Same difference

Version 2.0
Serebro Nabora conference, Moscow, 29 November 2014

This is a revised version of a talk originally delivered by John Hudson as the 2009 Beatrice Warde Memorial Lecture, at the Saint Bride Institute, London. The new version is shorter, more focused and contains revised and some new images. The text of the talk is significantly revised, although the content is similar to the (lost) original.

A large number of people contributed directly and indirectly to both versions of the presentation, providing the opportunity and impetus to gather these thoughts together, granting permission to reproduce images, discussing the ideas, and participating in a workshop at the University of Reading in 2009. They are all thanked here, although some of their contributions are not included in the necessarily shorter second version (notably, the student work):

Rob Banham, Catherine Dixon, Ben Weiner · Gayaneh Bagdasaryan, Maxim Zhukov, ParaType · Gillian Clayton, Kanak Chhatralia · Gerry Leonidas, Sue Walker, Mark Barratt · Craig Eliason, Peter Fraterdeus, Frode Helland, Kent Lew, Hrant Papazian, Brian Jongseong Park, Nina Stoessinger, Vladimir Tamari, David Yoon · David Berlow, David Březina, Cyrus Highsmith, Milo Ivir, Gerrit Noordzij, Andreas Stötzer, Mohamed Zakariya · Mary Dyson, Peter Enneson, Kevin Larson, Denis Pelli · Amélie Bonet, Camille Boulouis, Antonio Cavedoni, Chrystian Cruz, Gro Janarv, Julia Kaestle, Lilith Elise Laborey, Deema Merdad, Jonny Pinhorn, Annette Schmidt, Eben Sorkin, Gesine Todt

I would also like to thank the participants in the Serebro Nabora conference who contributed to the lively Q&A session following the presentation.
When I'm thinking about something, I like to understand the way I'm thinking about it. I like having a conceptual framework: it helps to clarify my ideas, and can encourage me to look at familiar things in fresh ways. I also like the fact that conceptual frameworks are disposable. When they cease to be useful, they're easier to replace than ideological commitments or unexamined prejudices.

A few years ago I began thinking about what I do, my work as a type designer, in terms of a conceptual framework. I asked myself, ‘What is the type designer doing?’ Can the work of typeface design be described conceptually, independently of a description of actions and products, and is such a description useful to that work?
Around the time that I began pondering this question, my colleague Cyrus Highsmith reminded me of this quote from Matisse:

“I do not paint things, I only paint the difference between things.”

— HENRI MATISSE

It seems to me that a large part of what I do when I am designing a typeface can be described conceptually as manipulating these relationships. There are things in a typeface that have to be differentiated from each other, and things that need to be made, in various ways, the same.
Most of the differences are, in fact, provided. They derive from the nature of writing systems, which need letters and other characters to be distinguishable from each other, and which provide multiple styles of script, each with their own differentiated characteristics. And of course they arise from the differences between writing systems even when—as in the case of Greek, Latin, and Cyrillic—they share some letterforms in common.

So the general question that I ask myself as a type designer usually regards how much of that difference to preserve. Most of what I do involves taking these things with their inherent differences and figuring out ways to introduce similarities to make them work together harmoniously.
Sometimes, though, one has to consider ways to make things different. A well-known example in the Latin script is the confusability of uppercase \( I \) and lowercase \( e \) in sans serif types, which type designers address by varying their heights, for example, or by adding distinctive features.

Similarly, the uppercase \( O \) and numeral zero can be confused, so special variant forms such as the slashed zero might be provided for circumstances in these characters must be unambiguous. This isn’t a new problem, or a new solution. The image in the middle right shows the variety of forms that Europeans in the 15th Century employed for the numeral zero.

The Arabic example in the lower line also involves confusability between a numeral, five, and a letter, \( hā’ \), but in this case the traditional solution is to employ a variant form of the letter, rather than of the numeral. So when the unconnected \( hā’ \) appears next to other letters it takes the simple ring shape \( ؤ \), but when in true isolation it takes this more complex looped shape \( ﮝ \).
Most of the time, though, I’m deciding how to make things that are different more the same. There are basically five kinds of sameness applicable in type design, a combination of some or all of which might be appropriate in resolving a particular design. [One might add ‘slant’ to this list, but the observations I would have to make about that are similar to those for featural sameness.]
The first two of these—weight and proportion—are best considered together. Weight must be balanced with proportion, as one sees in a well-designed type family with a range of widths and weights.
A lot of the work I do involves designing or collaborating on multiscript typefaces. This is a fairly old project for Microsoft: the Arabic Typesetting font from the late 1990s. The Arabic portion was designed by Mamoun Sakkal, based on a preliminary version by Paul Nelson, and I designed the Latin portion to harmonise in terms of weight and proportion. No attempt is made to make the Latin look ‘Arabesque’, or the Arabic look Latinised. They are the same only in terms of their general weight and proportion, within the graphic conventions of the individual scripts.
On the left in this slide is a Hebrew type that I designed for Adobe, along with the Adobe Thai and Arabic types on which I collaborated with Tim Holloway and Fiona Ross. In this project we didn’t design the Latin companions, but used Multiple Master sources for Adobe’s Minion type to interpolate scaled instances to harmonise with the weight and proportion of the non-Latin types. This is an ideal workflow for harmonising types of different origin and multiple weights for two different scripts.
Managing relationships of weight and proportion is the most important way to introduce a sense of sameness to the sets of inherently different symbols that make up a writing system. The overall importance of these relationships is demonstrated in the typeface used for this piece of text, which a typical reader would find unremarkable.
Portez ce vieux bon whiskys au juge blond qui fume. Tous les êtres humains naissent libres et égaux en dignité et en droits. Ils sont doués de raison et de conscience et doivent agir les uns envers les autres dans un esprit de fraternité.

In fact, this text isn’t set in one typeface, but in two quite different ones: a 16th Century French style, and an 18th Century English style. Because they have been manipulated to have similar proportions and weights, the two styles blend easily on the page, and a non-typographer would probably read a whole book typeset this way without noticing.
When you see the two complete typefaces side-by-side, it is easy to spot the difference that is submerged in their similarity of weight and proportion. This introduces the next kind of sameness that can be applied in type design, which regards stroke dynamic, expressed in the modelling of modulation of thick and thin elements in the letters.
The Dutch educator and theorist Gerrit Noordzij has systematically analysed the common types of stroke dynamics employed in writing and type design. I reckon that understanding his analysis is one of the most beneficial things you can do for yourself as a type designer, even if your instincts are to find ways to subvert the categories that he identifies.
Historically, these stroke dynamics are characteristic of particular script styles, writing systems, and periods, and are associated with particular writing tools and the way in which they are used.
It’s important to realise, though, that Noordzij’s analysis is of what is happening in the strokes that make up a letter, not in how that letter was made. This means it is tool independent, and can be used to talk meaningfully about typographic letters as well as written ones.

I expect some of you have seen the film Helvetica, which came out a few years ago. To me, it seemed strange that in 80 minutes of people talking about this one typeface nobody made this point: that Helvetica is, by virtue of its stroke dynamics, essentially a late 18th Century Romantic type with its stroke contrast reduced.
If this isn’t obvious from looking at these styles side-by-side, it can be demonstrated by algorithmically reducing the stroke contrast of a type in the Didot style—increasing the weight of the hairline strokes and decreasing that of the heavy stems and bowls—, or by increasing the stroke contrast of Helvetica. Each can be said to carry the other in its DNA.

Looking at this, you probably realise how consistent stroke dynamics can establish a relationship between typefaces of different weight, proportion, and contrast. It explains why certain serif and sans serif types might work well together, for example, and why others do not.
Ironically, this might be best illustrated by a notorious failure to use the same stroke dynamic in two weights of what is, nominally, the same type family. Here are the regular and bold weights of Times New Roman, which perversely display different stroke dynamics and, hence, characteristics of entirely different type styles.
The same stroke dynamic does not necessarily imply the same modulation axis; that is, if you imagine stroke dynamic as reflecting the path of a writing tool, you can also imagine that tool held at different angles. Since almost all the world’s scripts evolved in the practice of writing, most often by professional scribes within a culture of manuscript text manufacture, it shouldn’t surprise us that they tend to exhibit characteristic modulation axes, and that these are culturally varied.

In multiscript type design and typography, there can be a temptation to over-harmonise in this area, to apply a common modulation axis across writing systems, most often by ‘Latinising’ non-European writing systems. Not only does this tend to result in stiff, awkward forms with poor rhythm, it is unnecessary to achieving harmony in multiscript typography. In a well-conceived design such as the one shown here—Skolar Latin and Gujurati, by David Březina—the same stroke dynamic is used, but each script keeps its characteristic modulation axis.
STROKE DYNAMICS

Here’s another example: the pan-European typeface that I designed for the Dutch academic publisher Brill. Even though these are three related scripts, the Greek retains the modulation axis of the traditional lowercase forms, which evolved independently, during the Byzantine Empire. Again, the similarities of proportion, weight, and stroke dynamic are more than sufficient to ‘make same’ the three scripts while allowing each its particular cultural expression.

An interesting by-product of this approach is that certain features of Greek letterforms which, in a Latinised design, might be analysed as serifs, become more obviously part of the structural identity of the letter when written with the traditional modulation axis. This has implications for the design of sans serif Greek types.
These glyphs are also from the Brill types, and illustrate the role of featural sameness across a typeface, regardless of the writing system. In this case, the characteristic terminal shape found in the lowercase a is repeated or echoed in a wide range of letter shapes. In the Latin and Cyrillic sets, these terminals tend to be vertically turned, while in the Greek they are horizontal, conforming to the modulation axis.
In the same way that weight and proportion are closely related and can be considered together, featural sameness contributes strongly to establishing the idiom a typeface design, in the sense of traits that are appropriate to designs associated with particular periods or styles. Idiom can be understood as the way in which one typeface is similar to other types within a recognised style category. So, for example, if I mention a category such as ‘Geometric Sans’, one can form a mental picture of the features that such typefaces tend to include: their idiomatic traits.

Interesting things can occur when considering multiscript type design in terms of idiomatic sameness. This is the Nyala Ethiopic typeface, which I designed for Microsoft. The Ethiopic script has a long formal manuscript tradition, and Ethiopic text types tend to stick fairly closely to the scribal models. When I came to design the Latin companion for Nyala, I realised that echoing features of the Ethiopic counter shapes and stroke dynamic immediately suggested a particular style of Latin semi-serif calligraphy, which provided me with an idiomatic model through which to approach the design.
Since idiom is defined by similarities that constitute stylistic categories of design, it can provide a useful approach to determining appropriate features of script extensions to existing typefaces. Some years ago, I had the opportunity to design a new Cyrillic extension to Helvetica, which is shown here in the bottom half of the slide, compared with Linotype’s older Helvetica Cyrillic above.

When I asked Maxim Zhukov for his advice on approaching the new design, he immediately sent me photographs of historical Cyrillic sans serif types in the same idiom as Helvetica. This seems to me a very fruitful approach: addressing cross-script sameness not by copying as many shapes as possible from the Latin, but via the overall idiomatic style, and how this has been expressed in both cultures.

My title for this slide is, of course, ironic. It makes no more sense to ask what is the ‘real’ Helvetica Cyrillic than to ask what is the real English version of Anna Karenina. Extension of an existing typeface to another writing system is a kind of translation, and a typeface with some richness of expression will allow for more than one translation, each capturing different aspects of the original.
Those are ways in which the inherent differences between characters and even between different scripts can be harmonised or reconciled, can be made ‘the same’ while retaining those distinctions that are necessary to legibility and to cultural expression. The second question I posed was whether it was useful to think about type design in this way.

In practice, I find it most useful to apply this conceptual framework when I am designing new or unfamiliar characters that need, obviously, to work within the overall context of a typeface. The characters shown here are all extensions of the Latin script: letters for specialist orthographies such as the International Phonetic Association alphabet, or archaic and regional letters no longer in common use. Unless you’ve spent a lot of time designing such characters, so that they’ve become as familiar as abc, you need to develop a conscious strategy to harmonise them with the rest of the typeface, which means thinking systematically about the ways in which they can be made the same.
There are, of course, parallels in Cyrillic script: in the extended set of letters for various languages, and the historical letters for Old Church Slavonic. Convincingly incorporating the latter into modern styles of Cyrillic type is particularly tricky, but likely to be appreciated by scholars and specialist publishers.
Every few years, it seems, a new character comes along that is pretty much equally unfamiliar to everyone. I’m sure there were interesting discussions among Russian type designers about how to incorporate the new rouble symbol into various styles of typeface, as there were elsewhere for the euro, the Turkish lira, and the Indian rupee. It seems to be an age of new currency symbols.

Here’s a nice example of an analytical approach to another relatively new character: the uppercase form of the German sharp s or ‘eszett’. This document was prepared by my colleague Andreas Stötzer, as a contribution to a broad discussion among type designers about how to design this letter. In these images, Andreas considers features, weight and proportion, with reference to other uppercase letters.
Here Andreas considers idiomatic sameness, and precedence in the few older typefaces that incorporate this character. Note how referencing the hook of the uppercase J as a feature across a range of type styles suggests a variety of design solutions for the Eszett, each idiomatically consistent.
This way of thinking about difference and sameness can also be useful in coming up with ideas for new kinds of typefaces. Gabriola is a display type that I made for Microsoft. It consists of a single font that can be pushed in different directions using OpenType stylistic set features and a lot of complicated contextual programming. I was inspired by ancient Greek modal musical scales and liturgical chant, and the idea that the same text can be sung in different modes, each with its own tonal, melodic and emotional character.
These modes are implemented in the font using stylistic set features, for which I borrowed some Italian terminology. Each pair of stylistic sets consists of a formal, roman-like formata mode, and a cursive, more italic corsiva mode. Unlike a typical roman and italic, though, the pair is simply different expressions of the same design in terms of weight, proportion, and stroke dynamic. Only some letter shapes vary, as the styles play across different idioms. The alto styles have longer ascenders and descenders, and are also made use of contextually in the more elaborated rondo styles. The cancellaresca styles incorporate swashes, turned ascenders, and other elements of 15th and 16th Century Italian writing, while the rondo styles introduce flourishes inspired by 17th and 18th Century French and English calligraphy.
Taking this scheme, of moveable relationships of sameness and difference, and adapting it to multilingual and multiscript support was both challenging and enjoyable. The Latin and Cyrillic share similar conventions in distinguishing formal and cursive styles with different letter construction. In the Greek, there isn’t the same systematic variation, but instead there is a rich tradition of ligatures between letters that can be made more or less cursive.
Now that we’ve considered what the type designer is doing, and how it can be useful to think about type design in terms of sameness and difference, we’re left with the question Why? Are there benefits?

The obvious benefit is an aesthetic one. Culturally, and perhaps even biologically, we favour harmony over dissonance. We appreciate beauty in a page of even texture, which is not to say that other kinds of beauty don’t exist, but that as typographers we want to attract people to read text, and then we want it to be easy and pleasurable for them to do so.

Is there a benefit to the reader in all this ‘same difference’? 550 years of typographic book making and other kinds of publishing suggest there must be; otherwise, I think we would see more varied models of typography instead of what we do see: similar concern for even texture and harmony across the world’s typographic traditions, regardless of language and script.

I want to spend the last few minutes of this presentation discussing a phenomenon that I suspect accounts for this global typographic tendency, and which goes some way to answering the question Why?
The phenomenon I want to discuss is spatial frequency, and how we perceive it, and I’ll try to do this with as little scientific terminology as possible. What we’re looking at in this image is a progression, from right to left, in spatial frequency from high to low. High spatial frequency means more changes between dark and light bands, in the same distance, compared to fewer changes between dark and light bands at low spatial frequency. There is also a progression from bottom to top, between high contrast at the bottom to low contrast at the top.

There are a couple of interesting things to note about this image. One is that as the contrast of the high spatial frequency bands decreases it becomes harder to see them, so the upper right corner of the image appears to be just a uniform grey area. In the middle of the image, though, we can perceive the bands further up, which indicates that we are more sensitive to particular ranges of spatial frequency.

The other interesting thing is that this image looks different depending where you are sitting in the room. The further back you are, the more difficult it will be for you to distinguish the high frequency bands as the contrast decreases, so the larger that uniform grey area in the upper right will appear. If we were able to get far enough back, almost the whole rectangle would appear grey.
We can simulate this, of course, by making the image smaller. This is a useful reminder that when it comes to perception, size and distance are the same thing.
The most fun part of science is always the experiments, so let’s do one. Take a moment to look at this slide, and confirm for yourself that the pattern of bars at the top and bottom are the same. In a moment, I’m going to change the image, so that the red dot is replaced with a horizontal red line, and I want you to spend a few seconds moving your eyes back and forth along that red line.
Scan back and forth along that red line, and then when I change the image again I want you to fixate on the red dot in the centre.
If the experiment was successful, most of you should have experienced a brief perceptual distortion, in which your brain had trouble determining that the pattern of bars at the top of the image was the same as the bottom pattern. The reason for this is that the brain tunes to spatial frequency channels, and it takes time to switch between these channels. It’s like tuning a radio between different radio waves: it takes a while to manipulate the dial and move from one channel to another.

Reading is an activity that most human beings are very good at; so good, in fact, that once we’ve learned to read our conscious experience of reading is of something almost effortless. Indeed, most of the time, we’re not really conscious of the activity of reading at all, but lose ourselves in the meaning of text. For the brain, though, reading is a very intensive process: there’s a huge amount of activity to be coordinated between moving the eyes, processing the visual stimuli, recognising words, and comprehending syntax, morphology and all the other linguistic information that makes up meaning. The last thing you want to be doing while reading is spending time and resources tuning the dial on your mental spatial frequency radio.

So, we’ve determined that 1) we’re more sensitive to some spatial frequencies than others, 2) we adapt to the spatial frequency channel of what we’re looking at, and 3) it takes time to re-tune to a different spatial frequency channel.
This is an image from a Vision Research paper by Najib Majaj, Denis Pelli and others from 2002, entitled ‘The role of spatial frequency channels in letter identification’. This is another image that is going to appear different depending where you are sitting in the room. There are three Latin letters in this image, each occupying a different octave of spatial frequency. Depending how far away you are sitting, you will see a letter D, letter E, or letter F, or possibly some combination of two of these letters.

This indicates that letter identification is size-dependent (remember: size and distance are the same thing), and for a given size we are tuned to a particular spatial frequency channel. In the experiments reported in the paper, the authors were able to demonstrate that for each writing system, typeface, and text size, letter identification is mediated by a single, one or two octave-wide visual channel.

This, I think, suggests very good reasons why the type designer is working so hard to harmonise characters, especially with regard to weight, proportion and stroke modulation. In fact, what the type designer is doing is bringing all the different letters and symbols within a single spatial frequency channel, allowing the reader to engage with the text without needing to switch between channels. Put that way, it seems almost prosaic—certainly less profound than Mattise’s ‘painting the difference between things’—but in practical terms, it seems to me a useful insight. It doesn’t change what the type designer is doing: it affirms it.