Preference for Art: Similarity, Statistics, and Selling Price

Daniel J. Graham[*a], Jay D. Friedenberg[b], Cyrus H. McCandless[c] and Daniel N. Rockmore[a,d]
[a]Department of Mathematics, Dartmouth College, Hanover, NH, USA 03755;
[b]Department of Psychology, Manhattan College, Riverdale NY, USA 10471;
[c]Integrated Research, Merchant Mechanics, Inc., West Lebanon, NH, USA 03784;
[d]Department of Computer Science, Dartmouth College, Hanover, NH, USA 03755

ABSTRACT

Factors governing human preference for artwork have long been studied but there remain many holes in our understanding. Bearing in mind contextual factors (both the conditions under which the art is viewed, and the state of knowledge viewers have regarding art) that play some role in preference, we assess in this paper three questions. First, what is the relationship between perceived similarity and preference for different types of art? Second, are we naturally drawn to certain qualities—and perhaps to certain image statistics—in art? And third, do social and economic forces tend to select preferred stimuli, or are these forces governed by non-aesthetic factors such as age, rarity, or artist notoriety? To address the first question, we tested the notion that perceived similarity predicts preference for three classes of paintings: landscape, portrait/still-life, and abstract works. We find that preference is significantly correlated with (a) the first principal component of similarity in abstract works; and (b) the second principal component for landscapes. However, portrait/still-life images did not show a significant correlation between similarity and preference, perhaps due to effects related to face perception. The preference data were then compared to a wide variety of image statistics relevant to early visual system coding. For landscapes and abstract works, nonlinear spatial and intensity statistics relevant to visual processing explained surprisingly large portions of the variance of preference. For abstract works, a quarter of the variance of preference rankings could be explained by a statistic gauging pixel sparseness. For landscape paintings, spatial frequency amplitude spectrum statistics explained one fifth of the variance of preference data. Consistent with results for similarity, image statistics for portrait/still-life works did not correlate significantly with preference. Finally, we addressed the role of value. If there are shared "rules" of preference, one might expect "free markets" to value art in proportion to its aesthetic appeal, at least to some extent. To assess the role of value, a further test of preference was performed on a separate set of paintings recently sold at auction. Results showed that the selling price of these works showed no correlation with preference, while basic statistics were significantly correlated with preference. We conclude that selling price, which could be seen as a proxy for a painting's "value," is not predictive of preference, while shared preferences may to some extent be predictable based on image statistics. We also suggest that contextual and semantic factors play an important role in preference given that image content appears to lead to greater divergence between similarity and preference ratings for representational works, and especially for artwork that prominently depicts faces. The present paper paves the way for a more complete understanding of the relationship between shared human preferences and image statistical regularities, and it outlines the basic geometry of perceptual spaces for artwork.

Keywords: Perception, human vision, art preference, aesthetics, efficient coding, artwork, selling price, stylometry

1. INTRODUCTION

The question of what contributes to human judgments of preference remains controversial, despite years of study. With respect to vision, many attempts have been made to devise a "formula" for images that will be preferred, starting with Fechner's investigations of the golden ratio over a century ago. Though fully elucidating the components of visual aesthetics is perhaps an unattainable goal, it is appealing to imagine that the right combination of empirical data about art images—both quantities intrinsic to the image as well as extrinsic properties—could at least help guide typical viewers to images they will prefer. That is, by understanding shared biases in preference due to visual processing, along with

[*] daniel.j.graham@dartmouth.edu; phone 1 603 646-9020; fax 1 603 646-1312; math.dartmouth.edu/~dgraham

Human Vision and Electronic Imaging XV, edited by Bernice E. Rogowitz, Thrasyvoulos N. Pappas,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7527, 75271A · © 2010 SPIE-IS&T · CCC code:
0277-786X/10/$18 · doi: 10.1117/12.842398

emergent indicators of group preference (which one might assume reflect efficient competition in the broad "market" of art images), we could potentially develop an initial model for predicting preference for art. If successful, such models could be applied to images more generally, since it has been argued that art as whole, having been created by hand specifically for human viewing, might be expected to reflect essential features of human visual processing[1-6]. That is, by understanding preference for art, and by invoking the notion that artists strive to efficiently match their images to visual system processing strategies, we may come to understand shared aspects of visual preference itself more completely.

Here, we investigate the extent to which preference judgments are independent of other perceptual dimensions such as similarity, and the extent to which there exist regularities in paintings that are related to preference. Of course, semantic and social factors surely play important roles in preference. The presence of human faces, for example, likely introduces a host of biases. After separating paintings into three broad content categories, we test relationships between preference and image statistics previously shown to explain large portions of the variance in other perceptual judgments. This work shows that content is indeed confounding, but also that for certain types of images, image statistics can be quite useful for understanding preference.

While semantic factors (and especially emotional evocations) are difficult to quantify except in broad terms given the great number of dimensions over which they can vary, social factors may be more easily reducible to empirical measures. If one assumes that the net effect of a shared human aesthetics is reflected in the worth placed on an artwork by society, one might expect an artwork's value to align with its aesthetic appeal. If this were the case, it would indicate that even high-level properties of artwork such as value are influenced by the artist's efforts to appeal to aesthetics. On the other hand, economic factors could have an influence on value that overwhelms shared visual preferences. Therefore, we ask whether the perception of preference for art aligns with social measures of worth. Together, this work helps to elucidate what aspects of preference are universal, and the extent to which emergent measures of value track visual preference.

In Experiment 1, we investigate the relationship between perceived similarity and preference for three classes of paintings differing in terms of general content. This study is designed to give some insight into the effect of semantic content on relationships between basic perceptual judgments. We find that landscapes and abstract works show strong correlations between similarity and preference judgments, while images with prominent human faces do not show such correlations.

In Experiment 2, we consider the relationship between preference and image statistics relevant to early vision coding. While we expect that simple statistical measures will succeed in explaining substantial portions of data variance in preference rankings, we are also interested in understanding when such measures fail. Consistent with the results of Experiment 1, we show that the ability of coarse image statistics to correlate with preference is related to image content. Nevertheless, our simple statistical measures do find some success in explaining preference, and we believe that more successful models of preference can be constructed using probabilistic methods that exploit image statistics (as well as content tags and perceptual data such as similarity judgments), at least for some types of art. In particular, preference for abstract art could be amenable to prediction based largely on image statistics.

In Experiment 3, we investigate one prominent extrinsic factor that is often associated with preference, namely monetary value. We test preference of untrained observers for artworks recently sold at auction in order to assess whether individual biases in preference align with social biases in valuation. Results indicate that selling price shows no correlation with preference. However, we find that image statistics such as those found to be useful in explaining preference in Experiment 2 are correlated with preference for a new data set.

In general, we suggest that social context appears to act independently in assessing value of aesthetic objects, whereas individual humans show some regularity in both the relationship between similarity perception and preference, and in the relationship between image statistics and preference. However, the diversity of image content makes prediction of preference difficult, even within the broad content categories we test. But images lacking representational content could help reveal shared properties of preference. To the extent that universal aspects of compositional aesthetics exist, these might most profitably be investigated in non-representational art, as is suggested by our results.

We conclude that researchers may one day have a good understanding of the mechanisms of preference as they operate in individuals, and sophisticated models could be deployed to successfully guide viewers of large art databases to images they will prefer. However, emergent dynamics of preference (as reflected in value) in the context of group behavior may be dominated by factors besides those related to physiological aspects of human visual processing, and thus may be difficult to explain or predict in a generalized fashion. We begin with a cursory review of the literature for image preference.

2. BACKGROUND: STIMULUS FACTORS INFLUENCING IMAGE PREFERENCE

In the literature, a number of different stimulus features have been found to affect image preference. Following Fechner's work on the aesthetics of the golden ratio, which was later refuted, researchers focused a great deal of attention on complexity. Investigators often varied the complexity of simple figures like polygons and other geometric shapes to test relations with preference,[7,8] while some researchers also investigated the relation between complexity and aesthetic preference for more naturalistic stimuli such as landscape and abstract paintings.[9-11] Work in this spirit has continued to the present day.[12] However, there are problems in interpreting this research, since results differ from study to study. Some obtained a linear increase in preference with complexity, while others found an inverted-U shaped function where preference peaks at intermediate levels of complexity. A major concern is that complexity can be defined or manipulated in a variety of different ways from number of elements to information content.[13]

Other approaches to visual aesthetics have focused on symmetry and compositional balance, both of elements within paintings and of colored regions.[14-16] In this theme, researchers have examined preference for the positioning and facing direction of objects within rectangular frames.[17] Participants in this study preferred front-facing objects to be located at or near the center of the frame and left- or right-facing objects to face into rather than out of the frame. These findings were obtained using several different methods including two-alternative forced-choice, method of adjustment and in free choice photographs.

One interesting finding in the literature is that the more similar an image is to the observer, the more attractive it will be judged. Penton-Voak and colleagues created opposite sex facial images of participants and varied their similarity.[18] Attractiveness judgments increased with greater similarity but preference for average faces was stronger than preference for self-similar faces. Alexander and Marks hypothesized that viewers would find paintings attractive the more similar they were to both real and idealized versions of themselves.[19] The ratings data they obtained supported this notion, with a stronger relation between liking and the ideal self.

Work by Berlyne focused on the effect of culture, and this work concluded that there are relatively small differences between Eastern and Western subjects' similarity and preference ratings of Western art.[20] Other studies have generally confirmed this view, and they have shown similarly modest differences in basic perceptual judgments between experts and non-experts. However, Berlyne found that both similarity and preference depend somewhat on how representational and how bright the works are.

In the following experiments, we attempt to broaden the perspective of this field by investigating three aspects of human preference for art: perceived similarity, image statistics, and selling price. As visual objects that are generally designed to simulate aspects of the visual world, art has the potential to serve as a testing ground for more general theories of preference.

3. EXPERIMENT 1: PREFERENCE AND PERCEIVED SIMILARITY

Imagine a toy world where all visual stimuli vary along a single dimension, e.g., sine-wave gratings of a single orientation that vary in spatial frequency. Assuming one has normalized for contrast sensitivity and if one assumes linearity, one can make a basic assumption about the relationship between similarity and preference, namely that these judgments are likely to be linearly related. In other words, if humans consider two gratings—A and B—to be similar to one another, and humans prefer grating A to some other grating, C, which is judged to be not similar to either A or B, then we should find that grating B is preferred to grating C.

But what about artwork, which varies along a great many dimensions? Linearity is a basic assumption regarding relationships among perceptual dimensions, and in particular about how the human brain assesses similarity.[◊] Many models of object recognition require that perceptual spaces be linear. Perceived similarity for a given class of real-world objects—and the formal transformations necessary for relating members of that class—have been proposed as the general neural processing strategy for object classification.[21] Moreover, a standard modeling problem in natural vision is

[◊] It should be noted, however, that there is active debate about the respective roles of perception and cognition in assessing similarity.

to predict human similarity ratings of standard stimuli (e.g., Brodatz textures) using a collection of image statistics.[22] These models are generally linear in nature. One can suppose that visual stimuli live in a high-dimensional, linear state space—that is, each stimulus has some distance to all other stimuli, which determines perceived similarity.[§] Indeed, one proposal in this realm argues that the brain in fact represents only these similarity relationships, and not the set of features associated with any given stimulus.[23] Preference can be overlaid onto this space (assuming linearity): preference is simply a vector pointing in a certain direction. In this view, stimuli that are neighbors when projected onto this vector will be judged to have similar preference, while stimuli that have different projections will have different preferences.

There is reason to believe this can be a fruitful approach, at least in relation to visual stimuli. Much is known about the relationship between preference and similarity for certain stimuli, and especially for faces. In the case of faces, the vector of preference points towards the center of the similarity space, i.e., towards the average face. That is, degree of preference is aligned primarily with similarity to the mean, not to other similarity relationships (including similarity to one's own face, as noted above). However, for natural scenes, which have little structural regularity (i.e., they have little canonical form, except for sky/ground relations, but do contain a host of statistical regularities), the relationship between similarity and preference is less clear. Here we propose that examining art is a way to narrow the focus to only those images that are *meant* to be viewed by other humans. Though the matter of style can confound the analysis,[2] our previous work[1-3,6] and that of others[4,5,25] shows that art is quite relevant to natural vision, and notably, to aesthetics.

3.1 Methods

We used a paired comparison paradigm to assess preference within three groups of painting types, namely landscape, portrait/still-life, and abstract paintings. Images were classified into these groups using a prior three-alternative forced choice test.[2] Preference data were collected for each image set using a different group of observers, all of whom were undergraduate students from Manhattan College: 10 subjects evaluated 20 images of landscape paintings; 8 subjects evaluated 20 portrait/still-life paintings, and 15 subjects rated 18 abstract paintings.

Images were drawn from the digital collection of the Herbert F. Johnson Museum of Art, Cornell University, and they span a diverse range of contents, styles and provenances (see cited references for more details on the image collection[1,2]). Each subject was presented with all image pairs (190 pairs for landscapes and portrait/still-life, 153 for abstracts) in random order. Image pairs were presented at roughly .5 m on 17-inch computer displays surrounded by a neutral grey background. Both images in each pair were scaled to roughly the same horizontal size and randomly placed on the left- or right-hand side. Images were displayed such that they were centered vertically and their horizontal centers were equidistant from the center of the screen. Observers were asked to choose which image they preferred and to rate their degree of preference by inputting integer ratings from 1 (prefer a little) to 9 (prefer a lot) on a computer keyboard for each pair. The test was self-paced and participants were permitted to change their response if desired before proceeding to the next trial.

Note that all behavioral experiments described in this paper were carried out using "lay" observers (viewers lacking substantial expertise in art), and all images tested were extremely unlikely to have been viewed previously by participants. This was done to ensure that connoisseurship and familiarity played no role in perceptual judgments.

3.2 Results

As in similar studies by other groups,[17] we used two methods for computing preference rankings: average probabilities of preference calculated across participants and pairs,[‡] and a Bradley-Terry-Luce probabilistic model of preference. Both ranking methods were found to be in strong agreement with each other, showing typical Pearson correlations of r = .86 or greater in all experiments (all significant at the $p < 0.05$ level). Preference data (i.e., which image was preferred) were analyzed with and without degree-of-preference data, and results were found to be nearly identical for these two conditions across all experiments described in this paper. For brevity, we present only our results for averaged preference probabilities, and we omit degree-of-preference data.

We compared the preference rankings to previously collected similarity data for the same images using different observers drawn from the same subject pool.[6,26] Details of the methods and results for similarity judgment experiments

---

[§] Work is ongoing to determine the appropriate distance metric for the perceptual space of images. For example, Russell and Sinha found that "city-block" distance in terms of pixel similarity is a better predictor of preference than is Euclidean distance.[24]
[‡] This can be seen as akin to adding a 1 to entry p_{ij} in a N x N preference matrix, where N is the number of images, each time image i is preferred to image j. Matrices are averaged across participants, and then row-wise.

can be found in the cited references; briefly, the procedure for similarity judgment was much the same as the test in Experiment 1 except that participants rated only the degree of similarity (1 to 9 scale) for each pair. Multidimensional scaling (MDS) was applied to similarity ratings averaged across participants. We compared the first two MDS dimensions for each image set (which together explained 35-40% of overall variance in similarity data) to corresponding preference rankings.

For abstract art, we found that MDS dimension 1 of similarity was strongly correlated with preference (R^2 = .397 Pearson's r = -0.630; p = 0.0051). Given that the first MDS dimension explained 21% of similarity data variance, this result implies that around 10% of the variance in preference can be explained by similarity judgments. MDS dimension 2 showed no correlation with preference.

For landscapes, only MDS dimension 2 of similarity was strongly correlated with preference (R^2 = .374, Pearson's r = -0.612, p = 0.0042), with preference explaining around 5% of overall similarity data variance. For portrait/still-life works, neither of the first two MDS dimensions was correlated with preference rankings.

These results are consistent with those of our previous experiments showing that similarity judgments are strongly influenced by both image content and style for portrait/still-life paintings, and are somewhat influenced by the presence of humans in landscapes.[6] In particular, it was found previously that MDS dimension 1 of similarity perception for portraits was strongly correlated with provenance (i.e., in which hemisphere the work was produced), and MDS dimension 2 for portrait/still life paintings was strongly correlated with the presence of people. For landscapes, MDS dimension 1 was strongly correlated with the presence of human forms. Similarity for abstract works, on the other hand, proved to be more predictable in terms of basic image statistics.[26] The results of these previous experiments, along with the results of Experiment 1, are used to motivate further exploration of image statistics and their relation to preference in Experiment 2.

3.3 Discussion

Together, these results suggest that nonlinearities in the primary relationships between similarity and preference due to semantic content are most apparent in portraits, less so in landscapes, and less still for abstract works. Our results suggest further that "intrinsic aesthetics" (the kind associated with visual patterns lacking human forms) may operate in a different fashion from aesthetics associated with human forms. Moreover, as will be suggested in work described below, non-representational art possesses properties useful for the analysis of relationships between regular statistics and aesthetics. First, it possesses spatial structure but generally lacks recognizable objects (especially human forms) and is therefore usually free of confounding perceptual dimensions. Second, given the widespread acceptance of abstract art in the art world, one may expect that many of the same compositional and aesthetic imperatives that guide representational art are also present in abstract art. While some abstract works do not strive for aesthetic appeal, abstract art in general would be expected to contain regularities that map to shared aspects of visual preference.

4. EXPERIMENT 2: PREFERENCE AND IMAGE STATISTICS

Much evidence supports the notion that basic statistical regularities that are "diagnostic" of scene attributes could in principle be extracted by the early visual system. Studies of categorization for natural images have demonstrated relationships between basic spatial statistics (e.g., spatial frequency amplitude) and classes of natural scenes,[27] and between basic intensity statistics and perceptual judgments.[28] These same statistics are relevant to efficient visual coding in primates. Our previous work has suggested that sparseness (i.e., non-gaussianity of response in a given representation for a given class of images[29]) is likewise informative in terms of perceptual judgments for art. As shown in past work, though perception of similarity is weakly related to sparseness for portraits (and presumably for faces in general), the influence of sparseness is apparent for landscapes (and presumably natural scenes) and, most prominently, for abstract works.[6,26] For the latter, which lack human forms and most any content, humans perceive similarity in a way that aligns somewhat with variations in basic statistics of scenes. Moreover, in our studies, sparseness is only weakly correlated with spatial frequency amplitude spectrum slope. In a linear space of similarity perception, then, sparseness and amplitude spectrum slope could comprise two principal axes of the space. This is sensible because these statistical properties are tied to separate efficient coding strategies in the early visual system (amplitude spectrum regularities have been tied to retinal ganglion cell coding, while sparseness has been tied to V1 simple cell coding; see[30] for a review). Intuitively, how sparse your environment is can be quite influential in your perception of how similar that environment is to other environments. Because the visual system is in some ways built around the assumption of sparseness, it follows that gross changes in this statistical regularity will result in changes in perception. One would expect the same of amplitude spectra, as has been shown in past work by Torralba and Oliva.

Following on the results of Experiment 1 showing correlations between similarity and preference for some art types, we would expect preference to also be related to sparseness and amplitude spectrum slope, at least for abstract art.

4.1 Methods

In this experiment, we compare the preference data collected in Experiment 1 to a host of basic image statistics relevant to early visual coding, which have been examined in the past.[6,26] Given the past success of two measures, namely spatial frequency amplitude statistics and pixel intensity sparseness, in explaining similarity judgments, we focus on these statistics. Briefly, these two statistics gauge separate (though not independent) image dimensions: the spatial frequency amplitude spectrum (in particular, the slope of the spectrum plotted in log-log space) describes the proportion of large scale spatial structure and fine detail in an image, while pixel sparseness (in particular, the "activity fraction"[31]) describes the extent to which intensity distributions are non-Gaussian. Activity fraction sparseness is calculated as the square of the sum of pixel values divided by the sum of the squares of pixel values; it ranges from 0 (highest sparse) to 1 (lowest sparseness).

To ensure that images used were not subject to artifacts due to image acquisition and compression, we employed uncompressed images that were standardized in terms of acquisition, and which were found to be linear in terms of luminance.[1]

4.2 Results

For abstract works, pixel sparseness was significantly correlated with preference and accounts for 25.0% of preference data variance (Pearson's r = -0.477; p = 0.045). For landscapes, the amplitude spectrum slope was significantly correlated with preference (R^2 = .208; Pearson's r = -0.445; p = 0.049). In addition, if we compare preference to a statistic composed of the sum of the amplitude spectrum slope and the activity fraction, we find that this statistic can explain 24.3% of landscape data variance (r = -0.493; p = 0.027). More modest gains in correlation are observed for this statistic over sparseness alone for abstract art. Portrait/still life paintings showed correlation between preference and amplitude spectrum slope, but it did not rise to the $p < 0.05$ level of significance (Pearson's r = -0.419; p = 0.066).

4.3 Discussion

First, we note that we focus on just two statistics, one spatial statistic and one intensity statistic, though we acknowledge that color, and, in turn, color statistics, play an important role in the perception of art images (see, e.g., work by Wallraven and colleagues[32]). For example, higher color diversity in art due to different simulated lighting conditions has been shown to be preferred to lower color diversity.[33]

Consistent with our results in Experiment 1 and in previous studies of similarity, we conclude from these results that both abstract paintings and landscapes are quite amenable to study in terms of image statistics. Indeed, the ability to predict around a quarter or more of preference data variance for these image categories suggests that basic spatial and intensity statistics (especially those that have relevance to visual coding) may play a role in human judgment of aesthetics.

However, altogether different mechanisms could be at work when human forms—and especially faces—are prominent in a painting. In this case, emotional, social, and contextual cues may obscure underlying biases in visual preference. Indeed, the literature does suggest that faces are special. The well-known finding is that inversion disrupts recognition for faces but not for other comparable stimuli. The explanation is that face perception is configural, i.e., it encodes relations between features that are disrupted by inversion, while "normal" perception is more analytical, taking only individual features. One could assess the idea of separate systems of aesthetics by testing similarity and preference judgments for inverted portraits. Since the image statistics we measure are not affected by inversion, we would expect that configural information—and related emotional cues—would play a less prominent role in similarity and preference judgments, and that image statistics may therefore emerge as stronger correlates of these judgments for inverted faces compared to upright faces. If this were the case, separate models could be constructed for art images with faces, and without faces: for images with faces, models based on face averageness could be employed, while for images without faces, basic image statistics such as those tested here (along with color statistics), could be used.

We note that a number of past studies have investigated basic statistics in art and their relation to perception.[1-6] This work builds on findings in visual neuroscience showing that vertebrate visual system coding is shaped by statistical regularities in the visual environment, a notion termed the efficient coding hypothesis.[34] One can extend this idea to artwork, which is itself representative of the environment, and specifically of the aspect of the environment of greatest interest to humans. A general conclusion of this work is that artists include many of the same regularities found in natural scenes in their works, and it has been suggested that this is done in order to effectively stimulate the human visual system.[1] However, some regularities, such as those related to luminance, are drastically transformed by artists.[2]

Some work in this area has specifically addressed the role of statistics in aesthetics. Redies and colleagues found that artists portray human faces using more shallow amplitude spectra than those that are possessed by images of actual faces.[25] Redies argues that because complex natural scenes have similarly shallow spectra (falling as approximately 1/f), this finding constitutes evidence that artists modify the statistics of faces in order to achieve "aesthetic resonance" with the human visual system. That is, since visual coding is adapted to the basic statistics of natural scenes, such a shift for faces could suggest that artists perform this shift to achieve an aesthetic result. However, it should be noted that cross hatching and other fine details that are typical in the images Redies studied (a collection of etchings, engravings, woodcuts, lithographs, charcoals, and pencil, pen, and brush drawings,) could have biased the amplitude spectrum calculation. Note for example that those techniques that lack strong cross-hatching and have larger typical stylus widths (charcoals, brush drawings and woodcuts) were closest in terms of amplitude spectrum slope to actual faces, while techniques that employ cross-hatching and have smaller stylus widths were furthest away (pencil and pen drawings, lithographs, engravings and etchings; paintings were not tested). Therefore, materials must be taken into consideration when assessing these results. Given that it remains an open question why artists would as a group change the basic spatial statistics of faces, one could argue conversely that if paintings more closely match the basic statistics of human faces, such images would be better objects of study in terms of human aesthetics compared to media that utilize a great number of fine marks (and especially cross-hatching).

Alternative proposals regarding the relationship between statistical regularities and artwork have been offered.[1] A notion related to the efficient coding hypothesis termed the "efficient artist hypothesis" proposes that since art is made expressly for viewing by the human eye, it collectively may contain regularities above and beyond those found in scenes, including ones that are related specifically to *human* visual processing.[35] For example, whatever regularities are associated with human recognition of general categories (such as "landscape") and of specific objects or scenes, must in this view be well isolated in art. Consider that artists create works by hand and are therefore able to manipulate every element of their visual representation to suit their needs—in a way difficult or impossible in photography, even with advanced image processing techniques. As a result, artists as a group may select only the minimum features necessary to elicit a given perception, and may also follow universal aesthetic "rules" of composition for these features. To the extent that we can gauge such "rules" using basic statistics relevant to vision coding, we may find that we can predict preference and other perceptual judgments. The results of the present experiment show that this is indeed a possibility.

5. EXPERIMENT 3: PREFERENCE AND VALUE

A final consideration that has so far been unexplored in the art preference literature is the role of a work's monetary value. To our knowledge, no behavioral tests have been undertaken to assess whether price is predictive of preference. We have performed an initial experiment to address this question. We acknowledge that selling price is affected by numerous factors–for example, the formation of artistic canons plays a role in determining which works sell for the most money,[36] and contemporary art prices can be biased by current fashion, as well as other factors. Also, one must also acknowledge that art of the modern era, and especially abstract works, are not necessarily intended to be aesthetically pleasing. However, it seems reasonable that efforts to appeal to a shared human aesthetics are responsible for at least part of a work's value. Bearing in mind that connoisseurship and other factors bias selling price, one might expect that selling price would explain at least some portion of the variance of human preference ratings. One might even expect that price is a better indicator of preference compared to basic image statistics, such as those tested in Experiment 2.

5.1 Methods

We gathered a new set of images of paintings recently sold at auction (N=40), along with their selling price (in US Dollars). We asked non-expert observers (N=17, all undergraduate students at Manhattan College) to rate preference in paired comparisons as in Experiment 2. Initially, 300 images labeled "Paintings" were chosen pseudorandomly using the search function of the auction house website. Forty of these images were selected for testing. Works span five centuries and include paintings by Old Masters, works by Georges De La Tour, Diego Rivera, Roberto Matta, and Milton Avery, as well as a number of contemporary Chinese paintings. Images were downloaded from the auction house website and were converted to greyscale, normalized by the mean intensity of the group (to eliminate effects due to differences in overall brightness), and presented in the same manner as in Experiments 1 and 2. Because of the increased number of images, and in order to test all pairs, we randomly assigned each possible pair (780 total) to one of three ordered sets. This was done a total of 5 times to generate 15 possible random orderings, each of which had approximately uniform probabilities of containing each pair. Each participant was randomly presented with one of these 15 random orderings (i.e., each participant viewed 260 pairs).

5.2 Results

We found that selling price showed no significant correlation with preference (Pearson's $r = -0.139$; $p = 0.39$). This finding could in theory be due to inefficiencies in the market, or it could indicate that price largely reflects extrinsic attributes of paintings, rather than the visual aspects of the image itself.

However, preference was significantly correlated with the sum of the amplitude spectrum and the activity fraction, and this statistic explained 10% of preference data variance (Pearson's $r = 0.317$; $p = 0.046$). While this relationship is modest, it shows that the same basic statistics shown to correlate with preference in Experiment 2 are correlated with human perceptual judgments over more than one image set, and even when the set contains portraits. Moreover, we did not examine the role of color, which would likely explain further portions of preference data variance, as noted above.

On the other hand, while monetary value is not predictive of preference, semantic information that would be difficult to model using basic statistics does play a major role in preference. We note, for example, that the most preferred painting showed a young girl cradling a small dog, while the least preferred painting depicted two politicians shaking hands in a rather stiff pose. Clearly, shared principles of aesthetic perception are subordinate to semantic considerations, at least at the extrema of the distribution of preference.

6. GENERAL DISCUSSION

The success of basic statistical measures relevant to vision coding as correlates of perceptual judgments makes them useful for approaching a more complete understanding of preference, and also for the design of image retrieval systems. Our basic statistical measures explained substantial portions of data variance in preference rankings for abstract works, and somewhat smaller portions for landscapes, while preference for portrait/still-life paintings showed insignificant correlations with basic statistics. These findings align with results in Experiment 1 suggesting that abstract and landscape works vary along fewer perceptual dimensions, such that similarity judgments are strongly aligned with preference for landscapes and abstract works. In addition, we showed that these same statistics are better predictors of preference compared to an emergent measure of value (selling price) for a separate set of images of mixed content.

In terms of perceptual processing, the basic statistical measures we employed can be seen as complex scene analogs of pixel-based similarity measures used in models of object recognition in the brain[21]: for example, squared pixel intensity difference is a metric typically applied to simple 2D and 3D object stimuli, and one shown to have high correlations with computed distances between voxel responses in imaging studies of brain regions involved in object recognition.[37] It follows that responses in areas responsible for the determination of scene perception could vary as a function of their location in the similarity space defined by basic image statistics.

Models of preference for art images should be viewed as a special case in the realm of models of preference (see e.g., the work of Tversky[38]). But since art is designed for human viewing and may thus occupy a subset of possible stimuli that are well matched to shared biases in preference, this work could have wider applicability. However, content and context remain important stumbling blocks. Our results in Experiment 1 showing that similarity ratings do not fully align with preference ratings—and our observation in Experiment 3 that image content plays an important role especially at the extrema of preference—demonstrate that low-level image statistics can only take us so far. But while portrait paintings in our experiments showed no significant correlation with image statistics, successful models of preference for faces already exist: as noted above, average faces are strongly preferred. This fact could be used to improve models of preference in portraits. But context (e.g., the instructions for how to evaluate preference or similarity), which was not manipulated here, requires testing if we are to develop general models.

Note that we did not include color in these experiments. Given the wealth of findings showing that color statistics can aid in image search[39,40] and that art can be classified into perceptual categories using computational models trained on color statistics,[32] we expect that the explanatory power of our collection of image statistics would significantly improve with the addition of such color metrics. Regularities in Fourier phase statistics may also help in this regard. Basic statistical properties of images could form the basis of Bayesian models of search, along with other measures. For example, one could consider a vector containing basic statistics, as well as content tags, and also a work's location in a simple perceptual space (consisting of similarity and perhaps other dimensions), as prior information when users search for images by preference. Ensembles of statistics could be weighted according to their empirical ability to predict preference across large, standardized databases. Viewers could initially rate a diagnostic set of images, and a preference vector could be updated as more images are rated.

More work will also be needed to investigate individual differences. Nevertheless, individual preference for content in other media has shown itself to be quite amenable to probabilistic modeling based on accumulated linkage data and on manually-applied tags, as shown in algorithms such as PageRank (Google), Cinematch (Netflix), and Music Genome (Pandora). The key for such algorithms is to establish common metrics upon which to model individual choice. While these systems do not generally employ statistical features of the information itself in their models, we suggest that basic statistics of images would be useful complements to tag-based strategies for search within art databases. We conclude that hybrid probabilistic models of this sort stand a good chance of providing non-expert users with accurate suggestions for images they will prefer. Moreover, the fact that basic statistics relevant to vision coding succeed to the degree they do suggests that they may be related to underlying psychological processes that determine preference.

Acknowledgements: This work was supported by a National Science Foundation Small Grant for Exploratory Research (DMS-0746667)

REFERENCES

[1] Graham, D. J. and Field, D. J., "Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities," Spatial Vision 21, 149-164 (2007).
[2] Graham, D. J. and Field, D. J., "Variations in intensity statistics for representational and abstract art, and for art from the eastern and western hemispheres," Perception 37, 1341-1352 (2008).
[3] Graham, D. J. and Field, D. J., "Global nonlinear compression of natural luminances in painted art," Proc. SPIE 6810, 189-200 (2008).
[4] Redies, C., Hasenstein, J. and Denzler, J., "Fractal-like image statistics in visual art: similarity to natural scenes," Spatial Vision 21, 137-148 (2007).
[5] Redies, C., "A universal model of esthetic perception based on the sparse and efficient coding of natural stimuli," Spatial Vision 21, 97-117 (2007).
[6] Graham, D. J., Friedenberg, J. D., Rockmore D. N. and Field, D. J., "Mapping the similarity space of paintings: image statistics and visual perception," Visual Cognition, in press.
[7] Eysenck, H. J., "The empirical determination of an aesthetic formula," Psychological Review 48, 83-92 (1941).
[8] Eysenck, H. J., "An experimental study of aesthetic preference for polygonal figures," Journal of General Psychology 79, 3-17 (1968).
[9] Wohlwill, J. F., "Amount of stimulus exploration and preference as differential functions of stimulus complexity," Perception and Psychophysics 4, 307-312 (1968).
[10] Nicki R. M. and Moss, V., "Preference for non-representational art as a function of various measures of complexity," Canadian Journal of Psychology 29, 237-249 (1975).
[11] Neperud R. W. and Marschalek, D. G., "Informational and affect bases of aesthetic response," Leonardo 21, 305-312 (1988).
[12] Rigau, J., Feixas, M. and Sbert, M., "Informational aesthetics measures," IEEE Computer Graphics and Applications, 28, 24-34 (2008).
[13] Roberts, M. N., "Complexity and aesthetic preference for diverse visual stimuli," Doctoral dissertation, Universitat de les Illes Balears-TDR (2008).
[14] Locher, P. J., "An empirical investigation of the visual rightness theory of picture perception," Acta Psychologica 114, 147-164 (2003a).

[15] Locher, P. J., “Experimental techniques for investigating the contribution of pictorial balance to the creation and perception of visual displays,” Empirical Studies of the Arts 21, 127 – 135 (2003b).
[16] Locher, P. J., Stappers, P. J. and Overbeeke, K., “Spatial balance of color triads in the abstract art of Piet Mondrian,” Perception 34, 169-189 (2005).
[17] Palmer, S. E., Gardner, J. S. and Wickens, T. D., “Aesthetic issues in spatial composition: effects of position and direction on framing single objects,” Spatial Vision 21, 421-449 (2008).
[18] Penton-Voak, I. S., Perrett, D. I. and Peirce, J. W., “Computer graphic studies of the role of facial similarity in judgements of attractiveness,” In: [Validation in Psychology: Research Perspectives], Eds. H. Ellis and N. Macrae, Transaction Publishers, New Brunswick, NJ, 133-149 (2001).
[19] Alexander, B. and Marks, L. E., “Aesthetic preference and resemblance of viewer's personality to paintings,” Bulletin of the Psychonomic Society 21, 384-386 (1983).
[20] D. E. Berlyne, “Similarity and preference judgments of Indian and Canadian subjects exposed to Western paintings. International Journal of Psychology 1, 43-55 (1976).
[21] Cutzu, F., and Edelman, S., “Representation of object similarity in human vision: psychophysics and a computational model,” Vision Research 38, 2227-2257 (1998).
[22] Rao, A. R. and Lohse, G. L., “Towards a texture naming system: identifying relevant dimensions in texture,” IEEE Visualization 1993, 220-227 (1993).
[23] Edelman, S., “Representation is representation of similarities,” Behavioral and Brain Sciences 21, 449-498 (1998).
[24] Russell, R. and Sinha, P., “Perceptually-based comparison of image similarity metrics,” MIT AI Memo 2001-014 (2001).
[25] Redies, C., Hänisch, J., Blickhan, M. and Denzler, J., “Artists portray faces with the Fourier statistics of complex natural scenes,” Netw.: Comput. Neural Syst. 18, 235-248 (2007).
[26] Graham, D. J., Friedenberg, J. D. and Rockmore, D. N., “Efficient visual system processing of spatial and luminance statistics in representational and non-representational art,” Proc. SPIE 7240, 1N1-1N12 (2009).
[27] Torralba. A. and Oliva, A., “Statistics of natural image categories,” Network: Comput. Neural Syst. 14, 391-412 (2003).
[28] Motoyoshi, I., Nishidam, S., Sharan, L. and Adelson, E. H., “Image statistics and the perception of surface qualities,” Nature 447, 206-209 (2007).
[29] Hyvarinen, A., Hurri, J. and Hoyer, P. O., [Natural Image Statistics], Springer-Verlag, New York (2009).
[30] Graham, D. J., and Field, D. J., “Efficient coding of natural images,” In: [Encyclopedia of Neuroscience] L. R. Squire (Ed.), Academic Press, Oxford (2008).
[31] Rolls, E. T. and Tovee, M. J., “Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex,” J. Neurophysiol. 73, 713-726 (1995).
[32] Wallraven, C., Fleming, R., Cunningham, D. W., Rigau, J., Feixas, M., and Sbert, M., “Categorizing art: comparing humans and computers,” Computers and Graphics, 33, 484-495 (2009).
[33] Pinto, P. D., Linhares, J. M. M., Carvalhal, J. A., and Nascimento, S. M. C., “Psychophysical estimation of the best illumination for appreciation of Renaissance paintings,” Visual Neuroscience 23, 669-674 (2006).
[34] Simoncelli, E. P. and Olshausen, B. A., “Natural image statistics and neural representation,” Ann. Rev. Neurosci. 24, 1193-1216 (2001).
[35] Graham, D. J., “The relationship between efficient coding of natural scenes in the human visual system and statistical regularities in art,” Doctoral dissertation, Cornell University (2008).
[36] Cutting, J. E. “Gustave Caillebotte, French Impressionism and mere exposure,” Psychonomic Bulletin and Review 10, 319-343 (2003).
[37] Op de Beeck, H. P., Torfs, K. and Wagemans, J., “Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway,” J. Neurosci 28, 10111-10123 (2008).
[38] Tversky, A. and Shafir, E., [Preference, Belief, and Similarity], MIT Press, Cambridge (2003).
[39] Swain, M. J. and Ballard, D. H., “Color indexing,” Intl. J. Computer Vision 7, 11–32 (1991).
[40] Neumann, D. and Gegenfurtner, K. R., “Image retrieval and perceptual similarity,” ACM Trans. Appl. Perception 3, 31-47 (2006).