

Typography for academics

Pär Nyman

October 8, 2026

It has never been easier to typeset documents. Modern software allows us to almost instantly make typographic adjustments that used to take hours for even the most skilled craftsman. Ironically, as the capability to produce documents has become ubiquitous, the typographic quality of published texts has worsened dramatically.

In academia, we are all expected to self-publish documents, without previous training in typographic principles or the knowledge of best practices that professional typesetters developed over generations. Perhaps it should not surprise us then that academics regularly produce documents of poor typographic quality. What is worse, the problem is not isolated to self-published material; many publishers of scholarly material hold standards that are lower than what would be expected of publishing houses outside academia.

Personally, I think that the do-it-yourself attitude is one of the most attractive features of life in academia, and something that fundamentally separates our work from the high degree of specialization that characterizes modern work life. Because of the way our work is structured, we get to learn the basic principles of everything from programming and communication to teaching and organizational management. I find this constant learning immensely satisfactory, but I also think that this privilege comes with a responsibility to actually learn these things.

This text is written to those who – without too much effort – wants to improve the typographic quality of their academic texts. The ideas and principles presented are not new. On the contrary, the aim has been to summarize typographic principles developed over generations together with practical advice for how to accomplish these things with the help of modern software.

Academic texts are usually written in Microsoft Word (or similar software that can create .docx files) or TeX (with that term I refer to the family of macro packages and TeX derivatives known as LaTeX, pdfTeX, LuaTeX, etc.). Unfortunately, there is a lot of elitism surrounding these alternatives, but no, you are not brighter or a better researcher because you use the latter. However, from a strictly typographic perspective, TeX has important advantages. There are other reasons for why someone may prefer Microsoft Word – e.g., it is easier to learn, it provides better tools for documents with multiple authors, or it may be required by a co-author or journal – but matching TeX's finer typographic refinements is difficult. Then again, typography is not everything, and this text will include instructions for how to incorporate different techniques in both Microsoft Word and TeX. After all, a properly typeset Word document will look much better than most texts written in TeX.

This text is organized into two main sections, followed by shorter discussions of specific typographic topics. The first section deals with the design of the text, or macro-typography. Most of these aspects are usually directly visible to the reader, like the typeface, page margins and whether paragraphs are separated by vertical space or indentation. To a large extent, these are also matters of taste and style. The second section deals with micro-typography, which primarily is about getting the space right between characters, words and lines. The aim of micro-typographical refinements is to reduce the number of disturbances that face the reader, and to simply make the text look better, regardless of stylistic preferences. Successful micro-typesetting is therefore only noticed by a trained eye, but it will make the text more pleasing to anyone who reads it.

I have tried to mix words that have an intuitive meaning with the corresponding typographical terms. So for example, line-height is sometimes referred to as leading, and letter spacing as tracking. The purpose of the more technical language is to make it easier for the reader to search and find relevant content on the web, or elsewhere. At the end of the document is a short dictionary over all the technical terms used.

Designing the document

When writing academic texts, there is not too much one can do about the layout. Compared to other publications, not to mention web pages, the degrees of freedom are limited and most academic books and papers tend to follow more or less the same template. Yet, both researchers and academic journals often demonstrate an impressive level of creativeness when it comes to producing bad typography with limited means. Table X in the end of this document provides a list of the most important things to avoid, but a more constructive approach may be to first consider what it is that we want to achieve with our page layout. There are many technical terms that describe different typographical goals, but in a text of this kind I think it helps to collapse these different terms into three intuitive categories: structure, readability and appearance.

First, a good page layout creates a structure that is easy for the reader to understand and navigate. The most important part is that the document has a clear *hierarchy* (sometimes this is referred to as clarity), in which the visual weight given to an element corresponds to its hierarchical level. At the most basic level, this just means that higher level section titles should be easily distinguishable from lower level titles. Visual weight can be achieved by changing typeface, increasing the font size, using boldface or italics, or increasing the white space above and below the element. Moreover, elements that belong together should be visually separated from other content. This means that captions should be closer to their figures or tables than to the surrounding text, and that the white space above a section title should be larger than the space below it. Typographers sometimes refer to the grouping of elements as *alignment* and the amount of between them as *separation*, but because space and closeness is always relative I prefer to see them as two sides of the same coin. Lastly, it is important to consider in what order your readers will digest different parts of your document. We tend to write as if the text will be read from the first to the last word, but most readers of an article will see and process elements with lots of visual weight before they read

the rest of the text. You should use that to your advantage by letting some pretty graphs communicate your most important messages while making sure that someone who just browses your tables gets a fairly accurate idea about the empirical content in your publication.

A second aspect to keep in mind is that the text should be easy to read. Typographers often distinguish between *readability* – how easy it is to read blocs of text – and *legibility*, which is an inherent characteristic of the typeface and concerns how easy it is to distinguish and recognize individual glyphs. Both are of course important to consider when designing a document. For academic publications it is important that reading a long text does not cause eye-fatigue. This tends to go hand-in-hand with readability and legibility, but some text (as many high-contrast sans serif typefaces) can be easy to read while also eye-fatiguing in the long run.

Lastly, we all want our documents to be aesthetically pleasing as well as to communicate the right feelings to the reader. I did not use the default LaTeX font for this text, simply because I did not want it to look as just another empirical study filled with equations and regression tables. And if we are writing an early draft, we may avoid justified text and multicolumn layouts just because we do not want to give the impression of a finished and professional-looking product.

But how does one go about to create a beautiful document, without letting the aesthetic aspirations harm the readability of the text? I have three recommendations that I think are helpful in most typographic situations.

- Less is more. This is an often repeated truth, and to some extent it is possible to improve a document by simply removing unnecessary elements and variations. For example, there is never a good reason for using both indentation and vertical space to indicate the separation of paragraphs. But for a truly clean and minimalist layout to work, it is important to work with all aspects of the document. How little you can get away with is dependent on your typographic consistency, your micro-typesetting (especially the consistency of inter-word spacing) and how much visual weight that you have used in other places (within-elements or to separate elements from each other). When creating a page layout, it is therefore a good strategy to start with the body text – and choices like typeface, point size, line length and line spacing – before proceeding with the items that require more visual weight.
- Don't fear white space. It is recommendable to have wide page margins, if it improves the readability by reducing the number of words per line. And white space is in many instances preferable to the use of lines when separating elements from each other, partly because it looks less cluttered and partly because it is so flexible – it corresponds to an easily interpretable hierarchy (more space, more separation), it can be used for almost all elements and it can be stretched (see the section about optimizing a document).
- In the end, logic and rules of thumb only takes you so far, and there is no substitute to printing your document using different settings and comparing them in printed form. Do not forget to ask your colleagues which layout they like the most.

The details of Microsoft Word's default template vary between versions, but its choices of typeface, heading style, margins and paragraph spacing have often conflicted with the typographic recommendations outlined here. Most people who write academic texts in Word would therefore benefit from creating their own template. Defining formatting choices through styles is important, as it makes it easier to formatting the text, maintaining a consistent visual hierarchy and experiment with different typographic choices.

The details of Microsoft Word's default template vary between versions, but it has often been a good representation of violations against typographic principles, with sans-serif body text, blue-colored serif headings, narrow page margins, no vertical space after headings, etc. Most people who write academic texts in MS Word would therefore benefit from creating their own template.

I think the opposite is true for LaTeX. Here, the default page layout is pretty much in line with typographic principles, and it is reasonable to think twice before making any major modifications to the default settings. In fact, most academic papers I have read are of lower typographic quality than the article template with default settings. However, it is boring with papers that all look the same, and the typographic style in the default template may not give the right associations for the text you are writing, so I would encourage everyone to experiment. Just make sure that you know what you do.

Designing the body text

The most important parts of the body text is the typeface, the font size, the line-height and the line width. The typography of a body text can therefore be summarized by specifying these parameters, as for example 10/12x22 Garamond (10 pt Garamond with 12 pt line height and 22 pica line width).

Typeface

Many people think first of the choice of typeface when they think of typography. Since this choice is largely a matter of style and expression, and since I know too little about typeface design to offer much useful advice, I will give it relatively little attention here. Most importantly, the typeface should suit sustained reading and the character of the document. Judge it using several paragraphs at the intended size, and evaluate its appearance both on paper and on screen.

It is easier to say something about practical considerations. Choose a family with well-designed italic and bold styles and the characters required by your text. For academic writing, useful features include kerning, standard ligatures and tabular figures, which give all numerals the same width. If mathematical notation is prominent, consider both the availability of the required symbols and how well the mathematical typeface matches the body text.

For a new choice intended to work across several applications, a modern desktop font in OpenType format is generally a good starting point. OpenType can accommodate extensive character sets and typographic features, but the format does not

guarantee that a particular font includes them. The features and compatibility of the particular font matter more than its filename extension.

Also consider whether the typeface can be used throughout your workflow. Word and many statistical programs can use fonts installed on your computer, although support varies between programs, operating systems and export formats. A typeface available through a TeX package is not necessarily installed as a desktop font that your statistical software can use. If you want matching text in the document and its graphs, check that a suitable desktop version is available and inspect an exported graph before committing to it.

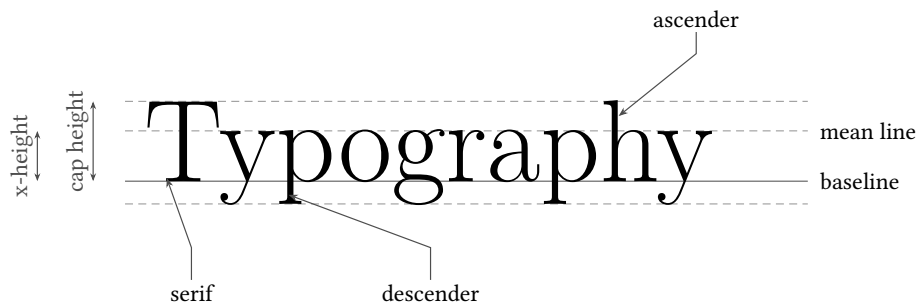
To find out which features a font supports, consult the documentation or specimen supplied by its designer or distributor, preferably for the exact version you are using. A font-inspection tool such as FontDrop can also show the characters, ligatures and OpenType features contained in a font file.

Font size

Personally I think 11 pt is a suitable point size for common serif types as Garamond or Times New Roman, but it is important to note that the appropriate font size crucially depends on your choice of typeface. Font size refers to the height of the em square – a reference box used when designing the glyphs of a digital typeface – which roughly corresponds to the distance from the top of the ascenders to the bottom of the descenders, usually with some additional space. However, that does not mean that font sizes are directly comparable between typefaces. First, font sizes should be seen as approximations, and there is nothing that prevents a font-designer from deviating from this convention. In practice this means that one font may simply be larger than another font, even at the same point size, just as two shoes of the same size can fit differently on your feet. For example, Century Schoolbook and Palatino are larger than most other typefaces, while Garamond is on the smaller end. Second, the perceived size of a font, as well as its readability at different point sizes, is dependent on how the typeface is designed. The most important characteristic in this regard is the height of the midline, or the upper part of most lower case letters, which is often referred to as the x-height. Typefaces with a high midline are usually easier to read, especially with small font sizes. Third, the font size may also depend on the medium(?) through which the text will be produced for the reader. Academic papers are often printed on low-resolution office printers, which can make it fatiguing for the eyes to read long texts in small point sizes.

Line-spacing

Typographers like to refer to line-spacing as *leading*, which initially referred to the size of the strips of lead that were inserted between lines of type, but today both terms usually refer to the vertical distance between two adjacent baselines (the height of a line plus the distance between the lines). A common recommendation is to let the line spacing be 120 percent of the point size (this is what most people think about as single-spacing), or 2 pt of space between the lines. However, remember to use such numbers as a starting point when you experiment, rather than as some kind of



answer. A single point of leading may be enough for the narrow measure found in many newspapers, while other texts can benefit from leadings of up to 160 percent of the point size. A larger line spacing makes the text appear lighter to the reader, which means that dark and heavy fonts may need more leading to not appear intimidating. More leading is also advisable for typefaces with large x-heights, which is often true for sans-serif types, as well as for larger text widths. Text set with larger point sizes need less leading in relative terms (as percent or em), but more leading in absolute terms (as pt).¹

Everyone is familiar with concepts with like single- and double-spacing, but these concepts lack clear definitions. Single-spacing usually refers to a line-height that is 120 percent of the font size, e.g., 10 pt type over 12 pt lines, in either nominal or actual terms. Double-spacing can either refer to a line-height twice the single-spacing (240 percent of the font size), or twice the font size, but regardless of definition it is neither good for aesthetics or readability. Journals sometimes require authors to submit their manuscripts with double-spacing. That is just dumb, so resist that if possible. If they do not leave you a choice, make sure to keep the preferred line-height for other versions of your manuscript, as well as to go over the complete layout of your paper; dramatic changes to the line-height will probably call for other changes too.

In Microsoft Word, single-spacing refers to a typeface-specific 120 percent-rule, where the line-height is set to 120 percent of the space needed to accommodate the chosen typeface, rather than a fixed percentage of the nominal font size. This is probably a good default for most documents and users, but if you are mixing typefaces – which is common if you are using in-line mathematical notation – or including other forms of elements within the body text, you are better of setting the line height to an exact number of points. The double-spacing setting results in lines that are twice as high as the single-spacing and thus 2.4 times the actual font height. When working with leading less than single-spacing, be aware that some letters may appear on screen as if they are cut off, but they usually look fine when printed or converted to pdf.²

In TeX, the default line height is defined by the document class and the nominal

¹ Felici 2012 proposes a formula which acknowledges the relationship between measure and point size to calculate how much extra lead to add. He simply divides the measure (in picas) by the size of the type (in points) to get the extra lead required.

² Professional typographers dislike this kind of automatic line-spacing, partly because it results in non-transparent leadings that are sometimes difficult to express in points, but such objections are hardly relevant for non-professionals

font size. So for example, the article class uses 12/13.6/14.5 pt leading for a font size of 10/11/12 pt. If you want to modify the leading in the entire document, the best way is to type `\linespread{#}` somewhere in the preamble, with the # replaced by the appropriate scaling factor. For 10/13 typesetting, you would use a scaling factor of 1.0833 (13/12). Note that many typeface packages adjust this factor, and that by including a `\linespread` command in your pre-amble you are overriding the setting that came with the package, which may have been smaller or larger than 1. If you only want to adjust the line-height for the body text and bibliography, which is usually the case when a journal requires you to use double-spacing, it is better to use the `setspace` package and the `\setstretch{#}` command. You can also use the package's `\doublespacing` command, which sets the scaling factor to 1.667/1.618/1.655 for a font size of 10/11/12 pt, effectively generating a line height that is 200 percent of the nominal font-size.

Measure and margins

Most software packages let us set the margins directly, which then gives us the text width, or measure, as a residual. This is unfortunate, because it leads us to think about margins being more important than the length of lines, when it is the other way around.³

We have all grown up with the inch-wide margins that is the default of Microsoft Word and similar word processors. However, with single columns on A4 papers, this typically results in a text with more than 100 characters per line. While studies on reading speed and accuracy differ in their results, a useful typographic rule of thumb is to aim for 45–75 characters per line, including spaces. Academic texts with single-column layouts should aim for the upper part of this interval, or the similar guideline of 2–2.5 times the length of the lowercase alphabet, unless there are special reasons for having narrow margins – e.g., page restrictions, environmental considerations or large tables that need narrow margins to fit. The smaller the type, the larger the margins should be, but if they become too wide the extra white space may look offsetting to the reader.

Another way of reducing the text width is to use a layout with multiple columns. This can reduce paper consumption when the document is printed, while allowing wide tables to span all the columns. However, narrow columns require careful attention to hyphenation and inter-word spacing to avoid excessive gaps between words. They can also make screen reading less convenient: if the page does not fit comfortably on the screen, readers may have to scroll down one column and then back up to the next. Multi-column layouts also give the document a journal-like appearance, which may or may not be what you are looking for.

There is not much to say about how to adjust the margins in Microsoft Word. In TeX, the margin width is automatically adjusted depending on the chosen font size. It is also possible to set it manually, using the `geometry` package, but that is usually not necessary.

³ *Measure* and *line length* is sometimes used interchangeably, but technically the line length refers only to the text and not to the white space that would be found between the text and the right margin if the text is set with a ragged right margin.

Paragraph separation

There are two common ways of separating paragraphs. The first is to indent the first line of each paragraph, an approach often used with justified text. I recommend omitting this indentation immediately after section headings, displayed equations and other elements that already provide a clear visual break, even when the following text begins a new paragraph. For single-column layouts, an indent of approximately 1.5–2 em (i.e., 1.5–2 times the font size) is a useful starting point, while 1 em often works well in multi-column layouts.

The other alternative is to separate the paragraphs by inserting white space between them, which is most common for texts with a ragged right margin. The appropriate amount of white space could be somewhere between 0.5–1 times your line height. And most importantly: do not use both kinds of paragraph separation at the same time.

In Microsoft Word, the default setting is nowadays to insert a 10 pt vertical space between paragraphs. However, this is easily changed in the paragraph settings. Remember that it is always better to create this space through the paragraph settings instead of using consecutive carriage returns, because it gives you unlimited flexibility to change the number if you would like and you do not run the risk of beginning a new page with an empty line.

In TeX, the default paragraph indent is dependent on the document class and font size, but it tends to be around 1.5 em. If you want to replace the indentation with vertical white space, you could add the following lines in your preamble. The last arguments allow the space to be up to 2pt smaller or larger than target, if it improves the placement of other elements (see Optimize the document).

```
\setlength{\parindent}{0em}  
\setlength{\parskip}{0.75\baselineskip plus2pt minus2pt}
```

Placement of figures and tables

As a rule of thumb, graphs and tables should be placed at the top of a page, where they give the layout a clear and orderly structure. There are cases when a small, unobtrusive figure may fit naturally between two paragraphs – much like an equation – especially when the reader benefits from seeing it alongside the relevant discussion. Larger graphs and tables are less suited to this placement, as having text above them tends to make the page look cluttered.

Some people are keen on placing their tables exactly where there are referred to in the text, so they can control what the reader has read before they look at the table, and what they will read after. However, this may not be the order that readers approach the document. On the contrary, many readers will begin by looking at graphs or tables, or browse between tables to quickly compare results. And when they do, they will search the captions and notes for the information they need, rather than the text immediately preceding the table.

Following this rule means that there must be some flexibility with regards to where figures and tables are placed in relation to the surrounding text. If you are not yourself responsible for the final page layout, it is therefore bad practise to write about the “table below”, as it actually may appear above that sentence.

Notes

Notes of all kinds – footnotes, endnotes and notes below figures and tables – are typeset with a smaller font size than the body text. A two-point difference is usually about right. The leading is either similar to the body text, or more compact. I think that 110–120 percent of the point size tends to look good. The aim is to create a clear hierarchy and visual separation without making the notes unnecessarily difficult to read.

There is perhaps more disagreement about the indentation. It is my impression that footnotes are most commonly typeset with first-line indents, although there appears to be variation across disciplines in this regard. However, indentation at the top of a footnote section is not very appealing, and because the footnote section is already separated from the main text there is no need for additional separation. In texts with few and short footnotes, in particular those that use the Oxford system of placing references in footnotes, I would advise against first-line indentation. Personally I prefer to use a hanging indent, with only the numeral flushed to the left margin, similar to how lists are usually typeset. This is common in books, but rarely encountered in academic journals.⁴ A third alternative is to simply ignore indentation. Even though is rarely done by professional book publishers, it is rather common in academic journals and I do not think it looks too bad.

The default behaviour of Word is to typeset footnotes without any indentation, but adding an indent is easily done in the paragraph settings. If you are using justified text, you may want to replace the space between the numerator and the text with a fixed-width space. This is obviously not needed for hanging indentation, where this space is given by the indent.

The default behaviour of TeX is to indent the footnote. The code below shows how to use hanging indents instead, as well as how to adjust the size of the indentation. A narrow indent is always recommended, but just make sure that there is enough room in the margin to accommodate enumerators with at least two-digits.

```
\usepackage[hang,flushmargin]{footmisc}  
\setlength{\footnotemargin}{.75em}
```

Optimize your document

In practise, it is not possible to achieve a perfectly typeset document. Just consider the trade-off between consistency and well-placed page breaks. On the one hand, we would like to see consistent spacing between words, between paragraphs and other elements as well as between the page bottom and the last line of text. On the other hand, we also want to avoid page breaks after the first (orphans) or before the last line of paragraphs (widows), just before an equation, after hyphenated lines as well as in the middle of a footnote. The best way to balance this trade-off is to find a good compromise where we allow some spacing inconsistency, but not so much that it becomes a disturbance, and thereby avoid the worst page breaks.

⁴ To some extent, this may reflect a preference against hanging indents in multiple-column layouts.

Word is not built to handle trade-offs like these. With some notable exceptions, like the inter-word spacing in justified text, the amount of spacing between elements is fixed, and adjustments need to be done on a manual basis. I regard this as a major weakness, but there are also positive aspects: you never have to wait for a document to compile, and for very small documents it is actually rather convenient with only fixed spacing. To avoid the most disruptive page breaks, Word users should specify which elements must remain together. In the paragraph settings, “Keep with next” can keep a heading with the beginning of the following paragraph. For a figure placed in line with the text in its own paragraph, applying the same setting to each paragraph in the sequence except the last keeps the image, caption and notes together.

TeX formalizes its treatment of such trade-offs by defining a set of penalties for a variety of imperfections, and then tries to minimize various sums of these penalties. There are two ways of adjusting these parameters. First, whenever we define a space or a distance between two elements – TeX users sometimes refer to the elements as boxes and the distance as glue – we can also allow this distance to be stretched or shrunk up to a pre-determined amount. Second, we can also adjust the penalties associated with different imperfections. The TeX defaults are reasonable, but I have sometimes increased the penalties associated with orphans and widows.

Micro-typography: Typesetting the text

We now turn to details that most readers rarely notice individually, but that contribute to an overall impression of an orderly, attractive and comfortable text. Micro-typography is less about stylistic preferences or the character of the document and more about refining its typesetting.

Kerning and ligatures

The procedure of adjusting the horizontal placement of characters based on what pairs of characters that are adjacent to each other is called kerning. Without such adjustment, the distance between some pairs of characters will appear too short or too long. While it is not overly important for the general impression of a text, it can strongly affect the perception of a title with an unfortunate sequence of letters, or the reference list, where the large number of dots and commas make kerning more important. Most TeX users have probably encountered bad kerning when words or abbreviations are used in equations and therefore, or rather unless the proper adjustments are made, the words are typeset as math instead of text.

In TeX, kerning is activated by default. In Microsoft Word, kerning needs to be activated in the font settings dialogue box (CTRL+D). The examples below show what the text name of William Hof’s fictional audio- and videostore looks like with and without kerning (Microsoft Word, Garamond). The text with kerning is much more appealing to the eye, but this is of course an exaggerated example with several combinations of characters that are known to require heavy kerning.

Combinations of characters can also be typeset using ligatures, which are glyphs that represent the combination of two or more characters. For example, an f followed by an

W. Hof's AV W. Hof's AV

Figure 1: Without kerning (left) and with kerning (right)

i tends to give a jarring look, because the bulb on the f and the dot above the i crowd into each other. In a ligature, this can be solved by connecting the two characters, so that the bulb and the dot completely overlaps, or by increasing the space between them. The design of the ligature, as well as whether a special ligature is needed or if the two characters look good together, depends on the typeface.

The range of available ligatures varies between fonts. Some include only common combinations such as fi and fl, while others also provide ff, ffi, ffl and discretionary ligatures. OpenType fonts can use glyph-substitution rules to select ligatures automatically, provided that the software supports and enables the relevant features. The example below is from Garamond in Microsoft Word.

fi fi

Figure 2: Separate characters (left) and ligature (right)

In Microsoft Word, automatic ligatures can be turned on through the advanced tab in the font dialogue (CTRL+D). If the chosen font does not support automatic ligatures, they can be inserted manually through the *Insert Symbol* dialog, in the subset *Alpha-betical Presentation Forms*. However, this is considered bad practice, as it can cause complications when searching, editing and copying text. The spell-checking function in Word has some kind of support for manually inserted ligatures, but I am not sure how well it works.

The default behaviour of TeX is to replace certain character combinations with ligatures and the user does not need to make any adjustments for the ligatures to look good. Unfortunately, ligatures can sometimes cause problems when searching or copying text from a PDF if the file lacks correct mappings between glyphs and Unicode characters. This is the very reason for why the character combination fi, as well as both en- and em-dashes, sometimes are transformed into strange characters when copy-pasted into another software – a problem I think virtually everyone in academia has encountered. In most instances, this problem is solved by simply setting the font encoding to T1:⁵

```
\usepackage[T1]{fontenc}
```

There are some other risks with ligatures that should be mentioned. First, if a font only has fi and fl ligatures, but no ligatures for ffi or ffl, the latter combinations will look unbalanced if the last two characters are replaced with a ligature. In those instances, words like affirmation or baffling should be typeset without ligatures. Second, do not

⁵ If the problem persists, check that the PDF contains correct glyph-to-Unicode mappings; with pdfLaTeX, these can be generated using glyphtounicode or the cmap package.

use ligatures if you are adjusting the tracking (letter-spacing), since it would lead to inconsistent spacing between the characters.

Justification, hyphenation and inter-word spacing

Academic texts tend to contain long words which, unless they are properly hyphenated, easily impair the appearance and readability of a text. If words are hyphenated too often, or at the wrong places, the text will be difficult to read. And if they are not hyphenated often enough, the right-side margin will give a very ragged appearance (if the text is left-aligned) or the inter-word spacing will be inconsistent (if the text is justified).

Good hyphenation is primarily achieved with the help of dictionaries in your software. As long as you have the right language installed and chosen in Word, or the right babel option in TeX, things usually look pretty decent. It is also possible to manually control hyphenation from within the document, by inserting ‘soft’ or ‘optional’ hyphens where line breaks would be appropriate. While it can substantially improve the visual impression of a text, it can also be time-consuming and clutter the text (for TeX users and Word users who like to have formatting symbols on). Unlike most typographic aspects dealt with in this text, this is a case where the author faces a clear trade-off between typographic quality and productivity. In Word, soft hyphens are inserted using CTRL+- (hyphen), while TeX users type \-.

Similarly, use non-breaking spaces to prevent line breaks between elements that belong together, such as a number and its unit or a table number and its label. In TeX, use a tilde instead of an ordinary space, as in 10~kg or Table~1. In Word for Windows, press Ctrl+Shift+Space. Use these spaces selectively, since unnecessarily restricting line breaks can make it harder to achieve even spacing in justified text.

As you have probably realized yourself, manual hyphenation works better in Word than in TeX. To begin with, TeX users need to compile the document before they can see what the text looks like, while the end result is immediately visible in Word. I usually do not have a problem with this – on the contrary, I like to be able to type the text without being disturbed by what it looks like – but it is less than ideal if you are manually hyphenating a close to a finished product. Moreover, most Word users hide formatting symbols like soft hyphens, while TeX users have no choice but to see their text cluttered with special characters.

I am a fan of justified texts. When properly typeset, they look so much better than texts with a ragged right margin. This is especially true when paragraphs are separated with indentation instead of vertical space. However, to make each line cover the full textwidth, one must instead allow the spacing between each word to differ across lines. And unless long words are hyphenated at suitable places, there is a risk that the text will be plagued by confusing hyphenations or too much variation in the inter-word spacing. This risk is even larger when the lines are short.

In TeX, one can also force a part of a text to be placed on the same line by nesting it within a mbox. The most common use is probably to make sure that the asterisk key below a regression table (* p osv) is not split over multiple lines.

The finer details

In TeX, there are many ways of refining the aspects dealt with above. Anyone with a special interest can read the package documentation, but for most occasions it suffices to just put the following in the preamble (or read Section 9 in the documentation for the most common problems that the default settings may cause):

```
\usepackage{microtype}
```

By default, this package enables character protrusion and font expansion where supported by the typesetting engine. Character protrusion, also known as margin kerning, allows selected characters, especially punctuation, to extend slightly beyond the left or right margin, creating a more even optical alignment. Font expansion makes very small horizontal adjustments to glyph widths, allowing more consistent inter-word spacing and potentially better line breaks. The package also provides additional kerning and inter-word spacing adjustments, but these are not enabled by default.

If you face a page limit, you can use the options ‘stretch=0’ (character width may not increase) and ‘shrink=50’ (character width may decrease by up to 5 percent; default is 20/2) as a quick solution. But this is not necessarily good typographic practice and the result must be checked visually.

Using the correct characters

Small distinctions between characters can make a noticeable difference to the appearance and clarity of a text. Hyphens, dashes and minus signs serve different purposes, as do opening and closing quotation marks, yet these distinctions are too often overlooked. Using the correct characters is a relatively easy way to improve the typography of an academic document.

Hyphens and dashes

Hyphens join words in compounds such as *well-known* and mark words divided across lines. En dashes are longer and serve several other purposes:

- In ranges, such as 1849–1912 or Tables A9–A12.
- To indicate relationships, contrasts or pairings, as in *left–right scale*, *cost–benefit analysis*, Djurgården–Hammarby or the score 9–1.⁶

⁶ Note that the choice between a dash and a hyphen also effects the meaning of a concept, at least for those who are not familiar with it since before. Let us consider two theoretical frameworks that are common among economists and political scientists: the citizen-candidate and the principal-agent models. Because the first model is written with a hyphen, we can assume that this model is saying something about citizens as a form of candidates. Had it been written citizen–candidates, it would have been about the relationship between the two groups. Conversely, the principal-agent model is about the relation between the principal and the agent, while a principal-agent model assumingly would be about the most important agent. Hence, we can communicate key dimensions of the framework, solely by distinguishing between dashes and hyphens. Similarly, Swedish–American relationships are relationships between Sweden and the US, while Swedish-American relationships supposedly are relationships that Swedes who have emigrated to the US may have.

- Between names in jointly named concepts, as in *Huber–White standard errors*.
- In compound modifiers where one component consists of several words, as in *New York–based researcher*. The en dash indicates that *based* relates to the whole expression *New York*. This works well for capitalized proper nouns, but in more complicated constructions it is often better to rephrase.

Dashes can also mark a break in a sentence or set off an explanation or aside. Many British styles use a spaced en dash for this purpose – as illustrated here – while American styles often use an unspaced em dash. Either convention works, but use it consistently.

In TeX, type – for a hyphen, -- for an en dash and --- for an em dash. In Word for Windows, type CTRL+minus or ALT+0150 for an en dash, CTRL+ALT+minus or ALT+0151 for an em dash.

Minus signs

For negative numbers and subtraction, use a mathematical minus sign rather than a hyphen or an en dash, as the minus sign is designed to match other mathematical symbols in height and weight. In TeX, hyphens are automatically typeset as minus signs when written in math mode. In ordinary Word text, type ALT+8722 for a minus sign. Also be aware that tables and numbers copied from Excel unfortunately come with hyphens instead of minus signs.

Quotation marks and apostrophes

In English text, opening and closing quotation marks have different shapes: “like this”, or ‘like this’. Use typographic quotation marks rather than straight typewriter quotes. Whether the outer quotation marks should be double or single depends on the style you follow; quotations within quotations normally use the other kind.

In TeX, type ``quoted text' ' for double quotation marks and `quoted text' for single quotation marks. The opening character is a backtick, while the closing character is the ordinary keyboard apostrophe. That apostrophe also produces the correct character in words such as don’t and author’s.

In Word, enable the replacement of straight quotes with smart quotes under File > Options > Proofing > AutoCorrect Options > AutoFormat As You Type, and set the text’s proofing language to English. Check the result, especially after pasting text from other sources. An apostrophe should have the same shape as a closing single quotation mark, including when it appears at the beginning of an abbreviated word or date, as in ’90s. Automatic formatting sometimes gets this wrong. To insert the correct apostrophe manually, type ALT+0146.

Figures and tables

Figures and tables deserve at least as much attention as the body text. They often contain the most important information in a paper, and their design affects both how

easily readers can understand that information and which aspects they notice first. Poorly formatted figures and tables are also often what lets down the overall impression of an otherwise carefully typeset document. Many of the principles discussed above apply here too: clear hierarchy, consistent formatting and careful use of space.

As discussed in the section on page layout, readers often inspect tables and figures before reading the surrounding discussion, and some may never read that discussion at all. Tables and figures should therefore be self-contained: their titles, labels and notes should provide enough information to understand what is shown and how to interpret it without searching through the main text.⁷ This is not primarily a typographic issue, however, so I will not pursue it further here.

Tables

The typography of a table should make comparisons easy. Numerical columns should normally be aligned on the decimal point, so that digits representing the same place value line up vertically. Columns containing only integers can be right-aligned. Where available, use tabular number spacing: numerals that all occupy the same horizontal space. This gives the numbers an orderly appearance without requiring a monospaced typeface for the rest of the table. Text labels should generally be left-aligned. Use a consistent and reasonable number of decimal places for comparable quantities.

Avoid vertical lines and use horizontal lines sparingly. In most tables, a line above the column headings, one below them and one at the bottom will be enough. Additional lines can sometimes be used to distinguish major groups of columns or rows, but alignment and spacing should do most of the work of revealing the table's structure. This is also the approach encouraged by the LaTeX package `booktabs`.

Figures

There are a few principles for good graphs that I try to follow in my own work. First, I think most scientific graphs should be created with one of two purposes. A communicative graph should emphasize the message it is intended to communicate, rather than trying to convey as much information as possible. An exploratory graph should instead provide as much relevant information as possible to let readers examine patterns, identify unusual observations or assess the analysis. The former benefits from a clear focus, while the latter may justify a much denser display. These purposes can overlap, but deciding which is primary helps determine what to include and how much visual emphasis to give it.

Second, variation in the display should correspond to meaningful variation in the data. Colours, markers and line styles should highlight the groups and contrasts most relevant to the analysis, while the perceived magnitude of a difference should reflect the underlying numbers. A common misconception is that the y-axis should always start at zero, but doing so can sometimes obscure substantively important variation.

⁷ Useful devices include rows reporting the number of observations, fixed effects and other model features; informative notes explaining the sample, units, uncertainty estimates and abbreviations; and variable labels that make clear what larger values represent, such as "Support for redistribution (higher = more support)".

When graphs or panels invite direct comparison, use common scales where appropriate and make any differences in scale explicit.

Third, removing unnecessary borders, heavy gridlines and decorative effects often strengthens a graph by directing attention towards the data. This is closely related to Edward Tufte's concept of the data-ink ratio: the share of a graphic's ink devoted to the non-redundant representation of data. I would use this as a reason to question every graphical element, while retaining only those that help interpretation. My personal preference is to remove both borders and grid lines.

Beyond these general principles, a few practical details make a substantial difference to how graphs look in the finished document. I generally recommend wide landscape proportions, although a square format sometimes suits the content better. The dimensions should accommodate the data and labels while allowing the figure to fill as much as possible of the intended text or column width without taking up unnecessary vertical space. Remember, however, that the relationship between width and height also affects the apparent steepness of trends, so proportions should be judged by how clearly they show the data as well as by their appearance.

The typography should be consistent across figures. Matching the typeface of the body text is a good starting point, although a clear sans-serif typeface can also work well alongside serif body text. Font sizes should be judged at the figure's final size in the document. For precise control over font sizes, specify them in points in your statistical software, set the graph's physical dimensions to match its intended size in the document, and insert it without resizing. Markers, line styles and colours should likewise represent the same groups consistently across graphs. Choose colours that work for readers with colour-vision deficiencies and in greyscale, reinforcing distinctions with markers, line styles or direct labels and using an ordered progression in lightness for sequential colour scales.

Last but not least, make sure to export graphs in a vector format, such as PDF or SVG, whenever possible. This keeps text and lines sharp when the figure is resized. If a raster format is necessary, use sufficient resolution at the size at which the figure will appear in the document.

The obvious

For those who do not want to read the whole text, here is a list of the ten most disturbing typographical sins that academics tend to conduct. Avoid these you will be fine.

- Do not underline.
- Do not combine indentation and vertical space between paragraphs.
- Do not doublespace. Especially not tables and footnotes.
- Keep the measure (line width) reasonable, even if your margins look wide.
- Avoid multi-line captions and titles. Place necessary information in notes.
- Do not indent the first paragraph after a section.
- Do not use the wrong characters.
- Give enough care to your figures and tables.

Small typographic dictionary

- ascender** The portion of a lower-case letter which extends above the mean line.
- descender** The portion of a letter which extends below the baseline.
- doublespacing** Vaguely defined concept, usually referring to a line-height of either 2 em or approximately 2.4 em.
- font** A variant of a typeface. So for example, Arial Black and Arial Narrow are two fonts within the same typeface.
- kerning** Adjustment of the spacing between specific pairs of letters.
- leading** Historically the distance between two lines of text, but now usually synonymous with the lineheight.
- legibility** How easy it is read and recognize individual characters.
- ligature** A combination of letters represented by a single glyph.
- mean line** An imaginary line which runs along the top of non-ascending lower-case letters.
- measure** The width of a block of text, measured in physical units or characters per line.
- orphan** Jargon for a paragraph-opening line that appears by itself at the bottom of a page or column.
- readability** How easy it is to read blocs of text.
- tracking** Uniform adjustment of the spacing between letters.
- typeface** A family of fonts, such as Arial or Adobe Garamond.
- widow** Jargon for a paragraph-ending line which falls at the beginning of a new page or column, and sometimes also used about paragraph-ending lines consisting of only one word.
- x-height** The distance between the baseline and the mean line, which typically is the same as the height of lower-case x.