TAKUYA SATOMURA, MICHEL WEDEL, and RIK PIETERS*

The authors propose a method and metric to quantify the consumer confusion between leading brands and copycat brands that results from the visual similarity of their packaging designs. The method has three components. First, image processing techniques establish the objective similarity of the packages of leading and copycat brands on the basis of their colors and textures. Second, a perceptual decision task (triangle test) assesses the accuracy and speed with which consumers can identify differences between brands from rapidly (300 milliseconds) flashed images of their packages. Third, a competing accumulator model describes the buildup of evidence on each of the alternative brands during consumers' perceptual decisions and predicts the accuracy and response time of brand identification. Jointly, these components establish the impact of copycat packaging's visual features on consumer confusion. The method is applied in a test of experimentally designed copycats and market copycats in 15 product categories. A three-tiered metric ("copy alert," "copy watch," and "copy safe") establishes the extent to which copycat brands imitate the package designs of target brands and identifies which visual features are responsible.

Keywords: copycat, brand confusion, packaging, perceptual decision, image processing

Copy Alert: A Method and Metric to Detect Visual Copycat Brands

Consumers are often duped by copycats of leading brands. As many as two-thirds of shoppers report that copycat packaging confuses them, and one-third admit to have even bought the wrong brand because of its similar packaging (Poulter 2009). Leading brands develop unique visual identities for their package designs to communicate with customers and stand out on the shelf (Warlop and Alba 2004), but copycat brands continue to hijack these identities by copying visual features such as colors and textures (Jacoby and Morrin 1998; Kapferer 1995; Miceli and Pieters 2010). Leading brands are rightly concerned about these copycat strategies. Copycat brands reduce the effectiveness of leading brands to stand out in cluttered retail environments, dilute their trademarks, hurt their equity, and erode the return on their financial investments. In many cases, this leads manufacturers of the target brands to bring the copycat brands to court (Zaichkowsky 2006).

Although such infringement court cases are mounting, in practice there is no single standard method of assessing the potential for trademark dilution (Bird 2007). Academic marketing research has investigated brand confusion caused by copycats (Kapferer 1995; Miaoulis and D'Amato 1978; Miceli and Pieters 2010; Warlop and Alba 2004) and has documented the harm from trademark dilution (Morrin and Jacoby 2000; Morrin, Lee, and Allenby 2006; Pulling, Simmons, and Netemeyer 2006; Van Horen and Pieters 2012). Despite their insights, these studies have provided neither objective measures of the visual similarity between leading brands and copycat brands nor quantitative metrics to assess the risk of brand confusion. These shortcomings have prevented wide-scale application of academic copycat research in managerial decision making and brand infringement liti-

*Takuya Satomura is Professor of Marketing, Faculty of Business and Commerce, Keio University (e-mail: satomura@fbc.keio.ac.jp). Michel Wedel is PepsiCo Professor of Consumer Science, Department of Marketing, Robert H. Smith School of Business, University of Maryland (e-mail: mwedel@rhsmith.umd.edu). Rik Pieters is Professor of Marketing, Tilburg School of Economics and Management, Tilburg University (e-mail: pieters@uvt.nl). The authors appreciate valuable comments from two anonymous *JMR* reviewers and acknowledge financial support from the Yoshida Hideo Memorial Foundation Research Grant 2009. The first author also acknowledges financial support from the Grants-in-Aid for Scientific Research (No. 23530542) in Japan. Teck-Hua Ho served as associate editor for this article.

© 2014, American Marketing Association
ISSN: 0022-2437 (print), 1547-7193 (electronic)

Journal of Marketing Research
Vol. LI (February 2014), 1–13

gation. Moreover, prior research typically has relied on similarity judgments after long exposures to brand packages, from several seconds up to one minute (Morrin and Jacoby 2000; Morrin, Lee, and Allenby 2006; Pulling, Simmons, and Netemeyer 2006; Warlop and Alba 2004). Such long exposures do not reflect real-life exposure conditions well, especially for packaged goods categories in which copycats are pervasive. In front of the shelves in retail stores, consumers often do not pay careful attention to the visual information on the packages of all stockkeeping units in the category. They quickly search for a product on the shelf and rely on first impressions to pick up their favorite brand. On the shelf, many packages may not receive more than a single eye fixation (a period of up to half a second in which the eye is still; Chandon et al. 2009; Van der Lans, Pieters, and Wedel 2008). Therefore, exposures to brand packages during choice in front of the shelf are often much shorter than a second. Methods to assess the effects of copycatting need to mimic such exposure conditions.

The present study caters to these needs. We propose a method and metric to assess the confusion between leading brands and copycat brands that results from the visual similarity of their package designs. Consumers' perceptual decisions about brands form the core of this approach. It has three components. First, image processing techniques determine the objective visual similarity of packages on the basis of their colors and textures. Second, a perceptual decision task determines the accuracy and speed with which consumers can identify the unique package from a set of three rapidly shown (300 ms) packages, of which two are identical. Third, a competing accumulator model describes the accumulation of evidence on each of the alternative brands during the decision task to establish accurate brand identification by consumers and identify which visual information accounts for it. Using this method, we propose a quantitative metric for copycat confusion.

It is important to note that our research pertains to confusion caused by copycats imitating visual features of the packages of leading brands. There are also instances in which copycat brands more subtly imitate aspects of the visual appearance of a leading brand with the aim of exploiting positive associations related to it (Van Horen and Pieters 2012; Warlop and Alba 2004), without aiming to create brand confusion. In addition, copycats can imitate more abstract themes and messages of a leading brand without directly copying visual features (Miceli and Pieters 2010). Although these copycat strategies are also important, this study does not address them. Rather, we focus more narrowly on copycat confusion caused by similarity of the visual design of packages, that is, on "attribute-based copycats" that imitate perceptual features of the packaging of a leading brand, such as its colors, shapes, and textures. These copycats seem to be the most often encountered on the market and dominant in infringement cases (Zaichkowsky 2006).

THEORETICAL FOUNDATION

Consumers make a perceptual decision when they determine whether they are viewing the package of a leading brand or the package of a copycat brand. Perceptual decisions involve discriminating between a set of alternatives on the basis of some visual input (Heekeren, Marrett, and Ungerleider 2008). To mimic conditions under which consumers make decisions in front of a shelf, our experimental setup includes a perceptual decision task in which participants identify, after a brief exposure, which package is different (the target) out of a set of three packages, two of which are identical. Such so-called triangle tests are common in perceptual identification research (Ennis and Jesionka 2011). We question how participants determine which of the three alternatives the target is, how accurate and fast their decision is, and which visual features of the brands contribute to their choice.

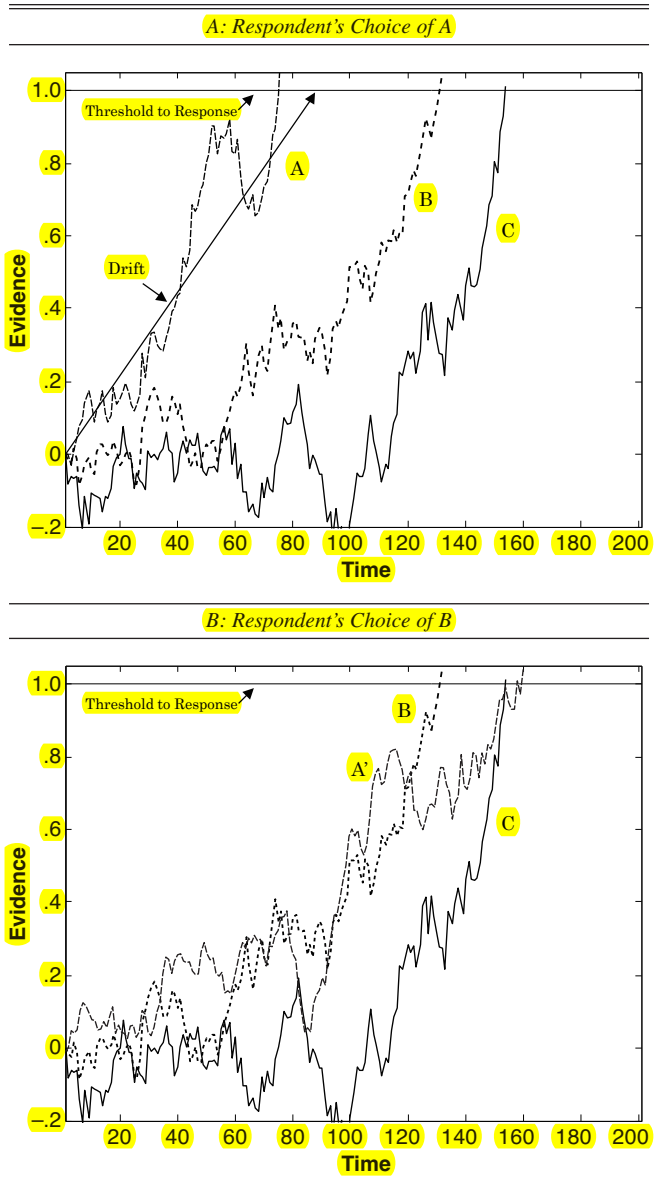
Competing accumulator models have been developed to capture the process underlying perceptual decisions (Smith and Ratcliff 2004; Usher, Olami, and McClelland 2002). Researchers have predominantly applied these models to two-alternative forced choice tasks and shown them to explain two essential properties of the data obtained from those tasks: that reaction times are right-skewed and that people make correct responses more quickly than incorrect ones (Smith and Ratcliff 2004). These competing accumulator models operate on the following assumptions about the underlying mechanisms: (1) in continuous time, the sensory input on each of the alternatives provides noisy information for the decision, (2) this information is integrated (accumulated) over time in a mental "accumulator" that is specific for each alternative and yields evidence favoring that particular alternative, and (3) a decision is made when the amount of accumulated evidence for one of the alternatives reaches a threshold. Figure 1, Panel A, illustrates this accumulation process for three hypothetical alternatives (A, B, and C).

Because sensory input is noisy, the evidence favoring each of the alternatives is stochastic rather than deterministic, which results in the so-called directed random walk or drift-diffusion models. For clarity, we refer to these models as "competing accumulator models" because they are based on the assumptions that the evidence for each alternative is stored in a separate neural accumulator and that these accumulators compete in reaching the threshold. Competing accumulator models provide descriptions of decision processes from first principles because they reflect the neurological processes underlying decision making (Heekeren, Marrett, and Ungerleider 2008). Specifically, in the case of perceptual decision tasks, researchers have shown competing accumulator models to describe the gradual increase in neural activity in competing groups of neurons in higher visual areas that represent the alternatives (Heekeren, Marrett, and Ungerleider 2008). Importantly, these models have been proven theoretically to result in optimal decisions. Although most extant work has focused on forced choices between two alternatives, scholars have developed extensions to more than two alternatives as well (Usher, Olami, and McClelland 2002), in which the model typically describes the difference in evidence favoring the alternatives. The deterministic "drift" of the process affects how quickly information accumulates and reaches a decision threshold.

Models that assume nondecaying accumulating information provide good descriptions of perceptual decisions (Leite and Ratcliff 2010), and the model by Usher, Olami, and McClelland (2002) is particularly attractive because it is amenable to Bayesian estimation. (We describe their

Figure 1

DYNAMICS OF THE COMPETING ACCUMULATOR MODEL



Notes: A, B, and C are brands competing for identification. The respondent chooses the brand for which the evidence crosses the threshold first (i.e., in Panel A, Brand A, and in Panel B, Brand B).

model in more detail subsequently.) An important aspect that is missing from Usher, Olami, and McClelland's (2002) model and other competing accumulator models is that they do not allow for exogenous variables to influence the rate of drift of the accumulating evidence on the alternatives. However, such exogenous variables, in the form of the visual features of the alternatives, are available and relevant in the context of copycat detection. Specifically, the similarity of the leading brand's and copycat brand's packaging, in terms of their visual features, is likely to influence the rates of drift (Van Vugt et al. 2013). Colors and textures of brand packages are a common focus in copycat litigations (Zaichkowsky 2006) and are among the basic features extracted by visual areas in the brain. Colors and textures provide viewpoint-invariant descriptors of scenes and objects, and their spatial distribution is widely employed in computer vision to mimic the way the human visual system processes visual information (Choi et al. 2002).

In summary, when consumers are exposed to the three packages in our task, the visual brain extracts color and texture information from the alternatives. The strength of the evidence that an alternative is the target (i.e., the package of the three that is different) is a function of its similarity with the other packages in terms of colors and textures. These similarities are the input to the perceptual decision, and evidence on each alternative accumulates until a threshold is reached. The proposed model represents this process through a race of accumulation processes. The visual information extracted from each alternative in the set is assumed to be noisy and nonnegative and to follow a diffusion process with drift. The drift is a function of the dissimilarity of an alternative with the others in terms of its colors and textures, and it determines how quickly information accumulates. Evidence accumulates over time for each of the alternatives separately until one reaches a threshold (Brand A in Figure 1, Panel A). Higher thresholds result in more accurate, but slower, responses. Because of random fluctuations in the strength of the visual input, evidence for a nontarget package may reach the threshold first, resulting in an error. Slower accumulation is associated with more difficult decisions: if a copycat is more similar to the target, its drift will be smaller, resulting in a longer time for the process to reach the threshold as well as a higher error rate (Brand A′ in Figure 1, Panel B). The model thus predicts correct responses by identifying the unique brand to accumulate more quickly and the incorrect brands to accumulate more slowly (Leite and Ratcliff 2010). We provide mathematical details of the model and our extension after describing the data.

DATA

Products and Image Processing

We selected 15 product categories (adhesive bandages, allergy medicine, canned noodles, chewy bars, corn flakes, crackers, dishwashing liquids, hand sanitizers, honey, ketchup, mayonnaise, mouthwash, raisin bran, sports drinks, and tissues). For each category, we selected two existing brands: the leading brand and a visually similar (national or store) brand. Although the latter brands are not strictly "copycats" (i.e., they are not necessarily known to have copied visual features of the leading brand with the aim of confusing consumers), for ease of terminology, we loosely refer to them as "market copycats." We obtained images of the packages with a digital camera under standardized distance, angle, exposure, and lighting conditions. To produce an image, we selected the foreground by using Adobe Photoshop CS5 and resized the image to 200~400 pixels per side.

In addition to the leading brands and market copycats, we also experimentally designed copycats by manipulating the images of the leading brand in each category with Adobe Photoshop CS5. We created these "experimental copycats" by substituting details (e.g., logo, label, pictures) of the original image of the leading brand in each category with those of the market copycat brand. In many cases, this process created experimental copycats that were visually

very similar to the leading brand, which is desirable to test the proposed methodology and provide a benchmark for comparison. Appendix A presents the images for the three options in each category.

Participants and Tasks

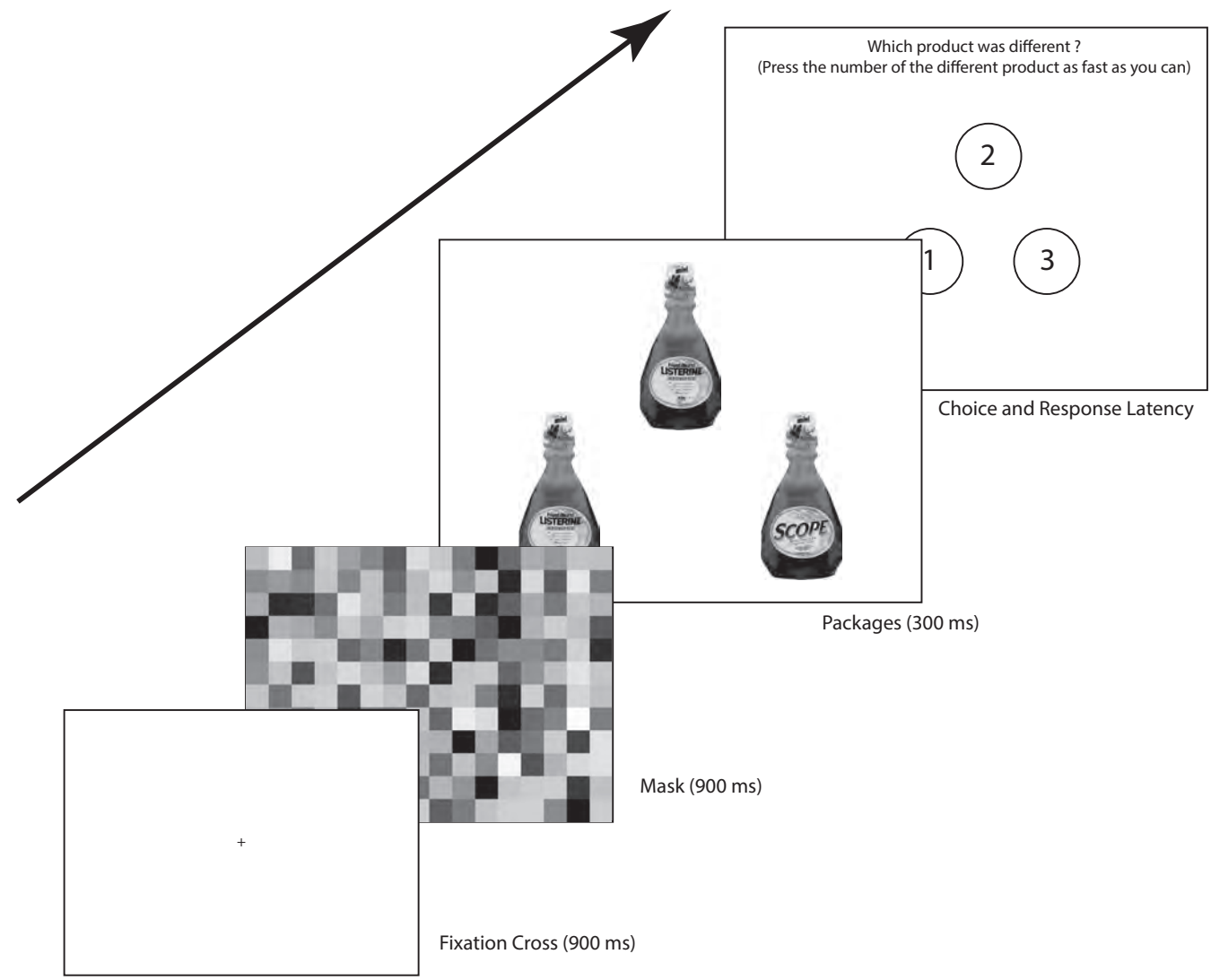
Ninety-nine paid undergraduate students participated in this study. Data collection took place individually in the Behavioral Lab of Tilburg University in the Netherlands. Instructions and stimuli were presented on 19-inch LCD monitors in full-color bitmaps with a 1,280 × 1,024-pixel resolution.

We propose a perceptual decision task based on the triangle test from sensory research (Ennis and Jesionka 2011). In the triangle test, a participant inspects three alternatives (two being the same and one being different) and identifies the alternative that is different. In our task, participants are simultaneously presented with images of three packages for 300 milliseconds. This duration prevents participants from making eye movements. The participant's task is to decide as accurately and as quickly as possible which of the three packages is the different one. This modification of the triangle test exposes participants briefly to multiple packages, similar to exposures to multiple packages on shelves in real-life conditions.

After a practice trial, the main test involved 30 trials. For each of the 15 categories, participants were exposed to two sets of three packages. One set included a leading brand and the market copycat brand. The other set of packages included the leading brand and the experimental copycat. We randomized order of categories and sets of packages for each participant.

Figure 2 shows the flow of the perceptual decision task. First, a fixation cross displayed for 900 milliseconds, followed by a visual mask for 900 milliseconds. Then, a set of three packages displayed for 300 milliseconds. Finally, the (blank) locations of the three packages displayed, and participants indicated which of the packages just shown was

Figure 2
THE PERCEPTUAL DECISION TASK: TRIANGLE TEST



different by pressing the corresponding number key. We repeated this procedure for all package triplets so that the task consisted of 30 trials for each participant. We excluded eight participants, whose mean response time was less than 300 milliseconds or who had any response time less than 100 milliseconds. We recorded the accuracy (1, 0) and latency (in milliseconds) of the response for each of the 30 trials. Appendix B provides the observed accuracies and latencies of correct and incorrect responses for each category.

Visual Similarity of Images

We computed the visual similarity between each pair of candidate brands on the basis of color and texture descriptors of their images. We then used these visual similarities as input to the competing accumulator model. We selected commonly used color and texture descriptors in copycat cases (Zaichkowsky 2006). Specifically, a color histogram descriptor (CHD) describes the entire global color layout of the image, and a dominant color descriptor (DCD) captures the most representative colors in the image. In many categories, one or two colors are used by most brands as "category codes"; for example, most orange juice packaging contains the color orange. Distinguishing these two types of color similarities enables us to establish whether copycat confusion occurs through usage of the overall colors and/or through the dominant color. Texture histogram descriptors reflect the orientation and layout of textures in the image. These texture descriptors capture detailed visual information on the packages, including package labels, logos, shapes, and forms. Next, we describe details of the specific measures.

Color histogram descriptors mimic the way the human visual system processes color information (Shapiro and Stockman 2001). We divided each image into a 4 × 4 grid, yielding 16 subimages. We then categorized colors in each of the 16 subimages into 64 bins on the basis of their red, green, and blue (RGB) values, resulting in a total of 1,024 histogram bins. Let a_m and b_m be the number of pixels at the mth bin of the histogram of images A and B, respectively. If M is the total number of bins in the histograms, we obtain the similarity between two images A and B (D_{AB}) by computing the following:

$$D_{AB} = \sum_{m=1}^{M} \min\left(\frac{a_m}{\sum_{j=1}^{M} a_j}, \frac{b_m}{\sum_{j=1}^{M} b_j}\right). \quad (1)$$

The DCD provides a description of the most representative colors in an image (Deng et al. 2001). The DCD is denoted as $F = (c_r, p_r)$, $(r = 1, 2, ..., R)$, where R is the number of dominant colors. For an image, each dominant color is represented by c_r, a vector of color space component values,[1] and p_r, the fraction of pixels corresponding to c_r, where $\Sigma_{r=1}^{R} p_r = 1$. We used standard methods to compute R, c_r, and p_r for each image. Let F_A and F_B be the DCD of images A and B: $F_A = (c_{Ar}, p_{Ar})$, $(r = 1, 2, ..., R_A)$, and $F_B = (c_{Bs}, p_{Bs})$, $(s = 1, 2, ..., S_B)$. Then, the DCD similarity $E(F_A, F_B)$ is defined as follows:

$$E(F_A, F_B) = \sum_{r=1}^{R_A} p_{Ar}^2 + \sum_{s=1}^{S_B} p_{Bs}^2 - \sum_{r=1}^{R_A}\sum_{s=1}^{S_B} 2a_{Ar,Bs} p_{Ar} p_{Bs} \quad (2)$$

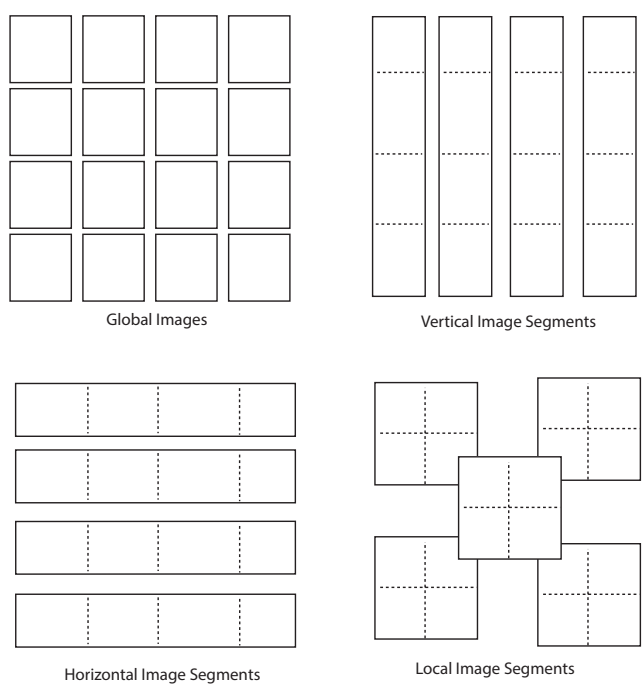
$$a_{r,s} = \begin{cases} 1 - d_{r,s}/d_{max} & d_{r,s} \le T_d \\ 0 & d_{r,s} > T_d \end{cases},$$

where $d_{r,s} = \|c_r - c_s\|$ is the Euclidean distance between colors c_r and c_s, T_d is the maximum distance for two colors to be considered similar, and $d_{max} = \alpha T_d$. Using prior conventions (Deng et al. 2001), we set $T_d = 10$ and $\alpha = 1.2$.

To compute texture similarities between the images of packages, we counted the directions of edges in each image, produced a histogram that summarized the frequencies of these patterns, and calculated the distance between the histograms of the images of different packages as follows: We first computed the edges in the images of the packages using the canny edge detector (Shapiro and Stockman 2001). Then, we divided the images into subimages on the basis of their spatial layouts. Figure 3 shows the u = 1, ..., 4 spatial image layouts we used for this purpose, which we refer to as global, horizontal, vertical, and local spatial layouts. We used 16 subimages for the global spatial layout, 4 subimages for the vertical and horizontal spatial layouts, and 5 subimages for the local spatial layout. For each subimage in each of the four image layouts, we categorized patterns of edges into one of all possible 25 types, representing all the types that are possible based on "masks" of 3 × 3

Figure 3

IMAGE LAYOUTS TO COMPUTE TEXTURE SIMILARITIES



[1]Color space component values are the values of the elements in a system defining the colors of each pixel in the image. For example, the common RGB color space indicates how much red, green, and blue make up the color of each pixel. The color can be expressed as an RGB triplet, each component of which can vary from zero to a defined maximum value. In this study, we use a 3-D vector in the CIELAB color space, which is a perceptually uniform color space (Deng et al. 2001).

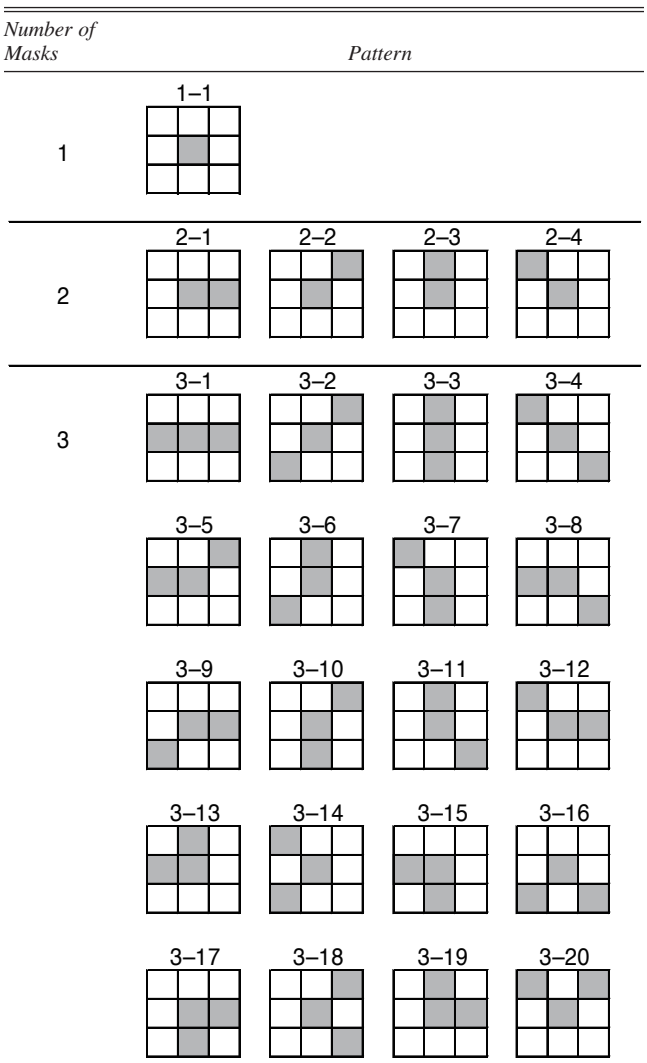
pixels (see Figure 4). We scanned each of the subimages with each of the 25 masks (i.e., moved each mask virtually across the image), counted the 3 × 3 pixel areas in the image that matched each mask, and summarized the counts in a histogram.

This texture histogram descriptor captures the orientation of detailed information in the visual image of each package (Choi et al. 2002). Because there are 16 subimages for the global spatial layout, there are a total of 400 histogram bins for the global layout. Similarly, there are 4 subimages and 100 histogram bins for the vertical and horizontal spatial layouts and 5 subimages and 125 histogram bins for the local spatial layout. Let $a_{m,u}$ and $b_{m,u}$ be the number of pixels in the mth bin of the histogram for spatial layout u. If M is the number of bins in the histograms, we obtain the texture similarity $F_{AB,u}$ as follows:

$$F_{A,B,u} = \sum_{m=1}^{M} \min\left(\frac{a_{m,u}}{\sum_{j=1}^{M} a_{j,u}}, \frac{b_{m,u}}{\sum_{j=1}^{M} b_{j,u}}\right). \tag{3}$$

Figure 4

MASKS USED TO DEFINE TEXTURES FROM EDGES



Notes: We modified these masks in line with Kubo et al. (2003).

All texture and color similarity measures take on values between 0 (different) and 1 (same) so that their coefficients in the accumulator model are directly comparable. For each target image (A), we computed each of the similarities, D_{AB}, E_{AB}, and $F_{A,B,u}$ (for u = 1, ..., U) with the other two packages (B) shown simultaneously in the triangle task. Appendix B shows the global color and texture similarities between the packages in all categories. The measures are correlated and range from .52 to .93, which may result in some collinearity when including them simultaneously in our model.

COMPETING ACCUMULATOR MODEL

Figure 5 provides a flow diagram of the entire approach. In this section, we present details of the model representing the perceptual decision process. The model extends Usher, Olami, and McClelland's (2002) work by modeling the drift as a function of the objective visual similarity of the images and by including unobserved heterogeneity in the drift parameters across participants. There are k = 1, ..., K alternatives (packages) in category j (K = 3). Let t denote decision time and let $w_{i,j,k}$ be the amount of visual information extracted by participant i from package k in category j during exposure to the three alternatives. The term $w_{i,j,k}$ follows a Wiener diffusion process (Brownian motion), which means that infinitesimal changes in $w_{i,j,k}$ are assumed to consist of a drift term $\mu_{i,j,k}$ and random noise that follows a normal distribution with mean 0 and variance σ_i^2:

$$dw_{i,j,k} = \mu_{i,j,k}dt + \sigma_i\sqrt{dt}. \tag{4}$$

It can be shown that the distribution of the time t it takes the accumulated information to reach an individual-specific threshold θ_i follows an inverse Gaussian distribution (Chhikara and Folks 1988). Formally,

$$f\left(t\middle|\theta_i, \mu_{i,j,k}\right) = \frac{\theta_i}{\sqrt{2\pi t^3}}\exp\left[-\frac{\left(\theta_i - \mu_{i,j,k}t\right)^2}{2t}\right]. \tag{5}$$

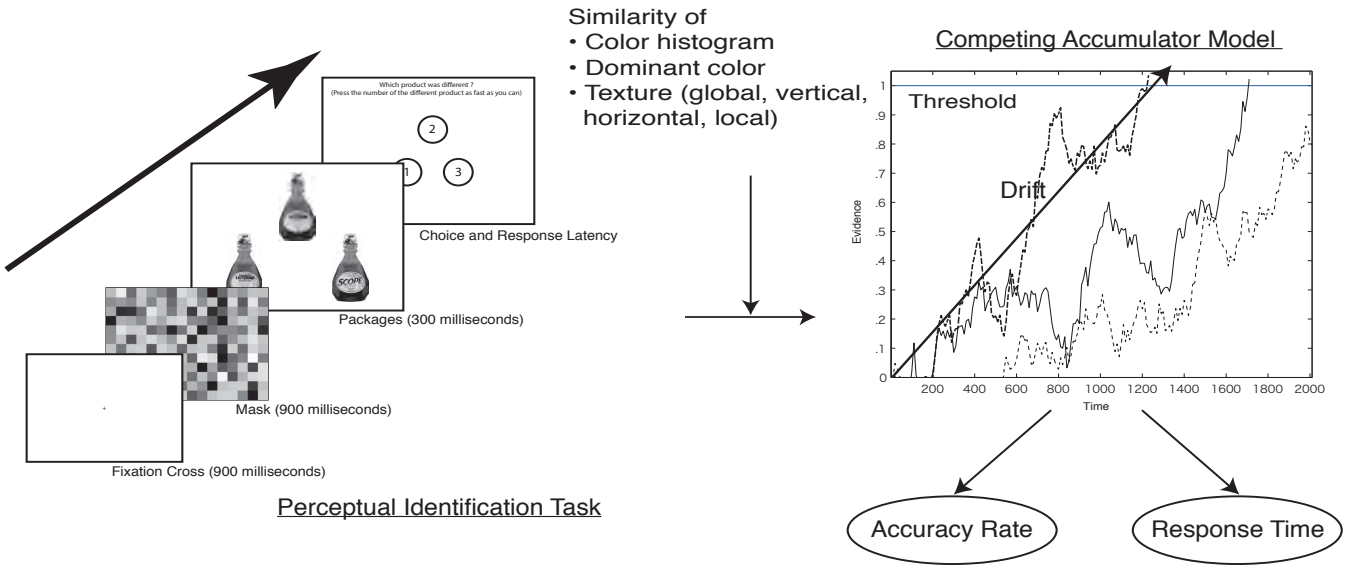
Here, σ_i is set to 1 for identification. A larger value of the threshold θ_i implies that it will take longer for the accumulated information to reach the threshold so that responses will be slower but more accurate. We parameterize the threshold as $\theta_i^* = \log(\theta_i)$ so that it is constrained to be positive during the estimation of the model.

We assume that the rate at which information accumulates (drift) is a weighted sum of the visual similarities of the image of the target with those of the other two packages. Thus,

$$\log\left(\mu_{i,j,k}\right) = \beta_i' x_{j,k} + \gamma_{i,j}. \tag{6}$$

Here, β_i is an individual-specific random parameter. We use a log-link to constrain the drift to be positive. Furthermore, $\gamma_{i,j}$ is a category- and individual-specific random effect so that the model accommodates unobserved differences between categories and participants. A larger value of this parameter increases the drift so that information accumulates faster, which leads to faster and more accurate responses. This result could be due to, for example, (unobserved) familiarity with or frequency of use of category j. In Equation 6, $x_{j,k}$ contains $D_{j,k}$, the similarity of the image in

Figure 5
COMPONENTS OF THE METHOD FOR VISUAL COPYCAT DETECTION



terms of overall color composition; $E_{j,k}$, its similarity in terms of dominant color composition; and $F_{j,k,u}$, its texture similarity for the global, horizontal, vertical, and local layouts. Thus, the drift is a weighted sum of these similarities (Van Vugt et al. 2013). The terms $D_{j,k}$, $E_{j,k}$, and $F_{j,k,u}$ are mean-centered in each triangle test. A larger value of a specific element of β_i increases the drift as a function of the corresponding similarity measure, which then contributes to faster responses. Thus, the speed of the accumulation of information is a function of the similarity between one package and the other packages in terms of color and texture features. All parameters are individual specific and assumed to arise from normal population distributions:

$$(7) \qquad \left(\beta_i', \gamma_i', \theta_i^*\right) = \Omega_i \sim N\left[\left(\bar{\beta}', \bar{\gamma}', \bar{\theta}^*\right), V_\Omega\right].$$

Note that our focus is on the likelihood of an *incorrect* perceptual decision, which indicates brand confusion. The probability that an incorrect answer has been provided by time t can be derived as

$$(8)\ G_k\left(t\middle|\theta_i, \mu_{i,j,k}\right) = \int_0^t g_k\left(t\middle|\theta_i \mu_{i,j,k}\right)dt$$
$$= \tfrac{1}{2}\left\{1 + \mathrm{erf}\left[\frac{\theta_i}{\sqrt{2}t}\left(\frac{\mu_{i,j,k}t}{\theta_i} - 1\right)\right]\right\}$$
$$+ \tfrac{1}{2}\exp\left(2\mu_{i,j,k}\theta_i\right)\left\{1 - \mathrm{erf}\left[\frac{\theta_i}{\sqrt{2}t}\left(\frac{\mu_{i,j,k}t}{\theta_i} + 1\right)\right]\right\},$$

where erf represents the error function. The joint probability for a participant i to choose alternative k in category j after duration t is

$$(9)\ \rho\left(y_{i,j} = k, t\right) = g_k\left(t\middle|\theta_i, \mu_{i,j,k}\right)\prod_{m \neq k}\left[1 - G_m\left(t\middle|\theta_i, \mu_{i,j,m}\right)\right].$$

The probability that participant i chooses k in category j is therefore the probability function $\rho(\cdot)$ in Equation 9 integrated throughout the distribution of the decision time t: $P(y_{i,j} = k) = \int_0^\infty \rho(y_{i,j} = k, t)dt$. In the experiment, participants engage in H = 2 perceptual decision tasks for each category. Thus, $y_{i,j,h}$ is the choice outcome for the hth task in category j. Assuming independence of the choices in each category, the likelihood of the sequence of choices for the two tasks and the J = 15 categories of participant i is

$$(10) \qquad L_i = \prod_{j=1}^{J}\prod_{h=1}^{H}\rho\left(y_{i,j,h} = k, t_{i,j,h}\right),$$

and the overall likelihood is $L = \prod_i L_i$.

Appendix C provides details of the Markov chain Monte Carlo (MCMC) estimation procedure and the prior settings (for Bayesian estimation of accumulator models, see also Vandekerckhove, Tuerlinckx, and Lee 2011). We run the MCMC algorithm for 120,000 iterations, checking convergence from the iteration plots, and use every tenth draw of the last 20,000 iterations to compute the posterior means and standard deviations of the parameters. To predict the accuracy rate, we obtain $P(y_{i,j} = k)$ by numerical integration $p(y_{i,j} = k) = \int_0^\infty \rho(y_{i,j} = k, t)dt$, using adaptive quadrature. We calculate this probability for each of the target draws from the posterior distribution of the parameters. Let $P^q(y_{i,j} = k)$ be the probability obtained from the draws of the parameters in the qth MCMC step, with q = 1, ..., Q. The expected probability is then approximated by $P(y_{i,j} = k) \approx \frac{1}{Q}\Sigma_q P^q(y_{i,j} = k)$.

RESULTS

We estimated our competing accumulator model on the accuracy and response-time data from the study. We estimated a sequence of 12 models to establish which color and texture descriptors significantly affect accuracy and response times in the decision task on the basis of their log-marginal likelihoods (Chen 2005). Table 1 presents the results.

Models 1–3 comprise all possible combinations of the two CHDs. Model 1 includes only the global color similarity (CHDs). Model 2, which has a lower log-marginal likelihood, includes only the DCD, and Model 3 includes both color similarity measures and has a higher log-marginal likelihood than Models 1 and 2. However, the effect of DCD in Model 3 is insignificant, and only CHD has a significant effect. Therefore, in the subsequent models, we always include the global color similarity (Model 1) and add various combinations of the texture similarity measures. In Models 4–7, we add the texture similarity measures to Model 1 one by one. Model 4, with global texture similarity next to global color similarity, has the highest log-marginal likelihood of these four models. To establish whether combinations of texture similarities account for the data better than Model 4, Model 8 adds both vertical and horizontal texture similarity to Model 1, and Model 9 further adds local texture similarities to Model 8. Both have a lower log-marginal likelihood than Model 4. Therefore, the subsequent models all contain global color and global texture similarity and expand on Model 4. Model 10 adds vertical and horizontal texture similarity to Model 4, and Model 11 adds local texture similarity to Model 4, but both log-marginal likelihoods are lower than that of Model 4. Furthermore, we estimated the model with all color and texture similarities (Model 12), but this led to the lowest log-marginal likelihood. This model also suffers from some collinearity, which affects the stability of the estimates. Therefore, we selected Model 4, with global color and global texture similarity, as the best model.

We also estimated a model in which the value of random category effect, $\gamma_{i,j}$, was set to 0 to investigate the effects of this rich parameterization of the intercept on the estimates. The log-marginal likelihood of this model is –5,429.7, which is lower than that of Model 4.[2] Finally, for comparison, we estimated a simpler model that describes the accuracies with a multinomial and the log-response times with a normal distribution and includes the global color and texture similarities as predictors.[3] The log-marginal likelihood of this model is –5,871.4, which is much lower than that of Model 4, testifying to our assertion that the competing accumulator model better captures the data-generating mechanism.

Table 2 gives the posterior means and 95% credible intervals for the thresholds ($\bar{\theta}^* = \log \bar{\theta}$), the effects of the two similarity measures ($\bar{\beta}$), and the covariance matrix of the random effects (V_Ω). To save space, estimates of the random category effects ($\bar{\gamma}$) are not shown. The estimate of the global color and global texture similarity is positive and significant (Table 2). This indicates that participants rely on the global distribution of colors and textures in the package design to identify the candidates. There is substantial heterogeneity in the estimates (see Table 2), reflecting individual differences in the extent to which colors and textures drive copycat identification. The estimate of texture similarity is larger, and there is more heterogeneity in its effect compared with that of color similarity, indicating that individual participants rely to a greater or lesser extent on this visual feature.

Table 3 presents the observed and the predicted (in-sample) accuracy rates and response times and their 95% credible intervals. The estimated accuracy rates and response times

[2]The estimates of this model are log θ = .74, global color = 1.12, and global texture = 1.63, which are close to the values of Model 4 (log θ = .75, global color = 1.00, global texture = 2.17), as we discuss subsequently.

[3]$\Pr(y_{i,j} = k) = \exp(\delta_i\beta_i'x_{j,k})/\{\Sigma_m \exp(\delta_i\beta_i'x_{j,m})\}$, $\log(t) = \beta_i'x_{j,k} + \gamma_{i,j} + \varepsilon_{i,j}$, $\varepsilon_{i,j} \sim N(0, s_i^2)$. The term δ_i is the scale parameter (Wedel and Pieters 2000).

Table 2

ESTIMATES OF THE PARAMETERS OF THE COMPETING ACCUMULATOR MODEL

		Similarity of Visual Features (β)	
	Threshold (log θ)	*Global Color*	*Global Texture*
Parameters	**.75**	**1.00**	**2.17**
	[.64, .86]	**[.33, 1.67]**	**[1.23, 3.22]**
(Co)variance (V_Ω)			
Threshold	**.24**	–.04	–.03
	[.22, .38]	[–.17, .08]	[–.21, .13]
Global color		**1.00**	–.30
		[.60, 1.82]	[–1.19, .13]
Global texture			**1.48**
			[.78, 3.43]

Notes: Posterior means with 95% credible intervals in brackets. Boldfaced estimates indicate that the credible interval does not cover zero.

Table 1

MODEL COMPARISONS

	Visual Features That Contribute to Brand Similarity						
	Color Similarity		*Texture Similarity*				
Model	*Global*	*Dominant*	*Global*	*Vertical*	*Horizontal*	*Local*	*Log-Marginal Likelihood*
1	X						–5,414.6
2		X					–5,468.1
3	X	X					–5,383.9
4	X		X				–5,381.8[a]
5	X			X			–5,419.5
6	X				X		–5,411.7
7	X					X	–5,653.7
8	X			X	X		–5,414.0
9	X			X	X	X	–6,983.0
10	X		X	X	X		–5,565.2
11	X		X			X	–5,414.6
12	X	X	X	X	X	X	–1,1438.8

[a]Denotes the highest log-marginal likelihood value.

are close to the observed accuracy rates not only overall but also for both the experimental and market copycats individually. The observed and predicted accuracy rates for the experimental copycats are lower than those for the market copycats, as we expected, and the response times are longer. The 95% credible interval of the predictions contains the observed values in all cases, with the exception of the accuracy for the market copycats. The model underestimates the accuracy for the market copycats by approximately 5% and overestimates the response times by a little less than 100 milliseconds. These differences may have been caused by some participants' prior familiarity with the leading brands and market copycats, for which the model did not accommodate. Figure 6 plots the observed and predicted accuracy rates at the brand level. It shows that the model predicts the brand-level accuracies well but slightly underpredicts the high accuracy rates for market copycats in the tissues and adhesive bandages categories.

Copycat Metric we like aspects of visual sim. design by aor confusion?

We propose a copycat metric (CC metric hereinafter) as a diagnostic of the perceived similarity of the package design of a copycat to that of a leading brand. First, we obtain $p_{a,j}(y_{i,j} = \mathcal{T}) = \int_0^\infty \rho(y_{i,j} = \mathcal{T}, t)dt$, the predicted accuracy of identifying the target package $\mathcal{T}$ for category j, the one in the triplet that is different (by numerical integration for each of the draws of the parameters). Let $p_{a,j}$ be the observed accuracy rate in the sample and $p_{d,j}$ be the true proportion of people in the population who discriminate the packages. The relation between $p_{a,j}$ and $p_{d,j}$ is $p_{a,j} = \frac{1}{3} + \frac{2}{3}p_{d,j}$, because the probability of identifying one of them as different when all three alternatives are identical is $\frac{1}{3}$. The CC metric is then defined as $CC = 1 - p_{d,j} = \frac{3}{2}(1 - p_{a,j})$, where $0 \le CC \le 1$; $CC = 1$ when the alternatives in the triplet are identical (i.e., $p_{a,j} = \frac{1}{3}$), the case of a copycat that exactly replicates the leading brand.

We develop critical limits for the CC metric using the power analysis for the triangle test (Ennis and Jesionka 2011). The number of correct responses in the triangle test is a random variable $Y_j = \Sigma_{i=1}^{N} I(y_{i,j} = \mathcal{T})$ that is binomially distributed: $Y_j \sim B(p_{a,j}, N)$. We aim to test H_0: $p_{a,j} = \frac{1}{3}$. From the smallest threshold T_α that satisfies $P(Y_j \ge T_\alpha) \le \alpha$ under H_0, the critical value CC_α can be calculated as $CC_\alpha = \frac{3}{2}(1 - T_\alpha/n)$. Here, α is the probability of incorrectly deciding that there is no copycat in the set. Some experimentation with our results enables us to propose a three-tiered classification based on $\alpha = .1$ and $\alpha = .001$, resulting in the cutoff values $CC_\alpha = .841$ and $CC_\alpha = .742$, and we label the corresponding critical regions of the CC metric as "copy alert," "copy watch," and "copy safe." In practical applications, users can choose to specify other values of α depending on their desired risk of leaving a copycat unidentified (Type I error).

Table 4 represents the CC metrics as well as the copycat classification of the market copycat and the experimental copycat in our application. Within the 15 categories, we observe eight copy alerts and one copy watch among the experimental copycats. These results provide external validity for our procedure and copycat metric because the procedure identifies two-thirds of the experimental copycats to be too similar to the leading brand in the category. Note that we have set a strict cutoff value to prevent false positives. Among the market copycats there are no copy alerts, but a copy watch arises in four instances. Again, these results

Table 3

OBSERVED AND PREDICTED ACCURACY AND RESPONSE TIME

	Overall	*Experimental Copycat*	*Market Copycat*
Observed accuracy rate	.520	.446	.593
Predicted accuracy rate	.499	.457	.540
	[.472, .524]	[.436, .478]	[.508, .570]
Observed response time (in milliseconds)	1,160	1,226	1,094
Predicted response time (in milliseconds)	1,213	1,240	1,185
	[1,050, 1,396]	[1,073, 1,428]	[1,026, 1,362]

Notes: 95% credible intervals are in brackets.

Figure 6

OBSERVED AND ESTIMATED ACCURACY RATES FOR EACH PRODUCT CATEGORY

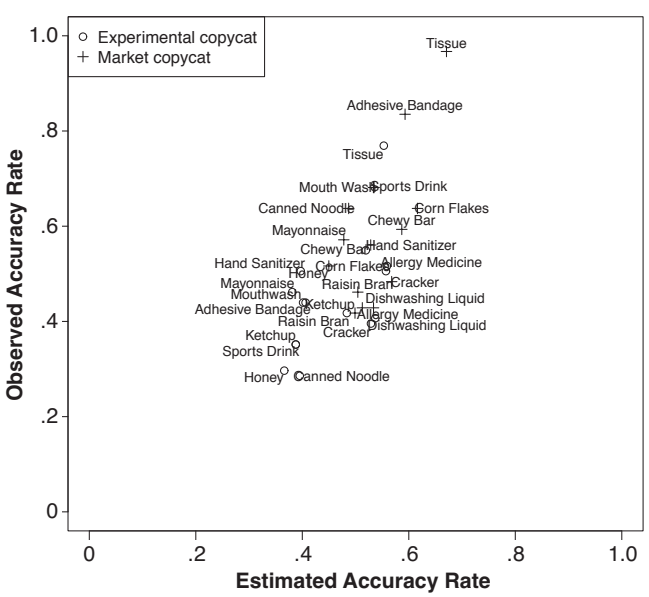


Table 4

CC METRICS

Category	*Experimental Copycat*	*Market Copycat*
1 Adhesive bandages	.894[b]	.638
2 Allergy medicine	.662	.723
3 Canned noodles	.901[b]	.769[a]
4 Chewy bars	.717	.624
5 Corn flakes	.662	.588
6 Crackers	.695	.640
7 Dishwashing liquids	.685	.689
8 Hand sanitizer	.901[b]	.721
9 Honey	.949[b]	.830[a]
10 Ketchup	.917[b]	.755[a]
11 Mayonnaise	.925[b]	.788[a]
12 Mouthwash	.887[b]	.710
13 Raisin bran	.770[a]	.741
14 Sports drinks	.915[b]	.709
15 Tissues	.664	.515

[a]Copy watch ($CC_\alpha = .742$).
[b]Copy alert ($CC_\alpha = .841$).

show face validity because we selected mature categories in which copycats of leading brands are likely to already have been eliminated through litigation or mediation.

Figure 7 shows the images of the leading brand and the market and experimental copycats in categories in which the CC metric of an experimental copycat is greater than .9 (resulting in a copy alert in all cases). The figure shows face validity for our CC metric because it indeed identifies cases in which the copycat is visually very similar to the leading brand.

Sensitivity Analysis

The key settings in computing the image similarities are the 64 color histogram bins and the 16 subimages used to compute global color and texture similarity (Figure 3; the size of the texture mask is the largest feasible size, and the number of texture histogram bins follows directly from it). We investigated the extent to which these settings affect the CC metric. We used 32 (= 2^5), 64 (= 2^6), and 128 (= 2^7) color histogram bins and 16 (4 × 4), 64 (8 × 8), and 256 (16 × 16) subimages in a fractional factorial design (one-third replicate of a 3^3 design). For each of the resulting nine combinations, we computed the global color and texture similarity measures, estimated the model, and computed the CC metric. Table 5 reports the mean absolute errors (MAEs) of the CC metric relative to the preferred settings. It shows that these settings do not have much impact on the CC metric because the average value of the MAE equals .015 and its maximum is .021. This finding supports the settings we chose in our study.

Figure 7
LEADING BRAND VERSUS EXPERIMENTAL AND MARKET COPYCATS

Category	Leading Brand	Experimental Copycat	Market Copycat
Honey			
Mayonnaise			
Ketchup			
Sports drinks			
Hand sanitizers			
Canned noodles			

Notes: CC metric of experimental copycat > .9.

CONCLUSION

Imitation, lookalike, and copycat brands flood national and international markets and threaten the equity of leading brands. Despite the prevalence of these practices, research on the topic is still limited, and established methodologies are lacking. We propose a method and metric to detect copycat brands on the basis of the visual features of their packages. The method consists of three integrated components: a triangle decision task to assess the speed and accuracy of identifying copycat packages; image processing to identify the visual features that are involved; and a competing accumulator model to describe the decisions made during the task.

Methodologically, our approach extends prior accumulator models of perceptual decisions in the mathematical psychology literature by formulating the drift of the accumulation process as a function of visual similarity and accommodating unobserved heterogeneity of parameters. Our method provides a metric for copycat detection that can be used in managerial decision making and litigation. Note that our research focuses on brand confusion caused by similarity of the visual design of packages and not on cases in which a copycat imitates abstract themes and messages of a leading brand without directly copying visual features or cases in which the copycat brand imitates the appearance of a leading brand with the aim of exploiting its positive associations. Our study of experimental copycats supports the validity of the proposed approach and CC metric, which is relatively insensitive to the settings of image processing methods used in extracting visual features. Incentive alignment, in which participants are rewarded for correct responses as in practice, could further enhance the external validity of our procedure.

Empirically, our results shed light on the role of different types of visual information in copycat confusion and reveal the importance of color. Yet the dominant color in a package design seemed unimportant in distinguishing copycats from target brands. This may be due to the prevalence of visual category codes, colors shared by all brands in a product category (e.g., red for ketchup, green for dishwashing liquid, yellow for honey). In those categories, consumers might not rely on the dominant color to distinguish brands. Rather than the dominant color, it was the entire global spatial color layout that contributed to the identification of the copycat. This finding may have resulted from the specific categories and copycats used in our study. In categories in which leading brands have a single unique color (e.g., the purple color of Milka chocolate bars), the dominant color might drive copycat detection more strongly. We leave this issue for further study.

Global textures also played an important role in copycat confusion. Different spatial layouts of textures, including vertical, horizontal, and local layouts, did not add to the explanatory power of global textures in our model, which may have been caused in part by some correlation between

Table 5
MAES OF THE CC METRIC FOR VARIOUS SETTINGS OF THE IMAGE SIMILARITY MEASURES

Number of color bins	32	32	32	64[a]	64	64	128	128	128
Number of color subimages	16	64	256	16[a]	64	256	16	64	256
Number of texture subimages	64	16	256	16[a]	256	64	256	64	16
MAE for CC metric	.010	.011	.021	0[a]	.015	.017	.016	.016	.013

[a]The setting used in the present study.

these measures. The present results are consistent with the finding that during brief exposures, people accurately identify advertising images on the basis of mostly coarse visual information (Pieters and Wedel 2012). Our findings are also consistent with the literature on gist perception, which has uncovered the role of the global layout of stimuli in identification (Oliva and Torralba 2006).

Some packages of leading brands stood out because of their colors and textures. Yet even though we studied mature categories, in which litigation and mediation have likely weeded out clear-cut copycats, we still observed copy watches, thus emphasizing that leading brands should remain vigilant. Copycatting is widespread not only in package designs but also in website, advertising, logo, store, and product design. Further research could apply the proposed approach and the CC metric to these other stimuli. The proposed three-tiered classification may help brand managers monitor the competition, preserve the integrity of their products, and determine which copycats are "too close for comfort" and should be challenged in court. The CC metric captures a critical aspect of brands' visual performance at the "moment of truth," which previous copycat research seems to have neglected: how brands can stand out to consumers during the very short exposures they have on the shelf.

REFERENCES

Bird, Robert C. (2007), "The Impact of the Moseley Decision on Trademark Dilution Law," *Journal of Public Policy & Marketing*, 26 (Spring), 102–117.

Chandon, Pierre, J. Wesley Hutchinson, Eric T. Bradlow, and Scott H. Young (2009), "Does In-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase," *Journal of Marketing*, 73 (November), 1–17.

Chen, Hing-Hui (2005), "Computing Marginal Likelihoods from a Single MCMC Output," *Statistica Neerlandica*, 59 (1), 16–29.

Chhikara, Raj S. and J. Leroy Folks (1988), *The Inverse Gaussian Distribution*. New York: CRC Press.

Choi, Yanglim, Chee Sun Won, Yong Man Ro, and B.S. Manjunath (2002), "Texture Descriptors," in *Introduction to MPEG-7: Multimedia Content Description Interface*, B.S. Manjunath, Philippe Salembier, and Thomas Sikora, eds. New York: John Wiley & Sons, 213–30.

Deng, Yining, B.S. Manjunath, Charles Kenney, Michael S. Moore, and Hyundoo Shin (2001), "An Efficient Color Representation for Image Retrieval," *IEEE Transactions on Image Processing*, 10 (1), 140–47.

Ennis, John M. and Virginie Jesionka (2011), "The Power of Sensory Discrimination Methods Revisited," *Journal of Sensory Studies*, 26 (5), 371–82.

Heekeren, Hauke R., Sean Marrett, and Leslie G. Ungerleider (2008), "The Neural Systems That Mediate Human Perceptual Decision Making," *Nature Reviews Neuroscience*, 9 (6), 467–79.

Jacoby, Jacob and Maureen Morrin (1998), "'Not Manufactured or Authorized By...': Recent Federal Cases Involving Disclaimers," *Journal of Public Policy & Marketing*, 17 (Spring), 97–107.

Kapferer, Jean-Noël (1995), "Brand Confusion: An Empirical Study of a Legal Concept," *Psychology and Marketing*, 12 (6), 551–68.

Kubo, Masaaki, Zaher Aghbari, KunSeok Oh, and Akifumi Makinouchi (2003), "Image Retrieval by Edge Features Using Higher Order Autocorrelation in a SOM Environment," *IEICE Transactions on Information and Systems*, E86-D (8), 1406–1415.

Leite, Fabio P. and Roger Ratcliff (2010), "Modeling Reaction Time and Accuracy of Multiple-Alternative Decisions," *Attention, Perception & Psychophysics*, 72 (1), 246–73.

Miaoulis, George and Nancy D'Amato (1978), "Consumer Confusion and Trademark Infringement," *Journal of Marketing*, 42 (April), 48–55.

Miceli, Nino and Rik Pieters (2010), "Looking More or Less Alike: Determinants of Perceived Visual Similarity Between Copycat and Leading Brands," *Journal of Business Research*, 63 (11), 1121–28.

Morrin, Maureen and Jacob Jacoby (2000), "Trademark Dilution: Empirical Measures for an Elusive Concept," *Journal of Public Policy & Marketing*, 19 (Fall), 265–76.

——, Jonathan Lee, and Greg M. Allenby (2006), "Determinants of Trademark Dilution," *Journal of Consumer Research*, 33 (2), 248–57.

Oliva, Aude and Antonio Torralba (2006), "Building the Gist of a Scene: The Role of Global Image Features in Recognition," *Progress in Brain Research*, 155, 23–39.

Pieters, Rik and Michel Wedel (2011), "Ad Gist: Ad Communication in a Single Eye-Fixation," *Marketing Science*, 31 (1), 59–73.

Poulter, Sean (2009), "Shoppers 'Conned' by Raft of Cheap Copycat Versions of Popular Brands," *The Daily Mail*, (May 11), (accessed October 22, 2013), [available at http://www.dailymail.co.uk/news/article-1180258/Shoppers-conned-raft-cheap-copycat-versions-popular-brands.html].

Pulling, Chris, Carolyn J. Simmons, and Richard G. Netemeyer (2006), "Brand Dilution: When Do New Brands Hurt Existing Brands?" *Journal of Marketing*, 70 (April), 52–66.

Shapiro, Linda G. and George C. Stockman (2001), *Computer Vision*. New York: Prentice Hall.

Smith, Philip L. and Roger Ratcliff (2004), "Psychology and Neurobiology of Simple Decisions," *Trends in Neuroscience*, 27 (3), 161–68.

Usher, Marius, Zeev Olami, and James L. McClelland (2002), "Hick's Law in a Stochastic Race Model with Speed–Accuracy Tradeoff," *Journal of Mathematical Psychology*, 46 (6), 704–715.

Van der Lans, Ralf, Rik Pieters, and Michel Wedel (2008), "Competitive Brand Salience," *Marketing Science*, 27 (5), 922–31.

Van Horen, Femke and Rik Pieters (2012), "When High-Similarity Copycats Lose and Moderate-Similarity Copycats Gain: The Impact of Comparative Evaluation," *Journal of Marketing Research*, 49 (February), 83–91.

Van Vugt, Marieke K., Robert Sekuler, Hugh R. Wilson, and Michael J. Kahana (2013), "An Electrophysiological Signature of Summed Similarity in Visual Working Memory," *Journal of Experimental Psychology: General*, 142 (2), 412–25.

Vandekerckhove, Joachim, Francis Tuerlinckx, and Michael D. Lee (2011), "Hierarchical Diffusion Models for Two-Choice Response Time," *Psychological Methods*, 16 (1), 44–62.

Warlop, Luk and Joseph W. Alba (2004), "Sincere Flattery: Trade-Dress Imitation and Consumer Choice," *Journal of Consumer Psychology*, 14 (1/2), 21–27.

Wedel, Michel and Rik Pieters (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19 (4), 297–312.

Zaichkowsky, Judy L. (2006), *The Psychology Behind Trademark Infringement and Counterfeiting*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Appendix A

IMAGES FOR LEADING BRAND, EXPERIMENTAL COPYCAT, AND MARKET COPYCAT PACKAGES WITH GLOBAL COLOR AND TEXTURE SIMILARITIES TO THE LEADING BRAND

1: Adhesive Bandages

CE: .901 CM: .309
TE: .930 TM: .674

2: Allergy Medicine

CE: .684 CM: .902
TE: .858 TM: .859

3: Canned Noodles

CE: .920 CM: .653
TE: .922 TM: .772

4: Chewy Bars

CE: .751 CM: .600
TE: .871 TM: .788

5: Corn Flakes

CE: .669 CM: .621
TE: .856 TM: .745

6: Crackers

CE: .883 CM: .673
TE: .788 TM: .793

7: Dishwashing Liquids

CE: .787 CM: .708
TE: .816 TM: .860

8: Hand Sanitizer

CE: .935 CM: .619
TE: .946 TM: .804

9: Honey

CE: .976 CM: .753
TE: .962 TM: .855

10: Ketchup

CE: .949 CM: .713
TE: .945 TM: .790

11: Mayonnaise

CE: .970 CM: .702
TE: .938 TM: .808

12: Mouthwash

CE: .948 CM: .727
TE: .921 TM: .723

13: Raisin Bran

CE: .786 CM: .710
TE: .891 TM: .877

14: Sports Drinks

CE: .957 CM: .618
TE: .946 TM: .779

15: Tissues

CE: .806 CM: .487
TE: .854 TM: .763

Notes: The leading brand, experimental copycat, and market copycat products are presented in order from left to right. CE = global color similarity of the experimental copycat. CM = global color similarity of the market copycat. TE = global texture similarity of the experimental copycat. TM = global texture similarity of the market copycat.

Appendix B

ACCURACY RATE AND RESPONSE TIMES FOR CORRECT AND INCORRECT RESPONSES

			Response Time (milliseconds)	
	Product Category	*Accuracy*	*Correct*	*Error*
15	Tissues	.868	855	1,078
1	Adhesive bandages	.637	1,070	1,353
5	Corn flakes	.577	1,168	1,440
4	Chewy bars	.571	967	1,134
12	Mouthwash	.560	1,066	1,336
8	Hand sanitizer	.533	935	1,251
11	Mayonnaise	.516	1,058	1,291
14	Sports drinks	.516	992	1,321
2	Allergy medicine	.467	1,208	1,350
3	Canned noodles	.462	1,141	1,207
6	Crackers	.440	1,344	1,321
13	Raisin bran	.440	957	1,231
7	Dishwashing liquids	.418	982	1,128
9	Honey	.407	948	1,499
10	Ketchup	.385	1,232	1,156

Notes: Categories are sorted by accuracy rate.

APPENDIX C: MCMC ESTIMATION ALGORITHM

We set prior distributions of parameters as follows:

$$\left(\beta_i', \gamma_i', \theta_i\right)' = \Omega_i \sim N\left(\Delta' Z_i, V_\Omega\right),$$

$$\mathrm{vec}(\Delta) \middle| V_\Omega \sim N\left(0, V_\Omega \otimes A^{-1}\right), \text{ and}$$

$$V_\Omega \sim IW\left(\nu_0, S_0\right).$$

The subscript i represents a participant, $Z_i = 1$, and Δ, A, V_Ω, S_0, and ν_0 are hyper parameters:

$$A = \frac{1}{n}, \quad \nu_0 = d + 4, \quad S_0 = \nu_0 I_d,$$

where n is the number of participants and d is the dimension of Ω_i. The MCMC estimation comprises the following steps:

Step 1. Draw Ω_i by a Metropolis–Hastings algorithm. Let Ω_i^{old} be a current Ω_i.

a. Generate a proposal Ω_i^{new} through the random walk method:

$$\Omega_i^{new} = \Omega_i^{old} + \xi_i, \quad \xi_i \sim N\left[0, s^2\left(H_i + \Sigma^{-1}\right)^{-1}\right],$$

where s is the scaling constant, H_i is the Hessian of the ith participant's unit likelihood evaluated at the maximum likelihood estimation, and Σ is a positive-semidefinite matrix.

b. Accept Ω_i^{new} with probability $\min\left[1, \frac{L_i\left(\Omega_i^{new}\right)p\left(\Omega_i^{new}\right)}{L_i\left(\Omega_i^{old}\right)p\left(\Omega_i^{old}\right)}\right]$.

Step 2. Draw Δ and V_Ω by Gibbs sampling.

a. Draw Δ:

$$\mathrm{vec}(\Delta) \middle| V_\Omega \sim N\left[\mathrm{vec}\left(\tilde{\Delta}\right), D\right],$$

$$\text{where } \tilde{\Delta} = \left(Z'Z + A\right)^{-1}\left(Z'Z\hat{\Delta} + 0\right),$$

$$\hat{\Delta} = \left(Z'Z\right)^{-1} Z'\Omega, \text{ and}$$

$$D = V_\Omega \otimes \left(Z'Z + A\right)^{-1},$$

b. Draw V_Ω:

$$V_\Omega \sim IW\left(\nu_0 + n, S_0 + S\right),$$

$$\text{where } S = \sum_i \left(\Omega_i - \Delta' Z_i\right)\left(\Omega_i - \Delta' Z_i\right)'.$$

Copyright of Journal of Marketing Research (JMR) is the property of American Marketing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.