The good, the bad, and the ugly: Influence of aesthetics on product feature judgments

JoAndrea Hoegg a,⁎, Joseph W. Alba b, Darren W. Dahl a

a Sauder School of Business, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z2
b Warrington College of Business Administration, University of Florida, P.O. Box 117155, Gainesville, FL 32611, USA

Available online 4 August 2010

Abstract

As goods commoditize more quickly in terms of functionality, design is increasingly becoming a critical point of differentiation. The present research examines the interaction of aesthetic design and product evaluation, testing the conventional wisdom among practitioners that “what is beautiful is good.” Three studies examine how design influences feature processing when aesthetics and feature performance conflict. Study 1 reveals a bias in the direction of the unattractive product—a negative aesthetic effect—and provides initial evidence that this bias stems from thoughtful reconciliation of incongruous information and a consequent elaboration of the conflicting dimension. Studies 2 and 3 examine boundary conditions. © 2010 Society for Consumer Psychology. Published by Elsevier Inc. All rights reserved.

Keywords: Aesthetics; Product design; Visual and verbal information

It is amazing how complete is the delusion that beauty is goodness. ~Leo Tolstoy

Introduction

Product design has become an important tool for marketers (Kalins, 2003). Indeed, as product categories rapidly commoditize, design can act as a point of differentiation and serve as a source of competitive advantage (Page & Herr, 2002). Although design is itself multifaceted (Bloch, 1995; Hoegg & Alba, 2008), much of the current discussion with respect to design centers on the importance of visual aesthetics and external appeal. One avenue of research has focused on stimulus factors that drive aesthetic impressions, such as physical size (Silvera, Josephs, & Giesler, 2002), prototypicality and unity (Kumar & Garg, 2010; Veryzer & Hutchinson, 1998), and the interaction of design complexity and exposure frequency (Cox & Cox, 2002). Other research has examined individual level differences in response to visual design (Bloch, Brunel, & Arnold, 2003; Yang, Zhang, & Peracchio, 2010). At a more global level, research has investigated ways in which aesthetics contribute to overall product evaluation. Consistent with intuition, for example, design can direct choice at the margin when performance information is absent or ambiguous (Yamamoto & Lambert, 1994). Less intuitive is the finding that aesthetics can also exert a non-normative influence on overall judgment, altering product liking even in situations where design should be irrelevant (e.g., Raghubir & Greenleaf, 2006; Madzharov & Block, 2010; Townsend & Shu, 2010).

The present research is concerned neither with the manner in which design elements influence aesthetic response nor with the influence that design exerts as an independent product attribute. Instead, we examine the less frequently posed questions of whether and how aesthetics alter processing and evaluation of objective product information. Although prior research has given extensive consideration to the processing of visual and verbal information (e.g., Bagozzi, 2008; Childers & Jiang, 2008; Wyer, Hung, & Jiang, 2008), little empirical effort has been toward understanding how conflicting pieces of visual and verbal information are reconciled. We investigate how consumers make functionality judgments when visual information, in the form of product design, conflicts with verbally presented feature performance information. In so doing, we uncover an aesthetic phenomenon that challenges the implicit and commonsensical assumption among design practitioners and scholars that being attractive is preferable to being unattractive.

In a general sense it is difficult to argue against the virtues of attractiveness. We are drawn to people and objects that are aesthetically pleasing, and attractiveness can compensate for—or
bias the interpretation of—other decision-related information (e.g., Landy & Sigall, 1974; Raghurib & Greenleaf, 2006; Townsend & Shu, this issue). In the commercial world, aesthetics may enjoy an additional advantage. That is, although there is no structural reason to expect a correlation between attractiveness and functionality (Tractinsky, 1997), the salience of very attractive designs in existing high-performance products, (e.g., Apple iPod, Mini Cooper, Method cleaning products) may create an expectation among consumers that more attractive products are also functionally superior (Creusen & Schoormans, 2005). This may occur even if the design itself does not communicate any specific information about functionality.

Despite consumers’ likely expectations regarding a correspondence between aesthetics and performance, the true ecological correlation is surely less than unity (Hassenzahl, 2004). Dyson vacuum cleaners are thought to be unattractive but superior in performance, and unattractive website designs may provide functional advantage (Scobleizer, 2008). As noted, our interest resides in this situation in which functional and aesthetic cues conflict and, therefore, consumer expectations are violated. Operationally, we examine how consumers judge relative feature performance of two alternatives when both options are characterized by an aesthetic-functionality conflict (i.e., one is superior to the other on a functional attribute but inferior in visual attractiveness). With regard to consumer response, at least three outcomes seem plausible.

First, given that the task is evaluation of relative performance of a functional feature that cannot be accurately assessed through visual inspection of the designs, the normative outcome would be a null effect of aesthetics. Individuals should be able to compare the objective product information and accurately judge functionality regardless of whether the respective designs are attractive or unattractive. A second possibility is that judgments will be biased in the direction of the more attractive, but inferior product. If, as we suspect, consumers expect a correlation between aesthetics and performance and are unmotivated to process carefully, they may simply assume that the more attractive design is functionally superior (Chaiken & Maheswaran, 1994). In a related vein, an initial affective reaction to the product designs also may create a positive halo that alters how other attribute information is processed (Nisbett & Wilson, 1977). If people develop an initial preference for the more attractive design, judgments of objective feature information may shift in the direction of the more attractive product.

While not denying the situational validity of these outcomes, consumer psychology should not be blind to a third possibility. Specifically, when aesthetics and objective feature information are in conflict, feature judgments may be biased in the direction of the less attractive product. Research on inconsistency reconciliation suggests that consumers are often motivated to form an integrated evaluation and hence consider conflicting pieces of information in relation to one another (Maheswaran & Chaiken, 1991; Sengupta & Johar, 2002). In the present context, if consumers naturally assume that a more attractive design will offer superior functional performance, incongruity is created when the expected consistency between aesthetics and performance is violated by the conflicting feature information (i.e., the unattractive product having better feature functionality). Assuming that consumers possess the cognitive resources to recognize the inconsistency, they should be motivated to resolve it and, in so doing, elaborate on the conflicting, verbally described feature (Maheswaran & Chaiken, 1991; Sengupta & Johar, 2002). Such elaboration is effortful and systematic and may render consumer beliefs about the feature more evaluatively consistent and result in a polarized assessment (Tesser & Leone, 1977). The final outcome in such a stimulus and processing environment is that the presence of aesthetic information will bias relative performance judgments of the verbally presented functional feature in the direction of the less attractive design. We refer to this as the negative aesthetic effect.

It is important to note that polarization is most likely to be observed when consumers elaborate on a single element or several highly correlated elements (Millar & Tesser, 1986). If our hypothesis is correct and elaboration on the conflicting functional feature leads to the negative aesthetic effect, polarization of other features that are uncorrelated with the conflicting feature should not be observed. Thus, if consumers are presented with written information about several product features but only one of these features is in conflict with aesthetic information, the conflict should result in polarization of that feature only.

Cognitive resources

Product performance judgments have been argued to be slower, more thoughtful, and more deliberate than liking or other affective judgments because they require integration of several pieces of information (Page & Herr, 2002). In the present paradigm, in which systematic and effortful processing is required to reconcile the conflicting cues, the negative aesthetic effect should arise only when consumers have the resources to recognize and attempt to resolve the inconsistency. Thus, constraining such resources should mitigate the effect. Moreover, inasmuch as cognitive constraints have been shown to lead to the dominance of automatic, affective processes over thoughtful, cognitive ones (Shiv & Fedorikhin, 1999), imposing such constraints may lead to an advantage for the more attractive but less functional option.

Overview of studies

Across three experiments we pit product aesthetics against verbal product information to examine how conflicting visual and verbal information are reconciled to form feature evaluations. In all studies, participants were provided with reviews about two brands and asked to make relative performance judgments of a particular feature based on the reviews. The reviews made it clear that one brand was substantially better than the other brand on that feature. Accompanying the reviews were pictures of the designs of the two brands. One design was more attractive than the other. The pictures and reviews were mismatched on valence such that the less attractive picture was paired with the superior target feature review and the more attractive picture was paired with the inferior target.
feature review. Study 1 tested for the presence of a negative aesthetic effect and examined whether the effect extends to other features. In addition, study 1 investigated the role of cognitive resources. In study 2 we examined a boundary condition for the negative aesthetic effect: the order of presentation of the conflicting visual and verbal information. Study 3 considered the role of brand name and examined the different roles that aesthetics and brand name can play in the formation of product judgments.

Across all studies, we deliberately employed the context of product comparison to investigate the negative aesthetic effect. It is in this context that the salience of relative attractiveness should be highest and the advantage for the unattractive product greatest.

Study 1

Participants were presented with fictitious Consumer Reports reviews of two brands, which were designed to be equivalent on all features except the critical one. On the functionality of this feature, hereafter referred to as the target feature, one brand was superior to the other. We then paired these reviews with photographs of the two product designs. One visual design was more aesthetically pleasing than the other. Superiority of the target feature could not be determined by a visual inspection of the product designs and no brand labels were visible.

The key manipulation involved the pairing of the visual and verbal information. Participants were presented with one of three stimuli configurations: (1) mismatched visual and verbal information (i.e., the more attractive visual design was paired with the inferior target feature review, and the less attractive design was paired with the superior target feature review), (2) matched visual and verbal information (i.e., the more attractive visual design was paired with the superior target feature review, and the less attractive design was paired with the inferior target feature review), or (3) no visual information (i.e., a verbal-information-only control condition). Configuration 1 is the condition in which we expect the negative aesthetic effect to obtain. If, as we predict, aesthetics have a reverse effect when aesthetics and functionality conflict, judgments should move in the direction of the objectively inferior target feature (i.e., the more attractive design) when accompanied by conflicting visual information.

For the matched configuration, strong ratings for the target feature should also be observed. If consumers believe that attractive products perform better, the sufficiency principle should be applied when a product’s design and functionality are of the same valence (Maheswaran & Chaiken, 1991). That is, a quick confirmation that aesthetics and functionality are conforming to expectations allows for more efficient processing of attributes and the easy judgment that the more attractive brand is superior on the target feature.

The second manipulation involved cognitive resources. According to our argument, the negative aesthetic effect requires thoughtful elaboration of the reviews, and therefore should be dependent on processing resources. Applying cognitive constraints to participants viewing configuration 1 should mitigate the advantage for the unattractive product. In contrast, because the matched configuration is consistent with expectations and hence can be easily and efficiently assessed, cognitive constraints should not affect the judgments in this condition. Similarly, for the verbal-information-only condition, no effect of processing resources was expected. The product descriptions were pretested so that the difference on the target feature could be easily assessed. Even when less attentive, participants should be able to determine that one brand is superior to the other on the target feature.

An additional goal of study 1 was to determine whether the negative aesthetic effect extends to other product features. The premise behind the negative aesthetic influence is that violation of expectations regarding a positive relationship between aesthetics and functional performance makes the relative superiority of the verbally described functional feature more salient—which, in turn, should result in elaboration and subsequent polarization of relative performance judgments of the target feature. If the focus is on that single feature, polarization should not occur for features that are not in conflict with the visual information. We test this proposition by examining two other features that were pretested to be equivalent in the written reviews.

When the visual and verbal information are mismatched, we expect participants’ feature ratings to move in the direction of the less attractive product when cognitive load is low and in the direction of the more attractive product when cognitive load is high. However, when the visual and verbal information are matched, we expect that the consistency of the information and its conformity with expectations will allow people to make their judgments with less effort, resulting in support for the superior target feature, regardless of cognitive load. We compare these conditions to the no-visual-information control conditions.

We have argued that the negative aesthetic effect should only obtain when consumers attempt to reconcile the inconsistency between the valence of the designs and the valence of the reviews. If consumers have the cognitive capacity to consider the two pieces of information, they should take a longer amount of time to make judgments when evaluating mismatched visual and verbal information than when evaluating matched visual and verbal information. To test this possibility we also examined the time participants took to provide their relative feature ratings.

Method

Study 1 conformed to a 3 (visual—verbal information: matched information, mismatched information, no visual information) × 2 (cognitive load: low vs. high) × 2 (product replicate: cookware vs. mixers) × 3 (feature: target, non-target 1, non-target 2) mixed design with visual—verbal information and cognitive load manipulated between subjects and product replicate and feature within subject.

Stimuli

Two product categories, cookware and electric mixers, were selected as stimuli. We chose product categories for which aesthetics could be a consideration but would not be paramount. Designs were selected that ensured the brands would appear equivalent or simply uninformative on the features discussed in the reviews (i.e., functional superiority of the target feature could not be determined by a visual inspection of the designs) yet would be significantly different in attractiveness. Twenty-four participants rated pairs of pictures of product designs on relative
measures were seven-point scales from −3 = “Brand A is far superior” to +3 = “Brand B is far superior” (A and B brand labels counterbalanced across participants). A rating of zero would indicate the participant believed the brands were equivalent. The ratings of the selected pictures on relative performance of the target feature were 0.21 for cookware and 0.38 for electric mixers, neither significantly different from 0, \(t(23) = -0.75\), and \(t(23) = 1.06\), both \(p > 0.30\). Thus, for both products, participants could not discern a difference in performance of the target feature based solely on the two designs. Attractiveness ratings were also assessed on seven-point scales. Relative attractiveness was 1.79 for cookware and 1.04 for electric mixers, both of which were significantly different from 0 (\(t(23) = 5.26, p < 0.001\) and \(t(23) = 2.51, p = 0.01\) for the cookware and mixers respectively).

Paragraph descriptions of the products were in the format of Consumer Reports reviews, each describing the functional performance of several product features. The verbal information was pretested so that one review was significantly better than the other on a single feature (referred to as the target feature) and equivalent on two other non-target features. Thirty-two participants evaluated the relative performance of the products on the target feature on nine-point scales where −4 = “Brand A is far superior” and +4 = “Brand B is far superior” (A and B brand labels counterbalanced). Ratings were −3.00 and −3.41 for the cookware and electric mixers respectively (\(t(31) = -5.53, p < 0.001\) and \(t(31) = -21.15, p < 0.001\)). Another 36 participants rated the relative performance of the non-target features. For the cookware, ratings were −0.06 and −0.06 for each non-target feature, both \(t < 1\). For electric mixers, ratings were 0.20 for non-target 1 (\(t < 1\)) and 0.31 for non-target 2, \((t(34) = 1.25, p > 0.20)\).

\textbf{Appendix A} provides the specific product features.

\textbf{Procedure}

Participants were told they would be reading Consumer Reports reviews for pairs of brands in two product categories and that they would be asked to assess the functional performance of the products based on the reviews. After this introduction, participants were told that the study was also investigating consumers’ ability to attend to multiple pieces of information and therefore they would be asked to remember a number while completing the study. Depending on cognitive load condition, participants were asked to memorize a 2 or 11-digit number, and were given 20 s to do so. After the cognitive load manipulation, the first product category was displayed in a randomly determined order. Participants in the mismatched visual–verbal information conditions saw the two Consumer Reports reviews side-by-side on the screen, position counterbalanced. Above each review was the conflicting photograph of the product. That is, above the review indicating a superior target feature was an unattractive picture and above the review indicating an inferior target feature was an attractive picture (see Appendix B for an example). Participants in the matched visual–verbal information conditions also saw the two Consumer Reports reviews side-by-side with pictures of the designs above. However, for these individuals, the pictures were reversed so that the more attractive picture was above the review describing the superior feature and the less attractive picture was above the review describing the inferior feature. Participants in the no visual-information control conditions saw only the two written reviews side-by-side.

Following presentation of the options, participants proceeded to a new screen and were asked to indicate the degree to which one or the other brand was superior on each of the three features (i.e., the target and the two non-targets), order randomly determined. They provided their responses using a 9-point scale from −4 (Brand A is far superior on this feature) to +4 (Brand B is far superior on this feature), position counterbalanced. A positive rating would mean that the brand with the objectively superior target feature was superior. A rating of 0 would indicate perceived equivalence between the two brands on the performance of the feature. A negative rating would mean that the participant perceived the brand with the objectively inferior target feature as better on the respective feature. To avoid evaluative inferences related to the scale points, the numbers on the scale were redacted so that participants only saw nine radio buttons ranging from “Brand A is far superior” to “Brand B is far superior.” The second product category then appeared and the process was repeated. The amount of time participants took to provide their judgment of each feature was also recorded. After providing ratings for both categories, participants were asked questions regarding their knowledge of the products, how seriously they took the task, and their demographic characteristics. A total of 214 undergraduate students from a west coast university participated in the study for partial course credit.

\textbf{Results}

All studies reported in this research involved reading detailed paragraph-format product reviews and making judgments of relative feature functionality. Due to the visual nature of the stimuli and the attentional requirements of the reviews, attention checks were performed in all studies. In the present study, 18 participants failed an attention check and were removed. Across all experiments, 22 of 358 of participants (6\%) were similarly identified and removed. Inclusion of these participants does not materially alter the pattern of results.

A 2 (visual–verbal information) \times 2 (cognitive load) \times 2 (product replicate) \times 3 (feature) mixed ANOVA on relative feature superiority revealed no four-way interaction (\(F < 1\)), but did reveal the critical three-way interaction of visual–verbal information, cognitive load, and feature (\(F(4, 380) = 3.63, p < 0.05\)). Main effects of visual–verbal information (\(F(2, 190) = 3.99, p < 0.05\)), cognitive load (\(F(1, 190) = 7.09, p < 0.05\)), and feature (\(F(2, 380) = 129.51, p < 0.001\)), an interaction of visual information and cognitive load (\(F(2, 190) = 3.07, p < 0.05\)), and an interaction of product replicate and feature (\(F(2, 380) = 20.89, p < 0.001\)) were also observed. This latter interaction reflects a difference in magnitude across the two replicates but is not relevant to our key hypotheses and is not analyzed further. No other effects were significant.

The three-way interaction was further analyzed by feature. As suggested by Fig. 1, when ratings are collapsed across product replicate, a significant interaction of visual–verbal information and cognitive load is observed for the target feature (\(F(2, 190) = 8.12, p < 0.001\)), but not for either of the two equivalent features.
more extreme rating in the presence of mismatched visual and verbal information. However, when participants in this stimulus environment were cognitively constrained, ratings of the target feature were significantly lower than when unconstrained ($M_{Mismatch-HiLoad} = 1.03$ vs. $M_{Mismatch-LowLoad} = 3.10$, $F(1, 190) = 27.14, p < 0.01$). That is, ratings gravitated toward the more attractive product under high cognitive constraints.

When the visual and verbal information were matched in valence (i.e., the superior review was paired with the more attractive product and the inferior review was paired with the less attractive product), ratings in the unconstrained condition were higher than in the control condition ($M_{Match-LowLoad} = 2.64$ vs. $M_{Control-LowLoad} = 1.93$, $F(1, 190) = 3.06, p < 0.05$) but not significantly higher than in the mismatched condition ($M_{Match-LowLoad} = 2.64$ vs. $M_{Mismatch-LowLoad} = 3.10$, $F(1, 190) = 1.27, p > 0.05$). This result is unsurprising, inasmuch as both the visual and verbal information were evaluatively consistent, and therefore pushed participants in the direction of the objectively superior verbal information. In contrast to the mismatched conditions, participants in the visual-information-matched conditions were unaffected by cognitive load, rating the superior feature as equally superior regardless of cognitive constraints ($M_{Match-HiLoad} = 2.64$ vs. $M_{Mismatch-HiLoad} = 2.27, F < 1$).

The non-target features were similar to one another in pattern but differed from that of the target feature. As noted above, the visual–verbal information × cognitive load interaction for both non-target features was not significant in either case, both $p's > 0.20$. Moreover, for both non-target features, there was no significant effect of visual–verbal information, ($F(2, 190) = 1.44, p > 0.20$ and $F(2, 190) = 1.85, p > 0.15$ for non-target 1 and non-target 2, respectively), nor of cognitive load (both $F's < 1$). Descriptive details for all three features are presented in Table 1.

**Reaction times**

To provide additional evidence that the relative performance judgments in the context of incongruity were due to thoughtful reconciliation of the conflicting elements, we also measured the time it took for participants to provide their judgment of the relative superiority of the target feature. If the conflict between the visual and verbal information compels consumers to consider the information more carefully and skeptically (Maheswaran & Chaiken, 1991), judgments should take longer than when the two pieces of information are matched in valence. The critical comparison involves the visual–verbal information-matched condition vs. the visual–verbal information-mismatched condition, both under low cognitive load.² A planned comparison of the two key cells revealed that when aesthetics and reviews were matched, participants took significantly less time to

² Because the cognitive load manipulation makes it more difficult to comprehend the written reviews, and makes people less likely to take the time to process them carefully, the cognitive load manipulation overwhelms the effect of the visual–verbal information match on reaction times under high cognitive load, making the respective comparisons under high load uninformative. Examining the main effect of visual–verbal information on reaction times, the difference between the match and mismatch conditions is attenuated (4.49 s when mismatched vs. 4.31 when matched $p > 0.20$).
make their judgments than when aesthetics and reviews were mismatched ($M_{\text{Match-LowLoad}}=4.04$ s vs. $M_{\text{Mismatch-LowLoad}}=5.01$ s, $F(1, 63)=3.63$, $p<0.05$). In contrast, the non-target features were processed at the same speed, regardless of congruency ($M_{\text{Match-LowLoad}}=4.75$ s vs. $M_{\text{Mismatch-LowLoad}}=5.44$ s, $F(1, 63)=1.84$, $p>0.15$ for non-target 1; $M_{\text{Match-LowLoad}}=4.82$ s vs. $M_{\text{Mismatch-LowLoad}}=5.35$ s, $F(1, 63)=1.16$, $p>0.25$ for non-target 2).

**Discussion**

This study provides evidence of a novel aesthetic effect. The visual information presented to participants provided no information regarding the functionality of the target feature and, in principle, should not have influenced relative judgments. An aesthetic effect was nonetheless observed. When aesthetics and feature functionality were in conflict, consumers’ relative judgments of feature performance were biased in the direction of the unattractive product as compared to the no-picture control condition. Such an outcome is consistent with our proposition that the unexpected negative correlation between aesthetics and feature functionality led to increased elaboration of the target feature and subsequent polarization of performance ratings. The result obtained only when participants had sufficient cognitive resources to consider the information, suggesting that the negative aesthetic effect stems from a thoughtful and deliberate process.

A comparison of the information-matched and information-mismatched conditions provided additional insight into the process underlying the negative aesthetic effect. Participants who viewed products for which the valence of the visual information was mismatched with the valence of the verbal information rated the unattractive product’s target feature more positively when cognitively unconstrained than when constrained. In contrast, participants who viewed products for which the valence of the visual information was matched with the valence of the verbal information gave equivalent relative target feature ratings regardless of cognitive capacity. We have shown that consumers generally expect good products to be attractive. When the attractive product was indeed superior on the target feature, participants were easily able to compare relative performance across the two brands and make their judgments regardless of cognitive capacity. In contrast, when the aesthetics and reviews were in conflict, the burden of cognitive constraints rendered reconciliation of the visual and verbal information difficult, negating the advantage for the unattractive product.

Additional evidence that the negative aesthetic effect stems from elaboration is provided by the response times in the mismatched and matched conditions. Given that the only difference between the two conditions was whether the superior target feature review was matched with the attractive or unattractive product, one might anticipate no difference in judgment latency. However, if a mismatch in valence between the visual and verbal cues fosters elaboration on the target feature (as it should if consumers expect aesthetics and functionality to be correlated), judgment latencies should be greater in the incongruent than congruent condition. Such a difference was observed.

Also critical to our argument is the finding that polarization is restricted to the conflicting target feature. We included measures for two additional features that were equivalent in the reviews to determine whether they would be subject to influence from the aesthetic difference in a manner similar to that of the target dimension. They were not. Rather, the ratings directionally reflected a slight tendency toward the more attractive product, consistent with prior reports of a positive halo effect of aesthetics in ambiguous environments.

**A test for robustness**

Inasmuch as the remaining studies are designed to test boundary conditions, we deemed it important to demonstrate...
that the negative aesthetic effect observed in study 1 is reliable. Hence, we briefly describe a second experiment that uses a sub-
design of study 1 focusing only on the mismatched-visual-
information and no-visual-information conditions. We asked 114 participants to make relative evaluations of two sets of cookware
and two sets of stereo speakers. All participants were cognitively
unconstrained. The stimuli presentation format and procedure
followed that of study 1. Participants read reviews of two brands in
the absence of accompanying visual designs or in the presence of
designs that were mismatched in valence with the accompanying
reviews. The dependent measure was the relative functionality
rating of the target feature, as described in the main study. Results
showed that when the written reviews of the two options were
paired with pictures that conflicted in valence, participants rated
showed that when the written reviews of the two options were
paired with pictures that conflicted in valence, participants rated
the absence of accompanying visual designs or in the presence of
designs that were mismatched in valence with the accompanying
reviews. The dependent measure was the relative functionality
rating of the target feature, as described in the main study. Results
showed that when the written reviews of the two options were
paired with pictures that conflicted in valence, participants rated
the absence of accompanying visual designs or in the presence of
designs that were mismatched in valence with the accompanying
reviews. The dependent measure was the relative functionality
rating of the target feature, as described in the main study. Results
showed that when the written reviews of the two options were
paired with pictures that conflicted in valence, participants rated

Study 1 and its follow-up demonstrated an advantage for
unattractive products that offer superior feature functionality.
Insofar as consumers have sufficient cognitive resources to
recognize and elaborate on the conflict between aesthetics and
performance, they should judge a product feature to be better than
they otherwise would. In the remaining studies we examine the
negative aesthetic effect in two contexts that provide boundary
conditions. Study 2 investigates aesthetic influence in a consump-
tion scenario in which individuals view the products prior or
subsequent to having an opportunity to read about their functional
features. Study 3 examines the context in which participants can
rely on brand names, in addition to aesthetic information, in
forming their judgments.

Study 2

In study 1, participants were simultaneously exposed to both
visual and verbal information. Under such conditions, neither a
cognitive nor an affective evaluation has yet been formed and
both types of information can be considered in the formation of a
judgment. In the marketplace, however, consumers are often
exposed through advertising or store displays to visual appear-
ances prior to learning about product features (Yeung & Wyer,
2004). This lack of simultaneity may provide an opportunity for a
design-driven affective reaction to arise, which may then bias the
interpretation of subsequent objective information in the direction
of the more aesthetically pleasing product. Moreover, without an
opportunity to consider the conflicting visual information and the
verbal information simultaneously, it is less likely that the
inconsistency would be recognized, which in turn would reduce
the likelihood of an elaborative reconciliation of the conflicting
elements. If so, the negative aesthetic effect should be mitigated
and a positive aesthetic effect may even emerge. To test this
possibility, the current study varied the order of presentation of the
visual and verbal information across three conditions. In the
“simultaneous” condition, participants were presented with visual
and verbal information concurrently, as in study 1. In the “visual-
first” condition, participants viewed the pictures of the two
product designs prior to reading the written reviews. The
advantage for the unattractive product should be attenuated
when consumers consider the product designs before reading
about their features. In the “verbal-first” condition, participants
read the reviews of the products prior to seeing the designs. In this
condition, participants could form a judgment of the target feature
in the absence of conflicting information; hence neither a bias in
the direction of the more attractive product nor polarization in the
direction of the unattractive product should be expected.

Method

The study conformed to a 3 (presentation order: visual-first,
verbal-first, simultaneous)×2 (product replicate: gas barbeque
grills and ice cream makers) mixed design, with presentation
order manipulated between subjects and replicate within subject.
As in study 1, the visual designs were pretested to be significantly
different in attractiveness but equivalent with regard to
performance of the functional features, and the corresponding
written reviews were pretested to be significantly different in
functionality of the target feature. For all participants, the designs
and reviews were mismatched in valence; that is, the unattractive
design was paired with the review of the functionally superior
target feature and the attractive design was paired with the review
of the functionally inferior target feature.

Participants were randomly assigned to a computer station. All
participants were informed that they would be reading reviews of
pairs of brands in different product categories, and would render
relative judgments of the brands based on the reviews.
Participants were then presented with the first product category,
randomly determined. Those in the simultaneous condition
saw the reviews and corresponding conflicting pictures of the two
brand options, side-by-side, position counterbalanced, as in study
1. Participants in the visual-first condition viewed the pictures of
the two brands displayed side-by-side on screen for 20 s. The
pictures then disappeared from the screen and were replaced by
the written reviews, also side-by-side, location corresponding to
the previously shown picture. Participants in the verbal-first
condition were presented with the reviews first, and subsequently
were exposed to the pictures of the product designs. After
presentation of both the visual and verbal information, all
participants rated the relative superiority of the target feature on a
9-point scale, where −4 indicated support for Brand A (the more
attractive brand with the objectively inferior target feature), 0
indicated perceived equivalence between the two brands, and +4
indicated support for Brand B (the less attractive brand with the
objectively superior target feature). The procedure was then
repeated for the second product category. Finally participants
completed an attention check and were dismissed. A total of 45
undergraduate students completed the study.

Results and discussion

Descriptive results are presented in Table 2. A 3 (presentation
order)×2 (replicate) mixed ANOVA on the target feature ratings
revealed no effect of replicate and no interaction of order and
replicate (both \( F < 1 \)), but did reveal a main effect of presentation
order \( (F(2, 42)=7.79, p<0.01) \). Collapsed across replicate,
participants’ relative performance ratings for the target feature moved in the direction of the inferior feature (i.e., the more attractive design) when the designs were presented prior to the reviews ($M_{\text{Visual-First}} = -0.60$) relative to when they were presented simultaneously ($M_{\text{Simultaneous}} = 1.33, F(1, 42) = 13.40, p < 0.001$). When reviews were presented prior to the designs, feature ratings were significantly higher than when the designs were presented first ($M_{\text{Visual-First}} = 0.53$ vs. $M_{\text{Visual-First}} = -0.60, F(1, 42) = 4.93, p < 0.05$), but significantly lower when than when the reviews and designs were presented simultaneously ($M_{\text{Visual-First}} = 0.53$ vs. $M_{\text{Simultaneous}} = 1.33, F(1, 42) = 5.18, p < 0.05$).

The results demonstrated that when consumers are provided visual exposure to products prior to learning about them, an initial affective reaction can be resistant to subsequent conflicting verbal information. When the designs were seen prior to the reviews, participants’ judgments of the target feature’s functionality moved in the direction of the objectively inferior, but more attractive brand. Consistent with prior research showing the durability of an affective evaluation made prior to a cognitive one (Edwards, 1990), the aesthetic reaction to the product designs muted the effect of the subsequent written reviews when the designs were viewed first and alone. In contrast, when the verbal information was presented first, and participants were able to judge the functionality of the two options without conflicting visual information, there was no positive effect of the subsequent verbal information. Consistent with our explanation that it is the effortful reconciliation of conflicting elements that drives the advantage for the unattractive product observed in study 1, the verbal-first condition also showed no evidence of a negative aesthetic effect. Only when consumers could consider the visual and verbal information in tandem, thereby providing opportunity to notice the evaluative inconsistency between the designs and reviews, did the negative aesthetic effect obtain.

**Study 3**

The preceding studies demonstrated that when presented with products for which aesthetics and functionality are in conflict—and consumers have the opportunity to elaborate on the conflicting information—relative feature performance judgments will polarize in the direction of the unattractive product. However, our studies thus far have controlled for other marketing variables that could alter the impact of product aesthetics. In real shopping environments, the decision milieu includes many other cues that could influence judgment, including price, store reputation, salespeople, and brand names. This final study examined the potential influence of one of these cues, brand name, that has been shown previously to influence product perceptions. We investigated whether prior beliefs about brand names can alter the effect of aesthetics on feature judgments.

When products are manufactured by established firms, the brand name is a strong signal of performance that can shape product evaluations (Hoyer & Brown, 1990; Page & Herr, 2002). In the realm of product design, Page and Herr (2002) found that aesthetics can play a stronger role than brand name in terms of liking ratings but that, in the case of overall quality judgments, brand names overwhelm aesthetics. Their research did not consider judgments of specific functional features, however. We suspect that unlike its effect on overall quality judgments, a brand name will not overwhelm the negative aesthetic effect in the judgment of relative feature performance. In studies 1 and 2, consumers were provided with objective information regarding the functionality of product features. Respondents could clearly discern that Brand B was better than Brand A on the target feature. The difference in attractiveness was also apparent. Unless a particular brand name is synonymous with visual design (e.g., Apple) or with functionality of a particular feature (e.g., Bose sound quality), the presence of brand information should not alter consumers’ evaluations of the target feature, inasmuch as the brand provides no assistance in reconciling the conflicting visual and verbal information. Given our premise that the negative aesthetic effect is driven by consumers’ attempt to reconcile the unexpected mismatch between the valence of the aesthetic information and the valence of the feature information in the written review, we anticipate feature evaluations to be relatively unaffected by the presence of known brand names. In contrast, because brand names do provide information about overall quality, we expect overall evaluations to be driven by brand perceptions, as reported in prior research.

Study 3 employed a paradigm similar to the one adopted in the preceding studies. Participants were again provided with pictures of product designs and written reviews for two brands. In all conditions, the designs and reviews were mismatched so that the unattractive design was paired with the review of the superior target feature and the attractive design was paired with the inferior target feature. The key manipulation was the manner in which we assigned brand labels to the two options. For half the participants, the less attractive option (with the superior target feature according to the written reviews) was assigned a strong brand name, and the more attractive option (with the inferior target feature) was assigned a weak brand name. For the other half of the participants, the brand assignment was reversed. If brand labels have no influence, feature ratings should favor the unattractive brand in both cases. The other factor examined in this study was brand knowledge. For the brand labels to have an effect they must be known to participants, and the relative difference between the strong and weak brand names must be understood.

**Method**

The study employed a 2 (brand-aesthetic pairing: strong brand paired with unattractive design vs. weak brand paired with unattractive design) × 2 (brand familiarity: yes vs. no) between subjects design. To manipulate brand-aesthetic pairing, we

<table>
<thead>
<tr>
<th>Product</th>
<th>Visual-first</th>
<th>Verbal-first</th>
<th>Simultaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas BBQ grills</td>
<td>-0.90 (2.60)</td>
<td>0.47 (2.58)</td>
<td>1.39 (2.64)</td>
</tr>
<tr>
<td>Ice cream makers</td>
<td>-0.30 (2.54)</td>
<td>0.59 (2.40)</td>
<td>1.28 (2.08)</td>
</tr>
<tr>
<td>Average</td>
<td>-0.60 (1.39)</td>
<td>0.53 (1.05)</td>
<td>1.33 (1.33)</td>
</tr>
</tbody>
</table>

Standard deviations appear in parentheses.
presented brand names along with the product information. Initial testing of awareness of brand names in each of the previously used product categories revealed substantial variance within and across participants. Cookware showed the highest level of brand awareness and was chosen as the target category.

In the strong-brand/unattractive-design condition, the less attractive design (with the superior target feature based on the written reviews) was accompanied by a strong and well-known brand name, T-Fal, and the more attractive design (with the inferior target feature) was accompanied by a weak brand name, Lucky Goldstar. In the weak-brand/unattractive-design condition, brand assignments were reversed.

Brand knowledge was a measured variable. We asked participants two yes/no questions: “Prior to participating in today’s study had you heard of T-Fal before?” and, “Prior to participating in today’s study had you heard of Lucky Goldstar before?” All participants either knew both brands or knew neither brand. We then asked follow-up questions as a brand-strength manipulation check: “Prior to participating in today’s study, how would you have rated the quality of T-Fal cookware?” and, “Prior to participating in today’s study, how would you have rated the quality of Lucky Goldstar cookware?”

The general procedure mimicked previous studies. Participants were presented with the visual and verbal information for the two brands. In this study, however, each product design was accompanied by a strong or weak brand label, depending on condition. After reading the information and viewing the product designs, participants judged the relative performance of the target feature. In addition, participants provided two overall quality judgments (i.e., overall quality and overall superiority), both on 9-point scales anchored by the respective brand names. Ninety-nine undergraduate students completed the study for partial course credit.

Results and discussion

Of the people who were familiar with the brands, four rated Lucky Goldstar as equal or superior in quality to T-Fal and were removed from further analysis. A quality index was created using the superiority and quality measures (α=0.93). An ANOVA on the relative performance rating of the target feature revealed no effect of brand–aesthetic pairing, no effect of brand knowledge, and no interaction (all Fs<1). Participants who did not know the brands judged the objectively superior feature (based on the reviews) as superior regardless of whether a strong or weak brand name was assigned to the unattractive product ($M_{StrongBrand-Unattractive}=1.89$ vs. $M_{WeakBrand-Unattractive}=2.46$, $F<1$). Participants who knew the brands were similarly unaffected by the brand labels, providing similar ratings regardless of the brand–aesthetic pairing ($M_{StrongBrand-Unattractive}=2.41$ vs. $M_{WeakBrand-Unattractive}=2.20$, $F<1$).

In contrast to the target feature ratings, for the overall quality index, we observed a significant interaction of brand–aesthetic pairing and brand knowledge ($F(1, 91)=4.65$, $p<0.05$). Participants who were not familiar with the brands provided equivalent quality ratings in both brand–aesthetic pairing conditions ($M_{StrongBrand-Unattractive}=1.09$ vs. $M_{WeakBrand-Unattractive}=1.47$, $F<1$). Both of these ratings were significantly higher than 0 and suggest that the advantage for the unattractive product carried over to overall judgments. Inasmuch as these individuals were unfamiliar with the brands, the labels could exert no influence. For participants familiar with the brands, on the other hand, the brand labels trumped aesthetics for the overall quality index, consistent with prior research. When the strong brand name was paired with the unattractive design (i.e., with the superior target feature), participants rated overall quality 1.62 but, more important, when the weak brand label was paired with the unattractive design, they rated overall quality 2.20 ($F(1, 91)=7.40$, $p<0.01$). In other words, the strong brand label overwhelmed the negative aesthetic effect and altered overall quality judgments.

General discussion

Three studies demonstrate that product design can alter the manner in which a functional feature is processed. We examined how design influences feature processing when aesthetics and functionality conflict, a situation that consumers do not typically expect, but one that does often arise (e.g., Dyson vacuums, SONY VAIO Ultra Mobile PC, and craigslist.com are thought to be unattractive designs but perceived to have high performing features). Our research uncovered a new aesthetic phenomenon—a bias in the direction of the unattractive product—which we term a negative aesthetic effect. We provide evidence that this bias stems from a thoughtful reconciliation of inconsistent visual and verbal information and elaboration on the conflicting dimension.

Each study pitted product appearance against performance information to address whether aesthetics could influence how a person evaluates feature performance. Study 1 revealed that when consumers have sufficient resources, they will rate a functional feature as better when paired with an unattractive product design than they would without accompanying pictures. The study also showed that the increase in perceived performance occurs only when consumers encounter a conflict between design and functionality. The results supported the premise that it is the violation of expectations regarding the relationship between aesthetics and functionality, along with consequent elaboration on the conflicting functional feature, which leads to the influence of aesthetics on functional performance perceptions. Moreover, study 1 showed that the negative aesthetic effect does not extend to other functional features that are not in conflict with aesthetic information. This result ruled out the possibility of a negative halo and provided further evidence that it is an additional thought about the single conflicting dimension that results in attitude polarization (Tesser & Leone, 1977; Millar & Tesser, 1986).

Studies 2 and 3 examined boundary conditions for the observed effect. Study 2 provided additional evidence that the negative aesthetic effect occurs only when consumers have an

---

1 For comparison purposes we also ran a condition with no brand labels under high cognitive load. We compared this group to the average of the two conditions in which participants were unfamiliar with the brands. This comparison essentially replicates the incongruous visual information conditions of study 1. Consistent with that study, we found that participants who did not know the brands provided significantly higher ratings of the target feature than participants in the comparison condition, $F(1, 108)=4.67 p<0.05$. 

opportunity to reconcile conflicting visual and verbal information. By separating the presentation of visual and verbal information the conflict between design and functionality was less salient and the negative aesthetic effect was attenuated. Moreover, the findings showed that initial exposure to designs prior to learning about their features moved feature judgments in the direction of the more attractive product, that is, the product with the objectively inferior functional feature. Study 3 extended prior research on the interaction of brand name and aesthetics, demonstrating that although brand names dominate overall quality judgments (see Page & Herr, 2002), they do not alter the influence of aesthetics on judgments of feature functionality.

Together the studies provided evidence that aesthetics can play a role in feature performance judgments, and showed that in contrast to previous research and conventional wisdom, there are times when being unattractive can be advantageous. This phenomenon is, we believe, unlikely to occur beyond the realm of product aesthetics. There are many cues that consumers utilize to make product judgments, such as source likeability, word of mouth, and price, and it is not inconceivable that in such cases incongruity could arise. Nonetheless, we argue that a negative value on these other cues is unlikely to lead to a positive outcome, because it would be difficult to reconcile or rationalize the conflict. For example, an unlikeable source should not enhance product judgments. With aesthetics, however, we speculate that reconciliation may stem from consumers’ attempts to rationalize the superior performance of a less attractive product. They may assume that product designers were focused on performance rather than wasting resources on aesthetic appeal. Such a rationalization would not be possible with most other negative cues. This account is, however, merely speculative. A similar possibility is that when consumers focus on the functional aspects of the product, as they are instructed to do in our studies, they may adopt the mindset that an unattractive product form is associated with greater functionality, thereby reinforcing the functionality of the superior functional feature.4 Although we provide evidence that the effect stems from thoughtful reconciliation of conflicting cues, additional research is needed to understand the precise nature of the rationalization that occurs.

Our research focused on the role of aesthetics, but product design can play many other roles, including attracting attention, aiding in categorization, and communicating product functionality (Creusen & Schoormans, 2005). We pretested our designs to ensure they provided no information as to the functionality of our chosen features; nonetheless, an alternative explanation for our findings is that the designs may have communicated some our chosen features; nonetheless, an alternative explanation for our findings is that the designs may have communicated some.

4 We thank one of the anonymous reviewers for this suggestion.

### Acknowledgments

Financial support from the Social Sciences and Humanities Research Council of Canada awarded to JoAndrea Hoegg is gratefully acknowledged.

### Appendix A. Product categories and target features

<table>
<thead>
<tr>
<th>Study</th>
<th>Product category</th>
<th>Target feature</th>
<th>Non-target 1</th>
<th>Non-target 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cookware</td>
<td>Ease of cleaning</td>
<td>Evenness of heat distribution</td>
<td>Well fitting lids</td>
</tr>
<tr>
<td></td>
<td>Electric mixers</td>
<td>Durability</td>
<td>Quietness of operation</td>
<td>Ease of use</td>
</tr>
<tr>
<td>2</td>
<td>Gas barbeque grills</td>
<td>Evenness of heat distribution</td>
<td>Quietness of operation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ice cream makers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Cookware</td>
<td>Ease of cleaning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B. Stimuli examples

**Stimuli Examples**

### Unattractive Design

**Good Target Feature**
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 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.

### Attractive Design

**Bad Target Feature**
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.

### Unattractive Design

**Good Target Feature**
This gas grill is extremely stable and easy to assemble. It is very convenient to operate, but durability can be a question mark. It has a large cooking capacity for outdoor entertaining. The BTUs on this unit are a little low for the size, but because of the efficient design there is little difference in searing quality. This grill will reach grilling temperatures in just a few minutes of preheat time and the control valves and the three position cooking grates make it possible to have great control over the grilling process. The elements are well balanced leaving no cold corners. It is easy to move around and is relatively lightweight to be transported when necessary.

### Attractive Design

**Bad Target Feature**
This grill can be preheated quickly and offers complete infrared power from three burners. The 50,000 BTU output allows cooking at extremely high temperatures giving excellent searing power as long as items are kept toward the center of the grilling area. Control valves enable simultaneous high and low temperature grilling when desired. With a full rotisserie, multiple cooking grates and good controls you get everything with this grill, but the thickness of the metal casing of the unit limits the long term durability. Assembly is straightforward, and once in place, the unit is simple with a push button ignition and easy to read dials.

*Note: The titles for the stimuli elements (“Unattractive Design,” “Attractive Design,” “Good Target Feature,” “Bad Target Feature”) are included here for clarity but were not presented in the actual studies.*

**References**


