The Rhetoric of Typography: The Awareness and Impact of Typeface Appropriateness

EVA R. BRUMBERGER

INTRODUCTION

The field of technical communication is experiencing a growing emphasis on visual rhetoric. We have an emerging body of theory and research on which to build our practice, but little attention in that body of work has as yet been given to typography as a component of visual communication. The majority of studies focused on typography have examined the role of typography in readability and legibility (see, for example, Felker 1980; Tinker 1963; Sutt 1959). Only a handful of studies have looked beyond typography’s role in the physical act of reading to its role in the rhetorical act of meaning-making.

As design decisions become more and more the domain of the technical writer, we cannot afford to overlook such a central component of visual communication. The studies presented here extend the discussion of typeface persona begun in my article “The rhetoric of typography: The persona of typeface and text” (2003); they examine the role of typeface appropriateness in readers’ interactions with a print document.

RELATED LITERATURE

Schriver and others have stressed the role that type plays in “seeing the text”—a phrase borrowed from Bernhardt (1986)—in clarifying its purpose and organization. Schriver (1997) also discusses the role of typeface mood, personality, and tone, emphasizing the importance of connecting these to a document’s genre, purpose, and context. As she argues,

... designing legible documents is not enough. A second important characteristic of well-chosen typography is rhetorical appropriateness—the relationship between the typeface, the purpose of the document, its genre, the situation, and the audience’s needs, desires, and purposes. (p. 203)

Brin. (1996) similarly states, “The moment a text and a typeface are chosen, two streams of thought, two rhetorical systems, two sets of habits, or, if you like, two personalities, intersect” (p. 22). Even Warde (1956) has supported this idea, observing that “The best part of typographic wisdom lies in this study of connotation, the suitability of form to content” (p. 148).

Empirical evidence does exist for the intuitive notion that typefaces have personas. The earliest of this research identified “atmosphere” values for typefaces (see, for example, Poffenberger and Franken 1923). Subsequent studies investigated whether professionals and laypersons perceive typefaces differently (see Brin. 1961; Tannenbaum, Jacobson, and Norris 1964). The most recent studies have identified personality profiles for typefaces (Brumberger 2003; Bartram 1982; Rowe 1982).

Additionally, a group of studies have investigated typeface suitability. The earliest project of this nature was conducted in 1920, when Berliner had subjects sort 18 hand-drawn letterforms according to their appropriateness for use in advertising four different products. She found that participants agreed about the appropriateness of a particular letterform for a particular product (cited in Davis and Smith 1933). This study established the pattern for subsequent studies that used typefaces instead of hand-drawn forms.

Poffenberger and Franken (1923) required subjects to rank typefaces in order of perceived appropriateness for a
particular purpose—either for advertising a specified commodity (such as coffee) or for expressing a particular abstract quality (such as dignity). Like Berliner, they focused on the use of type for advertising. Poffenberger and Franken's data demonstrate clearly that typefaces do vary in appropriateness, and that individuals sense the appropriateness or lack thereof.

Ovink similarly elicited subjects' judgments regarding typeface appropriateness. Participants judged the appropriateness of a set of typefaces for eight different literary subjects and also for eight different advertising themes (cited in Spencer 1969; Wendt 1968).

Haskins (1958) investigated the hypothesis that typefaces vary in appropriateness with changes in text subject and attempted to determine whether there were any "all-purpose" typefaces that could be used appropriately in a variety of contexts. Haskins gave participants the first page of several articles, varying only the typeface used for the main titles. Participants rated the appropriateness of the title typeface for each article; according to Haskins, the data suggest the hypothesis that appropriateness is more of an issue for texts involving "some kind of conflict or tension" (p. 191); typeface appropriateness appeared less important for entertainment pieces.

Finally, in a much more recent study, Walker, Smith, and Livingston (1986) investigated the hypothesis that the appropriateness of a typeface is in part determined by the extent to which it shares qualities with the ideas that it is being used to represent. Using semantic differential scales (lists of paired opposite terms such as masculine/feminine and strong/weak), Walker, Smith, and Livingston determined that most typefaces are characterized by a range of qualities (such as "hard," "unfriendly," "happy," "loud," and so forth), as are most professions. They concluded that the appropriateness of a typeface for a particular profession was determined by the number of qualities that it shared with that profession.

Each of these studies demonstrates readers' awareness of typeface appropriateness and their ability to make judgments regarding the appropriateness of typefaces for a specific purpose or text. However, the most recent of these investigations is now over 15 years old, and the world of typography has changed dramatically in that time, with advances in desktop publishing and significant shifts in reader characteristics. Additionally and equally importantly, none of these studies examined the impact of typeface (in)appropriateness on readers' interactions with a document; the studies that I have conducted begin to address this larger issue.

RESEARCH METHODOLOGY
This research project consisted of a pair of studies. The first study investigated the hypothesis that readers are aware of (in)appropriateness in typeface/text pairings, using current typefaces to verify and extend the results of previous studies. The second study investigated the extent to which dissonance between typeface and text persona affects readers' perception of the text itself.

Research participants
Participants in both studies were undergraduate students enrolled in an introductory psychology course, all of whom were required to participate in four hours of research studies. Participation in one of my studies precluded participation in the other.

Because a substantial body of research suggests there are significant differences in the ways that males and females use and interpret language (Crawford and Chaffin 1986; Edelsky 1977), gender was a potentially important variable in this project. Thus, half of the participants in each study were male, and half were female. This division allowed me to perform statistical analyses for each group separately and to make comparisons between groups.

I collected other demographic information as well, specifically participants' ethnicity and first language, World Wide Web usage, age, and academic major. Differences due to ethnicity and first language seem likely, as there are substantial cultural differences in language usage (Beamer 1992; Thrush 1997). Additionally, World Wide Web usage may affect individuals' perceptions of visual rhetoric since the Web is a particularly visual medium and one that is used extensively by many college students. Finally, age and academic major could conceivably also impact participants' interactions with both typefaces and text passages.

While it was not feasible to treat each of these factors as an independent variable because the resulting groups would have been too small to allow for meaningful statistical analysis, I did conduct secondary analyses to get a sense for whether any of these demographic factors affected participant responses.

STUDY 1: AWARENESS OF TYPEFACE APPROPRIATENESS
The first study in the project addressed the following research question: Do readers identify certain typefaces as more (or less) appropriate for certain texts? I selected the typefaces based on personas identified in a previous study (Brumberger 2003). In that study, three distinct categories of typeface persona were identified: "elegant," "direct," and "friendly." The typefaces chosen for the current study were the strongest representatives of each category. Thus, three typefaces were used in the study: 
- Caslon Regular ("elegant"),
- Arial ("direct"), and
- Bouchouh Md BT ("friendly").

I also selected text passages based on my previous study (Brumberger 2003), which identified three distinct categories of text persona: "professional," "violent," and "friendly." The texts chosen for the current study were the
The Rhetoric of Typography: Typeface Appropriateness

Brumberger

TABLE 1: TYPEFACE/TEXT PAIRS USED IN STUDY 1

<table>
<thead>
<tr>
<th>Text Category</th>
<th>Elegant</th>
<th>Direct</th>
<th>Friendly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional</td>
<td>Cognitive psychology</td>
<td>Cognitive psychology</td>
<td>Cognitive psychology</td>
</tr>
<tr>
<td>Rainbow six</td>
<td>Arial</td>
<td>Gray</td>
<td>Arial</td>
</tr>
<tr>
<td>Neusweek</td>
<td>Arial</td>
<td>Gray</td>
<td>Arial</td>
</tr>
</tbody>
</table>

The selection of typefaces and texts resulted in nine conditions (see Table 1); each participant received all conditions. Thus, participants received each of the three texts in each of the three typefaces. The texts were displayed in 11.5- to 12-point type to control more closely for the variations in size among the typefaces.

Materials and procedure

The 36 participants (18 male, 18 female) were given a packet containing a demographic data form, instructions, and three passages of text; each text was presented three times, each time in a different typeface, as described above. Because participants were judging appropriateness and were not reading for comprehension, repeating texts was not problematic. The text passages were randomly distributed in each packet to avoid any effects due to order. At the bottom of each page was the rating scale shown in Table 2.

Participants were instructed to rate the appropriateness of a typeface for a particular text based on how well the typeface was suited to the ideas expressed by the content, tone, and style of the text passage. The task took approximately 30 minutes for each subject to complete and was administered to groups of up to eight people at a time. Because participants were seated around a conference table, I could observe them while they completed the task.

Methods of analysis

I performed a two-way analysis of variance (ANOVA) on the data from Study 1 to identify the font-by-text interactions—to identify whether certain typefaces are consistently deemed more (or less) appropriate for certain texts. An ANOVA identifies whether observed differences are statistically significant, as measured by the probability (p) of their occurring by chance. Traditionally, a value of p ≤ 0.05 is considered significant (that is, the results would occur by chance only five or fewer times in 100 trials—or 5% of the time), whereas a value of p ≤ 0.01 is considered highly or very significant.

In addition, I analyzed the data to identify any effects associated with age, academic major, extent of use of the World Wide Web, ethnicity, or gender.

Results

Typeface appropriateness. The ANOVA revealed highly statistically significant main effects of both text (p < 0.01) and typeface (p < 0.01), as well as a very significant text-by-typeface interaction (p < 0.01). Participants had very clear and consistent opinions about typeface appropriateness. They rated Arial—the “direct” typeface—as generally more appropriate regardless of text persona. Not surprisingly, however, Arial was rated as most appropriate for the “professional” text. Also as one would expect, Bouhous Md BT the “friendly” typeface, was rated more appropriate for the Neusweek passage, the “friendly” text, than for the other two texts. Finally, Counselor, the “elegant” typeface, was rated as the least appropriate typeface for overall use, but it was seen as more appropriate for a professional text than for either of the other two text types.

Figure 1 provides a comparison of the mean appropriateness ratings for each typeface and text pairing. Figure 2 presents the same data in a different form to more clearly illustrate the text-by-typeface interaction (note that similar patterns are observed for the “professional” and “violent” texts, whereas a different pattern is exhibited for the “friendly” text).

Demographic differences. I also analyzed the data to identify any effects linked to the demographic factors discussed previously. It was not feasible to identify any effects linked to age, as only six of the participants were outside the 18–23 age range. Additionally, it was not possible to identify effects that might be due to academic major, as participants reported a wide range of majors. While there were some majors with several participants, the groups were not sufficiently large to allow for a statistically reliable or valid analysis.

Almost two-thirds of the Study 1 participants (22) used the World Wide Web only 5 hours or fewer per week.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
One-third of the participants used the Web 6–10 hours per week, with two participants using the Web more than 15 hours per week. I performed an ANOVA on the low usage (5 hours or fewer) and high usage (5+ hours) groups to determine whether Web use affects readers’ perceptions of typeface appropriateness. There were no significant effects due to World Wide Web use.

English was the first language for most of the participants in Study 1; only 5 participants were nonnative speakers. Removing the data of these nonnative speakers did not alter the results significantly. Just over half (19) of the participants were White, 14 were Hispanic, two were American Indian/Alaskan Native, and one was Asian. An ANOVA performed using only the data from White and Hispanic participants revealed no significant effects due to ethnicity.

Surprisingly, male and female participants did not differ significantly in their ratings; an ANOVA revealed no significant effects due to gender. In short, analysis of the data revealed no statistically significant effects linked to demographic factors.

Discussion
The Study 1 data indicates that readers are quite aware of dissonance between typeface and text personas. Participants clearly perceived typefaces with particular personas to be more appropriate for texts with particular personas. Not surprisingly, a typeface whose persona matched closely that of the text was seen as more appropriate for that text (for example, the “friendly” typeface was perceived to be more appropriate for the “friendly” text than for other texts).

However, the data also suggests that some typefaces may be considered “all-purpose”—appropriate for most text types. This result may be a result of usage; participants may rate a typeface that they are accustomed to seeing used for a variety of text types as being generic or all-purpose—as universally appropriate.

Study 2: Typeface Appropriateness and Text Personality
Study 2 investigated the extent to which the visual tone of a text—the persona established by its typeface—affects readers’/viewers’ perception of its verbal tone—the persona established by the style and content of the text itself. This study addressed the following research question: To what extent does the persona of the typeface affect readers’ perception of the persona of the text itself? Study 2 thus investigated the impact of readers’ perceptions of typeface appropriateness; however, it did so without directly asking participants for their opinions.

Instead, Study 2 used typeface/text combinations like those in Study 1 and asked participants to rate each text passage on a list of attributes, thereby indirectly assessing whether perceived clashes in typeface and text persona affected the way that participants viewed the text itself.

Data Collection Method
The data collection method was a modification of the attribute scaling methodology known as the semantic differential (Osgood, Suci, and Tannenbaum 1957). The semantic differential approach entails presenting participants with a series of paired opposite terms (such as masculine/feminine, strong/weak), referred to as semantic differential scales. For each concept being judged (in this case, a text passage), participants are asked to indicate the point on
Study 1 because that study asked participants to rate appropriateness directly; the text passages were presented in all possible orders.

Thus, this study used a counterbalanced between-subjects design (each participant received a subset of the possible conditions) with two independent variables. The rating scales were displayed on a separate page, so I was able to increase the typeface sizes for the text passages slightly; the texts were displayed in 12- to 12.5-point type.

**The Nelson-Denny reading comprehension test**

Because there may be a correlation between reading ability and perceptions of visual and verbal persona, I administered part two of the Nelson-Denny Reading Test (1993 version, Form H) to all participants to identify participants’ reading levels. The Nelson-Denny is a nationally standardized and normed multiple-choice comprehension test for high school students, college students, and adults. The Reading Comprehension and Rate portion of the test takes 20 minutes, the first minute of which is used to determine reading rate.

The Nelson-Denny contains seven reading passages and a total of 38 comprehension questions, each with 5 answer choices. The text passages are selected from current, widely used, high school and college texts.

**Materials and procedure**

The initial step in Study 2 was administration of the Nelson-Denny reading comprehension test. Participants were then presented with a packet containing passages of text; each text was shown on a separate page, with the attribute list displayed on the facing page.

A total of 72 participants (36 male, 36 female) were given two sets of materials: part I was the Nelson-Denny test; part II included a demographic data form and instructions, followed by three text passages and the attribute lists. Participants ranked each text passage on a 7-point scale for each of the 20 attributes. They were asked to work quickly, basing their responses on their first impressions.

The tasks took approximately 40 minutes for each subject to complete and were administered to groups of up to 8 people at a time. Because participants were seated around a conference table, I could observe them while they completed the task.

**Methods of analysis**

I performed a two-way ANOVA on the data from Study 2 to determine whether the mean attribute ratings for each text differed significantly as a result of typeface. I also analyzed the data to identify any effects due to gender.

**Results**

**Effects of typeface persona.** I used a significance level of $p < 0.01$ for the analyses. The ANOVA revealed a highly
TABLE 3: TYPEFACE/TEXT PAIRS USED IN STUDY 2

<table>
<thead>
<tr>
<th>Text Category</th>
<th>Elegant</th>
<th>Direct</th>
<th>Friendly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional</td>
<td>Cognitive psychology</td>
<td>Cognitive psychology</td>
<td>Cognitive psychology</td>
</tr>
<tr>
<td>Rainbow six</td>
<td>Rainbow six</td>
<td>Rainbow six</td>
<td>Rainbow six</td>
</tr>
<tr>
<td>Neuwaveek</td>
<td>Neuwaveek</td>
<td>Neuwaveek</td>
<td>Neuwaveek</td>
</tr>
<tr>
<td>CounselorScript</td>
<td>Times New Roman</td>
<td>Times New Roman</td>
<td>Times New Roman</td>
</tr>
<tr>
<td>Bouhaus Md BT</td>
<td>Bouhaus Md BT</td>
<td>Bouhaus Md BT</td>
<td></td>
</tr>
</tbody>
</table>

significant effect of text passage ($p < 0.001$). In other words, not surprisingly, there were significant differences in attribute ratings among the text passages. There were three attributes for which there was no text effect: "confident," "pretentious," and "straightforward," meaning that the texts were not seen as differing significantly from one another on these qualities.

The ANOVA revealed no significant differences due to typeface; nor, with one exception, were there any significant typeface-by-text interactions. In other words, attribute ratings for each text passage did not vary significantly due to typeface persona.

There was a highly significant typeface-by-text interaction on the "serious" attribute ($p = 0.004$). While the "professional" text (Cognitive psychology) was rated as equally serious regardless of typeface, the "violent" text (Rainbow six) was rated as most serious when presented in Times New Roman (the "direct" typeface) and least serious when presented in CounselorScript (the "elegant" typeface). I observed a similar pattern for the "friendly" text (Neuwaveek), though it was not as pronounced (see Figure 4).

Figure 4. Study 2, text-by-typeface interaction. The lines illustrate patterns of ratings of seriousness; note the different pattern exhibited by the "professional" text.

Demographic differences. The only demographic factor analyzed for Study 2 was gender because previous related studies have revealed very few differences linked to other demographic factors. Some gender differences were observed.

I performed an ANOVA to determine whether the typeface-by-text interaction described above was linked to gender in any way. The ANOVA revealed a highly significant effect of gender ($p = 0.004$). Male and female participants exhibited different patterns of ratings for each of the texts (see Figure 5). Their responses were most similar for the "friendly" text (Neuwaveek). Males and females both rated the "friendly" text as more serious when it was presented in Bouhaus Md BT (the "friendly" typeface); this result was more pronounced for males than for females.

Males also rated the "professional" text (Cognitive psychology) as most serious in Bouhaus Md BT and least serious in CounselorScript. In contrast, females perceived this text as marginally more serious in CounselorScript and slightly less serious in Times New Roman.

Finally, males rated the "violent" text (Rainbow six) as most serious when it was presented in Times New Roman and least serious in Bouhaus Md BT. Females similarly rated the text as least serious in Bouhaus Md BT, but their ratings were not markedly different for any of the three typefaces.

Figure 5. Study 2, gender differences.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Discussion

The data indicates that, for the tasks, texts, and typefaces used in this study, typeface persona did not have a significant impact on readers’ perception of text persona. That is, the visual personality of a text did not have a large impact on readers’ perception of its verbal personality. One possible explanation for this data may be that the texts chosen for the study all exhibited very strong personas themselves; the typeface persona may have been unable to overpower the text persona. A similar study with more neutral texts (and/or stronger typeface personas) may reveal that the visual tone has more of an effect. A second possibility is that the dissonance—the conflict—between the typeface and text personas was not pronounced enough to have a sizable effect.

Similarly, a different group of texts may result in larger gender differences. It is interesting that the text from Study 2 with the largest gender differences was the “violent” text. Males’ ratings of the seriousness of this text varied with typeface much more than did females’ ratings. It may be that females found the violent nature of the text strong enough to render the visual tone irrelevant, whereas males were less sensitive to the violence in the text and therefore less affected by the text persona. The data does seem to support the notion that males and females differ in their use and interpretation of language (see Crawford and Chaffin 1986; Edelsky 1977), and this difference appears to hold for both verbal and visual language.

CONCLUSIONS

This project began with two hypotheses:

1. Readers are aware of typeface/text matches and mismatches; that is, they are aware of the (in)appropriateness of a particular typeface for a particular text.

2. Typeface persona colors readers’ perceptions of text persona.

The data from Study 1 provides clear support for the first hypothesis. Participants in the study had strong opinions about the appropriateness of particular typefaces for particular text passages, and they were aware of typeface/text mismatches.

These results align for the most part with results from previous research on typeface appropriateness (see Poffenberger and Franken 1923; Haskins 1958). However, they do not support Haskin’s contention that appropriateness is perceived as more important for texts involving “some kind of conflict or tension” (p. 191) and less important for texts whose purpose is to entertain, possibly because this classification is somewhat of an oversimplification (for example, the Clancy text used for this project is intended to entertain, but it certainly involves conflict and tension as well).

The Study 2 data failed to provide substantial evidence for the hypothesis that typeface persona colors readers’ perceptions of text persona. For the most part, typeface/text dissonance did not significantly affect readers’ perceptions of text persona. However, it is quite possible that results would be very different with more dissonant typeface/text pairings or with pairings in which the type persona is substantially stronger than the text persona. The identification of typefaces and texts with strongly opposing personas would allow further investigation of the impact of such dissonance on the reading process.

Similarly, presenting a text that is neutral or ambiguous in a typeface with a strong persona (for example, a gender-neutral text in a highly feminine—or masculine—typeface) would allow further investigation of the impact of typeface persona on text persona.

A final issue that was not addressed in this project but is certainly worthy of investigation is whether readers’ awareness of typeface appropriateness affects other interactions with a document, including reading comprehension, reading time, or readers’ perceptions of the writer’s ethos and credibility. Further exploration of these issues would be particularly helpful in clarifying the role of typography as a rhetorical element.

ACKNOWLEDGMENTS

I would like to thank the Society for Technical Communication for supporting this project through a research grant. Thanks also to Dr. Douglas Gillian, of the New Mexico State University Department of Psychology, for his assistance with statistics, and to Dr. Stephen A. Bernhardt.

REFERENCES


**EVA BRUMBERGER** is an assistant professor at Virginia Tech, where she teaches in the professional writing program. She completed the studies discussed in this article as part of her dissertation research in the Rhetoric and Professional Communication program at New Mexico State University. Before pursuing her doctorate, she taught at the University of Wyoming. Her prior experience also includes adjunct teaching at three other institutions and working as a technical writer and editor on a full-time and freelance basis. Her research interests include visual rhetoric, document design, visual communication pedagogy, and technical communication pedagogy. Contact information: ebrumberg@vt.edu