Does ugliness affect judgment?
The effect of typography on movie choice

Sushi Subburamu
Bronx High School of Science
Advisor: Richard Lee

Mentor: Professor Denis G. Pelli
Psychology and Center for Neural Science
New York University

ABSTRACT

Beauty is a quality that gives pleasure to the senses. This may shape the decisions we make. While it is often supposed, in literary criticism, that we can ignore the beauty of appearance to make a pure judgment about content, in typography it is a widely accepted principle that the text design should be neither pretty nor ugly; it should be “transparent” so as not to distract from the content. We tested this directly, measuring the effect of font of the movie summary on the reader’s rating of the movie, based on reading the summary. We tested a common and an ugly font. We wondered whether the ugly font would distract the reader enough to make them a poor judge of the text’s content. We find that the ugly font substantially reduced the movie ratings, but did not reduce the differences. Thus, the ugly font did bias the movie rating, but, contrary to conventional typographic wisdom, did not impair judgment of content. In real life, the movie-summary font would affect whether the reader goes to the movies, but would not affect, once there, which movie is chosen.
INTRODUCTION

Be it a building or a website, most manmade objects are designed to be both functional and aesthetically pleasing. An object is an identifiable portion of an image that can be interpreted as a single unit. Words are objects. Printed words, text, are written in a typeface, which affect the reader’s perception of the text. Text conveys meaning to the reader. The writer wants to persuade with her words and good typography may help.

Publishers choose their fonts carefully. Recent studies have shown typefaces to convey mood, attitude, and tone (Shaikh, Chaparro, & Fox 2006). In 2005, The Guardian switched from “the tired old gentleman’s mixture of Miller, Helvetica and Garamond” to Guardian Egyptian. In 2006, The New York Times hired former Face magazine artistic director and font designer Neville Brody to design the Bespoke font. And in 2007 the Globe and Mail called upon Nick Shinn to create Globe and Mail Text (Gaertner 2008; Thangaraj 2004). Corporations use fonts in corporate branding. Advertisers use fonts to persuade consumers to buy a product. It is claimed that the most pervasive design element in marketing materials is typestyle (Hutton 1987; McCarthy et al. 2002; Pan et al. 1996).

Typography is the design of a page of text, including the choice of font. Aaron Burns, founder of ITC, the first American font retailer, comments, "In typography, function is of major importance, form is secondary and fashion is meaningless." Graphic artists rely on their individual perception of the page. Scientists focus on legibility and readability determined by size of type, style of type, quality of printing surface and other human factors. With the introduction of digital fonts, designers began creating fonts without any limits, and the number of typefaces increased. With hundreds of thousands of fonts available, it can be daunting to pick the “right” font.
Figure 1: A summary of the movie *Yesterday* written in the Times New Roman font. This is a half-size reproduction of the actual stimulus used in the experiment. Times New Roman is a serif typeface designed in 1931 for the *London Times* newspaper. It is a very common font and most people are familiar with it and find it ordinary. It is one of the nine web-safe fonts, because practically every computer has it.

Figure 2: A summary of the movie *Day Break* written in the Blackmoor LET font. This is a blackletter font, a kind of lettering originally designed to reproduce monks’ handwriting. Though blackletter was once common (e.g. the Gutenberg Bible), today it is rarely used, except by Heavy Metal bands, and most people find it ugly and unfamiliar.

Fonts have often been studied in terms of reading efficiency (reading time divided by accuracy). Results show no significant effects for font size or typeface (Bernard et al. 2002). However there are obvious aesthetic differences. Studies have compared observers reading text with good or bad typography, in general, or a good or bad font, in particular. Those who read
material in a good font performed better on cognitive tasks (Larson et al. 2005). Font can improve the efficiency and effectiveness of task performance (Bernard et al. 2002). Task abandonment and erroneous response are correlated with aesthetic preference. Cawthon et al. (2007) concluded that if the user likes an object, then our brains are encouraged to think creatively in order to solve any problem in which the object might be present. Reading preferences are usually linked to what we are used to and what we expect something to be. The font Georgia was more attractive than Arial, Courier, and Comic Sans, while Times New Roman was more attractive than Courier. It is possible that Georgia and Times New Roman were attractive because of their widespread use in both print and on computer screens (Bernard, et al. 2002).

People might believe they can discount other subjective factors from their judgment. Some art critics may claim that they can judge art independent of the space it is shown in. However the rendered judgment is more likely a result of a combination of internal and external factors. When user are surfing the web, they spend less than 45 seconds on each page. During this time, they decide whether to look to further into the website or move on.

Here we ask whether its font affects our ability to judge the content of the text. This can provide insight into how we combine factual knowledge with our own visual perceptions to make a judgment. Typography should be invisible. The mental eye should focus through type and not upon it. Initially typography must often draw attention to itself in order to draw attention to the text. But, it must relinquish the attention it has drawn. Warde (1927) argues that "if a typeface through any arbitrary distorting of design or color, gets in the way of the mental picture that is to be conveyed then it is fundamentally a bad type." Braley (1915) expressed this popular principle in rhyme: Type should be “simple, legible, quiet, plain, A joy alike to the eye and brain.” Typographers firmly believe this to the present day (e.g. Lupton,
Finding that our ability to judge content depends on the font we read it in would reveal an important practical consequence of typography. In our study we ask teenagers to rate movies based on reading movie summaries in two different fonts. Judging movie summaries is important component of the social life of our observers, all students in public high school. A third of all movie-goers are teenagers. Teenagers are exposed to movie summaries almost every day and they usually judge them before they go see the movie.

**METHODS**

*Participants:* Fourteen high school observers (five males and nine females) with normal or corrected-to-normal vision participated in the experiment. All were high school students aged 16-17 years old. All observers were fluent in English. The observers read and signed consent forms but were not made aware of the purpose of the experiment.

*Stimuli:* The movie summaries were typeset in Microsoft Word 2007. Thirty movie summaries were obtained from the internet (blockbuster.com). Movies with less than 2.5 stars were picked in order to avoid popular movies that the observers would be likely to have seen already. The summaries were approximately 150 words each (mean=148, standard deviation=35). The names of actors and the director were removed from the summaries. Two fonts, Times New Roman and Blackmoor LET, both in size 17 points were used. Each summary was printed on an 8.5 by 11 page, one for each font. We created two sets of stimuli. In the first set, 15 randomly chosen movies were in the font Blackmoor LET and the remaining 15 in Times New Roman. The second set used the other font for each movie summary. Within each set, we ordered the printed summaries so that the font alternated from page to page. Seven
observers were randomly chosen to read the first set; and the other seven observers read the second set. The average rating (waiting time) for each movie was calculated and a paired Student’s t-test was used to assess statistical significance.

Task: The printed summary was placed at a comfortable viewing distance, approximately 40 centimeters away. Summaries were randomly ordered, but were in the same order (ignoring font) for all observers. Each observer was asked to read the summary at his or her own pace. Most observers finished each summary within three minutes. After reading, they were asked “How long would you wait in line to see this movie, from zero to thirty minutes?” They were then asked if they had seen the movie before. None of the observers had seen any of the movies before. Finally observers were asked to rate how much they liked each font on a scale ranging from −10 to +10.

Second trial: Two weeks later, all observers were asked to take a second test, identical to the first, though they were merely told that it would be “similar”. Ten of the 14 observers complied. The differences in the ratings of each movie were taken between the first and second trial. The standard deviation of the movie rating differences was calculated for each fonts.
RESULTS

Font ratings: The observers were asked much they liked each font on a scale from -10 to 10. The average rating for Blackmoor LET was 2.1 (standard error = 1.4) and for Times New Roman 7.3 (standard error = 0.9). This documents the ugliness of Blackmoor LET for our observers.

First trial: We collected movie ratings based on reading of summaries printed in either Blackmoor LET or Times New Roman. See Fig. 4. On average, the ratings indicate that our observers were willing to wait 12 minutes for movies whose summaries were printed in Time Roman, but only 8 minutes for movies whose summaries were printed in Blackmoor LET. A paired Student’s t-test shows that this is a statistically significant effect, $t = 3.4$, df = 29, p-value = 0.002.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Blackmoor LET</th>
<th>Times New Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>8.4</td>
<td>11.8</td>
</tr>
<tr>
<td>Standard error</td>
<td>1.1</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 1: Average and standard error of the movie ratings (waiting time in minutes), for each font.

Figure 3: Average movie ratings based on movie summaries viewed in each font. The error bars indicate plus or minus one standard error. Based on 14 observers.
Figure 4: Average rating for each movie averaged across fourteen observers. 25 of the 30 movies were more highly rated when viewed in Times New Roman than in Blackmoor LET.

Second trial: The second trial, identical to the first, but two weeks later, assessed the test-retest reliability of the movie ratings. We computed the differences between test and retest (the two trials), for each summary and observer, and computed the overall standard deviation for each of
the two fonts. The two fonts, Times New Roman and Blackmoor LET, yielded practically the same standard deviation: 6 minutes. This indicates that the reliability of the movie ratings was unaffected by font.

<table>
<thead>
<tr>
<th></th>
<th>Blackmoor LET</th>
<th>Times New Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation</td>
<td>6.0</td>
<td>5.9</td>
</tr>
</tbody>
</table>

**Table 2**: Standard deviation of the differences in ratings between the first and second trial. Based on 10 observers.

**DISCUSSION**

The goal of the study was to see whether the font affected the judgment of content. We asked observers how long they would wait in line to see a movie, based on reading the movie’s summary in one of two fonts, one ordinary and the other ugly. The results from the first trial of this experiment show that observers’ movie rating (expressed as waiting time) does depend on the font in which they read the movie summary. Of course, the ratings varied between movies and among observers. This is expected as each movie summary has its own storyline and as each observer has their own personal taste in movies. For 25 of the 30 movies, observers ratings, on average, indicate that they would wait on line longer for the movie if they read its summary in the ordinary font rather than the ugly font. People might believe they can easily ignore nominally irrelevant subjective factors but our result shows otherwise.

The two fonts chosen for this experiment are very different: one ordinary and familiar, the other unfamiliar and ugly. The results are interesting because the movie ratings logically should ignore the font, which is irrelevant to the movie itself. This shows that our observers were unable to ignore an obviously irrelevant aspect of the printed summary when rating the movie itself. This is exciting. Many past scientific studies of reading have failed to find any large effect of good vs. poor typography on basic performance measures such as reading speed,
comprehension and even individual preferences. Many elements of typography that are deemed important by typographers go unnoticed by observers. Previous work has shown that observers do not usually differentiate between other elements that typographers consider to be important, such as kerning (adjusting the amount of space between characters), leading (spacing between line), and boldness (Larson et al. 2005).

Figure 5: The right-hand sample paragraph includes ligatures, kerning, small caps, old style numerals, and sub/superscript features. There were no reading speed, comprehension, or preference differences between these two conditions (Larson et al. 2005).

When these factors are pointed out, observers then agree that they are better. However typefaces are an important element of typography that observers do show concrete aesthetic preferences for. On a scale of -10 to 10, our observers gave a much higher rating of liking to Times New Roman (7) than to Blackmoor LET (2).

The second trial found practically no effect of font on the standard deviation of test and retest. This means that when observers were re-tested the amount of variability between the first and second trials was the same for both fonts. This indicates that while font does affect the average rating, it does not affect the relative ratings. Thus the judgment of content in
unimpaired by the ugly font. We had hypothesized that the ugly font might distract an
observer enough to make them a poor judge of content. This would predict that the standard
deviation of repeated testing would be much higher with the ugly font than with the ordinary
font. However our results show that that the standard deviation is practically the same for the
two fonts: 6.0 for Blackmoor LET and 5.9 for Times New Roman.
CONCLUSION

We measured the ratings of movies based on reading summaries written in two fonts, one ordinary and one ugly. Though logically irrelevant to the movie itself, the font did affect the ratings. The average rating was much lower for movies whose summary was presented in the ugly font. However, test-retest reliability was unaffected, indicating that the font did not affect ability to judge content. Thus, the ugly font reduced the movie rating, but, contrary to conventional typographic wisdom, did not impair judgment of content. In real life, the movie-summary font would affect whether the reader goes to the movies, but would not affect, once there, which movie is chosen.
ACKNOWLEDGEMENTS

My deepest thanks go to Dr. Denis Pelli for welcoming me to his lab and for his advice and extensive help in editing my paper. Many thanks to Nathaniel Blanco and Sarah Rosen and other valued members of the lab who made helpful suggestions during meeting and made this process even more enjoyable. I would also like to thank Mr. Richard Lee for embodying the definition of a true mentor through his weekly weekend meeting. And finally a thank you to all those who participated in my experiment.
REFERENCES


brothers.


Thangaraj, John (2004) "Fascinating fonts: is the power of typography a marketing myth" Bond University. Web. 

