INFLUENCE OF PRODUCT DESIGN ON INFORMATION PROCESSING

By

JOANDREA HOEGG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2006
Copyright 2006

by

JOANDREA HOEGG
This dissertation is dedicated to my parents.
ACKNOWLEDGMENTS

I would like to thank my advisor, Joseph W. Alba, to whom I am extremely grateful for guidance, encouragement, and patience. Joe was incredibly generous of his time, and I would not have succeeded without his support. I would also like to thank Alan Cooke, Rich Lutz, and Mike Lewis for their kind and valuable assistance throughout my time at the University of Florida. I thank Chris Janiszewski for his excellent advice, and my external committee member, Jason Colquitt, for his insightful comments on my work. I also express my gratitude to Mickey and Joe Belch, who encouraged me to pursue a PhD at the University of Florida.

I am blessed with many supportive friends both inside and outside the PhD program. At Florida, I would particularly like to thank Elise Chandon, Qiong Wang, Julia Belyavsky, Dan Rice, Anne Magi, and Wouter Vanhouche. Outside of the university, I thank Jackie Levy, Ann Lane, and Leah Thomson for many years of wonderful friendship. Also I am deeply grateful to my family for supporting my decision to return to school to pursue my PhD. Finally, I would like to thank Timothy Silk for believing in me all those times that I did not believe in myself, and for teaching me always to look on the bright side. I could not have done this without him.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 REVIEW OF RESEARCH ON APPEARANCE</td>
<td>4</td>
</tr>
<tr>
<td>Aesthetic Influence</td>
<td>5</td>
</tr>
<tr>
<td>Evidence from Research on Attractiveness</td>
<td>5</td>
</tr>
<tr>
<td>Evidence from Other Domains</td>
<td>7</td>
</tr>
<tr>
<td>Affect versus evaluation</td>
<td>7</td>
</tr>
<tr>
<td>Motivated reasoning</td>
<td>11</td>
</tr>
<tr>
<td>Halo effects</td>
<td>17</td>
</tr>
<tr>
<td>Awareness of affective and motivational influences</td>
<td>17</td>
</tr>
<tr>
<td>Non-Aesthetic Influence of Product Appearance</td>
<td>18</td>
</tr>
<tr>
<td>Summary</td>
<td>21</td>
</tr>
<tr>
<td>3 OVERVIEW OF EXPERIMENTS</td>
<td>23</td>
</tr>
<tr>
<td>4 PILOT STUDY</td>
<td>26</td>
</tr>
<tr>
<td>Method</td>
<td>27</td>
</tr>
<tr>
<td>Pretest</td>
<td>27</td>
</tr>
<tr>
<td>Design and Procedure</td>
<td>28</td>
</tr>
<tr>
<td>Results</td>
<td>30</td>
</tr>
<tr>
<td>Backpack Ratings</td>
<td>30</td>
</tr>
<tr>
<td>Running Shoe Ratings</td>
<td>31</td>
</tr>
<tr>
<td>Discussion</td>
<td>32</td>
</tr>
<tr>
<td>5 EXPERIMENT 1</td>
<td>36</td>
</tr>
<tr>
<td>Method</td>
<td>36</td>
</tr>
<tr>
<td>Stimuli</td>
<td>36</td>
</tr>
<tr>
<td>Design and Procedure</td>
<td>36</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1. Pilot Study Stimuli</td>
<td>34</td>
</tr>
<tr>
<td>4-2. Pilot Study Backpack Feature Ratings</td>
<td>35</td>
</tr>
<tr>
<td>4-3. Pilot Study Choice Proportions of Superior Backpack</td>
<td>35</td>
</tr>
<tr>
<td>4-4. Pilot Study Running Shoe Feature Ratings</td>
<td>35</td>
</tr>
<tr>
<td>4-5. Pilot Study Choice Proportions of Superior Running Shoe</td>
<td>35</td>
</tr>
<tr>
<td>5-1. Running Shoe Feature Descriptions Used in Experiment 1</td>
<td>40</td>
</tr>
<tr>
<td>5-2. Experiment 1 Running Shoe Feature Ratings</td>
<td>40</td>
</tr>
<tr>
<td>5-3. Experiment 1 Choice Proportions of Superior Running Shoe</td>
<td>40</td>
</tr>
<tr>
<td>6-1. Pretest Results for Performance Pictures</td>
<td>46</td>
</tr>
<tr>
<td>6-2. Pretest Results for Feature Descriptions</td>
<td>46</td>
</tr>
<tr>
<td>6-3. Results of Experiment 2</td>
<td>46</td>
</tr>
<tr>
<td>6-4. Manipulation Check for Experiment 2</td>
<td>47</td>
</tr>
<tr>
<td>7-1. Dimensions Used in Experiment 3</td>
<td>51</td>
</tr>
<tr>
<td>7-2. Results of Experiment 3</td>
<td>51</td>
</tr>
<tr>
<td>7-3. Manipulation Check for Experiment 3</td>
<td>52</td>
</tr>
<tr>
<td>7-4. Results of Experiment 3 with Helmet Replicate Removed</td>
<td>52</td>
</tr>
<tr>
<td>8-1. Pretest Picture Ratings</td>
<td>60</td>
</tr>
<tr>
<td>8-2. Results of Experiment 4</td>
<td>60</td>
</tr>
<tr>
<td>8-3. Manipulation Check for Experiment 4</td>
<td>60</td>
</tr>
<tr>
<td>9-1. Results of Experiment 5</td>
<td>64</td>
</tr>
</tbody>
</table>
INFLUENCE OF PRODUCT DESIGN ON INFORMATION PROCESSING

By

JoAndrea Hoegg

August 2006

Chair: Joseph W. Alba
Major Department: Marketing

Despite the growing importance of product design in the marketplace, relatively little consumer research has been conducted addressing the influence of product appearance on the processing of other product information. The current investigation presents the results of five experiments that pit product appearance against functional feature descriptions to examine how conflicting visual information and verbal information are reconciled. In all studies participants were asked to compare the quality of two brands of a given product based on a reading of allegedly unbiased professional reviews. The primary experimental manipulation was that for some participants the reviews were accompanied by pictures of the two brands. For each pair of products, one brand was superior visually but was actually inferior based on the objective review, and one brand was inferior visually but was objectively superior. In some experiments the visual superiority was aesthetic; that is, one product was more attractive than the other. In other experiments the visual superiority was performance based: one product looked like it would offer superior performance on a given dimension. The primary finding was
that product appearance exerted an influence on people’s tendency and ability to evaluate product quality, but only when the visual difference was performance based and not aesthetically based. Diagnosticity appeared to be the driving force.
CHAPTER 1
INTRODUCTION

Product design is of critical importance for managers. In an era of sophisticated manufacturing processes, where it is difficult for companies to gain competitive advantages in other areas such as price or reliability (Jordan, Thomas and McClelland 1996; Kalins 2003), design can act as a “strategic tool” that firms can use as a point of differentiation (Page and Herr 2002, p. 133). Moreover, with designer styles available at affordable prices (e.g., Isaac Mizrahi in Target; Tommy Hilfiger in JC Penney), the average consumer is now able to purchase and hence demand products that are not only functionally sound, but also aesthetically pleasing (Miller and Adler 2003).

Marketers and academics agree that product design is important, and there exists a large body of managerial and quantitative research on design as a component of the new product development process. Consumer research on design, however, has been somewhat limited. Consumer behavior investigations into product appearance have included examinations of behavioral responses to product form (Silvera, Josephs, and Giesler 2002; Veryzer and Hutchinson 1998), the interaction of brand strength and design (Page and Herr 2002), the effects of repeated exposure to designs (Cox and Cox 2002), and the role of individual differences in response to design features (Bloch, Brunel and Arnold 2003; Holbrook and Schindler 1994). Bloch (1995) developed a model of consumer response to product form that outlined several stages in the design process from managerial goals through the creation of the product’s form to the consumer’s response. Bloch’s model addressed the possibility of both cognitive and affective
responses to product form, but did not speak specifically to how variations in product design might influence the nature of the cognitive or affective reactions. In a discussion of the different roles of product design, Creusen and Schoormans (2005) similarly mentioned the psychological role that product appearance might play but did not offer predictions as to the nature and direction of the influence.

The current research contributes to the growing literature on product design by investigating how a product’s appearance can influence the processing of functional product information. I examine the extent to which design signals quality, and how such a signal might interfere with the evaluation of more objective product information. Prior consumer research in product design has examined how aesthetics could influence quality evaluations in the absence of other information (Page and Herr 2002). Other work has also looked at differences between visual and verbal information, primarily in terms of picture and word use in advertising (e.g., Hirschman 1986; Holbrook and Moore 1981). However, no research has specifically addressed how the appearance of a product might alter the evaluation of objective product information. In addition, prior research has not addressed the possibility of different types of design influence. The research reported here compares the influence of aesthetic design, or design for the sake of beauty, to the influence of performance-based design, or design as a signal of product performance, and investigates how these two roles might exert differential influence on consumer evaluations.

Five experiments were conducted that pitted design against written product information to examine how conflicting visual and verbal information would be reconciled. In all studies participants were provided with objective reviews about two
brands and were asked to evaluate the brands on particular features based on the reviews. Accompanying the two reviews were pictures of the brands that conflicted with the review. That is, the brand that was better in terms of feature descriptions was inferior visually, and the brand that was worse based on feature descriptions had the superior design. The visual superiority had two forms. In some studies the superiority of the design was aesthetic: one product was more attractive than the other. In other experiments the superiority of the design was performance based; one product looked like it would offer superior performance along certain dimensions.

The paper is organized as follows. Chapter 2 provides a review of literature on visual effects in marketing and psychology, as well as other relevant research to outline the scope of the investigation and inform the key questions. Because the studies distinguish aesthetic design from performance-based design the literature review is divided into two sections based. First is a review of aesthetic effects, where the influence of an object’s appearance relates specifically to its attractiveness. Second is a review of performance-based appearance effects; that is, where an object’s appearance communicates some level of performance or quality. Chapter 3 provides an overview of the studies. Chapters 4 through 9 present the methodology, data analysis, and findings of a pilot study and five experiments. Chapter 10 includes a general discussion of the key results, along with marketing implications, limitations, and avenues for future research.
CHAPTER 2
REVIEW OF RESEARCH ON APPEARANCE

Dating back to classical times, scholars have considered the aesthetics of the visual form (Dickie 1971), mostly attempting to determine what is considered beautiful and why. In more recent years, psychologists have attempted to determine experimentally what elements of an object make it attractive or unattractive (e.g., Berlyne 1974; Green 1995; Langlois et al. 2000; Martindale 1984; Martindale and Moore 1988; Rhodes et al. 2002; Rubenstein, Langlois and Roggman 2002; Zebrowitz 1997), and within marketing, a small but growing number of researchers have joined the quest to determine what design elements consumers find attractive (e.g., Cox and Cox 2002; Raghubir and Greenleaf 2006; Veryzer and Hutchinson 1998).

Separate from the literature on aesthetics per se, several streams of research exist that examine, directly or indirectly, ways in which appearance might influence the processing of other information. Research addressing this issue comes primarily from face perception, which has demonstrated that appearance can influence ratings of personality (Ambady and Rosenthal 1993; Eagly et al. 1991), job suitability (Hassin and Trope 2000) and even work quality (Landy and Sigall 1974; Mueller and Mazur 1996). In the product realm, there is evidence that even clearly irrelevant design elements like an invitation card or a web page background can have an influence on product preference (Mandel and Johnson 2002; Raghubir and Greenleaf 2006). Although such effects have been observed on preference rather than more objective ratings, the presence of influence in an unequivocally irrelevant context suggests the potential power of design.
Whether there are conditions under which design might exert influence over objective product quality judgments remains largely unresolved. Some research in marketing has shown that the attractiveness of a product design can influence product quality ratings, although the influence appears to occur at the margins (Page and Herr 2002; Yamamoto and Lambert 1994).

To help illuminate the issue, the review presented here will focus on research in person and object appearance, but will also borrow from several literature streams not centered on the visual form, notably work in motivated reasoning, halo effects, the interplay of affect and cognition, and accessibility-diagnosticity. Although these literatures do not specifically address appearance, they are useful inasmuch as they speak to the potential impact a stimulus can have on information processing.

This review divides the topic of the influence of appearance into two parts. The first section focuses specifically on aesthetics, that is, on appearance as a source of beauty or pleasure. The second section focuses on appearance as a signal of quality or performance.

**Aesthetic Influence**

**Evidence from Research on Attractiveness**

There is a fairly large body of research demonstrating effects that physical attractiveness can have on judgments of other people’s personality traits and capabilities (Eagly et al. 1991; Landy and Sigall 1974; Rubenstein et al. 2002; Zebrowitz 1997). For example, attractive people are perceived as having superior skills such as leadership, social ability and job competence (Langlois et al. 2000). There is even some research suggesting that physical attractiveness can influence evaluations of work quality. Landy and Sigall (1974) had participants rate essays supposedly written by students whose
pictures accompanied the essays. Despite the obvious irrelevance of appearance to the quality of an essay, participants rated both the essays and the writing skills of more attractive people as superior to those of less attractive people.

To a far more limited extent, attractiveness effects have been found in the consumer realm. Raghubir and Greenleaf (2006) demonstrated a dramatic attractiveness effect in an experiment that investigated aesthetic preference for the golden ratio of rectangles. Participants were presented with an invitation to a concert. The only difference between experimental conditions was that the invitation was printed on rectangular cards of one of two ratios: 1.38:1 or 1.62:1 (the golden ratio). Participants receiving the invitation conforming to the golden ratio rated the concert as more harmonious and indicated a higher likelihood of purchasing a CD of the performance than did participants receiving the invitation with a 1.38:1 ratio.

Page and Herr (2002) have suggested that the attractiveness of a product design produces an affective response associated with feeling-based, relatively automatic evaluative processes, while the quality of the product is based more on higher-order cognitive evaluation. In support of their results, Page and Herr (2002) found that liking judgments occurred faster than quality judgments. More importantly, product liking was driven by reaction to the product’s design and was not influenced by brand name or functionality, whereas evaluation of the quality of the product seemed to integrate design, function and brand information.

In terms of product evaluation, the design of a product may be impossible to ignore, even when the evaluation of the product should be totally independent of attractiveness. Some evidence for this comes from a study on industrial products
conducted by Yamamoto and Lambert (1994). They had business people in various organizational positions from sales to engineering indicate preferences for various industrial products. Products were presented on cards consisting of product attribute specifications and a photograph. Through a conjoint analysis technique Yamamoto and Lambert found that the attractiveness of the product did have an influence, even for engineers and buyers who were well informed about product characteristics. Across seven industrial products the more attractive product was significantly favored. This influence did not appear to be non-normative, however, since attractiveness was one of the lowest variables found to affect choice and the influence was only at the margins. Nonetheless, an influence of attractiveness was present despite overwhelming self-reports that attractiveness should play no role in the selection of a superior product. Considering that the products were industrial and the participants were experts, the finding that attractiveness played any role at all is striking, and suggests that such an influence may exist for consumers.

**Evidence from Other Domains**

There are several streams of research outside the aesthetics literature showing that an initial reaction to a stimulus can influence its subsequent evaluation. Three areas are directly relevant to the current discussion: the interplay of affective reactions and cognitive evaluations, motivated reasoning, and halo effects. Each of these is reviewed briefly.

**Affect versus evaluation**

Many authors have argued that affect and evaluation are distinct concepts, with affect referring to feeling states that arise due to interaction with a stimulus and evaluation referring more to a cognitive process of judgment based on beliefs (Breckler
and Wiggins 1989; Cohen and Areni 1991). Both can occur for a given object, although some have argued that affective reactions can occur entirely without cognitive evaluations (Berkowitz 1993; Zajonc 1980; Zajonc and Markus 1982, 1985).

In an attempt to disentangle affective reaction from a more cognitive appraisal, Breckler and Wiggins (1989) asked participants to evaluate blood donations on different dimensions. They asked for a series of relatively cognitive evaluations of blood donations, using semantic differential scales of the extent to which donating blood was useful, good, wise, etc. They also asked a series of feelings-based questions regarding how people felt when they gave blood. The key finding was among non-donors, who had positive cognitive evaluations of blood donation, but negative affective reactions toward blood donation, indicating that their affective and cognitive evaluations were distinct.

There is evidence to suggest that an affective reaction to a stimulus is resistant to change (Edwards 1990) and may be a better predictor of the type of thoughts people have about products than cognitive evaluations (Pham et al. 2001). As suggested by Zajonc (1980) and Edwards (1990), a strong affective response may be highly resistant to product information that contradicts the individual’s feeling about the product. Edwards (1990) examined the resistance to change of both cognitively and affectively formed attitudes, looking specifically at how attitudes formed through affective means or through cognitive means could be altered using affective or cognitive persuasion. She demonstrated that the effectiveness of a rational (cognitive) or emotional (affective) persuasive message depended on the nature of the attitude’s origin. Attitudes originally formed through an affective process were resistant to rational persuasive attempts and could only be altered through affective means, supporting Zajonc’s (1980) argument that
affective judgments are not easy to change. Cognitive attitudes (those formed through rational means) on the other hand, could be altered through both cognitive and affective means of persuasion.

Edwards (1990) used primacy to manipulate whether an attitude was formed through affective or cognitive means, assuming that the attitude formation occurred with the first persuasive exposure. If valid, this possibility has relevance when considering the affective reaction to product design versus a cognitive appraisal. In the marketplace it is common for consumers to see a product, through advertising or store displays, prior to learning about its features. If a person were to encounter a product’s design first, and then later read information about product function, the attitude formed during the observation of the design could be resistant to contrary information in the written description. On the other hand, if a person read the product information before seeing the design itself, any conflicting affective reaction to the design might still influence the attitude formed during the reading of the information, although it is likely the influence would be relatively weaker.

Pham et al. (2001) examined the role of affective reaction to a product and compared that reaction to a more reasoned, cognitive analysis. They predicted that feeling responses (affective reactions) would be better predictors of the number and valence of thoughts people have about products than reason-based responses. In accordance with models of affective response (e.g. Berkowitz 1993; Cohen and Areni 1991), Pham et al. proposed that feeling-based responses are generally faster than reason-based responses, and that an initial affective response prompts thought generation through “both automatic and controlled processes” (p. 170). In study 1 (Pham et al. 2001), participants evaluated
pictures from magazines using either a reason-based analysis or a feeling-based analysis. The authors found that feeling responses were faster than cognitive responses, suggesting that the pictures triggered a relatively automatic affective reaction. In study 2 Pham et al. (2001) examined the number of valenced thoughts generated through feeling-based or reason-based analysis. The valence of subjects’ spontaneously generated thoughts was more strongly related to the valence of the feeling-based responses than the reason-based responses.

Pham et al. (2001) also found that feeling-based assessments showed more interpersonal agreement than reason-based assessments. They concluded that feeling-based evaluations are more natural and more consistent. This suggestion is consistent with Takahashi (1995), whose work in the psychology of aesthetics found that people were very good at guessing the emotion portrayed through a drawing, and also with research in facial attractiveness showing that people generally agree in their ratings of attractive and unattractive faces (Rubenstein et al. 2002).

The affect-as-information model proposes that people use their feelings during judgment because they believe their feelings to be relevant to the judgment task (Schwarz 2002; Schwarz & Clore, 1983). Such feelings could stem from the product itself, with the product’s design acting as a form of mood manipulation (Norman 2004). A visually pleasing design could enhance mood, causing people to be more creative problem solvers and more willing to ignore details or small problems. An unattractive design could depress mood, leading people to be more analytical and heightening their tendency to expect and address problems. According to this logic, visually pleasing products could actually appear to work better.
Shiv and Fedorikhin (1999) demonstrated how an affective reaction to a stimulus can dominate a cognitive evaluation when cognitive resources are constrained. In accordance with Berkowitz (1993) the authors proposed that upon exposure to a choice, both affective and cognitive processes may be engendered; the affective reaction occurring relatively automatically and the cognitive processes occurring in a more controlled, resource-taxing manner. Due to this automatic nature of the affective reaction, when processing resources are constrained, only cognitive processing is impaired; thus, affective processes dominate.

The interplay of affect and cognition can be considered in the context of product aesthetics. A product’s design is quickly and easily perceived and may prompt an immediate affective reaction prior to a cognitive appraisal of product features. The affective reaction may be consistent or inconsistent with a later more conscious or controlled evaluation. The extent to which the later reasoned analysis would correct for the initial reaction is uncertain. The feeling-based reaction may be difficult to alter (Edwards 1990), or, in accordance with recent discussions of dual processes, the cognitive appraisal may be able to review and correct an initial assessment (e.g., Kahneman and Frederick 2002).

Motivated reasoning

Motivated reasoning is said to occur when a personal desire for a particular outcome interferes with the tendency or ability to evaluate choices objectively. The notion that motivations (hopes, fears, desires, etc.) can affect judgments has been studied and debated for decades (e.g. Ditto and Lopez 1992; Kruglanski 1990; Miller and Ross 1975; Pyszczynski and Greenberg 1987; Tetlock and Levi 1982). Although some have argued that what is called motivated reasoning can easily be explained entirely in non-
motivational, or purely cognitive terms based on prior beliefs or expectations (e.g., Miller and Ross 1975), the motivational camp continues to produce compelling evidence that motivation does play a role in judgment (e.g. Boiney, Kennedy and Nye 1997; Ditto and Lopez 1993; Markus and Kunda 1986). For example, in a study on business decisions, Boiney et al. (1997) reported that people tended to bias their judgments more or less as needed to support their desired conclusions. If people were acting in a purely cognitive manner, and simply coming to the most probable result given prior beliefs and expectations, it is unlikely they would have adjusted the degree of bias depending on what was needed to arrive at the desired conclusion.

In her review of the literature, Kunda (1990) suggested that motivation and cognition interact, with motivation affecting reasoning through reliance on a biased set of cognitive processes for accessing, constructing and evaluating beliefs. Kunda (1990) separated motivated reasoning into two categories: accuracy goals, for which the motive is to arrive at an accurate conclusion and directional goals, for which the motive is to arrive at a particular desired conclusion. She argued that both kinds of goals influence the choice of beliefs and strategies applied to a given problem. Accuracy goals lead to the use of beliefs and strategies considered most appropriate for the task; directional goals lead to the use of those that are considered most likely to result in the desired conclusion. Thus, as suggested by Kruglanski (1996) there is interplay of cognition and motivation: Motivation determines which cognitive processes will be used in a given situation.

In a product context, when one choice is aesthetically preferable to another there may be a motivation, perhaps below awareness, to under-critique the attractive product and over-critique the unattractive product, resulting in a choice consistent with the
aesthetic preference, but possibly non-normative. This type of motivated reasoning would be driven by a directional goal created by the product aesthetics. However, as noted by several authors (Boiney et al. 1997; Hsee 1995; Kunda 1990), people generally want to make rational, correct decisions; thus, their desires are constrained by reasonableness. The extent to which motivations influence judgments can be largely dependent on the strength and clarity of the information to be judged. When ambiguity is present motivation may play a larger role.

Ambiguity of product information. Hsee (1996) has reported that if decision criteria are somewhat ambiguous, people will use factors that they should not in order to achieve a motivated outcome. In his elasticity hypothesis (Hsee 1995; 1996; Schweitzer and Hsee 2002), Hsee defines two types of judgment factors, justifiable factors, which are relevant to the task and whose use can be justified, and non-justifiable factors, which are not relevant to the task but are personally appealing. When two options are presented, where one option is superior on a justifiable factor and the other is inferior on the justifiable factor but is superior on an unjustifiable factor, people’s choices will depend on the extent to which the justifiable factor’s values are difficult to evaluate. That is, when the values of the justifiable factor are fixed and clear, people will choose the option that is superior on that factor; however, when the level of the justifiable factor is not clearly indicated (e.g., is presented in a range), people will use the unjustifiable factor to choose the more personally attractive option. This occurs because without ambiguity in the justifiable factor, people have no excuse for considering the unjustifiable factor in their decisions. When the justifiable factor is presented in a range, people can distort their evaluation of this factor in the direction of the unjustifiable factor. This makes their
choice appear to be based on the justifiable factor when it is actually based on the unjustifiable factor.

Hsee (1995) demonstrated elastic justification in an experiment using a cover story of a proofreading assignment. Participants proofread ads and were paid on a per error basis. The ads were of two forms: furniture ads, which were uninteresting but had more errors per page, and personal ads, which were much more interesting, but had fewer errors per page, and participants could choose which type of ads to read. The dependent measure was the number of pages read from each type. When the number of errors per page was fixed, participants behaved normatively and chose to read the furniture ads. However, when the number of errors per page was presented in a range, proofreaders were more likely to choose the personal ads. Presumably, presenting the justifiable factor in a range made it easier to choose the more personally appealing but objectively inferior option.

Russo and his colleagues (Russo, Medvec, and Meloy 1996; Russo, Meloy, and Medvec 1998) used the term “predecisional distortion” to describe a form of biased reasoning that occurs when evaluating products on attributes presented in a sequence. The authors demonstrated that a preexisting preference for one option could lead to the distortion of ambiguous information to conform to the prior preference.

In the typical predecisional distortion paradigm, an initial preference for one of two products would be created through some kind of personal relevance manipulation. Thus, upon engaging in the evaluation task, subjects would already be motivated to choose a particular product. Relatively ambiguous product information would then be evaluated and (in most cases) a final choice would be made. Results indicated that feature
information was distorted to conform to the prior preference. As noted by the authors, predecisional distortion results are consistent with literature on confirmation bias, where people distort evidence to support an initial hypothesis or prior beliefs (e.g. Darley and Gross 1983; Deighton 1984; Hoch and Ha 1986; Nickerson 1998). Although Russo et al. (1996, 1998) support a cognitive explanation for predecisional distortion, they acknowledge that their original manipulation was motivational and cannot rule out a motivational explanation. An analysis of their evidence, in particular their second experiment (Russo et al. 1998), suggests that both motivational and cognitive factors were operating. The basic motivated reasoning process appeared to operate initially, with directional motivation influencing beliefs, but once the ambiguity was removed the motivational bias disappeared and a standard confirmation bias obtained.

In some cases the distortion of judgments based on prior beliefs can be justified from a Bayesian perspective (Koehler 1993; Lord, Ross and Lepper 1979; Russo et al. 1998). It may not always be easy to determine the diagnostic value of product attributes, particularly when they are ambiguous, so given uncertainty, it may be perfectly logical to use one’s prior belief in the superiority of one brand (or one’s preference for it) as part of an overall evaluation. Ignoring this prior information would contradict a Bayesian updating approach where prior beliefs are updated (but not necessarily discarded) when new information is presented.

**Individual differences.** Individual differences may influence the form and extent of motivated reasoning. The notion that some people are more oriented toward aesthetic qualities than others was originally suggested by Loewy (1951) and continues to permeate current aesthetics research (e.g. Bloch et al. 2003). Bloch et al. suggest that
people can be measured for their degree of centrality of visual product aesthetics (CVPA). CVPA is the overall level of significance that visual aesthetics hold for a particular consumer in his or her relationships with products. Consumers high in CVPA have greater concern for visual aesthetics and may place higher weight on aesthetics as a product dimension. These individuals assign high value to product appearance in enhancing their personal well-being. They perceive themselves to be better able to recognize or evaluate product designs, and they respond more strongly than others to aesthetic elements. The level of consumers’ CVPA may influence the extent to which they are directionally motivated by product aesthetics. Because these individuals experience stronger reactions to aesthetics, they may be more likely to engage in biased processing of other product attributes.

Another individual difference that may moderate the aesthetic effect is expertise. It is likely that for experts in a product category, the aesthetic appeal of the product does not exert a motivational influence because knowledge of functional product information is well established. Experts are less influenced by external factors and tend to process information more deeply (Alba and Hutchinson 1987). On the other hand, the motivated reasoning literature has demonstrated that people selectively retrieve information from memory that supports their desired goal (Kunda 1990), suggesting that when faced with an aesthetically pleasing product experts may selectively recall product information supporting their choice. Indeed, recent research in website atmospherics has demonstrated that product choices by both novices and experts are influenced by web page background primes (Mandel and Johnson 2002), so the effect of product design on experts is uncertain.
Halo effects

The halo effect, where global impressions of another person can unconsciously influence subsequent evaluations of that individual’s personality traits or abilities is a well-established phenomenon (e.g. Goldman, Cowles and Florez 1983; Nisbett and Wilson 1977). Research on halo effects has mostly focused on person perception, but it has been demonstrated that halo effects can and do exist for products (e.g. Beckwith and Lehman 1975; Holbrook and Huber 1979). Holbrook and Huber (1979) suggest that because of the complexity of some products’ attributes and the difficulty of calculating an overall evaluation using something like the Multiattribute method, people often work in reverse, letting their overall affective impression of the product guide their judgments of the product’s individual attributes. In terms of product aesthetics, an attractively or unattractively designed product could create a halo that would influence subsequent product evaluation.

Awareness of affective and motivational influences

Relatively little work has directly addressed the extent to which motivated reasoning occurs beyond conscious awareness, although a variety of findings suggest it does. For example, Boiney et al. (1997) found that the bias exhibited in motivated reasoning had no impact on judgment confidence; that is, biased decision makers were no more or less confident than unbiased decision makers. A lack of confidence could have suggested awareness of the bias. Other indirect evidence comes from Koehler (1993), who found that scientists were biased in evaluating the quality of research that agreed or disagreed with their prior beliefs. Despite displaying an agreement bias, the respondents believed they could evaluate the evidence in an unbiased manner and moreover, strongly believed that it was important to do so. This research was focused on prior beliefs rather
than directional motivations, but it does hint at a lack of awareness of the extent to which biased evaluation occurs.

Researchers in predecisional distortion have also claimed that awareness of the source of their bias is unlikely (Russo et al. 1998). They argued that in their paradigm, “if some of these sources of initial preference were recognized, consumers might dismiss them as unacceptable bases for brand comparison” (p. 449).

Halo effect studies also suggest that people are not aware they are being influenced. Nisbett and Wilson (1977) had college students evaluate a psychology instructor who was speaking in a videotaped interview. Half the subjects saw an interview where the instructor was warm and approachable, and half saw an interview where the instructor seemed cold and unfeeling. After watching the video, participants were asked to rate the instructor on various characteristics. Consistent with halo effects, participants rated the individual traits more positively for the warm teacher. However, the interesting finding is in regard to awareness. Most participants believed that their overall liking for the teacher had not influenced their ratings, and actually believed that, if anything, the reverse occurred, with the instructor’s traits influencing overall liking (Nisbett and Wilson 1977).

**Non-Aesthetic Influence of Product Appearance**

Independent of attractiveness, there is research to suggest that the extent to which a person looks appropriate for a role influences evaluations of objective criteria such as their resume or performance. In a study on physical traits and personal characteristics, Hassin and Trope (2000) found that people allowed inferences made on the basis of a person’s appearance to influence evaluations of the person’s suitability for particular jobs. The pictures used in the study controlled for attractiveness and rather focused on differences in physiognomy, or the extent to which a particular face looks like it would
possess particular character traits. Of course, such ratings are necessarily correlated with perceptions of attractiveness, but since attractiveness was controlled for during pretests, it cannot explain the results. Hassin and Trope found that people were generally unable to avoid using physical characteristics in evaluating personality or ability traits, despite believing that they could. In one study the authors specifically instructed participants to ignore accompanying photographs when evaluating their resumes as potential job candidates, and most (60%) indicated that they were able to do so. Nonetheless, experimental evidence showed that participants were influenced by the photographs, suggesting physical characteristics can play a role in evaluations beyond awareness and perhaps beyond attempts to correct for them.

Other studies on faces have demonstrated biasing effects of appearance independent of attractiveness. Mueller and Mazur (1996) report that male military officers who appeared dominant-looking (as defined by a high forehead and a square jaw) were more likely to be promoted within military ranks than those who appeared less dominant-looking. Todorov et al. (2005) found effects of appearance in the political real. The authors had study participants rate pictures of congressional candidates along seven personality dimensions, including competence, intelligence, charisma, trustworthiness, honesty, likeability, and leadership. They found that the competence ratings could predict with 68.8% accuracy the winner of the election.

In a product context, products can communicate a particular level of performance through their design. Certainly additional features visible to the eye would inform the consumer of product performance, but even when the number of features is equivalent across products, the design of those features can communicate different information.
Heavier construction materials are often equated with durability; size can be a signal of power; a sleek shape can suggest aerodynamics. Such design elements could influence the processing of functional product information in different ways. Performance-based product design could operate much like the effects observed in physiognomy experiments and distort the processing of other more reliable product information in a confirmation-bias type manner. Alternatively, the information communicated through design could provide quality cues that reduce the extent to which other product information is weighted.

The extent to which design would be used and perhaps overweighted as a decision cue may be a function of the extent to which it were perceived to be diagnostic to the judgment (Feldman and Lynch 1988; Slovic and Lichtenstein 1971). If the visual dimensions of the product, i.e., its design elements, communicate a particular level of quality, it might be relied upon at the expense of other cues. Unlike a halo or confirmatory processing effect, this suggests a relatively deliberate use of design, and in general this could be relatively normative. In the market there is often a correlation between the extent to which products looks like they will perform well and the extent to which they actually do, so from a cost/benefit perspective, using design as a quality heuristic may be perfectly reasonable. However, there are cases where appearance is not an accurate signal of quality, such as with copycat brands (Hutchinson and Alba 1991; Loken, Ross and Hinkle 1989). In such situations the question is whether consumers are able to disregard design and attend to truly diagnostic information.

Whether through distortion or a shift in dimension weights, the potential implications of performance-based design bias are dramatic. An example is the recent
decision by the US Department of Defense to choose Lockheed Martin over Boeing as the manufacturer of its new X-plane. At an estimated $225 billion, with another $175 billion in potential sales to foreign allies, the contract for the next-generation fighter jets was the largest defense contract in US history. This contract was awarded after a long competition between Boeing and Lockheed Martin. Both companies designed fighter jets to meet the requirements of the US Navy, such as acceleration and vertical lift capability. The clear performance criteria should have facilitated an unbiased comparative evaluation; however, the two companies achieved the Navy’s requirements in very different ways. One area in which they differed significantly was aesthetics. The Lockheed Martin plane looked aerodynamic and very close to the prototypical fighter jet shape. The Boeing plane had an unusual design that was described by a Pentagon source as looking like “a flying frog with its mouth wide open” (Prasso 2001). Although the Pentagon flatly denied that design played any role in the decision and there was no evidence that the Boeing plane was in fact functionally superior, the mere suggestion made by some industry experts (cited in Nova video 2003) that the Pentagon was influenced by the appearance of the planes implies the potential seriousness of design influence.

**Summary**

This literature review examined research highlighting different roles product appearance could play in information processing. The first section discussed ways in which a reaction to the aesthetic appeal of a product might bias product evaluation. Evidence from literature in halo and attractiveness effects, affect and cognition, and motivated reasoning seems to support the notion that product appearance could distort the processing of functional product information, impairing consumers’ ability to make
objective product judgments. On the other hand, literature focused specifically on product aesthetics has thus far suggested that aesthetic influence on quality judgments occurs primarily at the margins, influencing processing only in the absence of concrete information. As yet it is unclear under what conditions product aesthetics might bias quality judgments in the face of objective product information. Such influence may depend on several factors, including the strength of the affective reaction to the design, the strength of the functional product information, or individual differences.

The second section reported research suggesting that product appearance could have an effect on judgment independent of an aesthetic influence. Most of this work has focused on evaluations of people’s faces, where recent research has indicated that job performance by people who appear to fit a particular profile is judged differently than that of people who do not fit the profile. The same may be true of products. That is, design could appear to be a source of diagnostic information about the quality of the product, and as such could be relied upon, even in cases where it should not be.
CHAPTER 3
OVERVIEW OF EXPERIMENTS

Previous research has demonstrated that product design can be a cue to quality and that in the absence of other information such as brand name or product specifications, product design can influence choice (e.g. Creusen and Schoormans 2005; Page and Herr 2002). However, little product design research has been conducted on how design influences the processing of other product information, and most previous studies have found a fairly high degree of rationality in the use of design in decision making. There is evidence from research on facial appearance that appearance can bias evaluation of personality traits and performance; however, under what conditions this might in the product realm is unclear. It may be that people are deliberate in how they weight (or overweight) design, but given the wealth of research on how people are often unaware of how they use a variety of decision cues, it is plausible that design could exert a non-normative effect on decision making.

The studies presented in the next chapters extend the findings of the visual effects seen with faces into the product realm. The experiments investigate several questions: (1) does design bias information processing, (2) are there different types of design biases and do they exert differential influence, and (3) where in the decision process does design have the most/least influence.

The goal of the initial experiments was to determine whether design interferes with product quality judgments. To investigate this issue, the study employed a modified version of the predecisional distortion paradigm (Russo et al. 1996; 1998). In the studies
by Russo and his colleagues, participants were asked to compare the quality of two brands of backpacks. Prior to the evaluation task participants were given a personal reason to hope that one brand was superior to the other (i.e., an external motivation). They were then presented with descriptions of the product features for each brand and evaluated them one by one. Over time a bias toward their initial preference emerged, despite their being asked to evaluate each feature on an individual basis, rather than the backpack overall.

The pilot study and experiment 1 of the current investigation used a similar procedure. However, in place of an external motivation, pictures of the designs of the product were used to create the initial leaning toward one of the two brands. Prior to evaluating the feature descriptions, experimental participants saw pictures of the two brands, and one picture was significantly more attractive than the other. The subsequent written information suggested that the unattractive product was in fact superior. Comparing the ratings of participants who initially saw the pictures to the ratings of those who did not see the pictures made it possible to determine whether product appearance biased processing of feature information.

After establishing the presence of the product appearance bias, the remainder of the dissertation focuses on how the bias occurs and on some of the conditions of its influence. To do this, experiments 2 through 5 used a slightly different paradigm. The largest change was that rather than presenting the feature information one attribute at a time, in experiments 2 through 5 all attribute information was presented at one time in paragraph format. In studies 2 and 3 the primary comparison was still between ratings with pictures and ratings without pictures. In studies 4 and 5 comparison groups saw both
the pictures and the written information. In these experiments the key manipulation consisted of altering the order of presentation of the visual and verbal information. Comparing the evaluations of those who saw the pictures prior to the written reviews to the evaluations of those who saw the pictures after the reviews provided insight into the process underlying the design influence and helped to rule out some alternative explanations.
The pilot study used a modified version of the predecisional distortion paradigm (Russo et al. 1996) to investigate whether a product’s appearance leads to biased processing of product features. In Russo’s experiments, participants were told they would be evaluating two brands of backpacks and were given a personal motivation to hope that one brand was superior to the other. They were then presented with descriptions of the products’ features, one feature at a time. For each feature, they saw a description for each brand. Both descriptions were presented on the screen at the same time and participants indicated which backpack was favored based on each feature. After evaluating the two backpacks on the particular feature, participants were asked which backpack appeared to be superior overall, based on the information they had seen up to that point. This was referred to as the “horse race,” with the backpacks representing the horses. Participants indicated which “horse” was in the lead at that point.

The pilot study replicated the design of the Russo predecisional distortion studies. In place of a prior personal motivation to choose one of the brands, participants in the experimental condition were presented with pictures of the two brands before reading the feature information. One of the brands was significantly more attractive than the other.

Two product categories were used: backpacks, to mirror the Russo studies, and running (racing) shoes. Running shoes were chosen as a second category because they represented a product where looks should not matter to a serious runner who is shopping for a performance shoe.
Method

Pretest

Pictures of backpacks and running shoes were pretested to obtain pairs of brands in which one brand was significantly more attractive than the other. To prepare for the pretests, 24 pictures of backpacks and 24 pictures of running shoes were taken from internet websites. Using Photoshop, the pictures were manipulated to remove brand labels and to make them uniform in size. Two rounds of pretest ratings were conducted. First, 20 participants rated the backpacks individually on a scale from 0 (very unattractive) to 100 (very attractive). The three most attractive and unattractive products were selected. Next, 30 participants rated the relative attractiveness of the nine possible combinations of pairs of unattractive and attractive products, and the pairs with the largest difference were selected. Pictures of the final pairs of backpacks and running shoes appear in appendix a.

In terms of feature descriptions, the goal was to have one product be superior to the other overall, but not overwhelmingly so. For each feature, two descriptions, one for each brand, were created and pretested. For backpacks, two of the actual feature descriptions used by Russo were selected to mirror his procedure as closely as possible. Pretests resulted in stimuli lists that when considered overall, favored one brand over the other, but not always when the features were considered individually. The unattractive picture was then paired with the superior product list, and the attractive picture with the inferior list.

Because the attribute lists did not favor the unattractive product on every feature, the order of presentation of the attributes was controlled so that initially there was some ambiguity as to which product was superior. Forcing a single attribute order could have
resulted in generalizability concerns but for the purposes of the pilot study it was considered worthwhile. The feature lists and order of presentation appear in table 4-1.

**Design and Procedure**

The study had two between subjects conditions, no-picture and picture. There were two product replicates, backpacks and running shoes (order counterbalanced). Participants in the no-picture condition saw only the written feature information. Picture participants saw the feature information along with pictures of the two brands, which remained on the screen throughout the procedure.

The study was conducted on computer. All information was provided in written form on the computer screen. It was important to encourage participants to focus on quality rather than aesthetics, so before presentation of the feature information, specific instructions were provided. For the backpacks, participants were told that they were shopping for a new backpack to replace one they had bought previously. The previous one had been poor quality and this time they were determined to get a better backpack. They had narrowed the search down to two models, both of which had the basic features they were looking for, so their decision had come down to determining which pack was better quality.

For running shoes, participants were told to imagine that they were serious, experienced runners who regularly participated in road races. They were to imagine that they were buying shoes for an upcoming race and had narrowed the search to two models, both of which were made for their type of foot and both of which were equally comfortable. Their decision had come down to determining which shoe was better quality.
After the introductory screen, participants received feature information about the
two available brands, one feature at a time. Each feature had a description for brand A
and one for brand B. For example, for backpacks, the first feature described was the
durability of the material, and the two descriptions were the following:

Brand A: Has a layer of leather sewn inside the bottom of the pack to enhance its
integrity. In addition, a layer of water repellant material has been sewn between the
interior layer and the exterior.
Brand B: Double-stitched and double-lined for durability. In addition the material has
been chemically treated to enhance its strength and water repellency while retaining its
suppleness.

Both descriptions appeared on the screen at the same time, one on each side of the
screen (counterbalanced). For picture participants, the corresponding picture of the
backpacks appeared above the feature descriptions. Participants read both descriptions
and evaluated which brand was favored based on the two descriptions using a 9-point
scale where 9 = “description favors Brand A” (the objectively superior but less attractive
brand) and 1 = “description favors Brand B” (the objectively inferior but more attractive
brand). Consistent with Russo’s horse race paradigm, immediately after rating the
favorability of a given feature, participants indicated which brand they thought was
leading based on all feature information viewed up to that moment. This was a binary
choice. They then saw the next feature description and completed the same two tasks.
This continued until all five features were presented. After seeing and rating all five
features, participants were asked, “You have now seen all the differentiating features of
the two [backpacks/running shoes]. Considering all the information you have seen, which
[backpack/running shoe] do you think is better quality?” Participants indicated their
response to this binary choice. They were then asked how confident they were that they
had selected the objectively superior brand and indicated their response on a 9-point scale
from “not at all confident” to “extremely confident.” After providing their confidence rating, they then completed the entire task again for the second product category. Finally they then completed a product knowledge question and answered some general demographic questions. One hundred thirty-eight undergraduate students participated in the study in exchange for course credit.

**Results**

**Backpack Ratings**

Eighteen participants failed an attention check and were removed, leaving 120 in the analysis. A significant order effect of replicate required that only the first product replicate for each participant be analyzed. Table 4-2 presents the mean feature ratings for backpacks. Higher numbers indicate the belief that the information supports the objectively superior (less attractive) item. There is a significant effect of the picture, $F(1, 41) = 5.64, p < .05$. There is also an unsurprising effect of attribute, $F(4, 164) = 9.69, p < .001$, and no interaction, $F < 1$.

After each attribute rating, participants indicated which brand they believed was leading using a binary choice measure. Table 4-3 presents the proportion of participants choosing the objectively superior backpack as the leader after the presentation of the respective attribute. A Chi-square test of independence was conducted on each measure. The third and fifth measures were significant ($p$-values < .05), and the fourth measure was directional, but not significant ($p = .10$)

On the final choice measure, participants who did not see the picture selected the superior backpack 78% of the time (18 out of 23), whereas those who saw the pictures selected the superior (less attractive) backpack 40% (8 out of 20) of the time, $p < .05$. 
Note that this question was essentially identical to the horse race question after the fifth feature, so the choice proportions are the same as those of the fifth horse race question.

After completing the final backpack choice, participants were asked how confident they were that they had selected the higher quality backpack. On a 9-point scale, no-picture participants reported confidence of 6.35 and picture participants reported confidence of 6.20, indicating no difference in confidence as a function of the presence of the pictures, $t < 1$.

**Running Shoe Ratings**

Table 4-4 presents the mean ratings for the running shoe features. There is a marginal effect of the picture, $F(1, 75) = 3.98, p < .06$, but in contrast to the backpack results, the effect is in the reverse direction. Participants seeing the pictures rated the feature descriptions as more supportive of the superior shoe than did participants who did not see the pictures.

After each feature rating, participants indicated which brand they believed was leading using a binary choice measure. Table 4-5 presents the proportion of participants choosing the objectively superior shoes as the leader after the presentation of the respective attribute. Similar to the feature ratings, they are all in the opposite direction. There is a significant effect in the reverse direction on feature 1, $\chi^2 = 5.74, p < .05$, and a directional difference on feature 5, $\chi^2 = 2.64, p = .10$. For the final choice, people who did not see the picture selected the superior running shoe 59% of the time (23 out of 39), while people who saw the pictures selected the superior shoe 76% (29 out of 38) of the time.

After completing the final choice measure, participants were asked how confident they were that they had selected the higher quality running shoe. On a 9-point scale, no-
picture participants reported confidence of 6.54 and picture participants reported confidence of 5.95. This difference was not significant, $t(75) = 1.48, p > .14$.

**Discussion**

The bidirectional effect of the pictures across the two products is perplexing. The backpack results suggest an attractiveness effect, a finding that is consistent with research in facial attractiveness. The question is why the shoe results went the other way. One possibility is that the pictures of the shoes differed on some additional dimension besides attractiveness. In the face perception literature, physical characteristics other than attractiveness have been found to influence judgments of competence, job suitability, and job performance (Mueller and Mazur 1996; Todorov et al. 2005). Such characteristics have been shown to be extremely difficult to ignore (Hassin and Trope 2000). If physical characteristics other than attractiveness can signal individual performance, it is plausible that design elements other than aesthetic ones could signal a certain level of product performance. Perhaps the unattractive shoe appeared superior along one or more performance dimensions, and participants used this information in their evaluations. If so, it may be that this appearance of performance dominated the aesthetic difference and biased the processing of the feature descriptions in the opposite direction. This possibility is explored in experiment 1.

Another possible explanation for the surprising running shoe results is a problem with the stimuli. For the backpacks, none of the feature descriptions could be confirmed with a visual comparison of the brands. However, with the shoes, several of the verbal attributes mentioned features such as “mesh upper” or “striped rubber” could be confirmed or disconfirmed by the pictures. It may be that participants did not believe the
written descriptions because the pictures contained conflicting information. In experiment 1 the written descriptions were altered to avoid this problem.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Brand A (superior)</th>
<th>Brand B (inferior)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Backpacks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Durability</td>
<td>Has a layer of leather sewn inside the bottom of the pack to enhance its integrity. In addition, a layer of water repellent material has been sewn between the interior layer and the exterior.</td>
<td>Double-stitched and double-lined for durability. In addition the material has been chemically treated to enhance its strength and water repellency while retaining its suppleness.</td>
</tr>
<tr>
<td>2. Exterior material</td>
<td>Exterior material is made from microfiber and nylon, giving the bag both flexibility and durability.</td>
<td>Exterior material is a combination of duralite and nylon for strength with the ability to stretch.</td>
</tr>
<tr>
<td>3. Compartments</td>
<td>The main compartment is large enough to fit books or a laptop computer. Smaller interior pockets are organized for keys, pens, etc.</td>
<td>Has a large main compartment, plus smaller compartments and several accessory pockets inside to store small items.</td>
</tr>
<tr>
<td>4. Carrying comfort</td>
<td>Contoured load lifter shoulder straps ease the weight on the shoulders.</td>
<td>Dual-density foam in the shoulder straps increases carrying comfort.</td>
</tr>
<tr>
<td>5. Zipper</td>
<td>Has a heavy-weight reinforced zipper treated with lubricant for smoother running. In addition a flap of material ties down over the top of the backpack</td>
<td>Has a plastic zipper, coated with a thin layer of black plastic for smoother running.</td>
</tr>
<tr>
<td><strong>Running Shoes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Upper material</td>
<td>Ventilated upper and midsole with micro mesh.</td>
<td>Mesh used throughout for maximum breathability.</td>
</tr>
<tr>
<td>2. Weight</td>
<td>Lightweight to optimize running efficiency.</td>
<td>Lightweight while maximizing energy transfer properties.</td>
</tr>
<tr>
<td>3. Midsole</td>
<td>Low-to-the-ground midsole profile, 3D Ultralyte one piece a stable two-density midsole, and a midsole/outsole material for flexible, lightweight upper and outsole. Midsole and outsole components are engineered to set the foot up in an efficient, balanced position from heel strike to toe-off, offering stability and flexibility.</td>
<td>3D Ultralyte one piece midsole/outsole material for superior fit. Midfoot Torsion support device that adds medial support and stability.</td>
</tr>
<tr>
<td>4. Outer Sole</td>
<td>High-density carbon rubber increases traction and extends the life of the shoe.</td>
<td>Striped rubber outside for grip and durability.</td>
</tr>
<tr>
<td>5. Interior comfort</td>
<td>Mesh inner sleeve that stretches to adjust to your foot over long runs</td>
<td>Mesh upper and a lacing system that wraps around the arch for comfort and support.</td>
</tr>
</tbody>
</table>
Table 4-2. Pilot Study Backpack Feature Ratings

<table>
<thead>
<tr>
<th>Condition</th>
<th>Durability</th>
<th>Exterior Material</th>
<th>Compartments</th>
<th>Carrying Comfort</th>
<th>Zipper</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>4.65</td>
<td>5.52</td>
<td>6.48</td>
<td>5.78</td>
<td>7.22</td>
</tr>
<tr>
<td>Picture</td>
<td>3.70</td>
<td>4.85</td>
<td>4.85</td>
<td>4.10</td>
<td>6.40</td>
</tr>
</tbody>
</table>

Table 4-3. Pilot Study Choice Proportions of Superior Backpack

<table>
<thead>
<tr>
<th>Condition</th>
<th>After First Attribute</th>
<th>After Second Attribute</th>
<th>After Third Attribute*</th>
<th>After Fourth Attribute*</th>
<th>After Fifth Attribute**</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>30%</td>
<td>39%</td>
<td>70%</td>
<td>70%</td>
<td>78%</td>
</tr>
<tr>
<td>Picture</td>
<td>30%</td>
<td>40%</td>
<td>35%</td>
<td>45%</td>
<td>40%</td>
</tr>
</tbody>
</table>

* p = .10
**p < .05

Table 4-4. Pilot Study Running Shoe Feature Ratings

<table>
<thead>
<tr>
<th>Condition</th>
<th>Upper Material</th>
<th>Weight</th>
<th>Midsole</th>
<th>Outer Sole</th>
<th>Interior Comfort</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>3.67</td>
<td>4.38</td>
<td>5.62</td>
<td>6.90</td>
<td>5.41</td>
</tr>
<tr>
<td>Picture</td>
<td>4.37</td>
<td>4.47</td>
<td>6.82</td>
<td>7.11</td>
<td>6.08</td>
</tr>
</tbody>
</table>

Table 4-5. Pilot Study Choice Proportions of Superior Running Shoe

<table>
<thead>
<tr>
<th>Condition</th>
<th>After First Attribute**</th>
<th>After Second Attribute</th>
<th>After Third Attribute</th>
<th>After Fourth Attribute</th>
<th>After Fifth Attribute*</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>31%</td>
<td>38%</td>
<td>54%</td>
<td>64%</td>
<td>59%</td>
</tr>
<tr>
<td>Picture</td>
<td>58%</td>
<td>50%</td>
<td>68%</td>
<td>71%</td>
<td>76%</td>
</tr>
</tbody>
</table>

* p = .10
**p < .05
CHAPTER 5
EXPERIMENT 1

In the pilot study the superior shoe was pretested to be less attractive, but may have looked like it would perform better, which consequently influenced feature ratings. That is, rather than an attractiveness bias, there was a performance-based design bias. To address this possibility, experiment 1 used pictures that were pretested to be equally attractive, but significantly different on the extent to which they looked like high or low quality products. In addition, the written descriptions were altered so that a close inspection of the pictures could not confirm or disconfirm the written feature information.

Method

Stimuli

The experiment was conducted using running shoes only. The written descriptions of the attributes were altered to contain only information that could not be assessed visually. Table 5-1 presents the revised feature descriptions. New running shoe pictures were pretested to create a pair of brands that were equally attractive (56.63 and 61.29 on a 0-100 scale, \( p > .73 \)) but appeared significantly different in terms of the extent to which they looked like superior performing running shoes (relative performance rating of -37.22 on a scale from -100 to +100, where 0 means the shoes are equivalent, \( p < .002 \)). See appendix b for pictures of the shoes used in the study.

Design and Procedure

In contrast to the pilot study, experiment 1 had three between subjects conditions. The no-picture condition from the pilot study was retained as the control group. In
addition there were two levels of the picture condition: one from the pilot study, where the inferior picture was paired with the superior attributes (hereafter called the inferior picture condition), and a new condition in which the superior picture was paired with the superior attributes (hereafter called the superior picture condition). In all other respects the design and procedure mirrored the pilot study. Ninety-three undergraduate students participated in the study in exchange for course credit.

Results

Twenty participants were removed for failing an attention check measure. Table 5-2 presents the mean feature ratings for the remaining 73 participants by condition. As in the pilot study, there is an effect of the picture, $F(2, 70) = 4.04, p < .05$ but no picture x attribute interaction, $F < 1$. There is a picture bias in the inferior picture condition. People are biased toward the superior looking shoe when it is paired with inferior features. There was no bias in the opposite direction when the high performance shoe was paired with good features. Contrasts performed on the data revealed that the feature ratings in the no-picture control condition were significantly higher than the inferior picture condition for features 1, 2 and 4 (one-tailed $t$-tests, $p < .05$), and marginally higher on feature 5 ($p < .06$). There were no significant differences between the control condition and the superior picture condition, although feature 2 was marginally significant in a direction inconsistent with expectations ($p < .10$).

The proportions of participants who chose the objectively superior shoe as leading after the presentation of each feature are presented in table 5-3. None of the choice shares were significantly different from one another, but across all five choice opportunities, the inferior-picture choice proportion was lower than the no-picture condition. Moreover, in
three of the five choices, the superior-picture choice proportion was higher than the control condition, suggesting some influence of the pictures in both directions.

As in the pilot study, there was no difference in how confident participants were that they had selected the superior shoe, $F < 1$. On a one to nine scale where 1 = “not at all confident” and 9 = “completely confident,” the mean ratings were 6.18, 5.73, 5.52 for the no-picture, superior-picture, and inferior-picture conditions respectively.

**Discussion**

Experiment 1 demonstrated a performance-based design bias when controlling for attractiveness. When the objectively inferior shoe was paired with a picture that suggested superior performance, participants’ ratings were biased in the direction of that shoe. These results support the contention from the pilot study that the bias in the direction of the unattractive shoe was the result of its appearing to be functionally superior.

When the superior-looking shoe was paired with superior features, no bias was observed in the direction of the superior shoe for the feature ratings. Nonetheless, the choice shares, while not significant, were directionally supportive of an inclination toward the superior shoe beyond that of the control condition.

The feature ratings are interesting because there is no justifiable reason to rely on the picture when rating each feature. However, the use of the Russo paradigm made the participants’ major goal to find the superior shoe overall. Thus, it is possible that the picture was distorting the ratings in an overall halo-like fashion, particularly when the features were relatively ambiguous, rather than in a specific feature-by-feature fashion. Moreover, the pictures were pretested to ensure that one shoe appeared superior overall, but it is possible that it might have looked superior only on certain dimensions. This lack
of specificity may have influenced results. To focus more narrowly on individual
dimensions, subsequent experiments used products that were pretested to dominate only
on one specific dimension rather than overall.

Predecisional distortion is a form of confirmation bias, where sequentially
presented information is distorted in a desired direction. As acknowledged by Russo and
his colleagues (Russo, Medvec, and Meloy 1998), this type of sequential rating format
can lead to bias even for a control condition. To better isolate a design effect and ensure
that it is not an artifact of the experimental format, experiment 2 moves away from the
Russo paradigm and focuses on a single feature rating.

When considered in conjunction with the running shoe results from the pilot study,
the results of the current experiment call into question the backpack results observed in
the previous experiment. If the design effect for running shoes was due to the appearance
of performance, could it have been that the backpack results were also due to a
performance-based design bias rather than an attractiveness bias? The backpack pretests
did not control for the appearance of performance, so it may have been that the more
attractive backpack was also the one that appeared superior on performance. To examine
this possibility, a follow-up test was conducted on the backpack stimuli measuring the
correlation between perceived performance and perceived attractiveness. There was a
marginally significant positive correlation between attractiveness and performance, $r =
.24$, $p = .10$; thus, it is impossible to conclude that the results for the backpacks in the
pilot study were due entirely to attractiveness. Because there is no strong evidence of
biased processing due to aesthetics, the next two experiments focus only on the
appearance of performance. However, the issue of aesthetic influence remains
unresolved, and in experiment 4 the potential for an attractiveness bias is reexamined in a
different context.

Table 5-1. Running Shoe Feature Descriptions Used in Experiment 1

<table>
<thead>
<tr>
<th>Feature</th>
<th>Brand A (superior)</th>
<th>Brand B (inferior)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Weight</td>
<td>Lightweight to optimize running efficiency.</td>
<td>Lightweight while maximizing energy transfer properties.</td>
</tr>
<tr>
<td>2. Midsole material</td>
<td>Ethylene Vinyl Acetate air infused foam midsole provides superior cushioning without sacrificing durability.</td>
<td>Midsole is made of high quality polyurethane foam, which is dense and stable.</td>
</tr>
<tr>
<td>3. Midsole Design</td>
<td>Midsole components are designed to keep the foot in the most efficient position, maintaining balance but allowing it to flex naturally.</td>
<td>Midsole components are compression molded, with solid construction that ensures that the shoe will maintain its structure and shape over time.</td>
</tr>
<tr>
<td>4. Outer Sole</td>
<td>Durable high-density carbon rubber outsole extends the life of the shoe.</td>
<td>Blown rubber traction enhancing outsole increases grip on softer surfaces.</td>
</tr>
<tr>
<td>5. Interior comfort</td>
<td>Double strike insole is made with orthotic-type material that adjusts to the foot shape over time to maximize shock absorption.</td>
<td>Well cushioned insole provides good arch support.</td>
</tr>
</tbody>
</table>

Table 5-2. Experiment 1 Running Shoe Feature Ratings

<table>
<thead>
<tr>
<th>Condition</th>
<th>Upper Material</th>
<th>Weight</th>
<th>Midsole</th>
<th>Outer Sole</th>
<th>Interior Comfort</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>5.68</td>
<td>7.68</td>
<td>6.18</td>
<td>5.45</td>
<td>7.73</td>
</tr>
<tr>
<td>Superior Picture</td>
<td>5.23</td>
<td>6.88a</td>
<td>5.85</td>
<td>5.08</td>
<td>7.38</td>
</tr>
<tr>
<td>Inferior Picture</td>
<td>4.40a</td>
<td>6.84a</td>
<td>5.92</td>
<td>4.24a</td>
<td>6.96</td>
</tr>
</tbody>
</table>

*aCell mean differs from control group mean at p < .05 significance.

Table 5-3. Experiment 1 Choice Proportions of Superior Running Shoe

<table>
<thead>
<tr>
<th>Condition</th>
<th>After First Attribute</th>
<th>After Second Attribute*</th>
<th>After Third Attribute</th>
<th>After Fourth Attribute</th>
<th>After Fifth Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Picture</td>
<td>45%</td>
<td>82%</td>
<td>73%</td>
<td>59%</td>
<td>91%</td>
</tr>
<tr>
<td>Superior Pic</td>
<td>58%</td>
<td>92%</td>
<td>88%</td>
<td>54%</td>
<td>85%</td>
</tr>
<tr>
<td>Inferior Pic</td>
<td>36%</td>
<td>68%</td>
<td>82%</td>
<td>56%</td>
<td>80%</td>
</tr>
</tbody>
</table>

*p < .10
EXPERIMENT 2

Experiment 2 marks a small shift in the format of the studies, with four key changes. First, based on the findings of experiment 1, experiment 2 focused solely on a visual difference in performance, rather than a visual difference in attractiveness.

Second, whereas the pilot study and experiment 1 made use of a sequential presentation format, with feature information presented one at a time, in experiment 2 all product information was presented at once in paragraph format. Participants saw two paragraphs on the screen, one describing each brand. Each paragraph provided information on several product features. In addition, for participants in the picture condition, the written information was presented simultaneously with the picture. These changes eliminated the possibility of distortion occurring over time in a confirmation bias manner.

Third, to make the visual difference in performance as clear and as specific as possible, the study focused on only one dimension per product. Pictures were chosen to reflect superiority along one given dimension, rather than overall quality, and participants were only asked to rate the product on that single dimension. The corresponding paragraphs suggested that the inferior-looking brand was in fact superior on the given dimension.

Finally, to broaden generalizability, and to avoid some of the difficulties of the stimuli from the pilot study, experiment 2 and all subsequent experiments used five different product categories: cookware, stereo speakers, in-line skates, bicycle helmets,
and electric mixers. These categories were selected because they were products for which appearance could potentially provide information about performance, but could not unequivocally signal high or low quality.

Method

Stimuli

Pretests were conducted to select pictures and feature information for the new set of products. The goal for each pair of pictures was to have one product that appeared superior along a specific dimension, yet appeared equally or less attractive than the other product. Several pictures for each of the five product categories were selected for pretesting. Twenty-four participants rated pairs of pictures from the five categories on relative attractiveness and superiority on the key performance dimension. The question order was counterbalanced across participants. The measures were seven point scales from “Brand A is far superior” to “Brand B is far superior.” One-sample $t$-tests were conducted on the attractiveness and dimension ratings. Products chosen were those for which the attractiveness rating was not significantly different from the mid-point rating (“pictures are equally attractive”) but the performance dimension was significantly different from the mid-point. The pretest led to the selection of one pair of pictures for each product category. The final picture pairs selected appear in appendix c, and the ratings for attractiveness and performance along the focal dimension, mean centered for ease of comparison, appear in table 6-1.

Paragraph descriptions of product features were written in the format of an expert review, supposedly by Consumer Reports. The written information was pretested so that on the focal dimension, one review was significantly better than the other. The reviews were evaluated in pairs on a seven-point scale where $1 = “Brand A is far superior”$ and 7
“Brand B is far superior.” In addition care was taken to ensure that the focal dimension could not be confirmed by a visual inspection of the pictures. The product categories, key dimension, and pretest ratings, mean centered, appear in table 6-2. The paragraph descriptions used in the study appear in appendix d.

**Design and Procedure**

Experiment 2 had two between subjects conditions (picture and no-picture) and five product replicates. For picture participants, the written descriptions that indicated the product was superior along the key dimension were matched with the picture that suggested the product was inferior on the dimension.

The experiment was conducted on computer. Participants saw an introductory screen that informed them they would be reading Consumer Reports reviews for pairs of brands in five different product categories and would be asked to evaluate the quality of the products based on the reviews. After this introductory screen, they were presented with the first product category, in a randomly selected order, and saw two paragraphs side by side on the screen (counterbalanced), one describing each brand. Participants in the picture condition also saw the two pictures above the corresponding reviews. After the reviews had been on the screen for 40 seconds, participants were asked the degree to which one product or the other was superior along the key dimension. This was indicated on a 9-point scale from 1 (Brand A is far superior on this dimension) to 9 (Brand B is far superior on this dimension). A rating of 9 would mean that the objectively superior brand, based on the written information, was superior. To avoid any evaluative inferences related to the scale points the numbers on the scale were blacked out so that participants only saw nine radio buttons ranging from “Brand A is far superior” to “Brand B is far superior.” The picture and written information did not disappear until after subjects
provided their rating. The next product category then appeared on the screen and the process was repeated. After providing ratings for the five product categories, participants were given an unrelated filler task that lasted approximately ten minutes. Upon completion of the filler task a manipulation check was conducted. Participants were shown the picture pairs again. They were asked to imagine they had not read anything about the products previously and, based solely on a visual inspection of the products, to indicate which looked superior on attractiveness and on the focal dimension. Of course, despite the filler task, a consistency bias was a serious concern for the picture participants, so the usefulness of the manipulation check was somewhat limited. After the manipulation check, participants were asked questions regarding their knowledge of the products, as well as some demographic information. They were then thanked and dismissed. Eighty-three undergraduate students participated in the study in exchange for course credit.

Results

The mean ratings for the five products appear in table 6-3. Higher ratings indicate that the objectively superior brand according to the reviews was rated as superior. As the table indicates, both groups were relatively accurate; both were able to differentiate and tell which review was superior. Nonetheless, the significant effect of picture, $F(1, 81) = 4.28, p < .05$, indicates that the picture did have an effect. Across all replicates, the mean rating for the picture group was lower than the mean rating for the no-picture group.

The manipulation checks confirmed that the stimuli satisfied the criteria for the study. A rating of 0 would indicate that the two pictures appeared equal. A negative value would indicate that the picture pretested to be superior along the focal dimension was perceived as superior, and a positive value would indicate that the picture was perceived
as inferior. Thus, a successful manipulation would have attractiveness ratings equal to or greater than 0 and performance ratings significantly less than 0. Table 6-4 presents the mean ratings for the manipulation check. All five picture pairs had significant differences in the desired direction for performance and had either no differences or differences in the reverse direction for attractiveness.

Because participants in the picture condition might have been prone to consistency bias in their picture ratings, the manipulation check was performed a second time with only no-picture participants. The results mirrored those that included all participants.

**Discussion**

The results of experiment 2 suggest that people are using the pictures in their evaluation of the products. However, the reason for this is still unclear. It may be that participants had an initial reaction to the pictures that distorted subsequent evaluation of product dimensions. This is consistent with the results of pilot study and experiment 1. In experiment 2, all visual and verbal information was presented simultaneously, which could minimize the effects of an initial reaction to the visual stimuli. However, a visual reaction to design is usually rapid and relatively automatic (Page and Herr 2002), and it is highly unlikely that participants read the paragraphs before looking at the pictures. Thus, the notion that the effect is a function of an initial reaction influencing subsequent evaluation in a halo-like manner is still a real possibility. If the result were due to a halo effect, then the reaction to the visual information should carry over to other dimensions on which the picture does not dominate.

Another possibility is that the pictures are being relied upon because they appear diagnostic of quality. If this is the case, then if participants were asked about dimensions on which the superior picture does not dominate, there should be no effect of the picture.
Experiment 3 addresses these two possibilities by replicating the format of experiment 2 and adding questions about two non-focal dimensions.

Table 6-1. Pretest Results for Performance Pictures

<table>
<thead>
<tr>
<th>Category</th>
<th>Attractiveness Rating</th>
<th>Key Dimension</th>
<th>Performance Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>-0.37</td>
<td>Ease of cleaning</td>
<td>1.08*</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>0.54</td>
<td>Power</td>
<td>1.08**</td>
</tr>
<tr>
<td>In-line skates</td>
<td>-0.33</td>
<td>Speed</td>
<td>1.13**</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>0.50</td>
<td>Comfort</td>
<td>1.08***</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>0.13</td>
<td>Power</td>
<td>2.00***</td>
</tr>
</tbody>
</table>

*p < .05
**p < .01
***p < .001

Table 6-2. Pretest Results for Feature Descriptions

<table>
<thead>
<tr>
<th>Category</th>
<th>Key Dimension</th>
<th>Dimension Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>Ease of cleaning</td>
<td>-2.08**</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>Power</td>
<td>-1.71*</td>
</tr>
<tr>
<td>In-line skates</td>
<td>Speed</td>
<td>-2.62**</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>Comfort</td>
<td>-2.87**</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>Power</td>
<td>-2.71**</td>
</tr>
</tbody>
</table>

*p < .07
**p < .01

Table 6-3. Results of Experiment 2

<table>
<thead>
<tr>
<th>Category</th>
<th>Picture N=43</th>
<th>No Picture N=40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>7.30</td>
<td>7.58</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>7.30</td>
<td>7.45</td>
</tr>
<tr>
<td>In-line skates</td>
<td>6.60</td>
<td>7.83</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>6.86</td>
<td>7.55</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>7.93</td>
<td>8.18</td>
</tr>
<tr>
<td>Average</td>
<td>7.20</td>
<td>7.72</td>
</tr>
<tr>
<td>Category</td>
<td>Attractiveness Rating</td>
<td>Performance Rating</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Cookware</td>
<td>0.07</td>
<td>-1.04***</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>1.46*</td>
<td>-0.83*</td>
</tr>
<tr>
<td>In-line skates</td>
<td>0.99*</td>
<td>-1.80***</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>-0.60</td>
<td>-1.02***</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>0.61</td>
<td>-2.49***</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$
Experiments 1 and 2 demonstrated a performance-based design bias, but the nature of the effect was unclear. Experiment 1 suggested that people had a general reaction to the picture that colored their evaluation of features. Experiment 2 focused on one dimension only and confirmed the existence of a design bias, but did not help to clarify whether it was due to a halo effect of the picture or a result of the picture appearing diagnostic of quality on the focal dimension. In experiment 3 participants again saw all the product information on the screen at one time and again evaluated dimensional performance. However, in this case, participants evaluated the products along three dimensions, the focal dimension on which the products differed visually and two additional dimensions on which the products did not differ visually.

**Method**

**Stimuli**

The product pictures from experiment 2 were used in experiment 3. Two non-focal dimensions were added for each product, and feature descriptions were altered slightly to provide some information about the non-focal dimensions. The additional dimensions appear in table 7-1, and the revised paragraphs appear in appendix e.

**Design and Procedure**

The design was similar to experiment 2. There were two between subjects conditions, picture and no picture. There were five product replicates, but in this study
each replicate was accompanied by three dimension questions, one about the focal dimension and two about non-focal dimensions.

Participants were given the same initial instructions as in experiment 2. They were then presented with each product replicate, in random order. For each replicate, paragraph descriptions of the two brands appeared on the screen, position counterbalanced. For picture participants, a picture of each brand accompanied the descriptions. After 40 seconds the first of the three dimension questions appeared. The order of presentation of the three questions was determined using a Latin Square. Once participants responded to the first question, it disappeared and was replaced by the second and then the third question. As in experiment 2, the pictures and descriptions remained on the screen until all three questions were answered, at which time they disappeared and were replaced by the next product replicate. After answering the dependent measures for all five replicates, participants completed a filler task and responded to the manipulation check questions. Finally, they answered product knowledge and demographic questions, and were dismissed. Eighty-nine undergraduate participants completed the study.

**Results**

Eight participants failed an attention check measure and were removed, leaving 81 in the analysis. Mean ratings for the focal and non-focal dimensions appear in table 7-2. The dimension x picture interaction was not significant but was directionally supportive, $F(2, 158) = 2.19, p < .12$. However, a manipulation check on the focal dimensions revealed that for this group of participants, bicycle helmets did not constitute a successful manipulation of perceived performance on the key dimension. In contrast to the other four replicates, the two helmets did not differ significantly in the extent to which they
appeared comfortable for this group of participants. The manipulation check ratings, mean centered, appear in table 7-3.

Dropping helmets from the analysis and rerunning it on the remaining four replicates resulted in a significant interaction of dimension and picture, $F(2, 158) = 3.16, p < .05$. For completion there was also a significant effect of replicate, $F(2, 158) = 166.39, p < .001$ and no effect of picture, $F(1, 79) = 2.47, p > .05$. The overall means with helmets removed appear in table 7-4. Follow up contrasts on the revised means revealed that the difference between the picture and no-picture conditions was significant for the focal dimension ($p < .05$) but not for either of the non-focal dimensions.

**Discussion**

Experiment 3 was designed to determine whether the picture effect observed in previous studies was the result of a general halo or the result of the picture being perceived as diagnostic. The results of experiment 3 indicate the latter. The pictures provided information about the focal dimension but not about the other dimensions, and the picture only had an effect on the ratings of the focal dimension. For the focal dimensions, the presence of the picture lowered the ratings of the strength of the verbal information, but for non-focal dimensions, the presence of the picture had no effect. This suggests that participants used the pictures as a source of information only when they perceived them to be diagnostic. The helmet replicate did not conform to this pattern, but it also did not pass the manipulation check for this group of subjects.

Thus far, all the studies have compared information processing in the presence of visual information to processing in the absence of visual information. However, with such designs the pictures are an additional source of information, so assuming an averaging model it is not entirely surprising that there was an effect of the picture. A
A stronger test of the use of visual information would be to compare conditions that have the same information but vary the order in which the information is presented. Experiments 4 and 5 involved such tests.

In addition, although it appeared that the pictures were not operating as halos and coloring evaluation of all features, it may still have been that their influence was the result of an initial reaction to the key dimension. Since the pictures were pretested to differ significantly along one dimension, a relatively automatic reaction to this difference could have influenced the ratings of that dimension. Manipulating the order of presentation of the visual and verbal information in experiments 4 and 5 enabled a test of this possibility.

Table 7-1. Dimensions Used in Experiment 3

<table>
<thead>
<tr>
<th>Product</th>
<th>Focal dimension</th>
<th>Non-focal dimension 1</th>
<th>Non-focal dimension 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>Ease of cleaning</td>
<td>Fit of lids</td>
<td>Heat distribution</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>Power</td>
<td>Sound clarity</td>
<td>Sound distribution</td>
</tr>
<tr>
<td>In-line skates</td>
<td>Speed</td>
<td>Comfort</td>
<td>Ease of use</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>Comfort</td>
<td>Protection</td>
<td>Aerodynamics</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>Power</td>
<td>Loudness</td>
<td>Ease of use</td>
</tr>
</tbody>
</table>

Table 7-2. Results of Experiment 3

<table>
<thead>
<tr>
<th>Product</th>
<th>Condition</th>
<th>Focal Dimension</th>
<th>Dimension 2</th>
<th>Dimension 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>Picture</td>
<td>7.58</td>
<td>3.79</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>8.00</td>
<td>3.92</td>
<td>4.47</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>Picture</td>
<td>7.05</td>
<td>5.16</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>7.53</td>
<td>4.45</td>
<td>4.68</td>
</tr>
<tr>
<td>In-line skates</td>
<td>Picture</td>
<td>6.63</td>
<td>3.19</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>7.55</td>
<td>3.45</td>
<td>5.58</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>Picture</td>
<td>7.02</td>
<td>6.16</td>
<td>6.33</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>6.26</td>
<td>5.74</td>
<td>5.03</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>Picture</td>
<td>6.65</td>
<td>4.47</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>7.71</td>
<td>3.95</td>
<td>5.58</td>
</tr>
<tr>
<td>Average</td>
<td>Picture</td>
<td>6.99</td>
<td>4.55</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>7.41</td>
<td>4.30</td>
<td>5.07</td>
</tr>
</tbody>
</table>
Table 7-3. Manipulation Check for Experiment 3

<table>
<thead>
<tr>
<th>Category</th>
<th>Attractiveness Rating</th>
<th>Performance Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>0.51</td>
<td>-0.73*</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>1.25**</td>
<td>-1.10**</td>
</tr>
<tr>
<td>In-line skates</td>
<td>0.12</td>
<td>-0.68*</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>0.38</td>
<td>-0.42</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>0.31</td>
<td>-2.41**</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .001$

Table 7-4. Results of Experiment 3 with Helmet Replicate Removed

<table>
<thead>
<tr>
<th>Product</th>
<th>Condition</th>
<th>Dimension</th>
<th>Dimension 2</th>
<th>Dimension 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>Picture</td>
<td>6.98</td>
<td>4.15</td>
<td>4.90</td>
</tr>
<tr>
<td></td>
<td>No Picture</td>
<td>7.70</td>
<td>3.94</td>
<td>5.08</td>
</tr>
</tbody>
</table>
CHAPTER 8
EXPERIMENT 4

In previous studies feature evaluation without a picture was compared to feature evaluation with a picture, and a significant effect of the presence of the picture was observed. Certainly this suggests an influence of product design but assuming an averaging model, the use of the picture as additional information does not provide insight into the role the picture played. Thus, in experiment 4, the key comparisons occur between groups that saw both pictures and written information.

Experiment 3 suggested that the pictures had influence because they were perceived to be diagnostic. The picture indicated superiority along the focal dimension, so even though participants were encouraged to focus on the Consumer Reports reviews, the design may have provided what was perceived to be useful information. This suggests a relatively deliberate use of the picture as a source of information. Another possibility is that people had an immediate reaction to the dimensional information in the picture, which distorted evaluation of the corresponding feature information.

To address these possibilities, participants in experiment 4 received both pictures and written information, but the order of presentation of the visual and verbal information was varied. Half the participants saw the pictures first and then the reviews. The other half saw the reviews first, and then the pictures. If the results were due to an initial reaction to the information provided by the picture and an insufficient adjustment upon seeing the verbal information, then the picture should have an effect only when it comes
first. If the results of the previous studies were due to diagnosticity, as suggested by experiment 3, there should be no difference as a function of the order of presentation.

Another issue that is addressed in experiment 4 is one that arose in the pilot study. The pilot study was originally designed to examine attractiveness effects in a product context, but the results mapped more closely onto a performance-based design bias than an attractiveness bias. However, research in facial attractiveness does report a biasing effect of aesthetics. Most consumer research on product design has focused on aesthetics and has demonstrated a variety of effects on liking ratings or product evaluations in the absence of reliable information. In the current series of studies, the key measure has been a quality rating of particular features, and prior aesthetics research in consumer behavior is mute on whether attractiveness might bias such evaluation. Thus, to address this possibility and to confirm the results from the pilot study, an additional factor was included in experiment 4. Half the experimental participants saw the same pictures as those used in experiment 3; that is, pictures in which the products looked equally attractive but one looked superior on the key dimension. For the rest of the experimental participants, pictures were used where the opposite was true; that is, the products looked equal on performance but one was superior on attractiveness. Such a design has inherent difficulties because the pictures are not uniform across conditions. However, the use of five product replicates helped to reduce the magnitude of this design issue.

**Method**

**Stimuli**

The Consumer Reports reviews used in this study were the same as those used in experiment 2. Similarly, the performance pictures were also the same. The only change was the addition of pairs of pictures for each of the five categories in which the pairs had
equal ratings on performance along the key dimension, but significantly different ratings in attractiveness. The pretest results for these additional pictures are presented in table 8-1. For ease of comparison, the table also includes the pretest results from the performance pictures originally presented in experiment 2. The pictures of the new stimuli appear in appendix e.

**Design and Procedure**

The study was a 2 (order of presentation: picture first versus review first) x 2 (type of pictures: performance difference versus aesthetic difference) between subjects design. In addition there was a review-only condition as a control group. The experiment was run on computer, and all instructions were given on screen.

As in experiments 2 and 3 participants were told they would be reading Consumer Reports reviews for pairs of brands in five product categories and that they would be asked to assess the quality of the products based on the reviews. After the introductory screen the first of the five categories was displayed in a randomly determined order. The presentation of each product category proceeded as follows. Participants in the picture-first conditions saw pictures of the two brands. Participants looked at the pictures for 20 seconds before proceeding. The pictures then disappeared from the screen and were replaced by the written reviews. Participants could not proceed further until 40 seconds had passed. Participants in the review-first conditions saw the two reviews first. After 40 seconds participants could proceed from the written reviews to the pictures, which remained on the screen for 20 seconds. Participants in the baseline control condition saw only the reviews. After the presentation of the pictures and descriptions, participants proceeded to a new screen and were asked to indicate the degree to which one or the other brand was superior along the key dimension. This procedure was then repeated for
each of the five product categories. The rest of the procedure, including the filler task, manipulation checks, product knowledge and demographic questions, mirrored experiments 2 and 3. One hundred eighty-seven undergraduate students participated in the experiment for class credit.

Results

A computer error caused 19 participants to be unable to complete the study. An additional 23 participants failed an attention check and were removed, leaving 145 in the analysis. Table 8-2 presents the overall results for the five product replicates.

A 2 (order) x 2 (picture type) x 5 (replicate) mixed ANCOVA using knowledge of the products as covariates was performed on the four experimental conditions (the baseline condition was not included in this analysis). The analysis revealed no three way interaction ($F < 1$), a marginal interaction of picture type and replicate, $F(4, 448) = 2.39$, $p < .06$, no interaction of order and replicate, $F(4, 448) = 1.12$, $p > .34$, no effect of picture type, $F(1, 112) = 1.34$, $p > .25$, and no effect of order, $F < 1$. Importantly, however, there was a significant interaction of order and picture type, $F(1, 112) = 4.47$, $p < .05$. As the table indicates the pictures had little effect when they differed on aesthetics, regardless of whether the picture came before or after the written information. When the pictures differed on performance, they had an effect, but only when they followed the reviews.

Pairwise comparisons on the overall means confirmed that there was a significant difference between the picture-first and review-first conditions for the performance pictures, ($p < .05$). The difference between the aesthetic and performance conditions when the review preceded the pictures was marginally significant, ($p < .07$), as was the
difference between the review-first performance group and the baseline control condition, 
\( p < .08 \). All other comparisons were non-significant.

The manipulation check results appear in table 8-3. For the performance pictures, 
the goal was to have mean-centered ratings that were significantly lower than 0 on 
performance (indicating superiority for the superior picture), and equal to or higher than 0 
on attractiveness. For the aesthetic pictures the goal was the opposite. As indicated in 
table 8-3, all products met this goal for this group of subjects except in-line skates. The 
removal of this replicate did not materially change the results. An ANCOVA on the 
remaining replicates again revealed the critical significant interaction of order and picture 
type, \( F(1, 113) = 6.45, p < .05 \). For completion there was no three way interaction \( F < 1 \), a significant interaction of picture type and replicate, \( F(3, 339) = 3.53, p < .05 \), no 
interaction of order and replicate, \( F(3, 339) = 1.51, p > .20 \), no effect of picture type, \( F(1, 
113) = 1.23, p > .20 \), and no effect of order, \( F < 1 \). The revised overall means with the in-line skate replicate removed were slightly higher than the original means, but maintained 
the same directional patterns. For the picture-first conditions, mean ratings were 7.56 for 
performance pictures and 7.21 for aesthetic pictures. For the review-first conditions, 
mean ratings were 6.77 for performance pictures and 7.53 for aesthetic pictures. The 
review-only control condition had a mean rating of 7.23. Follow-up contrasts again 
revealed a significant difference between the two performance picture conditions \( p < 
.05 \) and a marginal difference between the review-first conditions, \( p < .07 \). The only 
notable change was that the difference between the review-first performance condition 
and the baseline review-only group became non-significant after the removal of the in-line skate replicate.
Discussion

What is immediately notable from experiment 4 is that there was little effect of aesthetics. The pilot study suggested an aesthetic bias for backpacks. Although the running shoes demonstrated a performance-based design bias, it was presumed that both effects could have existed but the aesthetic bias may have been dominated by the performance bias. However, the results of experiment 4 suggest that aesthetics may not influence quality evaluations. Of course, one could argue that this could have been a calibration issue; indeed, the mean ratings for speakers and to a lesser extent helmets, did suggest an attractiveness effect when the pictures preceded the reviews. A comparison of the pretest and manipulation check values for attractiveness versus performance renders this possibility somewhat less likely. As shown in table 8-1, the pretest ratings for the pictures tended to have similar sized differences for the performance and aesthetic pictures. The manipulation check ratings provide a similar picture. As indicated in table 8-3, for this group of participants, the difference ratings for the attractiveness pairs were generally much larger than for the performance pairs. Of course the scales may have been quite different, but it at least suggests that the result was not solely due to calibration.

Prior product design research would suggest that aesthetics might have an influence on information processing, although most such research has focused on aesthetics biasing preference or evaluations in the absence of other information (Page and Herr 2002). The expectation of an aesthetic influence is that people would have an automatic affective reaction that would color further processing. It is possible that at least in some cases aesthetics did have an influence when the picture came first, but overall at least, no strong aesthetics bias emerged overall. The possibility of an aesthetic bias cannot be ruled out here, and further studies would be needed to address this issue. However, based on the
results of the current study, the focus of the dissertation will remain on the performance-based design bias.

An additional surprising result was that for the performance pictures, the effect of design was observed when the picture followed the review but not when it preceded the review. Thus the notion that the influence of design is because people have an immediate reaction which biases subsequent evaluation appears incorrect.

The results for the performance conditions in experiment 4 seem consistent with a recency effect. In the experiment participants knew they would be making some kind of quality rating, but they did not know ahead of time what dimension they would be evaluating. Thus, upon seeing the question, which appeared after the product information had been removed from the screen, they had to rely on their memories to respond. The most recently viewed information was probably the easiest to recall, and so it was used, leading to the divergent results. However, recency cannot entirely account for all experimental conditions. The accuracy of the review-first/aesthetics condition demonstrates that people were able to recall the earlier written information. For these participants, the pictures were the most recently viewed pieces of information and were certainly accessible, but they were not diagnostic, so it was necessary to delve further into memory to retrieve task-relevant information. For participants in the performance/review-first condition, the pictures were also highly accessible in memory. Unlike the aesthetic picture groups, however, for these participants the pictures were diagnostic, so people could use them in their judgment. Thus, rather than being merely a recency effect, it appears the picture results are more consistent with a combination of accessibility and diagnosticity (see Feldman and Lynch 1988). If this is the case, then alerting participants
to the judgment task prior to presenting the product information might alter the influence of the picture and written information. This possibility is explored in experiment 5.

Table 8-1. Pretest Picture Ratings

<table>
<thead>
<tr>
<th>Picture Type</th>
<th>Performance Difference</th>
<th>Aesthetic Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attractiveness Rating</td>
<td>Performance Rating</td>
</tr>
<tr>
<td>Cookware</td>
<td>-0.37</td>
<td>1.08*</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>0.54</td>
<td>1.08**</td>
</tr>
<tr>
<td>In-line skates</td>
<td>-0.33</td>
<td>1.13**</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>0.50</td>
<td>1.08***</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>0.13</td>
<td>2.00***</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001

Table 8-2. Results of Experiment 4

<table>
<thead>
<tr>
<th>Picture Type</th>
<th>Performance Pictures</th>
<th>Aesthetic Pictures</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Picture First</td>
<td>Review First</td>
<td>Picture First</td>
</tr>
<tr>
<td>Cookware</td>
<td>7.43</td>
<td>6.45</td>
<td>7.52</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>7.86</td>
<td>7.61</td>
<td>6.87</td>
</tr>
<tr>
<td>In-line skates</td>
<td>6.09</td>
<td>6.06</td>
<td>6.52</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>6.89</td>
<td>6.26</td>
<td>6.16</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>8.06</td>
<td>6.77</td>
<td>8.29</td>
</tr>
<tr>
<td>Average</td>
<td>7.26</td>
<td>6.63</td>
<td>7.07</td>
</tr>
</tbody>
</table>

Table 8-3. Manipulation Check for Experiment 4

<table>
<thead>
<tr>
<th>Picture Type</th>
<th>Performance Difference</th>
<th>Aesthetic Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attractiveness Rating</td>
<td>Performance Rating</td>
</tr>
<tr>
<td>Cookware</td>
<td>1.01***</td>
<td>-1.19***</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>1.68***</td>
<td>-0.56**</td>
</tr>
<tr>
<td>In-line skates</td>
<td>0.74**</td>
<td>-1.18***</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>-0.19</td>
<td>-0.94***</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>0.68**</td>
<td>-2.12***</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001
Experiment 5 focused solely on the performance-design bias. The experiment attempted to address the order effect that was observed in the performance-picture conditions in experiment 4. In experiment 4 participants in the picture-first condition did not show a design bias, but those in the review-first condition did. However, participants in neither group knew ahead of time what dimension they would be evaluating. All they knew was that they would be evaluating some element of quality. Thus it is plausible that participants in the picture-first group did not focus on the key dimension, but rather on some overall quality judgment. The reviews then provided information on the key dimension, which was subsequently used to make the judgment. For participants in the review-first condition, the review focused participants on the key dimension, which may have then alerted them to the conflicting information provided by the picture. Thus, when it came to making a judgment, they were influenced by the picture.

In experiment 5 participants were told what dimension they would be evaluating prior to viewing any information about the product. The final question still occurred after both pieces of information were removed from the screen, but by providing participants with the dimension ahead of time, it was more likely that both the picture and written information were equally accessible at the time of judgment.
Method

Stimuli

The same set of pictures used in the performance conditions of experiment 4 were used in experiment 5. The same product information was also retained.

Design and Procedure

Experiment 5 consisted of one between-subjects factor of order, with three levels: picture first, review first, and a review-only control group. Again there were five product replicates.

The procedure was similar to that of experiment 4. The only difference was that participants were told what dimension they would be evaluating prior to seeing the pictures or the written reviews of the products. Ninety-one undergraduate students participated in the experiment for partial class credit.

Results

Eleven participants were removed for failing an attention check, leaving 80 in the analysis. The mean ratings appear in table 9-1. There was a significant effect of replicate, $F(4, 308) = 12.50, p < .01$ and no interaction of order and replicate, $F(8, 308) = 1.45, p > .15$. There was a significant effect of order, $F(2, 77) = 4.83, p < .05$, but the direction of the order effect was reversed from that observed in experiment 4. In the current study, when people were instructed ahead of time which dimension to focus on, the picture had a stronger effect when it preceded the review. Follow up comparisons revealed a significant difference between the picture-first and review-first conditions, $t(77) = 2.55, p < .05$ and between the picture-first and review-only conditions, $t(77) = 2.78, p < .05$, but no difference between the review-first and review-only conditions, $t < 1$. 

The manipulation checks revealed that for this group of participants, both the stereo speakers and the bicycle helmets were unsuccessful manipulations of performance design. Removing these two replicates from the analysis resulted in revised overall means of 7.07, 7.55, and 7.98 for the picture-first, review-first, and review-only conditions respectively. The effect of order was still significant, $F(2, 77) = 3.48, p < .05$, but the pairwise comparisons resulted in only a significant difference between the picture-first and review-only conditions, $p < .05$.

**Discussion**

By providing participants with the focal dimension before the information, experiment 5 enabled a test of the effect of order that avoided the problem of recency. In this study the appearance of the question after the presentation of the picture was no surprise, so participants did not have to search their memories for relevant judgment information.

The key result from experiment 5 is that the picture had the greatest influence when it preceded the written review, a reversal of the order effect observed in experiment 4. In this study, participants knew what feature they would be evaluating. When the picture appeared, it seemed diagnostic of quality and therefore was used in arriving at a judgment of the feature. The subsequent written review was of course also diagnostic. However, when it came time to supply the evaluation, the judgment made based on the picture was probably still highly accessible, and since it was also perceived to be diagnostic it influenced the rating.

When the review preceded the picture, participants were easily able to assess the quality of the two products along the dimension of interest before seeing the picture. Thus, even though the picture provided conflicting information, its influence was weaker,
perhaps because the judgment made based on the written reviews was very clear. The picture did still appear to have some influence, however. An examination of the overall means after the removal of the stereo speakers and bicycle helmets does suggest an influence of the picture when it followed the review, albeit a weaker one than was observed in the previous study.

Experiments 3 and 4 suggested that the picture effect was due to diagnosticity. However, in the current study, a diagnosticity result should have led to equal ratings for the picture-first and review-first conditions. The review-first condition showed some effect of the pictures, particularly for the in-line skates and electric mixers, but overall the results were not supportive. If the current results are reliable, the design effect may not be solely due to diagnosticity, but rather may be the result of an immediate reaction to the picture biasing subsequent evaluation, but only when attention is drawn to dimensions where the picture dominates. Additional data are needed to confirm the results of the review-first condition.

Table 9-1. Results of Experiment 5

<table>
<thead>
<tr>
<th></th>
<th>Picture First</th>
<th>Review First</th>
<th>Review Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookware</td>
<td>6.39</td>
<td>7.65</td>
<td>7.50</td>
</tr>
<tr>
<td>Stereo speakers</td>
<td>6.54</td>
<td>7.96</td>
<td>7.19</td>
</tr>
<tr>
<td>In-line skates</td>
<td>7.14</td>
<td>7.19</td>
<td>7.77</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>7.04</td>
<td>7.65</td>
<td>7.42</td>
</tr>
<tr>
<td>Electric mixers</td>
<td>7.68</td>
<td>7.81</td>
<td>8.69</td>
</tr>
<tr>
<td>Average</td>
<td>6.95</td>
<td>7.65</td>
<td>7.71</td>
</tr>
</tbody>
</table>
CHAPTER 10
GENERAL DISCUSSION

The goal of this paper was to examine some of the ways product design might influence judgment. A pilot study and five experiments pitted product appearance against conflicting written information to address whether design could influence a person’s tendency or ability to objectively evaluate objective product information.

Consistent with prior research in psychology as well as consumer research on product design, the initial pilot study focused on attractiveness as a source of bias. An attractiveness effect obtained for one of the two product replicates, but attractiveness was confounded with performance for that stimulus. For the other stimulus, no attractiveness effect was observed. Rather the results suggested a bias in the direction of the unattractive brand, which subsequent analysis revealed looked like a superior performer. This discrepancy was resolved in experiments 1 and 2, which confirmed that the appearance of performance biased the processing of other information.

Experiments 3 through 5 attempted to assess the source of the performance-design bias. Experiment 3 investigated whether design operated as a type of halo that influenced ratings of other attributes. Results indicated that design only influenced ratings for features on which it appeared to dominate, suggesting the effect was not due to a halo effect. Experiments 4 and 5 examined whether the bias was a function of an initial reaction to the visual information coloring subsequent processing of written information. Experiment 4 did not support such an explanation, and instead appeared more consistent with a accessibility-diagnosticity explanation. In that experiment, participants were not
focused on the key dimension prior to viewing the information. When asked the
dependent measure, participants used the most accessible information that was diagnostic
for the task. This tended to be the most recently viewed information. Experiment 5 was
less conclusive, but still hinted at a diagnosticity result. When participants were focused
on the key dimension ahead of time, the picture had an influence when it appeared prior
to the written information because the task was known at the time and the pictures were
diagnostic. When the picture followed the review, however, less influence of the picture
was observed. This result is inconsistent with a diagnosticity explanation and more in line
with an early reaction distorting subsequent evaluation. However, stimuli calibration
concerns may have been partially responsible.

Overall the results of the experiments suggest that for products, when it comes to
being evaluated for quality, seeing is believing. When a product appears to offer superior
performance along a given dimension, people believe that it does, and reading
information that suggests otherwise may be insufficient to counteract that belief.

Whether there is an aesthetics bias is unclear. Experiment 4 addressed the issue of
attractiveness as a source of bias initially investigated in the pilot study, and found no
effect of aesthetics. However, the study did not ask for choice or preference judgments.
Such judgments, which are inherently more personal, may be more susceptible to
aesthetic effects. Certainly if one considers the research in facial attractiveness and the
dramatic effects that have been reported (e.g., Landy and Sigall 1974), it is likely that an
aesthetic bias in products would occur under some conditions. The results of some of the
replicates in experiment 4 hint that there may be aesthetics bias; further research is
needed to confirm this and to understand the conditions under which an aesthetic design bias might obtain.

**Implications**

Product design and styling are often among the most important determinants of new product success (Cooper and Kleinschmidt 1987; Vriens et al. 1998). A good design can attract attention, communicate information about a product, and influence preference. Thus, understanding how consumers process and react to product design, particularly with regard to new product designs used in market testing, is of great importance. For managers, the key takeaway from this research may be that consumers believe what they see. If the goal of a new product is to communicate some level of performance, it is critical that this is communicated visually. Otherwise marketing efforts aimed at convincing consumers that the product excels on a given dimension may fall flat. In terms of design team goals, the research suggests that a company may be better off expending efforts on making a product look like it works well rather than simply making it look good.

From a consumer perspective, of particular importance is a better understanding of the role that product appearance plays in decision making. The experiments demonstrate that consumers are able to discount attractiveness in the evaluation of product quality, and yet seem unable to ignore appearance when design suggests performance. It may be that this is a deliberate choice and that consumers consciously discount the written information when their eyes tell them otherwise. However, if consumers are convinced by what they see rather than other information they encounter and are unable to ignore visual information when it should be ignored, there is potential for non-normative decision making. As an extreme example, products that appear to dominate on safety,
such as life jackets, flashlights, or cars, but actually do not provide the level of performance that their appearance suggests, could result in seriously detrimental consequences for the consumer.

**Limitations**

The research reported here focused on a fairly small number of products that were chosen because they were products where design could be important but was not critical to the function of the product. Because of these narrow selection criteria, it is unclear whether the results observed in the dissertation would extend to a wider array of product categories. Moreover, in several experiments not all the picture manipulations were successful. It will be important to improve the stimuli calibration and replicate the studies to confirm the validity of the results.

The research was conducted using an undergraduate student subject pool. This is a relatively homogeneous group and not representative of the typical American consumer. Although this is a concern common to nearly all consumer behavior research, it makes it difficult to assess the generalizability of the results to the population.

The studies presented in this dissertation suggested that there is little influence of attractiveness in the processing of product information. It is certainly possible, however, that under some conditions, the aesthetics of a product may indeed interfere with a consumer’s ability or inclination to evaluate product quality. It is possible that the products used in the current study were not sufficiently attractive or interesting to the student participants, and had more attractive or student-relevant products been used, an attractiveness effect might have emerged.
Future Research

Experiments 3 and 4 suggested that the effect of product design observed in this research were due to a diagnosticity effect. The results of experiment 5, however, left the issue somewhat unresolved. Experiment 5 had some evidence consistent with a diagnosticity effect, but also suggested that the design effect could have been the result of an initial reaction coloring subsequent evaluation. The study had two failed stimuli manipulations, which would need to be corrected in a future experiment. In addition, a study which manipulates order of information presentation, as in experiment 5, but also has participants rate features on which the pictures are not diagnostic, as in experiment 3, could help to resolve the issue.

The research reported here demonstrated that consumers’ ability to process written information is influenced by conflicting visual information. The evidence showed that when it comes to quality judgments, consumers are not biased by attractiveness, but are biased by the appearance of performance. This suggests that the effect may be relatively deliberative and conscious, but it may not be. Additional research is necessary to determine the degree to which people are aware of the effect of design and the extent to which people can correct for it. Consistent with research in mental contamination (Wilson and Brekke 1994) it could be that people are aware of the biasing effect of design and yet are still unable to counteract it.

As stated above, the dissertation found little aesthetic influence. Prior research has found that attractiveness does play a role in liking evaluations and may influence quality evaluations in the absence of other concrete functional information (Page and Herr 2002). The current research is consistent with these findings, but does not conclude that aesthetics cannot interfere with information processing. Research in face processing has
demonstrated that under certain conditions, an attractive face can alter the manner in which functional information is assessed (Landy and Sigall 1974). There may be some cases in which a sufficiently attractive product can have the same effect. For example, Norman (2004) suggests that product design can act as a mood manipulation, with good design elevating mood and making people less likely to notice small functional problems. If a product is attractive enough to cause an elevation in mood, then it may lead to motivated reasoning with regard to functional attribute information. Future research is needed to investigate the conditions under which product aesthetics will and will not interfere with processing of product information.
APPENDIX A
PRODUCT PICTURES USED IN PILOT STUDY

Unattractive Shoe    Attractive Shoe

Unattractive Backpack    Attractive Backpack
APPENDIX B
PRODUCT PICTURES USED IN EXPERIMENT 1

High performance shoes

Low performance shoes
APPENDIX C
PRODUCT PICTURES USED IN EXPERIMENTS 2 THROUGH 5

High Performance Brand

Low Performance Brand
# APPENDIX D

## PRODUCT DESCRIPTIONS USED IN EXPERIMENT 2, 4 AND 5

<table>
<thead>
<tr>
<th>Product Category</th>
<th>Low Performance Brand</th>
<th>High Performance Brand</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cookware</strong></td>
<td>This cookware set includes a large stock pot, three pots, and two frying pans. The pots have well fitting lids that keep heat in nicely and sturdy handles that stay cool to the touch. The set is durable and can be used at the highest heat, even when cooking with a gas stove. Food cooks very evenly because of the flat bottoms that sit squarely on the elements, maximizing heat distribution and minimizing burning. If pans sit out for an extended period, remaining food can be difficult to remove. The lids for the stock pot and medium pot also fit the two frying pans.</td>
<td>This cookware has good heat distribution; food generally cooks evenly. The pots and pans have a solid base, durable construction and lids that seal in heat very well. The stay-cool handles are ergonomically designed and comfortable to hold. Very little will stick to these pans, regardless of cooking method, so they can be washed out with dish soap and a cloth; it's rarely necessary to use a scrub brush. The set comes includes three handled pots of varying sizes with lids, a large stock pot with lid, and two frying pans that fit the lids of the stock pot and mid-sized pot.</td>
</tr>
<tr>
<td><strong>Stereo speakers</strong></td>
<td>These speakers pack several patented technologies into their slender cabinets to ensure strong results. They have a straight-line signal path crossover network which distributes the signal with minimal processing. At relatively low volumes the speakers provide clear and accurate sound, but slight distortion may occur at higher volumes. They can be used with a computer, a stereo or as part of a home theatre system.</td>
<td>These magnetically shielded speakers are appropriate for an apartment or dorm, but would be inadequate for a larger home. The speakers house two angled, full-range drivers engineered to generate more air distribution from a small enclosure than conventional drivers. They provide consistent sound at both low and high volumes. You can almost tell the relative position of the different instruments in the recording studio.</td>
</tr>
<tr>
<td><strong>In-line skates</strong></td>
<td>This in-line skate is solid and durable. It offers an easy-entry system making for very quick changes. It uses a TrueFit interior, wrapping a nylon and plastic film around the foot for extra stability and control. The fit is snug and well cushioned. There are tiny slits in the front which interfere slightly with aerodynamics, but which help cool the feet. The wheels are slightly softer than on other models, making for a more comfortable ride by increasing shock absorption. A rigid, well-aligned frame improves maneuverability.</td>
<td>This in-line skate has a closer fit in the ankle collar than older models, giving stability while still allowing the skater to flex forward. Lighter materials throughout give the skate an impressive weight of only three pounds, which increases ease of movement. The wheels are hard, to last longer and improve speed, but as a result, do not offer that much shock absorption. This is compensated for somewhat by the interior comfort. The skates have a thick liner and a soft foam interior to increase comfort and keep feet cool.</td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>This helmet has a roll-cage interior bracing somewhat reminiscent of roll bars in jeeps, made to protect the rider on impact. Despite this additional bracing, the helmet is reasonably light. The ventilation is adequate. Although it only comes in one size, the helmet is suitable for most head sizes. The foam on the interior lining is a little stiff, but the pads can be adjusted to suit your head size. The buckle is specially designed to stay secure but allow for a quick release in case of emergency.</td>
<td></td>
</tr>
<tr>
<td>Electric mixers</td>
<td>This helmet is durable with a reinforced rib cage frame. Of course, the reinforced structure adds some weight to the helmet, making it a bit heavier than most. But the lining is soft, thick and well padded with foam, so the extra weight is not much of a burden. Superior fit is provided by having three sizes. Also, it has a sleek wind funneling technology to pull cool air in and blow hot air out, cooling the head during rides. Straps are sturdy and will stay in place on impact.</td>
<td></td>
</tr>
<tr>
<td>Bicycle helmets</td>
<td>This mixer is relatively compact, fitting easily on most counters, but the bowl is deep enough to handle large sized jobs. It has a gradual starter, minimizing splatter. It changes speeds very smoothly, and its plastic gears make the machine surprisingly quiet. It has an auto shut-off motor protection feature if the motor gets overloaded, which does happen occasionally. Beaters snap in and out very easily and are dishwasher safe.</td>
<td></td>
</tr>
<tr>
<td>Electric mixers</td>
<td>This is an efficient mixer with a wide-set bowl that can be snapped in or lifted out of the cradle quickly and easily. The machine has all steel gears and a twin-motor design that rotates the bowl while mixing. It can handle everything from small, high speed tasks to large heavy doughs. Occasionally the batter will cling to the beaters, which requires stopping the machine to scrape the contents back into the bowl to complete the mixing.</td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX E
PRODUCT DESCRIPTIONS USED IN EXPERIMENT 3

<table>
<thead>
<tr>
<th>Product Category</th>
<th>Low Performance Brand</th>
<th>High Performance Brand</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cookware</strong></td>
<td>This cookware set includes a large stock pot, three pots, and two frying pans. The pots have well fitting lids that completely seal in heat and sturdy handles that stay cool to the touch. The set is durable and can be used at the highest heat, even when cooking with a gas stove. Food cooks evenly because of the flat bottoms that sit squarely on the elements, maximizing heat distribution and minimizing burning. If pans sit out for an extended period, remaining food can be difficult to remove. The lids for the stock pot and medium pot also fit the two frying pans.</td>
<td>This cookware has good heat distribution; food cooks very evenly. The pots and pans have a solid base, durable construction and lids that keep heat in nicely. The stay-cool handles are ergonomically designed and comfortable to hold. Very little will stick to these pans, regardless of cooking method, so they can be washed out with dish soap and a cloth; it’s rarely necessary to use a scrub brush. The set comes includes three handled pots of varying sizes with lids, a large stock pot with lid, and two frying pans that fit the lids of the stock pot and mid-sized pot.</td>
</tr>
<tr>
<td><strong>Stereo speakers</strong></td>
<td>These speakers have a straight-line signal path crossover network which divides the signal with minimal processing. Bass response is adequate but not strong, making these speakers perfect for a smaller space like a dorm. The sound is clear and accurate, very easy to listen to. You can almost tell the relative position of the different instruments in the recording studio. The balancing technology blends direct and indirect sound almost anywhere in the room, giving the feeling of surround sound with only two speakers.</td>
<td>These magnetically shielded speakers house two angled, full-range drivers engineered to generate more air movement from a small enclosure than conventional drivers. They deliver clean, uncolored, realistic sound, which is maintained at higher volumes, provided of course that the accompanying stereo system is adequate. They would be suitable for a home theater system because the direct/reflect system avoids the common problem of a speaker sweet spot and ensures the same quality of sound throughout the room.</td>
</tr>
<tr>
<td><strong>In-line skates</strong></td>
<td>This in-line skate is solid and durable. It offers a one strap easy-entry system making for very quick changes. It uses a TrueFit interior, wrapping a nylon and plastic film around the foot for extra stability and control. The fit is snug and very well cushioned. There are tiny slits in the front which interfere slightly with aerodynamics, but which cool the feet. The wheels are slightly softer than on other models, making for a more comfortable ride by increasing shock absorption.</td>
<td>This in-line skate has a close fit in the ankle collar, giving stability while still allowing the skater to flex forward. Lighter materials throughout give the skate a weight of only three pounds, increasing ease of movement. The wheels are hard, to last longer and improve speed, but as a result, do not offer that much shock absorption. This is compensated for very well by a thick liner and a soft foam interior. No laces and easy-tug entry grips make the skate very simple to put on and take off.</td>
</tr>
</tbody>
</table>
**Bicycle helmets**

This helmet has a roll-cage interior bracing made to protect the rider on impact. Despite this additional bracing, the helmet is not too heavy. A streamlined design and well positioned vents optimize speed. Although it only comes in one size, the helmet is suitable for most head sizes. The foam on the interior lining is a little stiff, but the pads can be adjusted to suit your head size. The buckle is specially designed to stay secure but allow for a quick release if necessary.

This helmet is strengthened with a reinforced rib cage frame for safety. Of course, the reinforced structure adds a bit of weight to the helmet. But the lining is soft, thick and well padded with foam, so the extra weight is not much of a burden. Superior fit is provided by having three sizes. Strategically positioned air vents and advanced wind funneling technology help to cool the head without sacrificing speed. Straps are sturdy and will stay in place on impact.

**Electric mixers**

This mixer is relatively compact, fitting easily on most counters, but the bowl is deep enough to handle large sized jobs. It has a gradual starter, minimizing splatter. It changes speeds very smoothly, and its plastic gears make the machine surprisingly quiet. It has an auto shut-off motor protection feature if the motor gets overloaded, which does happen occasionally. Beaters snap in and out very easily and are dishwasher safe.

This is an efficient mixer with a wide-set bowl and beaters that can be snapped in or lifted out of the cradle quickly and easily with one hand. The machine has all steel gears and a twin-motor design that rotates the bowl while mixing. The gears and motors are insulated to minimize noise. It can handle everything from small, high speed tasks to large heavy doughs.
APPENDIX F
PRODUCT PICTURES USED IN EXPERIMENT 4

Attractive Brand

Unattractive Brand
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

JoAndrea Hoegg earned bachelor’s degrees in English and education from the University of British Columbia in Vancouver, Canada. In 1998 she moved to San Diego, where she completed a Master of Business Administration at San Diego State University. In 2001 she entered the Ph.D. program in marketing at the University of Florida. She joined the University of British Columbia as an assistant professor of marketing in July 2006.