

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Aesthetics of spatial composition: Facing, position, and context, and the theory of representational fit

### Permalink

<https://escholarship.org/uc/item/8wf7d7r9>

### Author

Gardner, Jonathan Sammartino

### Publication Date

2011

Peer reviewed|Thesis/dissertation

Aesthetics of spatial composition:  
Facing, position, and context, and the  
theory of representational fit

By  
Jonathan Sammartino Gardner

A dissertation submitted in partial satisfaction of the  
requirements for the degree of

Doctor of Philosophy in Psychology  
in the Graduate Division of the  
University of California, Berkeley

Committee in charge:  
Professor Stephen E. Palmer, Chair  
Professor John F. Kihlstrom  
Professor Gerald A. Mendelsohn  
Professor Brian A. Barsky

Spring 2011

## Copyright Page

Aesthetics of spatial composition: Facing, position, and context, and the theory of representational fit.

Copyright © 2011, All Rights Reserved  
Jonathan Sammartino Gardner

## Abstract

Aesthetics of spatial composition:

Facing, position, and context, and the theory of representational fit

by

Jonathan Sammartino Gardner

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Stephen E. Palmer, Chair

Artists who work in two-dimensional visual media regularly face the problem of how to compose their subjects in aesthetically pleasing ways within a surrounding rectangular frame. We performed psychophysical investigations of viewers' aesthetic preferences for the position and facing direction of single, directed objects (e.g., people, cars, animals, teapots, and flowers) within such rectangular frames and examined the role that meaning and context play in some of these preferences.

For the horizontal placement of objects, preferences were measured using two-alternative forced-choice preference judgments, the method of adjustment, free choice in taking photographs, and an analysis of stock photography. In front-facing conditions, preference was greatest for pictures whose subject was located at or near the center of the frame and decreased monotonically and symmetrically with distance from the center (the center bias). In the left- or right-facing conditions, there was an additional bias toward preferring objects that face into rather than out of the frame (the inward bias). Similar biases were evident using a method of adjustment, in which participants positioned objects along a horizontal axis, and in free choice photographs, in which participants were asked to take "the most aesthetically pleasing picture" they could of everyday objects. The results are discussed as affirming the power of the center and facing direction in the aesthetic biases viewers bring to their appreciation of framed works of visual art (e.g., Arnheim, 1988; Alexander, 2002).

Next, aesthetic preference for the vertical composition of single-object pictures was studied through a series of four two-alternative forced-choice experiments. The results reveal the influence of several factors, including spatial asymmetries in the affordances of the object and the typical position of the object relative to the observer. With asymmetric side views of objects, people generally prefer objects typically below the observer's viewpoint (e.g., a bowl and a swimming stingray) to be located below the center of the frame and objects located above the observer's viewpoint (e.g., a light fixture and a flying eagle) to be located above the center of the frame. In addition, people generally prefer symmetric views of those same objects from directly above or directly below to be closer to the center of the frame. We suggest that these results can be unified by the hypothesis that observers prefer the object's "affordance space" to be centered within the frame.

Finally, four experiments examined the role of contextual meaning (through titles) on preference for images with objects at different positions and perspectives. As predicted by a theoretical account in terms of “representational fit,” people prefer standard (default) compositions with a neutral title that merely describes the content of the picture, but prefer non-standard compositions when they “fit” a title that has corresponding spatial implications. Thus, there is greater aesthetic value in novel compositions that violate expectations if the outcome is meaningful and coherent. These results support representational fit as an aesthetic theory that unifies fluency accounts, in which the default context prevails (e.g., Reber, Schwarz, & Winkielman, 2004), with classic aesthetic accounts in terms of novelty and expectation violation, in which a nonstandard meaning is intended or inferred.

## Acknowledgements

- First and foremost, to Dave Carter (1938-2008), my mentor and friend whose undergraduate course at Georgetown University, “Psychology, Photography, and the Visual Arts” first introduced me to the fascinating overlap between psychology and the visual arts. He made me a better psychologist and a more successful photographer, and taught me how to see art and beauty in daily life. He made me realize that upon graduating, I didn’t have to choose between my two passions of psychology and photography -- I could do both. Without him I would not have pursued a PhD in psychology in the first place.
- To my advisor and committee chair, Steve Palmer, who taught me that it is possible to study aesthetics in the first place; who taught me how to think and conceptualize problems large and small in an empirically-minded way; and who provided thorough and meaningful feedback on literally countless drafts (I’ve counted 53 drafts so far, but I’m sure that I’m leaving some out) of my research -- much of which is presented here. I can proudly call myself an experimental psychologist, and I owe that status to him.
- To the other members of my committee: John Kihlstrom, Jerry Mendelsohn, and Brian Barsky, who have all provided useful feedback and support throughout this process, and who suggested countless books, articles, and lines of research to include that make my work here much more readable, interesting, and worthwhile than it would have otherwise been.
- To Bruce Mangan, my friend and colleague, who has taught me much about philosophy, aesthetics, and consciousness, and who functionally served as a fifth committee member for my dissertation. Our many conversations and his insightful comments -- especially on the definition of aesthetics and the history of aesthetics in philosophy and psychology -- have been invaluable.
- To Tom Wickens, without whose patience, understanding, and knowledge, several of the complicated statistical analyses presented here would have remained incomprehensible to me.
- To the research assistants and lab managers who have helped me with scheduling participants, data collection, and data analysis, and without whom much of this work would have taken me much longer to complete: Gary Hackett, Hye-Lim Jeon, Daisy Liu, Christie Nothelfer, Rosa Poggesi, Michael Vattuone, Frances You, and Emily Zuckerman.
- To teachers and friends who have encouraged me along the way: Roberto Bocci, Victor Boft, Calvin Custen, Tate Guelzow, Fred Loya, Karen Schloss, Barry Shapiro, and Bijan Yashar.
- To Karen Schloss, whose company, comments, and commiseration over the course of my time in grad school was -- and continues to be -- indispensable.
- To Jessica Thierman, whose love and support can help me get through anything.
- Last, but certainly not least, to my family, who have supported me through everything, including my 22+ years of education: Mom, Joseph, Pa, and Grandma.

Some of my research has been supported by National Science Foundation Grant No. 0745820 and a Google Gift to my advisor, Steve Palmer.

# Introduction

## Aesthetics

Met with the term “aesthetics,” many people are liable at first to conjure up an image of a person standing in front of a painting at an art gallery, struck with awe and wonder at the colors, composition, and expressive content before them. However, we can have aesthetic experiences in response to a much broader variety of stimulation: a natural scene we see during our day (e.g., the view from a mountain peak after a long hike, or the view of a tranquil lake from the kitchen window), an object we see in the world (e.g., an unusually well-designed car driving on the freeway), or even an attractive person that we see walking down the street.<sup>1</sup> It seems reasonable, in fact, to say that we have aesthetic experiences of varying degrees on a daily basis, even if we don’t go anywhere near a museum or even view an image of any sort.

For the purposes of the research I present here, I will define aesthetics as the study of principles relating to the sense of beauty and “rightness” – beauty in the pleasure that we feel in response to images, objects, people, and events; and rightness in the pleasure that is often felt in terms such as “it fits,” or “it’s right” or “it works” very well in its context. These pleasures are two aspects of aesthetic experience which – especially in extreme cases of being moved to tears of awe or joy – is often ineffable and difficult to describe thoroughly. However, my research will focus on images and will be limited to the low, but positive, end of the scale of aesthetic experiences that take the form of judgments of liking.

Beauty is often referred to as the antecedent of aesthetic experiences. But there is a difference between the “beauty” that encompasses aesthetic experiences, and the “beauty” that we colloquially talk about in everyday language. Harold Osborne (1970) points out that there are really two meanings of beauty. One is the soft, simple, pleasing notion of beauty that became dominant in Renaissance and post-Renaissance Western culture. The other is the notion of beauty-with-a-capital-B that is shared with many other cultures and times: e.g., Ancient Greece, China, and others. This second concept is not always even translated as “beauty” in English for the very reason of avoiding confusion. While Osborne does prefer the translation of *to kalon* -- the original word found in Plato’s writings -- as “beauty,” he points out that when this term is used to describe an object, “we mean that [the object] is well made for a particular purpose but we do not necessarily imply that it is good to look at or that it has beauty of appearance...” (Osborne, 1970). Immanuel Kant, in his treatment of aesthetics in the *Critique of Judgment*, differentiates between “beautiful,” referring to the simple, pleasing quality, and “sublime,” referring to the more moving, powerful quality that can also include the functional component that Osborne mentions (Kant, 1790/1987). In *Meaning and the structure of consciousness*, Bruce Mangan lays out an argument for how the “rightness” of an aesthetic experience is not only in concert with the entire range of shallow to deep aesthetic experiences, but also can serve to unify our phenomenology of aesthetic experiences with notions of aesthetics in philosophy as well as psychological and neuropsychological data about the structure of consciousness.

Aesthetic experiences are complex psychological phenomena that bear importantly on

---

<sup>1</sup> These examples are all of the visual domain, however. To be thorough, it should be noted that aesthetic experiences can result in any sensory or non-sensory (e.g., cognitive) domain. These experiences also fit with the use of “rightness” as the content of an aesthetic experience. However, such non-visual experiences are beyond the scope of this dissertation and will not be discussed in any detail.

studies of consciousness (see Mangan, 1991), emotion (see Winkielman, et al., 2002), cognition, and perceptual processing (see Reber, et al., 1998, 2004; Winkielman, et al., 2006). Dutton (2009) argues that people across many, if not all, cultures and backgrounds have aesthetic experiences, even if they are not in response to the same kinds of objects or events. Insofar as psychologists are interested in universal and culturally specific characteristics and determinants of behavior, they ought to be interested in aesthetic experience and all it involves.

## Images

Our visual world teems with images. Television shows us thirty of them each second, movies, twenty-four; newspapers and magazines pair almost every article with at least one photograph or graphic, and the internet, which is working its way more and more into the center of our lives, is filled with image content, often surrounded by even more images, including those in the form of advertising banners. This is to say nothing of the billboards, bus ads, posters, and other images that we view on a regular basis. While it would be impossible to give a specific number, it is easy to say that we see thousands of images each day in one form or another.

Some of these images are pleasing to us and engender in us a feeling of rightness and beauty; others we could take or leave after we have gleaned from them the relevant information; and, still others we find downright unpleasant. These reactions can be brief or lasting, but regardless of their duration, they are unified by placement on the scale of aesthetic responses.

Underscoring the importance of images in today's modern world is the fact that it is increasingly easy to create images, particularly photographs. Whereas only ten years ago, good digital cameras were prohibitively expensive, dedicated image-creating devices such as point-and-shoot or single-lens-reflex digital cameras have become commonplace. Moreover, many other devices with disparate primary purposes, such as laptops, MP3 players, and cell phones, today come equipped with built-in cameras. It is easy, with this plethora of devices, for almost anyone to take hundreds of photographs in a very short time and amass thousands of them in just a couple of years or even less.

## My Goal

Because of both the pervasiveness of images and people's aesthetic responses to them, the time seems ripe for a principled study of images and the factors present in them that elicit aesthetic responses. There are many factors that might be relevant to various subsets of images, such as some explicit style or some particular content, but one important factor that is relevant for virtually all two-dimensional images is spatial composition. All images, by definition, have their content -- whatever it may be, abstract or representational -- arranged in a specific way within the frame, and this is what I mean by spatial composition.

Over the course of this dissertation, numerous studies of spatial composition will be presented with an eye towards complementing prior research on aesthetics as well as demonstrating the power of psychophysical methods for studying aesthetic preference. The results will be discussed in terms of their immediate and general ramifications.

First, a brief discussion of aesthetics in philosophy will be presented. It is followed by a more thorough overview of aesthetics in the field of psychology. The principles of balance and fluency will be discussed, and theoretical and empirical work relevant to these topics will be reviewed. Next, I will set forth the empirical rationale for my research, followed by numerous



experiments, broken down into three sets. The first five concern preference for the composition of objects in the horizontal dimension. The next four experiments concern analogous preferences in the vertical dimension. The final four experiments demonstrate that preference for compositions -- and aesthetic experiences in general -- can be dramatically affected by different meaningful context.

# Chapter I

## Relevant Work in Philosophical Aesthetics

It would do a disservice to the rich history of the study of aesthetics, which began as a branch of philosophy, to fail to mention some of the work on philosophical aesthetics that is especially germane to my research. There is no good reason to reinvent this wheel, even if it was not empirically tested by philosophers. It is well beyond the scope of this dissertation, however, to provide a thorough analysis of even the primary contributions to the field of philosophical aesthetics. The notions of wholeness and coherence, however, and the arguments that Plato and his successors have made about them are highly relevant to this dissertation.

### Plato

The earliest explicit theory of aesthetics is found in the writings of Plato. Although there are bits and pieces of Plato's theory of art, aesthetics, and beauty spread across his *Symposium*, *The Republic*, *Hippias Major*, and other works, his most thorough account of aesthetics, at least for present purposes, is in *Phaedrus*.

### *Phaedrus*: Plato's Theory of Aesthetics and Beauty

Rhetoric was the most highly regarded art form in Ancient Greece (Osborne, 1970). Though its origins were political, it became common for people to demonstrate their talent for rhetoric by giving speeches on topics with which they had little or no experience. Artful rhetoric demonstrated the power of the speaker not only to think logically and carefully, but to present an aesthetically crafted argument, political or otherwise, in the form of a compelling speech. In *Phaedrus* Plato uses Socrates's critique of a speech by Lysias to provide an assessment of proper aesthetic qualities:

Socrates: Then again, don't the various parts of his [Lysias'] speech give the impression of being thrown together at random? Do you see any intrinsic reason why the second topic, rather than any of the others, should be placed second? I am an ignoramus, of course, but it seemed to me that the writer showed a fine carelessness by saying whatever occurred to him. Can you point out any compelling rhetorical reason why he should have put his argument together in the order he has?

Phaedrus: You do me too much honour if you suppose that I am capable of divining his motives so exactly.

Socrates: But I think you would agree that any speech ought to have its own organic shape, like a living being; it must not be without either head or feet; it must have a middle and extremities so composed as to fit one another and the work as a whole.

Phaedrus: Of course.

Socrates: Well, now look at your friend's speech and see whether it conforms to this criticism. You will find that it is no better than the epitaph said to have been inscribed on the tomb of Midas the Phrygian.

Phaedrus: What epitaph is that and what is the matter with it?

Socrates: It goes like this:

'A girl of bronze on Midas' tomb I stand,  
As long as water flows and trees grow tall,  
Remaining here on this lamented tomb,

I'll tell to all who pass, "Here Midas lies"  
You notice, I am sure, that it is of no consequence what order these lines are  
spoken in.  
Phaedrus: You are making fun of our speech, Socrates. (Plato, *Phaedrus*, 264)

The way in which the parts fit together -- their order and structure, which is, in effect, the speech's composition -- is of paramount importance to Plato. Many important characteristics emerge from this principle. First and foremost, small alterations to an entity -- be it a speech, a painting, or some other object that is viewed in aesthetic terms -- can make very large differences to the aesthetic experience, to the beauty of the entity.

The most famous example of this principle comes from William Strunk and E.B. White's analysis of four variations on a statement by Thomas Paine:

Variant 1: Times like these try men's souls.  
Variant 2: How trying it is to live in these times!  
Variant 3: These are trying times for men's souls.  
Variant 4: Soulwise, these are trying times. (Strunk & White, 2000)

The original sentence, of course, was: "These are the times that try men's souls." As Strunk & White point out, this is a clear demonstration that the four variants lack the same "organic shape" or beauty of the original -- a perfect validation of Plato's argument. The same content and nearly identical words do not yield sentences that are equally pleasing, beautiful, organic, or whole. This meshes well with the notion of rightness that I mentioned in the introduction. To use a word and metaphor that will come up later in this dissertation: there is a certain *fit* between the words of the original sentence that holds them all together, and it is this fit that gives the sentence its aesthetic beauty.

## Variations on a Theme: Kant and Mangan

Many other philosophers, using many different terms -- wholeness, coherence, synthesis, gestalt, significant form, and others -- have made similar points about rightness being a key ingredient for beauty or, more generally, a pleasing aesthetic experience (Mangan, 1991; Osborne, 1970). This is quite possibly the oldest and most universal criterion for aesthetic experience.

In *Meaning and the structure of consciousness*, Mangan provides an insightful analysis of aesthetics in philosophy and its relation to the structure of consciousness, most notably in the work of Immanuel Kant. Giving a comprehensive summary of Mangan's analysis is beyond the scope of the present discussion. I will therefore merely distill the most useful and relevant points for my present argument.

Over the course of his analysis, Mangan proposes a view of aesthetic experience that is cognitively oriented, and is centered around the notion of "rightness" or "meaningfulness" as opposed to a simple notion of beauty, as was discussed in the introduction. He points to Kant's discussion of the human drive for order, which consists first and foremost of "Zweckmassigkeit [Kant's term], or purposiveness [its translation]" (Mangan, 1991). Mangan, reinforcing Osborne (1970), notes that purposiveness could also be translated as "'appropriateness' or 'suitability,'" again resonating with the notion of rightness or meaningfulness. He states:

We are predisposed to find as much order as possible in the world, and Kant thinks that a given set of elements will seem more ordered if they seem to cohere into a unified whole, and more ordered still if the unity appears to satisfy an aim or purpose (rather than the same structure resulting from a chance occurrence)... *Zweckmassigkeit* emphasizes Kant's attempt to deal with the cognitive principle which is always trying to maximize order in every possible way, and predisposes (but does not compel) our cognitive apparatus to interpret all order-laden situations as having the richest possible order. (Mangan, 1991)

Mangan paraphrases Kant's definition of beauty as "the feeling of purposiveness in the absence of any concepts adequate to the experience." Our aesthetic feeling is thus tied to our much more inclusive and -- from a scientific standpoint, biological -- drive for order and rightness. Kant's purposiveness is "the conscious experience signaling that order has been discovered by nonconscious processes."<sup>2</sup> This categorization efficiently unifies Kant's theory of human cognition with aspects of aesthetic experience that we cannot describe with precise concepts and which, to some degree, always elude precise sensory specification. (This ineffability is rooted in the same phenomenology as the *je ne se quois*, a phrase which, Mangan points out, was popularized, but not coined, by Leibniz). This argument supports the notion of "rightness" that I mentioned in the introduction. It is also supported by relevant work in psychology, which is the topic of the next chapter.

---

<sup>2</sup> Kant uses the term "transcendental" -- not meant to be religious in any way, despite its current colloquial usage -- to refer to any processes that take place outside of consciousness. This is a direct precursor to Hermann von Helmholtz' notion of unconscious inference, though whether Helmholtz was explicitly influenced by Kant is a topic of much debate (e.g., Riehl, 1876, Cassirer, 2005).

## Chapter 2

# A Brief History of Psychological Aesthetics

The study of psychology had its beginnings, in most universities around the world, in departments of philosophy. Over time, theorizing gave way to empirical study, and psychology as a discipline came to be more of a science and separated from philosophy. Nonetheless, its philosophical origins led it to import some of the questions that had previously been examined only by philosophers. Most importantly for present purposes, this included aesthetics.

### Fechner

Gustav Theodor Fechner, the scientist and mystic whose name is most commonly associated with the field of psychophysics, is singularly responsible for the importation of the field of aesthetics into the domain of psychology. Fechner performed hundreds of studies of aesthetics and published two books on his research -- *Zür Experimentalen Aesthetik* (Towards Experimental Aesthetics) in 1871 and the more expansive *Vorschule der Aesthetik* (School of Aesthetics) in 1876. Unfortunately, although his *Elements of Psychophysics* has been available in English translation for over 50 years, both *Zür Experimentalen Aesthetik* and *Vorschule der Aesthetik* remain available only in German, a sad reminder of the relative decline of attention given to aesthetics, especially in modern times.<sup>3</sup> The present analysis of his principles and experiments will therefore be drawn from secondary sources.

*Zür Experimentalen Aesthetik* was a short volume that laid forth Fechner's three major methodologies for his studies of aesthetics (Green, 1995). The first method -- the method of choice -- consisted of providing a participant with two or more choices of items and having them pick the one that they found most aesthetically pleasing. The second -- the method of production -- required participants to create an aesthetic object that had characteristics pleasing to them. The third -- the method of use -- was theory-based and consisted of analyzing existing works of art with regard to one or more dimensions of relevance described by an aesthetic theory to test the validity of the theory. Fechner used all three of these methods to study the aesthetic power of the Golden Ratio, an important piece of his contribution that will be addressed in the section on balance in Chapter 3.

*Vorschule der Aesthetik* was a much longer and grander treatment of aesthetics, consisting of two volumes totaling almost 600 pages. He began by eschewing any attempt to tackle the concept of beauty in and of itself, instead using it:

...as an auxiliary term to find a brief designation, in the sense of linguistic usage, of things that unite in themselves conditions that lead to general liking... (Fechner, 1876, as cited in Allesch, 2001).

Christian G. Allesch points out that to the “philosophical scientists” of the day (of which psychologists and people who studied aesthetics made up two sub-groups), the framework that Fechner put forth in *Zür Experimentalen Aesthetik* and *Vorschule der Aesthetik* was perceived as a provocation of sorts (Allesch, 2001). Like most of Fechner's research on the various fields he pioneered and studied, it was radically insightful for its time.

---

<sup>3</sup> Until the early/middle part of the 20<sup>th</sup> Century, the majority of PhD candidates in the field of psychology were required to learn to speak a foreign language. This language was often German, since so many of the foundational psychologists wrote in this language. Thus, until this point in time, translation of Fechner's work into English was largely unnecessary.

Fechner foresaw an entire science of aesthetics that would eventually include some objective definition of beauty. However, he also firmly believed that “higher-level philosophical aesthetics” needed to be grounded in a thorough understanding of what was “likely to trigger a higher pleasure than only sensual lust directly from sensorial impressions and under which circumstances this was likely to occur with the probability of an intrinsic psychological law” (Fechner, as cited in Allesch, 2001). In the absence of such understanding and preliminary experimentation, Fechner believed that “all our systems of philosophical aesthetics” would amount to nothing more than “giants with feet of clay.” After this bold yet empirically-minded introduction in *Vorschule der Aesthetik*, Fechner proceeded to build the foundation of aesthetic science from the ground up, to use his own metaphor..

Although Fechner went on to explicate a theory that included twelve specific aesthetic principles,<sup>4</sup> I will argue that Fechner’s most important contribution to the empirical study of aesthetics came from his methods, rather than from the particular conclusions he derived from their use. His three methods (see above) are nothing less than a paradigm shift from the ways that philosophy and the history of aesthetics had proceeded until that time. Fechner’s prescient claim in 1876 was simply that the nature of aesthetic experience and the study of aesthetic preference can be subjected to the same rigorously empirical methods as sensory perception, behavior, and all of the other aspects of experience that psychology deems fit for experimental study. This dissertation is, in part, an attempt to show that he was correct in this assertion.

### Post-Fechner, Pre-Freud: The First Dry Period

After Fechner introduced the study of aesthetics to psychology, many researchers took it upon themselves to perform at least a study or two in the field. As I will discuss in Chapter 3, many of these psychologists chose to follow his lead by pursuing the Golden Ratio. However, their studies were very limited in both scope and impact; they rarely built upon each other; and no one attempted to put forth a coherent theory of aesthetics based on their results.

### Sigmund Freud and Psychoanalytic Aesthetics

Prolific as he was, Sigmund Freud wrote very little about aesthetics per se. He did produce two essays in the vein of art criticism -- *The Moses of Michaelangelo* (Freud, 1914/1989) and *Leonardo da Vinci and a Memory of his Childhood* (Freud, 1910/1989) -- though neither paper laid forth Freud’s thoughts on what made the specific artworks he discussed aesthetically pleasing or displeasing. Instead, Freud’s theoretical stance on the psychological function of art is woven into other writings on jokes, dreams, fantasies, and the like.

Despite his questionable status in cognitive and clinical psychology today, Freud’s theory of aesthetics has flourished, perhaps better than any other aspect of his work. His psychoanalysis has “spread to every nook and cranny of the culture, from Salinger to ‘South Park,’ from Fellini to foreign policy...” (Cohen, 2007). Even seventy years after Freud’s last original publication, one need not look hard to find that journal articles on psychoanalytic aesthetics and Freudian

---

<sup>4</sup> These principles were derived from his thoughts on sensory inputs and the workings of the mind. The first five were primary principles: “of an aesthetic threshold; of aesthetic help or increase; of the unified combining of the manifold; of noncontradictability, of agreement or truth; and of association” (Fechner, 1876, as translated by Martin, 1906). Seven “subordinate” principles were also put forth: “...that of contrast, sequence and reconciliation; of summation, practice, blunting, habit or custom and of satiety; of persistence, of change, and of the amount of occupation; of the expression of pleasure and displeasure; of secondary pleasurable and displeasurable ideas ; of the aesthetic mean; and of the economical application of the means.” The validity of these principles is fairly questionable; indeed, there has been only one attempt to critically examine them, and it concluded that the principles thoroughly failed to be supported by the data collected (Martin, 1906).

interpretations of works of art are still published regularly.<sup>5</sup>

Freud's theory is rooted in his conception of human desire and his three-tiered structure of the mind. The lowest tier, the "id," houses our most basic, biological desires; the highest tier, the "superego," contains the culturally- and societally-based rules that limit our behavior and compel us not to act on the desires of the id; finally, the "ego" mediates between the other two, arriving at a compromise that, if effected properly, results in mental health as well as societal acceptance (Freud, 1933/1989; Westen, 1998). Art appreciation and creation then fit into this framework because the desires of the id that are denied by the ego and superego can bubble up in works of art -- though symbolically, so as to prevent the "latent" (i.e., repressed) content of the desires from manifesting directly and causing embarrassment, anxiety, or other social repercussions (Freud, 1911/1989). This leaves unaddressed, however, the large numbers of artworks that depict situations that are not likely desirable on any level. In his *Psychoanalytic Explorations of Art*, Ernst Kris draws on Freud's writing of "mastery" of anxiety to justify exactly these works, and their place in a psychoanalytic theory of art. Specifically, Kris postulates that by depicting unpleasant scenes, the artist -- and by proxy, the viewers -- are able to process and alleviate their anxiety surrounding these scenes (Kris, 1952).

One relevant distinction between the different psychoanalytic theories of art is that of form versus content. From Freud's own writings, one of the most important aspects of one's reaction to a scene (be it in a dream or in a work of art) is that of the associations that it brings up for the viewer. For example, if a person has a dream in which a pig is a central element, it is not the actual physical form of a pig but rather the symbolic significance of the pig that dictates the appropriate interpretation. Put another way, our reaction to an abstract work of art is based around the associations we have with the colors and forms, not something intrinsic about the colors or forms themselves.<sup>6</sup>

## Gestalt Psychology

Though the Gestalt school of psychology is primarily associated with matters of perceptual organization, their notions of grouping, perceptual coherence, and "gestalt" translate very naturally into the arena of aesthetics, especially in the visual domain.

Wertheimer's classic work on form and visual perception, "Laws of Organization in Perceptual Forms," illustrates principles of vision that could apply equally well to art, retinal images, or other forms of visual sensation. Beginning with a now-classic observation, he says:

I stand at the window and see a house, trees, sky. Theoretically I might say there were 327 brightnesses and nuances of color. Do I have "327"? No. I have sky, house, and trees. It is impossible to achieve "327" as such. And yet even though such droll calculation were possible and implied, say, for the house 120, the trees 90, the sky 117 -- I should at least have *this* arrangement and division of the total, and not, say, 127 and 100 and 100; or 150 and 177. The concrete division which I see is not determined by some arbitrary mode of organization lying solely within my own

---

<sup>5</sup> According to the ISI Web of Knowledge, there have been 97 articles published since 2000 that included "Freud," and either "aesthetics" or "art" as their topic.

<sup>6</sup> Several authors have pointed out that in this way, psychoanalytic theories of art seem incompatible with formalist accounts of aesthetics (e.g., Marshall Bush, 1967), but the details of this debate extends beyond the relevance of Freud's theory to my survey of the history of aesthetics in psychology.

pleasure; instead I see the arrangement and division which is given there before me. (Wertheimer, 1923/1938; emphasis original)

Wertheimer then puts forth the original set of what have come to be known as Gestalt grouping principles. It could be said that Wertheimer did not do any experiments on his grouping principles, but it would be more accurate to say that he did not conduct any *formal* experiments: everyone who reads his papers and looks at his figures becomes a participant, and their perceptions become (virtual) data that prove the validity of these principles.

Entire dissertations have been written about these grouping principles: their nature, their validation, and their ramifications for human perception. However, for present purposes, the crucial point is not the specific laws, but rather what they demonstrate about the inherent propensity of human beings to group parts into wholes. This propensity, taken to its logical conclusion in the following paragraph, can give us some insight into aesthetic experience that will later be relevant to the present analysis of aesthetic response.

According to Wertheimer, the units of visual sensation “usually combine in some ‘spontaneous,’ ‘natural’ articulation -- and any other arrangement, even if it can be achieved, is artificial and difficult to maintain.” Paraphrased within a Platonic framework, one might say that the organic whole is more natural for perception than the fragmented pieces. If it is more natural, it is reasonable to say that it would feel more “right.” On the other hand, if we are presented with pieces that do not make up a coherent whole, we would be unable -- or at least less able -- to combine them in this way, and our perception would feel less natural or right. So to the extent that a positive aesthetic experience is based on our feeling of rightness, as claimed by authors from Plato to Mangan, a configuration or ‘whole’ that is coherent should be preferred to one that is incoherent or less coherent.

## Rudolf Arnheim

Rudolf Arnheim was very much a part and product of the Gestalt school of psychology, having been trained by Max Wertheimer, Wolfgang Köhler, and Kurt Levin at the University of Berlin. The fact that he warrants his own separate section here may therefore seem puzzling. However, his name has come to be identified above all others with the psychology of art. Indeed, his teaching career culminated with a position at Harvard as “Professor of the Psychology of Art.” One could reasonably argue that Arnheim was the single most influential psychologist of art and aesthetics in the 20<sup>th</sup> Century.

Much like the Gestalt psychologists before him, Arnheim analyzed art in terms of the principles that underlie perception. In *Art and Visual Perception*, he describes his goal as “making visual categories explicit, ... extracting underlying principles, and... showing structural relations at work.” He states that “[T]his book deals with what can be seen by everybody... [a] survey of formal mechanisms,” (Arnheim, 1974). These mechanisms include (but are not limited to): balance, shape, form, space, light, and color. Each of these factors gets an entire chapter of explanation and examples, all informed by Gestalt psychology. The line-and-dot images drawn by Arnheim to illustrate so many of the principles he unpacks would make a fitting sequel to Wertheimer’s “Laws of Organization.”

Indeed, the influence of Gestalt psychology on Arnheim is visible in almost every detail of his work. It is perhaps best exemplified both thematically and stylistically in Arnheim’s explanation of “perceptual dynamics” in *Art and Visual Perception*, which begins with the basic Gestalt principle of simplicity and leads him to “A Diagram of Forces” (a subheading in his chapter on



dynamics) which in turn harks back to the Gestaltist notion of the perceptual “field” -- analogous to electromagnetic force-fields -- laid out by Wertheimer (1923) and elaborated by Köhler (1947):

In trying to discover what makes a visual object or event look the way it looks, we have been so far safely directed by what I have called the principle of simplicity. This principle, a basic guideline of gestalt psychology, holds that any visual pattern will tend toward the simplest configuration available to the sense of sight under the given circumstances. It has explained to us why certain shapes or colors fuse into units or come apart, why some things look flat while others have volume and depth; it has enabled us to understand the rationale of completeness and incompleteness, whole and part, solidity and transparency, motion and standstill. If one basic principle elucidates so many different phenomena, we owe it gratitude. However, at this point it is necessary to acknowledge that the tendency toward simplicity alone cannot do justice to what we see...

... dynamic properties, inherent in everything our eyes perceive, are so fundamental that we can say: *Visual perception consists in the experiencing of visual forces...* Directed tension, then, is what we are talking about when we discuss visual dynamics. It is a property inherent in shapes, colors, and locomotion, not something added to the percept by the imagination of an observer... (Arnheim, 1974)

Arnheim went on to expand this notion of dynamics in his work on visual “centers,” which he described to some degree in almost all of his works, culminating in *The Power of the Center*, his twelfth book. This work on centers and balance will be discussed in Chapter 3.

Arnheim almost never conducted proper behavioral experiments. His theory is built on his analysis of case studies of individual works of art. This limits the impact of his work insofar as he does not provide proof that a principle that works with image A also works with image B; he leaves it up to his reader to make that connection, if it is warranted.

Nonetheless, Arnheim’s impact on the psychology of aesthetics and the subfield of psychology of art is immeasurable. He made this area of study appealing and accessible to a very broad audience. Though his work lacks empirical support, the principles described in his vivid prose are still viable and testable, providing an important foundation for the field to this day.

## Daniel Berlyne

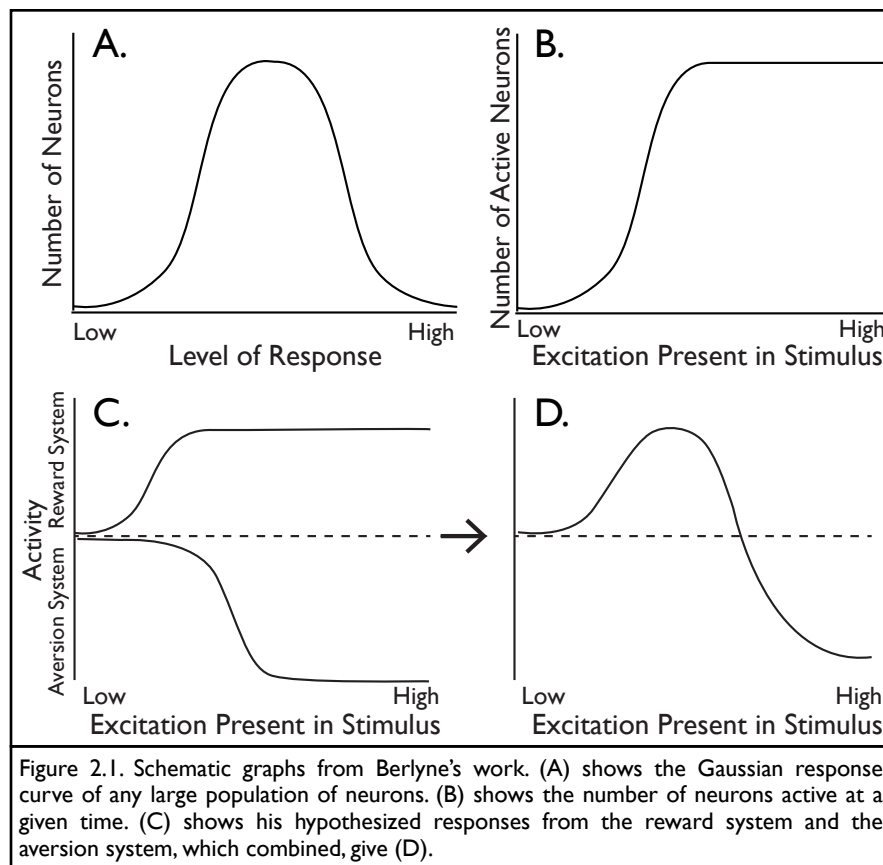
Daniel Berlyne came from a very different tradition within Psychology, being a product of the behaviorist and neo-behaviorist schools against which the Gestaltists railed. He was familiar with the work of Clark Hull and was “above all, a motivation theorist” (Konecni, 1978). His interest in motivation eventually brought him to the field of aesthetics, on which he wrote his justly famous 1971 book, *Aesthetics and Psychobiology*. Indeed, as a result of his academic history and context, his theory of aesthetics contained some remnants of behaviorism: he viewed aesthetics in terms of the physiological systems in our bodies that respond to visual and auditory stimulation.

### Pleasure and Aversion System

To Berlyne, the aesthetic response was the result of a combination of arousal and activity in

our pleasure and aversion system in the brain. With regard to the underlying neuronal responses, Berlyne made three assumptions on which he built many predictions about aesthetic experiences. Taking for granted that the pleasure system in the brain is made up of neurons, he postulated that over a large population of neurons, the average response of the population would approximate a Gaussian curve (Figure 2.1A), such that a few neurons would have very low responses, a few would have very high responses, and the majority would fall somewhere in the middle. It follows from this that the number of neurons active at any given time would be approximated by a cumulative Gaussian (Figure 2.1B). From here, he made two more debatable assumptions: that the aversion system would have a higher threshold than the pleasure system, and that the aversion system would plateau at a higher level of activation than the pleasure system (see Figure 2.1C). The combination of all of these statements results in the so-called “Wundt curve” shown in Figure 2.1D, which originated with some of Wundt’s work on stimulus intensity and sensation. Berlyne stated that arousal above a certain level is inherently negative simply because it is overstimulating. Although this conclusion seems valid, it follows that the most arousing positive display (i.e., highest peak with lowest trough in Figure 2.1D) is less arousing overall than the most arousing negative display. This is a hard conclusion to justify either empirically or phenomenologically.

Berlyne proposed that the activation of this aesthetic system was modulated by both arousal-increasing devices and arousal-moderating devices which, when combined, would elicit a pleasurable or aversive response, depending on their relative levels of activation.



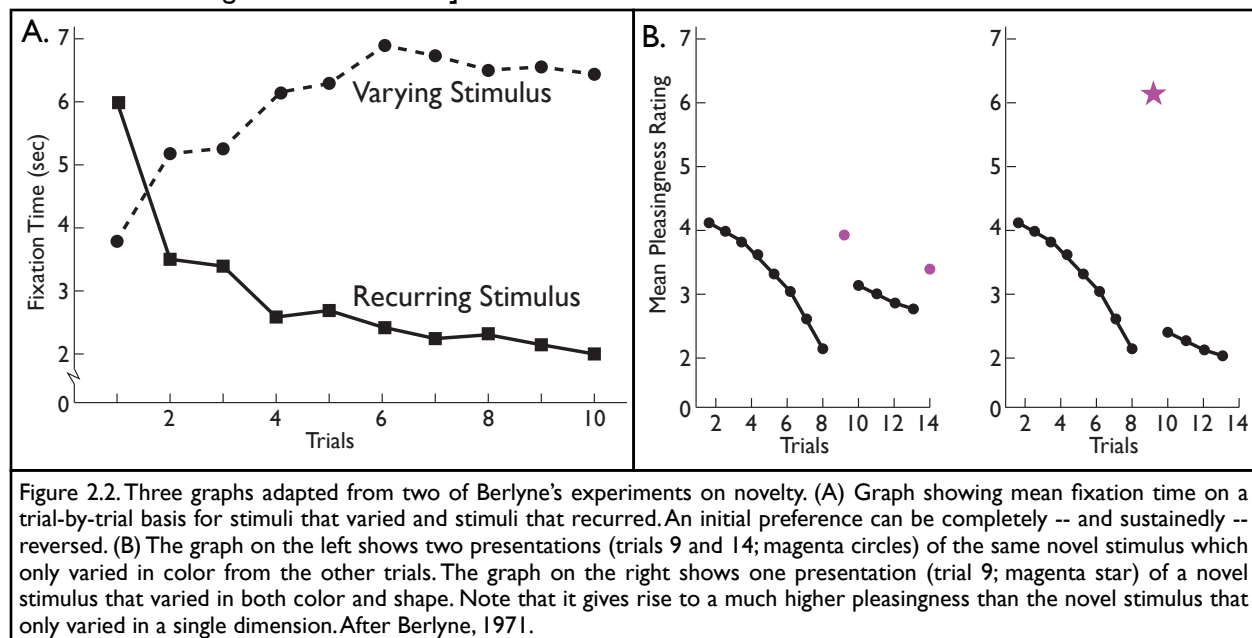
The arousal-increasing devices he proposes are: dishabituation, novelty, surprise and incongruity, uncertainty, absence of clear expectations, complexity, conflict, ambiguity and multiple meaning, and instability. The arousal-moderating devices he proposes are fewer: familiarity, exemption from inhibition and exertion, grouping and patterning, and dominance. Berlyne performed many behavioral tests exploring each of these devices. I will focus on his experiments on novelty and complexity, as they will be the most relevant for later consideration.

## Novelty

Novelty is a topic that has often been discussed in art theories.<sup>7</sup> Berlyne's treatment of novelty is arguably the most thorough, both theoretically and experimentally, of anyone who ever addressed the topic. He recognized that, rather than being a simple characteristic of anything new, there were several different kinds of novelty, each of which could have relatively different effects on people's aesthetic experiences. A stimulus that was different from everything that a person had ever seen before, fell into the category of "absolute novelty." He pointed out, however, that in the truest sense, this is "highly unlikely to be the case except when a newborn baby or somebody just cured of blindness first receives visual stimulation." All other such cases Berlyne terms "relative novelty," which is to say "they consist of previously experienced elements in unprecedented combinations..." Berlyne also distinguishes between "short-term novelty" and "long-term novelty," with the key difference being how long it has been since the viewer has seen something similar: a few minutes, hours, days, or even longer. Berlyne performed a battery of experiments on novelty, using a variety of stimuli, including color patches, patterns, and even sounds. He presented viewers with the same patterns on successive trials in some cases, as well as with each presentation separated in other cases by several trials, and in still others by several days. (See Figure 2.2A for a graph demonstrating the change in fixation time on varying and recurring patterns, and 2.2B for the pleasingness ratings on subsequent trials for two stimuli that varied in both shape and color). The experiments provide a strong case for novelty as one of the major determinants of aesthetic preference.

Berlyne presents three key findings associated with these experiments:

I. "Subjective novelty declines gradually as a stimulus is repeated several times in succession." [Note the general shapes of the solid curves in Figure 2.2A and 2.2B]



<sup>7</sup> For the sake of completeness, it is worth mentioning several key authors. Alexander Gerard was one of the first people (as well as Henry Home, 1765) to recognize that novelty could play a role in aesthetic preferences. He placed novelty first in his set of "simple principles" of taste, saying: "new objects...exalt and enliven the frame of mind, make it receive a strong impression from them, and thus render them in some measure agreeable" (Gerard, 1764). Other noteworthy authors to recognize the value of novelty include Lanz (1931/1968), Peckham (1965), Martindale (1969), Dix (1986), and Cupchik (1992).

2. "A stimulus is rated less novel when it resembles one that has been perceived within the last few minutes than when it does not." [Note the first magenta circle and second magenta circle in Figure 2.2B]
3. "A stimulus is rated as more novel the more it differs from what has just been experienced." [Compare the ratings of the magenta circles with the rating of the magenta star in Figure 2.2B]... (Berlyne, 1971).

Though it is not explicitly stated, Berlyne's earlier Wundt curve derivation for hedonic value also applies to novelty -- we prefer stimuli that are more novel, but only up to a point, after which they will be "too novel" and will fall into the negative tail of the graph of hedonic value shown in Figure 2.1D. Additionally, Berlyne demonstrates that the ratings of novelty in his experiments vary directly with ratings of interestingness (and pleasingness, as shown in the data reviewed here), making a direct link between his theory of novelty and his theory of aesthetic preference. Novelty thus deserves a place in any theory of aesthetic preference, a point that will arise in the analysis and critique of fluency theories of aesthetic preference to be presented later.

### Complexity

Berlyne demonstrated that on the initial test of "mean pleasingness" ratings of two complex and two simple patterns (as defined by the number of angles in the shape), most people strongly preferred the simple pattern. However, over 6 subsequent tests, separated by 8 presentations of each pattern, the preferences crossed over, and complex patterns became more preferred than the simple one (see Figure 2.3). This fairly straightforward experiment demonstrated that aesthetic preferences unfold and change over time in systematic ways that depend importantly on their complexity. The results also suggest why there might be discrepancies between people's

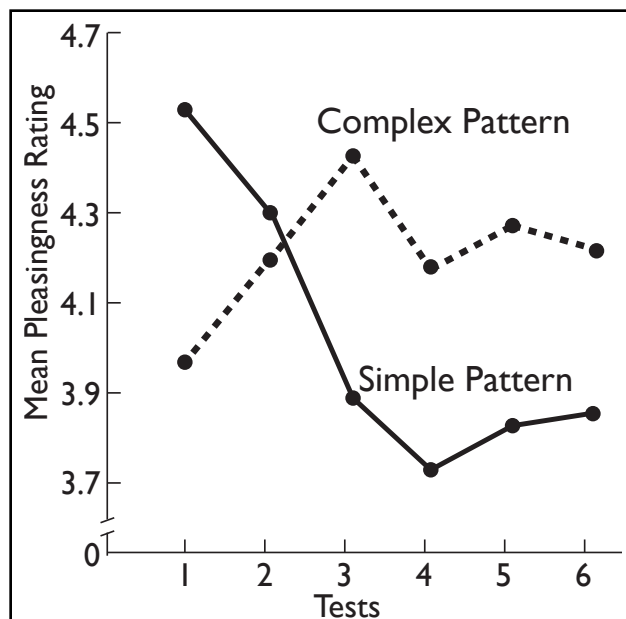


Figure 2.3. A graph adapted from one of Berlyne's experiments on pleasingness of simple and complex patterns, and how that pleasingness changes with subsequent exposures. Between each test, patterns were presented eight times.

aesthetic response to the different kinds of images that were sometimes chosen in laboratory settings versus those that are typically viewed in art galleries or people's homes: images that are chosen as more pleasing after a brief viewing are likely to be less complex than those chosen after prolonged viewing in a gallery or daily exposures in a home.

### Berlyne's Impact

Berlyne's impact on the field of aesthetics is hard to overstate. His theory touches on many key components of aesthetic experience, and almost all of it is supported by empirical results. He was the last prominent psychologist to devote a significant portion of his career so thoroughly to research on aesthetics. This is not to say that aesthetic research ended with him, for many more recent studies in psychology will be described over the course of this

dissertation. Nevertheless, he was the last major figure in the field of aesthetics who developed a systematic body of research together with a novel theoretical approach.

### Current Empirical Research: Another (Relatively) Dry Period

The study of aesthetics in psychology in recent years has not been as much of a sub-field of psychology as it has been seemingly random and unsystematic sprinkling of experimental studies. As with the first “dry period” I described, these studies do not especially build upon each other, even when conducted by the same author or group of authors.

A primary exception to this generalization was a coherent line of aesthetic research on directional preference: the study of systematic differences in people’s preference for images versus their left-right mirror reversals. It was sparked by an early claim by Heinrich Wölfflin (1928), as reported in Gaffron (1950), who postulated that aesthetically pleasing paintings generally have their principle figure or major area of interest located distinctly to the *right* of the physical center of the picture. Wölfflin and Gaffron suggest that this effect arises because people tend to scan pictures in an arc from lower left to upper right, so that content right of center is perceptually emphasized and therefore more salient. Although their claims were purely phenomenological, subsequent empirical work lends some credence to a rightward bias for major content to be toward the right side of complex pictures. These experiments typically investigated which of two complex photographs, paintings, or drawings people prefer between exact mirror-reflections of each other (e.g. Levy, 1976; McLaughlin, 1986; McLaughlin et al., 1983; Nachson et al., 1999). The results show that there is a relatively small but consistent preference for the version of the picture whose more significant content is on the right side, as Wölfflin (1928) suggested. The effect is not universal, however, being more pronounced for right-handed participants and even reversing somewhat for left-handers (Levy, 1976; McLaughlin, 1986). This finding has been interpreted as reflecting asymmetries in visual processing by the left *versus* right cerebral hemispheres (Levy, 1976). More recent research has examined cultural influences due to reading directions, however, which is more consistent with Wölfflin (1928) and Gaffron’s (1950) scanning direction hypothesis. A cross-cultural study of the asymmetry effect in picture preference found that viewers who read left-to-right (Russian) showed a right-side bias, whereas those who read right-to-left (Hebrew and Arabic) showed a left-side bias (Nachson et al., 1999). Interestingly, this emphasis on asymmetries in off-center compositions is somewhat at odds with the literature on balance and the aesthetic importance of centers, which will be the topic of the next section.

## Chapter 3

# Balance and Fluency

Two concepts in the modern literature bear special mention because of their general importance and specific relevance to the research reported here on spatial composition: balance and fluency.

### Balance

Balance is one of the most often discussed principles of spatial composition in art. Talk about composition inevitably leads to a physical, mechanical analogy: art that is balanced creates a perception of stability and equal spatial distribution of “weight.” To Rudolf Arnheim, balance is of such paramount importance to the study of visual aesthetics that it is the very first chapter of his seminal book, *Art and Visual Perception*. Arnheim is not able to define balance cleanly, however, in that he provides no essentialist definition. Rather, balance is discussed as being self-evident when present and as an amalgam of spatial position, size, color, and importance, depending on the example at hand.

Christopher Alexander (2002), who wrote the loftily-titled “The Nature of Order: An Essay on the Art of Building and the Nature of the Universe,” has an entire first volume that focuses on centers in art, architecture, and nature. In his theory, centers are “the building blocks of wholeness,” and he gives example after example, in the style of Arnheim (especially his *Power of the Center*), that he claims prove his point.

### Defining Balance Experimentally

At the other end of the spectrum from Arnheim and Alexander’s armchair aesthetics lie empirical researchers, such as Daniel Berlyne, Paul Locher, and Chris McManus. All of them agree with Arnheim about the importance of balance in art but are more experimentally oriented. Indeed, Locher and McManus have performed controlled studies aimed at understanding the nature of balance in art and aesthetic preference. Although Berlyne did not do empirical work on this subject, he did review and critique it extensively.

As mentioned earlier, there are two broad categories into which Berlyne places aesthetic principles: arousal-moderating and arousal-increasing. He categorized balance as an arousal-moderating device; i.e., because balance imparts a sense of stability, it is boring and decreases our enjoyment of a work of art. At a first blush, it would seem that Berlyne is saying that people do not prefer works of art that are balanced. Although this would seem to place him in conflict with Arnheim, Locher, and McManus, Berlyne is actually saying only that balance, *in isolation*, would decrease arousal. In conjunction with other arousal-increasing devices such as variety, novelty, conflict, and others, however, the stability balance provides can act as a powerful factor that causes other factors to be perceived in a coherent manner. For this reason, Berlyne termed balance a “special problem” in his text, along with proportion.

In a review of experiments on balance and art, Berlyne analyzes balance as one form of symmetry, a property often found in life and nature. In art, he notes, true symmetry is rare, but balance is ubiquitous. He describes an experiment by Pierce (1894) that utilized only three lines to examine the property of balance: one line, always the same length, was placed vertically in the center of the display to divide it; a second line, always fixed in a given display, but of varying length, width, and color, was placed in one of the resulting halves of the display, and a third line, also varying in length, width, and color was restricted to the other half of the display and was

movable by the observer. The stated task was to place the test line such that the entire display appeared balanced. Understandably from the analogy of physical balance, the length, width, and distance of the fixed line from the center all had an effect on where the test line was placed, and there were interactions among the four variable factors in the fixed line (length, width, distance from the center, and color) and the same four variables in the test line. All else being equal, test lines were placed further from the center than the fixed line if they were shorter or thinner than the fixed line, but they were placed closer to the center if they were longer or thicker (Pierce, 1894). Although these results are consistent with physics, color also had effects, thus defying a complete explanation by this mechanistic analogy of perception, because colors have no physical weight. This experiment was later replicated without the dividing line, and the results were the same (Puffer, 1903). From these experiments, Berlyne notes an important characteristic of balance: it cannot be explained in purely physical terms, but must be explained psychologically, as characteristics that have no physical weight (e.g., color) can alter the perceived balance of a display.

In "Balance in Pictures" (1985) Chris McManus went a step beyond Pierce by examining the issue of balance in art. Participants were asked to move a fulcrum left and right under each image until the image appeared balanced with respect to the fulcrum. In his first experiment, for which the displays were works of art, there were large differences between images, but decent agreement between participants. In the second experiment, McManus used abstract displays of one or two colored squares in an extension of the displays used by Pierce. In agreement with Pierce's results, McManus found that position was very important for balance judgments but that there were also significant interactions with color.

### Balance and Centers

It is nearly impossible to discuss balance in composition without talking about centers because balance implies equal distribution of elements around a center. Indeed, after writing *Art and Visual Perception*, Arnheim went on to write an entire book related to balance and centers, entitled *The Power of the Center*. Here he discusses at length how the center of a work of art can act as a fulcrum, weighing the pictorial elements on each opposing side of the frame (top/bottom as well as left/right), or as a hub, from which all points of interest emanate. He cites many paintings from many different regions, cultures, and time periods, pointing out cases of successful use of balance as well as cases where something -- at least in his view -- went awry.

In a demonstration of how sensitive we are to balance and centers (See Figure 3.1), Arnheim places within a square a black dot that is slightly off-center, and comments that its deviation is easily noticed. He goes on to argue that the same small deviation from another position (say, the halfway

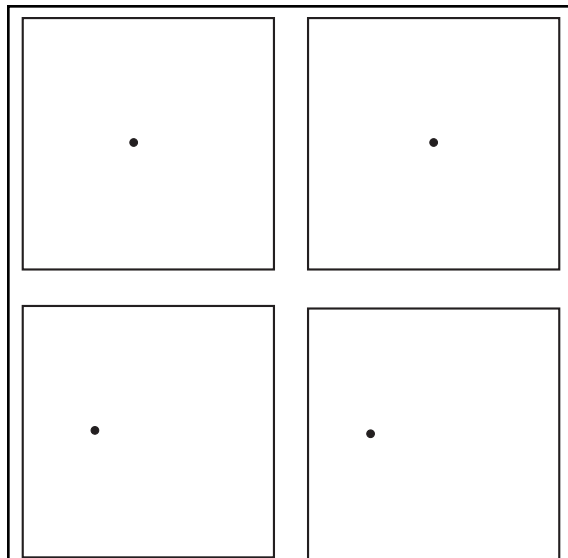


Figure 3.1. Four squares with a single dot in each one, after Arnheim, *The Power of the Center* (1988). One of the top two squares has a dot placed at the exact center; this is an easy distinction to make. One of the bottom two squares has a dot placed at the exact halfway point between the center and the left edge; this is a much more difficult distinction to make, illustrating our heightened sensitivity to centers.

point between the center of the square and one of the edges) would easily go unnoticed. He did not perform an empirical study to support his argument, but first-hand phenomenology can easily verify this fact by inspection of Figure 3.1. As in all of his writings, his examples have a self-evident quality that largely obviates the need for empirical research to support his claims.

Further evidence of the importance of centers in aesthetic response comes from Tyler (1998a, 1998b; 2003), who reported a strong bias for the placement of one of the two eyes in non-profile portraits of human faces to lie along the vertical midline of the frame. He found this central bias to be much more pronounced for the eye than for the face as a whole, the mouth, or even the single eye in profile portraits. This finding, although surprising, does not itself lend strong support to the aesthetic relevance of the center so much as it presupposes the importance of the center and uses it to support the special relevance of the eye (as opposed to the mouth or the whole face) to an aesthetically successful portrait from certain perspectives. Moreover, his conclusion was undermined when McManus and Thomas (2007) showed that people did not prefer the portrait in which one eye was centered to the same portrait in which neither eye was centered in a two-alternative forced choice task. They also performed a statistical simulation of portrait creation using Tyler's criteria and concluded that Tyler's analysis of existing portraits likely turned out as it did "because of geometric constraints on the placing of a relatively large object, the head, within a pictorial frame." Tyler (2007) rebutted their objections, saying that their statistical simulation lacked "appropriate assumptions to match the properties of the observed distributions in historical portraiture." He also provided data showing that when people were given freedom to draw heads on blank sheets of paper, more than half of the resultant drawings had one eye centered, an experiment that he points out is "much closer to the portraitists' state of mind."

A more obliquely relevant finding that nevertheless provides clear support for the perceptual salience of the center of a rectangular frame -- although not for its aesthetic salience -- was reported by Palmer (1991) in a series of studies on symmetry. Participants were asked to rate the 'goodness of fit' between a single small circular probe figure and a surrounding rectangular frame when the circle was located at one of 35 equally spaced positions in a 5 x 7 grid within the rectangle. Participants' average fit ratings are represented in Figure 3.2 by the diameter of the circles located at the corresponding position within the frame. By far the

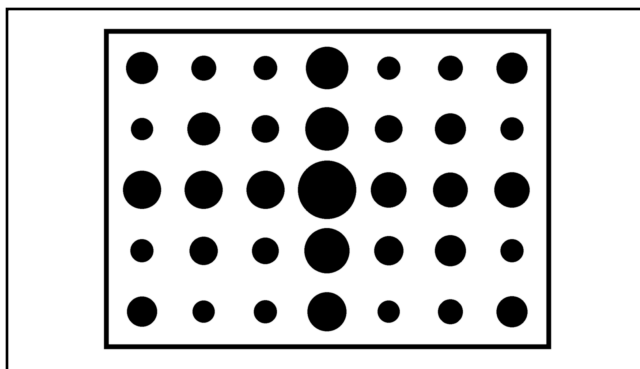


Figure 3.2. Goodness ratings of positions within a rectangular frame. Participants rated images of a single small circle at each of these 35 locations within a rectangle. The diameters of the circles depicted are proportional to the average 'goodness' rating on a 1-7 scale. The central, most symmetrical location was by far the 'best' position for the circle, with ratings diminishing for lesser degrees of symmetry.

highest ratings occurred when the circle was located at the center of the rectangle, where the rectangular frame is globally symmetric by reflection about both its vertical and horizontal axes. Indeed, the pattern of goodness ratings seems to be driven almost exclusively by symmetry structure. The next-highest ratings occurred when the probe circle lay on a single global axis of symmetry, with locations on the vertical axis being rated higher than those on the horizontal axis, consistent with the greater perceptual salience of vertical than horizontal symmetry (e.g., Palmer & Hemenway, 1978). Next highest were goodness ratings of locations along extended axes of local symmetry (the



angle bisectors), with the lowest ratings of all occurring when the circle lay on no symmetry axis at all. Similar results were obtained when a small circle was located at analogous positions within a trapezoidal shape, including the fact that the highest ratings occurred at the center. Although the relationship between these ratings of 'goodness of fit' and explicit judgments of aesthetic preference is not entirely obvious *a priori*, subsequent research from Palmer's laboratory has shown that goodness of fit relations between an object and its surrounding frame show strong positive correlations with aesthetic responses for similar visual displays (Griscom & Palmer, 2010). The dominant role of the center will also figure prominently in the research findings presented below.

### Balance, but not centers: The "Rule of Thirds"

A discussion of balance and composition is not complete without mentioning the "rule of thirds." This rule is a widely recommended heuristic for spatial composition that is most frequently discussed in photography. It states that when composing an image, the photographer should divide the frame into equal thirds horizontally and vertically and place the subject (i.e., the focal object) at one of the four points of intersection of the divisions (e.g., Smith, 1797; Field, 1845).<sup>8</sup> The rule of thirds clearly implies that the subject should not be placed at or even very near the center of the frame either horizontally or vertically to produce the most pleasing aesthetic effect, but at or near the third-points, which are distinctly off-center.

The rule of thirds has not yet been tested empirically but is nonetheless frequently endorsed in statements about pleasing compositions and is employed in many different user interfaces relating to image creation. For example, in the latest version<sup>9</sup> of Adobe Photoshop, indisputably the most widely-used photography editing program, clicking on the "crop" tool automatically places a rule-of-thirds grid within the cropping box, presumably so that the user can take it into account when cropping their image. Similarly, Canon, the largest manufacturer of point-and-shoot and single-lens-reflex (SLR) digital cameras worldwide, has a rule-of-thirds grid as an overlay for all images viewed on the LCD screen of the camera (but not through the viewfinder for the SLR cameras), again encouraging the user to place their region of interest (i.e., the focal object) at one of the intersection points. In fact, in the manuals for its latest cameras (both point and shoot and SLR), Canon provides the following information about composition and the rule of thirds:

Composition is how a shot is put together, the arrangement of visual subjects in the picture area. Determining where you put your subject(s) can dramatically change the quality of your shots. Composition is often based on an old rule called The Rule of Thirds. Here is how to use the Rule of Thirds: Take the rectangle of your LCD monitor and... draw the equivalent of a tic-tac-toe board.

You can activate two different types of Grid display in the Move Mode menu (such as a 'Grid 1'). This is a useful composition tool for practicing the 'Rule of Thirds' -- Note how the important elements of the frame (the girl, the base, the bat) are positioned on or near the grid lines. The

---

<sup>8</sup> In its original formulation by Nathaniel Smith, the rule is put forth as applying to the ratio of land or water to sky; that is, one-third to two-thirds, in either direction, and is vaguely expanded, as paraphrased by George Field in his book on color, to apply to "the crossing and breaking of lines and objects, etc." (Field, 1845).

<sup>9</sup> Creative Suite 5 (Version 12), as of this writing.

most important element, her face, is placed near one of the intersecting crosses (grid is simulated [in Figure 3.3]...).

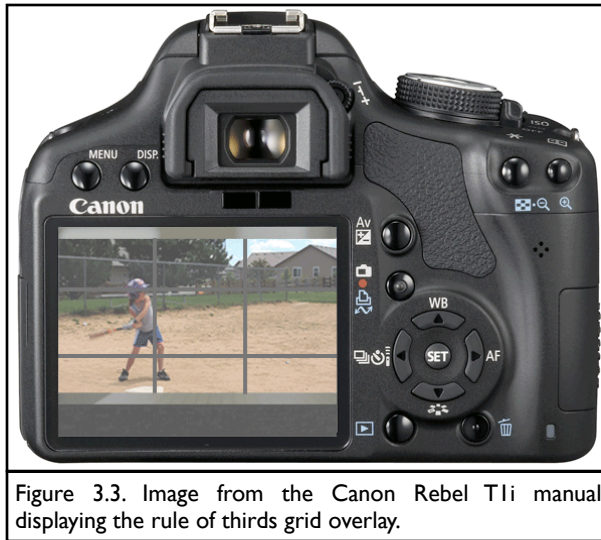


Figure 3.3. Image from the Canon Rebel T1i manual displaying the rule of thirds grid overlay.

Wherever the lines intersect is the best place for your subject. Good, dynamic, interesting photography rarely has the main person, place or thing dead center. Framing it off-center provides more depth to most compositions, be it a still photo or video footage.

In fact, the [Canon camera] makes using the Rule of Thirds easy! You can create these lines for reference on the LCD screen. (Canon Digital Learning Center, 2010).

With certain compositional principles the lack of experimental evidence has not proven problematic in light of later research; e.g., Arnheim's principles have by-and-large been validated. The same cannot be said for the rule of thirds, which is inconsistent with some of the results to be discussed later in this dissertation.

### Does balance affect aesthetic preference?

In addition to describing various forms of balance, it must be determined whether balance indeed affects aesthetic preference for pictures, as Arnheim and his followers would predict. Paul Locher has performed many such experiments with works of art. In one study, Nodine, Locher, and Krupinski (1993) recorded eye movements of participants either trained or untrained in the visual arts. One of their hypotheses was derived from Berlyne, who talked about two different forms of exploration in perception: one that actively explores and seeks out new stimulation and the other that fixates on a given portion of the display as the viewer contemplates the content or attempts to understand ambiguous content. These two forms of exploration make clean, distinct, and testable predictions about differences in the kinds of eye movements people will make in viewing pictures. The first kind of exploration implies eye movements that move around the picture space, not stopping in any one place for very long (or at least not at first), and the second implies eye movements that would be fewer and longer-lasting. Training in the visual arts was hypothesized to teach the participants how to view art "well," which is claimed to be in a manner corresponding to the first kind of exploration. With this in mind, it becomes clear why Berlyne viewed balance as a collative variable: it will cause the viewer to explore the picture space more completely and to form associations between all elements of the picture rather than focusing in on a single element and isolating it from the rest of the picture.<sup>10</sup> It is noteworthy that this conceptualization of balance has a strong resonance with Plato's concept of coherence and organic wholeness that was discussed earlier.

The displays used by Nodine, Locher, and Krupinski (1993) consisted of many works of art

---

<sup>10</sup> This statement seems likely to be composition dependent (e.g., an object with a single focal object placed in the center is balanced, even though such a composition would not likely encourage participants to examine the edges of the image), but Berlyne makes no such qualification of his statement.

from several time periods, each of which had two versions: the original and a modified version in which the balance of the work was altered either by removing an element altogether or by rearranging some of the elements. Indeed, the author's predictions about the effects of training were validated: trained artists made more eye movements covering more of the picture space than did untrained participants. They also more reliably preferred the originals, which the authors claim were the better or more carefully balanced of each pair.<sup>11</sup> They concluded that Berlyne was correct in his theory of perceptual exploration as it relates to aesthetics, that training in the visual arts does indeed affect perception of works of art in a way that makes people able to better appreciate the spatial composition of works of art, and that judgments of aesthetic preference are affected by both the method of exploratory perception and the viewer's level of training.

There are significant flaws in this study, however. First, the manner in which the original works were altered was not consistent. Sometimes the image was cropped, other times an element was removed, and still other times an element was relocated. . This mixture of different alteration strategies clouds the presumption of equality among the images in the "less-well-balanced" set, a point that will be revisited in describing the rationale for the present research methods. Secondly, it is possible that the only actual effects of training in the visual arts are to give the artist the ability to tell an original work from an altered one and a bias to prefer the original. Despite these problems, the study stands as one of the few rigorously empirical examinations of balance in art. Moreover, the effect of training on eye movements and exploratory perception is solid and arguably more important than the other more questionable conclusions that were drawn.

Another study by Locher, Overbeeke, and Stappers (1998) on balance took quite a different approach. Five years after the eye movement study just described, Locher set out to examine whether balance was a meaningful "organizing design principle" by making video recordings of adults (some who were trained in graphic design and others who were naive) while they constructed aesthetically-pleasing displays out of paper cutouts. Several different shapes were used -- such as leaves, circles, rectangles, and squares, all of different sizes -- although only one category of shapes was used for each display to be constructed. The participants were told to create the display that was most aesthetically pleasing by arranging the shapes within an empty frame on top of a piece of glass. Underneath the glass a video camera recorded the creation process. Locher then analyzed whether balance had played a role either in the process or the final product. It turned out that it had: about half of the displays displayed symmetry at some point in the creation process, and the vast majority of final designs had the physical weight evenly distributed about the center.

### Balance in Proportion: The Golden Ratio

Balance can be thought of as a stable spatial relationship among objects distributed within a frame, but there is the further question of how balanced or stable an object is in itself. This issue can be related to the Golden Ratio, which is a proportion thought to hold a special stability in the shape of an object, and thus for people's aesthetic response to it. The Golden Ratio 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. If the

---

<sup>11</sup> A potential confound of this result is that there was no explicit control of whether or not trained participants were familiar with the original images. The researchers briefly state that for some of the other comparisons, trained participants' judgments were split between originals and altered versions, but this is not clear evidence that familiarity did not play any role.

two parts are used as the length and width of a rectangle, they define a particular rectangular shape that should be particularly aesthetically pleasing because of its length-width balance (or aspect ratio). The Golden Ratio dates back to the time of Ancient Greece, where Euclid is often credited with the first explicit reference. Others noted its significance even earlier, however, including the Ancient Egyptians, who appear to have used the Golden Ratio in the construction of the Great Pyramids of Giza, and Pythagoras, who discovered that the human body contains several instances of the Golden Ratio in different places (Livio, 2002).

The claim of present interest is that the Golden Ratio is the most aesthetically-pleasing proportion to human observers. In addition to the Pyramids, the Golden Ratio is found in the structure of the Parthenon, the face of Leonardo's Mona Lisa, the compositions of Georges Seurat, Salvador Dali, and many others (Atalay, 2004; Livio, 2002). Indeed, it is an intriguing idea that we are aesthetically sensitive to such a specific proportion. In its strongest form, the claim is that if the Golden Ratio is present in a work of art, observers' response to that work will be more positive than it would otherwise be.

Any scientific, historically-minded discussion of the Golden Ratio begins with Fechner. His *Vorschule der Aesthetik* included an entire chapter on the Golden Ratio, which is the only portion of the volumes ever translated into English. Going about his studies in the same rigorously empirical manner with which he conducted psychophysical research, his first examination of the Golden Ratio employed his aforementioned "method of choice," in which participants viewed a set of ten rectangles with proportions ranging from square to 2.5:1, and chose the most aesthetically pleasing rectangle (Fechner, as cited in Green). The most preferred was indeed the Golden Rectangle, and preference diminished in both directions as the proportions deviated from it. In a less well-controlled study, Fechner gathered preference data on some 22,000 paintings to see whether those that were framed by Golden Rectangles were preferred to those that were not. This study did not yield any demonstrable preference for the Golden Ratio. Fechner performed several more experiments testing the Golden Ratio, and concluded his book chapter by stating that there is indeed something special about this proportion.

As the Golden Ratio became a popular topic of study among psychologists more confusion arose. One psychologist (Angier, 1903) attempted to replicate Fechner's findings and failed, stating that while the average preference was for the golden ratio, it was not preferred by individuals. Another replication (Haines & Davies, 1904) showed no preference standing out above any other in a large range surrounding the Golden Ratio. Lalo (1908) decided that neither replication had followed Fechner's original procedure closely enough; he did so to the letter, and once again replicated the finding. Over the next hundred years, studies of so-called "simple figure" aesthetics appeared often: many psychologists known for their work in unrelated fields examined people's aesthetic preferences for proportions. Perhaps the most notable of these was E.L. Thorndike, known for his puzzleboxes and studies of animal learning, who performed experiments on the Golden Ratio's presence in rectangles, triangles, crosses, and line patterns (Thorndike, 1917). He found a lack of evidence for the golden ratio in rectangles, crosses, and line patterns, but did find it in triangles. Thorndike points out that unlike some of the past research supporting an effect of the Golden Ratio, he did not present his objects in isolation, and that under such circumstances, participants' responses may have been different. However, no follow-up with the same objects or methods was ever performed.

To date, after 125 years of psychological study, the aesthetic salience of the Golden Ratio in human perception remains a controversial topic. Sets of experiments are performed that are interpreted one way by the experimenter -- either in support or in refutation of the Golden

Ratio hypothesis -- and in the opposite way by others who analyze the results differently. (Compare, for example, Green, 1995, on Boselie's research, with Boselie's own analysis in Boselie, 1984a, 1984b, 1992). Regardless, it has become one of the most widely cited spatial factors that may influence aesthetic preference.

## Fluency

Fluency is a term that originated with the study of language, where it loosely means how easily language can be processed by an individual (e.g., "Joe is fluent in German"). Within modern psychology, however, the term is much more broadly construed and is used to refer to the subjective feeling of ease of processing (see Alter & Oppenheimer, 2009, for a metareview). Additionally, some psychologists have claimed that fluency underlies aesthetic preferences. One of the earliest mentions of the concept that fluency represents in its current form -- a sort of pleasure from metacognition -- came more than 200 years before it emerged in the psychological literature. In his *Essay on Taste* (1764), Alexander Gerard began his discussion with the simple assertion that "The mind receives pleasure or pain, not only from the impulse of external objects, but also from the consciousness of its own operations and dispositions." However, this notion was not followed up for over a century, and when it was, it took a different form -- that of pleasing physiological rather than psychological processes -- before it returned to the form envisioned by Gerard.

### Precursor to Fluency: Physiological Aesthetics

The first cohesive thread of thought and research on fluency (though it was not yet called that) came from a group of scientists and philosophers who believed that ease of processing -- though in physiological rather than metacognitive terms -- played a role in aesthetic preference. Among the first of these was Grant Allen, who published a book in 1877 entitled *Physiological Aesthetics*, in which he stated that "beauty of Form is chiefly concerned with the muscular sweep of the eye in cognizing adjacent points," (Allen, 1877). Although Allen did not dwell on this point for long, it planted the seed for later cultivation by other scientists and philosophers. In his ground-breaking volume on beauty and aesthetics, George Santayana expanded on this idea:

In the curves we call flowing and graceful, we have... a more natural and rhythmical set of movements in the optic muscles; and certain points in the various gyrations make rhymes and assonances, as it were, to the eye that reaches them... It is easy to understand by analogy with the superficially observed conditions of pleasure, that such rhythms and harmonies should be delightful," (Santayana, 1896).

Santayana's assumption bears a striking resemblance to the current stance of fluency in aesthetics: that our perception of visual information (and by extension, other sorts of mental processing as well) can give a viewer pleasure if the perception is rhythmic and harmonious. There are two differences, though. One is a difference of mechanism: instead of metacognition, he designates physiology via eye movements, something that was untestable at the time that he wrote. The second difference is the cause of the pleasure: Santayana spoke of rhythms being pleasing, not of easy processing being pleasing.<sup>12</sup>

---

<sup>12</sup> Later research on reaction time with button-presses or other sorts of responses would demonstrate that rhythmic patterns or embedded sequences lead to faster response times. See, for example, Nissen & Bullemer, 1987, for the original serial reaction time task; also see, e.g., Cohen, Ivry, & Keele, 1990 for sequence learning.

Empirical tests of this assumption linking rhythm and harmony with preference came less than ten years later, with a study of eye tracking. George M. Stratton's 1902 experiments employed a coarse method of tracking eye movements in the service of aesthetic preference. He took long-exposure photographs of the eye and tracked a dot of light that was reflected from the cornea, a technique still used today, albeit employing more advanced infrared technology. This provided Stratton with a photographed line that could be interpreted using geometry and optics to extract the eye's movement over the time of the exposure. Stratton applied this technique to test empirically whether or not smooth eye movements really resulted in more pleasing images, and whether people actually even traced smooth lines smoothly at all (Stratton, 1902). It turned out that people's eye movements were never without at least some jaggedness. Figure 3.4 shows some of Stratton's original figures that demonstrate how smooth and jagged lines result in nearly identical eye movements.

The inescapable jaggedness of eye movements is no great surprise today to anyone who knows about saccades, but what makes Stratton's contribution especially useful is a brief comment he makes about phenomenological reports from his participants. Though the physiology betrays no difference whatsoever, the *phenomenology* was dramatically different when tracing a smooth line versus a jagged line with their eyes (see Figure 3.3) One participant, upon seeing her eye movement records, "was greatly surprised at their irregularity. She had felt sure, she said, that she had followed the lines with the greatest exactness," and Stratton notes that other subjects had similar reactions. This finding provides the connection that brings the early physiological literature on aesthetics into register with current notions of fluency in aesthetics. Objectively, one's eye movements tracing a smooth curved line may be no different from one's eye movements tracing a more jagged line,<sup>13</sup> but one's corresponding subjective experience varies dramatically.

### Fluency: Migration to Aesthetics

In the psychological literature, the beginning of fluency's conceptual transition into the domain of aesthetics began with Robert B. Zajonc's first studies on the "mere exposure effect." In a landmark paper, Zajonc (1968) demonstrated that "mere exposure of the individual to a stimulus object enhances his attitude toward it." Thus, the more often a participant saw a word, image, or other stimulus, the more they liked it. To connect this finding with fluency and fluency with aesthetic response, however, one final link was required: increases in exposure also had to produce increases in subjective ease of perception. This link was demonstrated by much of the literature on attribution and mere exposure (Bornstein, 1984; Bornstein & D'Agostino, 1994; Kunst-Wilson & Zajonc, 1980).

Factors other than mere exposure, and their relation to fluency and aesthetic response were first tested empirically by Reber, Winkielman, and Schwarz (1998). They presented novel line drawings of objects that were degraded somewhat in order to make identification difficult. These images were the "targets" and participants' were instructed to make recognition judgments as well as preference judgments, counterbalanced across participants. The targets were preceded by other line drawings serving as primes, which were so degraded that the participants were unaware that the primes had been presented at all. When the prime depicted the same object as the target drawing, not only did recognition time decrease, but so did preference. They concluded that fluency had influenced aesthetic preference.

---

<sup>13</sup> It is not unreasonable to say that at some point of high jaggedness, our eye movements would become more crooked as well.

Recent research has provided more evidence for a link between fluency and aesthetic preference. In a paper with the subtitle “Is beauty in the perceiver’s processing experience,” Reber, Schwarz, and Winkielman (2004) presented experiments that tested the effects of semantic and visual priming, figural goodness, symmetry, and other factors on preference. As with their earlier experiments, they showed that fluency, as measured primarily by reaction time, is highly predictive of aesthetic preference. Winkielman and colleagues (2006) performed further experiments in which participants were required to discriminate different, artificially constructed categories of dot patterns and faces. They then showed that the prototypes of these categories were both recognized faster and rated as more pleasing, again linking fluency and aesthetic preference.

The aforementioned studies by Winkielman, Reber, and their colleagues provide good evidence that fluency can explain at least some of people’s aesthetic preferences. However, fluency -- especially in reference to aesthetic preference -- has become a very broad term. A distinction between different kinds of fluency was suggested by Whittlesea (1993): perceptual fluency for ease of perceiving the perceptual characteristics of an object and conceptual fluency for ease of understanding, which could be broader than simple perceptual components. Reber and Winkielman, in their studies, state that their results are applicable to “processing fluency,” which they say includes both perceptual and conceptual fluency. This distinction and its ramifications will be discussed later in light of the results of the research I will present in the remainder of this dissertation.

## Chapter 4

# Research Rationale and Goals

As diverse as the research in the previous chapters has been, there are some very large gaps in our knowledge of aesthetic preference for spatial compositions. Most notably -- and most germane to the research at hand -- there have been no attempts thus far to explore preferences for spatial composition in a rigorously systematic and principled way. Indeed, what seems to be the simplest way to assess preference for different spatial compositions -- taking an object, placing it at different positions in a frame, and seeing which positions are most preferred -- is largely absent from the myriad methodologies used in previous aesthetic research. In a field that originated with the founder of psychophysics, why might this be the case?

### Prescriptive Approaches to Aesthetics

Most sources of aesthetic principles are incredibly ambitious. Many authors provide formulations of what viewers *ought to prefer* (e.g., Greenberg's 1982 essay, "Modernist Painting"). While this is a useful way to describe one's own theory of aesthetics, it is not a productive way to get at what principles are actually at work as typical viewers look at an image. The problem is that such treatises typically focus on of the success of examples that conform to the author's formulated principles while ignoring the counterexamples that do not.

A recent example of prescriptive theorizing about art and aesthetics in the field of psychology is the influential article "The Science of Art: A Neurological Theory of Aesthetic Experience," by V.S. Ramachandran and William Hirstein. Citing research on neural responses and the peak-shift effect, the authors state that "All art is caricature," with a necessary but disingenuous disclaimer -- "This is not *literally* true, of course, but as we shall see, it is true surprisingly often" (Ramachandran & Hirstein, 1999; emphasis original). They proceed to delineate many examples, ranging from political cartoons to sculptures to François Boucher's paintings of cherubs, which they deem to be caricatures of various types, and they conclude that the reason that we like the examples on their cherry-picked list is precisely *because* they are caricatures. The downfall of their theory is threefold: First, they do not give any account for the myriad caricatures that people typically do *not* like. Second, they have little, if anything, to say about pleasing images that are purely documentary and not caricatured in any way (e.g., photography). Third, by the end of the paper, they have crafted such a broad definition of "caricature" that they could likely wriggle out of virtually any counterexample to their statement. At this point, their term of interest -- *caricature* -- becomes functionally useless, a problem that will be revisited with the theory of fluency later on.

Some corpora of aesthetic theory are prescriptive to a lesser degree, analyzing existing works of art, and attempting to shed light on some of the principles that may underlie the aesthetic success of the works. Much of Rudolf Arnheim's writing falls into this category, as do large portions of Ernst H. Gombrich's *Art and Illusion* (1960). These two authors are much more cautious about their statements, and the principles that they examine are much more universal than Greenberg's statements about modernism and abstraction or Ramachandran and Hirstein's claims about caricatures.

As opposed to previous prescriptive approaches, the theoretical proposals developed in the present research are firmly grounded on measured preferences of a subset of the general population (i.e., undergraduates at UC Berkeley). I do not make claims about what *should* be preferred, only about what a representative sample of this population of viewers *does* prefer.



## Existing works of art vs. original images

The overwhelming majority of current literature (e.g., Konecni, 2003; Konecni & Cline, 2001; Levy, 1976; Locher, et al., 2005; McLaughlin, 1986; McLaughlin et al., 1983; Nodine, et al., 1993), and even some early studies (Fechner, 1876) use works of fine art as their displays. An important limitation of this constraint is that such studies do not afford access to a full range of aesthetic responses because such displays are (presumably) limited to the upper end of the spectrum. That is, it is reasonable to expect that effects of aesthetic preference will be greatly restricted if the set of images displayed to participants are all already of a certain minimum quality and kind, as opposed to the entire set of all possible images with regard to the factors being studied. Put another way, if one uses only images from art galleries or museums, one leaves out all of the less successful images that never make it into art galleries or museums in the first place.

With regard to spatial composition, this point is especially important. If one wants to use existing works of art to study spatial composition, there are two primary options, each with its own not-insignificant pitfall. The first option is to leave the original works untouched, and study compositional differences between images (e.g., Fechner's study of the Golden Ratio in paintings' aspect ratios and Konecni's (2003) studies of the Golden Ratio), however, this has the obvious and relatively insurmountable confound of content differences between images. In the case of one of Fechner's Golden Ratio experiments, for example, if a large group of people is presented with one hundred paintings from a museum, it would be fairly difficult to assert that any difference in preference would be due to the aspect ratio of the images and not to the large and uncontrolled differences among them in content. (It would be equally difficult to argue that people liked an image with a certain content simply because the image was framed by a Golden Rectangle). Additionally, the literature on mere exposure effects demonstrates that prior viewing of the same art works is likely to have some influence on their subsequent preferences, introducing another confound, even if participants do not remember seeing such images previously. (See Zajonc, 1968, for the original research; see Bornstein, 1989 for an excellent metareview of the mere exposure effect).

The second option circumvents this issue by using a single image and altering it to arrive at different compositions of the same content. In this case, however, two new problems arise: First, mere exposure effects imply that the original may be preferred simply because it has been seen before. (See Cutting, 2006, for an excellent survey and explication of this problem in canons of artwork, specifically, the Impressionist canon). Second, altered versions of the same image having the "same content" can become a difficult claim to make when dealing with detailed works of art. Take for example, Seurat's *Les Poseuses* (1888), which Nodine, Locher, and Krupinski altered

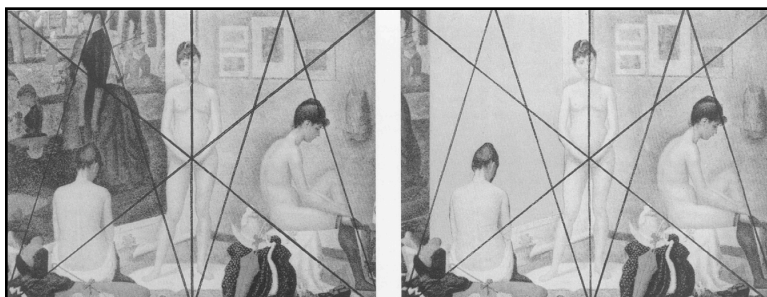


Figure 4.1. Two versions of Georges Seurat's *Les Poseuses* (1888): Original (left), with compositional overlay; revised version (right), also with compositional overlay. Taken from Nodine, Locher, & Krupinski (1993).

in their study of eye movements, aesthetic preference, and spatial composition. Figure 4.1 shows the two versions of the painting that are treated as equals in the study (original on the left, altered on the right). It does not take long to see that the two images are not, in fact, equal on the basis of content, as the original contains an entire figure (a painting on a wall) that is

absent in the altered image. More generally, unless the image in question is made up of interchangeable puzzle pieces, it is impossible to move or remove any elements without at the very least changing the content to the extent that more or less background will be visible in one of the versions, and this content change could easily alter preference.

I therefore think it best to create novel images for use in studies of spatial composition. This allows control over every aspect of the image -- content, composition, color -- while still varying whatever factors are deemed important to study.

### Some Thoughts on “Artistic” Images

A few remarks are in order about how the present results are (or might be) relevant to the scientific study of art. First, I have quite consciously avoided claims about the relevance of these findings for deciding how ‘artistic’ various images are, focusing instead on simpler claims about people’s preferences based on aesthetic aspects of experience. There are many reasons for this, not the least of which is the thorny question of how ‘art’ should be defined in the first place. I particularly intend to exclude, insofar as possible, the role of cultural/institutional factors in adjudicating issues of artistic merit, such as the opinions of art critics, museum curators, and other art experts (e.g. see Cutting, 2006). The participants in the present studies have well-defined and surprisingly consistent aesthetic biases, even though none of them would likely be counted as an expert in the realm of visual art. I am indeed interested in how people’s aesthetic preferences might differ as a function of expertise and formal artistic training, but I do not address that topic in the research described here. Nevertheless, I do believe that aesthetic response plays an important role in the understanding and assessment of art, and to that extent, the present findings are at least relevant to a scientific understanding of art.

Second, I readily acknowledge that the pictures that participants are asked to judge aesthetically are not beautiful or artistic in any strong sense. Indeed, most people would probably describe them as relatively uninspiring, if not downright ugly. Nevertheless, no participant ever indicated any difficulty in making the aesthetic judgments they were asked to make, and their preferences exhibited striking commonalities. If the pictures had been more aesthetically pleasing, they would certainly be judged as more ‘artistic’ overall, but I expect that the results corresponding to those reported here would be, for the most part, unchanged.

### Building a Science of Aesthetics

I agree with Fechner’s assessment of how a science of aesthetics ought to be built. It must have as its beginning simple and clean studies of images and the relatively low-level properties that give rise to different aesthetic preferences. Only with this solid foundation can a science of aesthetics be constructed that has the hope to address more challenging and deeper questions of aesthetic experience.

My goal in this dissertation is to provide a portion of this foundation through the empirical research I report below. In Chapters 5 and 6, I will present nine experiments that serve as the foundation of an aesthetic theory, “representational fit,” which will come into its fullest description with the four experiments that I present in Chapter 8. In the end, I will argue that representational fit is able not only to account for a wider range of aesthetic experiences than other current theories, but also to unify elements of the literature on aesthetics, visual perception, Gestalt psychology, and even philosophy.

# Chapter 5

## Experiments on Horizontal Position

The research I report in this section is a series of four studies designed to understand some of the principles that underlie aesthetic response to two important variables in spatial composition: the horizontal position and facing direction of a single meaningful object relative to a surrounding rectangular frame. Experiments 1.1 and 1.2 illustrate the primary method used to investigate such compositional issues: two-alternative forced choices (2AFC) of aesthetic preference. Participants were shown two pictures that differ only in the spatial framing variable (s) of interest and were asked to indicate which picture they prefer aesthetically. In this way all other differences were neutralized – particularly aesthetic response to the object(s) depicted – isolating the effects of compositional variables. These precise 2AFC measures were augmented with other tasks allowing greater freedom of choice, such as the method of constrained adjustment in Experiment 1.3 and free-choice in framing photographs in Experiment 1.4. Finally, Experiment 1.5 is an analysis of stock photographs, where pre-existing image segmentations of actual photographs were used to see whether the results found in Experiments 1.1-1.4 would generalize. Because all of the present measures are specifically designed to eliminate the effects of content, the research strategy differs radically from, but is complimentary to, research aimed at determining what perceptual content participants find pleasing (e.g., Biederman & Vessel, 2006). Both kinds of research are clearly necessary to understand why people prefer the pictures they do.

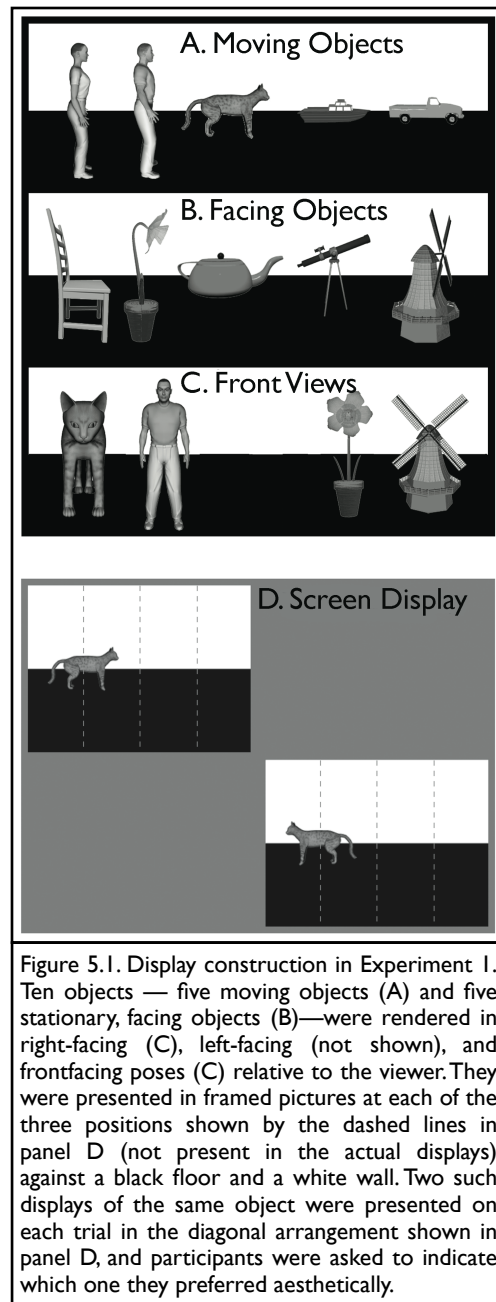
### Experiment 1.1: Position and Direction of Moving vs. Facing Objects

The first experiment was an exploratory study aimed at finding out whether a psychophysical approach to studying the aesthetics of spatial framing was even viable. Jaded by the old adage “there’s no accounting for taste,” there was an initial concern that huge individual differences might swamp any systematic effects. This did not turn out to be a problem, because the results were both orderly and robust.

Starting from first principles rather than well-known heuristics, such as the *rule-of-thirds*, two variables of obvious interest were examined: the location of a single object and the direction in which it faces (if it has a perceptual front) relative to a surrounding rectangular frame. We studied the effects of these variables on preferences for the composition of pictures depicting directed objects of two kinds: objects that move in a particular direction and those that merely face in a particular direction (see Figure 5.1). The “moving” objects were chosen to be representative of objects that typically move horizontally toward their front: a man, woman, car, boat, and cat. The “merely facing” objects were typically stationary, but nevertheless have a well-defined, canonical front and back: a chair, teapot, flower, windmill, and telescope. We thought that moving objects might exhibit a stronger directional bias because participants might expect the corresponding real object’s motion to take it to or toward the center of the frame, whereas the merely-facing objects would not. The operational definition of the “location” of an object was the location of its central point (midway between its left and right extremities) relative to the center of the frame, and “facing into the frame” was defined to mean that the direction the object faces (i.e., the direction from the object’s center to its front) is the same as the direction

from the object's center to the frame's center.<sup>14</sup>

A rectangular frame with a 4:3 aspect ratio – the same as a standard television screen – was used and objects were placed at three locations along the horizontal midline: in the geometric center of the frame and at its quarter points, as illustrated by the dashed lines in Figure 5.1D. We studied front views of the same objects, which were roughly symmetrical and thus not



laterally directed, to get a measurement of positional preferences unaffected by directional preferences. We expected the results to show a *center bias*: i.e., that participants' preferences would be strongly peaked at the center and approximately symmetrical, although in light of the previous research reviewed above, they might be somewhat skewed toward the right side. We also studied left- and right-facing views, which we expected to show both an approximately symmetrical center bias and a strongly asymmetrical *inward bias*: i.e., that participants would prefer pictures in which the object faces into, rather than out of, the frame. To avoid complications arising from possible preferences for front- versus side-facing views, the 2AFC pairs always contained the same view of the same object, with front views paired only with other front views and side views paired only with other side views.

### Method

**Participants.** All nine participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age was approximately 19 years. All were naïve to the purpose and nature of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

**Design.** There were 60 paired comparisons of front-view images, resulting from the orthogonal combination of 10 objects (5 moving and 5 facing objects) and 6 image pairs defined by the permutations of 3 frame positions taken 2 at a time. There were also 300 paired comparisons of side-view images, resulting from the orthogonal combination of the same 10 objects and 30 image pairs defined by the permutations of 6 frame

<sup>14</sup> Another way of characterizing the facing bias is to say that the direction the object faces is opposite the direction from the frame's center to the object's center. This alternative characterization of the facing bias thus can be viewed as emphasizing directional opposition and balance: i.e., the object's unbalanced position to one side is balanced by its facing in the opposite direction. These formulations cannot be disambiguated in the present research, and choose to discuss the hypothesis in terms of directional consistency because of its more natural mapping to the usual characterization in terms of the object "facing into the frame."

positions and directions taken 2 at a time. The screen locations of the two images in each trial were always upper-left and lower-right to reduce possible alignment effects and were counterbalanced by the just-described design of image pairs. The order of the trials was randomized by Presentation software (see <http://www.neurobs.com>) that controlled the experiment.

Displays. The three frame locations were defined by the precise geometry of the frame: the points at which the left quarter-line, the vertical midline, and the right quarter-line intersected the horizontal midline of the rectangle. The centers of the objects were defined as the central point horizontally between the most extreme points at the left and right sides of the object.

Each screen consisted of two grayscale images of an object on a black ground plane against a white background, with the horizon placed along the horizontal midline of the frame. One image was located in the upper-left corner of the computer screen and the other was in the lower-right corner so that the images were not aligned on either the horizontal or vertical dimension (see Figure 5.1D). Each image was separated from the edges of the screen by approximately .75 cm, and placed on a neutral gray background, as shown in Figure 5.1D. Objects were modeled and rendered using Poser 6 software, and the resulting images and screens were constructed using Adobe Photoshop CS2. The display was 18" diagonally and the resolution was 640 x 480 at a refresh rate of 85 Hz.

Procedure. Participants viewed the computer screen from approximately 60 cm. They were instructed to look at each screen and to press a button (left or right) indicating which image they "liked more." They proceeded at their own pace and were given the opportunity to take a short break after every 60 trials.

## Results and Discussion

Participants' responses were scored for the probability with which they chose each picture in each of the 36 2AFC pairs of pictures for each object (i.e., the 6 pairs of center-facing views and the 30 pairs of side-facing views). To create a composite measure of the aesthetic response to each picture, the average probability of choosing that picture across all of its pair-wise comparisons was computed. The resulting probabilities, averaged over participants and objects, are plotted in Figure 5.2 as a function of object location for the center-, right-, and left-facing objects. Because of concern about statistical assumptions for probabilities, we also analyzed the choice by computing Bradley-Terry-Luce (BTL) scale values from the 2AFC data for each participant (Bradley & Terry, 1952; Luce, 1959).<sup>15</sup> Unsurprisingly, these values were very highly correlated with the probability data ( $r=0.96$ ), but they allowed us to use a somewhat more

---

<sup>15</sup> In statistical analyses of binary choice data, such as those plotted in Figure 5.2, caution must be exercised because the variances of probability distributions do not conform to standard variability assumptions of the analysis of variance, being necessarily lower at extreme values (i.e., near 0.0 and 1.0) than at moderate values. We therefore analyzed our results using both the probability data plotted in Figures 3, 5, 6 and 7 and Bradley-Terry-Luce (BTL) scale values derived for each individual participant from their 2AFC choice data. The scaling algorithm was written in the statistical programming language R (R Development Core Team, 2006) using the Bradley-Terry package of Firth (2005). These values were then analyzed with a repeated-measures analysis of variance. The BTL scaling procedure derives scale values from binary choice data such that the scaled values have more equal variances across the scale, thus overcoming potential problems due to unequal variances in probability data. In fact, the original probabilities and the BTL scaled values were essentially the same ( $r = .96$  in Experiment 1 and  $r = .97$  in Experiment 2), because the data include relatively few points in the extreme ranges. We chose to report the data in the figures as untransformed probabilities – i.e., percentages of trials on which a given image was chosen over all others – because they are more intuitively meaningful than BTL scaled values. We report statistical analyses for both the untransformed average probabilities and the transformed BTL values (the latter in square brackets) for the overall analyses of variance. Given the close correspondence between the probability data and the BTL scale values, the middle range of most of the probabilities, the magnitude of the effects, and the similarity of the outcomes, we did not repeat the statistical tests with BTL values for subsequent specific comparisons, but simply report the analyses based on mean probability measures.

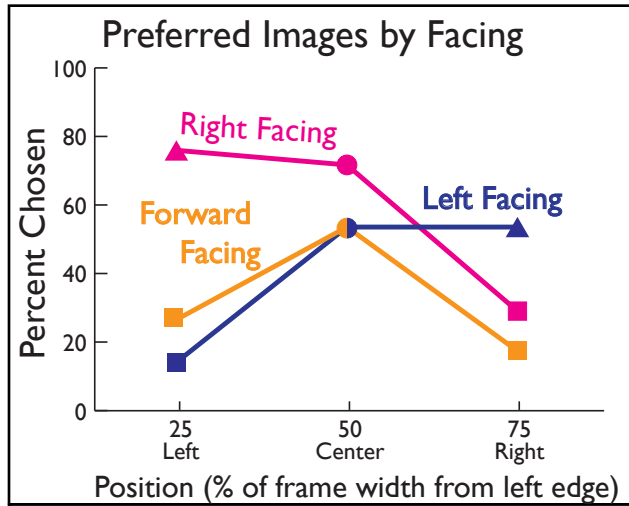


Figure 5.2. Results of Experiment 1.1. The average percentage of times that the given image was preferred over all possible comparisons is plotted as a function of the position of the object's center for left-, right-, and forward-facing views.

cautious statistical analysis. The results of analyses of variance based on the BTL scaled data are reported below in square brackets following those based on the probability data.

The overall within-participants analyses of variance showed main effects of position ( $F(2,16) = 10.36 [4.35], p < .01_{.03}$ ), facing condition ( $F(2,16) = 33.62 [5.94], p < .001 [.01]$ ), and their interaction ( $F(4,32) = 25.16 [3.98], p < .001 [.01]$ ). The center-facing views, which were only compared with other center-facing views of the same object, demonstrate a strong, symmetrical preference for the center position, which was chosen more frequently than either the left-side or right-side positions ( $F(1,8) = 11.99, p < .01$ ), which did not differ reliably from each other ( $F(1,8) = 2.33, p > .10$ ).

The left-facing and right-facing conditions produced a dramatically different and highly asymmetrical pattern of results. The central position (plotted as circles in Figure 5.2) was strongly preferred to the side position for which the object faced out of the frame (plotted as squares in Figure 5.2) for both the left- and right-facing views ( $F(1,8) = 56.87, 71.33$ , respectively,  $p < .001$ ) but did not differ from the side position at which the object faced into the frame (plotted as triangles in Figure 5.2) ( $F < 1$ , in both cases). The two lateral positions differed significantly as well, with the view facing into the frame (triangles) being preferred over the view facing out of the frame (squares) for both the left- and the right-facing objects ( $F(1,8) = 37.86, 22.72$ , respectively,  $p < .001$ ). The pattern of results is thus consistent with both of the initially hypothesized preferences: participants exhibit both a strong center bias and a strong inward bias.

The results also show a fairly clear preference for right-facing objects over left-facing ones. The rightward bias can be seen by comparing the corresponding conditions in Figure 5.2 for the side-facing conditions: The right-facing probability is greater than the left-facing probability at all three positions: the center position (circles), the inward facing position (triangles), and the outward facing position (squares) ( $F(1,8) = 11.46, 62.53, 7.10, p < .001, .001, .02$  respectively). This rightward bias suggests a preference for the object facing in a direction consistent with the left-to-right reading direction in English (cf., Nachson et al., 1999) and/or the bottom-left-to-top-right scan path hypothesized by Wölfflin (1928) and Gaffron (1950). It may also be related to hemispheric processing and handedness (cf., Levy, 1976; Mclaughlin, 1985).

There was, by definition, no main effect due to moving objects versus facing objects, because all comparisons were within-object. There was a marginal interaction between object type and facing condition ( $F(2,16) = 3.83, p < .05$ ), but it has no obvious interpretation: People slightly preferred the moving objects to face leftward and the merely-facing objects to face rightward in the side-view conditions. It is unclear why this might occur. The three-way interaction that would have indicated stronger facing effects for moving than facing objects was not present ( $F < 1$ ). It therefore seems unlikely that either of the facing effects is related to participants' expectations that the object is about to or could move in the direction it faces. The pattern of

results shown in Figure 5.2 thus appears to be robust with respect to moving versus merely-facing objects.

There are at least two plausible explanations of the inward bias we found in this experiment, which I will discuss as the “directional consistency” and “front position” hypotheses. The directional consistency hypothesis is that people prefer the intrinsic directedness of the object (i.e., from its center to its front) to be consistent with the direction from the object to the center of the frame (i.e., from the object’s center to the frame’s center). By this account, people prefer facing objects to be directed so that their front is in the same direction relative to the object-center as the frame-center is. An alternative hypothesis can be formulated in terms of the position of the object’s front: People may simply prefer the front of the object to be located as near the center of the frame as possible (i.e., there may simply be a center bias for the object’s front rather than its center). This possibility is consistent with the inward bias we obtained because, for any non-centered position, the front of the object will be closer to the frame-center when it faces into the frame than when it faces out of the frame.<sup>16</sup> The present data cannot discriminate between these two possibilities, but Experiment 1.2 provides a test that does.

### Experiment 1.2: Position and Direction of Objects with Different Aspect Ratios

In the second experiment, we examined more closely people’s aesthetic preferences due to the interaction between position and direction. We increased the spatial resolution by using seven equally spaced locations within the range covered in Experiment 1.1 and looked at possible shape-based directional effects by varying the aspect ratios of the objects depicted. We were particularly interested in whether the preference functions for left- and right-facing objects would continue to have their maximum at the center, or whether they might actually peak off-center on the side at which the object faces into the frame. This possibility bears directly on the front position account of the inward bias because it predicts that people should prefer an off-center position when it the object’s front at the frame’s center. (The directional consistency explanation does not necessarily make this prediction, although it can be compatible with it). We also increased the number of positions because we wanted to examine the precise shape of the preference functions in terms of the center bias, which should be a symmetrical, inverted U-shaped function that peaks at the central position, and the inward bias, which should appear as a monotonic function of position that increases toward the left side for right-facing objects and toward the right side for left-facing objects.

In addition, we varied the aspect ratio of the objects to see how this global shape parameter would affect the frame-relative facing effect. As illustrated in Figure 5.3, we included two tall, thin objects that were vertically oriented (a man and a flower), two objects that were about equal in height and width (a teapot and a rocking horse), and two short, wide objects that were horizontally oriented (a wolf and a jeep). The front position hypothesis predicts that the inward bias effect should be weakest for the tall, thin vertical objects (because the distance from frame center to the front of the object changes little when its facing direction is reversed), and strongest for the short, wide, horizontal ones (because the distance from frame center to the front of the object changes greatly when its facing direction is reversed).

---

<sup>16</sup> This hypothesis is similar in certain respects to Tyler’s (1998a, 1998b) finding that an eye of the subject typically falls on the centerline of a portrait, in that it defines the preference in terms of a specified part of the object being located in the center.

In order to reduce the pairwise comparisons to a manageable number in the face of the four additional positions, the forward-facing views were eliminated and comparisons were only made for each side-facing view at each position with (a) all other views that showed the same object facing in the same direction (the “same-facing” comparisons) and (b) the single view of the same object at the same position that faced in the opposite direction (the “opposite-facing” comparisons), as indicated in Figure 5.3.

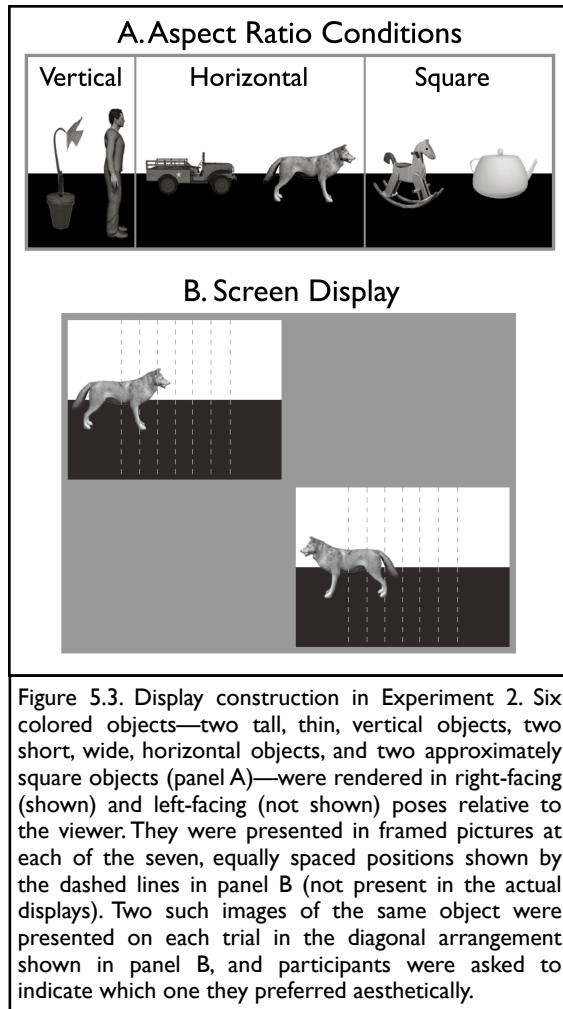


Figure 5.3. Display construction in Experiment 2. Six colored objects—two tall, thin, vertical objects, two short, wide, horizontal objects, and two approximately square objects (panel A)—were rendered in right-facing (shown) and left-facing (not shown) poses relative to the viewer. They were presented in framed pictures at each of the seven, equally spaced positions shown by the dashed lines in panel B (not present in the actual displays). Two such images of the same object were presented on each trial in the diagonal arrangement shown in panel B, and participants were asked to indicate which one they preferred aesthetically.

### Method

**Participants.** All but one of the 18 participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course, and the remaining participant was a lab manager in the psychology department. Their mean age was 19.6 years. All participants were naïve to the purpose and nature of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol. The data from one participant was eliminated due to his/her failure to follow the instructions.

**Design.** The experiment consisted of 588 trials, defined by the 98 pairwise comparisons for each of the 6 objects. The 98 pairwise comparisons for each object consisted of the 14 pairs of opposite-facing comparisons at the same position, the 42 pairs of left-facing comparisons at different positions (the permutations of 7 positions taken 2 at a time), and the 42 pairs of right-facing comparisons at different positions. These pairs are counterbalanced for screen position because the permutations necessarily contain both spatial arrangements (i.e., with each picture appearing once in the upper right and once in the lower left positions).

**Displays.** The objects in the pictures of Experiment 1.2 were rendered in color using Poser 6 and Adobe Photoshop software, but were still placed in front of a black ground plane and white wall plane. The monitor measured 19” diagonally, but the resolution and viewing distance were unchanged from Experiment 1.1.

**Procedure.** The procedure for Experiment 1.2 was identical to that in Experiment 1.1, except that participants were given a chance to take a break every 98 trials, resulting in 5 possible breaks during the experiment, rather than 6 as in Experiment 1.1.

### Results and Discussion

The results were computed, as in Experiment 1.1, using both average probabilities of aesthetic preference and Bradley-Terry-Luce scale values.<sup>4</sup> Once again, the two measures were



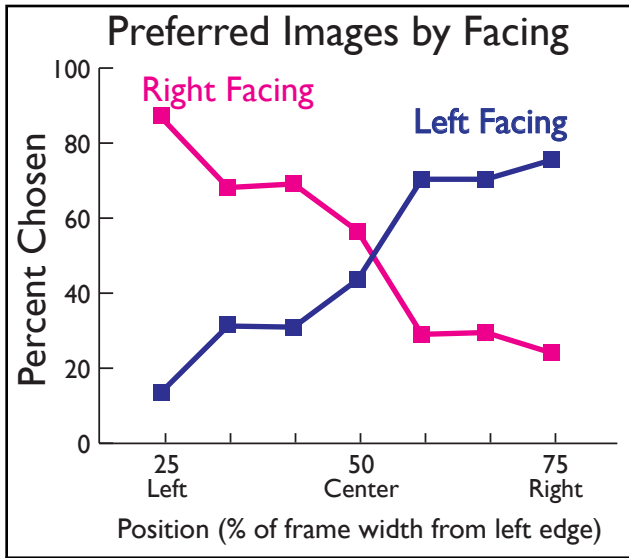


Figure 5.4. Results of Experiment 1.2 for opposite-facing conditions. The average percentage of times the given image was preferred over the same object at the same position but reflected about its geometric center is plotted as a function of the position of the object's center for left-facing and right-facing views. (Note that for each position, the two curves plotted here must sum to 1.0, because the participant was forced to choose one of the two views in each comparison. This fact explains the vertical symmetry of the two functions and dictates that only one function be analyzed statistically.)

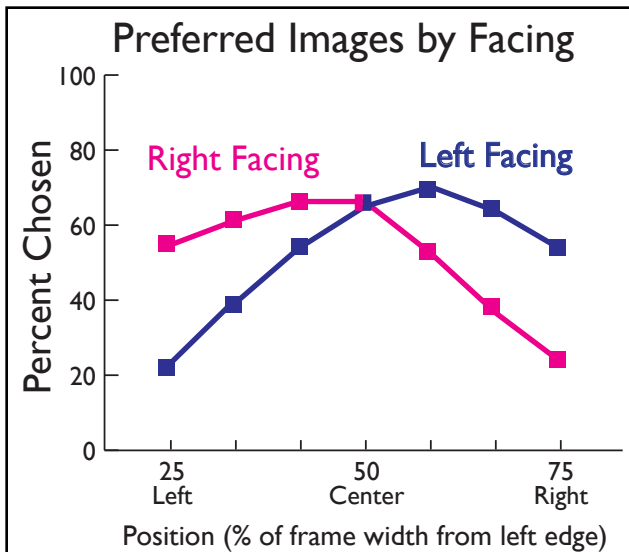


Figure 5.5. Results of Experiment 1.2 for same-facing comparisons. The average percentage of times the given image was preferred over all comparisons in which the same object faced in the same direction is plotted as a function of the position of the object's center for left-facing and right-facing views.

so strongly correlated ( $r = .97$ ) that I used the BTL values only in the overall analyses of variance and the tests of linear and quadratic trends, for which the quantitative structure of the data is particularly important.

The data for the opposite-facing conditions, averaged over participants and objects, are plotted in Figure 5.4 as a function of position for the left-facing and right-facing views. Because these data come from comparisons in which the global position of the object was the same in both pictures, it should reveal any directional component in relatively pure form. Indeed, there is a strong main effect of position ( $F(6,96) = 38.11, p < .001$ ) that increases dramatically and monotonically from left to right for the left-facing objects and from right to left for right-facing objects. Further analyses show that this function has a significant linear component ( $F(1,16) = 67.68, p < .001$ ) and no reliable quadratic component ( $F(1,16) = 1.39, p > .10$ ). These results are thus entirely consistent with the hypothesized inward bias for objects to face into the frame. There is also a slight bias toward preferring right-facing objects, as can be seen at the central and outermost positions, but this bias is not statistically reliable ( $F < 1$ ).

The data for the same-facing conditions were treated in the same way as the data in Experiment 1.1: they were averaged over all pairwise comparisons containing the given position and facing condition to arrive at a single measure of aesthetic preference for each condition and were subjected to BTL scaling. These data, averaged over participants and objects, are plotted in Figure 5.5 as a function of position. Overall within-participants analyses of variance indicated a large interaction between left/right facing condition and the seven positions ( $F(6,96) = 19.05 [32.81], p > .001 [.001]$ ), which is evident in the cross-over of the two functions in Figure 5.5. No rightward facing bias could possibly be reflected in these data

because they do not include any opposite facing comparisons. For both the left- and right-facing conditions, both the linear component ( $F(1,16) = 42.61$  [29.08], 34.35 [32.47],  $p < .001$  [.001]) and the quadratic component ( $F(1,16) = 43.34$  [30.76], 24.16 [19.73],  $p < .001$  [.001]) were statistically reliable.<sup>17</sup> This outcome indicates that the data contain, as expected, both an approximately linear inward facing bias and an inverted U-shaped center bias.

The functions for left- and right-facing objects do, in fact, appear to have their maxima somewhat off-center, at around 42 and 58 percent of the way from the right and left edges, respectively, but the curves are so broad that no statistical differences are evident between these points and their immediate neighbors. Note, however, that there is no evidence favoring strong peaks at the 33 and 67 percent positions, as predicted by the rule of thirds, both of which were explicitly present in this experiment. Moreover, the data clearly reinforce the conclusion from Experiment I that the rule of thirds is properly applied only when a single focal object is directed inward. Indeed, if it were applied such that the object faced outward, the aesthetic effect would be decidedly negative for most viewers.<sup>18</sup>

No main effects due to the aspect ratio of objects are possible in this experiment because different objects were never compared to each other. Interactions between aspect ratio and other variables are possible, however. To simplify these analyses, I first recoded the left/right facing factor to reflect whether objects face into or out of the frame, analogous to reflecting either the left-facing or the right-facing curve (but not both) in Figure 5.5 about a vertical axis. This recoding effectively eliminated any main effects and interactions due to the facing factor and revealed a small but significant interaction between aspect ratio and position ( $F(12,192) = 3.02$ ,  $p < .001$ ). The nature of this interaction can be seen in Figure 5.6: There are more extreme positional variations for the short, wide objects (the wolf and jeep) than for the tall, thin objects (the man and flower) ( $F(6,96) = 6.57$ ,  $p < .001$ ), with intermediate effects for the objects of intermediate aspect ratio (the teapot and rocking horse), being larger than the effects for the tall, thin objects ( $F(6,96) = 2.12$ ,  $p < .05$ ) and slightly, but not significantly, smaller than the effects for the short, wide objects ( $F(6,96) = 1.34$ ,  $p > .10$ ). Importantly, the peak position for the short, wide objects falls off-center toward the side at which it faces into the frame and received reliably higher ratings than the central position ( $F(1,16) = 9.50$ ,  $p < .01$ ). A similar trend is evident for the squarish objects, but it does not reach statistical significance ( $F(1,16) = 2.50$ ,  $p > .10$ ). This pattern of results is thus consistent with the predictions of the front position hypothesis, which postulates that the preference for a given object in a given position will be determined by the distance of its front from the center of the frame. Oddly, no corresponding differences due to aspect ratio were evident in the opposite-facing conditions, perhaps due to the smaller number of observations per data point.

To explore the extent to which the same-facing data can be predicted by variables relevant to the directional consistency and front position hypotheses, I performed a linear regression analysis of the positional effects for each of the six individual objects. The predictor variables I used were the distance of the object's center from the frame's center (the object center variable), the distance of the object's front from the frame's center (the front center variable),

---

<sup>17</sup> The Bradley-Terry-Luce model, as implemented, requires that the table of comparisons be complete. As a result, the BTL values reported are not broken down into separate same-facing and reverse-facing components, thus taking all of the data into account.

<sup>18</sup> This is not to say that having a single focal object face out of the picture always produces a poor aesthetic effect. In some contexts, violating expectations in this way may work quite well. This question will be addressed in the third set of experiments, Chapter 7.

and the object's directional consistency with respect to the frame (the directional consistency variable). I defined the distance of an object's front from the frame's center by subjectively determining the rectangle that bounded the apparent front portion of the object and then measuring the distance from the frame's center to the center of this rectangle. Directional consistency was coded as +1 if the direction from the object's center to its front was the same as that from the object's center to the frame's center (i.e., at the three positions plotted as "facing in" in Figure 5.6), -1 if these directions were inconsistent (i.e., at the three positions plotted as "out of the frame" in Figure 5.6), and 0 if they were neutral (at the center of the frame). The raw correlations of these three variables with the preference data for the 7 positions for each of the 6 objects (i.e., 42 observations) were: object center  $-.64$ , front center  $-.89$ , and directional consistency  $+.71$ . (The object center and frame center correlations are negative because the preference data increase near the center, whereas these distance measures decrease near the center.)

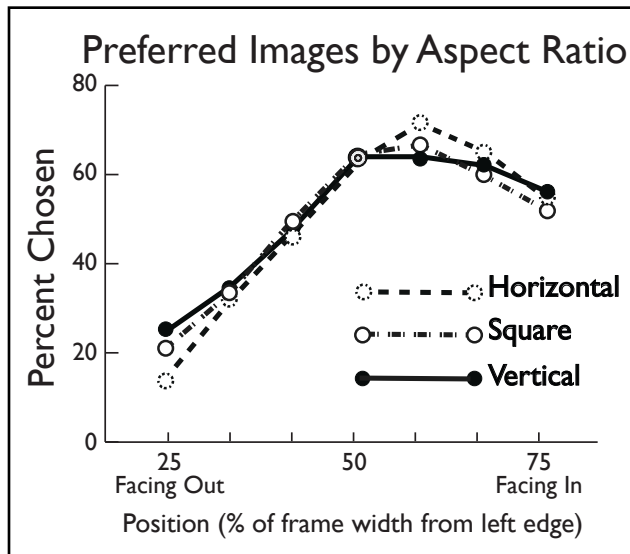


Figure 5.6. Results of Experiment 1.2 for different object aspect ratios. The average percentage of times the given image type was preferred is plotted separately for the three object aspect ratio conditions -- horizontal, vertical and square -- for the seven frame positions defined in terms of the object facing into versus out of the frame.

A stepwise regression in which the program determined the order of entering the predictor variables showed that front center was entered first (accounting for 79% of the variance;  $F(1,40) = 155.7, p < .001$ ), object center was entered next (accounting for an additional 4% of the variance;  $F(1,40) = 155.7, p < .001$ ), and directional consistency was entered last (accounting for an additional 9% of the variance;  $F(1,38) = 43.9, p < .001$ ). The final regression equation with these three independent variables had a multiple correlation of  $.96$  and accounted for 92% of the variance for 42 data points. That front center was the best predictor and was entered first in the equation lends support to the front position hypothesis. However, to determine its importance beyond object center and directional consistency, I also

performed a regression analysis in which object center and directional consistency were entered before front position. In this case, front center only accounted for an additional 1% of the variance, which did not quite reach statistical significance ( $F(1,38) = 3.94, p = .054$ ). The results are therefore ambiguous in the following sense: The single best predictor is front position, but the combination of the other two predictors (object center and directional consistency) fits the data well enough that adding front position does not significantly improve the fit.

The method that I used to determine the front position was entirely subjective. Moreover, because the regression program computed linear weighted combinations of the predictor variables, the object center and front center variables can actually be viewed as reflecting the influence of a single location whose distance from the frame's center provides the best fit to the data. Therefore, some modified estimate of the front's center accounts for 83% of the variance, with directional consistency accounting for an additional 9%, as reflected in the original stepwise analysis. It would be interesting to estimate this modified front location separately for each

object to see how much the fit could be improved and to see how closely it might correspond to the “perceptual center” of each object as determined by other methods. This question of “perceptual” versus physical centers is addressed in Experiment 1.5.

### Experiment 1.3: Position and Direction in a Constrained-Adjustment Task

Forced-choice methods provide exceptional precision in determining people’s preferences among the discrete alternatives chosen by the experimenter. Unfortunately, the combinatorial realities of fine-grained sampling are daunting, because the number of trials increases proportionally to the square of the number of sample values along the dimension(s) of interest. We therefore explored the more open-ended task of constrained adjustment in Experiment 1.3, in which participants used a computer mouse to drag the object along the horizontal midline until they found the most aesthetically pleasing position, at which point they clicked the mouse to record their preference. In other blocks of trials, we asked them to find the position that they found least aesthetically pleasing to provide data anchoring the other end of the aesthetic scale.

#### Method

Participants. All 9 of the participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. All were naïve to the purpose and nature of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol. Two participants were eliminated because they either did not understand or did not follow the instructions, as indicated by the fact that, unlike all other participants, their settings for the best positions were nearly identical to those for the worst positions.

Design. The experiment consisted of two blocks of 126 trials: one block in which participants placed the object in the *most* pleasing position, and one block in which they placed it in the *least* pleasing position. Within each block, each of 42 objects (the ten objects in each of three facings from Experiment 1.1, and the six objects in each of two facings from Experiment 1.2) were presented three times: once with the object’s starting position at the left edge, once in the center, and once at the right edge of the frame.

Displays. The objects in Experiments 1.1 and 1.2 were unchanged in Experiment 1.3 and were presented in the same frame with the same black ground plane and white background. The only differences were that there was only one frame in the center of each screen and that the participant could control the horizontal position of the object by moving the mouse laterally. The monitor and display settings were unchanged from Experiment 1.2.

Procedure. In one block of trials, participants were instructed to look at each frame as it was presented and to move the object horizontally (using the mouse) to the position where the object made the overall image “most pleasing.” In the other block, they were asked to position the object where it made the overall image “least pleasing.” Vertical displacements of the mouse were not considered, so the object’s center slid smoothly along the horizontal midline. The order of the blocks was counterbalanced so that half of the participants were instructed to place the object in the best position in the first block and half to place it in the worst position in the first block.

#### Results and Discussion

The position of the geometric center of the object was recorded for each trial as its

distance in pixels relative to the center of the frame, with positions to the left coded as negative and those to the right as positive. To enable rough comparison of the present data with those of previous experiments, the frame was divided into seven equal bins along the horizontal dimension, and the image from each trial was categorized according to the positional bin into which its center fell. The average percentage of trials on which the object center fell into each bin is plotted in Figure 5.7 for the “best position” instructions and in Figure 5.8 for the “worst position” instructions. In each case the data are shown for the leftward, rightward, and forward facing conditions.

Separate analyses were conducted on the data from the ten objects in Experiment 1.1 (Object Set 1, five of which implied motion and five of which were merely facing) and the six objects in Experiment 1.2 (Object Set 2, including two each at three aspect ratios), primarily because only the former were shown in forward (center) facing positions. The data from the left- and right-facing conditions for both object sets are combined in Figures 5.7 and 5.8 because there were no reliable differences between them,  $F(1,9) < 1$ .

The analysis of the best positions with Object Set 1 showed a strong main effect of facing conditions (left, center, and right),  $F(2, 18) = 136.90, p < .001$ , but no difference between objects that implied motion and those that were merely facing,  $F(1,9) < 1$ . The centers of the objects facing right were placed further to the left than the front-facing objects,  $F(1,9) = 90.86, p < .001$ , and those facing left were placed further to the right than the front-facing objects,  $F(1,9) = 85.90, p < .001$ . These results are in complete accord with those from 2AFC judgments in Experiments 1.1 except that the inward facing bias appears to be stronger in the present data, with more strongly lateralized maxima.

The best position results for Object Set 2 were similar to those for the left- and right-facing conditions of Object Set 1, with the centers of left-facing objects much farther to the right than those of right-facing objects,  $F(1,9) = 81.80, p < .001$ . These results are also in accord with those from Experiment 1.2, except that the inward facing bias was again stronger in the present

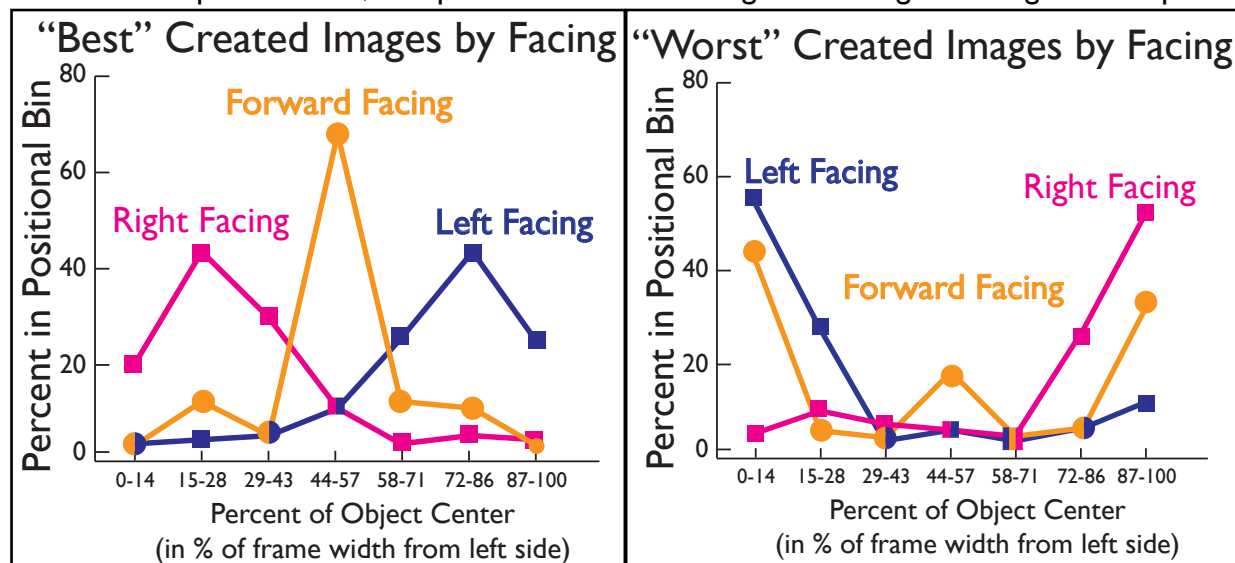


Figure 5.7. Results of Experiment 1.3 for the 'best' position. The percentage of trials in which the center of the object fell into each of seven positional bins when participants were asked to place it in the most aesthetically pleasing position for the center-facing, left-facing and right-facing images of the 16 objects used in Experiments 1 and 2.

Figure 5.8. Results of Experiment 1.3 for the 'worst' position. The percentage of trials in which the center of the object fell into each of seven positional bins when participants were asked to place it in the least aesthetically pleasing position for the center-facing, left-facing and right-facing images of the 16 objects used in Experiments 1 and 2.

data. Unlike Experiment 1.2, however, the results for objects with different aspect ratios did not reach significance in the present data ( $F(2,18) = 2.78, p = .088$ ). Indeed, they were not even ordered in the predicted direction, with the tall, thin objects producing the largest displacement (132 pixels) and the square objects and the short, wide objects producing smaller displacements (64 and 76 pixels, respectively). Given the lack of evidence for systematic effects of aspect ratio here and in the opposite-facing comparisons of Experiment 1.2, we have limited confidence in the aspect ratio effects observed in the same-facing conditions of Experiment 1.2.

The data from the worst positions are equally clear and compelling. The worst positions are clearly at or near the edges of the frame. This is the result of the center bias for the best positions to be at or near the center. The only exception is a small increment for the forward-facing objects at the central position, which may reflect a few participants' belief in an explicit "rule" that objects should not be placed at the exact center of a picture. For the forward-facing objects, the two edges are about equally bad, but for the left- and right-facing objects, there is a huge asymmetry: left-facing objects are least pleasing when they are on the left side of the frame and right-facing objects are least pleasing when they are on the right side of the frame. This fact is reflected in the average worst-positions of the objects in that participants located the right-facing objects farther to the right than the left-facing objects,  $F(1,9) = 121.28, p < .001$ . This pattern of results presumably arises from the joint operation of the inward and center biases, which together dictate that the worst compositional choice is for the object to face outward at the most extreme position.

The present data thus converge with the primary results of Experiments 1.1 and 1.2 in affirming the existence of powerful preferences for objects to be positioned toward the center of the frame and to face into the frame. It is unclear, however, why the constrained adjustment data for the best composition gave stronger evidence of the inward bias than the 2AFC data in the previous experiments. One possibility is that adjustment strategies in the present paradigm tended to magnify the inward bias effect. For example, if participants tended to move the object outward from the center in trying to find the best location and if hysteresis effects are present, they might move it farther outward than they would judge optimal in a 2AFC paradigm. Because there is no data on the trajectories of object positions, however, we cannot evaluate such hypotheses in the present data.

#### Experiment 1.4: Position and Direction in Free-Choice Photography

The results of the first three experiments clearly demonstrate the existence of the center and inward biases, but we wanted to see whether they would also be revealed under the more open-ended conditions of people taking actual photographs. Participants in the previous experiments might well have discerned the purpose of the studies from the nature of the displays they were shown or the nature of the adjustments they were allowed to make, and this might have influenced their choices, either consciously or unconsciously. In the present experiment participants were given a digital camera and simply asked to take the most aesthetically pleasing picture they could of a series of everyday objects. Under these conditions it seems unlikely that the participants would discern my underlying hypotheses.

Each participant was given a digital single-lens-reflex camera and asked to take the most aesthetically pleasing pictures they could of three everyday objects – a teapot, a tape dispenser, and a steam iron – in each of seven instructional conditions. The target object was positioned on a turntable so that participants could change its orientation, if they wished. The first instructional condition imposed no constraints at all: Participants were free to take whatever

picture they found most aesthetically pleasing. After doing this for each of the three objects, they were given the following series of six tasks that imposed specific constraints: the object must be located off-center and facing rightward (condition OCR) or facing leftward (OCL), the object must be partially out of the frame and facing rightward (OFR) or facing leftward (OFL), and the object must be entirely inside the frame and facing rightward (IFR) or facing leftward (IFL). Participants were told that “facing rightward” and “facing leftward” did not mean that the object necessarily had to be in full profile, but only that its front had to face right or left of directly toward (or away from) the camera. The image participants saw in the viewer was exactly the image that was recorded and analyzed.

The initial, unconstrained photographs were scored for both the central position and the direction of the object in the picture. We expected that there would be a bias toward placing the object at or near the center of the frame and that, if it were off-center, there would be a bias for the object to face into the frame. Because the rest of the conditions dictated the direction of the object, the sole dependent variable of interest was the location of the object’s center. If there is indeed a preference for objects to face into the frame, then right-facing objects should tend to be framed left of center and left-facing objects should tend to be framed right of center.

## Methods

Participants. All ten participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age was 19.9 years. All were naïve to the purpose and nature of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. The unconstrained (“best picture”) condition was completed first for all three objects, and the remaining six conditions were randomized. Object order, which remained consistent within participants for every condition, was also randomized across participants.

Materials. The three objects (a teapot, a tape dispenser, and a steam iron with its cord removed) were positioned on a white turntable 12” in diameter. The ground plane was a white, foam-core matboard and a second white matboard, perpendicular to the ground plane, stood behind the turntable against a wall.

Procedure. Each participant received a brief tutorial in using the Nikon D100 digital single-lens reflex camera. The camera was set for automatic exposure, so that participants only needed to position the camera and zoom the lens, which had an effective adjustable focal length of 42mm to 135mm. The participants could stand wherever they wanted to take the photographs, but were constrained by the size of the area in which the apparatus was located (approximately 5’ x 4’). The aspect ratio of the digital images taken by the camera was 3:2, and their size was 3008 x 2000 pixels.

For the three initial pictures, the experimenter placed the object on the turntable facing directly forward, and participants were then free to turn the object however they wanted. For subsequent pictures, participants were allowed to place the object on the turntable themselves, in accord with the facing instructions for that condition. The order of the left- and right-facing constraints was counterbalanced across participants.

## Results and Discussion

The image location of the center of the object and the object’s direction of facing were determined by eye for each digital photograph. In cases where the center of the object was

outside the frame, the position was coded as being at the edge of the frame closest to its center ( $\pm 1504$  pixels) rather than at its actual center, which was not visible in the photograph. Because this occurred frequently in the instructional condition in which participants were required to take a picture in which the object was partly out of the frame (i.e., the OFR and OFL conditions), the data from these conditions is not included below.

The initial three “best” photographs for each participant were the only ones in which they had free choice about how the object should face as well as where it was positioned. These images showed a strong rightward facing bias, with 80% of the objects facing right and 20% facing left ( $p < .001$  by a binomial test). It is noteworthy that this bias occurs despite the fact that a right-handed person would normally use both the iron and the teapot (but not the tape dispenser) in a left-facing position. This fact is inconsistent with any hypothesis that a rightward bias results from standard conditions of use or even the frequency with which the objects are seen in a right-facing orientation. There was also a strong inward bias with respect to the frame,

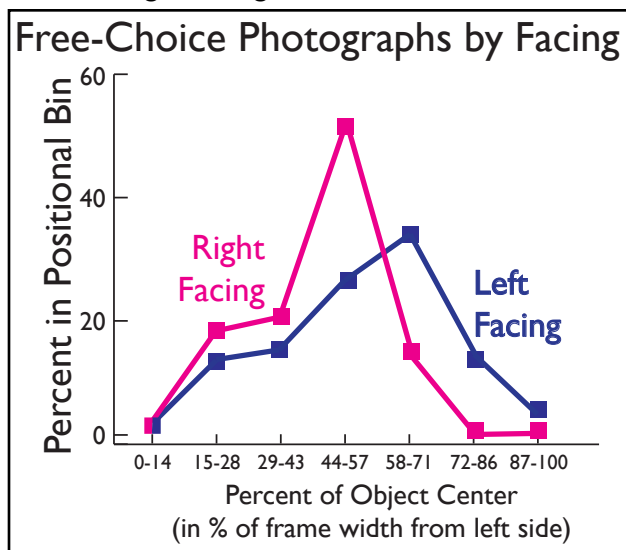


Figure 5.9. Results of Experiment 4. The percentage of photographs taken in which the center of the object fell into seven positional bins for the partly constrained left-facing and right-facing instructional conditions in a free-choice photography task.

with 77% of the images showing the object facing into the frame versus 23% showing it facing out of the frame ( $p < .001$  by a binomial test). The position of the object in the “best” photo condition was strongly biased toward the center, with a mean position approximately 35 pixels offset to the left of center, probably as a result of the right-facing bias, which was relatively strong in these data.

The rest of the photographs, which were constrained to be facing either rightward or leftward according to explicit experimental instructions, were only analyzed for the position of the object’s center. Figure 5.9 shows the resulting percentages of photographs in which the objects’ centers fell within each of the seven equally spaced positional bins, with separate functions for right-facing and left-facing instructional

conditions. The results of a two-way analysis of variance on the positions of the centers of the objects indicate a main effect of facing direction ( $F(1,9) = 6.99, p < .03$ ), but not of instructional condition,  $F(1,9) < 1$ , or their interaction ( $F(1,9) = 1.92, p > .20$ ). Although these data are not as orderly as those from the 2AFC and constrained adjustment tasks, the asymmetric signature of the inward facing bias was still clearly evident: People placed the right-facing objects toward the left side of the frame and the left-facing objects toward the right side of the frame.

### Experiment 1.5: Position and Direction in Stock Photography

The results of the previous four studies provide convincing evidence that the inward bias and the center bias are solid and robust phenomena when studied in the laboratory. A serious issue that needs to be addressed, however, is whether these biases show up in aesthetically pleasing images that are created by “real” photographers. Natural photographic images typically contain various elements -- other objects and surfaces in the background around the focal object -- that were not present in any of the images in Experiments 1.1-1.4. The presence of



such additional content could affect the nature and/or strength of the biases reported above. Additionally, as mentioned in the discussion of Experiment 1.2, it is not yet clear whether there was a difference between the physical center of an object, and its perceptual center.

Both of these issues were addressed by studying to stock photography. Stock photographs are shot by professional photographers, and their content is typically of everyday places, people, objects, and events. These photographs are then licensed by a company who sells them to individuals for use in advertisements, films, newspapers, books, and magazines. The original purpose of stock photographs was to provide available images to commercial sources who did not have photographers on staff. Because the photographs are taken by professional photographers, they are technically well-executed (with regard to focus, color balance, and other factors), and because they are intended for general consumption, they are (at least intended to be) aesthetically pleasing, though they seldom achieve the level of fine art.

Previous research by Martin, Fowlkes, Tal, and Malik utilized the Corel Image Database of Stock Photography to create “ground-truth” image segmentations of 800 photographs (Martin, et al., 2001). The original use of these data was to create a computerized segmentation algorithm that would approximate human performance and perception. The same data, however, provided a rich dataset of object positions, from which the presence or absence of the inward bias and the center bias could be assessed.

Three hundred photographs selected from the original 800 used by Martin and colleagues, were presented to participants, who made a judgment for each photograph about the number of “focal objects” it contained. Focal objects were defined as objects that the image was “about” or “without which, the photograph would not be worth taking”. After all photographs had been judged, each photograph that participants chose as only having one focal object was presented a second time, and participants were instructed to click on “the center of the visible portion of the focal object.” The horizontal position and facing direction of the focal object in the single-object photographs were then analyzed with respect to both their objective centers and the subjective centers provided by the participants’ designations.

## Method

**Participants.** The nineteen participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age was 20.3 years. All were naïve to the purpose and nature of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

**Design.** There were 300 trials in the first part of the experiment in which participants indicated the number of focal objects they saw in each picture: one, two, or more than two. In the second part of the experiment, the number of trials for each participant depended on how many photographs from the first portion were judged by that participant to have a single focal object. The order of presentation of the photographs in both the first portion and the second portion was randomized for each participant.

**Displays.** Each landscape-oriented photograph measured 480x320 pixels, and each portrait-oriented photograph measured 320x480 pixels. They were centered on a 19” monitor with a resolution of 1024x768 pixels.

**Procedure.** For the first part of the experiment, participants made a button response to designate whether there was one, two, or more than two focal objects present in the photograph presented. The photographs remained on the screen until a response was given. In

the second part of the experiment, participants used the mouse to click on the center of the visible portion of the focal object in each of the one-object photographs from the first part. As in the first part, the photographs remained on the screen until a response was given.

### Results and Discussion

The first analysis determined which photographs had relatively high agreement among participants about the number of focal objects. A photograph was included in further analyses if at least 15 people (of 19, or 79%) were in agreement about it having a single focal object. This resulted in 87 photographs being included in the analysis of object centers (See Figure 5.10). Next, the one-object images were coded for their facing direction. Two research assistants coded each image for whether the focal object (unambiguous in all 87 photographs) faced

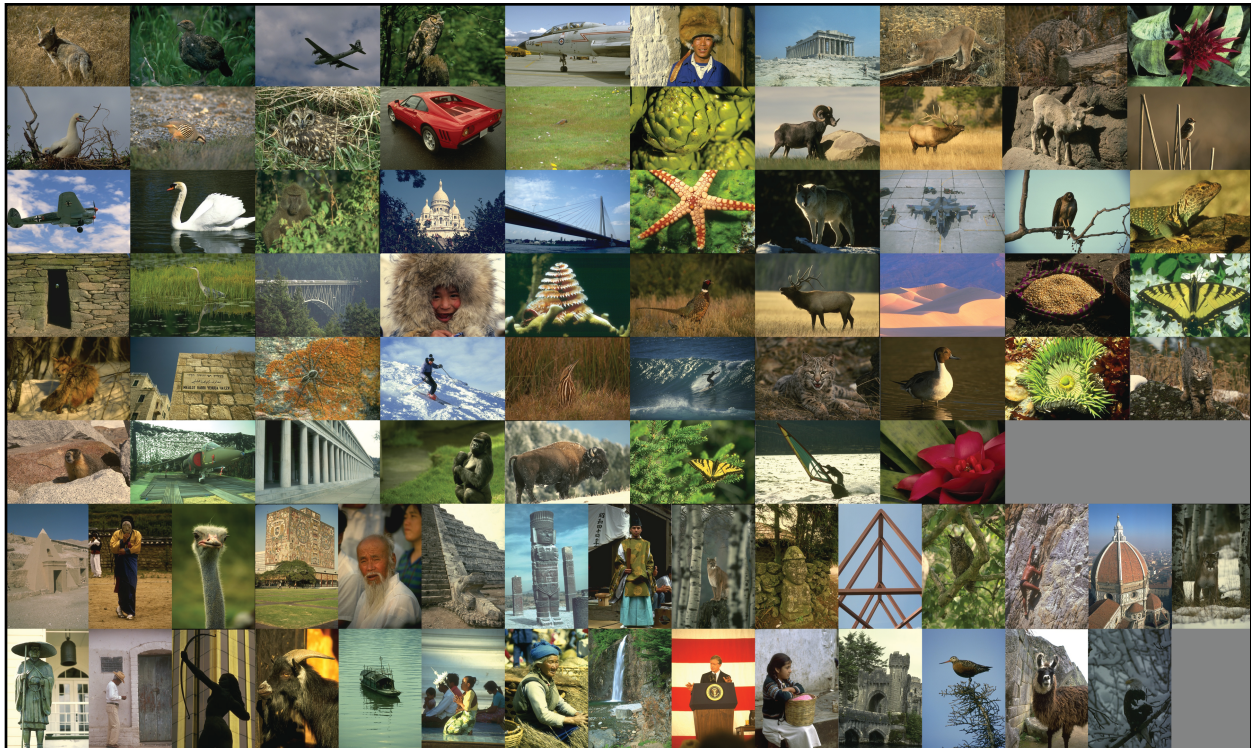


Figure 5.10. The 87 images judged to contain a single focal object by participants in the first part of Experiment 1.5. These images were then presented to the same participants who were instructed to click on the center of the visible portion of the focal object in each photograph.

leftward (facing 45° or more to the left), rightward (facing 45° or more to the right), or forward (between the other two conditions), providing data about which they both were in agreement.

To address the question of whether there was a difference between the perceptual (or subjectively judged) centers and the physical (or objectively derived) centers of objects, two measures of the physical center for each photograph was computed. Both measures were based on the parsed segments (image regions) that made up each focal object (see Figure 5.11B). In the bounding-box method, the smallest rectangle that included all object-pixels visible within the frame was determined, and its center was taken to be the center of the visible portion of the object (see Figure 5.11C). The geometrical method, the average x- and y-values of all the pixels included in the segmentations was computed and taken to be the center of the visible portion of the object (See Figure 5.11D). The histograms of the centers for the two objective methods were not significantly different using a two-tailed Kolmogorov-Smirnov test. Thus, only data from

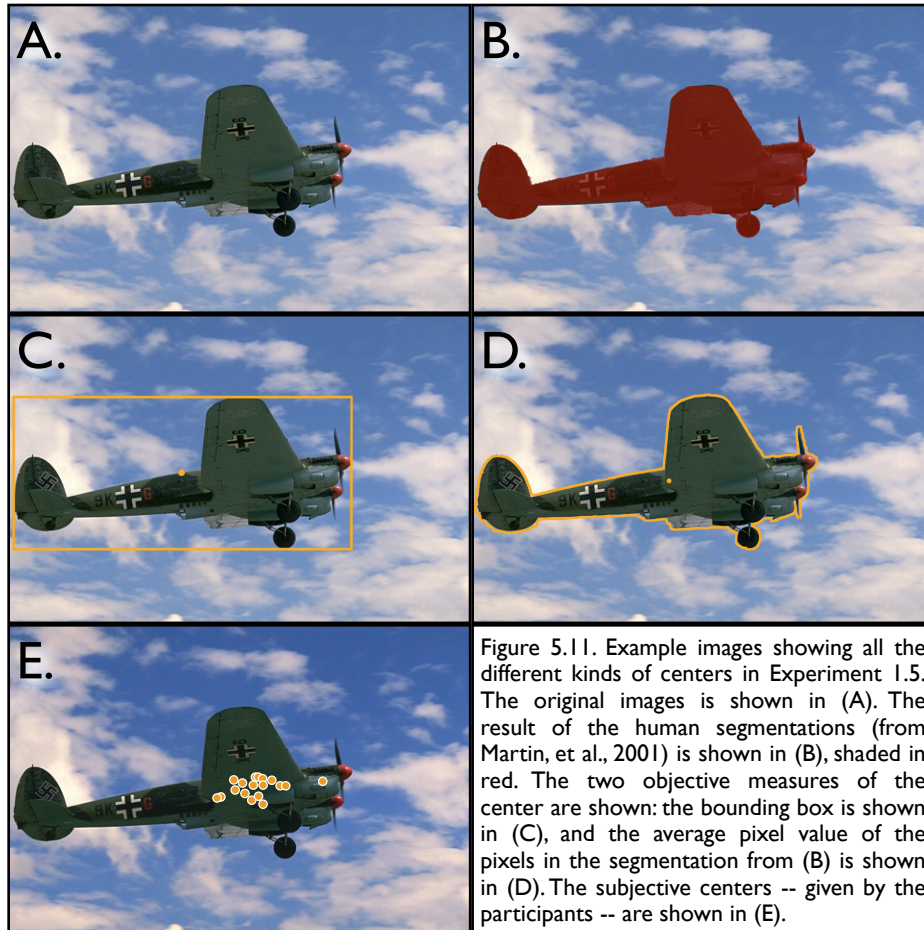


Figure 5.11. Example images showing all the different kinds of centers in Experiment 1.5. The original images is shown in (A). The result of the human segmentations (from Martin, et al., 2001) is shown in (B), shaded in red. The two objective measures of the center are shown: the bounding box is shown in (C), and the average pixel value of the pixels in the segmentation from (B) is shown in (D). The subjective centers -- given by the participants -- are shown in (E).

the bounding box method is used for the “objective center” data in subsequent analyses and graphs.

Nine equally spaced positional bins were created and the number of photographs whose focal objects’ centers fell into each bin was computed. For the objective centers, 76% of the forward facing objects were located in the center bin, significantly falling off further from the center (See Figure 5.12A;  $p < .001$  by a binomial test). This is a more pronounced center bias than had been found before. When the depicted objects were facing

leftward or rightward, the majority (48.5% and 61.8%, respectively) of the objective centers were placed in the off-center bin one position removed from the center in the direction opposite the facing direction, thus replicating the inward bias. The distributions for leftward- and rightward-facing objects were statistically significantly different from each other as measured by a two-tailed Kolmogorov-Smirnov test, as well as from the forward-facing distribution ( $p < .05$  in all cases).

The subjective centers presented primarily a center bias. When the depicted objects were facing forward, 64.5% of the subjective centers were located in the center bin, for both kinds of centers (See Figure 5.12A, 5.12B). The three distributions of subjective centers, however -- for leftward, rightward, and forward facing objects -- were not significantly different from each other (for leftward vs. rightward, rightward vs. forward, and forward vs. leftward,  $p > .4$  in all cases by two-tailed Kolmogorov-Smirnov tests). This demonstrates that there was no significant inward bias with respect to the subjective centers of the focal objects: more centers of rightward facing objects (49%), leftward facing objects (44.5%), and forward facing objects (64%) were placed in the center bin than in any other bin (see Figure 5.12B), and the percentages fell sharply in both directions further from the center. This demonstrates an interesting phenomenon: when an object was physically offset in the direction opposite its facing, viewers perceived the object as being centered. This could occur because the frame itself biases the perception of the object centers toward the frame’s center. This hypothesis could be tested by re-cropping the images in which the entire focal object was visible so that the position of its objective center varies

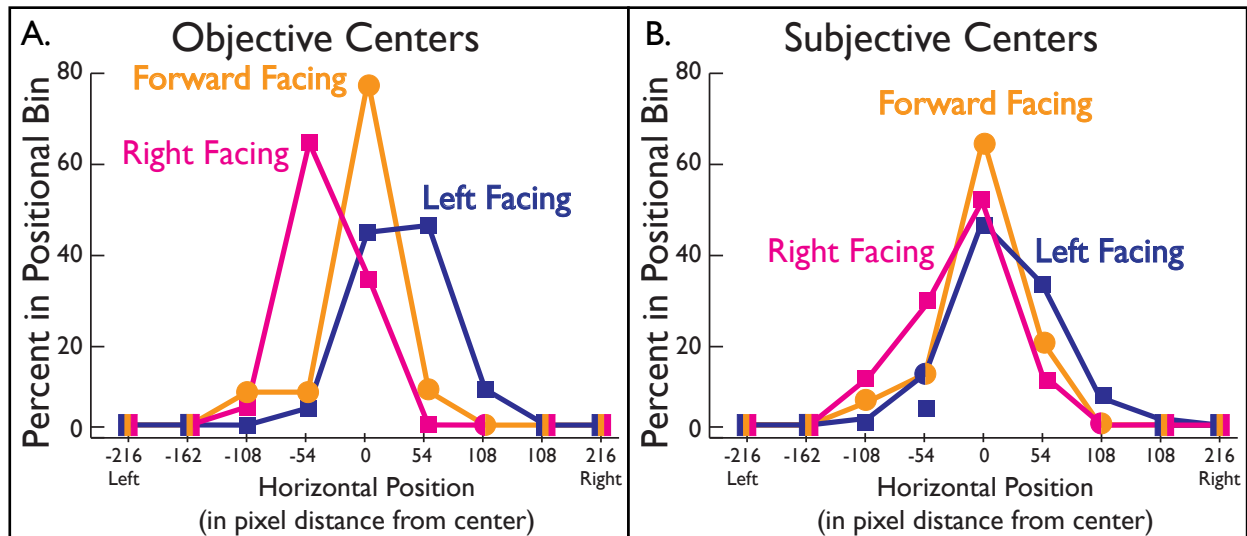


Figure 5.12. Results of Experiment 1.5. Objective centers -- measured by the bounding box -- are shown in (A). Note that there is a strong inward bias for the leftward and rightward facing images. Subjective centers -- measured by participants' judgments -- are shown in (B). Note the lack of an inward bias.

systematically relative to the frame's center and then asking different groups of viewers to indicate the position of the center of each focal object. If the frame exerts an influence on the perceived (subjective) center of the object, then the object-center judgments should vary in the direction of the frame-center. This particular follow-up experiment remains for future research.

### Conclusion for Experiments on Horizontal Position

Five experiments were conducted that investigated people's aesthetic preferences for the framing of simple, single-object pictures in rectangular frames. The results were consistent in revealing two powerful biases in aesthetic preference, one for objects to be positioned toward the center of the frame (the center bias) and the other for objects to face into the frame (the inward bias). Both produced consistent, statistically robust effects in every relevant comparison. A weaker third bias was sometimes evident for participants to prefer objects facing to the right (the rightward bias), but it was not as consistently observed as the other two principles, being present in Experiments 1.1 and 1.4 (and implied by the asymmetry in the histogram in Experiment 1.5), but not in Experiments 1.2 or 1.3.

Experiment 1.2 provided further evidence about why people prefer objects to face into the frame. By sampling object positions more densely and by varying the aspect ratio of the objects, we found evidence for both the directional consistency hypothesis (that people prefer the object to face in the same direction as the direction from the object's center to the frame's center) and the centered front hypothesis (that people prefer the object's salient front region to be as close to the center as possible). The lack of corresponding aspect ratio effects in Experiment 1.3 casts doubt on the generality of the centered front hypothesis, however. Even so, the overall results strongly support the paramount importance of the frame's center in aesthetic considerations for these simple compositional issues (cf. Alexander, 2002; Arnheim, 1988): The most important variables can be specified in terms of the frame's center, the object's center, and the relation between the two.

The relation between the rule of thirds and these results is complex enough to merit a final summary. First, the rule of thirds appears to conflict rather dramatically with people's aesthetic preferences for the position of single symmetrical, forward facing objects, which people strongly

avored to be located at or very close to the center of the frame. For left and right facing objects, it is clear that a more lateralized position is preferred, but only on the side for which the object faces into the frame. Even here the precise location of the preference peak corresponds only roughly to the one-third or two-thirds lines, being closer to the center in Experiment 1.2 and more peripheral in Experiment 1.3. Composing the frame so that the center of the object conforms to the rule of thirds was aesthetically displeasing to the overwhelming majority of the observers when the object faced outward. If equal weight is given to these three cases --forward-, inward- and outward-facing objects -- then the results can be summarized as being consistent with the rule of thirds, at best, one-third of the time.<sup>19</sup>

Clearly, people's aesthetic responses to spatial composition cannot be based solely on the kinds of biases reported here. Indeed, images that deliberately violate standard expectations, such as the center and inward biases that are documented here, can produce quite positive aesthetic responses, particularly if the violation is integral to conveying an intended message or mood. A picture of a solitary person on a deserted beach, for example, might be judged more aesthetically pleasing if the subject were positioned decidedly off center and facing out of the frame, violating these conventions to convey feelings of isolation, loneliness, and/or longing. This pleasingness of violations of these biases -- and the contexts in which they can occur -- will be more thoroughly addressed in Chapter 7.

The first four experiments in this dissertation, people's aesthetic preferences for spatial composition have some marked commonalities. People have a strong bias to prefer images that contain objects that are either in the center of the frame, if they are facing forward, or near the center of the frame, but offset in the direction opposite facing, if they are facing leftward or rightward. The center bias and inward bias, as we have termed these findings, demonstrate that aesthetic preferences can be clear and orderly if sufficiently robust and studied in a rigorous way that allows them to be manifest. Having explored the horizontal dimension of object placement in this portion, the next set of experiments addresses the vertical dimension and our attempts to map out the preference space in a similar fashion.

---

<sup>19</sup> Two notes must be made for completeness' sake. First, the rule of thirds, as usually stated, does not include the points we studied along the horizontally oriented midline of the frame. We have supposed that it would apply to such cases as well. Further tests of the rule would therefore be desirable using the four points it explicitly prescribes. Second, to refer back to the Canon camera manual that I described in Chapter 3, they do approximate the inward bias later on, though only in reference to video, and without mention of the rule of thirds: "Try to give your video 'leading' room. If someone is running left to right, keep them in the left side of the screen. It will feel like there is room for them...If your subject is looking to the left...[and] you place the subject all the way to the left, there is nowhere for your viewer's eyes to go..." (Canon Digital Learning Center, 2010).

# Chapter 6

## Experiments on Vertical Position

The first set of experiments showed that people systematically prefer a vertically symmetrical, forward facing object to be centered in the frame (a “center bias”), and an asymmetrical left- or right-facing object to face into, rather than out of, the frame (an “inward bias”). The second set of experiments take the obvious next step of exploring preference for spatial composition in the vertical dimension. A center bias in the vertical dimension seemed likely because few objects seem to “face” upward or downward in the same way that they face leftward or rightward. It turned out, however, that there is indeed an inward bias in the vertical dimension, but the issues are much more complicated. Along the way, some interesting additional biases based on the ecological statistics of object positions in the world became evident that are not apparent in the horizontal dimension. These complexities arise primarily from the fundamentally anisotropic nature of the gravitational field and its pervasive effect on the vertical location of objects in the environment.

The logical progression of the research is as follows: Experiment 2.1 showed a strong lower bias in people’s preference for a bowl resting on a horizontal surface to be positioned lower in the frame. Various hypotheses about why this lower bias might exist were formulated and tested in subsequent experiments. Experiment 2.2 replicated the lower bias for the bowl, but found an equally strong upper bias for a light fixture attached to a ceiling. Experiments 2.3 and 2.4 further explored these effects to find out whether they were affected by viewing direction (perspective), spatial asymmetries in the functional properties of objects (functional asymmetries), ecological statistics about the location of objects relative to human viewers (ecological statistics), or some combination of these factors. Biases arising from both functional asymmetries and ecological statistics were clearly evident. The results are discussed in terms of people preferring images that have a spatial composition that reflects the most important spatial information about the object, including their affordances relative to the observer and the ecological statistics of its likely vertical position relative to human observers.

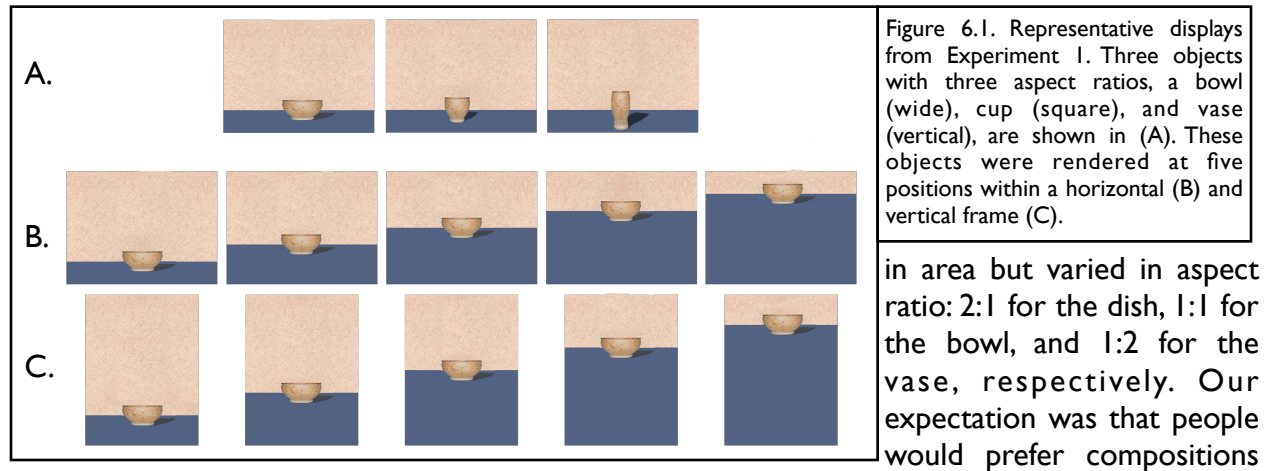
### Experiment 2.1: Objects Supported by a Horizontal Plane Below Eye Level

Following the logic of the previous work on horizontal placement of objects, a two-alternative forced-choice (2AFC) design was used to get accurate measures of people’s preferences for different vertical placements of an object within a frame. On each trial, two images were presented that differed only in the vertical position of the identical object and that of the horizontal edge of its supporting ground plane. As before, participants were simply asked to select the picture that they found more aesthetically pleasing, a method that has the advantage of allowing participants to make a direct comparison of the two competing images without the confounding effects of memory, anchoring, or criterion shifts. Participants made 2AFC preference judgments for all possible pairs of images of the same perspective view of the same object. The average probability of choosing each image versus all others as an estimate of people’s overall preference for it was then computed.

We decided to begin with an almost-pure-side view of a ceramic bowl resting on a table-like horizontal surface as a canonical single-object scene in which the vertical position of the object might vary. Such a scene is, in effect, a two-object scene, because explicitly or implicitly, the vertical position of the horizontal edge of the ground surface must be considered as well as that of the focal object itself. For our initial exploratory study, we decided to simplify the image set

by always placing the horizontal edge of the supporting surface at the same level as the center of the object. This pair of image features was positioned at five equally spaced heights in the picture, with the third position being at the exact center of the frame.

In the prior experiments, the horizontal position of objects had been varied while holding constant their vertical position at the center. In this experiment we did the reverse. We also varied the aspect ratio of both the frame and the object, creating a wider (dish), neutral (bowl), and taller (vase) version of the same object (see Figure 6.1A) within a horizontal or vertical frame (See Figure 6.1B, 6.1C). The frames were equal in area but varied in aspect ratio: 4:3 for the horizontal frame and 3:4 for the vertical frame. The objects were also approximately equal



in area but varied in aspect ratio: 2:1 for the dish, 1:1 for the bowl, and 1:2 for the vase, respectively. Our expectation was that people would prefer compositions in which the aspect ratio of the frame was most similar to the aspect ratio of the object, but with the caveat that there might be an overall bias toward the horizontal frame shape due to familiarity. Each object shape was compared to the same object (aspect ratio) in both frame shapes, to detect possible interactions in preference between the aspect ratios of the object and frame.

### Methods

**Participants.** All 15 participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age was 19.3 years. All were naïve to the purpose of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

**Design.** The experiment consisted of 420 trials, resulting from the pairwise comparisons of the same object (dish, bowl, or vase) located at each of 5 vertical positions within both of the frame shapes, balanced for screen position (i.e., the permutations of ten items – one of the objects at the five positions in the two frame shapes – chosen two at a time, for each of the three objects). The 5 positions were located at 20%, 35%, 50%, 65% and 80% heights in each respective frame, such that the horizontal edge of the ground plane and the object center were at that height (Figure 6.1B, 6.1C).

**Displays.** The monitor measured 19" diagonally, and had a resolution of 600 x 800 pixels. Each screen consisted of a pair of colored images on a gray background. The two different frame shapes (360 x 270 pixels for the horizontal frame and 270 x 360 pixels for the vertical frame) always had the same center points on the screen: 200 pixels to the right or left of the center, and 150 pixels below or above the center such that one image was located in the upper left quadrant and the other in the lower right quadrant. The object was a ceramic bowl, which was

photographed from a side-view perspective and then imported into Adobe Photoshop CS2. The object was scaled into its different aspect ratios (100 x 50 pixels for the bowl, 75 x 75 for the cup, and 50 x 100 for the vase), with the texture on the bowl being reapplied at the original aspect ratio such that any resulting textural irregularities on the surface of the bowl were corrected (see Figure 6.1A).

**Procedures.** Participants viewed the computer screen from approximately 60 cm. They were instructed to look at each screen and to press either the left or right mouse button to indicate which image they “liked more.” They proceeded through the trials at their own pace.

### Results and Discussion

Participants’ responses were scored for the percentage of trial on which they chose each given image in all of its 2AFC comparisons. The resulting percentages are plotted in Figure 6.2, averaged over object shapes. There was a main effect of position ( $F(4,44) = 59.02, p < .0001$ ), which had statistically significant linear ( $F(1,11) = 90.35, p < .001$ ) and quadratic components ( $F(1,11) = 9.55, p < .01$ ). The most preferred position was that second from the bottom, with the bottom-most position being chosen almost as often. These two positions were reliably preferred to the top two positions ( $F(1,11) = 85.45, p < .001$ ). The slight preference for horizontal over vertical frames was not significant ( $F(1,11) = 1.108, p = .32$ ), but there was a reliable interaction between frame shape and position  $F(4,44) = 4.04, p < .01$ , which is small, but evident in Figure 6.2. There was no main effect of object shape ( $F(2,22) = 0.38, p > .50$ ) nor was there any interaction between frame shape and object shape ( $F(4,44) = 2.20, p > .13$ ).

The strong lower bias in the results was something of a surprise, because we expected to find primarily a center bias analogous to my results with the horizontal placement of symmetrical objects. That is, we thought that people would prefer the bowl to be at or near the center of the frame, but instead, the bottom two positions were most preferred. Several interpretations of this finding are possible. The *gravitational hypothesis* is that the lower bias might

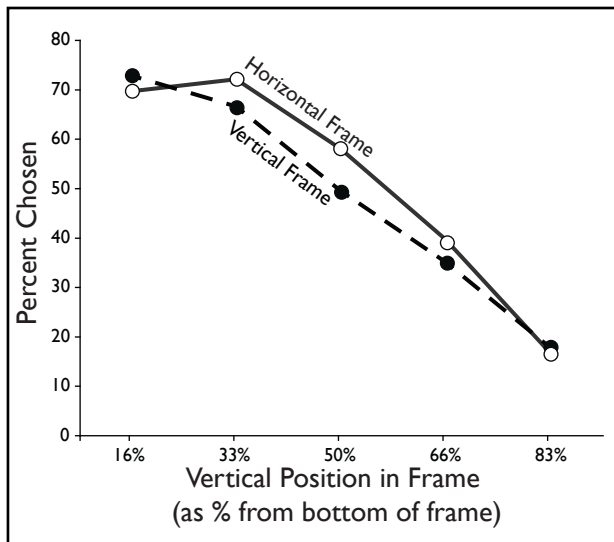


Figure 6.2. Results of Experiment 2.1. The average percentage of trials in which the given image was preferred over all possible comparisons is plotted as a function of the position of the object in the frame for the horizontal frame (solid line, open circles) and the vertical frame (dashed line, filled circles).

reflect a preference for the vertical position of the object and ground plane’s far edge to be close to the bottom of the frame, where they would lend greater overall gravitational stability to the image (cf. Arnheim, 1974, 1988). The *affordance asymmetry hypothesis* is that the present results are actually just the lower half of an inward bias in the vertical dimension, analogous to that the inward bias in the horizontal dimension, which we interpreted as being driven by asymmetries in their functional (affordance) structure. The idea is that because the top of the bowl is functionally its most salient part – playing the crucial role in how substances are placed into and taken out of it – the top of the bowl is analogous to the fronts of the objects we studied previously in the horizontal dimension (a person, a dog, a flower, etc.). If people prefer objects to “face into the frame” in the vertical dimension because of



such functional asymmetries, then they should prefer the bowl to “face” upward into the frame, such that its top is closer to the frame’s center than its bottom and there is more room above than below it. Even so, the lower bias found in the present experiment is a good deal stronger than the inward biases found with horizontal position, suggesting that some additional factor may be at work. Finally, the lower bias found here might be explained by an *ecological hypothesis* derived from environmental statistics about where bowls are usually located relative to human observers (and independent of the perspective from which it is seen). Note that these three hypotheses are not mutually exclusive: Indeed, all might be involved to some degree in producing the strong lower bias found for the bowl in Experiment 2.1. We address these and related issues in Experiments 2.2 and 2.3.

## Experiment 2.2: Objects Supported by Horizontal Planes Above versus Below Eye Level

We explored the three hypotheses just described of the strong lower bias found in Experiment 2.1 – gravitational stability, affordance asymmetries, and/or ecological statistics – through two manipulations in Experiment 2.2. First, a hemispherical light fixture was added, whose shape was similar to that of the bowl, but was supported by attachment to an overhead (ceiling) plane. If gravitational stability is the crucial factor, people should prefer the light fixture to be at or near the bottom of the frame, just like the bowl. Both of the other two hypotheses predict a reversal of the vertical bias for the light fixture, toward preferring it to be placed higher in the frame. Because the functionally salient part of the light fixture is the lower, light-emitting portion, the presence of a bias governed by affordance asymmetry, analogous to that in the horizontal dimension, predicts that the fixture should be preferred at or near the top of the frame (i.e., an upper bias for an object that “faces” downward). And because light fixtures are usually located above human observers, the ecological hypothesis also predicts that it should be preferred toward the top of the frame. The latter two hypotheses cannot be distinguished in the present experiment, but an attempt to disentangle them is given in Experiments 2.3 and 2.4.

The second manipulation included in the current experiment was decoupling the object’s vertical position from that of the horizontal edge. Within the constraints imposed by gravitational support of the bowl by an underlying surface and of the light fixture by an overhead surface, the vertical position of the horizontal edge was varied independently of the object’s position. By decoupling these two variables, we will be able to determine the extent to which the results of Experiment 2.1 might have been due to the height of the horizontal edge coinciding with the height of the bowl.

The bowl and light fixture were photographed, isolated, and placed at different vertical locations in a frame, all of which were consistent with attachment to a horizontal plane below or above them, respectively. There were five possible positions of the object and five possible positions of the horizontal edge. The five images of the bowl were similar to those in Experiment 2.1, except that the perspective was now slightly above it and the horizon also varied in vertical placement (Figure 6.3A, 6.3B). The five images of the light fixture were constructed in a manner analogous to the images of the bowl in every respect except that the relative vertical positions of the fixture and the horizontal edge were reversed.

### Method

**Participants.** All fifteen participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age

was approximately 19.7 years. All were naïve to the purpose of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. There were 210 paired comparisons resulting from all ordered pairs of the fifteen possible images of an object (either the bowl or the light fixture) on each plane, taken two at a time. The colors of the plane of attachment and that of the back plane were counterbalanced within subjects, resulting in a total of 420 trials. Presentation of the two objects was blocked, and the order of the blocks was balanced across participants.

Displays. The bowl images were created by photographing the same ceramic bowl as in Experiment 2.1 from approximately 18° above the horizontal (to make different placements of the horizontal edge plausible), and then isolating it in Photoshop and placing it in a frame with a separately created background consisting of a single horizon edge. The different positions were generated by shifting the bowl and the horizon line vertically in the frame. The light fixture images were created in the same way, except that it was photographed from approximately 18° below the horizontal. The five object and horizon locations were defined by equally spaced vertical intervals across the frame, with one position in the exact center of the frame, two positions above the center, and two positions below the center (See Figure 6.3A, 6.3B). The center of each object was defined as the center of the bounding box around the object.

Each screen consisted of two color images, measuring 480 x 360 pixels, placed on a background of neutral gray. The images were placed at the same 384-pixel height on the screen, and were separated by 100 pixels. The images were created in Adobe Photoshop CS3 (see above), and the trials were created, randomized, and displayed using Presentation ([www.neurobs.com](http://www.neurobs.com)). The CRT monitor measured 18" diagonally, had a resolution of 1024 x 768 pixels, and a refresh rate of 85 Hz.

Procedure. Participants viewed the computer screen from approximately 60 cm. They were instructed to look at each screen and to press either the left or right mouse button to indicate which image they preferred. They proceeded through the trials at their own pace.

## Results and Discussion

From the participants' left/right responses the percentage of trials on which they chose each given image in all of its 2AFC comparisons was computed. Because neither the order of object presentation (i.e., block order) nor the color of the attachment plane had any effect (all  $F_s < 1.5$ ,  $p_s > .10$ ), subsequent analyses were performed on the data averaged over these factors. The preference data are plotted separately for the bowl (Figure 6.3C) and the light fixture (Figure 6.3D) as a function of the height of the object relative to the frame (x-axis), the height of the horizontal edge relative to the frame (dashed curves), and the height of the object relative to the horizontal edge (solid curves).

The most obvious and important feature of these data is that the light fixture data are an almost perfect mirror image of the bowl data, with a remarkable correlation of  $+0.98$  ( $p < .001$ ) between corresponding points of the two data sets. This means that whatever is true of the bowl data is essentially true of the light fixture data, but with the height relations reversed. For example, the lower bias in the bowl data, with the bowl at 16% and 33% from the bottom of the frame being chosen significantly more often than the 66% and 83% positions ( $F(1,14) = 16.18$ ,  $p > .001$ ), is an upper bias in the fixture data, with the 66% and 83% positions being chosen significantly more often than the 16% and 33% positions ( $F(1,14) = 13.15$ ,  $p < .01$ ).

If one considers only the data in which the bowl and horizontal edge were at the same

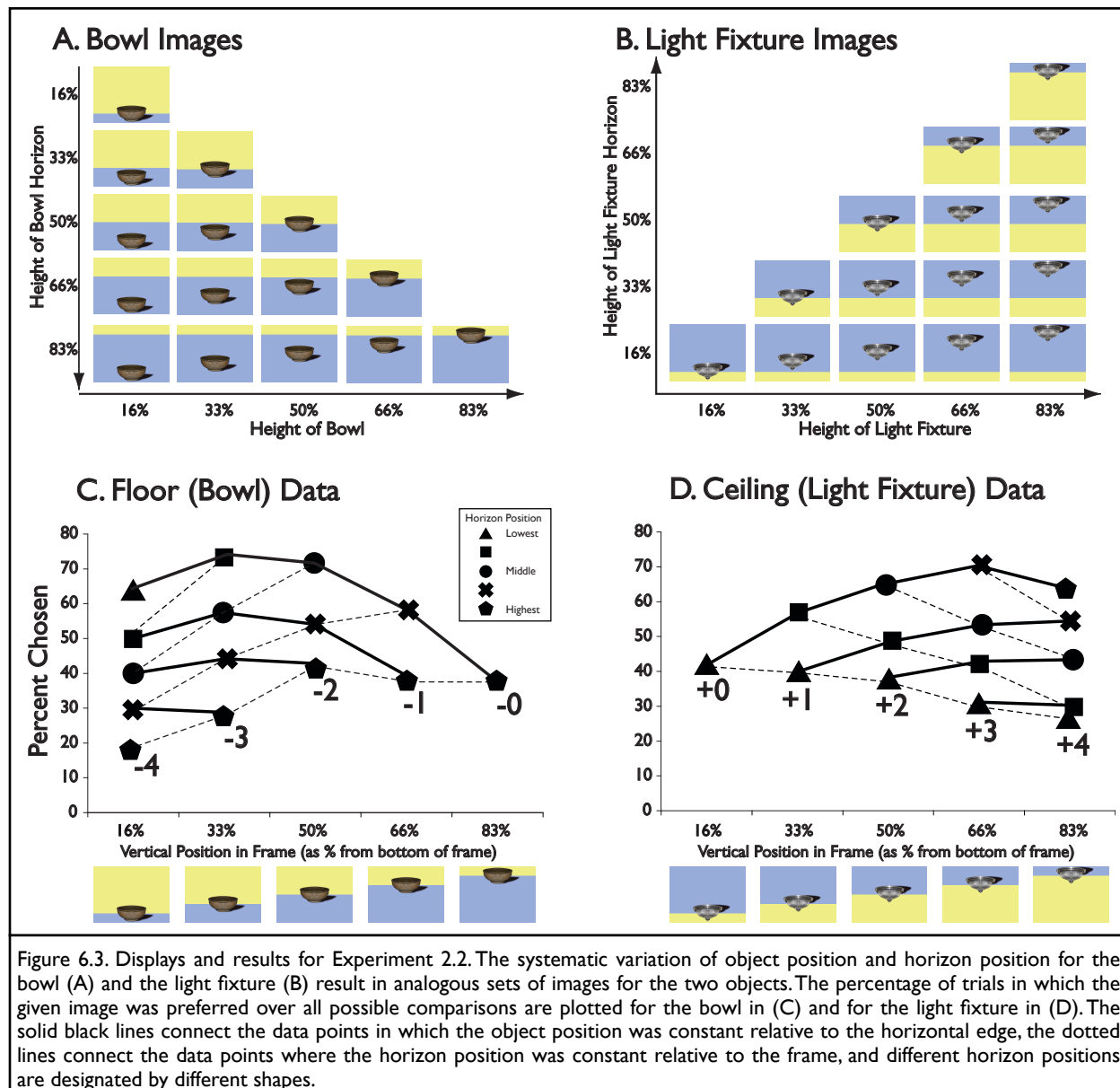


Figure 6.3. Displays and results for Experiment 2.2. The systematic variation of object position and horizon position for the bowl (A) and the light fixture (B) result in analogous sets of images for the two objects. The percentage of trials in which the given image was preferred over all possible comparisons are plotted for the bowl in (C) and for the light fixture in (D). The solid black lines connect the data points in which the object position was constant relative to the horizontal edge, the dotted lines connect the data points where the horizon position was constant relative to the frame, and different horizon positions are designated by different shapes.

vertical level (the top curve, labeled “0,” in Figure 6.3C), the lower bias for the bowl in Experiment 2.2 replicates the lower bias in Experiment 2.1 (Figure 6.2) almost exactly for the horizontally oriented frame ( $r = +.92$ ). Moreover, if one considers the data from both the bowl and light fixture when they are at the same level as the horizontal edge (the top curves, labeled “0,” in both Figures 6.3C and 6.3D), the pattern looks quite similar to the results for the horizontal positions of objects reported in Experiment 1.2 (Figure 5.4): There is an overall center bias that is strongly influenced by an inward bias toward objects “facing” into the frame. In the present case, we presume that the bowl “faces” upward and the light fixture “faces” downward because of the asymmetry of their primary affordances (cf., Gibson, 1977). In analogy to the horizontal case, an affordance bias predicts that people will prefer the location of the more functionally salient parts to be closer to the center of the frame, and indeed they do. The present data are therefore compatible with an affordance bias in the vertical dimension. They are also compatible with an ecological height bias, however, because the bowl is preferred lower

in the frame and is generally lower than a person's viewpoint when looking at it, and the light fixture is preferred higher in the frame and is generally higher than a human's viewpoint when looking at it. Of course, both affordance and ecological biases may contribute to the observed effects. This issue will be revisited in Experiment 2.4.

The observed pattern of near-perfect height reversal effectively rules out the gravitational hypothesis. The presumption behind it is that people preferred the bowl at its lowest position in Experiment 2.1 because that composition put the object and horizontal edge close to the bottom of the frame where the picture would be perceived as most gravitationally stable. If that were the relevant variable, people would also prefer the light fixture and horizontal edge to be close to the bottom of the frame. In fact, the fixture is preferred close to the top of the frame, where it appears least gravitationally stable. Given the near-perfect symmetry of the data, we conclude that gravitational stability is not a variable that affects people's preferences in these images strongly enough to be detectable.

Next we examined effects due to the relation between the height of the horizontal edge and the height of the object. The clear pattern that emerges from the data is that the most preferred image of both the bowl and the light fixture at each object height is the one in which the horizontal edge is at the same height as the object ( $F(1,14) = 8.36, p < .05$ , for images with the object at the horizon height versus the next closest horizon height). As the height difference between the object and horizon increases, the percentages of choices decrease regularly. Indeed, the decrease in the percentage of trials on which people chose each image (other than the one where they are at the same height) is highly correlated with the height difference between the object and the horizontal edge ( $r = +.96, p < .001$ ). There is thus a robust "same-height" bias toward preferring the object and horizontal edge to be at or near the same height and a systematic decrease in preference for increasingly large distances between them, at least for these particular perspective views of these two objects.

Why might there be such a strong bias toward preferring the horizontal edge to be at or near the same height as the object? One possibility for preferring the horizontal edge to be at the same level as the object is that the depth relation between the object and the edge is most evident in these images, because the edge is actually occluded by the object. This possibility does not account for the monotonic decrease in preference for the other edge positions. Perhaps the most satisfying explanation for the entire pattern of results is that it is related to the perspective from which the pictures of the objects were taken and the likely distance between the object and the far edge of the supporting surface. The present images were mainly side views taken at an angle of about  $18^\circ$  from horizontal. The supporting surface would thus be greatly foreshortened and its far edge would plausibly be relatively near the object in the projected image. For comparison, consider the images if the perspective on the objects had been pure side views or pure top (bottom) views. In a pure side view, the far edge of the supporting surface would have to be at essentially the same height as the bottom (top) of the bowl (light fixture), and in a pure top (bottom) view the horizontal edge would likely be entirely out of the picture.

We have argued that the lower and upper biases shown in Figure 6.3C and 6.3D may have arisen from an inward bias in the vertical dimension that is analogous to the inward bias in the horizontal dimension: Pictures of objects that "face" upward and are supported by a horizontal plane below them are preferred when they are located below the center of the frame (i.e., a lower bias), whereas objects that "face" downward and are attached to an overhead plane are preferred above the center (i.e., an upper bias). It seems most likely that this occurs, analogously

to the horizontal cases, because of functional asymmetries in the properties of the objects: Just as a person, car, or flower can “face” left or right in a picture, it seems plausible to claim that the bowl “faces” upward and the light fixture “faces” downward. However, unlike the horizontal cases, there are other confounding variables that prevent a definitive conclusion about the reason for the inward-facing pattern of preferences. First, bowls are typically located below the viewer, and light fixtures are typically located above the viewer. The upper and lower biases thus may arise from the aforementioned *ecological hypothesis* – a bias for the height of objects within rectangular images to reflect the height of the depicted objects relative to an observer in the real world. Second, the bowl was viewed from slightly above, and the light fixture was viewed from slightly below. The observed inward biases could result if people simply prefer objects seen from above to be low in the frame and objects seen from below to be high in the frame. This will be referred to as the *perspective hypothesis*. The next two experiments attempt to disentangle these possibilities.

### Experiment 2.3: Perspective Effects for the Bowl and Light Fixture

In Experiment 2.3 we attempted to distinguish among three different explanations for the inward bias found in Experiment 2.2: the effects of visual and functional asymmetries in the object (analogous to the horizontal inward facing bias), the effects of ecological height due to the positions of the objects relative to a typical human viewpoint, and the effects of perspective view due to the angle of the camera relative to a purely horizontal (side) view of the object. This

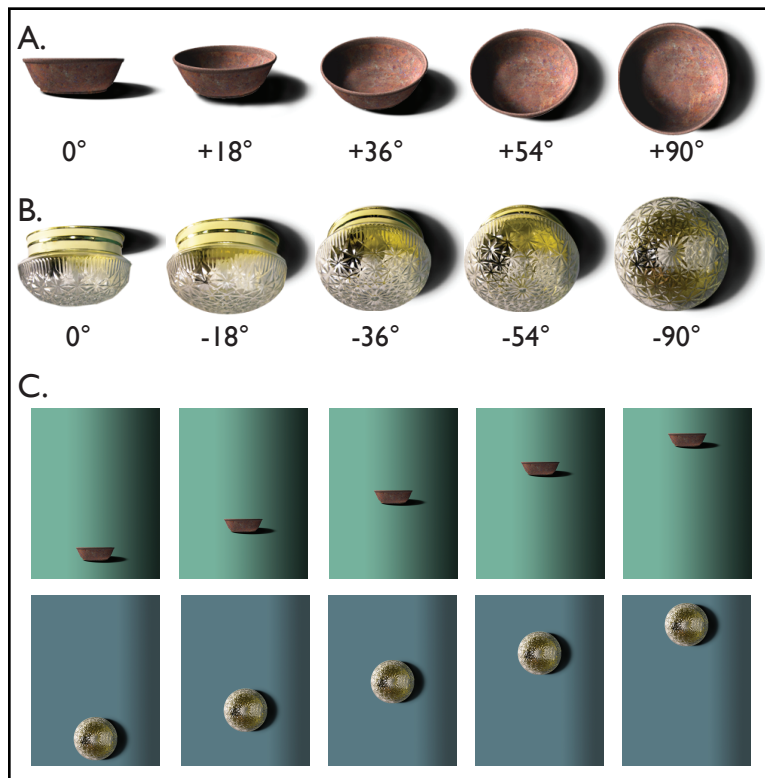


Figure 6.4. Displays for Experiment 3. A bowl (A) and a light fixture (B) were photographed from five different perspectives, (0° to 90° relative to a pure side view). Each object was then placed on a background at each of five different positions, as shown in panel C for the 0° view of the bowl and the -90° view of the light fixture.

was done by measuring preference effects in the vertical position of pictures of the bowl and light fixture taken from five different perspectives that formed a 90° arc around the object: 0° (a pure side view), 18°, 36°, 54°, or 90° (directly) above the bowl and 0°, -18°, -36°, -54°, or -90° (directly) below the light fixture (see Figures 6.4A and 6.4B). By analogy with the horizontal effects reported in the first set of experiments, a pure affordance asymmetry (“facing”) explanation predicts that the preference for low placement of the bowl and high placement of the fixture should be greatest for the pure side views (0°), where the object is functionally most asymmetrical, decrease monotonically as the angle increases from horizontal, and be entirely eliminated for the 90° above (below) views, from which the images of these object are

visually and functionally symmetrical when they directly “face” the camera. A pure ecological explanation implies that there should be no difference between the preferred height of an object depending on its perspective angle because all that matters is where they are located in the world: All of the bowls should be preferred equally low in the frame, and all the light fixtures should be preferred equally high. Finally, a pure perspective explanation – that pictures taken from above are preferred when the object is lower in the frame and those taken from below are preferred when the object is higher in the frame – implies that the perspective effects in the present experiment should be opposite those predicted from the functional asymmetry explanation: greatest from directly above (below), where the viewpoint is most completely from above (below) the bowl (light fixture), and eliminated for the pure side views, where the viewpoint is neither above or below. Again, it warrants reiteration that these alternatives are not mutually exclusive, and they may combine to give various mixtures.

To avoid the possible confounding of variables introduced by having a horizontal edge in the images, the objects were placed on a horizonless colored background with a soft lightness gradient that ran along the horizontal axis (see Figure 6.4C). The objects again appeared at 5 different vertical positions within the frame (see Figure 6.4C).

### Method

Participants. All seventeen participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. Their mean age was approximately 19.5 years. All were naïve to the purpose of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. Twenty paired comparisons resulted from all ordered pairs of the five different vertical positions of the same perspective view of the object, taken two at a time, for each of the two objects (bowl and light fixture). The five perspective views of the two different objects (at +/- 0°, 18°, 36°, 54°, and 90° from a pure side view) were never compared directly to each other, resulting in a total of 200 trials. The trials were blocked by object, so that half of the participants saw all of the light fixture trials first, and the other half saw all of the bowl trials first.

Displays. Each screen consisted of two colored images, measuring 450 x 600 pixels, placed on a neutral gray background. The images, shown in Figure 6.4, were placed at the same 450-pixels height on the screen, and were separated by 300 pixels. The images were created in Adobe Photoshop CS3. The LCD monitor measured 19” diagonally, and had a resolution of 1440 x 900 pixels and a refresh rate of 60 Hz.

Procedure. The procedure was identical to the Experiments 2.1 and 2.2.

### Results and Discussion

The results are plotted in Figure 6.5A and 6.5B for the bowl and fixture, respectively, as a function of object height relative to the frame and the angle of the depicted perspective relative to horizontal. The data were analyzed in terms of the percentage of trials on which each view of each object was chosen as preferred at each vertical position in the frame. Consistent with the results of Experiment 2.2, there was an overall lower bias for the bowl ( $F(1,16) = 21.11, p < .001$ ) (see Figure 6.5A) and an overall upper bias for the light fixture ( $F(1,16) = 32.96, p < .001$ ) (see Figure 6.5B), which resulted in a powerful interaction between object and vertical position ( $F(4,64) = 32.51, p < .001$ ). Smaller, but still significant, interactions were present between perspective and position ( $F(16,256) = 2.52, p < .001$ ) and among object, perspective, and vertical

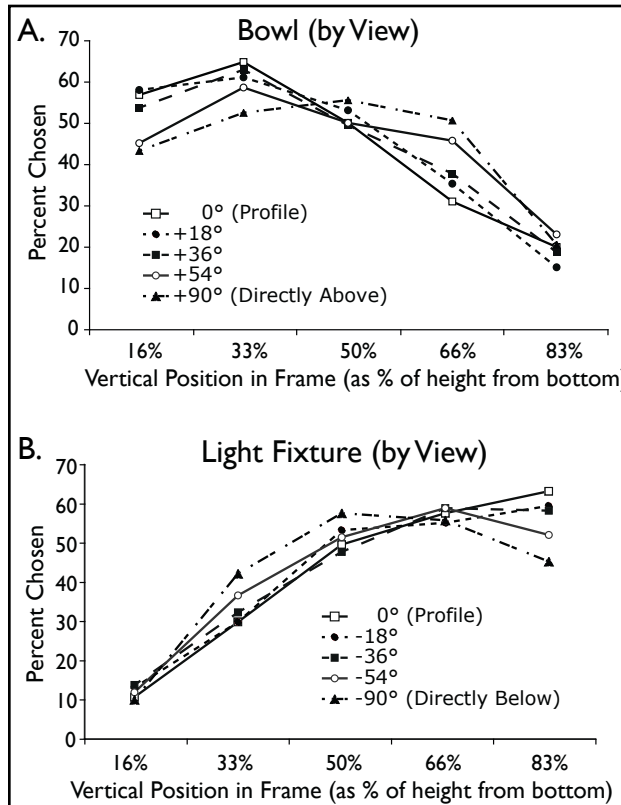


Figure 6.5. Results of Experiment 2.3. The percentage of trials in which the given image was preferred over all possible combinations of the same view of the object is plotted for the five perspective views of the bowl (A) and the light fixture (B).

position ( $F(16,256) = 3.61, p < .001$ ). No effect of block order or any of its interactions were statistically reliable (all  $F_s < 1.8$ , all  $p_s > .15$ ).

The interactions between vertical position and perspective for the bowl (Figure 6.5A) and light fixture (Figure 6.5B) arise from seemingly complex differences in the slopes and shapes of the preferences curves. The functions are relatively straight and steep for the  $0^\circ$  (side-view) perspectives, but transform rather smoothly to the  $90^\circ$  (top/bottom) perspectives, which are inverted-U shapes that are much more symmetrical. The structure of these interactions can be nicely captured by examining the amount of variance accounted for by the linear and quadratic components of the functions, which together account for an average of 92% or the variance in these curves, ranging from 82% to 99.8%. Consistent with the work on horizontal preferences in the first set of experiments, we interpret the linear trends as indicating the strength of inward biases (i.e., the lower bias for the bowl and the upper bias for the light fixture). We interpret the quadratic trends, which peak at the central position, as indicating a strong center bias.

Figure 6.6 shows the proportion of variance explained by the linear and quadratic components for the bowl and light fixture as a function of perspective. As the angle from horizontal increases (for the bowl) or decreases (for the light fixture) and the object's image becomes more symmetrical about a horizontal axis, the variance attributable to the linear trend decreases systematically and that attributable to the quadratic trend increases systematically. This pattern of biases is also characteristic of the results found in the horizontal preference functions: When the perspective view was strongly directed and asymmetrical about the relevant axis (i.e., a pure side view), people strongly preferred the object to "face" into the frame, but when it was facing the viewer and symmetrical, they tended to prefer the object to be centered in the frame. The present data thus strongly support the functional asymmetry explanation, which predicts the reduction of the inward bias with deviations from a purely side view in perspective. They do not appear to support the perspective explanation, which predicts the opposite trend (i.e., a larger inward bias as the perspective deviates from  $0^\circ$ ), but it is logically possible that a larger functional asymmetry component combines with a smaller perspective effect to produce the observed pattern. It is interesting that the linear component never disappears entirely, accounting for almost half of the variance even in the  $90^\circ$  perspective views from directly above the bowl and below the light fixture, so the existence of a perspective effect cannot be ruled out, but this might also be due to an ecological effect. Experiment 2.4 further investigates the possible influences of perspective, functional asymmetry, and ecological factors by using objects that are not constrained in the perspectives from which they can be viewed.

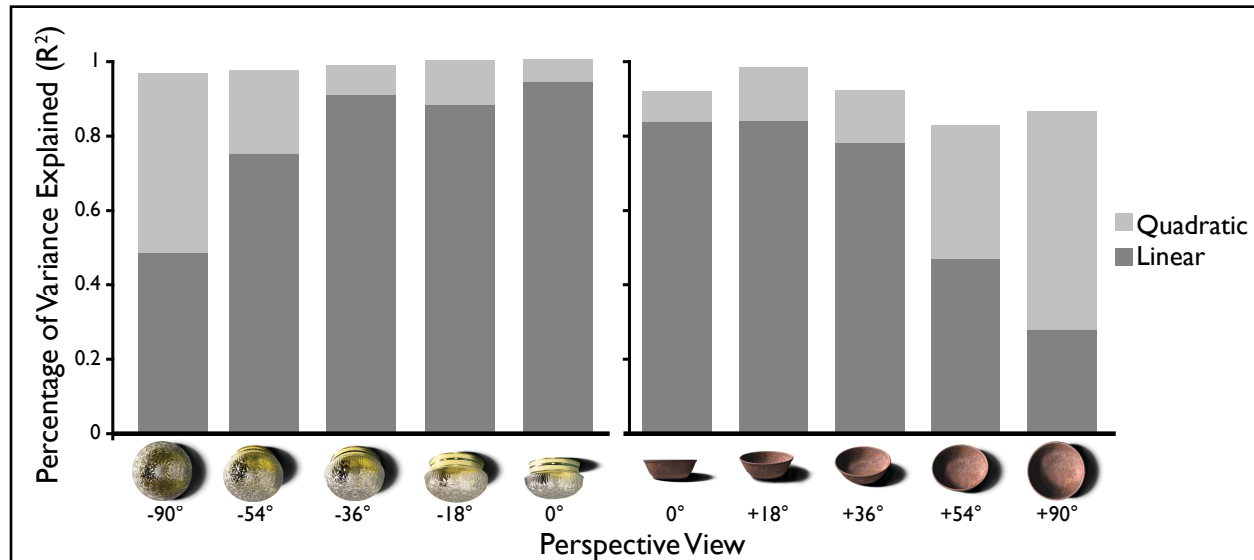


Figure 6.6. A stacked bar graph showing the percentage of variance in the data for the bowl and for the light fixture that are explained by a linear component (darker gray), and a quadratic component (lighter gray). Note that the patterns are similar for the two objects in terms of the deviation in perspective viewing angle from a pure side view (0°).

## Experiment 2.4: Objects from Unconstrained Perspectives: Flying Eagles and Swimming Stingrays

Experiment 2.4 investigated the upper and lower biases of Experiments 2.2 and 2.3 more extensively to shed further light on possible causes. As just discussed, the findings from Experiments 2.1-2.3 seem to arise from functional asymmetries in the objects themselves plus ecological effects due to their typical location relative to the viewer and/or perspective effects of whether the object is viewed from above or below. One drawback to the bowl and light fixture used in the previous studies is that, because they are supported by an opaque surface below and above them, respectively, not all perspective views are possible: The bowl cannot be viewed from below nor the light fixture from above. As a result, ecological effects are necessarily confounded with perspective effects: The bowl, typically located below the viewer and supported by a surface below it, is almost always seen from a vantage point at least somewhat above it, and the light fixture, typically located above the viewer and supported by a surface above it, is almost always seen from a vantage point at least somewhat below it.

To dissociate these effects, preferences were studied for vertical placement of two objects that could be viewed from any perspective angle, yet have characteristic positions above or below human observers: a flying eagle, which is typically located above human observers, and a swimming stingray which is typically located below human observers. By using such unattached objects, we generated believable images from viewpoints that covered an entire 180° arc around one side of the object – from directly above to directly below – for both an object that is typically located above the viewer and one that is typically located below the observer.

Two further issues are also addressed in this experiment: generalizing the results to objects radically different from a bowl and light fixture and the eliminating the possibility of “demand characteristics” from having the same observers viewing both objects. Experiments 2.1-2.3 all used a bowl and light fixture as representative objects located below and above the observer, respectively, but it makes sense to replicate the results with other objects. If the causes of the effects observed in previous experiments are general, the same pattern of results should be



found for the stingray and eagle, which are quite different from a bowl and light fixture. Experiments 2.2 and 2.3 also used a fully within-participant design, such that the same observers saw both the bowl and the light fixture images. Perhaps they preferred the bowl to be lower and the light fixture to be higher in the frame only because they saw both kinds of objects and felt pressured to make some kind of height distinction between them. In the present experiment, different participants made judgments about the eagle images and the stingray images to avoid this possibility.

### Method

Participants. All 24 participants were students at the University of California, Berkeley, who received partial course credit in their undergraduate psychology course. They were randomly assigned to the eagle and stingray conditions. Their mean age was approximately 19.1 years. All were naïve to the purpose of the experiment and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. There were 20 paired comparisons that resulted from all ordered pairs of the five different vertical positions in the frame of the same perspective view of one object, taken two at a time, for each of the two facing directions (left and right). There were four perspective views of the two objects – directly-above (+90°), side-above (+45), side-below (-45), and directly-below (-90°) relative to a pure side view – which were never compared to each other, resulting in a total of 160 trials. (The pure side view was not included to reduce the number of trials, because we were most interested in the cases in which the perspective includes an upward or downward component.) To avoid demand characteristics resulting from seeing both the eagle and stingray images together, the object (eagle or stingray) was a between-subjects variable. Within each group, all trials were randomized.

Displays. Each screen consisted of two colored images, measuring 400 x 400 pixels, on a background of neutral gray. The images were placed at the same 400-pixel height on the screen, and were separated by 200 pixels. The 3D models of the eagle and the stingray were purchased from Content Paradise ([www.contentparadise.com](http://www.contentparadise.com)), and were arranged in Poser 6. Both left-facing and right-facing versions of both objects were used. Backgrounds were added and the images were finalized in Adobe Photoshop CS3. For the eagle, the directly-above and side-above images had a background of mottled green that resembled ground or grass, and the directly-below and side-below images had a background of sky and clouds on a bright day (Figure 6.7A). For the stingray, the directly-above and side-above images had a background of dark turquoise that resembled a view into deeper water, and the directly-below and side-below images had a background of a lighter turquoise that resembled a view toward the water's surface (Figure 6.7B). In side-above and side-below views of the stingray, a vertical lightness gradient was imposed on the background to imply deeper water below the stingray and lighter water above the stingray. The LCD monitor measured 20" diagonally with a resolution of 1280 x 800 pixels and a refresh rate of 60Hz.

Procedure. The procedure was identical to the first three experiments.

### Results and Discussion

Again, the average percentages of trials on which each given image was chosen over all of its 2AFC comparison images were computed. These data were analyzed separately for eagle and stingray conditions. Because the left/right facing direction of the eagle and stingray had no effect on participants' preferences for their respective heights in the frame (Eagle:  $F < 1$ ; Stingray:  $F$

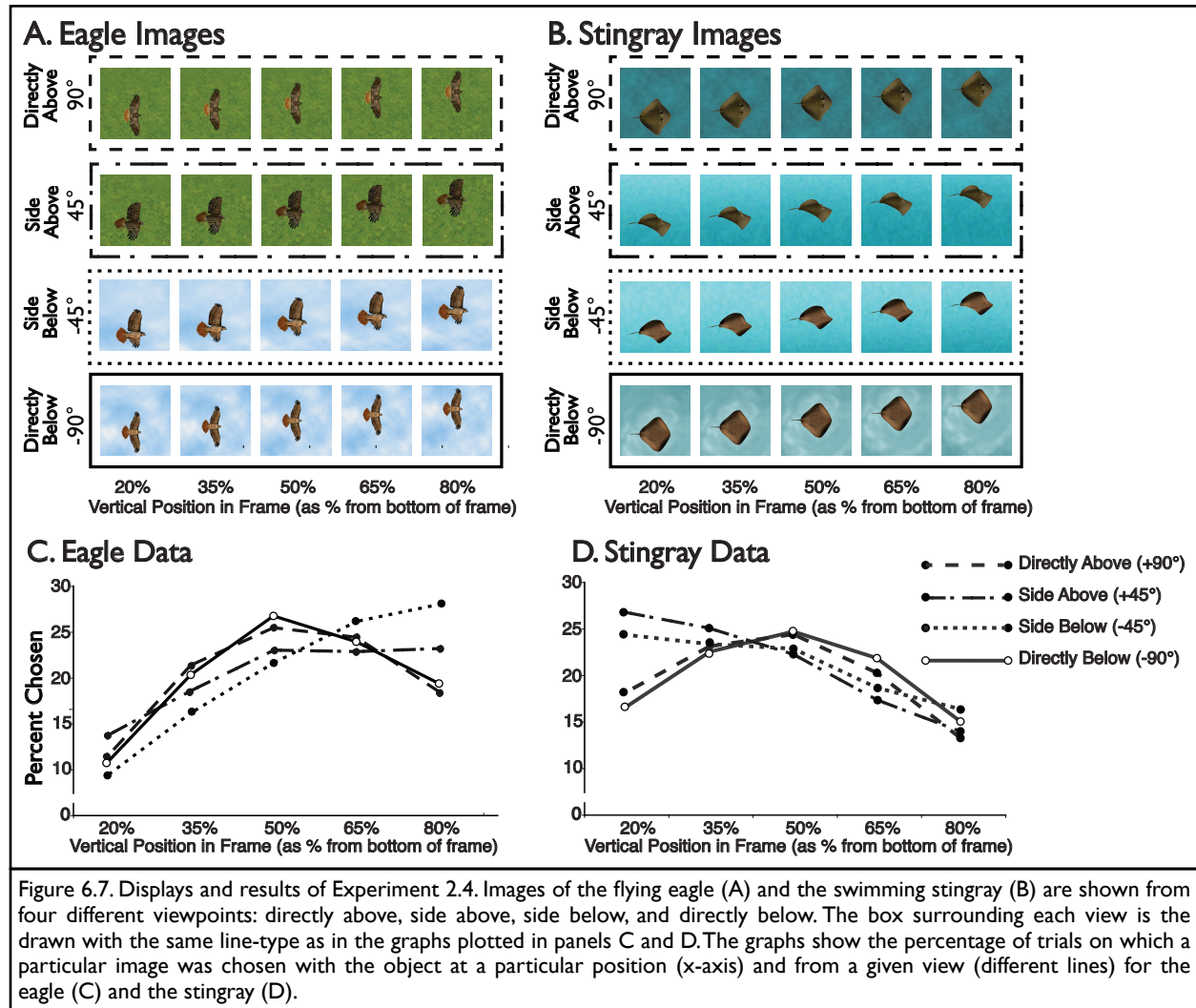


Figure 6.7. Displays and results of Experiment 2.4. Images of the flying eagle (A) and the swimming stingray (B) are shown from four different viewpoints: directly above, side above, side below, and directly below. The box surrounding each view is the drawn with the same line-type as in the graphs plotted in panels C and D. The graphs show the percentage of trials on which a particular image was chosen with the object at a particular position (x-axis) and from a given view (different lines) for the eagle (C) and the stingray (D).

(4,44) = 1.10,  $p > .3$ ), the data were averaged over facing directions in all subsequent analyses. There was a significant overall interaction between perspective viewpoint and vertical position of the object for both the eagle,  $F(12,132) = 5.23$ ,  $p < .001$  (see Figure 6.7C), and the stingray,  $F(12,132) = 4.80$ ,  $p < .001$  (see Figure 6.7D), arising from distinct patterns of positional preference for different perspective views. The percentages of variance accounted for by the linear and quadratic components of these curves are graphed in Figure 6.8.

Consider first the results in the directly-above and directly-below perspectives. For neither the eagle nor the stingray did it matter whether the view was from directly above or directly below (eagle:  $F(4,44) = 1.05$ ,  $p > .4$ ; stingray:  $F(4,44) = 1.11$ ,  $p > .3$ ). Both objects show a strong component of the inverted-U shaped (quadratic) functions characteristic of the center bias we have previously found with preferences for the horizontal position of symmetrical views of objects (eagle:  $F(1,11) = 8.60$ ,  $p < .02$ ; stingray:  $F(1,11) = 7.84$ ,  $p < .02$ ). In both present cases, however, there is also a linear component that was not found for the horizontal position of symmetrical objects: The eagle function was tilted upward, indicating preference for higher placements ( $F(1,11) = 11.23$ ,  $p < .01$ ), and the stingray function was tilted downward, indicating preference for lower placements ( $F(1,11) = 4.87$ ,  $p < .05$ ). The facts that these linear

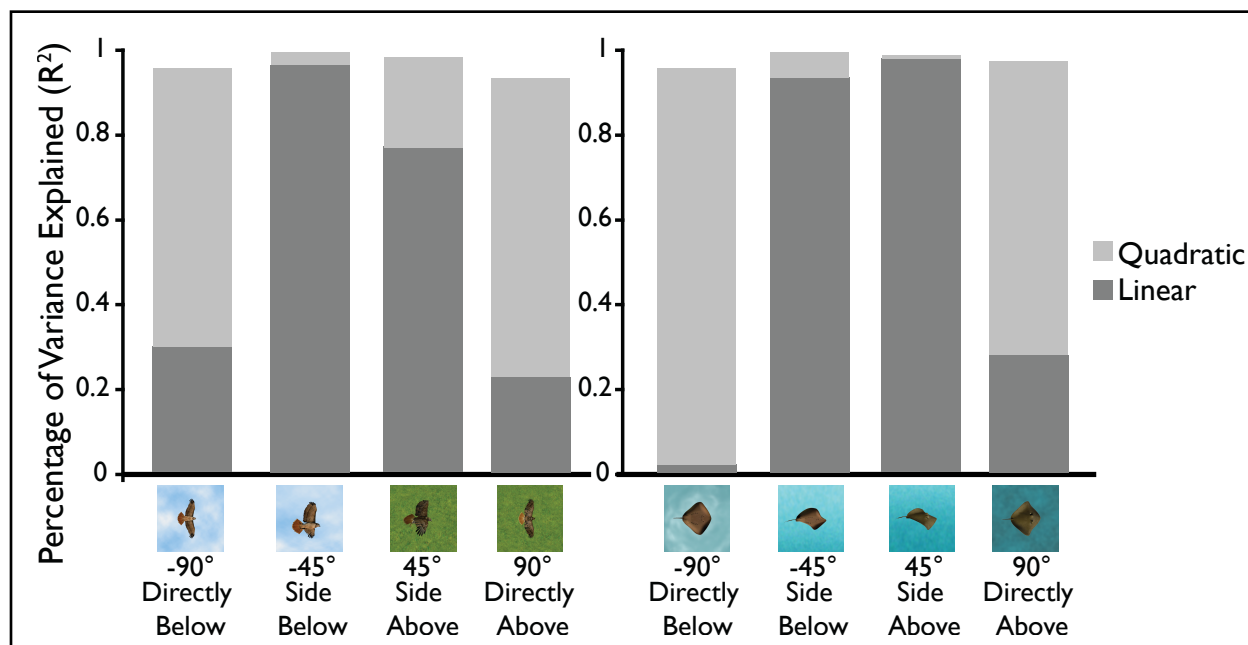


Figure 6.8. A stacked bar graph showing the percentage of variance in the data for the eagle and for the stingray that is explained by a linear component (darker gray), and a quadratic component (lighter gray). Note that the pattern is similar for the two objects in terms of the deviation in perspective viewing angle from a pure side view (0°).

components are positive in sign for the eagle and negative for the stingray and that they do not differ for the directly above versus directly below conditions for either object strongly suggests that they arise from an ecological bias, which predicts a positive slope (i.e., upper bias) for the eagle and a negative slope (i.e., lower bias) for the stingray. If these effects were driven by a perspective bias due to looking upward or downward at the object, there should be an upper bias for the objects viewed from below and a lower bias for objects viewed from above, independent of which object was depicted. Clearly, this is not what happened.

Both the side-above and side-below views also showed a strong upper bias for the eagle and a strong lower bias for the stingray, as evidenced by reliable linear components for both objects, though opposite in slope (eagle side-above:  $F(1,11) = 6.26, p < .05$ ; side-below:  $F(1,11) = 42.25, p < .001$ ; stingray side-above:  $F(1,11) = 13.75, p < .01$ ; side-below:  $F(1,11) = 6.191, p < .05$ ). In the case of the eagle, these two perspectives differed significantly from each other ( $F(4,44) = 4.68, p < .01$ ), because the upper bias was stronger when the eagle was viewed from below than when it was viewed from above. The same trend was evident in the linear  $\times$  linear component interaction in the stingray data ( $F(1,11) = 4.85, p < .05$ ). Given the ecological effects apparent in the views from directly above and below, the most consistent interpretation of these results is that the flying eagle images produced a strong ecological upper bias and a weaker inward bias, such that the view from below (showing the bird's belly) is preferred relatively higher in the frame than the view from above (showing the bird's back). As the graph in 6.7C shows, at positions that were 65% and 80% from the bottom of the frame, the eagle viewed from below was more preferred ( $F(1,11) = 11.50, 7.50, p < .01, .02$ , respectively), whereas at positions 20% and 35% from the bottom of the frame, the eagle viewed from above was more preferred ( $F(1,11) = 13.92, 4.85, p < .01, .05$ , respectively). Corresponding trends were apparent in the side-above and side-below views of the stingray, but the corresponding contrasts were not statistically significant ( $F(1,11) = 4.376, 3.702, p < .07, .08$ , respectively). This pattern of data suggests that, in addition to strong ecological effects, there may be either weak perspective

effects or weak functional asymmetry effects in these conditions because the upper bias is slightly stronger for the eagle when looking up at it, whereas the lower bias is slightly stronger for the stingray when looking down at it.

At least two, and possibly three, factors thus appear to be operating in producing the pattern of preferences evident in this experiment. First, there is a strong overall upper bias for pictures of flying eagles to be located toward the top of the frame and a lower bias for pictures of swimming stingrays to be located toward the bottom. These effects are reflected in the linear components of the perspective views for both objects (see Figure 6.8) and are presumably due to the fact that flying eagles tend to be located above human observers in the world whereas swimming stingrays are typically located below human observers in the world. Second, there is a relatively strong center bias for the views from directly above and directly below, which is still evident in the side-above view of the eagle (and perhaps weakly evident in the side-below view of the stingray), and nearly absent in the side-below view of the eagle and side-above view of the stingray (see Figures 6.7C, 6.7D, and 6.8). This center bias is presumably due to the symmetrical structure of the eagle's and stingray's left and right sides. Third, there may be an additional, but smaller, upper bias for the side-below view of the eagle and lower bias for the side-above view of the stingray that is likely to be due to some asymmetry in function or visual salience between the top and bottom surfaces of the eagle and stingray. This component statistically isolated from the ecological bias, because both are reflected in the linear components plotted in Figure 6.8, but its weakness in the directly above and below views implies that it is unlikely to be due to perspective effects produced by looking upward and downward per se. Below we suggest that the center bias arising from the left-right symmetries of the objects and the inward bias arising from top-bottom asymmetries may reflect a single underlying factor, which we will call affordance space asymmetries.

## Conclusion for Experiments on Vertical Position

Experiments 2.1, 2.2, 2.3, and 2.4 explored people's aesthetic preferences for spatial composition in the vertical dimension for single-object pictures within a rectangular frame. The results were partly analogous to our findings on preference for composition in the horizontal dimension in that they demonstrated the existence of a ubiquitous center bias and some sort of inward bias. The inward bias in the vertical dimension differs from that in the horizontal dimension, however, in that it arises for different objects rather than different facing directions of the same object: There was a lower bias for some objects (a bowl and swimming stingray) and an upper bias for others (a light fixture and a flying eagle). Biases in the vertical dimension also proved to be more complex than in the horizontal dimension in that the inward bias appears to arise from multiple factors, such as what we are calling affordance asymmetry effects, ecological effects, and possibly perspective effects. Below, each of these factors is discussed in greater depth.

What we are calling affordance asymmetry effects would arise because functional interactions with many objects are not equally distributed over the space surrounding them. The fact that fronts of objects are almost always more salient in interactions than the sides and backs of objects can explain the inward bias found in the horizontal placement of an object within a picture (see Chapter 5): People seem simply to prefer the more salient functional parts of the object to be closer to the center. By analogous logic, the primary affordances of a bowl concern the space inside and above it, because that is where substances are placed into it and removed from it. The sides and bottom of the bowl are certainly important for its ability to

contain those substances, to be moved from place to place, and to rest stably on a supporting surface, but the space above it seems to be far more salient for the bowl's most important functions than the space below it, thus producing a substantial functional asymmetry in the vertical dimension. If there is a bias toward having the more salient functional parts and area around them closer to the center of the frame, then a purely side-view of a bowl should be preferred when it is located toward the bottom of the frame. Note that there is no such affordance asymmetry in a purely top-view of the same bowl, because the bowl is both physically and functionally symmetric with respect to the area of space depicted around its sides, so any affordance asymmetry bias should diminish progressively as the perspective rises from a pure side view to a pure top view. Similarly, a ceiling-mounted light fixture's primary affordance is to emit light below it. A purely side-view of such a fixture thus exhibits a distinct functional asymmetry in which the space below it is more important than the space above it. The space around a light fixture in a purely bottom-view, however, is both physically and functionally symmetric and should therefore not be systematically biased toward a lower or higher position. This is precisely the pattern found in Experiments 2.3 and 2.4 for the percentage of variance accounted for by the linear components of the results (see Figures 6.6 and 6.8).

An even simpler explanation can be devised by positing the existence of a perceptually defined "affordance space" around an object that reflects the extent and/or importance of functions that take place in that region of the object. If such an affordance space is asymmetrical, as suggested above, then what we call an inward bias may actually be understood as a center bias that operates on an asymmetrical affordance space with more area on the functionally more salient side(s). That is, if viewers implicitly prefer the affordance space around an object to be centered in the frame, and if that affordance space is asymmetrical toward the front of horizontally facing objects (e.g., a person, chair, or vehicle), toward the top of "upward facing" objects (e.g., a bowl), and toward the bottom of "downward facing" objects (e.g., a light fixture), then most of the inward biases found in both horizontal and vertical dimensions may actually be center biases on affordance spaces. This is an empirical question which could be addressed by future research.

Perspective height effects in aesthetic preferences refer to possible differences in the preferred vertical placement of an object within a frame that depend on the direction from which the object is viewed. My hypothesis that such effects might exist arose from another possible explanation of the inward biases found in Experiments 2.2: People may tend to prefer an object typically viewed from above (e.g., a bowl) to be located below the center of the frame, and they may prefer an object typically viewed from below (e.g., a light fixture) to be located above the center of the frame. The implications of this possibility were investigated in more detail in Experiments 2.3 and 2.4, but the results posed several problems. First, the upper and lower biases in Experiment 2.3 are clearly strongest when viewing objects from the side, weakening considerably as the perspective moves toward views from directly above and below (see Figures 6.5 and 6.6), where one would logically expect perspective effects to be strongest. Second, when perspective effects were dissociated from ecological effects in Experiment 2.4, there was no difference between views of the same object (for either eagle or stingray) when viewed from directly above versus directly below (see Figures 6.7 and 6.8). Because any perspective effect would presumably have to be maximal such contrasts between purely upward and purely downward views of the same object, we conclude that, despite its intuitive plausibility, the vertical component of the perspective viewing angle is not a significant factor in driving esthetic preferences for vertical placement of a focal object within a picture.

Ecological effects are based on the fact that some objects tend to be located higher than the observer in the environment (e.g., flying eagles and light fixtures) and others tend to be located lower (e.g., bowls and stingrays). It appears to cause strong and pervasive inward height biases in the present results in that people prefer the vertical position of an object within the frame to be consistent with the vertical position of the object relative to the observer: i.e., eagles and light fixtures to be high and bowls and stingrays to be low. Although the magnitude of this bias is presumably constant across perspective conditions in Experiments 2.3 and 2.4, the magnitude of the inward bias varies because it seems to combine with affordance asymmetries. There were not any corresponding ecological biases in the horizontal position of objects, presumably because objects do not have characteristic positions to the left or right of mobile human observers in a three-dimensional terrestrial world.

The second set of experiments in this dissertation took the results of the first set of experiments on horizontal position and set out to see how they applied to the vertical dimension. The results were more complicated than originally anticipated, demonstrating influences of several factors, including spatial asymmetries in the affordances of the object and the typical position of the object relative to the observer. With asymmetric side views of objects, people generally prefer objects typically below the observer's viewpoint (e.g., a bowl and swimming stingray) to be located below the center of the frame and objects located above the observer's viewpoint (e.g., a light fixture and flying eagle) to be located above the center of the frame. In addition, people also generally prefer symmetric views of those same objects from directly above or directly below to be closer to the center of the frame. The discussion of these experiments was concluded by suggesting that these results can be unified by the hypothesis that observer's prefer the object's "affordance space" to be centered within the frame. Finally, these experiments join the first set of experiments to demonstrate that aesthetic preferences can be not only clean and orderly if studied in the right way, but also illustrative of principles that can help us to better understand how people view the world in general.

## Chapter 7

# Interpreting the Experiments on Horizontal and Vertical Position

The results of the two sets of experiments firmly establish that systematic compositional biases are present in people's aesthetic judgments of framed images. I would now like to briefly discuss the deeper question of *why* this might be the case. In the domain of color aesthetics, Palmer and Schloss (2010) make a strong case that preferences perform an adaptive "steering" function, biasing sighted observers to approach objects that are likely to be beneficial and to avoid objects that are likely to be detrimental. This explanation is satisfying from an adaptive, evolutionary perspective, because acting in accord with such aesthetic preferences would be beneficial for the organism to the extent that the preferences are correlated with (i.e., carry predictive information about) what is "good" versus "bad" for the observer. It is not as obvious what adaptive function might be served by compositional biases: the center bias, the inward bias, and the various ecological biases (height, perspective, and size). An unavoidable problem for any adaptive theory of these biases is that they apply to framed two-dimensional visual displays of objects that did not exist when humankind was presumably being shaped by evolution.

Even in modern times, the composition of rectangular, framed images seem to be largely irrelevant to people's lives, at least outside of art, wall decorations, website design, and other aesthetically relevant domains. To be more generally relevant, the domain to which these principles apply would have to be broadened to include more adaptive decisions and choice behaviors.

### Eye Movements and the Center Bias

One intriguing possibility is that the compositional biases discovered in the present research may be related to the processes of making optimal eye fixations and other bodily adjustments in actively exploring the world. People make thousands of eye movements every day to bring various ecological scenes, objects, and parts into view so that the observer can see them clearly.

Among the most general and ubiquitous purposes of eye fixations is to identify objects and assess their affordances in relation to the observer's current plans and goals (Gibson, 1977). As discussed previously, affordances could account well for the aesthetic effects found in spatial composition, which, in turn, may be rooted in principles that the visual system uses to optimize the visual field for perceiving the environment more quickly and accurately. What exactly needs to be seen, of course, varies greatly from one task to another, but the task of object identification will surely be among the most important.

The general idea is that if the composition of an image within a rectangular frame is conceived as roughly analogous to the positional structure of objects within the visual field, it would be adaptive for people to make fixations that make the most important information about the most relevant objects most readily available in the image. The center bias would be related to the strong foveation of retinal receptors and the cortical magnification of information at or near the central area of the visual field. The functional asymmetry biases would similarly be related to putting the most important and informative regions of the image at or near the fovea. Ecological biases would be related to providing proximal image features that are consistent with distal object features, depicting objects from canonical perspectives, and at sizes and positions within the frame that mirror the corresponding properties of the depicted object.

Indeed, the literature on eye-tracking demonstrates that there is a “central fixation bias” in scene perception: people tend to fixate on the center of a rectangularly-framed scene before making eye movements elsewhere. Once they make saccades to other parts of the scene, they tend to fixate the centers of whatever objects or groups of objects they are looking at. (See Rayner & Pollatsek, 1992, for a review of eye movements in scene perception, and Zelinsky et al., 1997, for further experiments demonstrating this).<sup>20</sup> In experiments designed to get at the origin of this bias, neither image location, nor image features, nor motor bias could adequately explain the bias (Tatler, 2007).

There are differences between rectangular frames and the field of vision, to be sure. One is that the frame of a picture is explicitly visible whereas the boundary of the visual field is not, being defined merely by the absence of sensory input. Another is that the shape of the visual field is roughly oval rather than rectangular. Such relatively minor differences aside, however, the “adaptive vision” hypothesis of such aesthetic effects provides a plausible, ecologically relevant rationale for why people might have these kinds of default biases. These ideas are as yet mere conjectures, of course, but that is always the starting point for the next round of empirical tests. They at least have the virtue of making a bridge between aesthetic preferences and adaptive aspects of real-world perception.

### Fluency: A Possible Explanatory Theory

The existence of ecological height effects in the vertical preference data suggests that aesthetic preferences are driven to some degree by how well the position of the focal object within the spatial structure of the framed image fits the observer’s expectations about the position of such objects within corresponding environmental scenes. In this regard, ecological effects are consistent with “fluency” accounts of aesthetic preference discussed earlier (e.g., Reber, et al., 1998, 2004; Winkielman, et al., 2002, 2006), according to which people prefer images that are more easily (fluently) perceived. (See Chapter 3 for a more thorough description of fluency theory). It has been known for many years, for example, that people are faster and more accurate at recognizing objects in appropriate locations within otherwise coherent contextual scenes (e.g., Biederman et al., 1982; Palmer, 1975). In these previous studies, the appropriateness of the target object’s position was defined in relation to other objects, such as a loaf of bread being appropriately located on a cutting board on a kitchen counter next to, say, a bread knife and a piece of cheese rather than perched atop a post in the front yard, as a mailbox might be. More recently, Estes et al. (2008) demonstrated ecological height effects on object recognition by showing that participants were faster at discriminating snakes from birds when the birds were high in the frame and the snakes were low than when the birds were low and the snakes were high. This pattern of results, in which subjective measures of aesthetic preference mirror objective measures of perceptual performance, thus supports the primary claim of fluency theory.

Additional evidence for a concordance between aesthetic judgments and recognition performance can be found for other compositional factors. In a pioneering study of “canonical perspective,” Palmer, Rosch, and Chase (1981) demonstrated that everyday objects are more quickly and easily perceived from some perspectives than others. Khalil and McBeath (2006) later reported that more canonical views are also rated as more aesthetically pleasing than less

---

<sup>20</sup> Thanks are due to Tim J. Smith of Birkbeck, University of London, for pointing out this body of literature and its relevance to my work.



canonical ones. Similar effects have recently been found to hold in the size domain. Konkle and Oliva (in press) have shown that people prefer pictures of objects of different sizes to be depicted in 2D images with corresponding differences in their depicted sizes relative to a surrounding frame: The image of a mouse, for example, should be smaller than the image of an elephant within the same-sized frame. Konkle and Oliva call this the “canonical visual size” effect. Linsen, Leysson, Gardner, and Palmer (submitted) have confirmed that this size preference holds for aesthetic judgments as well: People rate pictures of a butterfly as most aesthetically pleasing at a smaller size within its frame than they do for pictures of an elephant. Although no corresponding studies of the effects of size on recognition performance measures have yet appeared, it seems plausible that similar results would be obtained: A mouse would be recognized more quickly and accurately if its image is small (rather than large) within its frame, and an elephant’s if its image is large (rather than small). Thus, it appears plausible that people’s aesthetic judgments will generally be higher under the same conditions in which their recognition performance is better. This confluence of results, assuming it continues to hold, thus supports the fluency account of aesthetic response: People generally prefer the images in which the objects are most easily perceived and recognized.

### Three Problems With Fluency

As promising as it may seem at first, however, fluency accounts have at least three problems. A fourth, more severe problem will be presented in the next chapter.

#### 1. Novelty

To say that the easier something is to process and the more familiar it is to the viewer, the more pleasing it will be strongly contradicts Berlyne’s claim that novelty per se tends to increase aesthetic response as well as the sizable body of experimental work that supports this claim. This is a significant problem.

#### 2. Fluent Images We Don’t Like

Another problem with the fluency account of aesthetic response is that it would imply that people would have positive aesthetic responses much more frequently than they do in response to images that rarely -- if ever -- engender positive aesthetic responses. There are many images that people see over the course of their daily lives that are very easy to process both perceptually and conceptually, and thus highly fluent: traffic signs, corporate logos, cell phone interface screens, and the canonical stick figures that designate men’s and women’s restrooms, to choose a few very fluent images. It seems unlikely that any of these examples, especially the restroom signs, ever gave anyone a positive aesthetic experience, even though they are all quickly recognizable, perhaps the primary hallmark of a fluent image. This suggests that fluency is unlikely to be the root cause of people’s positive aesthetic responses, but rather just a factor that influences them. I will claim below that it is actually a byproduct of the true underlying factor: a proper fit between the image and its context.

#### 3. Disfluent Images We Do Like

If fluency is indeed the main determinant of aesthetic preference, then images that are not fluent ought not to be pleasing. This would be rather easy to test empirically (and will be examined in Chapter 8), but before delving into empirical examinations, it is worth exploring the work of three artists as exemplars: Pablo Picasso, Giorgio De Chirico, and Jean Paul Bourdier, all of whom have produced works that are purposely disfluent, yet, at the same time,

aesthetically pleasing.

Though Pablo Picasso (1881-1973) did, at the beginning of his career, produce fairly conventional paintings (e.g., *Femme Aux Bras Croisés*, 1902; *La Vie*, 1903), he is best known for his Cubist paintings, such as *Femme á la Mandoline* (1910), shown in Figure 7.1A. This image is not at all fluent. The woman it depicts has a face and body that has been rearranged, reimagined, and highly abstracted. Neither the colors nor the shading used to delineate the shapes is very true-to-life. In spite of that (one might even say, *because of that*), this image can be highly pleasing, perhaps even moreso than other images of women that are more realistic and therefore, more fluent (e.g., Vasily Surikov's *Woman With Guitar* (1882), shown in Figure 7.1B).<sup>21</sup>

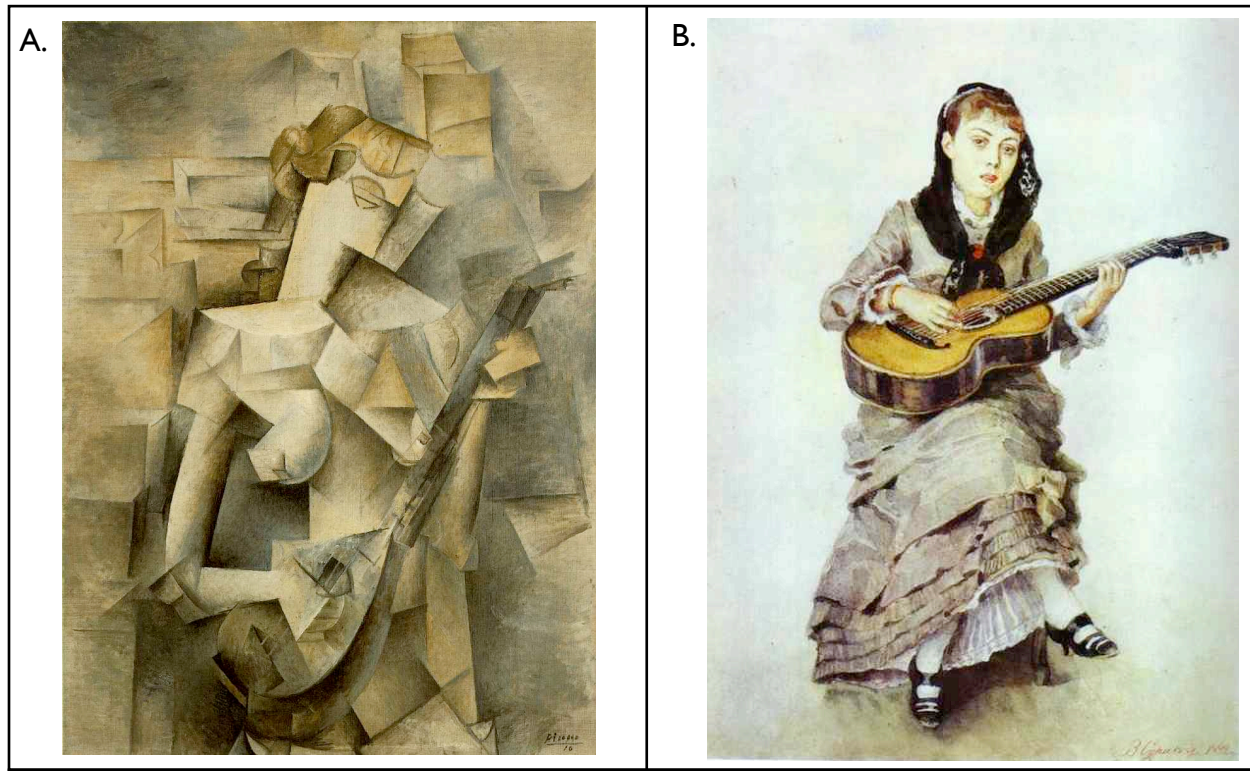


Figure 7.1. (A) A painting by Pablo Picasso, *Femme á la Mandoline* (1910), which is highly disfluent but can still be pleasing (if not more pleasing) compared to more fluent images depicting the same subject, such as Vasily Surikov's *Woman with Guitar* (1882).

In addition to the objects being fluent or non-fluent, artists can manipulate how fluent the depicted scene as a whole is. Giorgio de Chirico's surrealist paintings often contain arrangements of objects -- fluent in and of themselves -- in combinations that are often rather odd and unusual, such as in *Love Song* (1914) and *The Double Dream of Spring* (1915) (Figure 7.2A, 7.2B, respectively). These arrangements can hamper any sort of ease of perception of the scene (what easily recognizable sort of scene is being depicted that could simultaneously include

<sup>21</sup> On the topic of Picasso's depictions of the human form, E.H. Gombrich and Rudolf Arnheim, in comments exchanged over the course of their books and essays, got into a dispute over what constituted realism. Arnheim argued for a relativistic definition that left room for highly abstracted paintings to be included in it: "only a further shift of the reality level is needed to make the Picassos, the Bracques, or the Klees look exactly like the things they represent" (Arnheim, 1954; n.b., this particular quotation was removed in the second edition of Arnheim's book in 1974. The sentiment, however, remained intact.). Gombrich countered that such a definition of realism was so broad that it denied vision any sort of objectivity (Gombrich, 1960). In this particular case, I am inclined to agree with Gombrich; that being said, if such depictions as Arnheim mentions could ever truly be judged as realistic depictions of objects, then perhaps the images I have described could be deemed fluent. However, delving further into such stretched definitions of realism are beyond the scope of my argument here.

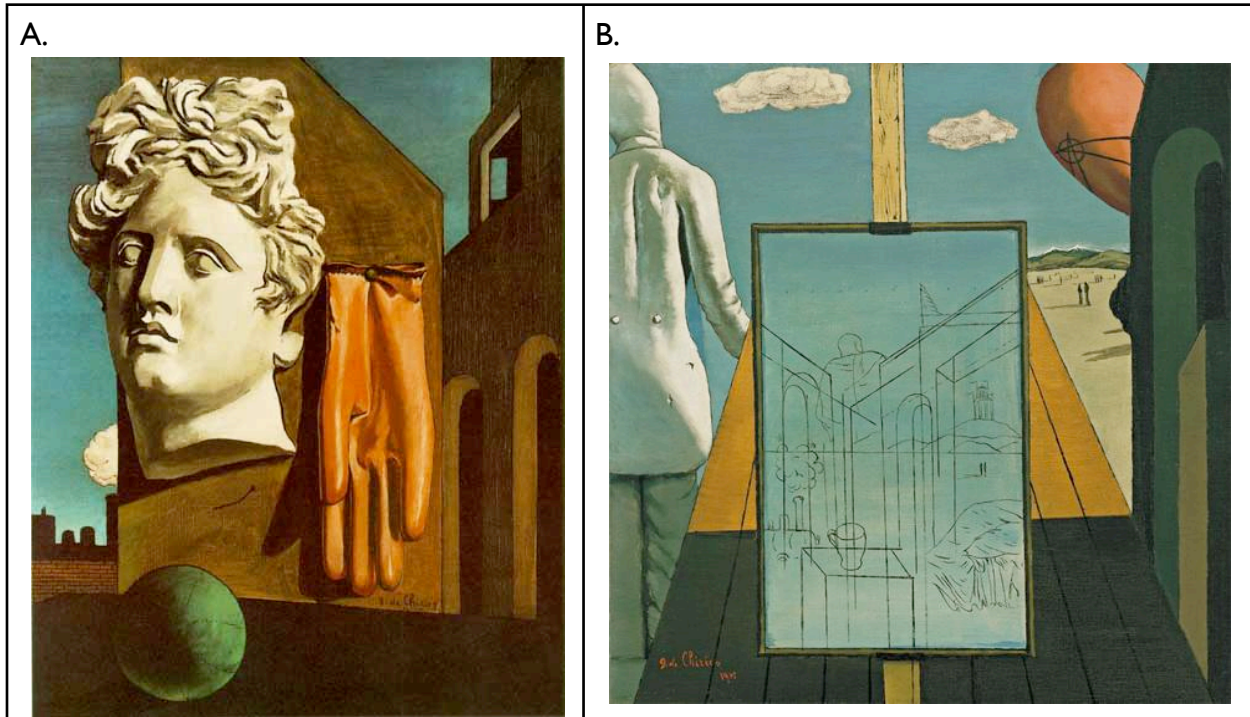


Figure 7.2. Two paintings by Giorgio De Chirico that are highly disfluent in terms of the depicted scenes and compositions as a whole. (A) shows *Love Song* (1910) and (B) shows *The Double Dream of Spring* (1915).

a glove that is two-thirds the size of a building, a gigantic head carved in a classical style, and a green sphere?), but as with Picasso's images, they are often aesthetically pleasing and may even succeed in this specifically because of their disfluency.

There are many potential reasons that people might get aesthetic pleasure from the Picasso and De Chirico paintings. The fact that one can find disfluent images that are aesthetically pleasing demonstrates that people are not *solely* tuned to the fluency of processing the focal object(s). There must be something else. To illustrate what this "something else" might be, I will present one more example.

Figure 7.3 is a striking photographic image by Jean Paul Bourdier from his book, *Bodyscapes*. Look at the photograph now, but do not read the caption below it that contains the title of the photograph. In this photograph, which initially appears to be simply a collection of reddish clay mounds, Bourdier has composed a female body within the frame so that it violates many of the biases mentioned above, including the center and inward biases both horizontally and vertically as well as ecological biases for canonical perspective, orientation, and color.

The photograph as presented in *Bodyscapes* without its title or any explanation. The title is found only in the appendix, and no further comments on the photograph are made). The viewer is thus presented simply with an unusually-composed image, and are left to form his or her own response to it over time. In light of Berlyne's experiments on simplicity and complexity, if we do not dwell on this image for very long, we may not find it pleasing. In fact, many might not even notice the presence of the female form at all. If the viewer spends enough time exploring the photograph, however, he or she is likely to find it quite pleasing aesthetically.

I will argue here and in the next chapter that what happens with this photograph -- and many other images people see -- is that over time, viewers begin to make a meaning for it. It is the fit between this meaning and its composition that is the most important component of our

aesthetic pleasure on viewing it, and for which fluency does not have a credible account. Our account that includes this component is called “representational fit,” since it is the fit between the meaning and the representation of the image that we believe affects people’s preferences.

To return to Bourdier’s photograph, once the female body is noticed and appreciated for how seamlessly it blends into the landscape, a message begins to emerge that people’s bodies are an integral part of the natural world and that even the boundaries between people and their



environment are unclear and permeable.<sup>22</sup> Moreover, once such a meaning is created for this photograph, it becomes clear that a stock photograph of a woman simply standing or lying down on a sandy landscape would fail to convey such a complex and subtle meaning. The consideration of meaning leads to the fourth and largest problem with fluency, one that is intimately related to people’s liking of disfluent images: how can the role of meaning, context, and interpretation in a theory of aesthetic preference be explained by fluency? This issue is the topic of the next chapter.

Figure 7.3. *Nature’s Writing*, Jean Paul Bourdier, (in *Bodyscapes*, 2007).

---

<sup>22</sup> I do not assume that everyone who views this photograph will detect or generate -- even over time -- the precise message I have presented here, but my example is intended as a *possible* example, not the *only* example, for such an image.

## Chapter 8

# The Case For Meaning

As I have interpreted fluency theory thus far (in terms of consistency between aesthetic judgments and recognition performance), it implicitly assumes that the “meaning” of a single-object picture is essentially to depict or portray that object. The aesthetic response to any image, then, should be proportional to its effectiveness in accomplishing this task, making it as easy as possible to recognize the depicted object. Although compositions of this sort are often used in pictures that produce a significantly positive aesthetic response – such as well executed “stock” photographs (see Gardner, Fowlkes, Nothelfer, & Palmer, 2008; www.corbis.com, www.gettyimages.com, or other similar sites) or naturalists’ drawings of plants and animals (e.g., Audubon’s images; Bertamini, Bennett, & Bode, 2010) – they are much less frequent in the kinds of images one sees in art museums and/or high-end galleries.<sup>23</sup> Such pictures are generally expected to have some significant meaning or to evoke some kind of emotional/expressive feelings in the viewer that are seldom achieved in stock photographs, no matter how easily recognized they may be (e.g., *Nature’s Writing* in Figure 7.3 of the previous chapter).

### Fluency’s Problem with Meaning

The main problem I see with a fluency account of aesthetic response arises from the lack of any explicit role for the *meaning*<sup>24</sup> of the image in the standard fluency account. I will illustrate this problem with a discussion of a photograph on the cover of TIME magazine, which falls somewhere between stock photograph and fine art. I will also argue that this problem is related to the central aesthetic principle of novelty (see Berlyne’s research in Chapter 2), which usually requires violating rather than conforming to expectations.

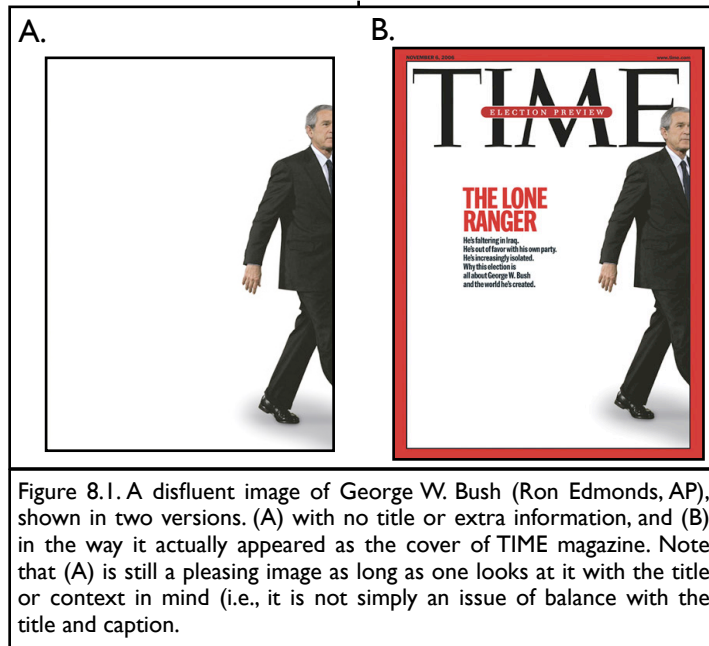


Figure 8.1. A disfluent image of George W. Bush (Ron Edmonds, AP), shown in two versions. (A) with no title or extra information, and (B) in the way it actually appeared as the cover of TIME magazine. Note that (A) is still a pleasing image as long as one looks at it with the title or context in mind (i.e., it is not simply an issue of balance with the title and caption.

### TIME Magazine and *The Lone Ranger*

On November 6, 2006, the cover of TIME magazine displayed a photograph by Ron Edmonds of the Associated Press. In the photograph, edited for the cover, then-President George W. Bush had been isolated from any contextual surroundings and was placed simply on a flat white background, in mid-stride (See Figure 8.1A). With no other objects present anywhere in the picture, he is clearly the focal object in the frame. Nevertheless, the composition does not present to the viewer a fluent image of Bush: he is facing to the right and

<sup>23</sup> This is not to say that aesthetically pleasing images are necessarily good “art” or vice versa. The connection between standard notions of aesthetic response and art has loosened considerably since the advent of Cubism, Abstract Expressionism, and the further development of modern art. Such matters are, however, beyond the scope of this dissertation.

<sup>24</sup> I use “meaning” here to refer to any sort of context or associations over and beyond the physical attributes of the depicted object(s).

positioned so far to the right of the center of the frame that he is not entirely in the frame with his body is cropped down the middle. By the fluency account, this should not be a very aesthetically pleasing image. However, the headline on the cover read “The Lone Ranger,” alluding to the facts that the Republicans had just lost six senate seats in the recent election, and that then-President Bush was seen as being out of step with the general population, too far to the right, and alienating even some members of his own party as he struck out on his own. With this title and political context, the image on the cover was not only aesthetically pleasing, but was also a much more fitting composition than a centrally-cropped, more fluent composition would have been. It is important to demonstrate that in this case, as in many others, the meaning of the image fits best with this novel composition, and this novelty further enhances the viewer’s aesthetic response to the final result, once the justification provided by the meaning has been recognized and interpreted by the viewer.

### The Case For Meaning, Part I: Context in Art

I will now build a “case for meaning” and will attempt to demonstrate not only that the context of an image (or sculpture, as in one of my examples) can indeed affect people’s responses and preferences for it, but also that this role has been well-known and well-used by artists and art historians for many years. As with the philosophical work on aesthetics discussed earlier, I believe that any thorough psychological theory of aesthetic preference ought to build upon this knowledge and include a role for meaning.

### Wind, Warburg, and Gombrich: Context and the Perception of Emotions<sup>25</sup>

Edgar Wind, the first professor of art history at Oxford University, wrote on an eclectic group of works over the course of his career. His books include *Pagan Mysteries of the Renaissance* (1968), *The Eloquence of Symbols: Studies in Humanist Art* (1983), and *Hume and the Heroic Portrait* (1986). He came to be known as the epitome of a “Warburgian scholar” in the United States (Eisler, 1969). Indeed, he was friends with Aby Warburg (the nominal referent of Eisler’s phrase), taught at the Warburg Institute in London, and helped found the *Journal of the Warburg Institute* in 1937. It is in the first issue of the first volume of this publication that he wrote about context in the work of Sir Joshua Reynolds, a portrait painter from the mid- to late 1700s. Wind’s article is subtitled “Comments on an observation by Reynolds.” The observation in question is found in Reynolds’ twelfth discourse on painting, in which he describes how an explicit manipulation of the context of different gestures and poses can dramatically change their emotional interpretation:

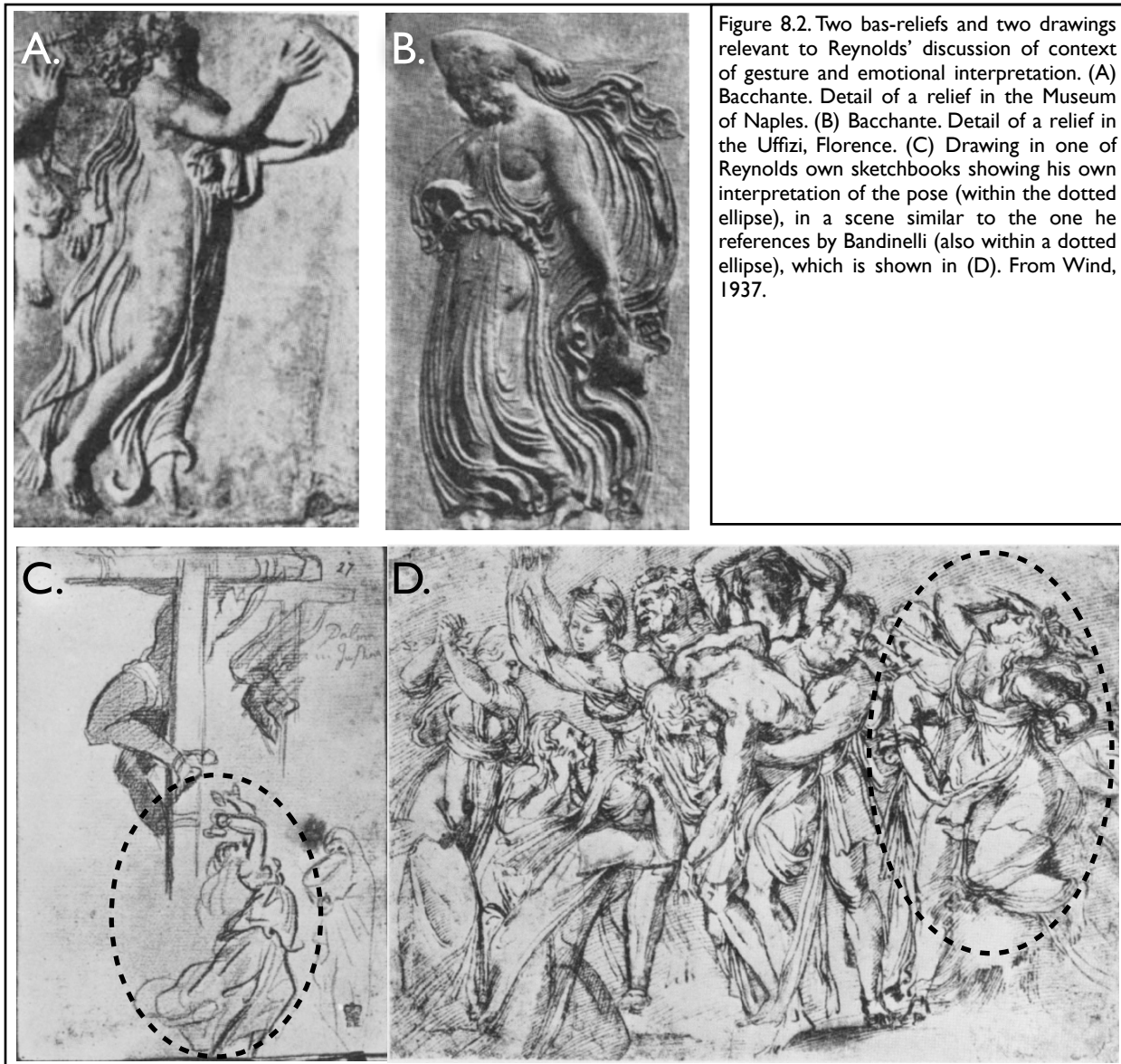
It often happens that hints may be taken and employed in a situation totally different from that in which they were originally employed. There is a figure of a Bacchante leaning backward, her head thrown quite behind her, which seems to be a favourite invention, as it is so frequently repeated in bas-relievos, cameos and intaglios [See Figure 8.2A, 8.2B]; it is intended to express an enthusiastick frantick kind of joy. This figure Baccio Bandinelli, in a drawing that I have of that Master, of the Descent from the Cross [See Figure 8.2C, 8.2D], has adopted, (and he knew very well what was worth borrowing,) for one of the Marys, to express

---

<sup>25</sup> Thanks are due to John Kihlstrom for bringing my attention to this particularly rich example of context and for pointing me to the appropriate sources for researching it.

frantick agony of grief. It is curious to observe, and it is certainly true, that the extremes of contrary passions are with vary little variation expressed by the same action. (Reynolds, 1798).

Wind calls this insight of Reynolds’ “the shrewdest advice given by Sir Joshua Reynolds to his students,” and says that with this rule, “Reynolds hit, apparently *en passant*, upon a fundamental law of human expression...” (Wind, 1937). Wind traces some of the relevant classical examples of the sort that Reynolds was referring to and points out that the same pose used to illustrate a “Bacchic frenzy” was “transposed by the early mannerist artists into an expression of religious ecstasy” (Wind, 1937). Thus, different circumstances or context, such as a scene of Christ’s burial versus a scene of pagan festivities, can completely alter our interpretation of near-identical -- if not identical -- physical poses.<sup>26</sup> According to Wind, Aby Warburg “without knowing of this



<sup>26</sup> This is an anticipation of the Schachter and Singer theory of emotion, which states that the same physiological states can be interpreted by people as different emotions depending on the circumstances in which they occur. (Schachter & Singer, 1962).

passage in Reynolds' Discourses, or of the drawing in Reynolds' sketch book, collected material which tended to show that similar gestures can assume opposite meanings" (Wind, 1937).

Ernst H. Gombrich, an art historian and critic who is likely second only to Rudolf Arnheim in popularizing the arts and psychological interpretations of them, revisited these findings in his "intellectual biography" of Aby Warburg (Gombrich, 1970). Gombrich writes of Warburg's research on the "history of classical 'engrams'" that "[the artist] can use ['engrams'] in a different context, 'invert' their original savage meaning, and yet benefit from their value as expressive formulae" (Gombrich, 1970). In reference to the larger significance of this scholarship, Gombrich notes:

Whatever the validity of Warburg's individual examples...they appeared to prove that, while the energy of past experience remains enshrined in the 'engrams' or symbols, this energy may be canalized into different themes of expression...The symbol... is a charge of latent energy...only through contact with the 'selective will' of an age does it become 'polarized' into one of the interpretations of which it is potentially capable. (Gombrich, 1970).

This discussion builds on Reynolds' (and Wind's) original point by hinting at the fact that even the same work of art could have different interpretations, if viewed at different times. This resonates with Berlyne's work on novelty and habituation as well as with many people's personal experience that truly great works of art produce new and different revelations on repeated viewings.

### Messerschmidt's Character Heads: Interpretations and Reactions

Another powerful example of context and its effect on perception of art lies in the "character heads" of Franz Xaver Messerschmidt (1736-1783). I will illustrate the role of context in our responses to these heads by tracing their history, and how we might likely see

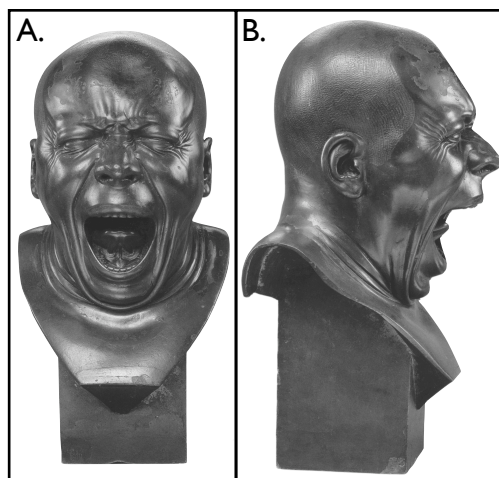


Figure 8.3. Two views of one of Messerschmidt's heads, sculpted between 1770 and his death in 1783. While it anatomically, anecdotally, and intuitively (from our own perception of it) depicts pain, it was titled *The Yawner* by an anonymous source in 1794, and that is the title it has carried ever since. From Bückling, 2006.

them and learn about them today, using one of his heads as an example. Towards the end of Messerschmidt's life, he had psychological problems, lost his job at the university where he worked, and contracted an incredibly painful disease, although current art historians and scholars still disagree about what disease in particular it was. As the story goes, in order to distract himself from these varieties of pain, he took to sculpting "character heads" of people with highly contorted, unpleasant expressions that he rendered in very realistic detail, often using his own face as a model (Pötzl-Malikova, 2010). At the time of his death in 1783, Messerschmidt had sculpted 49 heads, and not a single one of them had been given a title by its creator. It was not until eleven years later, when an anonymous writer was documenting Messerschmidt's work for an exhibition guide, "apparently geared to stimulate public interest in order to foster sales," that the heads had been distinguished by titles (Anonymous, 1794; Yonan,



2009). Because apocryphal titles came from a source other than the artist, Pötzl-Malikova (2010) argues that they can mislead our perceptions from the original intentions of Messerschmidt so seriously that they should be discarded. Willibald Sauerländer, an art historian, states bluntly that:

The trivial titles assigned to individual head-pieces by an anonymous writer ten years after Messerschmidt's death -- *Beak Head, The Furious and Vengeful Gypsy, The Incompetent Bassoonist* -- are nothing but an attempt to resist their social illegibility. (Sauerländer, 2010a)<sup>27</sup>

Similar sentiments are echoed in a recent exhibition catalogue, which, in reference to *The Yawner* (see Figure 8.3), says that while we might be inclined to view the head as yawning based on its title, there is nothing about the sculpture itself that “might be compelling proof the present head is yawning, especially since the tongue is drawn up so high, which does not occur in yawning” (Bückling, 2006).

## The Case For Meaning, Part 2: Context in Aesthetic Preference

There is little psychological research on the role of context on aesthetic preference. Most research on the effects of perceptual context has not made any explicit mention of aesthetic preference, though it does focus on comprehension (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982; Bower, Karlin, & Dueck, 1975; Bransford, 1973; Palmer, 1975), which can be linked to aesthetic preference through fluency theory, as argued above. However, there has been one set of experiments, especially germane to the research to be reported in the next chapter, which used titles to manipulate the context for aesthetic preference judgments.

Millis (2001) studied the influence of three different title categories: no title; descriptive titles that were “redundant with what can be readily observed by the viewer;” and metaphorical titles, that “provided an explanation or a metaphorical interpretation of the scene.” Aesthetic experiences were measured by combining scores on scales of interest, emotion, enjoyment, and number of thoughts about each image. Millis found that people had greater aesthetic experiences with the metaphorical titles than with either descriptive titles or no title at all, which he called the “elaboration effect”. Further, he found that “random” titles (metaphorical titles originally created for a different image) did not lessen aesthetic experiences compared to the descriptive or untitled conditions and that “titles increased aesthetic experiences only when they contributed to rich and coherent representations.” Additionally, Millis found that the elaboration effect occurred whether or not the participants were told that titles given to the images were true or false. In the false condition, participants were told that the experimenters had generated the titles and that they were not the “real” titles of the artworks, a manipulation that has significance in light of the Messerschmidt heads described above.

Clearly, the context provided with an image can have a powerful effect on people's aesthetic response to that image. With an eye toward testing this context and its effect on preference for different compositions, I will now describe the third and final set of experiments.

---

<sup>27</sup> Glenn F. Bengé points out, in a response to Sauerländer, that a manual of engravings of emotional expressions by Charles Le Brun in 1698 -- meant to be used as a teaching tool, and as an assessment of students' skill -- could shed some additional light on these heads. He theorizes that unawareness of such techniques and pedagogy could be what “led the Viennese artistic and social circles to see Messerschmidt's expressive heads as lunatic rather than virtuosic” (Benge, 2010). Sauerländer replied that Messerschmidt's heads are likely “not normative examples for representing the passions but emotional explosions in a private environment” (Sauerländer, 2010b).

## Chapter 9

# Representational Fit and the Role of Context

In the remainder of this dissertation I analyze the aesthetics of pictures in terms of the notion of “representational fit”<sup>28</sup>. The core hypothesis of representational fit, at least with respect to spatial composition, is that an image is aesthetically pleasing to the extent that the placement of objects within the frame is consistent with spatially relevant aspects of the meaning of the image. The four experiments described in this chapter test the representational fit hypothesis by manipulating meaning through assigning different titles to the same images. The default case of the meaning of an image is simply to portray the object depicted so that it is as easy to recognize as possible (see Chapter 5 and 6), as was presumably the case when no titles were given in first two sets of experiments. When a title is given explicitly, however, it is usually taken as an indication of the essential meaning the image-maker intends to convey. The results of the first two sets of experiments (Chapter 5 and 6) do not discriminate accounts in terms of fluency versus representational fit because they did not involve any non-default meanings. The following four experiments incorporate neutral (default) titles versus nonstandard titles for images, thus enabling such distinctions to be made. The two primary goals of the following experiments are to find out (a) whether preference for the same spatial composition would change as a function of the title and (b) whether spatial compositions that are normally judged as aesthetically unpleasing would be seen as more aesthetically pleasing when presented with a title that was explicitly designed to fit the nonstandard composition. The results provide considerable support for the predictions of representational fit.

### Experiment 3.1: Representational Fit between Titles and Object Positions

Images with a single focal object positioned at one of five horizontal locations within the frame (see Figure 9.1) were created along with three titles for each image. One title was a purely descriptive (neutral) title that simply described the content of the image. Because the neutral title essentially represents the default meaning that the picture would have with no title at all, viewers’ preferences with this title should conform to the prior findings when no titles were included (Chapter 5). The other two titles contained what Millis (2001) would call “elaborative” information and were explicitly designed to shift the preference function for the horizontal location of the focal object to one side of the frame or the other.

The elaborative titles fell into two categories: titles that implied unseen objects and titles that utilized spatiotemporal metaphors. The images and titles for the “unseen objects” category are shown in Figure 9.1A. The research described in Chapter 5 showed that an untitled image of, say, a running racehorse would be preferred slightly off-center in the direction opposite its facing (e.g., left of center if the horse is facing/running rightward). The same would presumably be true when it is given a purely descriptive, neutral title, such as “Racehorse.” However, when the same images of a racehorse are paired with the alternative titles that imply the existence of unseen objects behind or in front of the racehorse, the preference function should shift correspondingly to show the space between the seen racehorse and the unseen others. The title “Front Runner,” for example, implies the presence of a group of unseen horses behind the

---

<sup>28</sup> “Fit” alone, might be the more general and appropriate term, since “representational fit” is presumably only relevant to visual objects that qualify as representations, such as photographs or representational paintings. With that qualification, however, I will use representational fit here.

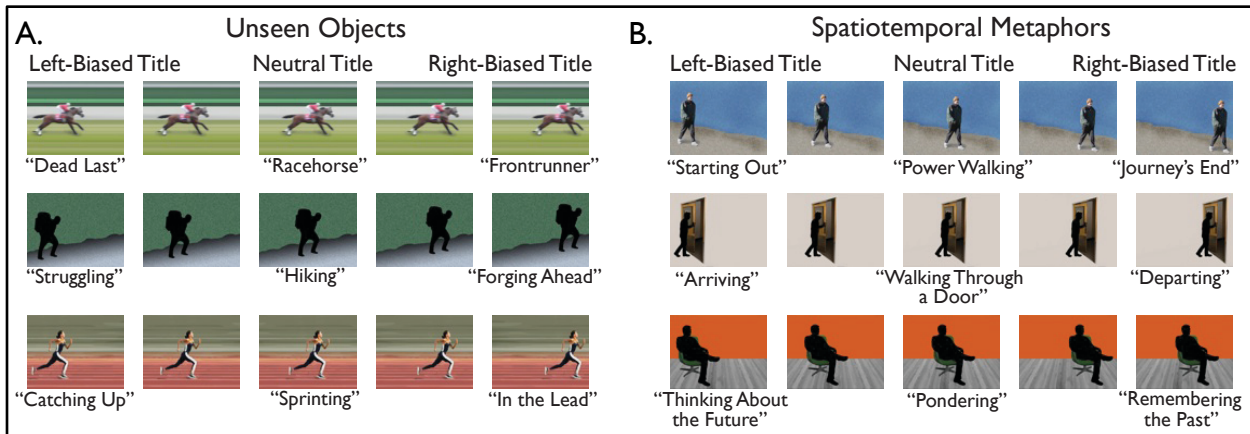


Figure 9.1. The images used in Experiment 1.1, grouped by the category of the title (A,B). The neutral title for each sequence is placed above the center image, and the left- and right-biased titles are placed below the left and rightmost images, respectively.

depicted one and outside the frame. If the fit between the title's spatial implications and the spatial structure of the image matters, viewers should prefer the depicted horse to be off-center toward its direction of motion to emphasize the space behind the depicted horse and to balance the unseen horses behind it, with the horse effectively running out of the frame. In contrast, the title "Dead Last" implies unseen out-of-frame horses that are in front of the depicted horse, in which case the horse should be preferred to be on the side of the frame opposite its facing/running direction. A similar analysis applies to the picture of the runner. The hiker on an upward slope, titled "Struggling," or "Forging Ahead," was intended to imply the existence of a force that is metaphorically akin to an unseen object. That is, in the "struggling" case, the force is in front of the hiker, keeping him from progressing, but in the "forging ahead" case, the hiker is ahead of the force.

The images shown in Figure 9.1B were titled using spatiotemporal metaphors. George Lakoff and Mark Johnson (1980) have argued that people have a strong bias to treat time spatially, as if they are moving through time in a forward direction. At least for English speakers, the future is typically conceived as being in front of them and the past as extending behind them, and time is typically conceived as moving from left to right (Boroditsky, 2000; Lakoff & Johnson, 1980). By titling an image of a man walking "Starting Out" or "Journey's End," the implied position of the man relative to the spatial extent of his journey is shifted and so, therefore, should the preference function the best-fitting spatial composition be shifted correspondingly. When starting out, the bulk of the implied journey is in front of the walker, and so he should be shifted away from center in the direction opposite his motion, so that he is clearly walking into the frame. At the end of the journey, he should be shifted away in the direction toward his motion, so that he is clearly walking out of the frame. Similar considerations motivated the "Arriving" and "Departing" titles for the person walking through a door. A metaphorical extension was included to mental events in the "Thinking about the Future" and "Remembering the Past" titles, where no physical motion was implied or depicted.

Figure 9.1 shows all six objects composed at all five horizontally distributed positions within the frame. These pictures were combined with each of the three title types (neutral, left-biased, or right-biased) that are shown in Figure 9.1 under the composition expected to be the most preferred with that title. Mirror-image versions, in which motion was depicted from left to right, were not studied because it would conflict with the standard interpretation of time going from left to right.

Before preference for titled images could be assessed, the title-image pairs needed to be evaluated to ensure that there was a title that indeed fit better with a specific spatial composition than the others and that there was enough difference between these fit judgments to provide a reliable test of the representational fit hypothesis on preference ratings in Experiment 3.2. In the present experiment, each participant was presented with two alternative titles for the same image and was asked to indicate which title they felt was more appropriate to the image. The results of this procedure, averaged over all pairwise combinations of titled images of the same focal object, provided a measure of the relative goodness-of-fit between the three titles and the images showing each object at each position.

### Method

**Participants.** The 12 participants were students at the University of California, Berkeley who received partial course credit in their undergraduate psychology course for their participation. Their mean age was 19.4 years. All participants were naïve to the purpose and nature of the experiment, and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

**Design.** There were three titles for each object: a title that would fit best with it being on the left side of the frame, one that would fit best with it being on the right side of the frame, and a neutral title that simply described the focal object. The center of the focal object was located at five equally-spaced positions (20%, 35%, 50%, 65%, and 80% of the frame from the left edge) within each source image (see Figure 9.1) and always faced to the right.<sup>29</sup> There were 180 trials, resulting from the permutations of three titles presented two at a time for each of the six objects at each of the five positions ( $3 \times 2 \times 6 \times 5 = 180$ ) (see Figure 9.2). This design includes left/right counterbalancing of the positions of the two titles relative to the picture. The order and presentation of the trials was randomized by Presentation software ([www.neurobs.com](http://www.neurobs.com)).

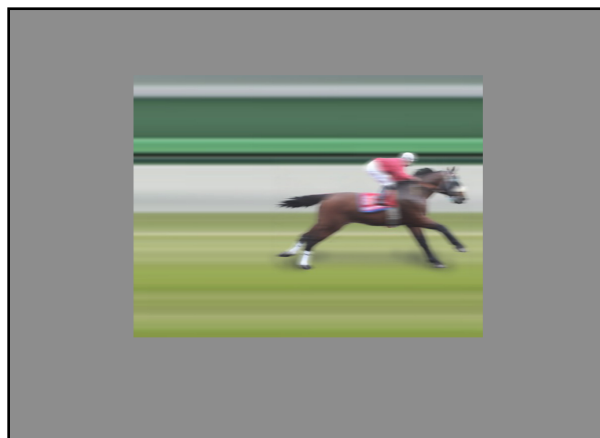


Figure 9.2. A schematic of the trial screen for Experiment 1.1. Note that this is a schematic and is not to scale.

**Displays.** The background of the screen was neutral gray. Each image measured 450 x 600 pixels and was centered on the screen. The two possible titles for each image were black text placed within a white rectangle at the bottom of the screen that ran its entire width (see Figure 9.2, which shows the racehorse image only for purposes of illustration). The images, frames, and titles were created in Adobe Photoshop CS3. The display measured 20" diagonally, and had a resolution of 1440 x 900 pixels.

**Procedure.** Participants viewed the screen from approximately 60cm. They were instructed to press either the left or right mouse button to

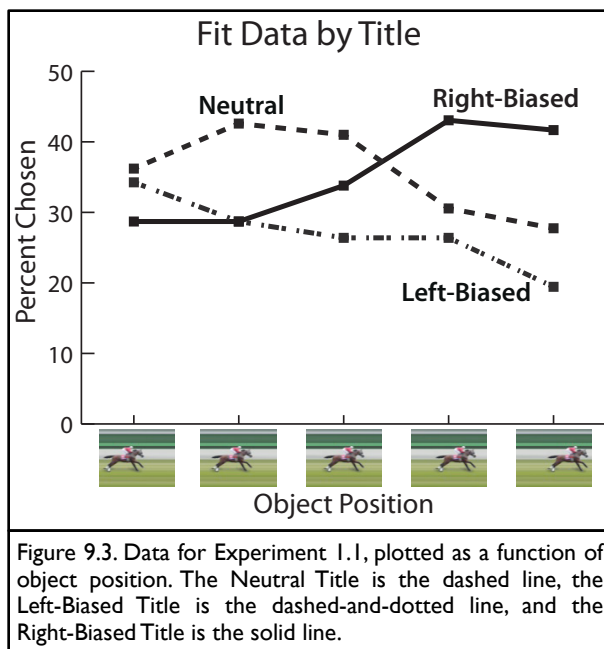
<sup>29</sup> The focal object in each image faced to the right in every case. Counterbalancing for object facing would have doubled the number of trials, which could prove problematic especially in Experiment 3.2. Furthermore, it would have introduced a conflict between the reading direction of the participants -- the same direction as they perceive time as moving (Lakoff & Johnson, 1980; Boroditsky, 2000) -- and the direction of implied time passage in the left-facing spatiotemporal metaphor images. The effects obtained in the experiments described here would likely be present in a study of left-facing images, but combined with opposite effects from reading direction and time passage, which would likely produce less pronounced or noisier effects.

indicate which title they thought “fit best” with the displayed image. They proceeded through trials at their own pace.

### Results and Discussion

Results were analyzed in terms of the percentages of trials on which each given title was chosen as fitting best with each particular object position. The data are plotted in Figure 9.3 as a function of position and title, averaged over all six different objects. There was no main effect of title category (spatiotemporal metaphor or unseen object), and none of its interactions with other factors was significant (all  $p$ 's > .5). The data were therefore averaged over that factor in subsequent analyses. Since there were equal numbers of comparisons done for each object at each of the five positions, there could not be any main effect of object, nor any interaction between object and position. The three-way interaction between object, title, and position was significant, but not large ( $F(40,440) = 1.42, p < .05$ ), indicating that the pattern of the two-way title  $\times$  position interaction differed only slightly across images of different objects. There was no significant interaction between title and object, however ( $F(10,110) = 1.393, p > .15$ ). The data plotted in Figure 9.3 have therefore been averaged over object types as well as title categories.

The effect of primary interest is the interaction between title and position ( $F(8,88) = 4.67, p < .001$ ) plotted in Figure 9.3. The “left-biased” titles fit better with images in which the focal object was at the left two positions than the right two positions, ( $F(1,11) = 6.248, p < .03$ ), the “right-biased” titles fit better when the object was at the right two positions than the left two positions ( $F(1,11) = 17.445, p < .01$ ), and the neutral titles fit better when the object was at the middle position or one position away in the direction consistent with the inward bias found in previous studies than the position one to the right and one to the left of those two positions ( $F(1,11) = 6.488, p < .03$ ). The leftmost, middle, and rightmost images of each object were chosen for further analyses to provide the clearest spatial differences in Experiment 3.2. Further specific comparisons indicated that these three positions fit best with the correspondingly biased title (averaged over objects): for the left-biased title, the leftmost position fit better than the middle position ( $F(1,11) = 4.901, p < .05$ ) and the rightmost position ( $F(1,11) = 11.932, p < .01$ ), for the



neutral title, the middle position fit better than the leftmost position ( $F(1,11) = 4.845, p < .05$ ) and the rightmost position ( $F(1,11) = 6.132, p < .05$ ), and for the right-biased title, the rightmost title fit better than the leftmost position ( $F(1,11) = 10.170, p < .01$ ) and the middle position ( $F(1,11) = 5.416, p < .05$ ). These three positions were paired with the three possible titles for use as the image-title pairs in Experiment 3.2.

The main effect of title was unexpectedly significant ( $F(2,22) = 3.93, p < .05$ ). The pattern apparent in Figure 9.3 indicates that viewers generally thought that the left-biased title fit less well with all images compared to the other two title groups (neutral title compared to left-biased title:  $F(1,11) = 5.432, p < .05$ ; neutral title compared to right biased title ( $F(1,11) < 1$ ; left

title compared to right title:  $F(1,11) = 5.061, p < .05$ ). This difference may have arisen from some overall preference for the right-biased and neutral-biased titles over the left-biased ones because of the negative semantic connotations of some of the left-biased titles (e.g., “Dead Last,” “Struggling,” and “Catching Up”). This possibility is examined empirically in Experiment 3.2.

These results demonstrate that the titles did indeed fit the images in the expected way. This set of titles and images, then, can be used to test the representational fit hypothesis that a titled image will be most preferred when the image’s spatial composition fits the elaborative title well (i.e., left-biased titles with left-balanced compositions and right-biased titles with right-balanced compositions) and least preferred when the image’s spatial composition fits the title poorly (i.e., left-biased titles with right-balanced compositions and right-biased titles with left-balanced compositions). The images with neutral titles were expected to replicate the previous findings for images without any title (Chapter 5).

### Experiment 3.2: Representational Fit and Preference for Position

Experiment 3.1 demonstrated that there are varying degrees of fit between different titles and different spatial compositions within a rectangular frame. In Experiment 3.2 the relation between goodness-of-fit and preference is evaluated using a 2AFC paradigm in which observers picked which of two image-title pairs they preferred aesthetically. The image-title pairs were created by combining the three titles for each object studied in Experiment 3.1 with the three compositions that fit those titles well and were most clearly different from each other, with the focal object positioned on the left side, the middle, or the right side of the frame. The representational fit theory predicts (a) that image-title pairs that fit better will be preferred to those that fit worse and (b) that the most preferred “elaborative” image-title pairs (the left-biased and right-biased ones) will be preferred to the most preferred “default” image-title pairs (the neutral ones). The latter prediction is commensurate with the findings of Millis (2001), common practice in artistic instruction, and the literature on novelty discussed in Chapters 2 and 7. It is not necessarily consistent with a perceptual fluency theory based on the ease of visual processing.

#### Method

Participants. The 12 participants were students at the University of California, Berkeley who received partial course credit in their undergraduate psychology course for their participation. Their mean age was 19 years. All participants were naïve to the purpose and nature of the experiment, and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. There were nine image-title pairs (three image compositions combined with three titles) for each of six objects, consisting of all possible combinations of the three titles for each object combined with the best-fitting position for each of the three titles for that object, as measured in Experiment 3.1. These nine image-title pairs for each object were shown in all possible pairwise comparisons, counterbalanced for left/right position of each image-title pair ( $9 * 8 = 72$  2AFC trials/object for each of six objects, resulting in 432 trials). The order and presentation of the trials was randomized by Presentation software ([www.neurobs.com](http://www.neurobs.com)).

Displays. The background of the screen was neutral gray. Each image measured 450 x 600 pixels, and had a 40-pixel black frame placed around it that had a beveled, three-dimensional appearance, to give the impression of a framed work. The two images were placed side-by-side

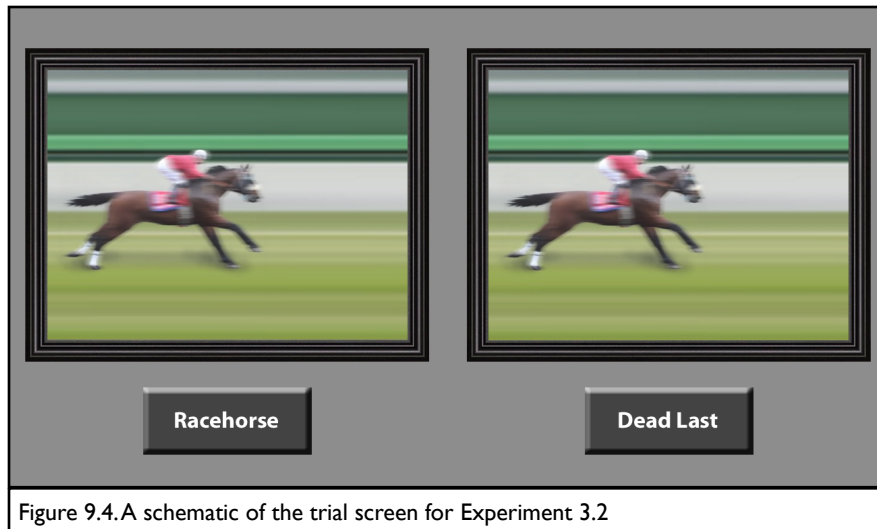


Figure 9.4. A schematic of the trial screen for Experiment 3.2

at the center of the screen. The title for each image was white text placed on a beveled black square that was centered horizontally below each image and was separated from the image by 100 pixels (See Figure 9.4). The images, frames, and titles were created in Adobe Photoshop CS3. As in Experiment 1, the display measured 20" diagonally and had a resolution of 1440 x 900 pixels.

**Procedure.** Participants viewed the screen from a distance of approximately 60cm. They were told that they would see two titled images on each screen and that their task was to pick the image-title pair that they liked better by using the left and right arrow keys on the keyboard.

### Results and Discussion

As in Experiment 3.1, participants' responses were scored for the probability that they chose each particular image-title pair. These probabilities, averaged over objects, are plotted in Figure 9.5A. The title x position interaction was highly significant ( $F(4,40) = 30.067$ ,  $p < .001$ ), and there was a smaller 3-way interaction among image, title, and position ( $F(20,200) = 4.430$ ,  $p < .001$ ), as well as a significant interaction between title and image ( $F(10,100) = 6.596$ ,  $p < .001$ ). There was also a main effect of title ( $F(2,20) = 13.469$ ,  $p < .001$ ). The left-biased title was chosen less than both the neutral title ( $F(1,10) = 27.804$ ,  $p < .001$ ) and the right-biased title ( $F(1,10) = 29.830$ ,  $p < .001$ ). The right-biased title and the neutral title were preferred equally and had identical means.

The main effect of title raised the question of whether participants' preferences might be influenced by how much they like the title on its own (e.g., they might well prefer the title "Front Runner" to the title "Dead Last" even if no image was present). To assess this possibility, a different group of 22 participants were asked to rank-order the three titles for the same object from most to least preferred. The content of the image was minimally described, but no image was shown: e.g., "Rank how much you like the following three titles for an image of a racehorse: 'Dead Last,' 'Racehorse,' and 'Frontrunner.'" The main effect of kind of title (left-biased, right-biased, or neutral) was trending towards significance ( $F(2,42) = 2.445$ ,  $p < .10$ ). The mean for the left-biased titles was 1.84, the mean for the neutral titles was 2.114, and the mean for the right-biased titles was 2.05. The neutral title was almost significantly preferred to the left-biased title ( $F(1,21) = 4.279$ ,  $p < .06$ ). The left-biased title and the right-biased title were preferred equally ( $F(1,21) = 2.072$ ,  $p < .20$ ), as were the neutral title and the right-biased title ( $F(1,21) < 1$ ). There was a significant interaction between kind of title and object ( $F(10,210) = 4.917$ ,  $p < .001$ ). For the images of the woman "sprinting" and the racehorse, the left-biased titles were less preferred than the other two, likely because of the negativity of being behind in a race and trying to catch up, or of placing "dead last." For the man "power walking," the man

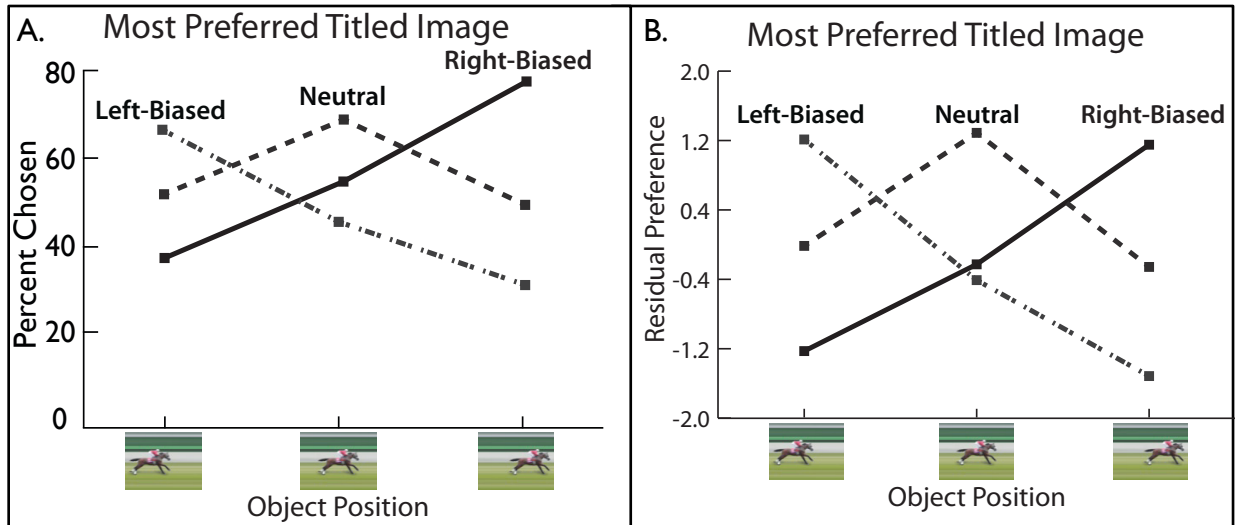


Figure 9.5. Data for Experiment 3.2, plotted as a function of object position. (A) shows the raw data, and (B) shows the data with the variance due to preference for titles alone removed. The neutral title is the dashed line, the left-biased title is the dashed-and-dotted line, and the right-biased title is the solid line. Data are averaged over image; the racehorse is simply to represent the object position. Note that the less-preferred endpoints of the neutral title line do not dip down as far as those for the left-biased or right-biased titles.

“hiking,” the man sitting in a chair “pondering,” and the man “walking through a door,” the three titles were preferred about equally.

The fit data between the title and image and the title-alone preference data from Experiment 3.1 were then used as two predictor variables in a regression analysis to predict the preference data from Experiment 3.2. When both factors were entered in a stepwise fashion, fit between title and image was entered first and explained 27.5% of the variance ( $F(1,52) = 19.755, p < .001$ ). Preferences for the titles was entered second and accounted for an additional 15.2% of the variance ( $F(2,51) = 19.025, p < .001$ ). This model thus accounted for a total of 42.7% of the variance.

After the variance accounted for by preference for the titles alone was removed, the data were plotted again as a function of title and object position, averaged over objects (see Figure 9.5B). The best image-title pair with the left-balanced title, the right-balanced title, and the neutral title did not differ significantly ( $F(2,20) < 1$ ), but the large title  $\times$  position interaction is even clearer. Image compositions that were least preferred with one title became most preferred with another title. The best-fit left-biased titled image, the best-fit right-biased titled image, and the best-fit neutrally-titled image were seen as no different from each other (best-fit left-biased vs. best-fit right-biased:  $F(1,10) = 2.845, p > .10$ ; best-fit left-biased vs. best-fit neutral:  $F(1,10) < 1$ ; best-fit right-biased vs. best-fit neutral:  $F(1,10) < 1$ ).

Even with the variance due to preference for titles alone removed the main effect of title was significant ( $F(2,20) = 6.648, p < .01$ ) in the preference data from Experiment 3.2. (see Figure 9.5B), as a result of the nature of the interaction between titles and positions. It is clear from Figures 9.5A and 9.5B that the neutral title was most-preferred for the default (middle) position and that preference for the neutral titled images was never as low as the lowest preference for left-biased and right-biased titled images. This effect presumably occurs because of constraints on the nature of the frame and the distance of each position from the best-fitting position for a title. For the left- and right-biased titled images, the compositions in which the object is two positions away from the best position receive the lowest ratings. For the neutral-titled images, the worst depicted positions are only one position away from the best position. This is



supported by the fact that there are no significant differences among all four positions that are one position away from the best position, two for the neutral title, and one each for the left- and right-biased titles ( $F(3,30) < 1$ ).

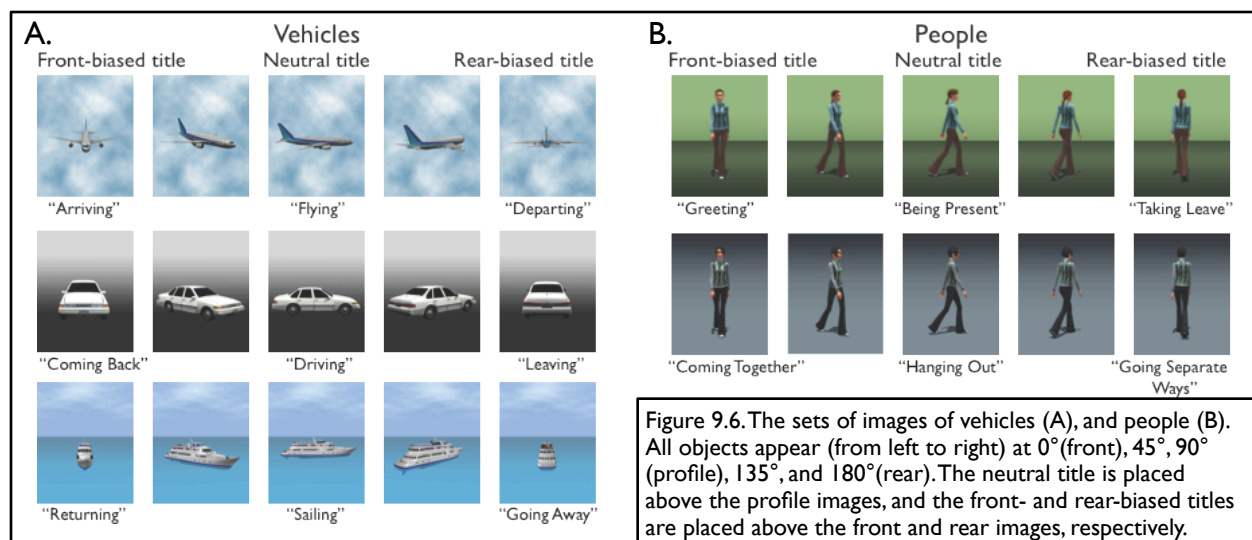
These results demonstrate conclusively that aesthetic preference for spatial composition is not absolute but depends on contextual meaning. With one title, a certain spatial composition can be most preferred and another least preferred, whereas with a different title, the opposite pattern can hold. This supports the claim that fluency -- the idea that people prefer images that are easier to perceive -- is only one piece of the puzzle. Representational fit provides a better explanation because it contains the right kind of structure to accommodate such contextual interactions. Indeed, fluency can be considered a special case of representational fit in which the context of the image dictates that the image will be highly preferred when its focal object is easily recognized.

### Experiment 3.3: Representational Fit between Titles and Perspectives

As mentioned in Chapter 7, previous research on canonical perspective by Palmer, Rosch, and Chase (1981) showed that many objects are recognized more quickly and accurately from some perspective views than others. Khalil and McBeath (2006) recently extended this work to aesthetic response by showing people canonical and non-canonical views of objects and finding that canonical perspectives were rated as more aesthetically pleasing than non-canonical perspectives.

Greater aesthetic response to canonical perspectives constitutes another example of a default case in which people prefer those views in which the object is most recognizable. By analogy with the positional effects due to representational fit in Experiments 3.1 and 3.2, similar differences in aesthetic preference should be induced for canonical versus non-canonical perspectives of objects by giving them titles that imply seeing them from the front versus behind the object. Five sets of images – three of vehicles (Figure 9.6A) and two of people (Figure 9.6B) – were generated from different viewpoints so that they appeared at five different perspectives relative to the camera: 0° (head on), 45°, 90° (profile), 135°, and 180° (facing directly away).

The task was analogous to that of Experiment 3.1. Participants were presented with a single image at a given perspective and were asked to decide which of two titles was more appropriate for it. For each image, three titles were generated: one purely descriptive title for the default



case, one elaborative title that biased a front view of the object, and one elaborative title that biased a rear view. The titles were all variants of “coming” versus “going” events, such as “Arrival” and “Departure” for a flying plane (Figure 9.6A) and “Greeting” versus “Taking Leave” for a walking person (see Figure 9.6B). The fit judgments for the neutral titles should conform to canonical perspective data (usually a side view of the vehicles and a front-side view of the people) and the elaborative titles should shift the peaks in the preference functions toward the less canonical views that were appropriate for the meaning of the non-standard title.

Participants. The 17 participants were students at the University of California, Berkeley who received partial course credit in their undergraduate psychology course for their participation. Their mean age was 20.3 years. All participants were naïve to the purpose and nature of the experiment, and gave informed consent in accord with with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

Design. As in Experiment 3.1, there were 180 trials, resulting from three titles presented two at a time for each of five perspective views for each of six objects ( $3 \times 2 \times 5 \times 6 = 180$ ). Again, the order and presentation of the trials was randomized by Presentation software (www.neurobs.com).

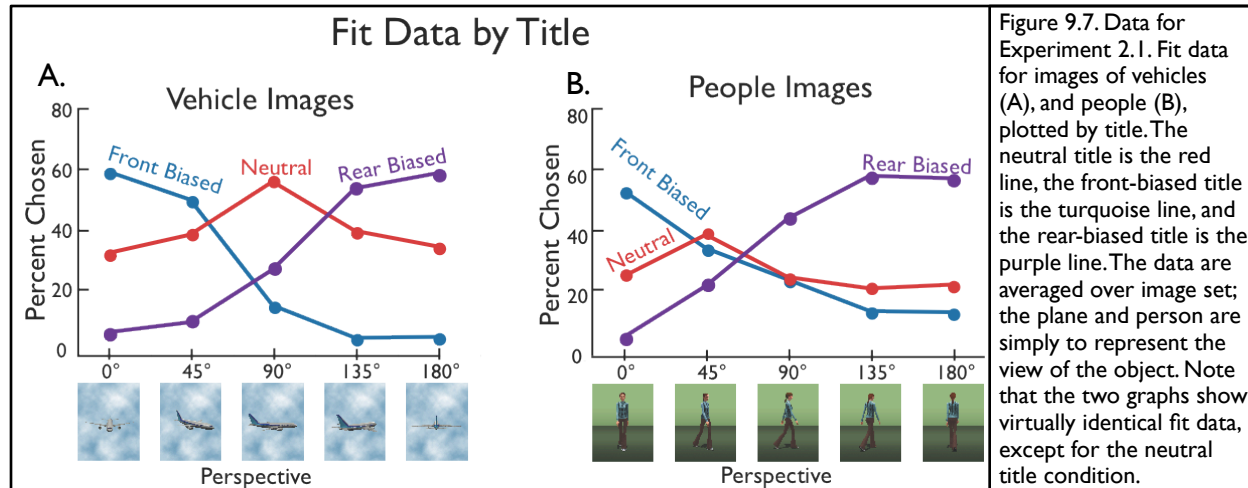
Displays. The displays were in the same format as those in Experiment 1, except that the images measured 526 x 667 pixels.

Procedure. The procedure used the same equipment, software, and instructions as in Experiment 3.1, except that the images were presented at a slightly higher 1620 x 1080 pixel resolution.

## Results and Discussion

As in Experiment 3.1, the data were analyzed in terms of the percentage of trials on which a given title was chosen as fitting best with each particular view of the appropriate object. The key interaction between title and view was highly significant ( $F(8,152) = 100.226, p < .001$ ), showing that participants did choose different images as fitting better or worse with different titles. The object x title x view interaction was also significant ( $F(32,608) = 5.718, p < .001$ ), as was the object x title interaction ( $F(8,152) = 11.449, p < .001$ ). These two interactions were likely driven by the difference between the neutral title images of people versus vehicles (see below).

The front-biased titles fit significantly better with the 0° and 45° views than with the 135° and 180° views ( $F(1,19) = 188.836, p < .001$ ). For rear-biased titles the 135° and 180° views fit significantly better than the 0° and 45° views ( $F(1,19) = 225.141, p < .001$ ). The pattern of fit judgments for the images of the two people did not differ from each other ( $F(2,38) = 1.409, p > .20$ ), nor did it differ for the images of the three vehicles ( $F(4,76) = 1.425, p > .20$ ), but the pattern did differ between the vehicles (Figure 9.7A) and the people (Figure 9.7B) ( $F(2,38) = 8.950, p < .01$ ). For images of vehicles with neutral titles, participants chose the pure side views (90°) as fitting best (90° vs. 45°:  $F(1,19) = 43.623, p < .001$ ; 90° vs. 135°:  $F(1,19) = 34.110, p < .001$ ), but for the images of people with neutral titles, they chose the front-side view (45°) as fitting best (45° vs. 0°:  $F(1,19) = 8.987, p < .01$ ; 45° vs. 90°:  $F(1,19) = 4.410, p < .05$ ). For this reason, the fit data for the images of people (see Figure 9.7A) have been plotted separately from the data for the images of vehicles (see Figure 9.7B). Because of this difference, the 45° views were chosen as the best-fitting images for the neutral titles with people in Experiment 3.4, and the 90° views were chosen as the best-fitting images for the neutral titles with vehicles. For all objects, the 0° and 180° views were chosen as the best-fitting images for the front-biased and



back-biased titles, respectively. Statistical comparisons indicated that these views fit best with the correspondingly biased title: For the front-biased title, the 0° view fit better than the 45° view ( $F(1,19) = 56.296, p < .001$ ), the 90° view ( $F(1,19) = 127.577, p < .001$ ), and the 180° view ( $F(1,19) = 190.382, p < .001$ ); for the rear-biased title, the 180° view fit better than the 0° view ( $F(1,19) = 211.196, p < .001$ ), the 45° view ( $F(1,19) = 129.353, p < .001$ ), and the 90° view ( $F(1,19) = 65.272, p < .001$ ).

These results demonstrate that the titles indeed fit reliably better with some images than others. The representational fit hypothesis is that a titled image will be most preferred when the depicted view of the focal object fits the elaborative title well (i.e., front-biased titles with front views of objects and rear-biased titles with rear views of objects) and least preferred when the image's view of the focal object fits the title poorly. These predictions are tested in Experiment 3.4.

### Experiment 3.4: Representational Fit and Preference for Perspectives

Analogous to Experiment 3.2, the present experiment was designed to determine whether people's aesthetic preferences varied according to the goodness-of-fit between the images and the titles when different perspective views were used. Would the same people who dislike highly non-canonical images of people and vehicles facing away from the camera when paired with neutral titles prefer them when the same images were given titles biased toward "leaving" events?

#### Method

**Participants.** The 22 participants were students at the University of California, Berkeley who received partial course credit in their undergraduate psychology course for their participation. Their mean age was 20.1 years. All participants were naïve to the purpose and nature of the experiment, and gave informed consent in accord with the policies of the University of California, Berkeley, Committee for the Protection of Human Subjects, which approved the experimental protocol.

**Design.** There were 360 trials, resulting from all possible pairwise comparisons of nine image-title combinations, counterbalanced for left/right position of each image-title pair ( $9 \times 8 = 72$ ) for each of five objects ( $72 \times 5 = 360$ ). The nine image-title combinations for each object were each title for the object paired with the best perspective views of the object for each of the three titles as determined by the results of Experiment 3.3: 0°, 90°, and 180° for the vehicle

images and 0°, 45°, and 180° for the people images. The order and presentation of the trials was randomized by Presentation software (www.neurobs.com).

**Displays.** The background of the screen was a neutral gray. Each image measured 526 x 667 pixels, and had a 40-pixel black frame placed around it that had a beveled, three-dimensional appearance, to give the impression of a framed work (See Figure 9.8). The two images were placed side-by-side at the center of the screen. The title for each image was white text placed on a beveled black square that was centered horizontally below each image, and was separated from the image by 125 pixels. All parts of the screen were created in Adobe Photoshop CS3. The display measured 20" diagonally and had a resolution of 1620 x 1080 pixels, as in Experiment 3.

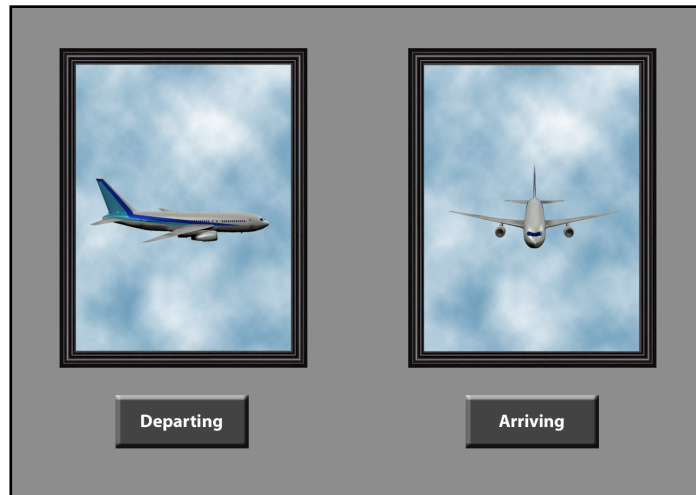
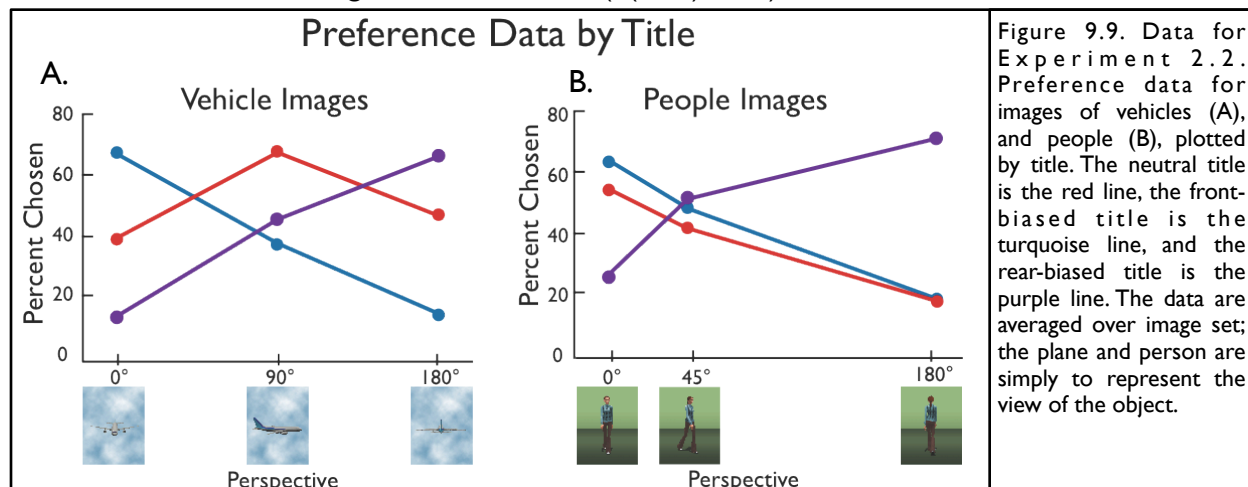


Figure 9.8. A schematic of the trial screen for Experiment 3.4

**Procedure.** Participants viewed the screen from a distance of approximately 60 cm. They were told that they would see two titled images on each screen, and that their task was to pick the image-title pair that they liked the best by using arrow keys on the keyboard.

## Results and Discussion

Because different perspectives were used for the images of vehicles and those of people (0°, 90°, and 180° for vehicles versus 0°, 45°, and 180° for people), these data were analyzed separately. For the images of vehicles, there was a main effect of title ( $F(2,42) = 18.463$ ). Across views, the neutral-titled images were chosen more frequently than the front-biased titled images ( $F(1,21) = 20.593, p < .001$ ) and the rear-biased titled images ( $F(1,21) = 28.484, p < .001$ ). The pattern of results plotted in Figure 9.9A indicates that this is due to the worse-fit neutrally-titled images being chosen more frequently than the worse-fit images for the elaborative titles rather than from the best-fit neutrally-titled image being chosen more frequently than the best-fit front-biased or rear-biased title images. Indeed, there was no statistically significant difference between the best-fit images for each title ( $F(2,42) < 1$ ). The interaction between title and



perspective view was significant ( $F(4,84) = 80.184, p < .001$ ), demonstrating that for all three objects (the car, the plane, and the boat) the different views were chosen more frequently with some titles than others. The object (car, plane, or boat)  $\times$  title interaction ( $F(4,84) = 1.730, p > .15$ ) and the object  $\times$  title  $\times$  view interaction ( $F(8,168) = 1.591, p > .12$ ) were not significant, and there was no main effect of object ( $F(2,42) < 1$ ).

The corresponding data are plotted in Figure 9.9B for the images of people. There was a main effect of title ( $F(2,42) = 10.533, p < .001$ ) driven by the fact that the front-biased title was slightly preferred to the neutral title ( $F(1,21) = 4.559, p < .05$ ). The rear-biased title was preferred to the front-biased title ( $F(1,21) = 7.274, p < .05$ ) as well as to the neutral title:  $F(1,21) = 17.015, p < .001$ . Unlike the vehicle images, this resulted in a statistically significant difference between the best fitting images for each title ( $F(2,42) = 17.974$ ). People preferred the best-fitting rear-biased title image to the best-fitting neutral title image ( $F(1,21) = 22.624, p < .001$ ), and the best-fitting front-biased title image to the best-fitting neutral title image ( $F(1,21) = 35.434, p < .001$ ). The best-fitting front-biased title image and the best-fitting rear-biased title image did not differ significantly, however ( $F(1,21) = 1.029, p > .3$ ).

It is worth highlighting three aspects of the results of the images of people. First, when images were given a rear-biased title, they liked the rear-view composition best, showing that in certain contexts, non-standard compositions are highly preferred to standard compositions. Second, there was a main effect of view ( $F(2,42) = 10.106, p < .001$ ), showing that the images displaying the  $0^\circ$  (front) view of the person were more frequently preferred than the  $180^\circ$  ( $F(1,21) = 13.115, p < .01$ ) and the  $45^\circ$  view was more frequently preferred than the  $180^\circ$  view ( $F(1,21) = 29.367, p < .001$ ), but the  $0^\circ$  view and the  $45^\circ$  view did not differ from each other ( $F(1,21) = 1.624, p > .2$ ). As with the vehicle images, the interaction between title and view was highly significant ( $F(4,84) = 56.492, p < .001$ ). There was no main effect of object ( $F(1,21) = 1.000, p > .30$ ), and the object category  $\times$  title and image-set  $\times$  view interactions were not significant ( $p > .30$  in both cases). Third, unlike Experiment 3.3,<sup>30</sup> the most preferred perspective for the front-biased title turned out to be the same as the most preferred perspective for the neutral title. With this in mind, the fact that the best-fit front-biased title image was preferred to the best-fit neutral title image is consistent with the results of Millis' (2001) study, in which people preferred the same image when it was given a relevant, elaborative title, than when it was given a descriptive title. This preference for "Coming Together" and "Greeting" could have been based on a similar sort of title-only preference as was found in Experiment 3.2.

Because preference for the title alone accounted for 16% of the variance in Experiment 3.2, we asked a different group of 22 participants to rank how much they liked each of the three titles, from most to least, for an unseen image. Again, we did not show the participants any image, but simply described the content of the image: "Rank how much you like the following three titles for an image of a plane: Arriving, Flying, and Departing." With these titles, the rankings were different from the titles used in the position experiments: For the perspective data, the main effect of kind of title was significant ( $F(2,42) = 3.903, p < .05$ ), such that the front-biased titles were preferred to the neutral and rear-biased titles. The interaction between kind of title and object was not significant ( $F(8,168) = 1.338, p < .3$ ).

The title ranking data described in Experiment 3.3 were used as a regression predictor, to see how much variance in the preference data for titled images it could account for on its own.

---

<sup>30</sup> In Experiment 3.3, the front-biased title had been chosen to fit best with the  $0^\circ$  (front) perspective, whereas the neutral title had been chosen to fit best with the  $45^\circ$  perspective.

Unlike Experiment 3.2, the preference for the titles alone accounted for a scant .5% of the variance, not a statistically significant amount ( $F(1,43) < 1$ ). The measured fit between the image and the title from Experiment 3.3, however, accounted for 84.5% of the variance ( $F(1,43) = 234.78, p < .001$ ), providing very strong support for the relevance of representational fit to aesthetic preferences for titled images.

### Conclusion for Representational Fit Experiments

Four experiments were conducted to examine the effects of context (by means of different titles) on people's aesthetic preferences for images. The results were consistent in demonstrating that people's aesthetic preference for titled images were largely driven by the goodness-of-fit between the title and the image. With the horizontal position of objects (Experiments 3.1, 3.2) and the perspective of objects (Experiments 3.3, 3.4), the fit between the title and the image accounted for 27.5% and 84.5% of the variance in people's preferences, respectively.

In the final chapter I will discuss the relation of these results on perspective to the results of the experiments on horizontal and vertical position (Chapters 5 and 6) and to the theory of fluency, which explained the results of experiments using untitled or neutrally-titled images very well, but not those using images with other titles.

# Chapter 10

## Interpreting the Contextual Effects of Titles

The results of Experiments 3.1 through 3.4 on contextual effects due to titles in Chapter 9 deal a significant blow to the claim that perceptual fluency is the best explanatory principle for aesthetic preferences. Whereas the notion of perceptual fluency could account very well for the results of the experiments with untitled images in Chapter 5 and 6, it cannot account for how preference for two identical images could be so drastically different when given different titles. It seems that the lack of any sort of account of meaning or context places a fairly tight boundary around the set of experiences that perceptual fluency can account for in the aesthetic domain. How often in the real world do we view images in a vacuum, with no context, be it from a title or from our personal experience?

### Conceptual Fluency

Before fluency is relegated to account for aesthetic preference only in the “default context” of viewing images for the purpose of identifying the depicted object(s), it must be acknowledged that there is another way in which fluency can attempt to account for a wider array of aesthetic experiences. As briefly mentioned in the discussion of fluency and aesthetics in Chapter 3, perceptual fluency, as it relates to the ease of perceiving an object, is not the only kind of fluency that its proponents have described. Whittlesea (1993) has also argued for the relevance of *conceptual* fluency: an ease of processing facilitated by semantic information. This makes the term fluency, without a qualifier, dangerously ambiguous, at least as it applies to aesthetics, because multiple conflicting fluencies could be present in the same image. One could view an image that was perceptually and conceptually fluent, perceptually fluent but conceptually disfluent, perceptually disfluent but conceptually fluent, or perceptually and conceptually disfluent. Reber, Schwarz, and Winkielman (2004) allude to this problem in one of their papers, and because there is no clear way to determine the relative contributions of different kinds of fluency; they simply leave the question open for future research. On a broader level, especially as the images in question get more similar to those that we see in the real world, including multiple levels of meaning and context, these components can be combined in rather confusing ways. For example, consider an artist who succeeds at his/her goal of creating an image that is purposely difficult to understand, such as *Nature's Writing* by Jean Paul Bordier (see Figure 7.3 in Chapter 7). Such an image would be perceptually disfluent on one level (because the woman's body is difficult to identify) and conceptually disfluent on another level (because the concept of a body being part of the earth is difficult to understand), but conceptually fluent on another, higher level, (because it illustrates well the concept of being difficult to understand).

There is a logical possibility that some kind of fluency theory might be able to account for the aesthetic appeal of an image whose goal is to be difficult to understand. However, it also seems that the point at which an image can be described as simultaneously conceptually fluent and conceptually disfluent is the point at which the meaning of the term “fluency” is dangerously unclear.

### Representational Fit as a Unifying Principle of Aesthetic Response

There seem to be two clear cases in which fluency theory makes concrete and testable predictions: perceptual fluency, in which the target object is easy to perceive, and conceptual

fluency, in which an image is fluently interpretable within a particular semantic context. Fluency theory cannot account for highly preferred images in which neither sort of fluency holds, however, and it does not have any clear-cut prediction when these two kinds of fluencies conflict. Moreover, representational fit can account for why fluency can account for a subset of aesthetic experiences in the first place: When a certain context dictates that images be easy to process, as with stock photography, images will be pleasing if they fit well with that context.

The notion of representational fit is also consistent a large portion of the classic literature on aesthetic response. This includes the experimental literature on novelty (e.g., the research by Berlyne and his followers reviewed in Chapter 2), as well as the arguments put forth by philosophers ranging from Plato (e.g., his treatment of beauty and wholeness in the *Phaedrus* mentioned in Chapter 1) to Kant (e.g., his work on our drive for order and structure) and Mangan (e.g., his argument that brings together philosophy, phenomenology, neural networks, and psychological data). For all these reasons, representational fit is a parsimonious, effective way to explain much of the data about people's aesthetic preferences.

At the same time, representational fit can shed some light on real-world experiences outside of the laboratory. While fairly specific in content and scope, the last four experiments suggest a principle of a good fit between the representation in the image and the meaning or context of the image, which can be employed to discuss a large variety of aesthetic experiences, including people's responses to art. It is my belief that an individual's response to a work of art depends on the fit between the image they see and their context of viewing it as art. This context, however, is in many ways different for different people and thus results in the large individual differences that are observed when studies are conducted on preference and actual works of art.

## Conclusion

Over the course of the history of psychology as a field, the study of aesthetics has waxed and waned. Its shortcomings are many: poor experimental design, limited scope, and faulty assumptions, to name a few. In general, the attitude of many people toward aesthetics seems to be rooted in the notion, "De gustibus non disputandum est," a Latin expression that roughly translates to "There can be no disputing about tastes." This aphorism is a sort of intellectual shrug espousing the view that aesthetic preferences are purely subjective, have no empirically detectable commonalities, and thus cannot be usefully subjected to the experimental rigor that is so highly valued in scientific psychology. There is also a difference between talking about a given topic and then dismissing it, versus not talking about the topic at all because it is not worthy of discussion. I believe that both of these views are present in the field and that neither is valid. Indeed, both do a disservice to the field of psychology insofar as they deprive various subfields of psychology of the wealth of information to be gleaned from studies of aesthetic preference that could be applied to other areas, such as perception, cognition, consciousness, and emotion.

Taking a cue from Fechner, who first realized that the components of psychological experience that were once thought to be too subjective for study could be studied rigorously to determine how a person's psyche related to the physics of the world, I've addressed the issue of aesthetic experience and preferences using psychophysical methods. Starting from first principles by using simple images and systematically varying their spatial compositions, striking similarities in people's preferences for images of objects have been documented in both the horizontal and vertical dimension of spatial composition. Moreover, the results of the final set of



studies on context have demonstrated that preference for composition is not absolute, but relative to a context that can be studied by manipulating image titles. In other words, the principles revealed in the first two sets of experiments do not provide a formula for creating aesthetically pleasing images. It is clearly not the case that an image will be pleasing if its focal object is always placed in a specific location (e.g., facing into the frame or showing the canonical perspective). All other factors being equal, the most preferred images are the ones that adhere to the biases described in Chapters 5 and 6, at least in the “default case” where the purpose of the image is simply to clearly show the depicted object. In other cases in which the images have different contexts, however, violations of these biases can be preferred at least as much, and sometimes more. Instead, it is the fit between the spatial composition of the image and the intended meaningful representation of its content that most strongly determines people’s aesthetic preferences.

## References

- Alexander, C. (2002). *The nature of order: An essay on the art of building and the nature of the universe, Volume 1: The phenomenon of life*. Berkeley: Center for Environmental Structure.
- Allen, G. (1877). *Physiological aesthetics*. London: Henry S. King & Co.
- Allesch, C.G. (2001). *Fechner's aesthetics – a provocation?* Lecture given at the symposium in honor of Gustav Theodor Fechner's 200<sup>th</sup> anniversary, October 19<sup>th</sup>-20<sup>th</sup>. Leipzig, Germany.
- Alter, A.L., & Oppenheimer, D.M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13, 219-235.
- Angier, R.P. (1903). The aesthetics of unequal division. *Psychological Review Monograph Supplement*, 4, 541-561.
- Anonymous. (1794). *Merkwürdige Lebensgeschichte des Franz Xaver Messerschmidt*. Vienna: n.p.
- Arnheim, R. (1969). *Visual thinking*. Berkeley: University of California Press.
- Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. 1<sup>st</sup> ed. Berkeley: University of California Press.
- Arnheim, R. (1974). *Art and visual perception: A psychology of the creative eye*. 2<sup>nd</sup> ed. Berkeley: University of California Press.
- Arnheim, R. (1988). *The power of the center*. Berkeley: University of California Press.
- Atalay, B. (2004). *Math and the Mona Lisa: The art and science of Leonardo Da Vinci*. Washington: Smithsonian Books.
- Bell, C. (1914). *Art*. London: Chatto and Windus.
- Benge, G.F. (2010). The Messerschmidt explosions. *New York Review of Books*, 57(19). Retrieved from <http://www.nybooks.com/articles/archives/2010/dec/09/messerschmidt-explosions/>
- Berlyne, D.E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.
- Bertamini, M., Bennett, K.M., & Bode, C. (2010). The anterior bias in visual art. *Perception*, 39. ECVF Abstract Supplement, 116.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14(2), 143-177.
- Biederman, I., & Vessel, E.A. (2006). Perceptual pleasure and the brain. *American Scientist*, 94, 247-253.

- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, 106(2), 265-289.
- Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, 63, 545-545.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1-28.
- Boselie, F. (1984a). Complex and simple proportions and the aesthetic attractivity of visual patterns. *Perception*, 13, 91-96.
- Boselie, F. (1984b). The aesthetic attractivity of the golden section. *Psychological Research*, 45, 367-375.
- Boselie, F. (1992). The golden section has no special aesthetic attractivity. *Empirical Studies of the Arts*, 10, 1-18.
- Bourdier, J.P. (2007). *Bodyscapes*. San Rafael: Earth Aware Editions.
- Bower, G.H., Karlin, M.B., & Dueck, A. (1975). Comprehension and memory for pictures. *Memory & Cognition*, 3, 216-220.
- Bradley, R.A., & Terry, M. E. (1952). Rank analysis of incomplete block designs, I. The method of paired comparisons. *Biometrika*, 39, 324-345.
- Bransford, J.D., & Johnson, M.K. (1973). Considerations of some problems of comprehension. In W.G. Chase (Ed.), *Visual Information Processing* (pp. 383-438). New York: Academic Press.
- Bückling, M. (Ed.). (2006). *The fantastic heads of Franz Xaver Messerschmidt: Catalogue of the exhibition at the Liebieghaus Skulpturensammlung Frankfurt am Main, 15 November 2006 to 11 March 2007*. Munchen: Hirmer.
- Canon Digital Learning Center. *EOS Rebel T1i: Shoot like you see, edit like you think*. Retrieved from <http://www.usa.canon.com/dlc/controller?act=GetArticleAct&articleID=3026>.
- Cassirer, E. (2005). Hermann Cohen and the renewal of Kantian philosophy. (L. Patton, Trans.). *Angelaki*, 10(1), 95-108. (Original article published 1918).
- Cohen, A., Ivry, R., & Keele, S.W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(1), 17-30.
- Cohen, P. (2007, November 25). Freud is widely taught at universities, except in the psychology department. *New York Times*. Retrieved from <http://www.nytimes.com/2007/11/25/weekinreview/25cohen.html>

- Cupchik, G.C. (1992). From perception to production: A multilevel analysis of the aesthetic process. In G.C. Cupchik & J. László (Eds.), *Emerging visions of the aesthetic process: Psychology, semiology, and philosophy*. New York: Cambridge University Press.
- Danto, A.C. (2003). *The abuse of beauty: Aesthetics and the concept of art*. Peterborough: Carus Publishing Company.
- Diderot, D. (1966). On the Interpretation of Nature. (L.G. Crocker, Ed., D. Coltner, trans.). In *Classics in the history of thought: Diderot's selected writings*. New York: Macmillan Publishing Company. (Original work published 1753).
- Dix, R. (1986). Addison and the concept of 'novelty' as a basic aesthetic category. *British Journal of Aesthetics*, 26(4), 383-390.
- Dutton, D. (2009). *The art instinct: Beauty, pleasure, and human evolution*. New York: Bloomsbury Press.
- Eisler, C. (1969). Kunstgeschichte American style: A study in migration. In D. Fleming (Ed.), *The intellectual migration: Europe and America: 1930-1960*. Cambridge: Harvard University Press.
- Estes, Z., Verges, M., Barsalou, L.W. 2008. Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, 19(2), 93-97.
- Fechner, G.T. (1871). *Zür experimentalen aesthetik*. Leipzig: S. Hirzel.
- Fechner, G.T. (1876). *Vorschule der aesthetik*. 2 vols. Leipzig: Breitkopf & Härtzel.
- Field, G. (1845). *Chromatics: The analogy, harmony, and philosophy of colours*. London: David Bogue, Fleet Street.
- Firth, D. (2005). BradleyTerry: Bradley-Terry models. R Package, Version 0.8-5.
- Freud, S (1989). *Introductory lectures on psychoanalysis*. (J. Strachey, Trans.) New York: W.W. Norton & Company. (Original work published 1933).
- Freud, S. (1989). Leonardo da Vinci and a memory of his childhood. In P. Gay (Ed.), *The Freud reader* (pp. 443-481). New York: W.W. Norton & Company. (Original work published 1910).
- Freud, S. (1989). The Moses of Michaelangelo. In P. Gay (Ed.), *The Freud reader* (pp. 522-539). New York: W.W. Norton & Company. (Original work published 1914).
- Freud, S. (2003). Formulations on the two principles of mental functioning. In R. Steiner, (Ed.), *Unconscious phantasy* (pp. 67-76). London: H. Karnak Ltd. (Original work published 1911).
- Gaffron, M. (1950). Left and right in pictures. *Art Quarterly*, 13, 312-331.
- Gerard, A. (1764). *An essay on taste*. London: Millar.

- Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (eds.), *Perceiving, acting, and knowing* (pp. 67-82). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gombrich, E.H. (1960). *Art and Illusion*. London: Phaidon Press.
- Green, C.D. (1995). All that glitters: a review of psychological research on the aesthetics of the golden section. *Perception*, 24(8), 937-968.
- Greenberg, C. (1982). Modernist Painting. In F. Francina & C. Harrison (Eds.), *Modern art and modernism: A critical anthology*. London: Sage Publications, Ltd.
- Griscom, W., & Palmer, S.E. "Individual Differences in Preference for Harmony." Paper presented at the 10th Annual Meeting of the Vision Science Society, Naples, FL, May 2010.
- Haines, T.H., & Davies, S.E. (1904). Psychology of aesthetic reaction to rectangular forms. *Psychological Review*, 11, 249-251.
- Home, H. (1765). *Elements of criticism*. London: Millar.
- Kant, I. (1790/1987). *Critique of judgment*. (W.S. Pluhar, Trans.). Indianapolis: Hackett Publishing Co.
- Khalil, S., & McBeath, M. (2006). Canonical representation: An examination of preferences for viewing and depicting 3-dimensional objects [Abstract]. *Journal of Vision*, 6(6):267, 267a.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: W.W. Norton & Company.
- Konecni, V.J. (1978). Daniel E. Berlyne: 1924-1976. *American Journal of Psychology*, 91 (1), 133-137.
- Konecni, V.J. (1984). Elusive effects of artists' messages. *Advances in Psychology*, 19, 71-93.
- Konecni, V.J. (1991). Portraiture: An experimental study of the creative process. *Leonardo*, 24(3), 325-328.
- Konecni, V.J. (2003). The golden section: Elusive, but detectable. *Creativity Research Journal*, 15 (2-3), 267-275.
- Konecni, V.J. (2005). The Aesthetic Trinity: Awe, Being Moved, Thrills. *Bulletin of Psychology and the Arts*, 5(2), 27-44.
- Konecni, V.J., Cline, L.E. (2002). The "Golden Woman:" An Exploratory Study of Women's Proportions in Paintings. *Visual Arts Research*, 69-77.
- Konkle, T., & Oliva, A., (In press). Canonical visual size for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance*.
- Koren, L. (2010). *Which aesthetics do you mean? Ten definitions*. Point Reyes: Imperfect Publishing.

- Kris, E. (1952). *Psychoanalytic explorations in art*. New York: International Universities Press.
- Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, 207(4430), 557-558.
- Lalo, C. (1908). *L'esthétique expérimentale contemporaine*. Paris: Alcan.
- Lanz, H. (1931). *The physical basis of time: An essay on the aesthetics of sound*. Stanford: Stanford University Press. (Reprinted: New York, Greenwood, 1968).
- Levy, J. (1976). Lateral dominance and aesthetic preference. *Neuropsychologia*, 14(4), 431.
- Livio, M. (2002). *The golden ratio: The story of phi, the world's most astonishing number*. New York: Broadway Books.
- Locher, P.J., Krupinski, E.A., Mello-Thoms, C., & Nodine, C. F. (2007). Visual interest in pictorial art during an aesthetic experience. *Spatial Vision*, 21(1-2), 55-77.
- Locher, P.J., Overbeeke, K., & Stappers, P.J. (2005). Spatial balance of color triads in the abstract art of Piet Mondrian. *Perception*, 34(2), 169-189.
- Locher, P.J., Smith, J. K., & Smith, L. F. (2001). The influence of presentation format and viewer training in the visual arts on the perception of pictorial and aesthetic qualities of paintings. *Perception*, 30(4), 449-465.
- Locher, P.J., Jan Stappers, P., & Overbeeke, K. (1998). The role of balance as an organizing design principle underlying adults' compositional strategies for creating visual displays. *Acta Psychologica*, 99(2), 141-161.
- Locher, P.J., Stappers, P.J., & Overbeeke, K. (1999). An empirical evaluation of the visual rightness theory of pictorial composition. *Acta Psychologica*, 103(3), 261-280.
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: J. Wiley.
- Mangan, B.B. (1991). *Meaning and the structure of consciousness: An essay in psychoaesthetics*. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Paper presented at the 4th Annual International Conference on Computer Vision, Vancouver.
- Martin, L. J. (1906). An experimental study of Fechner's principles of aesthetics. *Psychological Review*, 13(3), 142-219.
- Martindale, C.E. (1969). *The psychology of literary change*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- McLaughlin, J. P. (1986). Aesthetic preference and lateral preferences. *Neuropsychologia*, 24(4), 587.

- McLaughlin, J. P., Dean, P., & Stanley, P. (1983). Aesthetic preference in dextrals and sinistrals. *Neuropsychologia*, 21(2), 147.
- McManus, I. C. (1980). The aesthetics of simple figures. *British Journal of Psychology*, 71(4), 505-524.
- McManus, I. C., Edmondson, D., & Rodger, J. (1985). Balance in pictures. *British Journal of Psychology*, 76(3), 311-324.
- McManus, I. C., & Furnham, A. (2006). Aesthetic activities and aesthetic attitudes: Influences of education, background and personality on interest and involvement in the arts. *British Journal of Psychology*, 97(4), 555-587.
- McManus, I. C., & Thomas, P. (2007). Eye centring in portraits: A theoretical and empirical evaluation. *Perception*, 36(2), 167-182.
- McManus, I. C., & Weatherby, P. (1997). The golden section and the aesthetics of form and composition: A cognitive model. *Empirical Studies of the Arts*, 15, 209-232.
- Nachson, I., Argaman, E., & Luria, A. (1999). Effects of directional habits and handedness on aesthetic preference for left and right profiles. *Journal of Cross-Cultural Psychology*, 30(1), 106-114.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1-32.
- Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo*, 26(3), 219-227.
- Osborne, H. (1970). *Aesthetics and art theory: An historical introduction*. New York: E.P. Dutton & Co.
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, 3(5), 519-526.
- Palmer, S. E. (1991). On goodness, Gestalt, groups, and Garner: Local symmetry subgroups as a theory of figural goodness. In G. Lockhead & J. Pomerantz (Eds), *The Perception of Structure: Essays in Honor of Wendell R. Garner*. Washington: American Psychological Association.
- Palmer, S.E., Gardner, J.S., & Wickens, T.D. (2008). Aesthetic issues in spatial composition: effects of position and direction on framing single objects. *Spatial Vision*, 21(3-5), 421-449.
- Palmer, S. E. & Hemenway, K. (1978). Orientation and symmetry: effects of multiple, rotational, and near symmetries. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 691-702.
- Palmer, S.E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX*. Hillsdale, NJ: Erlbaum.

- Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences*, 107(19), 8877-8882.
- Peckham, M. (1965). *Man's rage for chaos: Biology, behavior, and the arts*. New York: Schocken Books.
- Pierce, E. (1894). Aesthetics of simple forms. I. Symmetry. *Psychological Review*, 1, 483-495.
- Plato. (1973). *Phaedrus*. (Walter Hamilton, Trans.). New York: Penguin Books.
- Plato. (1892). The Republic. In (B. Jowett, trans.), *The Dialogues of Plato* (3<sup>rd</sup> ed.). 5 vols. New York: Macmillan and Co.
- Pötzl-Malikova, M. (Ed.). (2010). *Franz Xaver Messerschmidt, 1736-1783: From Neoclassicism to Expressionism*. Milano: Officina Libraria.
- Puffer, E.D. (1903). Studies in symmetry. *Psychological Review, Monograph Supplements*, 4(1), 467-539.
- R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramachandran, V.S., & Hirstein, W. (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies*, 6(7), 15-51.
- Random House, Inc. (1997). Aesthetics. In *Random House Unabridged Dictionary*. Retrieved from <http://dictionary.infoplease.com/aesthetics>
- Random House, Inc. (1997). Beauty. In *Random House Unabridged Dictionary*. Retrieved from <http://dictionary.infoplease.com/beauty>
- Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology*, 46(3), 342.
- Reber, R. (in press). Processing fluency, aesthetic pleasure, and culturally-shared taste. In S.E. Palmer & A.P. Shimamura (Eds.), *Aesthetic science: Connecting minds, brains, and experience*. New York: Oxford University Press.
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9(1), 45-48.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364-382.
- Reynolds, J. (1798). *The works of Sir Joshua Reynolds, Knt. late president of the Royal Academy: containing his discourses, idlers, a journey to Flanders and Holland, (now first published,) and his commentary on Du Fresnoy's art of painting; printed from his revised copies, (with his last*



corrections and additions,) in three volumes. (3 vols.)(2<sup>nd</sup> ed.). London: T. Cadell, Jun. and W. Davies, in the strand. Retrieved from: <http://books.google.com/books?id=XotAAAAAYAAJ&dq=sir%20joshua%20reynolds%201798&pg=PP5#v=onepage&q&f=false>

Riehl, A. (1876). *Der philosophische criticismus und seine bedeutung für die positive wissenschaft. Vol 1: Geschichte und methode des philosophischen criticismus*. Leipzig: Wilhelm Engelmann.

Rosch, E., & Lloyd, B.B. (1978). *Cognition and categorization*. Hillsdale: Lawrence Erlbaum Associates, Inc.

Santayana, G. (1896). *The sense of beauty: Being the outlines of aesthetic theory*. New York: Charles Scribner's Sons.

Santayana, G. (1905). *The life of reason*. 5 vols. New York: Scribner.

Sauerländer, W. (2010a). It's all in the head. *New York Review of Books*, 57(16). (Dollenmayer, D., trans.). Retrieved from <http://www.nybooks.com/articles/archives/2010/oct/28/its-all-in-the-head/>

Sauerländer, W. (2010b). The Messerschmidt Explosions, Reply. *New York Review of Books*, 57(19). Retrieved from <http://www.nybooks.com/articles/archives/2010/dec/09/messerschmidt-explosions/>

Schachter, S., & Singer, J.E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379-399.

Seamon, J. G., Marsh, R. L., & Brody, N. (1984). Critical importance of exposure duration for affective discrimination of stimuli that are not recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(3), 465-469.

Smith, N. (1797). *Remarks on rural scenery, with some twenty etchings of cottages, from nature; and some observations and precepts relative to the picturesque*. London: Nathaniel Smith.

Strunk, W., & White, E.B. (2000). *The Elements of Style*. 4th ed. New Jersey: Pearson Education Company.

Tatler, B.W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14).

Thorndike, E. L. (1917). Individual differences in judgments of the beauty of simple forms. *Psychological Review*, 24(2), 147-153.

Tyler, C.W. (1998a). Painters centre one eye in portraits, *Nature*, 392, 877-878.

Tyler, C.W. (1998b). Eye placement principles in portraits and figure studies of the past two millennia, *SPIE Proc.* 3299, 431-437.

- Wertheimer, M. (1938). Laws of organization in perceptual forms. In *A source book of Gestalt psychology*, 71–88. (Original work published 1923).
- Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin*, *124*, 333-371.
- Whittlesea, B.W.A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1235-1253.
- Wind, E. (1937). Comments on an observation by Reynolds. *Journal of the Warburg Institute*, *1*(1), 70-71.
- Winkielman, P., Schwarz, N., & Nowak, A. (2002). Affect and processing dynamics: Perceptual fluency enhances evaluations. In S. Moore & M. Oaksford (Eds.), *Emotional cognition: From brain to behaviour*, (pp. 111-135). Philadelphia: John Benjamins Publishing Company.
- Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*(9), 799-806.
- Wölfflin, H. (1928). Über das rechts und links in bilde. *Gedanken zur Kunstgeschichte*, (pp. 82–90). Schwabe, Basel, Switzerland.
- Yonan, M. (2009). The man behind the mask? Looking at Franz Xaver Messerschmidt. *Eighteenth Century Studies*, *42*(3), 431-451.
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, *8*(6), 448-453.