A quantitative framework for analyzing distinctive features of typography

Christian Mosbæk Johannessen
University of Southern Denmark

Mads Qwist
Copenhagen Business School

A quantitative framework for analyzing distinctive features of typography

1 Introduction

Being able to describe typography accurately is relevant in as diverse areas of study as marketing and branding, education and literacy studies in order to understand, for example, how typography facilitates reading (Veytsman & Akhmadeeva, 2011), how it conveys personality (Ahmed, 2013), identity messages and style (Van Leeuwen, 2005a) or how it builds brand value (Hyndman, 2016). As pointed out by Bateman in this journal, many such fields attempt to evaluate the design and use of typography by correlating design differences with recipient experience, performance or behavior through various empirical and experimental methods (2019: 255). However, as Bateman also points out, in designing such empirical and experimental studies the best approach is to create systematically varying and appropriately contrasting stimuli. If we do not, we simply cannot know which differences in design make which differences in experience, performance or behavior. This has to be based on a description of the form level of typography that is both detailed and operational, and although such schemes for typography have been proposed (Van Leeuwen, 2006), their operationalization for empirical analysis has been remarkably difficult. To give just a single example, in their discussion of a recent study comparing the legibility of Gill Sans with a newly designed font, KBH, commissioned by the municipality of Copenhagen, Beier and Oderkerk (2019) reflect precisely over this matter:

> When fonts of different families […] are compared, the fonts vary on so many variables that the chances of finding a significant difference between their effects on performance is much greater than when we isolate just one variable. It is, however, difficult to tell how the variables are interrelated. On the other hand, isolating a single variable, such as when comparing KBH Text and KBH Display […], should, in theory, enable us to identify the specific feature of the fonts that causes the difference in performance. (Beier & Oderkerk, 2019: 62)

And this is a best-case scenario: Beier and Oderkirk compare only *sans serif* fonts, Gill Sans, KBH Text and KBH Display, which tend to have fairly uniform stroke widths. Imagine how much more complex their task would have been had it involved a *sans serif*, a *serif* and a *script* font? Serif fonts tend to have variability of width *between* the strokes in a letterform (what we will refer to as *contrast* with reference to Stötzner, 2003), and script fonts are even worse. They also have a high degree of variability *within* each stroke (which we, again with reference to Stötzner, refer to as *tension*). Such complex interactions between variables would have to be tickled out before we can be confident about which variable produces which performance.

Moreover, if we could agree on a description of typographic form it would be easier to compare results across disciplines, for example letting insights from branding inform our understanding of reading. Thus, our aim with this contribution is to present the kind of detailed and analytically operational account of the form level of micro-typography, which Bateman calls for.[1] A further aim is to demonstrate our procedure, on a diverse *n = 147* sample of unique fonts, to test if it is sensitive and reliable enough to pick up and quantify systematically occurring patterns of variation between *sans serif*, *serif*, *slab serif*, *handwritten* and *script* fonts.

2 Theoretical background

As discussed by Eva Brumberger (2003: 206), the rules of typographic practice are largely a kind of practitioners' lore "supported by intuition but lacking a theoretical and empirical foundation". Until recently, the typographic profession had an almost arcane, artistic aura, which came across in for example Robert Bringhurst's authoritative *Elements of Typographic Style* (1997). Here, he warned against uninitiated dabbling:

> Type design is an *art* practiced by few and mastered by fewer—but font editing software makes it possible for anyone to alter in a moment the widths and shapes of letters to which *an artist* may have devoted decades of study, years of inspiration, and a rare concentration of skill. The power to destroy such a type designer's work should be used with caution. (Bringhurst, 1997: 35, our emphasis)

In other words, until just a few decades ago the profession had no vocabulary for discussing the meaning potential of typography. In the words of Jonathan Hoefler, in Gary Hustwit's documentary *Helvetica*, "[…] typography has this real poverty of terms to describe things beyond x-height and cap height and weight and so on". (2007: 00.30.12)

The film was part of a surge in popular interest in typography shared by multimodal semioticians. Spurred by the ubiquity of user-friendly text editing and desktop publishing software, semioticians wondered what the meaning potential of typography might be and how it could be approached as a common means of communication (Stöckl, 2005; Stötzner, 2003; Van Leeuwen, 2005b, 2006). Of these contributions Theo van Leeuwen's (2006) Halle and Jakobson-inspired *distinctive feature* approach has had the more enduring impact in multimodal studies, especially in social semiotics (Armean, 2016; Nørgaard, 2009; Pantaleo, 2014; Rajabi, 2020; Ravelli & Starfield, 2008).

Van Leeuwen's thesis is that typographic form can be construed as a mix of formal distinctive features; weight, horizontal expansion, slope, curvature, connectivity, vertical orientation and regularity, which combine in appeals to our experiential basis[2] to form "connotational" meanings or "experiential metaphors" (Van Leeuwen, 2006: 147). For example, through experiential analogy, one might associate fonts of great weight, expansion and angularity with a certain unyieldingness, gravity and sharpness. The values thus connoted might be more appropriate in communicative contexts characterized by heteronormative masculinity, as for example corporate identities for manufacturers of heavy-duty construction machinery (e.g., Caterpillar, Kato, Komatsu, Liebherr, and Yanmar).

[1] The term micro-typography is discussed by Hartmut Stöckl (2005) and encompasses the structure and proportions of letterforms. It is distinct from meso-typography, which pertains to the structure of space between letterforms, words and lines, and macro-typography which pertains to text blocks and page lay-out.
[2] Van Leeuwen draws inspiration here from Lakoff and Johnson's orientational metaphors (1980: 14-15).

Van Leeuwen emphasizes that his "first attempt at identifying distinctive features" does not amount to a list of "*the* 'authoritative' meanings of letterforms" (Van Leeuwen, 2006: 147). Indeed, in our operationalization we have dropped regularity because any structural feature can be distributed with varying degrees of evenness, making 'regularity' a meta-feature that, for all its primacy in perception, remains operationally elusive. Furthermore, we have included features from Andreas Stötzner's suggestions for formal "signographic" description (2003: 291), specifically tension and contrast. We expect these features to be strong predictors of stroke dynamics, and so to be important for distinguishing fonts that emulate handwriting.

Although van Leeuwen's model has been criticized by Shane Morrissy for overlooking many aspects of experience that could animate the connotative and metaphorical potential of distinctive features in typography (Morrissy, 2017), we agree that the features themselves remain a "useful framework for examining common characteristics across a range of distinct typographical resources" (2017: 184). The extent to which they correspond with people's experience of typography, and how they correlate with our experiential basis to make meaning, is precisely the kind of empirical and experimental questions we hope our procedure can help with in the future.

Based on these considerations, we aim to answer the following questions:

1. How can we operationalize and quantify a formal distinctive feature framework in order to make comparisons across many instances and many classes of typography possible?
2. To which extent can a distinctive feature approach to typography, as a minimum requirement for usefulness, pick up differences between commonly accepted typographic classes such as *sans serif*, *serif*, *slab serif*, *handwritten*, *script*, etc.?
3. If so, what can we learn from it about how such distinctive features, or variables, interact?

3 Procedure

3.1 Data collection

In order to test the procedure, we collected a sample of fonts, *n*=147, from Monotype's online font retail site MyFonts (MyFonts, 2020). The site currently offers customers a total of more than 130,000 fonts in 6 categories: sans serif, slab serif, serif, display, handwritten, and script. We dropped the 'display' category because it transcends all the other five categories. From each of the remaining five categories we collected a gross sample of the 33 best-selling fonts but dropped duplets and erroneously categorized individuals (see figure 1 for a small selection). The net sample was distributed thus: sans serif (33), slab serif (33), serif (33), handwritten (21), script (28). Beyond these steps we did nothing to further sub-classify the sample according to e.g., 'regular', 'italic' or bold' fonts or according to functional considerations such as whether fonts were designed with running text or headlines in mind. The sample appears in the way the fonts' designers chose during upload to the site.

We used the so-called "pangram" feature on MyFonts, which allows users to type their own character string and assess its appearance in different fonts, to have the letters 'H', 'p', 'k', 'x', 'o' and 'a' rendered in each font. These characters were chosen to get a reasonable mix of upper and lower case letters and anatomical features (see Lupton, 2014: 36). We zoomed the maximized Google Chrome browser window (2560 x 1600 pixels) until the pangram filled it and took a screen shot. We stored these as PNG files and opened them in Adobe Photoshop for analysis with the measure and count tools.

Figure 1: A small selection of the images, five from each class, of fonts we collected from MyFonts.com. Even judging by a first, casual, visual inspection it is clear that handwritten and script fonts show much greater variability in their deployment of distinctive features than the other classes. The category with the least variance seems to be sans serif.

3.2 Measurements

With few exceptions[3] all measurements take *x-height* (the distance between the base line and median of the lower-case letter ‘x’ in a font, measured at an orthogonal to the base line) as a reference of scale in order to make our quantification of typographical features independent of the actual size or image resolution of a specimen[4]. Thus, we arrive at values that are proportional to the x-height. Because measuring takes quite some time, our trial application makes a few necessary compromises: We assume that, as a general rule, the look of all letters in a font follows the same overall pattern. If the letter ‘o’ in a font is narrow, all letters are equally narrow. If the ascender on ‘k’ rises high above the x-height, we assume the same of all other ascenders. Of course, this loses potentially important details but, if one’s research questions require a finer grain, making more measurements is straightforward. We take a total of 30 measures of each font. The first four are taken to characterize overall letter proportions, as shown in figure 2A. They are (1) *x-height*, (2) *v-height* (v for vertical), (3) *o-width* (the width of ‘o’ at its widest point) and (4) *slope angle* (the angle between the stem of ‘k’ and the baseline). The next 26 measures capture the more detailed appearance of strokes, e.g., H1, H2 etc. For each of the 13 strokes in the six letters, we take an orthogonal measure (or radial in the case of curves) at the widest (*w*) and narrowest (*n*) points, as shown in figure 2B.

[3] The exceptions are measures of slope (which outputs a value in degrees), counts of shape features (which takes regions as its unit of analysis, not strokes) and calculations of tension.

[4] In many cases of especially handwritten and script fonts we have to rely on judgment to establish best fit for baseline and median (see figure 2A)

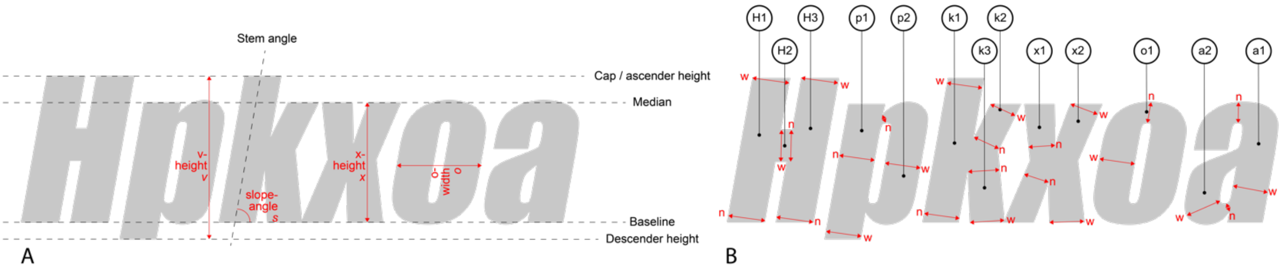

Figure 2. A: Illustration of the 4 measures taken to characterize overall letter proportions, orientation, expansion and slope. B: Illustration of the 26 measures taken to capture more detailed features of strokes: weight, tension and contrast.

3.3 Deriving values for distinctive features

3.3.1 Weight

Typographic weight is generally understood as "[…] the boldness of the strokes in relation to the size of the sign." (Stötzner, 2003: 291) Our operationalization of weight uses Johannessen's (2011: 244) *weight scale rating* (WSR), which relates the width of strokes to a font's x-height and outputs an average value between zero (a stroke of no width) and one (a stroke width equal to the font's x-height). In figure 3A, which illustrates the procedure on the stroke in 'o', the font's x-height is 384 pixels. The stroke's widest point (*w*) is 113 pixels and the narrowest is 70 pixels. These relations are averaged over the number of observations (one for *w,* one for *n*).

$$\frac{\left(\frac{w}{x}+\frac{n}{x}\right)}{obs} \rightarrow \frac{\left(\frac{113}{384}+\frac{70}{384}\right)}{2} = 0.24$$

For this particular stroke, WSR is 0.24. We repeat this procedure on all thirteen strokes in each specimen and average over the result. Average values above 0.5 are unlikely to occur in typography as such broad strokes would suffocate the apertures (loopholes) in letters such as 'p', 'o' and 'a'.

3.3.2 Tension

"Tension" refers to "[…] the dynamics caused by increase and decrease of a drawn stroke which in writing occurs rather spontaneously and contributes considerably to a sign's lively shape" (Stötzner, 2003: 291). Our operationalization picks up both such dynamical variations as well as *entasis* (convexity or concavity used for optical adjustment of straightness). In order to calculate a stroke's tension ratio, we relate the widest measure (*w*) to the narrowest measure (*n*) as shown in figure 3B.

$$\frac{w}{n} \rightarrow \frac{113}{70} = 1.61$$

For the stroke in 'o', the widest part (113 pixels) is 1.61 times thicker than the narrowest part (70 pixels). We repeat this for all thirteen strokes in each specimen and average over the results.

Figure 3: Illustration of the particular measures taken in the example to calculate distinctive features.

3.3.3 Contrast

Whereas "tension" refers to the variability of weight *within* a stroke, "contrast" refers to "a more or less pronounced gradation *between* thin and thick strokes". (Stötzner, 2003: 291) In order to calculate a contrast ratio, we identify the strokes with the highest (WSR_w) and lowest (WSR_n) weight scale ratings for each letter (using previously calculated WSRs, see section 3.1.1) and find their proportions.

$$\frac{WSR_w}{WSR_n} \rightarrow \frac{0.312}{0.27} = 1.15$$

In the example shown as figure 3C, the widest stroke in 'H' is H1 (WSR_w = 0.312). The narrowest stroke is H2 (WSR_n = 0.27). In other words, for the letter 'H' in this font, WSR_w is 1.15 times greater than WSR_n. We repeat this for all letters except 'o' (which has only a single stroke and so no contrast) and average over the results.

.

3.3.4 Orientation

Our operationalization of orientation and expansion are inspired by van Leeuwen's suggestions: "Typefaces may be either oriented towards the horizontal dimension, by being comparatively 'flattened' […] or towards the vertical dimension by being comparatively elongated, stretched in the vertical direction" (Van Leeuwen, 2006: 149). Because van Leeuwen's suggestions are open to interpretation, we have decided to let our characterization capture the relative stubbiness of the font's ascenders and descenders by relating the distance between ascender and descender height (v) to the x-height (x).

$$\frac{v}{x} \rightarrow \frac{522}{384} = 1.36$$

In our example, illustrated in figure 3D, the x-height is again 284 pixels, and the vertical height of the font, v, is 522 pixels. Thus, the vertical height is 1.36 times greater than the x-height.

3.3.5 Expansion

Van Leeuwen describes expansion thus: "Typefaces may be condensed, narrow, or they may be expanded, wide" (2006: 148). We choose to understand expansion similarly to what typographers call set width (INSERT LUPTON CITATION). On the assumption that all letters are of similar expansion, we calculate it as a proportion between the width of 'o' (o) and the font's x-height (x).

$$\frac{o}{x} \rightarrow \frac{266}{384} = 0.69$$

In the example, illustrated in figure 3E, the o-width is 266 pixels, and the x-height is 384 pixels, which means that 'o' is 0.69 times wider than 'x' is tall. Values greater than one denotes letters that are wider than the x-height.

3.3.6 Slope

Measuring slope is straightforward using the measure tool in Photoshop. We simply measure the angle between the stem of 'k' and the base line, as illustrated in figure 3F. A value of 90˚ means that the stem is orthogonal to the baseline. Values below 90˚ denote letters that slope to the left, values above 90˚ denote a slope to the right. The latter is more likely to occur because the ergonomics of righthanded writers favor gestures resulting in right-sloping letters

3.3.7 Connectivity

Our operationalization of connectivity simply relates the number of connections (c) between letters in the font sample to the number of letter spaces (s), illustrated in figure 3G. A connectivity value of 0 means no letters are connected and 1 means all letters are connected. Our example has a connectivity of 0.0 because it has 5 spaces and no connections.

$$\frac{c}{s} \rightarrow \frac{0}{5} = 0.0$$

3.3.8 Curvature

According to van Leeuwen, "[…] a letterform can stress angularity or it can stress curvature" (2006: 149). He discusses how loops, the shapes of ascenders and descenders as well as details of serifs etc. all contribute to a more or less rounded appearance.

Our operationalization of curvature draws on Johannessen's (2016) analysis of shape in corporate logos. Whereas, so far, we have taken staples of typographic terminology*, strokes* and *letters*, as the point of departure for our operationalization, our procedure for describing a typography's 'shape' calls for a different outlook. Whether thick or thin, a stroke or letter always occupies some region defined by a distinctly shaped border, which contains the territory inside it. Thus, the basic unit for our shape analysis is a two-dimensional *region*. From this perspective, we consider *apertures*, the loops inside some letters such as 'b', 'd', 'e', 'g' etc., as regions in their own right and treat them separately. In the example, there are 9 individual regions as depicted in figure 4.

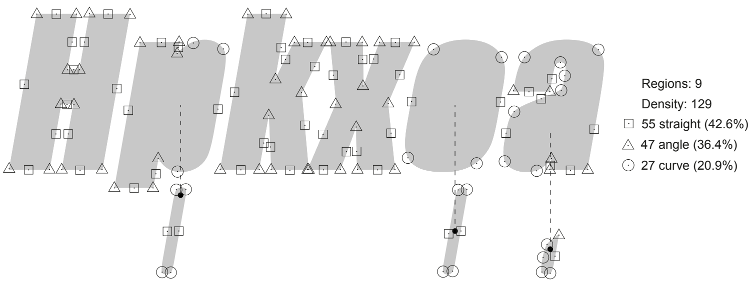

Figure 4: Illustration of the 9 regions in the example. Notice that the apertures in 'p', 'o' and 'a' have shapes that are analyzed separately.

We count all instances of three types of shape—straights, angles and curves—for all regions in the font specimens. Our approach is inspired by how shapes are described in vector graphics and is illustrated in figure 9. We count an occurrence of straight, angle or curve when we would expect them in a Bezier-based vector description. The actual annotation is done with the count tool in Photoshop. In the example, there are 55 instances of straight features, 47 instances of angles and 27 instances of curves.[5] In total, they amount to a *shape density* of 129 shape features, which is an important metric. In order to characterize the way shape is distributed in the samples, we derive the proportions of types of shape in relation to the shape density and call these metrics S/D (the proportion of straights to density), A/D and C/D. In the example, S/D = 42.6, A/D = 36.4, and C/D = 20.9. A final measure is the average shape density per region, D/r. This is an interesting measure because, for example, it will indicate whether letterforms have serifs or not, even if it does not explicitly classify them accordingly. It will also pick up whether the average shape of letterforms in a font is clean, as in many sans serif fonts, or distressed, as in many handwritten fonts.

In combination, our operationalization of curvature yields four variables, S/D, A/D, C/D and D/r, instead of one.

[5] It should be noted that, whereas annotating occurrences of straight and angle is straightforward, curves are harder to distinguish. Especially in cases when one curve transitions smoothly into another, we rely on judgement.

3.4 Principal Component Analysis (PCA)

After measuring the typefaces in our sample and calculating how they score on the above 11 variables, there are several avenues of statistical exploration we could pursue. Most straightforwardly we could simply compare a single variable at a time. This might in itself be revealing. However, we choose to pursue a fuller picture of the data and use a popular and well-understood multivariate statistics technique called *principal component analysis* (PCA) to explore how all 11 variables interact (see Abdi & Williams, 2010 for an introduction).

The aims of PCA are (1) to extract the most important information from a data set, (2) compress the dataset by keeping only this important information, (3) simplify the description of the dataset and (4) analyze the structure of the observations and the variables (Abdi & Williams, 2010: 434). The PCA technique does this by computing the extent to which our original 11 variables exhibit similar patterns of variation and then, on that basis, grouping them into new variables called *principal components*. In our case, PCA creates 11 principal components by mixing the information in the old ones by virtue of the strength of their correlations. When we sort these new variables according to how much of the variance in the data they contain, we can throw away the less useful ones and be confident that the ones we keep still map the structure in the original data. For this reason, the convention in PCA is to report the percentage of variance explained by the principal components. We aim to keep three dimensions, as this number can be conveniently visualized.

One drawback of PCA is the abstractness of its output, which makes the original variables in the data, for example “weight” in our typography data (which sems intuitive and corresponds well with a straightforward aspect of experience), much harder to interpret. This is because the PCA procedure mixes information about all the old variables into the new principle components. In order to help interpret PCA analyses, the new principal components are commonly inspected using a type of graph known as a “bi-plot”. Bi-plots mix simple scatterplots of the individual observations, for example with principal component one (PC1) along the x-axis and PC2 along the y-axis, with so-called “loading vectors”, which make the plot easier to interpret. Loadings are estimates of the information shared by an original variable and a principle component (Abdi & Williams, 2010: 438). These are illustrated in the plot by arrows that indicate the vector of the original variable in the plot (the arrow’s direction) and the importance of the variable (the arrow’s length) for that principal component. Thus, the longer a loading arrow, the more strongly that original variable is expressed in the plot. Bi-plots are important epistemological vehicles and, in section 4, we use them to visualize our analysis.

PCA is gaining traction as a technique for exploratory analysis in multimodal studies precisely because it can reveal patterns in larger samples of very complex multimodally constituted texts which might otherwise escape attention. Recently, PCA has been used in a diachronic study of page layout in comics (Bateman, Veloso, & Lau, 2019), which reduces 52 dimensions of variation to just five principle components that nevertheless capture 61% of the variation in the original data, and also a synchronic study of graphic style in corporate logos (Johannessen, Tvede, Claussen-Boesen, & Hiippala, in press), which reduces 14 dimensions of variation to five. These cover 76.4% variation.

4 Results

We begin with a summary of mean values in table 1 of all distinctive features across the five classes. To the extent our sample is representative of fonts in common usage, these can be thought of as a baseline for comparison to see to which extent a font deviates from a given class norm.

Class	Weight	Tension	Contrast	Orientation	Expansion	Slope	Conn'y	D/r	S/D	A/D	C/D
Sans Serif	0.157	1.092	1.071	1.826	0.950	90	0.000	12.607	**0.406**	**0.396**	0.196
Serif	0.149	2.251	**2.227**	2.041	1.006	90	0.042	**25.680**	0.297	0.299	0.403
Slab Serif	**0.194**	1.236	1.189	1.834	**1.022**	90	0.000	23.985	0.398	0.376	0.225
Handwritten	0.132	2.149	1.250	2.317	0.570	**109**	0.545	19.249	0.149	0.170	0.680
Script	0.134	***2.587***	1.345	**2.440**	0.574	106.4	**0.605**	19.347	0.085	0.128	**0.786**

Table 1: Overview by class of mean values for all 10 measured variables. Highest average for each value is emboldened.

These averages already suggest that our quantification of what was apparent during visual comparison of the fonts in figure 1 is reasonable: That sans serif fonts are characteristic by their absence of relative extremes, except in the case of *straight* and *angle*, slab serif fonts by their *weight* and *expansion*, serif fonts by their stroke *contrast*, and script fonts by their stroke *tension*, *orientation*, *connectivity* and *roundedness.* However, these averages say nothing about the variability exhibited in each class and only very little about how variables interact, which is why we performed principal component analysis (with normalization) to explore our data further. As illustrated in the scree plot in figure 5, the first five out of eleven principal components explain 84.43% of the variance in the dataset. We dismiss the remaining six as noise. The plot also reveals that 45.1% of the variance is explained by the first principal component alone and that the first three together explain 68.9% of the total variance in the dataset. Therefore, we find it reasonable to focus on the first three components for the exploratory analysis further below.

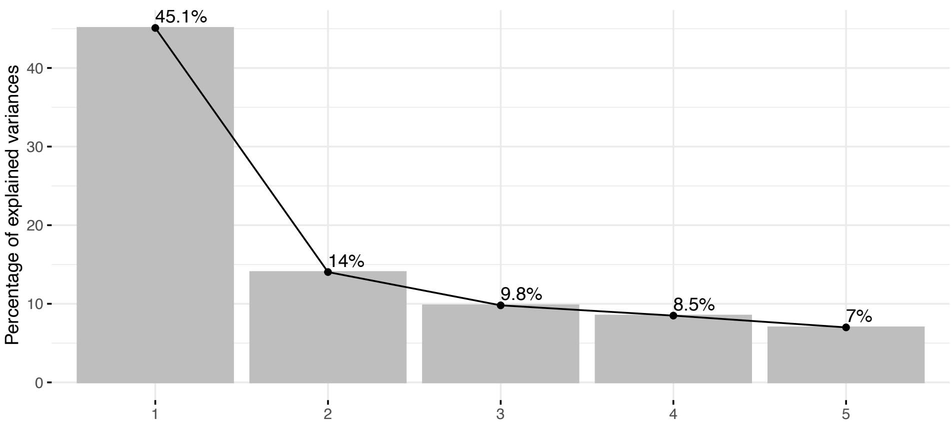

Figure 5. Percentage of variance explained (y-axis) by the principal components (x-axis). Total variance covered by the first five components: 84.43%.

Table 2 reports PCA loadings, i.e. how each of the original distinctive features correlate with the new principal components. The values run from -1 for perfect negative correlation over 0 for no correlation to +1 for perfect positive correlation.

	PC1	PC2	PC3	PC4	PC5
Slope	*-0.776*	*-0.097*	0.246	**0.213**	-0.013
Weight	0.430	0.051	0.176	**0.828**	**0.171**
Tension	-0.377	**0.651**	*-0.268*	**0.226**	**0.169**
Expansion	**0.811**	0.240	0.124	0.113	0.077
VertOrientation	-0.424	0.008	**0.521**	*-0.261*	**0.682**
Connectivity	*-0.816*	*-0.054*	**0.259**	0.103	*-0.101*
Contrast	0.016	**0.876**	*-0.200*	*-0.183*	0.109
D/r	0.040	**0.522**	**0.691**	-0.075	*-0.458*
S/D	**0.946**	*-0.051*	0.101	*-0.107*	0.055
A/D	**0.865**	-0.037	0.130	-0.047	0.044
C/D	*-0.936*	0.045	*-0.119*	0.079	*-0.051*

Table 2. Correlations between the original 11 variables and the new principal components. The top three positive and negative correlates for each component are marked in bold and italic, respectively.

The first principal component, PC1, reveals an interesting pattern of correlations. First of all, it picks up almost all the information about how shape is proportionally distributed in the fonts. Notice how strongly *straightness* (S/D = 0.946) and *angularity* (A/D = 0.865) are correlated. This suggests that those two features co-vary in the fonts. Notice also that *curvature* (C/D = -0.936) is strongly negatively correlated with them. This means, unsurprisingly, that fonts cannot simultaneously be characteristically angular and curvy. Secondly, it reveals that *curvature* co-varies with *connectivity* (-0.816) and *slope* (-0.776) and to a lesser degree with vertical *orientation* (-0.424) and stroke *tension* (-0.377). We will discuss the potential impact of this below.

The second principal component, PC2, is strongly associated with *contrast* (0.876) and *tension* (0.651) and also with *shape density* (D/r = 0.522). None of the original variables correlate negatively with PC2 to any particularly interesting degree.

The third principal component, PC3, is more diffuse: The two variables that load the most strongly onto it, *density* of shape (D/r = 0.691) and vertical *orientation* (0.521), cross load strongly onto other components.[6]

As we pointed out in section 3.4, the output of PCA is very abstract. In order to interpret it, figure 6 plots the relative position of the fonts on PC1 (x-axis) against PC2 and PC3, respectively (y-axis). Notice how the loading arrows represent the strength of correlation between an original variable and a principal component. The longer the arrow, the stronger the correlation. As a rule of thumb, the farther an individual font is plotted in the direction of a loading arrow, the higher its value is on the original variable. Class membership is used to color the individual groups and to calculate the 95% confidence ellipses[7]. The first thing to notice is how most fonts fall into three fairly distinct clusters that correspond one-to-one with font classes: Sans serif, slab serif, serif. A fourth large cluster corresponds with both the handwritten and script classes, which suggests that the two classes are in fact a single population and that MyFonts' users do not distinguish as clearly between them when uploading fonts as they do the other classes.

[6] D/r loads strongly (>0.5) onto both PC2 and PC3. Vertical orientation loads onto both PC3 and PC5. Ordinarily, we would take remedial action and remove the offending variables from the data set, but we allow it in this exploratory study. As our plots will reveal, they are important for interpreting the results.

[7] The ellipses predict, with 95% confidence, that a font in a class will fall within that space.

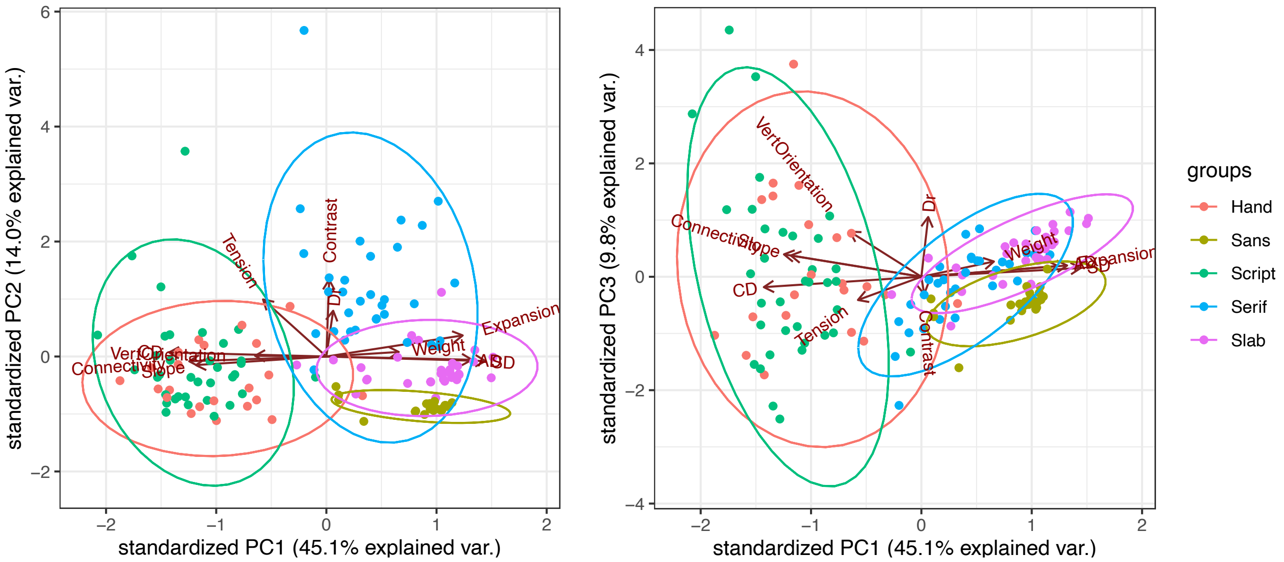

Figure 6. Plotting the principal components against one another. The area covered by the ovals corresponds to two standard deviations, which covers 95% of the data in a normal distribution. The size of the confidence ellipsis, in turn, indicates the variability of the data. Compare these with figure 1 and notice how the sans serifs show very little variability, slabs a little more, serifs more still and the two handwriting fonts the highest degree of variability.

The second thing to notice is the placement of the clusters in relation to the loading arrows. Notice how the handwritten/script fonts cluster on the left side of the plots, in the direction of *tension, vertical orientation, slope, connectivity* and *curvature.* And how all the non-script-like fonts cluster on the right, in the direction of *weight*, *horizontal expansion, straightness and angularity*. A possible interpretation would be that the classes split roughly along the mean on PC1 into fonts that have a 'manual' and 'artificial' appearance, and that narrow, sloping letters with very long ascenders and descenders, made of thin but tension-filled strokes that are connected, seem to be tell-tale signs of the presence of the hand in the font.

A last thing to notice is the angle between the arrows. The more acute the angle (the closer the arrows are together), the stronger the interaction between the original variables. For example, regardless of whether we plot PC1 against PC2 or PC3, *connectivity* and *slope* remain close and D/r remains almost orthogonal to S/D, A/D and C/D, which means D/r is uncorrelated with shape features.

To further explore the possibility that handwritten and script fonts in fact belong to the same population, we also performed *t-distributed stochastic neigbourhood embedding*, or t-SNE (Van der Maaten & Hinton, 2008). Whereas PCA is a linear dimensionality reduction method, t-SNE is non-linear and aims at preserving the relative distances between samples in the original feature space while performing the reduction. In figure 7 below, a pattern similar to that in the PCA analysis is repeated, but with an even clearer separation between the cluster of handwritten and script fonts and the rest. This substantiates our suggestion that font classes can be categorized as either manual or artificial.

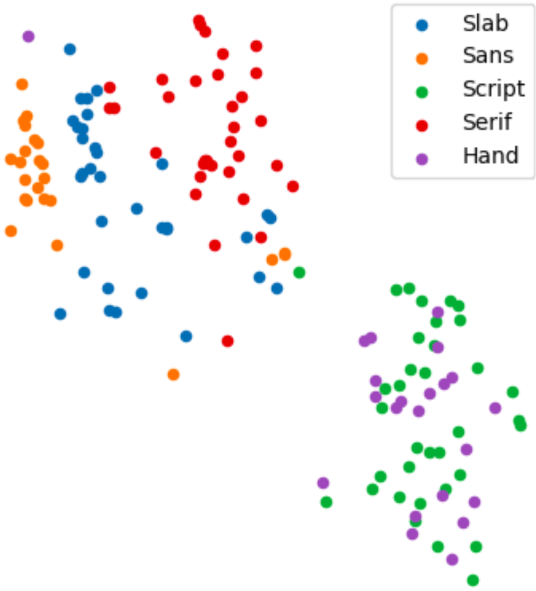

Figure 7. 2D plot of t-SNE as an alternative to PCA. The pattern in the PCA analysis is confirmed, however with a slightly stronger separation between the cluster containing handwritten and script fonts and the rest.

On the basis of our PCA and t-SNE analyses of the data, we hypothesize that the variables *tension, vertical orientation, slope, connectivity* and *curvature* are predictive of what one might call "the presence of the hand" in fonts, in other words that these variables contain information about the relationship between sensory-motor work and typographic trace.

5 Conclusion

We set out to (1) demonstrate how van Leeuwen's distinctive feature approach to typography can be quantified to make measures comparable across many instances and classes of fonts, (2) ask if such an approach can detect differences between common font classes as a test of minimal usefulness and (3) whether it can tell us anything interesting about interactions between distinctive features.

Starting from a measure of a font's x-height as a reference of scale, we have suggested and tested a quantitative approach to a mix of features. The procedure clearly passes our test of minimal usefulness by detecting categories that are well-established in the typographic profession, and we can be fairly confident that distinctive features are descriptively adequate and that our procedure is a valid operationalization of them. Had it failed this basic empirical test, it would obviously call the principles of operationalization into question, but also cause us to be more critical of the notion of distinctive features as a productive framework for describing how fonts look and understanding how they set up complex connotative meaning potentials. We can now begin to use the procedure to ask more interesting questions about the deployment of typographic resources in more specific social contexts.

The most predominant pattern of interactions between variables is one that unmistakably splits the fonts into two groups according to whether they retain traces of hand movements (fonts in the script and handwritten classes) or not (fonts in the sans serif, slab serif and serif classes). The variables *C/D (curvature), vertical orientation, connectivity, slope* and *tension* are very closely correlated and are highly predictive of manual looking fonts. The variables *weight*, *expansion, S/D (straight)* and *A/D (angle)* are very closely correlated and highly predictive of artificial looking fonts. The variables *D/r (density)* and *contrast* are the dimensions that interact the least. *D/*r seems to primarily predict the presence and type of serifs in the artificial fonts. Contrast seems to primarily predict fonts to fall in the serif class.

Such interaction effects will not surprise practitioners of type design and calligraphy. The practice of their craft has made them extremely sensitive to the qualities of strokes and proportions of letterforms. But for the rest of us, it is interesting to see such phenomena thus quantified and visualized so we can inspect their interactions in more detail. Our results in fact encourage further reflection on what patterns might lie beneath the within-class variability of, for example, the large handwritten/script cluster. Might a manual vs. artificial distinction also be systematically at work here, yielding more or less artificial-looking handwritten fonts? The low degree of between-class overlap in combination with high within-class variability suggests we could mine this data further to explore such possibilities.

While we certainly expect our study to be of interest to an audience of typographers and graphic designers, we suggest that it might also be useful outside a strictly typophile community. As we pointed out in the introduction, several fields of research could benefit from a more robust operationalization of font variables, which goes beyond the usual weight and x-height, for purposes of experimental and empirical research design, including education, marketing and branding.

Moreover, it is possible that predominantly qualitative studies of *visual style* (Van Leeuwen, 2005a: 147), *graphic ideology* (Spitzmüller, 2015) and *typographic landscaping* (Järlehed & Jaworski, 2015) might be complemented by the more qualitative approach to typographic description we suggest here. It only requires a small leap of imagination to hypothesize, for example, that typography deployed in social practices dominated by traditional heteronormative gender discourses of masculinity and femininity will cluster in a certain and distinct ways. Of course, this begs the question of how well our procedure will perform outside the highly controlled environment of this study where sampling of letters is homogenous and commensurate. The moment we start looking at typography in the wild, complexity is bound to increase. Specimens will contain different letters; some will be all caps and have no explicit x-height and so on. Such obstacles are impractical and are bound to incur extra work, but we are confident they can be overcome. One way of achieving this would be to automate the procedure using computer vision or deep learning in a way that balances the gains of decreased workload with the losses of accuracy of measurements.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics, 2*(4), 433-459.

Ahmed, M. M. (2013). Typeface persona: Investigating Gotham's suitability for Obama's 2008 presidential campaign. *Unpublished doctoral dissertation). Howard University, Washington, DC. Retrieved from http://search. proquest. com/docview/1512417400*.

Armean, O. (2016). New Media Typography. *Journal of Media Research-Revista de Studii Media, 9*(24), 28-38.

Bateman, J. A. (2019). Information design and multimodality: New possibilities for engagement across theory and practice. *Information Design Journal, 25*(3), 249-257.

Bateman, J. A., Veloso, F. O., & Lau, Y. L. (2019). On the track of visual style: A diachronic study of page composition in comics and its functional motivation. *Visual Communication*, 1470357219839101.

Beier, S., & Oderkerk, C. A. T. (2019). The effect of age and font on reading ability. *Visible Language, 53*(3), 50-68. Retrieved from https://search.proquest.com/docview/2330018843?accountid=14211 http://UL2NB3KN6E.search.serialssolutions.com/directLink?&atitle=The+effect+of+age+and+font+on+reading+ability&author=Beier%2C+Sofie%3BOderkerk%2C+Chiron+A+T&issn=00222224&title=Visible+Language&volume=53&issue=3&date=2019-12-01&spage=50&id=doi:&sid=ProQ_ss&genre=article

Bringhurst, R. (1997). *The elements of typographic style* (2nd ed.). Vancouver: Hartley & Marks.

Brumberger, E. R. (2003). The rhetoric of typography: The persona of typeface and text. *Technical communication, 50*(2), 206-223.

Hustwit, G. (Writer) & G. Hustwit (Director). (2007). Helvetica. In.

Hyndman, S. (2016). *Why fonts matter*: Random House.

Johannessen, C. M. (2011). *The Forensic Analysis of Graphic Trademarks. A multimodal social semiotic approach.* (Ph.d.). University of Southern Denmark, Odense.

Johannessen, C. M. (2016). Experiential meaning potential in the Topaz Energy logo: A framework for graphemic and graphetic analysis of graphic logo design. *Social Semiotics*, 1-20.

Johannessen, C. M., Tvede, M., Claussen-Boesen, K., & Hiippala, T. (in press). A Corpus-Based approach to Colour, Shape and Typography in Logos. In *Emprical Multimodal Research*.

Järlehed, J., & Jaworski, A. (2015). Typographic landscaping: Creativity, ideology, movement. In: Taylor & Francis.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago press.

Lupton, E. (2014). *Thinking with type: A critical guide for designers, writers, editors, & students*. New York, NY: Princeton Architectural Press.

Morrissy, S. (2017). Inexplicable in the last resort. *Information Design Journal, 23*(2), 184-193.

MyFonts. (2020). Display Fonts. Retrieved from https://www.myfonts.com/display-fonts?at=1

Nørgaard, N. (2009). The semiotics of typography in literary texts. A multimodal approach. *Orbis litterarum, 64*(2), 141-160.

Pantaleo, S. (2014). Elementary students consider the “what” and “how” of typography in picturebooks. *New Review of Children's Literature and Librarianship, 20*(2), 144-166.

Rajabi, E. (2020). Typography as a regime of reading/looking in Toni Morrison’s The Bluest Eye (1970). *Cogent Arts & Humanities, 7*(1), 1811580.

Ravelli, L. J., & Starfield, S. (2008). Typography and disciplinary identity in academic writing. *Information Design Journal, 16*(2), 133-147.

Spitzmüller, J. (2015). Graphic variation and graphic ideologies: A metapragmatic approach. *Social Semiotics, 25*(2), 126-141.

Stöckl, H. (2005). Typography: body and dress of a text-a signing mode between language and image. *Visual Communication, 4*(2), 204-214.

Stötzner, A. (2003). Signography as a subject in its own right. *Visual Communication, 2*(3), 285-302.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research, 9*(11), 2579-2605

Van Leeuwen, T. (2005a). *Introducing Social Semiotics*. Abingdon: Routledge.

Van Leeuwen, T. (2005b). Typographic Meaning. *Visual Communication, 4*(2), 137-143.

Van Leeuwen, T. (2006). Towards a semiotics of typography. *Information Design Journal, 14*(2), 139-155.

Veytsman, B., & Akhmadeeva, L. (2011). Towards evidence-based typography: Literature review and experiment design. In: TUGboat.