response based on a psychobiological conception that pleasure is a matter of the viewer’s degree of arousal while viewing an image (Berlyne 1957, 1960, 1971). Arousal, in turn, was proposed to depend on a complex psychobiological response to three types of variables—collative, psychophysical, and ecological—each of which produced responses in a primary reward system and a primary aversion system. The combined effect of these two systems was claimed to produce an inverted-U function of the underlying variable, with aesthetic pleasure first increasing as a function of arousal and then decreasing as arousal became too great. Collative variables are those related to the viewer’s expectations (e.g., novelty, complexity, uncertainty, conflict, surprisingness, unfamiliarity), psychophysical variables are those related to sensory dimensions of the stimulus (e.g., intensity, pitch,
brightness), and ecological variables are those related to meaningfulness and associations to environmental objects. According to Berlyne’s theory, as a relevant variable increases (e.g., complexity), a primary reward system becomes increasingly active and generates positive affect. As complexity continues to increase, however, an aversion system becomes active and generates negative affect. The reward system saturates before the aversion system does, leading to the classic inverted-U curve so closely associated with Berlyne’s theorizing. Berlyne claimed that collative variables were the most important, and the bulk of his empirical work was directed at understanding them. We will not attempt a full review of the theory or the evidence relevant to its evaluation. Suffice it to say that although Berlyne amassed much evidence in support of his ideas (e.g., Berlyne 1971, 1974), some of its key predictions have not been supported empirically (Martindale et al. 1990).

7.3. Prototype Theory
Rosch’s (1975) transformative research on prototype effects in categorization suggested a different avenue to explaining visual preferences that was developed primarily by Martindale and Whitfield. Simply put, people may prefer prototypical examples of categories to nonprototypical ones. Preferences for prototypes have been demonstrated for colors (Martindale & Moore 1988), furniture (Whitfield & Slatter 1979), faces (Light et al. 1981), exemplars of semantic categories (Martindale et al. 1988), and surrealist paintings (Farkas 2002).

This work can be viewed as an elaboration of the effects of certain ecological variables within Berlyne’s framework, with the caveat that many of the effects appear to show monotonic increases with prototypicality rather than the inverted-U-shaped function that Berlyne’s theorizing implies. The more recent ecological effects described above for preference effects due to canonical position in the world (Sammartino & Palmer 2012a), canonical perspective (Palmer et al. 1981, Sammartino & Palmer 2012b), and canonical size (Konkle & Oliva 2011, Linsen et al. 2011) can all be interpreted as supporting the importance of prototypes, albeit with respect to viewpoint-based features of images of objects relative to viewers rather than to category membership. Nevertheless, even preferences for both prototypical examples and canonical features seem too limited to be taken as a general theory of aesthetic response. They are more like ubiquitous factors that influence aesthetic preference. Moreover, prototype theory, by itself, does not clarify why prototypes should be preferred.

7.4. Fluency Theory
An interesting answer to the “why” question is provided by fluency theory, which is perhaps the single most general explanation of aesthetic preference (e.g., Reber 2012; Reber et al. 1998, 2004). Fluency theory posits that people prefer visual displays to the extent that they are processed more easily (or fluently). It does a good job of predicting aesthetic effects due to many low-level features (for a review, see Oppenheimer & Frank 2008), such as preferences for larger (Silvera et al. 2002), more symmetrical (Jacobson & Höfel 2002), more highly contrastive displays (Reber et al. 1998). It can also explain categorical prototype effects (see above) because prototypes are well known to be more easily and quickly processed than nonprototypical examples (e.g., Rosch 1975). The kinds of spatial composition effects described above for untitled, single-object pictures—namely, the fact that people tend to prefer the focal object to be at or near the center of the frame, to face inward both horizontally and vertically, and to be depicted at a height, perspective, and size within the frame that reflect their characteristic height, perspective, and size relative to a standard observer (Palmer et al. 2012b; Sammartino & Palmer 2012a,b)—are also consistent with perceptual fluency. The reason is that these compositional choices tend to make the object most recognizable as the kind of object it is (cf. Estes et al. 2008, Palmer et al. 1981) and therefore lead to the easiest (most fluent) processing for the purpose of object
identification. These compositional choices amount to selecting the most likely (or “default”) settings of the spatial variables associated with the representational schema for that kind of object (cf. Palmer 1975, Palmer et al. 2012b). Finally, fluency theory even provides a plausible causal explanation of mere exposure effects: The more often a given image is seen, the more easily and fluently it will be processed, and, by hypothesis, increased fluency leads to increased preference.

One problem with fluency theory is that it does not square well with Berlyne’s inverted-U results for preference as a function of complexity. More complex stimuli presumably always require more processing than simpler ones, so it seems that fluency theory should predict a monotonic decrease in preference as a function of complexity. The typical result, however, is an initial increase in preference with greater complexity, until some optimum level is reached, after which a decrease with additional complexity is found (see Berlyne 1971 for a review). It is also unclear how fluency would account for preferences for single colors, especially in light of the strong evidence of links to preferences for ecological objects (Palmer & Schloss 2010).

A deeper challenge to fluency theory as a complete account of aesthetic response is that its basic premise—that the most easily processed image will be the most aesthetically pleasing one—seems to contradict a central tenet of art theory: namely, that the aesthetic (or perhaps the artistic) impact of an image is increased by intentionally violating the viewer’s expectations. Indeed, the most fluent image of an object would seem to be the most boring image of it—analogous, perhaps, to the formulaic paintings created by Komar and Melamid as the most pleasing pictures for people in different cultures (Wypijewski 1997)—whereas aesthetically pleasing images are expected to be more interesting. Indeed, much of the recent history of art, from Impressionism onward, can be understood as a continual process of violating the rules and conventions of prior art practice. It is presently unclear how one should view the relation between art history and aesthetics, however. Perhaps art has simply become increasingly dominated by novelty and the quest to conceptually stretch the boundaries of art to the detriment of the aesthetic response such art evokes in many or most viewers.

An equally troubling, and possibly related, problem is that fluency theory seems to leave out a crucial variable in aesthetic judgment: namely, the meaning or message of the image as intended by the artist and/or as inferred by the viewer (which assuredly may not be the same). It is perhaps such semantic considerations that lie, in part, behind Kant’s suggestion that aesthetic judgments necessarily involve “free play of the imagination” in the beholder, and it is disturbing that such considerations seem to be either irrelevant or possibly even contradicted by perceptual fluency accounts. More recently, fluency theory has been extended to conceptual (rather than perceptual) fluency in an attempt to account for effects of such higher-level semantic variables (Reber et al. 2004, Whittlesea 1993). The common principle is that fluent (easy) processing is aesthetically valued in both perceptual and conceptual domains.

Sammartino & Palmer (2012b) have proposed an alternative to fluency theories of aesthetic response to pictures of objects and scenes based on their concept of representational fit. The suggestion is that people prefer images to the extent that their spatial composition optimally conveys an intended or inferred meaning of the image, thus including a semantic variable that is missing from fluency theory. Interestingly, it also makes the same predictions as fluency when the message intended or inferred for the image is simply the default one: to depict the focal object(s) or event(s), as it typically is in stock photography. Under such circumstances, the best horizontal and vertical position, size, and perspective of the object(s) will be those that make it optimally recognizable. When the intended or inferred meaning is different, however, as it is with the elaborated titles in the studies described above, the otherwise non-preferred compositional choices of position, perspective, and size that are most compatible with the message can produce the highest
aesthetic response in viewers (Sammartino & Palmer 2012b). Fluency and representational fit may also be viewed as compatible accounts, however, differing primarily in emphasis, with fluency theory focused on the ease of processing and representational fit theory focused on the reasons that processing is easy.

7.5. Multicomponent Theories of Aesthetic Response to Art

Although the focus of this review has intentionally not been on aesthetic responses to viewing art objects, it is useful to see how the kind of work reviewed above fits within theoretical frameworks designed to account for the aesthetic appreciation of art. Shimamura (2012), for example, has recently presented a very general approach to understanding people’s aesthetic responses to art. To the extent that part of people’s appreciation of art concerns aesthetic dimensions of their experiences, it is presumably relevant. Shimamura’s framework is summarized by its acronym, I-SKE, which stands for the intention (I) behind the artwork in relation to the viewer’s sensory (S), knowledge-based (K), and emotional (E) responses. I-SKE is thus a very general framework that seems to encompass much of what matters in people’s aesthetic responses to art objects.

The research reviewed in this article generally concerns the aesthetic effect of sensory-perceptual information and its relation to knowledge-based components. Most of the biases identified in preferences for spatial composition, for example, are knowledge based. Surprisingly, knowledge is relevant even for simple color preferences, because they turn out to depend strongly on people’s stored knowledge of their affective responses to the objects that are characteristically associated with those colors (Palmer & Schloss 2010).

Leder and colleagues (2004) have formulated a detailed information-processing model that attempts to analyze and synthesize these many diverse aspects of people’s appreciation of art, and especially of modern art. It consists of a series of five processes, each of which is influenced by many factors. The five processes are proposed to occur in the following order: (a) perception, which responds to stimulus factors, such as complexity, symmetry, color, contrast, and organization; (b) implicit classification, which involves integrating this perceptual information into related information stored in memory concerning familiarity, prototypes, and conventions; (c) explicit classification of the artwork in terms of its style and content; (d) cognitive “mastering” of its style and content by interpreting them within one’s knowledge about related art and with respect to the viewer’s self; and finally (e) evaluation of satisfaction in terms of both the work’s cognitive aspects (e.g., understanding of meaning and ambiguity) and the cumulative influence of all the stages of processing on the viewer’s affective state. Of the stages, the first two are deemed to be automatic, the last two to be deliberate, with explicit categorization being a mixture of both. In addition to these processes, however, there are many contextual considerations that influence the outcome of the processing, including the situation in which the viewer finds the artwork (museum, gallery, psychology experiment, etc.), what previous experiences the viewer has had with similar artworks, what social and linguistic information the viewer has surrounding the current experience, and what his/her affective state is coming into the current viewing experience.

Comparatively less attention has been paid to emotional components of aesthetic response. However, Silvia’s recently proposed “appraisal theory” of aesthetic response is based primarily on analyzing the diverse emotional responses a viewer might have to art objects (e.g., Silvia 2005a,b, 2006, 2012). He divides emotions into several kinds—positive emotions (happiness, enjoyment, pleasure),

Silvia does not mention negative emotions (sadness, misery, pain) as a category, but perhaps he intends them to be implied by virtue of being the inverses of the positive emotions.
knowledge emotions (surprise, interest, confusion), hostile emotions (anger, disgust, contempt), and self-conscious emotions (pride, shame, guilt, embarrassment)—and argues that all are relevant to understanding aesthetic responses to art. Positive emotions are classically emphasized in most aesthetic research (see section 2.1), but Silvia argues that the other classes of emotions are indeed relevant to people’s reactions to art. This is especially true for modern art, where many artists clearly intend their works to elicit surprise, interest, anger, disgust, shame, and/or guilt rather than pleasure.

8. CONCLUSIONS AND ISSUES FOR FUTURE RESEARCH

We have attempted to summarize, critique, and synthesize the emerging scientific literature concerning human aesthetic response to and preferences for visual stimulation using behavioral methods. We began with the claim that this body of research exemplifies a branch of aesthetic science from a psychological perspective. We submit that the results we have described constitute clear advances in the understanding of this domain. We hope to have convinced the reader that meaningful research in this field is not only possible, but also interesting, important, and fun.

Nevertheless, many critical issues need to be addressed in future research. One is how to connect this body of work with the ongoing literature on behavioral studies of the perception of art. As we said at the outset, aesthetic response and perception of art are conceptually distinct, but overlapping, domains. Can aesthetic preferences for simple features such as individual colors, color combinations, form, texture, and spatial composition somehow be combined to understand the far more complex appreciation of fine art, or do these features form truly emergent wholes (Gestalts) that transcend understanding in terms of their simpler components? Needless to say, this is a daunting task, but there are some bodies of artwork that lend themselves well to beginning to answer the question, such as Piet Mondrian’s “grid” paintings in red, yellow, and blue and Josef Albers’s Homage to the Square series.

A second critical issue is how to extend the present body of research more fully into the burgeoning topics of emotion and cognitive meaning. Much of the research described above concerns what most would call purely formal characteristics of aesthetic appreciation. Indeed, this is why the art of Mondrian and Albers is closer to the reviewed research than that of, say, Monet, Picasso, or Duchamps. Formal aspects are clearly important, but they are just one part of what one wants to know about aesthetic response. Although some inroads have been made recently in this direction both theoretically and empirically, the field is wide open for more systematic research that builds on the increasingly sophisticated understanding of both topics that has arisen in psychology.

A third critical issue, not unrelated to the first two, is how to connect behavioral research on aesthetic preferences with the emerging field of neuroaesthetics. As we also said at the outset, understanding the neural correlates of aesthetic response depends fundamentally on the development of behavioral methods and results. As this behavioral foundation becomes firmer, more precise and detailed studies of the brain regions that produce aesthetic responses can be undertaken with considerable prospects of success. We expect that eventually there will be a synergy between behavioral and neural studies of aesthetics that will help both enterprises reach the common goal of understanding visual aesthetics and human preference.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.
ACKNOWLEDGMENTS
The preparation of this article was supported in part by the National Science Foundation under grants #0745820 and #1059088 and a Google Gift to S.E.P. Any opinions, findings, conclusions, and/or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation or Google.

LITERATURE CITED
Allesch GJv. 1924. Die aesthetische Erscheinungsweise der Farben. Psychol. Forsch. 6:1–91
Angier RP. 1903. The aesthetics of unequal division. Psychol. Rev. Monogr. Suppl. 4:541–61
Fechner GT. 1876. Vorschule der Aesthetik. Leipzig, Germany: Breitkopf & Härtel
Fernandez D, Wilkins AJ. 2008. Uncomfortable images in art and nature. Perception 37:1098–113
Haines TH, Davies SE. 1904. Psychology of aesthetic reaction to rectangular forms. Psychol. Rev. 11:249–51
Moon P, Spencer DE. 1944a. Aesthetic measure applied to color harmony. J. Exp. Psychol. 34:234–42
Strauss ED, Schloss KB, Palmer SE. 2012. Exposure to colored objects influences color preferences. Manuscript submitted
Taylor C, Schloss KB, Palmer SE, Franklin A. 2012b. Adult color preferences are not evident in infants’ responses to color. Manuscript submitted