SCIENCE OF TYPOGRAPHY

Essay by Ellen Lupton, "Cold Eye: Big Science," Print magazine, Summer 2003.

Despite heroic efforts to create a critical discourse for design, our field remains ruled, largely, by convention and intuition. Interested in alternative attitudes, I recently set out to examine the scientific literature on typography. From the late nineteenth century to the present, researchers from various fields—psychology, ergonomics, human computer interaction (HCI), and design—have tested typographic efficiency. This research, little known to practicing designers, takes a refreshingly rigorous—though often tedious and ultimately inconclusive—approach to how people respond to written words on page and screen.

What did I learn from slogging through hundreds of pages photocopied or downloaded from journals with titles like Behavior and Information Technology and International Journal of Man-Machine Studies? Both a little and a lot.

Each study isolates and tests certain variables (font style, line length, screen size, etc.). Although rational and scientific, this process is also problematic, as typographic variables interact with each other—a pull on one part of the system has repercussions elsewhere. For example, in 1929 Donald G. Paterson and Miles A. Tinker published an analysis of type sizes—part of a series of studies they launched in pursuit of "the hygiene of reading."1 Texts were set in 6-, 8-, 10-, 12-, and 14-point type. The study emphatically concluded that 10 points is the "optimum size" for efficient reading—a result relevant, however, only for texts set at a particular line length (80 mm), in a particular typeface (not disclosed).

Another study by Paterson and Tinker tested ten different fonts, including traditional, serifed faces as well as the sans serif Kabel Lite, the monospaced American Typewriter, and the densely decorated, neo-medieval Cloister Black.2 Only the last two fonts—Typewriter and Cloister—caused any significant dip in reading speed. The authors' conclusion: "Type faces [sic] in common use are equally legible" (613). Science leaves the designer more or less at sea in terms of font choice.

A 1998 study testing fonts on the screen revealed conflicts between how users performed and what they said they liked. An interdisciplinary team at Carnegie Mellon University compared Times Roman with Georgia, a serif font designed for the screen.3 Although the team found no objective difference, users preferred Georgia, which they judged sharper, more pleasing, and easier to read. A second test compared Georgia with Verdana, a sans serif face designed for on-screen viewing. In this case, users expressed a slight "subjective preference" for Verdana, but they performed better reading Georgia. Once again, the study concludes with no definitive guide.

How is typographic efficiency judged? "Legibility" concerns the ease with which a letter or word can be recognized (as in an eye exam), whereas "readability" describes the ease with which a text can be understood (as in the mental processing of meaningful sentences). Designers often distinguish "legibility" and "readability" as the objective and subjective sides of typographic experience. For scientists, however, readability can be objectively measured, as speed of reading + comprehension. Subjects in most of the studies cited here were asked to read a text and then answer questions. (Speed and comprehension are factored together because faster reading is often achieved at the expense of understanding content.)

The literature on readability includes numerous articles on whether (and why) paper is preferred over screens. In 1987 researchers working for IBM isolated and tested variables that affect text on both screen and page, including image quality, typeface, and line spacing.4 While the team hoped to successfully identify the culprit behind the poor performance of the screen, they discovered something else instead: an interplay of factors seemed to be at work, each variable interacting with others. The screen itself proved not to be the root cause of its own inefficiency; fault lay, instead, in the way text was presented—in short, its design.

In a second paper the IBM team proved that the efficiency difference between page and screen could be erased entirely if the screen were made to more closely resemble the "normal" conditons of print.5 This study presented black, anti-aliased typefaces on a light, high-resolution screen—features that became more or less standard in the 1990s. The IBM research thus established that design conventions evolved for print effectively translate to the realm of the screen.

While such work confirms the commonality of design for page and screen, other research defies some of our most cherished assumptions. Consider the burning typographic questions of line length and the appropriate number of characters per line. The Swiss modernists have long promoted short, neat lines as ideal for reading, from Josef Müller-Brockman (seven words per line) to Ruedi Rüegg (forty to sixty characters). Such rules of thumb have become basic instinct for many designers.

Science, however, tells a different tale. One study determined that long line lengths are more efficient than shorter ones, concluding that columns of text should fill up as much screen real estate as possible.6 (Grotesque images swim to mind of marginless, unstructured pages of HTML, expanding to fill the screen with one fat column.)

Another study compared texts with 80 characters per line to texts with 40 characters per line. The 80-character lines were created—get this!—by collapsing the width of each letter, thus jamming more text into the same space.7 Despite this unforgivable crime against typography, the study found that subjects could read the denser lines more efficiently than lines with fewer—albeit normally proportioned—characters. Ugliness, we learn, does not always compromise function.

Upsetting assumptions is not a bad thing. Although the research cited here may not tell us exactly how to set type, its conclusions could be useful in other ways. For example, it was once progressive to promote the use of "white space" in all things typographic. Perhaps it is time to reconsider the value of density, from page to screen to urban environment. Down with sprawl, down with vast distances from a to b, and up with greater richness, diversity, and compactness among information and ideas, people and places.

What we might expect from the science of type is a seamless web of rules. Such is not forthcoming. In its drive to uncover fixed standards, the research has affirmed, instead, human tolerance for typographic variation and the elasticity of the typographic system. Science can help ruffle our dogmas and create a clearer view of how variables interact to create living, breathing—and, yes, readable—typography.

Notes

1. D. G. Paterson and M. A. Tinker, "Studies of Typographical Factors Influencing Speed of Reading: II. Size of Type," Journal of Applied Psychology, 13, 2 (1929): 120–30.

2. D. G. Paterson and M. A. Tinker, "Studies of Typographical Factors Influencing Speed of Reading: X. Style of Type Face," Journal of Applied Psychology, 16, 6 (1932): 605–613.

3. Daniel Boyarski, Christine Neuwirth, Jodi Forlizzi, and Susan Harkness Regli, "A Study of Fonts Designed for Screen Display," CHI 98, 18–23 (April 1998). Not paginated.

4. John D. Gould, Lizette Alfaro, Vincent Barnes, Rich Finn, Nancy Gischkowsky, and Angelo Minuto, "Reading is Slower from CRT Displays than from Paper: Attempts to Isolate a Single-Variable Explanation," Human Factors, 29, 3 (1987): 269–299.

5. John D. Gould, Lizette Alfaro, Rich Finn, Brian Haupt, and Angelo Minuto, "Reading from CRT Displays Can Be as Fast as Reading from Paper," Human Factors, 29, 5 (1987): 497–517.

6. Robert L. Duchnicky and Paul A. Kolers, "Readability of Text Scrolled on Visual Display Terminals as a Function of Window Size," Human Factors, 25, 6 (1983): 683–692.

7. Study by Kolers et al, cited in Carol Bergfeld Mills, and Linda J. Weldon. "Reading Text from Computer Screens," ACM Computing Surveys, 19, 4 (December 1987): 329–358.