Dressed for the Occasion: Font-Product Congruity in the Perception of Logotype

John R. Doyle and Paul A. Bottomley
Cardiff University

Styles of lettering (i.e., fonts) can differ in their appropriateness for describing certain types of brands and products. We use the Osgood dimensions of evaluation, potency, and activity to measure the connotative meaning of fonts and product categories. Judgments of the appropriateness of a font for a product partly depend on the consistency between the connotative meaning of the font and that of the product along dimensions of potency and activity. However, its perceived appropriateness also depends on the connotative meaning of the font per se, independent of the type of product being described. These differences are also evident when participants choose a company to call on the basis of ads similar to those found in the Yellow Pages. Finally, the effect of the meaning that people assign to a font is similar to the effect of the meaning they assign to the product's name itself.

In a recent review of consumer research over a 30-year period, Simonson, Carmon, Dhar, Drolet, and Nowlis (2001) noted a steadily increasing interest in cognitive processes and in subjective as opposed to rational bases of judgments and decisions. Consistent with this trend, marketing–consumer researchers have recently begun to consider the role of typography (principally typeface) in advertising and consumer contexts—a topic that had remained at the periphery of psychology for over 80 years (Childers & Jass, 2002; Doyle & Bottomley, 2004; Henderson, Giese, & Cote, 2004; McCarthy & Mothersbaugh, 2002; Pan & Schmitt, 1996). McCarthy and Mothersbaugh presented a broad framework to classify typographic factors in advertising. They identified three "typographic dimensions" (typeface, spacing, and layout) that are crossed with three "typographic outcomes" (semantic, legibility, and appearance). The research presented here belongs in the typeface–semantic cell of their grid. Specifically, we are interested in the meaning that typefaces convey and the congruity of its meaning with that of the product to which they refer.

Styles of lettering are not themselves neutral. They convey different messages of their own (Bartram, 1982; Brumberger, 2003; Morrison, 1986; Rowe, 1982; Wendt, 1986) through what Childers and Jass (2002) termed "typeface semantics." For instance, consumers might associate ornate fonts (and, by inference, brands displayed in these fonts) with elegance and style. All writing has to be written in a certain style of lettering. Consequently, brand names (logos), ad copy, and packaging may all convey covert messages through the choice of typeface they adopt.

This article examines the factors that make a font appropriate for describing a product or brand in a logotype, that is, a name that is set in a distinctive typeface but is devoid of secondary embellishments (Carter, 1999). Study 1 provides an understanding of the meaning present in fonts and also in products and determines how these factors combine to influence perceptions of the appropriateness of fonts for different products. Study 2 examines the effects of font-product congruity on brand choice. In this study, we also include brand name as a variable, enabling us to examine the consequences of not only font-product congruity but also of name-product congruity and font-name congruity.

THEORETICAL BACKGROUND

Conceptual Framework

Figure 1 is a diagrammatic representation of how research conducted in the area, including this article, fits together. Its role is to help visualize the structure of past research as we present it. (Omission of a link simply indicates that research has yet to be undertaken; it does not rule out such a relation.) Previous work on the influence of fonts can be conceptualized with reference to its implications for the variables described in this figure and the relations between them.
Effect of typeface characteristics on typeface semantics (Link 1). Tannenbaum, Jacobson, and Norris (1964) linked the standard physical characteristics of Serif-Sans Serif, Roman–italic, and upper–lower case to the typeface semantics measured using Osgood, Suci, and Tannenbaum’s (1957) three dimensions of connotative meaning: evaluation (good, pleasant, beautiful, happy), potency (strong, hard, rugged, potent, tough), and activity (active, fast, young, lively). They found, for instance, that italicized fonts were perceived as more active but less potent than regular Roman fonts. Furthermore, fonts presented in upper-case lettering were more potent than in lower case. However, serif and sans-serif typefaces did not differ in terms of evaluation, potency, or activity (EPA).

Henderson et al. (2004) provided a broader analysis of the relation between the physical (visual) characteristics and typeface semantics as well as the nature of these factors themselves (see Boxes A and B in Figure 1). Beginning with a set of adjectives useful for conveying a firm’s corporate identity (e.g., warm, interesting, honest, etc.), they derived four typeface semantic factors: pleasing versus displeasing, engaging versus boring, reassuring versus unsettling, and prominent versus subtle. Furthermore, they used two separate factor analyses to derive a typeface-specific set of visual factors (flourish, weight, and compressed) and a universal set of visual factors (elaborateness, harmony, and naturalness) that are applicable to symbols in general, whether alphabetic or not. Indeed, the universal factors are similar to those obtained by Henderson and Cote (1998) using non-alphabetic logos. To understand how the physical (visual) characteristics (Box A) influenced the semantic impressions created by the font (Box B), a series of regression analyses were conducted. We defer a discussion of these somewhat complex relations until the concluding section, where we also integrate our findings directly with theirs.

Henderson et al.’s (2004) four typeface semantic factors are optimized to capture the meanings that fonts convey. For instance, it was graphic designers who generated the starting set of adjectives that are considered to be useful for conveying a firm’s corporate identity. So it is not known how succinctly or how comprehensively their typeface dimensions would also capture the meanings that products or names convey. We must assume that graphic designers, similar to any other specialized group, have their own idiosyncratic view of the world. This uncertainty in the taxonomy’s scope is an important consideration for the work presented in this article, because proven cross-domain portability is precisely what we need in a classification system to determine the degree to which different fonts, products, and names are congruent with each other. For this reason, we will instead use Osgood’s three dimensions of connotative meaning (to be discussed shortly) that have a proven track record of portability.

It is important to note here that our work is particularly concerned with the role of font in the context provided by the product. This point differentiates our work from that of Henderson et al. (2004), whose conclusions regarding what makes an effective typeface do not take into account the product for which the typeface is used. It similarly differentiates us from Tannenbaum et al. (1964), whose work on the antecedents of Osgood’s dimensions in typeface was also devoid of context.

Effects of typeface semantics on brand preferences (Link 2). According to McCarthy and Mothersbaugh (2002), the semantic associations of a typeface are the connotations that consumers derive from the text that go beyond the
Effects of appropriateness on brand choice (Link 3). Font-product appropriateness has consequences of its own. In a field test, Doyle and Bottomley (2004) found that controlling for the influence of name and other presentational cues, people chose chocolate truffles more frequently from a box of chocolates that had an appropriate font for chocolates, 

\textit{Caslon}, than from a box with an inappropriate font, \textit{Salem}. Therefore, the logotype that a brand adopts may exert a continuous force for better (if well chosen) or for worse (if not).

Effects of typeface semantics on appropriateness and brand choice (Links 4 and 5). Despite the wealth of empirical studies on typeface semantics and appropriateness, the links between boxes have been less well researched. In particular, the link between typeface semantics and appropriateness has not yet been investigated. Although Doyle and Bottomley (2004) linked appropriateness to choice, appropriateness was effectively treated as a black box.

The appropriateness of typefaces is known to depend on the type of product involved, but an understanding of why this dependence exists is unclear. Our two studies help to provide this understanding. Study 1 examines the antecedents of font-product appropriateness (Link 4 in Figure 1). In the context of a more comprehensive model of product-font-name effects, Study 2 shows the same drivers account for people's choice of brands (Link 5), and a similar set of drivers apply to name-related effects.

Sources of Font-Product Congruity

We suggest that a font may be perceived as appropriate for an occasion (e.g., product, brand) when there is congruity between the meaning (or associations) conveyed by the product and the meaning (or associations) conveyed by the font. We identify three potential sources of shared meaning. First, there may be learned arbitrary associations. For instance, certain products are associated with certain countries, and so, by historical chance and conditioning, are certain fonts. Prominent examples are the \textit{Garamond} family of fonts associated with Ireland, and the \textit{Georg} family with Germany. Thus, there is a fit between these fonts and traditionally Irish or German products (e.g., linen and beer, respectively) based on mutual country-of-origin associations. Second, \textit{figurative} associations occur when a particular feature of a font depicts something in the real world. The typeface \textit{Snowdrift}, for example, which has snow-capped letters, would naturally suggest applications for frozen goods. Both of these sources of congruity are analyzed in Study 2 and labeled as \textit{obvious associations}. Based on our empirical evidence, however, it appears that there are relatively few obvious associations operating between font and product.

Third, fonts may evoke more \textit{abstract} meanings that, if shared with a product, could form the basis of congruity. These semantic qualities are processed automatically from the printed word. Lewis and Walker (1989) conceived of typeface as generating a second route to meaning, in addition to the word itself. They used the typeface \textit{Palatino Italic}, which was shown in a pretest to have connotations of fast, and \textit{Cooper Black}, which connoted slow. People responded faster in the experimental task (pressing one of two keys to a simple semantic judgment) when the connotative meaning of the word was congruent with that of the typeface (e.g., “cheetah” in Palatino Italic, “tortoise” in Cooper Black) than when animal and font were incongruent. Our research, similar to that of Childers and Jass (2002) and Henderson et al. (2004), focused on this third source of meaning.

Osgood's Componential Analysis of Connotative Meaning

A method of analyzing connotative meaning, which is known to be robustly portable was developed by Osgood et al. (1957). They factor analyzed connotative meaning over a large number of studies and found three orthogonal, bipolar dimensions: EPA. Although Osgood's three main dimensions of connotative meaning were originally derived using words, they have subsequently been used to analyze connotative meaning in a variety of different contexts, such as simple pictures (McMurray, 1958), color (Valdez & Mehrabian, 1994), and nonverbal communication (Mehrabian, 1972).
and Context-Sensitive Effects

When people encounter a new brand, they necessarily rely way a brand can do this is through the lettering it adopts. In

Recent neuropsychological research has demonstrated that connotative meaning builds up very quickly in the brain. Skrandies and Chiu (2003) presented words briefly on a screen and recorded changes in electrical potential on participants’ scalps. They contended that, “there is a significant effect of semantic meaning on brain activity occurring as early as 80 ms latency” (p. 48). Word classes were drawn from each end of the bipolar EPA dimensions: E+, E−, P+, P−, A+, and A−. Each of these six regions of connotative space exhibited its own distinctive electrical signatures, which were consistent whether the stimuli were Chinese characters or German words (Skrandies, 1998; Skrandies & Chiu, 2003).

The three EPA dimensions are not perfectly orthogonal. When the context is human character traits, for example, potency and activity can combine to create a super-ordinate personality dimension labeled surgency or dynamism (Goldberg, 1981). Cross-cultural variations can also exist. Schneider (1999; personal communication, October 21, 2002) had males rate various role identities on EPA. Respondents from the United States tended to evaluate potent role identities as good, as evidenced by a positive correlation between evaluation and potency (r = .40). However, respondents from Germany did not (r = -.15). Schneider (1999) interpreted this as evidence that German males had a much less accepting attitude to authority figures than males in the United States. These variations in the relation between EPA are all domain specific. Rather than suggesting the need for a multitude of EPA-like microtheories, it is more useful to see them against the stable background of a wider conception of EPA. In conclusion, access to the connotative meaning of a stimulus is fast, automatic, and specific in its action.

In summary, Osgood’s framework of connotative meaning can be applied to words, objects, pictures, and typefaces. It is therefore a lingua franca into which the connotative meanings in both typeface and product, and any future class of materials, may be usefully translated. The positions of product and typeface in the multidimensional space defined by EPA may then be manipulated relative to each other. Osgood’s EPA can be applied to products and fonts and names. For this reason, we chose to use it rather than Henderson et al.’s (2004) more recent, typeface-specific connotative framework, which was not developed for, and has not been applied to, other classes of materials.

Typeface Semantics: Context-Free and Context-Sensitive Effects

When people encounter a new brand, they necessarily rely heavily on what the brand is trying to signal about itself. One way a brand can do this is through the lettering it adopts. In Study 1, participants were presented with a scenario in which there are no particular brands to be considered. Instead, they were asked to rate the appropriateness of different fonts for a product category. This task invites people to address the following question: If I were purchasing in that product category, are the signals I get from the font consistent with what I would want from my purchase? If so, the font is considered appropriate.

In conceptualizing the processes that influence whether someone perceives a font to be appropriate for a product category, we distinguish between (a) a simple, context-free influence of font and (b) a more complex, context-sensitive influence that results from a comparison of the font’s meaning with that of the product. In the first regard, a font that has connotations of good, pleasant, and beautiful might signal that the product dressed in such a font will also deliver these qualities. We assume that people have a predisposition to seek good, pleasant, and beautiful qualities in products, rather than their opposite. Thus, they will judge fonts that signal that a product will deliver these qualities to be appropriate. Because “good, pleasant, and beautiful” is a description of the positive pole of Osgood’s evaluation dimension (it is E+), we hypothesize:

H1: The appropriateness of a font will be greater if the evaluative meaning it conveys is positive than if it is negative, and this will be true independent of its consistency with the product’s meaning along this dimension.

The influence of potency and activity is less clear-cut. Nonetheless, it is possible that people will have a bias to one or the other pole of potency, and also of activity in the qualities they seek in products. To this extent, fonts that signal these qualities would be considered more appropriate, independent of its consistency with the product’s meaning along these dimensions. This is an empirical question, and no formal hypotheses are offered.

Potency and activity are more likely to enter into context-sensitive processing. Some products, for example, may be desirable if their qualities are high in potency. For example, a hammer may be preferred if it is “strong, hard, and rugged.” Other products are more desirable when they are low in potency; a pillow may be preferred if it is “soft, delicate, and tender.” Similar considerations arise with respect to activity. In these circumstances, the qualities signaled by a particular font may be considered more appropriate if the potency and activity it conveys are congruent with the qualities that are considered desirable in the product being described. This suggests the following hypotheses:

H2: A font’s appropriateness for describing a product will be greater if its meaning along the dimension of potency is congruent with that of the product than if it is incongruent.
H3: A font’s appropriateness for describing a product will be greater if its meaning along the dimension of activity is congruent with that of the product than if it is incongruent.

We examined the impact of these typeface semantic effects on both perceptions of the appropriateness (in Study 1) and brand choice (in Study 2). In considering these effects, a question arises as to whether the typeface meaning has a direct impact on brand preferences or whether this impact is mediated by its influence on perceptions of appropriateness. These alternative possibilities could not be examined in the research we conducted. However, the possibilities are discussed more fully after the results of our studies are reported.

STUDY 1

Pretest
To select stimuli for use in the studies, we produced a database of EPA norms for both fonts and products. The databases were then used to select stimuli for Studies 1 and 2. To produce the norms for fonts, 142 participants rated 132 fonts on Osgood’s three dimensions of connotative meaning. Fonts were displayed using a selection of upper- and lower-case letters in alphabetical order in 24 point. Each font was rated using three 11-point semantic differential scales with endpoints –5 and +5. Each person rated one third of the fonts on all three dimensions. Following a similar design to the pretest of fonts, 69 participants rated the connotative meaning of 120 products. For fonts, the correlation was -.28 between evaluation and potency, .01 between evaluation and activity, and -.03 between potency and activity. For products, the correlations were -.38, .51, and .13, respectively.

To elicit EPA ratings for use in our experiments, we defined the endpoints of each scale by a cluster of adjectives that had loaded heavily on their respective dimensions in previous studies (cf. Heise, 1978; Osgood et al., 1957). Specifically, for evaluation the scale ranged from +5 (good, pleasant, beautiful, happy) to –5 (bad, unpleasant, ugly, unhappy). For potency the scale ranged from +5 (strong, hard, rugged, potent, tough) to –5 (soft, delicate, tender, weak, gentle). For activity the scale ranged from +5 (active, lively, young, fast) to –5 (passive, still, old, slow). Thus we used three single-item scales, although each of their endpoints was defined by a cluster of adjectives. This procedure has obvious economies over using separate scales for each bipolar pair of adjectives (“good” vs. “bad,” “pleasant” vs. “unpleasant,” etc.). Osgood himself seemed eventually to have favored the cluster approach (Heise, 1978, p. 64).

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>E</th>
<th>P</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>E+</td>
<td>Bath towels, sofas, sunglasses, video rental, wine glasses, yoghurt</td>
<td>2.77</td>
<td>-0.38</td>
</tr>
<tr>
<td>E−</td>
<td>Boxing gloves, cigarettes, herbal teas, knives, road haulage, sewing machines</td>
<td>-1.30</td>
<td>0.98</td>
</tr>
<tr>
<td>P+</td>
<td>Burglar alarms, butcher, crash helmets, electrician, hammer, whisky</td>
<td>0.15</td>
<td>3.30</td>
</tr>
<tr>
<td>P−</td>
<td>Fabric softener, florists, hairdresser, perfume, soft furnishings, specialty jams</td>
<td>2.31</td>
<td>-2.94</td>
</tr>
<tr>
<td>A+</td>
<td>Champagne, disco, health and fitness centers, record store, sports watches, trampolines</td>
<td>2.43</td>
<td>0.67</td>
</tr>
<tr>
<td>A−</td>
<td>Bookshop, carpet, insulation, laundry detergent, life insurance, storage services</td>
<td>0.65</td>
<td>-0.68</td>
</tr>
</tbody>
</table>

| E+      | Belez Thick Cond., Bodoni h, Cooper Black, Garamond It., Palatino It., Atthis Normal | 1.85  | -0.03 | -0.15 |
| E−      | big stripes, Chisel Stripe, ambition, just chroniic, traffic | -2.53 | 0.44  | 0.13  |
| P+      | Broadway, Circus Bold Cond., CUT ROCK, Salem, Stencil | 0.32  | 3.41  | -0.45 |
| P−      | Salem Light, Cond., Manic, Garamond, ISADORA CAPS, Kidprint | 0.03  | -3.79 | -0.14 |
| A+      | BERTRAN BBLADES, Greenlower, HANDY, Snap LrE, SALTERIT | 1.10  | -0.79 | **4.05** |
| A−      | Basque, Batik Regular Cond., Diagonal, Garamond, Metrnostyle Extended | 0.27  | 0.29  | -2.34 |

*Note.* Numbers are mean ratings for each condition, on a scale ranging from –5 to +5. A bolded number is significantly different from the number directly below it. Off-diagonal differences are potential confounds of the on-diagonal experimental contrasts.
with 2 representing the positive pole of the dimension and 2 representing the negative pole. The 4 products in each set were then paired with all 12 fonts selected to represent differences in meaning along this dimension, resulting in three sets of 48 font-product stimuli. This procedure was then repeated, randomly dividing the 12 products into another 3 sets of 4. A similar procedure was employed for each of the other dimensions. Thus, a total of 18 different sets of 48 font-product pairings were constructed, with 6 pertaining to each dimension of meaning.

Procedure. Participants consisted of 231 undergraduate students. Of these, 68 received stimuli varying in potency, 70 received stimuli varying in activity, and 93 received stimuli varying in evaluation. (An approximately equal number of participants received materials pertaining to each subset of 48 stimuli representing the dimension.) To introduce the study, participants were told, “Companies spend a lot of time, energy and effort in presenting what they hope is the right image to consumers,” and that the purpose of this study was to “…explore what people feel about different fonts for different types of products.” Given this orientation, they were asked to rate each of the 48 stimuli along an 11-point scale ranging from 0 (absolutely inappropriate) to 10 (absolutely appropriate).

A response sheet was constructed as a 12 row x 4 column grid. The row labels were the 12 fonts displayed by using the word “sample”. The column headings were the 4 products, printed in CG Omega. The word “sample” was chosen in preference to the conventional test string “The quick brown fox …” or the letters of the alphabet. Although the former has strong connotations of its own, the latter has too many letters, requiring a small point size, therefore minimizing perceptual differences between fonts. As it is clear that the word “sample” is not a brand name, it was obvious to participants they were rating fonts for product categories and not fonts for a particular brand.

Results

To reiterate, each participant rated 48 font-product pairs pertaining to a given meaning dimension, 12 of which represented each combination of font polarity and product polarity. Ratings representing each combination of these variables were averaged and entered as repeated-measures in a by-subject analysis of variance, polarity of font being crossed with polarity of product. Mean ratings of appropriateness along each meaning dimension are summarized in Table 2 as a function of font polarity and product polarity.

H1 contends that the appropriateness of a font will increase with its evaluative meaning independent of the product’s meaning along this dimension. This prediction was confirmed. Fonts were judged more appropriate if their evaluative meaning was positive than if it was negative (M = 5.30 vs. 3.52), F(1, 92) = 80.07, p < .001. In addition, however, appropriateness was generally perceived to be greater when the font was high in potency than if it was low (M = 4.75 vs. 4.41), F(1, 67) = 7.42, p < .01, and to be greater when the font was low in activity than when it was high (M = 5.27 vs. 4.72), F(1, 69) = 19.59, p < .001. Thus, fonts were not only considered more appropriate if they connoted pleasantness and beauty, but also if they connoted strength and ruggedness or, alternatively, stillness and passivity.

According to H2 and H3, a font’s appropriateness for describing a product should be greater if its meaning along the dimensions of potency and activity is congruent with the product’s meaning along these dimensions. These hypotheses were also confirmed. Perceptions of a font’s appropriateness were greater when its potency was congruent with that of the product than when it was incongruent (5.63 vs. 3.54, respectively), as evidenced by a significant interaction of Font x Product Polarity, F(1, 67) = 122.55, p < .001. Similarly, perceptions of appropriateness were greater when font and product were congruent in activity than when they were incongruent (5.63 vs. 4.36, respectively), F(1, 69) = 89.57, p < .001. However, the font’s congruity in evaluative meaning with that of the product had no impact on perceptions of their appropriateness (4.58 vs. 4.24), F(1, 92) = 2.32, p > .10.

Finally, appropriateness was generally perceived to be greater when (a) the product was high in evaluation than when it was low (M = 4.55 vs. 4.27), F(1, 92) = 6.02, p < .05; (b) the product was low in potency than when it was high (M = 4.72 vs. 4.44), F(1, 67) < 3.90, p > .05; and (c) the product was high in activity than when it was low (M = 5.18 vs. 4.81), F(1, 69) = 5.34, p < .05. In other words, people who contemplated E+, P−, A+ products tended to see them as appropriately dressed, whatever the font. The reason for these effects is unclear.

<table>
<thead>
<tr>
<th>Font</th>
<th>P−</th>
<th>P+</th>
<th>Product</th>
<th>A−</th>
<th>A+</th>
</tr>
</thead>
<tbody>
<tr>
<td>E−</td>
<td>3.54</td>
<td>4.99</td>
<td>P−</td>
<td>5.59</td>
<td>3.85</td>
</tr>
<tr>
<td>E+</td>
<td>3.49</td>
<td>5.61</td>
<td>P+</td>
<td>3.22</td>
<td>5.66</td>
</tr>
</tbody>
</table>

Note. P+, P−, A+, A−, E+, and E− relate to the connotative meaning of fonts and products measured along Osgood’s three dimensions of EPA.
Three additional analyses were conducted in which data were pooled over subjects and items were used as the unit of analysis, thus evaluating the generalizability of results over different randomly chosen sets of stimulus materials (Clark, 1973). These analyses also permitted us to evaluate effects of each dimension of meaning treating values along other dimensions as covariates. These analyses confirmed conclusions drawn from the first set of analyses. That is, analyses of evaluation yielded a significant effect of font, $F(1, 136) = 85.34, p < .001$, whereas analyses of potency and activity each yielded both a main effect of font—in each case, $F(1, 136) > 4.36, p < .05$—and a Font × Product interaction—in each case, $F(1, 136) > 15.63, p < .001$. However, no effects of product meaning per se were reliable, suggesting that the unexpected effects of this variable in the first set of analyses may have been spurious.

In summary, appropriateness was a positive function of the value of the font along the evaluation and potency dimensions and a negative function along the activity dimension. In addition, it was a positive function of font-product congruity along dimensions of activity and potency. The first effects suggest a simple, context-free means of processing the logo-type, whereas the second effects suggest a more complex, context-sensitive means of processing.

**STUDY 2**

Study 2 replicated the findings of the first experiment under conditions more similar to those that people would encounter outside the laboratory. We constructed a situation similar to those in which consumers use the Yellow Pages to decide on a firm or service to call when there is limited information about the choice alternatives. The study differed from Study 1 in other respects as well. First, we investigated the hypotheses tested in Study 1 at the brand rather than product category level. Second, we used an entirely different set of subjects, fonts, and products. Third, participants made choices among competing brands rather than rating appropriateness. Fourth, greater realism was introduced by including distracters such as name, address, phone number, and position on the page—any of which could serve as a competing reason to choose one brand over another.

Finally, Study 2 also extended the scope of the font effects to an analogous set of name effects. Each advertisement in the Yellow Pages scenario contained a family brand name (e.g., Smith or Jones). These names can each be thought of as lying at a point in Osgood's EPA space, in a similar manner to the fonts and products in Study 1. Because name is just another way of signaling what a brand might deliver, we expected that the effect of name and name-product congruity would parallel the effects of font and font-product congruity.

In this regard, Mehrabian and de Wetter (1987) found that preferences for a product with a given name were a function of the difference between (a) the desired emotional impact of a product along dimensions of pleasantness, arousal, and dominance and (b) the actual impact of the name-product configuration along these dimensions. But Mehrabian and de Wetter did not allow for the influence of the main effects of name, independent of product. We examined both context-free (main) and context-sensitive (congruity) impact of names on brand choice.

The effects of name-font congruity were also considered. Brands in a product category can be thought of as conveying their different positions in the market through both their names and the fonts that dress these names. However, the congruity between font and name per se signals nothing about whether a specific brand will have desirable qualities. To this extent, there would be no reason to expect an advantage to accrue as a result of this congruity. On the other hand, it is possible that congruity between font and name elicits positive affect. If this is so, and if the affect is transferred to the brand, a positive relation between font-name congruity and brand choice might in fact occur.

Study 2 examined the source of font-product congruity due to learned or figurative associations (together known as obvious associations), as distinct from semantic associations. Doyle and Bottomley (2004) determined that perceptions of the appropriateness of a font influenced brand choice, but did not investigate the source of appropriateness that might have been responsible. Typical font-product combinations judged appropriate were Snowdrift with ice cream, and Kendall Script with fountain pens, suggesting that obvious associations rather than congruity of connotative meanings, as investigated in Study 1, might be the basis of these perceptions. Both types of meaning are measured in Study 2.

One limitation of the study is worth mention. As Figure 1 indicates, the effect of typeface on product choices could either be direct or alternatively, mediated by its impact on perceptions of appropriateness. To evaluate this possibility, however, appropriateness estimates would have to be obtained from the same participant who made choices. Because of the heavy demands already imposed on participants by the study design, we were unable to examine this possibility.

**Pretesting**

*Obvious associations.* To model the source of congruity due to learned or figurative associations, 20 participants (similar to participants in the main study) viewed a 14 × 18 grid of fonts × products and identified those pairings where they perceived, "some obvious association between typeface and product ... that the majority of people would also see." These were the 14 fonts and 18 products to be used as materials in Study 2. A total of 23% of combinations were identified as obvious associations by at least 1 participant. However, no pairing met the stringent criterion that it should be seen by the majority of people. The maximum number of
people who agreed was just 5 of 20 (the typeface \textit{Cinema} with power tools). In fact, people’s obvious associations were highly idiosyncratic, because the majority of these combinations were not seen by anyone else!

Method

\textbf{Selection of stimuli.} Fourteen fonts and 18 products were chosen from their respective databases of EPA norms such that they had not been used in Study 1, they were drawn from all regions of the multidimensional space define by EPA, and all explanatory variables (see next) were as independent from each other as possible.

A total of 84 names was randomly selected from a larger database of 128 names that had been rated by 53 people to provide EPA norms. The original database was compiled from the telephone directories for three regions. The compilation of this database parallels those of fonts and products, with multiple pairs of bipolar adjectives characterizing each EPA dimension as descriptors, but in the interests of economy only a single rating being elicited.

\textbf{Construction of stimulus materials.} A total of 42 eighteen-page booklets was prepared. Each page of the booklet contained 14 ads pertaining to products in a particular category (see Table 3). The product (e.g., wallpaper, used cars) was written at the top of the page, as it might in the Yellow Pages or similar compendium. Each page had the same background of three columns of gray with a ragged right edge suggesting one-line directory entries. The foreground was 14 box ads, each of which contained a brand name and a unique telephone number and address. Each brand name on a given page was conveyed in a different 1 of 14 fonts (see Table 3).

The materials in the design were extensively randomized to ensure that there was no confounding between font, brand name, position on the page (box ads locations), and product. For instance, in the totality of the design, all fonts were associated with all 14 positions on the page, several brand names, and all 18 products.

The names of the brands on each page were presented in a different 1 of 14 fonts, in a position that varied over booklets so that each font appeared equally often in each position. (The other information, unique to each ad, was an address and telephone number in the typeface CG Omega in all cases.) Within a booklet the 84 brand names were each used three times, once in each set of six pages. Extensive counterbalancing ensured that, pooled over booklets, font, name, product, and position of the ad on the page were not confounded with each other.

\textbf{Procedure.} The main experiment and pretests of obvious associations were all conducted on distinct sets of participants. A total of 42 participants were run in small groups of between 1 and 6 with each participant receiving their own booklet. Instructions were as follows: “Imagine you are looking through the Yellow Pages, or similar compendium of ads, such as a specialized trade magazine … . You wish to purchase a product or service, or at least to find out more about it to inform a future purchase decision.” Participants were asked, “ … to draw up a short-list of companies, and identify four brands per product category, you would telephone to ‘get a quote’ or obtain more information from.” These firms make up their consideration sets. Each person worked through their booklet of 18 products, placing four circles per page to indicate the firms–brands to be contacted.

Results

The dependent variable was whether a particular brand was short-listed for a particular product and by a particular person. As such, there were 10,584 observations (18 products x 14 brands per product x 42 subjects), of which 3,024 were chosen (18 x 4 x 42). There were 16 independent variables. Obvious associations were constructed by counting the number of people in the pretest who had identified each font-product combination as having an obvious association. There were EPA measures for each of font and name, but not product, to capture the context-free influence of these variables on choice (see, conceptual framework). Note that products were treated as replicates, with participants constrained to identify four brands that they wished to consider further. Thus, each product received the same number of choices.

To model the context-sensitive route, we included measures of font-product congruity, name-product congruity, and name-font congruity for each of EPA (nine variables). Congruity was measured as the usual multiplicative interaction terms in regression. Recall that font, name, and product were each measured for EPA on a scale ranging from −5 to +5, so if a name and a product both had positive ratings, or both had

\begin{table}[h]
\centering
\caption{Fonts and Products Used as Stimulus Materials in Study 2}
\begin{tabular}{ll}
\hline
\textbf{Fonts} & \textbf{Products} \\
\hline
\texttt{Albertus Medium}, \texttt{Avalon Normal}, \texttt{Avalon Quest}, \texttt{CG Times Bd.}, \texttt{Cinema}, \texttt{Classica}, \texttt{Dal}, \texttt{Gill Sans}.
\texttt{Globus}, \texttt{Medium}, \texttt{Lucida Sans Bd}, \texttt{Lucida Sans It}, \texttt{Mythology}, \texttt{Tourney}. & \texttt{Aquarium}, \texttt{Bricks}, \texttt{Camera}, \texttt{Double glazing}, \texttt{Dating agency}, \texttt{Dry cleaners}, \texttt{Fast food}, \texttt{Guest house}, \texttt{Ladies watches}, \texttt{Microwave ovens}, \texttt{Motorbikes}, \texttt{Optician}, \texttt{Pest control}, \texttt{Power tools}, \texttt{Tennis club}, \texttt{Used cars}, \texttt{Vitamins}, \texttt{Wallpaper} \\
\hline
\end{tabular}
\end{table}
negative ratings (e.g., for the potency dimension, a “masculine” name with a “masculine” product, or a “feminine” name with a “feminine” product), they multiply together in the regression interaction term to yield a positive "fit" score for that combination, thereby indicating congruity of name with product on the potency dimension. However, combinations from opposite poles of a dimension (e.g., P+ with P-: a “masculine” name with a “feminine” product, or vice versa) multiply to give a negative fit score, indicating an incongruent pairing. Fit scores were constructed similarly for the evaluation and activity dimensions. Finally, there were 14 position-on-the-page dummy variables denoting the box ads' locations (e.g., top left).

**Font effects.** The results from the logistic regression, predicting choice of which brands to phone is presented in Table 4. Position-on-the-page dummy variables have been omitted for brevity. Nagelkerke $R^2$ for the model was .048 ($p < .001$). The font related effects confirm those from Study 1. Study 1 showed that appropriateness was generally perceived to be greater when the font was high in evaluation than when it was low, high in potency than when it was low, and low in activity than when it was high. Correspondingly, choice in this study was positively influenced by both evaluation (FontE: $B = 0.30$, $p < .001$) and potency (FontP: $B = 0.13$, $p < .01$) and negatively influenced by activity (FontA: $B = -0.18$, $p < .001$). Thus brands dressed in good-pleasant (E+), strong-hard (P+), and passive-still (A-) fonts were chosen more frequently, independent of their consistency with the connotative meaning of the product or connotative meaning of the name.

Study 1 found that if fonts and products were congruent rather than incongruent along the potency or activity dimension, they were rated as more appropriate. Likewise in Study 2, there were significant FontP × ProdE ($B = .27, p < .001$) and FontA × ProdA ($B = .08, p < .001$) interactions terms. As anticipated, however, the interaction of FontP × ProdE was not significant ($B = .02, p > .5$). Thus, the congruity between the connotative meaning of the font and the connotative meaning of the product along the potency and activity dimensions, which had an impact on perceptions of appropriateness in Study 1, was an important determinant of choice as well.

Study 2 also showed that participants were more likely to choose a brand if it was in a font that had an obvious association with the product category, where obvious associations measures the strength of learned or figurative associations, or both ($B = .08, p < .001$). Modeling obvious associations here not only demonstrates its existence, but also partials it out as a potential confound of congruity derived from the shared connotative meanings between font and product.

**Name effects.** The name effects followed the pattern of results for font in almost every respect. There were effects for NameE ($B = .15, p < .001$), NameP ($B = .08, p < .01$), and NameA ($B = -.06, p < .05$). Brands with good-pleasant (E+), strong-hard (P+), and passive-still (A-) names were more likely to be chosen, independent of their consistency with the connotative meaning of the product or font. These findings exactly parallel the context-free influence of font on choice. Also, congruity of name with product along the potency dimension was significant ($B = .07, p < .02$), whereas congruity along the activity ($p > .35$), and evaluation ($p > .05$) dimensions was not significant. For context-sensitive name-product combinations, only congruity along the potency, but not along the activity dimension, was important. So, these findings for name exactly parallel those findings for font in

<table>
<thead>
<tr>
<th>Effect</th>
<th>$B$</th>
<th>Standard Error</th>
<th>Wald</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FontE</td>
<td>0.301</td>
<td>0.043</td>
<td>49.78</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td>FontP</td>
<td>0.125</td>
<td>0.043</td>
<td>8.30</td>
<td>.004*</td>
</tr>
<tr>
<td>FontA</td>
<td>-0.184</td>
<td>0.026</td>
<td>49.07</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td>NameE</td>
<td>0.153</td>
<td>0.035</td>
<td>18.97</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td>NameP</td>
<td>0.080</td>
<td>0.030</td>
<td>6.93</td>
<td>.008*</td>
</tr>
<tr>
<td>NameA</td>
<td>-0.056</td>
<td>0.027</td>
<td>4.22</td>
<td>.040*</td>
</tr>
<tr>
<td>FontE × ProdE</td>
<td>0.016</td>
<td>0.026</td>
<td>0.36</td>
<td>.550</td>
</tr>
<tr>
<td>FontP × ProdP</td>
<td>0.267</td>
<td>0.027</td>
<td>94.80</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td>FontA × ProdA</td>
<td>0.077</td>
<td>0.023</td>
<td>11.13</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td>NameE × ProdE</td>
<td>0.045</td>
<td>0.026</td>
<td>3.03</td>
<td>.081</td>
</tr>
<tr>
<td>NameP × ProdP</td>
<td>0.065</td>
<td>0.027</td>
<td>5.73</td>
<td>.017*</td>
</tr>
<tr>
<td>NameA × ProdA</td>
<td>-0.021</td>
<td>0.023</td>
<td>0.85</td>
<td>.355</td>
</tr>
<tr>
<td>NameE × FontE</td>
<td>0.011</td>
<td>0.030</td>
<td>0.12</td>
<td>.728</td>
</tr>
<tr>
<td>NameP × FontP</td>
<td>-0.008</td>
<td>0.028</td>
<td>0.08</td>
<td>.781</td>
</tr>
<tr>
<td>NameA × FontA</td>
<td>0.016</td>
<td>0.025</td>
<td>0.42</td>
<td>.519</td>
</tr>
<tr>
<td>Obvious Associations</td>
<td>0.080</td>
<td>0.023</td>
<td>12.51</td>
<td>&lt; .001*</td>
</tr>
</tbody>
</table>

Note. The subscripts evaluation (E), potency (P), and activity (A) refer to Osgood’s dimension of connotative meaning. The interaction terms measure the congruity of meaning between two concepts. For instance FontE × ProdE measures the congruity between font and product along with evaluation dimension. Variables entered the logistic regression in standardized form.

* $p < .05$ (or better).
ascenders and descenders; and are sans serif. Products low in potency, such as a hairdresser or perfume, should benefit from the opposite qualities on these visual characteristics. Henderson et al. did not have a typeface semantics dimension corresponding to activity, but reviewing our own sample of fonts that vary across the activity dimension in Tables 1 and 3, we suggest that active fonts lack harmony (unbalanced, rough, asymmetric) and are elaborate (distinctive, ornate, convey meaning, and often difficult to read).

So, the goals of simultaneously increasing evaluation and potency; decreasing activity; and increasing font-product congruity across the potency and activity dimensions, as implied by our regression results, are subject to trade-offs among the variables. We cannot simply say that the best fonts are in a particular part of EPA space, because the presence of significant congruity terms implies that the level of potency and activity for the product must also be taken into account.

Our font-product congruity findings across the potency dimension at first seem contrary with the findings of Pan and Schmitt (1996). Their work was based on the observation that certain forms of writing, for instance Chinese, are entirely visual, whereas others that employ an alphabet, such as English, are largely sound-based. This led them to predict that Chinese readers would pay more attention to the visual form of a multimodal presentation, while Americans would pay more attention to the sounds. They explored products along the dimension of masculinity versus feminine, which approximates potency in Osgood’s conceptualization. Product names were presented simultaneously on tape and on a screen to match the product on sound (a man vs. woman spoke the name) or on screen (“masculine” vs. “feminine” typeface was used). Whereas Chinese attitudes were affected by visual (product–typeface) congruity, Americans were affected by sound (product–voice) congruity. Indeed, Americans experienced no visual congruity. However, our results suggest that English speakers do pay attention to visual congruity, at least when there is no voice to compete for that attention. Pan and Schmitt’s use of a voice with the visual presentation may have effectively primed the sounding-out route for English speakers and rendered the visual route ineffective. If this is so, we have identified an important boundary condition.

Finally, the extent that choice is influenced by appropriateness and whether font effects are direct or are mediated by appropriateness offers an interesting avenue for future research. The issue is complex because some effects of typeface may operate via direct effects and others via mediated effects. It is interesting to speculate that the feelings elicited through the evaluative meaning of fonts have a direct impact on choice, whereas the effects of font-product congruity, which result in perceptions that the font “fits,” are mediated by perceptions of appropriateness.

In conclusion, some marketing interventions are attention capturing, rather similar to the drama of open-heart surgery. Other interventions may escape attention altogether yet, similar to the benefits of a good diet, they may be just as effective. Using the right typeface belongs to the latter kind of benefit, contributing unobtrusively over a long period of time to the image and sales of the brand. If not exactly red-hot cognition, it will at least be a sustained warm glow. Needing no further investment, the right typeface in a logo or packaging will be an inheritance passed from one brand manager to another. There are signs that this neglected source of benefit is now beginning to receive the attention it deserves from consumer researchers.

REFERENCES


ascenders anddescenders; and are sans serif. Productslow in potency, such asahairdresser or perfume, should benefit from the opposite qualities on these visual characteristics. Henderson et al. did nothave a typeface semantics dimension corresponding to activity, but reviewing our own sample of fonts that vary across the activity dimension in Tables 1 and 3, we suggest thatactive fonts lack harmony (unbalanced, rough, asymmetric) and are elaborate (distinctive, ornate, convey meaning, and often difficult to read).

So, the goals of simultaneously increasing evaluation and potency; decreasing activity; and increasing font-product congruity across the potency and activity dimensions, as implied by our regression results, are subject to trade-offs among the variables. We cannot simply say that the best fonts are in a particular part of EPA space, because the presence of significant congruity terms implies that the level of potency and activity for the product must also be taken into account.

Our font-product congruity findings across the potency dimension at first seem contrary with the findings of Pan and Schmitt (1996). Their work was based on the observation that certain forms of writing, for instance Chinese, are entirely visual, whereas others that employ an alphabet, such as English, are largely sound-based. This led them to predict that Chinese readers would pay more attention to the visual form of a multimodal presentation, while Americans would pay more attention to the sounds. They explored products along the dimension of masculine versus feminine, which approximates potency in Usgood’s conceptualization. Product names were presented simultaneously on tape and on a screen to match the product on sound (a man vs. woman spoke the name) or on screen (“masculine” vs. “feminine” typeface was used). Whereas Chinese attitudes were affected by visual (product–typeface) congruity, Americans were affected by sound (product–voice) congruity. Indeed, Americans experienced no visual congruity. However, our results suggest that English speakers do pay attention to visual congruity, at least when there is no voice to compete for that attention. Pan and Schmitt’s use of a voice with the visual presentation may have effectively primed the sounding-out route for English speakers and rendered the visual route inoperative. If this is so, we have identified an important boundary condition.

Finally, the extent that choice is influenced by appropriateness and whether font effects are direct or are mediated by appropriateness offers an interesting avenue for future research. The issue is complex because some effects of typeface may operate via direct effects and others via mediated effects. It is interesting to speculate that the feelings elicited through the evaluative meaning of fonts have a direct impact on choice, whereas the effects of font-product congruity, which result in perceptions that the font “fits,” are mediated by perceptions of appropriateness.

In conclusion, some marketing interventions are attention capturing, rather similar to the drama of open-heart surgery. Other interventions may escape attention altogether yet, similar to the benefits of a good diet, they may be just as effective. Using the right typeface belongs to the latter kind of benefit, contributing unobtrusively over a long period of time to the image and sales of the brand. If not exactly red-hot cognition, it will at least be a sustained warm glow. Needing no further investment, the right typeface in a logo or packaging will be an inheritance passed from one brand manager to another. There are signs that this neglected source of benefit is now beginning to receive the attention it deserves from consumer researchers.

REFERENCES


Received: April 20, 2005
Revision received: July 9, 2005
Accepted: July 29, 2005