Font appropriateness and brand choice

John R. Doyle*, Paul A. Bottomley

Cardiff Business School, University of Wales, Aberconway Building, Cardiff CF1 3EU, UK

Abstract

This paper illustrates how font, viewed as a component of a brand's visual equity, can enhance a brand's identity and build its market share. In a two-alternative forced-choice task, brands, covering 10 different product categories, were chosen twice as frequently when they were in an appropriate font relative to when they were not. This finding was replicated with no diminution of the effect even when brand names were highly connotative. In a subsequent small field study, consumers chose chocolates from a box of chocolates having an appropriate font rather than one having an inappropriate font on 75% of occasions. Our study establishes the importance for marketers to adopt a coherent font policy to cover current and possible future brand extensions.
© 2002 Elsevier Inc. All rights reserved.

Keywords: Font; Brand; Choice

1. Introduction

Although the most obvious aspect of a brand's projection is its name (Hart, 1998), other aspects also merit attention: collectively, shape, symbol, color, and lettering contribute to what has been called *visual equity* (Lightfoot and Gerstman, 1998). Visual equity is the value derived from 'visual form', i.e. the 'look and feel' of the brand. It contributes to brand awareness (e.g. FedEx's purple and chrimson lettering), to brand image (Keller, 1998), and thus ultimately to sales. Kohli and LaBahn (1997) sum it up neatly for the naming process: "a carefully created and chosen name can bring *inherent* and *immediate* value to the brand" (i.e. without the need for prior conditioning). Likewise, visual form should bring value to the brand.

Many elements make up the visual form: pictogram, color, letter font, etc. In this paper, we explore the value of font, where 'value' is operationalized as an increased likelihood of the brand being chosen.

1.1. Review of literature

Three-quarters of a century ago, Poffenberger and Franken (1923) asked people to rank order the appropriateness of 29 fonts for each of five "commodities" (automobiles, building materials, coffee, jewelry, perfume) and five "abstract qualities" (cheapness, dignity, economy, luxury, strength). They concluded that "differing type faces do vary in appropriateness and that judges are able to 'feel' this appropriateness or lack of appropriateness" (p. 328). Grouping products together that shared appropriateness with a given font type, automobiles, building material, and coffee clustered strongly together, while jewelry and perfume formed a second tight cluster. The former cluster was most appropriately represented by fonts that were emboldened, simple and easy-to-read (e.g. Cheltenham Bold, Century Bold). These in turn were associated with the qualities of "cheapness," "economy," and "strength." The latter cluster was most appropriately represented by fonts that were italicized, scripted, ornate (e.g. Caslon Old Style Italic; Typo Slope), and these in turn were associated with the qualities of "luxury" and "dignity."

The finding that fonts differ in appropriateness for products was replicated by Davis and Smith (1933), and by Schiller (1935)—the latter a close replication in that many of the same qualities and products were used. Schiller found that in the dozen or so years intervening between her study and Poffenberger and Franken's, coffee and automobiles were no longer perceived as merely commodity-like products, but had begun to take on elements of the luxury products perfume and jewelry.

In subsequent studies, fonts have been found to differ in their perceived appropriateness to represent not just different kinds of products, but also different kinds of books (Ovink, 1938), newspaper stories (Haskins, 1958), and

* Corresponding author.
E-mail address: DoyleJR@cardiff.ac.uk (J.R. Doyle).

professions (Walker et al., 1986). A parallel research stream relates and distinguishes fonts by rating scale descriptors (Poffenberger and Franken, 1923; Schiller, 1935; Davis and Smith, 1933; Ovink, 1938; Tannenbaum et al., 1964; Bartram, 1982; Rowe, 1982). In studies conducted in and after the 1960s, Osgood's semantic differential (Osgood et al., 1957) has been frequently used to select rating scale descriptors. So, for instance, *italics* has been found to be associated with Osgood's *activity* dimension, whereas *bold* was associated with Osgood's *potency* dimension.

1.2. Consequences of font appropriateness

Interestingly enough, Poffenberger and Franken (1923) took a pessimistic view of the behavioral consequences of their findings, believing that the differences in appropriateness which they observed would have little import: "…it might be argued that such differences as do exist are too slight to warrant consideration for practical purposes" (p. 328). Our work will show otherwise.

One of the best designed studies in the area, and the one with the most relevant theoretical implications for this paper, was conducted by Lewis and Walker (1989). They theorized that image (of which font is an example) and word offer parallel routes to meaning. (In an analogous way, someone's intonation offers meaning, over and above the actual words spoken.) They pretested a number of fonts to find a pair, one of which connoted "heavy" and the other "light." People had to press the left key if words "heavy" or "slow" were presented, and the right key if "fast" or "light" were presented. People's response times were faster when the font was appropriate for the word being presented, e.g. "heavy" presented in the font that connoted heaviness.

A second experiment showed that response times were again affected even when the quality (e.g. "fast") was implicit in an animal name (e.g. "cheetah") rather than explicitly presented as an adjective/adverb, as in their first experiment. Typical contrasts used were: cheetah (fast) versus tortoise (slow); and elephant (heavy) versus hedgehog (light). Part of the semantic representation of cheetah is "fast": and according to people pretested before the main experiment, part of the semantic representation of Palatino Italic is also "fast," while part of the meaning of Cooper Black is "slow." When meanings are in consonance, response times are speeded up; when they conflict, response times are slowed down. These experiments proved that there *are* behavioral consequences to font–word pairings. But equally important to take away from this study is the idea that fonts generate their own connotative meaning, which is processed independently of the meaning generated by the word per se.

Of more immediate relevance is the work of Perfect and his colleagues. Perfect and Askew (1994) showed that prior exposure to printed ads may induce people to rate them more favorably on a second presentation, even if they are unable to consciously recognize the ad. In other words, they

have demonstrated the now well-established phenomenon of *implicit memory* (Jacoby and Witherspoon, 1982; Graf and Masson, 1993) applied to the recollection of ads. They argued that exposure to the ad leads to increased *perceptual fluency* on second presentation (i.e. they just see more quickly what is in the ad and what it is about). "Were it tested appropriately, this increased fluency would be expressed as quicker response time on a perceptual identification task" (Perfect and Heatherley, 1997, p. 803). However, under certain circumstances, people may misattribute their own fluency not to prior exposure, but to features inherent in the ad: "…as a result [the ads] are rated as being more memorable, more likeable, more distinctive, and more eye-catching." (The misattribution theory of implicit memory is due to Jacoby et al., 1989).

A follow-up study (Perfect and Heatherley, 1997) attempted to pinpoint the locus of the effect, using the principle that the greater the perceptual overlap between the first and the second presentation, the greater the perceptual fluency, the greater the misattributed *affect* (emotional 'charge') generated. They hypothesized that *affect* would be greatest when the second presentation was the ad itself (which included the company logo), and the least when the company name appeared in a standard font (Times New Roman). It would be somewhere in between when the company name (logo) appeared in its usual font, but in the absence of the ad. Although they did replicate their previous study, they found only weakly supportive evidence for the more specific predictions.

Returning to Lewis and Walker's study, congruence between font and name leads to perceptual fluency (response times were faster), which we hypothesize from Perfect's work, should lead to positive *affect*, which should lead to an increased likelihood to investigate the product further, or even choose it. Satisficing behavior implies that to have your product investigated first will improve its chances of being eventually chosen. Our experimental hypotheses are therefore:

H1: Products presented in an appropriate (consonant) font will be investigated further more frequently than when presented in an inappropriate font.

H2: Products presented in an appropriate (consonant) font will be chosen more frequently than when presented in an inappropriate font.

Note that in this case perceptual fluency is not derived from prior exposure, but from two systems in the brain (image processing and verbal processing) producing congruent meanings. The two etiologies of fluency mirror the two methods described above, by which visual forms may drive visual equity—either by prior exposure (cf. Perfect and Askew), or by carefully choosing the right form to give "inherent and immediate value to the brand" (cf. Lewis and Walker).

2. Experiment 1

2.1. Pretest of font appropriateness

Our first task was to identify a pool of fonts appropriate for a pool of products. The 2×2 FCB grid (Vaughn, 1980) distinguishes products by whether they are 'thinking' or 'feeling' products, and whether they are high or low involvement products. It was used as a sampling frame to select 32 products, eight from each quadrant of the grid. Next, 27 fonts were chosen from those bundled with Microsoft packages, supplemented by some freeware/ shareware obtained via the Internet, covering traditional standard as well as less easily classifiable 'handwriting' and 'display' fonts. Clearly, no one could be asked to rate all of the 32 product $\times$ 27 font matrix. It was therefore divided up into 12 subgrids, each having 8 products $\times$ 9 fonts. Approximately 13 people (undergraduates) rated each subgrid (155 raters in all). They were asked to view a fictitious brand 'Martingale', which was printed in each of the nine fonts (using 26 pt for Arial—the other fonts were subjectively matched for size by the authors), and were instructed to rate the appropriateness of each font for the specified product, on a scale from 0 (*entirely inappropriate*) to 100 (*entirely appropriate*). The task took about 10 min to complete, and 150 usable questionnaires were completed.

The average appropriateness ratings in each cell of the 32 product $\times$ 27 font matrix in the pretest were used to identify a set of product categories and associated fonts according to the following criteria: (i) For each product, two fonts could be found that were sufficiently different in their appropriateness ratings; (ii) For each product (P1), a shadow product (P2) could be found such that if font F1 was appropriate for P1 and font F2 was inappropriate, then for P2 the situation was reversed: F2 was appropriate for it, while F1 was inappropriate. Note that if P2 is a shadow for P1, then P1 must be a shadow for P2; (iii) No font was used for more than one pairing of products. Given these constraints, it was possible to isolate 10 products with associated fonts. Products and related fonts are listed in Table 1 with corresponding appropriateness ratings derived from the pretest.

2.2. Brand names

Twenty names (family names) were selected from the phone book such that they were uncommon (though not odd), each had a different initial letter, and held no particularly strong associations for the investigators. In order to show the font off properly, each name also had to have at least half a dozen letters, including descenders and risers, and few repeated letters. The selected names were randomly assigned to product categories and served as brands, as listed in Table 1.

Table 1
Fonts and names used in Experiments 1, 2 and 3

Expt.	Product Category	Scenario	Font 1	Font 2	Name 1	Name 2
1	Car Rental	IF	Bodoni (68.9)	Mistral (31.7)	Bamforth	Pettifer
1	Specialty Jams	P	Bodoni (49.2)	Mistral (66.3)	Quimby	Stephen
1	Ice Cream	P	Snowdrift (83.1)	Arial (36.2)	Farleigh	Galloway
1	Life Insurance	IF	Snowdrift (8.4)	Arial (81.2)	Montford	Newberry
1	Fountain Pens	IF	French Script (81.3)	traffic (32.2)	Hoyland	Timperley
1	Mountain Bikes	IF	French Script (22.0)	traffic (56.4)	Denbury	Jacoby
1	Airlines	IF	Caslon Antq It (73.8)	BERTRAM (36.7)	Irving	Leighton
1	Soft Drinks	P	Caslon Antq It (45.9)	BERTRAM (68.9)	Amstey	Ellaway
1	Chocolates	P	Coronet (78.8)	LwrWSide (17.2)	Kersley	Roycroft
1	Batteries	P	Coronet (31.9)	LwrWSide (75.7)	Colgrave	Oldroyd
2	Bottled Water		Garamond Italics	STENCIL	Crystal	Aqua Vitalis
2 & 3	Chocolates		Signet Roundhand	Salem	Tempatation	Indulgence

Numbers in parentheses are mean appropriateness ratings of fonts for different product categories, measured in the pretest of Experiment 1 on a scale of $0-100$.
IF = "Investigate First" scenario; P = "Purchase" scenario.

2.3. Questionnaire

A four-page questionnaire was used. As general orientation, the first page read: "Companies spend a lot of time, effort and money in presenting what they hope is the right 'image' to customers. This study is interested in your thoughts about a number of brands in a variety of product categories. . ." The second and third pages were each headed by a description of the two choice scenario relevant to the five questions (one for each product category) which followed.

In the "investigate first" (IF) scenario, the instructions read: "The purchase of expensive products is often both a time consuming and highly involving process which requires you to search extensively for information on possible alternatives. You might, for example, consult the Yellow Pages and telephone a company for a quotation, or simply request a catalogue. Imagine that YOU are planning to buy each of the following types of product, which company/brand would you *investigate* first?"

In the "purchase" (P) scenario, the instructions were: "Imagine that you find yourself in a small 'corner shop' which only stocks two brands of each particular type of product, i.e. two brands of coffee, two brands of soap, etc. As far as you can tell from their external appearance, there is little to distinguish each pair of brands apart, and that includes their prices. Assuming you had planned to buy each of the following types of product, which brand would YOU choose to purchase?" By chance, five of the products fell into each of the scenarios.

Questions on the IF page asked: "Which of the following two brands of *Car Rentals* would you investigate first: Bamforth/Pettifer?" On the P page, the questions asked: "Which of the following two brands of *Specialty Jams* would you choose: Quimby/Stephen?" Note that if one name of the two alternatives was in Bodoni font, the other had to be in Mistral font, and vice versa. Alternatives within a product never appeared in the same font. Presentation order was extensively counterbalanced. For car rentals, for instance, half the questionnaires had Bamforth in Bodoni, with the other alternative, Pettifer, in Mistral; and vice versa for the other half of the questionnaires. Half the questionnaires ran through the five IF products first, and then ran through the five P products: half the time it was the other way round. Finally, products were also presented in two different orders for each scenario separately.

2.4. Procedure

One hundred and twenty Open University students filled in the questionnaire, handed out in small groups of 14 or so. Questionnaires were administered by their tutors (who were unknown to the investigators, and who were blind to the purpose of the experiment). The Open University offers part-time distance learning degrees, mainly to UK residents. The students are drawn from all walks of life (from judges to prisoners), some having originally received little formal schooling. As such, Open University cohorts are generally acknowledged to be much more representative samples of the UK population than typical undergraduate students. Three bottles of wine were raffled to all respondents who had properly completed the questionnaire.

2.5. Results

Table 2 contains the number of times each combination of font and name was chosen, and simple tests resulting from that. For instance, when the product was car rental, 80 (=41+39) people chose the name that was in the Bodoni font, but only 39 (=20+19) when the name was in Mistral.

Table 2
Number of choices made to one of two brand names when in different fonts, and for different product categories

				Probabilities		
				Font	Name	F × N
Car rental	Bamforth	**Bodoni** 41	Mistral 20	<.001	n.s.	n.s.
	Pettifer	39	19			
Specialty jams	Quimby	Bodoni 30	**Mistral** 49	<.001	<.001	n.s.
	Stephen	12	29			
Ice cream	Farleigh	**Snowdrift** 33	Arial 15	<.001	<.05	n.s.
	Galloway	46	26			
Life insurance	Montford	Snowdrift 10	**Arial** 54	<.001	n.s.	n.s.
	Newberry	5	50			
Fountain pens	Hoyland	**Fr. Script** 49	Traffic 6	<.001	n.s.	n.s.
	Timperley	53	12			
Mountain bikes	Denbury	Fr. Script 31	**Traffic** 33	n.s.	n.s.	n.s.
	Jacoby	27	28			
Airlines	Irving	**Caslon** 37	Bertram 23	<.01	n.s.	n.s.
	Leighton	38	22			
Soft drinks	Amstey	Caslon 25	**Bertram** 35	<.05	n.s.	n.s.
	Ellaway	24	36			
Chocolates	Kersley	**Coronet** 48	LWS 12	<.001	n.s.	n.s.
	Roycroft	47	12			
Batteries	Colgrave	Coronet 25	**LWS** 35	<.05	n.s.	n.s.
	Oldroyd	25	34			

Font and Name columns are probabilities from binomial tests on marginal totals. F × N is the interaction of font and name, tested by $\chi^2(1)$ on 2×2 cells. All effects labelled n.s. are $P>.2$. The emboldened font is the more product-appropriate font.

This was statistically significant by the binomial test, as indicated in the column headed "Font," and it was in the direction predicted by H1 and H2 (see appropriateness ratings in Table 1). Inspecting the row totals, 61 (=41+20) people chose Bamforth, and 58 (=39+19) chose Pettifer. The implication is that one name was not significantly preferred to the other (P>.2), as shown in the next column. χ^2 tests on the 2×2 table, presented in the final column, show that font and name did not interact. Thus, for car rentals, the *proportion* preferring Bodoni to Mistral was no different when the name Bamforth was used (41:20) compared to when Pettifer's was used (39:19). All tests identified as nonsignificant were P>.2; also, font effects were tested using one-tailed tests, since we are making a prediction about their direction. The null hypothesis for the font effect is that brands in the appropriate font would not be chosen more frequently than those in inappropriate fonts, whereas H0 for the name effect is just that neither of the two names would be chosen more frequently than the other (two-tailed).

It can be seen that of the 10 products, all 10 font effects were in the direction predicted by H1 and H2, and 9 of the 10 were statistically significant differences, the exception being mountain bikes. Name was significant on 2 of the 10 occasions. Galloway was preferred to Farleigh (ice creams), and Quimby to Stephen (specialty jams). Name did not interact with font for any of the products. Finally, for each product, the *z*-score from the normal approximation to the binomial can be used as an index of the size of the font effect. There was no significant difference between the size of the font effects for the products in the IF and P scenarios [independent measures $t(8) < 1$], thus strengthening the likelihood that the findings will generalize to different tasks.

Returning to car rentals, in isolation our results could be interpreted as showing that brands in Bodoni are preferred to those in Mistral, rather than as evidence for H1 and H2. However, this cannot be so, because if we look at car rentals' shadow product, specialty jams, the situation is reversed: there it would appear that brands in Mistral are preferred to those in Bodoni. In fact, when each product is paired with its shadow product, the entire pattern of results conforms to that predicted by H1 and H2: it is the specific pairing of font with product that is important. Furthermore, this finding is present for both P and IF scenarios.

2.6. Discussion

An average of 67% of all choices were made to products in the appropriate font. Furthermore, although 9/10 font effects were significant, only 2/10 name effects were, so in this experiment, font was a more powerful determinant than name in influencing choice. However, while we expressly manipulated the appropriateness of font to get extremes, names were chosen to be neutral, and so

the relative power of font and name to influence choice cannot be generalized from these data. Although family names *do* sometimes appear in the market as brand names, especially for services and family-founded firms, a sterner test of the font effect would be in the context of more connotative brand names. The next experiment seeks just such a test.

3. Experiment 2

3.1. Pretest of names and fonts

Four product categories were initially chosen and several dozen names were generated by the experimenters. The two product categories where the experimenters felt they had been most productive in generating names were then retained (box of chocolates and bottled water). On the basis of advice from three other people, six names were retained for each of these products. They were, for chocolates: Enigma, Forbidden, Indulgence, Intrigue, Mystery, Temptation; and for bottled water: Aqua-Vitalis, Cascade, Cloudburst, Crystal, Fountain, Glacier. Using the pretest data from Experiment 1 as a basis, eight fonts were selected for each product. In the pretest to Experiment 2, each of the six names appeared in each of the eight fonts in two different random orders, and for both products. Twenty-three people were asked to rate (on a 0–10 scale) the appropriateness of each of the 6 name × 8 font combinations, for both products separately.

3.2. Procedure

From this pretest data, two names and two fonts were selected for each product. For chocolates, they were (with mean ratings for Signet Roundhand font followed by Salem font in parentheses): Temptation (7.13, 4.00) and Indulgence (7.63, 3.88). And for bottled water they were (with means for Garamond Italics font followed by Stencil font): Aqua Vitalis (6.09, 4.17), and Crystal (6.30, 4.13). The print sizes for Signet, Salem, Garamond Italics, and Stencil were: 36, 24, 32, and 26 pt, respectively—perceptually approximately equally sized. These fonts and names met the criteria of having a good separation between font appropriateness (for the particular product category, and those particular names), while names were as appealing as possible (again for the product category, and those particular fonts), and equally as appealing as each other.

Respondents were students, similar to those in Experiment 1. The questionnaire asked them to look at the logos for bottled water: Aqua Vitalis (stylized in one of the fonts), and Crystal (stylized in the other font), then reply to the following:

[Q1] Please tick the logo you prefer (as a logo for bottled water).

[Q2] Please estimate what percentage of people in this room would agree with you. ___%
[Q3] Imagine you are a marketing director about to introduce a new bottled water onto the market. Please circle the logo you would choose for it.''

On the second half of the page, having exactly the same format, were questions about chocolates. Eight variants of the questionnaire were prepared by crossing all factors: so that each product appeared as the first product in half of them; so that each font appeared first in half of them; and so that each name appeared first in half of them ($2 \times 2 \times 2 = 8$). The eight variant questionnaires were randomly assigned among respondents. Eighty-one usable questionnaires were returned.

4. Results

Combining responses from Questions 1 and 2, it is possible to infer what percentage of people each respondent thought would choose the first alternative. This was used as the dependent measure in an analysis of variance. Three between-subject factors were: first alternative as name A or B (name), first alternative in font X or Y (font), and which of the two products appeared on the top half of the page (order).

For bottled water, the only significant result was for font: $F(1,73) = 28.42$, $P < .0001$, $\eta^2 = 0.28$. (The strength of effect statistic η^2 is a generalization of R^2). People thought others would choose the first alternative more often when it appeared in the appropriate font (Garamond Italics rather than Stencil), as hypothesized. One marginally significant results was that Aqua Vitalis was slightly preferred to Crystal as a name: $F(1,73) = 2.96$, $P = .09$, $\eta^2 = 0.04$. Similarly, for chocolates, font was significant: $F(1,73) = 62.51$, $P < .0001$, $\eta^2 = 0.46$. Again, people thought others would choose the first alternative more often when it appeared in the appropriate font (Signet rather than Salem). This result was tempered by an interaction of font with order: $F(1,73) = 4.40$, $P < .05$, $\eta^2 = 0.06$. The font effect was strongest when chocolates appeared on the top half of the page. However, font was still significant when analyzed for top and bottom halves of the page separately (both $P < .05$).

Responses to Question 3 were analyzed using the same structure of factors, and logistic regression (forward step using log-likelihood ratio criterion). Results revealed an even simpler pattern for both bottled water and chocolates. Only font was significant ($P < .0001$ for both products). No other factors or interactions were even marginally significant. For both products, the analysis of Q1 was virtually identical to Q3: $P < .0001$ for the font effect.

For bottled water, the alternative in the appropriate font was chosen 66/81 times (81%), and for chocolates 64/81 times (79%). The figure for chocolates compares almost

exactly with the proportion choosing the brand in the appropriate font in Experiment 1, namely, 95/119 (80%). Note also that Experiment 2 used different names (non-family brand names) and different fonts from those used in Experiment 1, yet the same end result was obtained. We therefore conclude that even when the brand names have meaning of their own, font can still have a powerful effect on people's choice.

5. Experiment 3

5.1. Field test

It must be admitted that many of the choices offered respondents in Experiments 1 and 2 are one step removed from actual choice situations. To remedy this shortcoming, a small field study was undertaken using the same names and fonts that were used for chocolates in Experiment 2. Round, metallic silver cardboard boxes (4 in. diameter, 2.25 in. in height) were used to contain chocolate truffles in a choice task. The box tops were decorated with silver string tied up in a shoelace bow, and a white label ($3 \times 5/8$ in.) attached, bearing one of the names ''Temptation'' or ''Indulgence'' in either Signet (32 pt.) or Salem (24 pt.)— i.e. approximately equal-sized. For each person, two boxes were placed on a blue tray. Half the respondents chose between ''Temptation'' in Signet versus ''Indulgence'' in Salem, and half chose between ''Temptation'' in Salem versus ''Indulgence'' in Signet. Each of these four possibilities appeared as often on the left as on the right. Forty subjects were convenience sampled by knocking on office doors of academic and support staff in a university. They were simply asked to choose a truffle from one of the *two* boxes that were on the tray in front of them. They did not see the other two boxes. Handedness of respondent was also noted, though since there were only two left-handers in the sample this was not carried forward as an explanatory variable. An equal quota of all distinct conditions was sampled.

Thirty of the 40 respondents (21 females and 19 males) chose to take a truffle from the box that had Signet font. This is significant by a simple binomial test ($P < .001$), but to check for possible interactions among the variables, logistic regression (forward step using log-likelihood ratio criterion) was used, with the between-subject factors of font, name, and gender as the explanatory variables of choice. The only factor to emerge significant was font ($P < .005$): all other factors and interactions were nonsignificant ($P > .2$).

The 75% of choices made to the appropriate font in this experiment are about the same proportion as the 79% and 80% found in Experiments 1 and 2, thus establishing that the results we find using paper-and-pencil tests can indeed be transferred to actual choice behavior, and with much the same degree of impact.

6. Summary and discussion

We have shown that the aptness of a font for a particular type of product does have consequences. In Experiment 1, brands presented in appropriate fonts were chosen more often than brands presented in inappropriate fonts (averaged over the 10 product categories, they were chosen in the ratio 2:1). One limitation of that study was that the brand names were connotatively neutral for the respondents. Highly connotative names, it might be argued, might overwhelm the font effect. In Experiment 2, therefore, brands were chosen to have connotatively rich names: nonetheless, in the two product categories examined, appropriate fonts were still chosen more often than inappropriate ones (in the ratio 4:1). Finally, in Experiment 3, rather than merely indicating what they *would* choose in certain scenarios, people were presented with an *actual* choice of chocolates from two chocolate boxes. Again, the box with the appropriate font was chosen more often (ratio 3:1).

One interesting finding is that, in both the main experiments and the pretests, we consistently found no interaction of gender with font. In particular, women do not prefer lighter, more scripted, scrolled (i.e. so-called "feminine") fonts (such as Signet). This equality between the sexes certainly should make life easier for the company that would use a font to project its brand(s) in mixed-gender markets.

One limitation is the reliance on pretesting materials before the main choice experiments. Not only is this procedure cumbersome, but it may seem to make the results a foregone conclusion. Some may see font–product appropriateness and font–product choice as just alternative ways of eliciting the same information. This view, of course, can only be taken with the benefit of hindsight. It is, in fact, a major result of our experiments. Nonetheless, the whole research stream would gain from being able to take off-the-shelf fonts, as it were, that had known qualities, independently measured. Unfortunately, the principal "dimensions of font connotation" are still not known.

An obvious next step in the drive towards realism is the introduction of color and image as criteria for exercising choice between brands in addition to that of name and font investigated here. This may be in the form of advertising, point-of-purchase materials, or packaging.

If name and font elicit dual routes to meaning, then it should be possible, in future work, to manipulate the relative contribution that each makes. For instance, we could hypothesize that when people expect to have to justify their purchase decisions, they will rely more heavily on verbal cues (i.e. name) relative to visual cues (e.g. font) in the product. Because names are easier to talk about than fonts, name-based choice is easier to justify. It is also possible to see a parallel here between the central and peripheral processing routes of the Elaboration Likelihood Model (Petty et al., 1983); it might be hypothesized that visual cues are less elaborated anyway, irrespective of manipulations of involvement.

Our results show that when names are highly appropriate for the brand, font may still have a sizeable impact. When the advantage of one name over another reaches a limit, font appears to offer a wholly new way to outperform the competition (or be outperformed). This means that marketers must pay close attention to the multidimensional nature of a brand's projection. Name is certainly important, but so too is font (and other elements of visual form, presumably).

The findings also have important consequences for companies that would extend their brand name into new product categories. When creating a cross-product brand identity, consideration must be given to the portfolio of product categories in which the brand will compete. As such, there may be occasions where it is wise to accept a slightly suboptimal font in the parent product category if the font travels better to other categories that will subsequently be entered. However, it is also possible that it would not be appropriate anyway to extend a brand into product areas where the font was inappropriate (e.g. life insurance companies extending into ice cream, or vice versa). Font appropriateness, and more generally visual identity, may be an effective marker for identifying just which categories are compatible.

We have shown that fonts are an important part of a brand's projection. In so doing, we have also shown that the topic is eminently researchable, which makes its neglect all the more puzzling.

References

Bartram D. The perception of semantic quality in type: differences between designers and non-designers. Inf Design J 1982;3(1):38–50.

Davis RC, Smith HJ. Determinants of feeling tone in type faces. J Appl Psychol 1933;17:742–64.

Graf P, Masson MEJ. Implicit memory: new directions in cognition, development, and neuropsychology. Hillsdale (NJ): Lawrence Erlbaum; 1993.

Hart S. Developing new brand names. In: Hart S, Murphy J, editors. Brands: the new wealth creators. Basingstoke (UK): Macmillan; 1998. p. 34–45.

Haskins JB. Testing the suitability of typefaces for editorial subject-matter. Journal. Q 1958;35:186–94.

Jacoby LL, Witherspoon D. Remembering without awareness. Can J Psychol 1982;36:300–24.

Jacoby LL, Kelley CM, Dywan J. Memory attributions. In: Roediger HL, Craik FIM, editors. Varieties of memory and consciousness: essays in honour of Endel Tulving. Hillsdale (NJ): Lawrence Erlbaum; 1989. p. 391–422.

Keller KL. Building, measuring, and managing brand equity. Upper Saddle River (NJ): Prentice-Hall; 1998.

Kohli C, LaBahn DW. Creating effective brand names: a study of the naming process. J Advert Res 1997;67–75 ([January/February]).

Lewis C, Walker P. Typographic influences on reading. Br J Psychol 1989; 80:241–57.

Lightfoot C, Gerstman R. Brand packaging. In: Hart S, Murphy J, editors. Brands: the new wealth creators. Basingstoke (UK): Macmillan; 1998. p. 46–55.

Osgood CE, Suci GJ, Tannenbaum PH. The measurement of meaning. Urbana: University of Illinois Press; 1957.

Ovink GW. Legibility, atmosphere and forms of printing types. Leiden: Sijthoff; 1938.

Perfect TJ, Askew C. Print ads: not remembered but memorable. Appl Cogn Psychol 1994;8:693–703.

Perfect TJ, Heatherley S. Preferences for advertisements, logos, and names: effects of implicit memory. Psychol Rep 1997;80:803–8.

Petty RE, Cacioppo JT, Schumann D. Central and peripheral routes to advertising effectiveness: the moderating role of involvement. J Consum Res 1983;10:134–46.

Poffenberger AT, Franken RB. A study of the appropriateness of type faces. J Appl Psychol 1923;7:312–27.

Rowe CL. The connotative dimensions of selected display typefaces. Inf Design J 1982;3(1):30–7.

Schiller G. An experimental study of the appropriateness of color and type in advertising. J Appl Psychol 1935;19:652–64.

Tannenbaum PH, Jacobson HK, Norris EL. An experimental investigation of typeface connotations. Journal Q 1964;41:65–73.

Vaughn R. How advertising works: a planning model. J Advert Res 1980;20(5):27–33.

Walker P, Smith S, Livingston A. Predicting the appropriateness of a typeface on the basis of its multi-modal features. Inf Design J 1986;5(1):29–42.