Beauty in a Bottle: Package Aesthetics Cues Efficacy Beliefs of Product Performance

Aparna Sundar, University of Cincinnati, USA
Karen Machleit, University of Cincinnati, USA
Theodore Noseworthy, University of Guelph, Canada

Beauty-in-a-bottle is a visual metaphor for an unspoken promise when objective evaluation is difficult. Three experiments and a field study demonstrate that packaging aesthetics serves to inform purchase intent of consumers at the point of decision. Efficacy beliefs inform purchase intent but this happens only in the beauty category.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1015605/volumes/v41/NA-41

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
Beauty in a Bottle: Package Aesthetics Cues Efficacy Beliefs of Product Performance
Aparna Sundar, University of Cincinnati, USA
Theodore J. Noseworthy, University of Guelph, USA
Karen A. Machleit, University of Cincinnati, USA

ABSTRACT

While consumers often shop in outlets with large product assortments, having no prior experience or knowledge to predict a product’s efficacy can pose a significant challenge at the point of purchase. Indeed, the functionality of most consumer goods tends to be a bit of a black box, at least until consumed. Imagine a consumer trying to select a new face cream. The majority of the ingredients may as well be illegible, and the promotional campaigns, brand awareness, and product claims tend to sit on par with several competing products in the same category. So which product does the consumer choose? This is where consumers make heuristic judgments. They infer quality from price and visual aesthetics. According to Proctor & Gamble (P&G) shoppers make up their mind about a product in three to seven seconds, just the time it takes to note a product on a store shelf. P&G refers to this time lapse as the “first-moment-of-truth.” Hence, companies like P&G spend considerable resources to enhance the appeal of product packaging.

Indeed, the role of package aesthetics in enhancing product preference is well-documented (Bloch 1995; Townsend, Montoya and Calantone 2011, Berkowitz 1987, Hertenstein, Platt and Veryzer 2004). When consumers are faced with a choice between two products of equal price and equivalent function, they tend to prefer the more attractive option (Kotler and Rath 1984). Hence, it is not surprising that in the absence of prior use or knowledge of the product, consumers will make initial impressions based on package aesthetics (Chandon, Hutchinson, Bradlow and Young 2009). However, one cannot help but question why this happens. Most day-to-day purchases are not just for hedonic gain, but utilitarian value. Given the sheer efficiency of the first-moment-of-truth heuristic, it would be surprising if the judgment had absolutely nothing to do with the product’s functionality, particularly given the strong relationship between form and function (Berkowitz 1987; Hertenstein et al. 2004).

In exploring this question, we make several advances. First, we demonstrate that while consumers show preferences for more appealing packaging in general, their belief in the product’s efficacy seems to be a function of their level of awareness of the category and the product’s aesthetic appeal. Specifically, consumers who have limited knowledge of a category tend to judge how well a product will function based on how pretty its packaging is. Second, we demonstrate that these inferences translate into predictable and meaningful differences in purchase intent and choice. Finally, we go one-step further by showing that this effect is most pronounced in what we refer to as credence categories (products that promote outcomes that cannot be objectively evaluated). Hence, the story comes together to illustrate that with little knowledge to the contrary, consumers tend to assume that the beauty products work better if the product’s packaging is congruent with its positioning. Hence, the first-moment-of-truth heuristic may have more to do with beliefs in efficacy than with aesthetic preference.

CONCEPTUAL BACKGROUND

Researchers have devoted considerable attention to models of processing that impact purchase intent (Bloch 1995, Park, Lennon and Stoel 2005, Orth and Stel 2005, Deng and Kahn 2009). When consumers are faced with a choice between two products of equal price and equivalent function, they tend to prefer the more attractive option (Kotler and Rath 1984). Hence, it is not surprising that in the absence of prior use or knowledge of the product, consumers will make initial impressions based on package aesthetics (Chandon, Hutchinson, Bradlow and Young 2009). It is well-documented that aesthetics in package design can augment product preference (Hoegg, Alba and Dahl 2010). Without prior knowledge, consumers use aesthetic cues to inform purchase intent (Raghunib and Greenleaf 2006). It has also been shown that aesthetic information plays a critical role in product trials (Childers and Jiang 2008, Wyer, Hung and Jiang 2008). Prior research has addressed several aspects of package design that influence purchase intention including packaging size (Do Vale, Pieters and Zeelenberg 2008), location (Deng and Kahn, 2009), color (Folkes and Matta 2004; Yang and Raghunib 2005), shape (Raghunib and Krishna 1999), graphics (Bone and France 2001), and feel (Eriksson and Larsson 2011, Krishna 2006). In the absence of any other information, or when prior information about a new brand is missing (Postrel 2003, Horsky and Honea 2009), visual information of a product or package (color, design, graphics, etc.) serves to set consumption expectations and make purchase decisions (Bloch 1995, Di Muro and Noseworthy 2013, Kotler and Rath 1984, Noseworthy and Goode 2011, Sujan 1985). Yet, when a consumer knows how a product functions, the symbolic, ergonomic and functional value as communicated by design (Hoegg and Alba 2011) becomes redundant. While research in this area has focused on the influence of package aesthetics on satisfaction, product usage, and consumption perception (Shiv et al. 2005), the effect of package appeal in determining efficacy of an unknown product has not been studied.

Past research categorizes products on store shelves in terms of how easy or difficult it is to evaluate the utility of the product (Hsieh, Chiu and Chiang 2005). Products that consumers have no object way to evaluate are termed credence goods (Lovelock 2008, Brush and Artz 1999, Klein 1998). The two other types of products are search and experience products. Search products are those that the consumer can obtain multiple levels of information on, upon which they would base evaluations. The experience products are those that consumers evaluate based on actual consumption experience (Darby and Karni 1973, Zeithaml 1988). Examples of credence products are beauty products whose utility is hard to evaluate. Despite evidence that packaging influences claim receptivity for credence goods, there is little empirical work examining how variations in packaging form affect the desirability of credence goods. The promise of “beauty in a bottle” is a compelling proposition to any consumer. We contend that, in the absence of opportunities for consumers to make quality evaluations, aesthetics in packaging cues the semantic link between the category and functional efficacy of beauty. We show that a key determinant of a product’s success is its efficacy or perceived product effectiveness, particularly when prior knowledge is either absent or low.

Brand awareness usually cues expectations of how the product will function when prior usage or trial is absent (Hoyer and Brown, 1990). Nedungadi and Hutchinson (1985) define awareness as the ability to bring a brand to mind. Awareness serves as a heuristic when consumers lack time or prior usage experience. Horsky and Honea (2009) note that in cases of inferior product quality, packaging biases a positive response to hedonic product experience in which consumers were more likely to accept credence claims (products whose utility impact was difficult for consumers to ascertain). Consumers show a marked preference for a high-awareness brand even when there were known differences in quality and price (Macdonald and

Sharp, 2000). Nedungadi and Hutchinson (1985) note that any new information about a category is usually generalized asymmetrically between more- and less-known brands.

We contend, however, that in the absence of consumer awareness, the aesthetics of the product or the beauty-in-a-bottle promise serves as the first-moment-of-truth heuristic. This hypothesis is tested in study 1. Our hypothesis is that, in the absence of any other cues to a brand’s functionality, consumers look to packaging information to form purchase inferences, and this is mainly due to functional inferences based on appeal. Study 2 demonstrates that awareness or expertise in the category moderates this effect as consumers use the beauty-in-a-bottle heuristic only for low-awareness products. Study 3 is a field experiment that demonstrates the robustness of this effect across a variety of brands in the beauty and household category. The hypotheses tested are summarized as follows:

**Hypothesis 1:** In the absence of any other cues to a brand’s functionality, consumers look to packaging information to form purchase inferences and indicate a higher preference for more aesthetically pleasing packaging.

**Hypothesis 2:** Expertise in the category moderates the above effect as consumers use the beauty-in-a-bottle only in case of low awareness products to make functional inferences.

**Hypothesis 3:** The beauty-in-a-bottle effect is moderated by product category and only informs products in the beauty category.

### STUDY 1

**Design and Procedures.** Participants (N = 75, 58% females; M<sub>age</sub> = 36.5) were recruited via Amazon.com’s crowd-sourcing website, Mechanical Turk, to complete a brief (10 min) questionnaire on “product preferences” in exchange for a nominal fee ($0.20; see Paolacci, Chandler and Ipeirotis [2010] for a validation of Mechanical Turk as a data source). Participants were told that we were interested in their shopping preferences, and were asked to imagine shopping in a store for a bottle of lotion. After simulating the scenario, participants were instructed that they would be asked to indicate their preferences for specific products. As before we captured purchase intention and perceptions of product effectiveness. The product form manipulation followed study 1. In addition to rating the appeal of various package form manipulations, participants were asked to list and rate the familiarity of brand names (anchored: 1 = not very familiar; 5 = very familiar). The list included fictitious and well-known brands in the category. The brand names with the highest and lowest means ratings were carried forward into the study. All other measures remained consistent with study 1.

**Results and Discussion.** An analysis of variance with purchase intent as the key dependent measure, yielded a significant awareness × appeal interaction (F(1, 125) = 4.35, p < .05). Simple effects confirmed that when awareness was high, participants did not show a significant difference in purchase intent between the more appealing (M = 3.27) and less appealing product (M = 3.51; F < 1). However, when awareness was low, participants preferred the product with the more appealing packaging (M = 3.06) over the product in the less appealing packaging (M = 2.12; F(1, 125) = 5.35, p < .05). To determine whether perceived efficacy once again accounted for the increased desirability of less aware products, we conducted a moderated mediation analysis (Preacher, Rucker and Hayes 2007; Model 2). As predicted, perceived efficacy mediated the positive relationship between appeal and purchase intention when the product had a low awareness (B = .86, SE = .29, z = 2.96, p < .005) but not when brand awareness was high (z = -.18, p = .85). This finding supports the prediction that when awareness is low consumers use heuristics such as package design to determine product preferences, as consumers make efficacy inferences based on the appeal of the product, but when it comes to high awareness products, the heuristic is rendered irrelevant. Hence, consistent with literature on expertise, product awareness washed the heuristic. This raises the distinct possibility that this heuristic, like many others, is domain specific. That is, the link between beauty and efficacy might only exist for a credence good that is promoting beauty, as opposed to functional utility. Study 3 was designed to test this prediction, and also to check the robustness of our findings.

### STUDY 3

This study uses a unique data set detailing 146 brand names of beauty and household products. It includes information about the products featured including images of the actual images of products.
used in store. In this study the objective was to evaluate the role that category plays in the beauty-in-a-bottle heuristic. We predicted that the link between beauty and efficacy exists only for products for which it is hard to ascertain functionality or utility ahead of time.

BrandSpark International, in collaboration with SSI, conducted a nationwide online survey of over 50,000 American consumers (age 18+) who are involved in household shopping and provided the data for our study. The survey was weighted by gender, age and region to conform to US MRI data and reflect the principal shopper. Of the individuals who participated in the nationwide survey 57,492 individuals (10.6% male and 89.4% female) participated in an additional survey that evaluated preferences to products that were either in the beauty (credence) or household (experience) categories. A total of 106 images of beauty products and household products were obtained from individual brand representatives who were Brandsparks clients. Images were all clients’ representations of the actual product package as used in store with the product against a white background. All images used were shown the using the same approximate size and with clear resolution.

Independent Measures. Each respondent saw one image with the product description and then answered a series of questions such as: How appealing is the package to you? Prior to this survey were you aware of this product? Have you or have you not purchased this product before? Approximately how many times have you purchased the product? The first time you purchased this product, please indicate what had the most influence on your decision to buy the product? How much does the packaging of this product encourage you to purchase this product? Respondents would then see a different product and repeat the series of question for the product presented. When repeated observations are collected on a set of individuals and the measurement occasions are not identical, such as measuring appeal, purchase intent and awareness of individual products, Raudenbush (2004) recommend treating multiple observations as nested within persons. We chose full maximum likelihood as the estimation method to compare the model fits across three nested models (Raudenbush, Bryk and Congdon 2002). In our model, each individual’s appeal rating is nested within products. We used hierarchical linear modeling (HLM; Raudenbush, Bryk and Congdon 2002) to deal with multilevel data as it allows for simultaneous estimation of the relationship of variables at two levels, using iterative maximum likelihood estimation. We estimated a main effect model, an interaction model with main effects, and second and third order interaction with a random intercept and finally an interaction model with main effects, second and third order interaction with a random intercept and appeal. This final model allowed for individual variation in the level of product appeal.

Results and Discussions. The parameter estimates for all three models are contained in Table 2. Note that we report unstandardized coefficients. Model 3 had the smallest fit indices (AIC=1081349), suggesting a good fitting, parsimonious model. Given the sample size, we assumed this to be a reliable (Diamantopoulos and Siguaw 2000). We found an effect of package appeal ($\gamma_{10} = 1.93$, $t = 157.82$, $p < .001$), awareness ($\gamma_{10} = .25$, $t = 16.15$, $p < .001$) and a marginal effect of category ($\gamma_{10} = -.02$, $t = -1.75$, $p = .08$). The main effects were qualified by a significant three way interaction ($\gamma_{107} = -.013$, $t = -6.56$, $p < .001$). In addition, a spotlight analysis at two standard deviations above and below the mean for appeal, and awareness were computed for beauty vs. household products. Consumers in the household category showed increased preference for high appeal package regardless of whether the awareness was high ($M_{low\text{ appeal}} = 2.15$ vs. $M_{high\text{ appeal}} = 5.64$) or low ($M_{low\text{ appeal}} = 1.58$ vs. $M_{high\text{ appeal}} = 4.75$). In contrast, while consumers in the beauty category showed a similar preference in a low awareness condition ($M_{low\text{ appeal}} = 1.73$ vs. $M_{high\text{ appeal}} = 5.12$) the effect of package appeal did not matter in the high awareness condition ($M_{low\text{ appeal}} = 3.81$ vs. $M_{high\text{ appeal}} = 3.81$).

This field study highlights the schematic concept of beauty or package aesthetics being important when awareness of a product is low and which is more the case for credence goods. For products associated with the beauty category, we find that package appeal plays a critical role in the absence of brand cues to the quality or other characteristics of the product. A summary of findings are in table 1.

GENERAL DISCUSSION

Our research contributes to the literature on product aesthetics by showing how consumers use inferences based on package appeal to determine preferences of other heuristics to base preferences. Research has shown package form to play a critical role in persuasion tactics that marketers use by varying shape, visual graphics, materials, and proportions (Bloch 1995). Product form takes on multiple roles in communicating to the consumer (Hoegg and Alba 2011), including attention, schematic mapping of product category, basic arousal, and usage-related and functional meaning. In this research, we demonstrate that inferences regarding performance are generated purely based on package appeal when considering products for which utility and hence functional value are difficult to evaluate. We also show that these inferences are critical in driving trial for products that are hard for consumers to evaluate, even after they have consumed or used them (credence goods). Thus, we contribute to existing knowledge in this area by evaluating the influence of appeal in consumer preferences for products that require expertise and are challenging to evaluate.

This research illuminates the nuances of inferences that consumers make based on appeal. We demonstrate that, although products in appealing packaging are preferred in general, it is only when there is a lack of expertise or awareness in a category that appeal cues functional inferences. Further, studies 1 and 2 show the utility of this heuristic with credence goods, such as those in the beauty category, wherein it is difficult to determine an objective means of evaluation. Finally, by combining empirical analyses of multiple brands across different product types, we bolster the generalizability of the results while highlighting the differences in implications for credence versus experience products. By drawing on the product aesthetics literature, including findings from statistical analyses of field survey data and controlled experiments, we further enhance our contribution to theory and marketing practice.

Certain limitations to this research should be considered. In studying credence goods, our research has focused on the beauty category, because the appeal metaphorically maps to the concept of beauty and hence to the “beauty-in-a-bottle” metaphor. This heuristic is certainly applicable to other areas of product aesthetics in which the limitations of credence goods or functionality of the product are difficult to evaluate. Further, while study 3 is useful in determining the limiting conditions of this phenomenon, future studies could explore the role that perceptions of efficacy play in different categories of credence goods. Future studies could also benefit by examining the moderating role of product awareness caused by lack of expertise, cross-cultural differences, gender differences, and cross-generational differences.

REFERENCES

Table 1  Mean Purchase Intent

<table>
<thead>
<tr>
<th>Study</th>
<th>High Appeal</th>
<th>Low Appeal</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>3.74</td>
<td>2.55</td>
<td>$F = 2.55, p &lt; .008$</td>
</tr>
<tr>
<td>Study 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Awareness</td>
<td>3.06</td>
<td>2.12</td>
<td>$F(1,125) = 5.35, p &lt; .02$</td>
</tr>
<tr>
<td>High Awareness</td>
<td>3.27</td>
<td>3.51</td>
<td>$F(1,125) = .37, p = NS$</td>
</tr>
<tr>
<td>Field Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beauty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Awareness</td>
<td>5.12</td>
<td>1.73</td>
<td></td>
</tr>
<tr>
<td>High Awareness</td>
<td>3.81</td>
<td>3.81</td>
<td></td>
</tr>
<tr>
<td>Spotlight Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Awareness</td>
<td>4.75</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td>High Awareness</td>
<td>5.64</td>
<td>2.15</td>
<td></td>
</tr>
</tbody>
</table>

Table 2  Results of the HLM model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1: $\gamma$ (SE)</th>
<th>Model 2: $\gamma$ (SE) Interaction model with random intercept only</th>
<th>Model 3: $\gamma$ (SE) Interaction model with random intercept and appeal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>2.07 (.018**)</td>
<td>1.91 (.020**)</td>
<td>1.87 (.019**)</td>
</tr>
<tr>
<td>Appeal ($\gamma_{01}$)</td>
<td>1.75 (.004**)</td>
<td>1.87 (.012**)</td>
<td>1.93 (.012**)</td>
</tr>
<tr>
<td>Awareness ($\gamma_{02}$)</td>
<td>0.20 (.004**)</td>
<td>.22 (.015**)</td>
<td>0.25 (.015**)</td>
</tr>
<tr>
<td>Category ($\gamma_{03}$)</td>
<td>-.23 (.005**)</td>
<td>-.047 (.012**)</td>
<td>-.020 (.011)</td>
</tr>
<tr>
<td>Appeal X Awareness ($\gamma_{04}$)</td>
<td>.081 (.017**)</td>
<td>.069 (0.017**)</td>
<td></td>
</tr>
<tr>
<td>Awareness X Category ($\gamma_{05}$)</td>
<td>-.007 (.018)</td>
<td>-.020 (.018**)</td>
<td></td>
</tr>
<tr>
<td>Appeal X Category ($\gamma_{06}$)</td>
<td>-.159 (.014**)</td>
<td>-.192 (.013**)</td>
<td></td>
</tr>
<tr>
<td>Appeal X Awareness X Category ($\gamma_{07}$)</td>
<td>-.151 (.021**)</td>
<td>-.138 (.021**)</td>
<td></td>
</tr>
<tr>
<td>Gender ($\gamma_{08}$)</td>
<td>-.005 (.009)</td>
<td>-.009 (.009)</td>
<td>.004 (.009)</td>
</tr>
</tbody>
</table>

Fit Statistics

-2ResLL=1086552  AIC=1086566  BIC=1086629
-2ResLL=1085781  AIC=1085803  BIC=1085901
-2ResLL=1079925  AIC=1081349  BIC=1081465

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


Horsky, Sharon and Heather Honea (2009), “Do We Judge a Book by its Cover and a Product by its Package? How Affective Expectations are Contrasted and Assimilated into the Consumption Experience,” *Advances in Consumer Research, 36*.


