Improving PDF readability of scientific papers on computer screens

Khaled Moustafa[i]

Abstract

Reading scientific papers in PDF-columned formats on computer screens is sometimes daunting and unfriendly. This is mainly due to the fact that PDF columned-texts are often truncated from the top or the bottom of pages so that readers have to scroll up-down repeatedly in the same page to get the whole text read. To make the reading process as smooth as possible, PDF viewers should be designed as responsive tools with responsive layouts that should automatically adapt the amount of readable text to the sizes of the displaying monitors so that readers will not waste time in repetitive vertical and horizontal scrolling movements within the same page. Here, I discuss such a problem-solving proposal that could be implemented in PDF viewers to improve the readability of PDF-columned texts and to make the reading process as flexible and painless as possible.

In order to improve text reading and enrich text composition in digital displaying platforms, most publications in scientific journals are currently written in different formats, mainly HTML (hypertext markup language) and PDF (portable document format)[1]. However, whatever the document format, several factors influence the legibility and reading speed and performance, such as the font size and interline spacing (Kwon et al. 2013), the print size and distance (Legge and Bigelow 2011), the letter-stroke boldness (Bernard et al. 2013), and the displaying screen sizes and their resolutions. Other factors such as the intra- and inter-word spacing (Slattery and Rayner 2013) (Hsu and Huang 2000) (Chung 2004), the crowding and eccentricity (Pelli et al. 2007), the visual span (Risse 2014) and the prevailing display conditions such as the character shape, color, luminous contrast, and text layout have also been reported (Nes 1984). For a good text displaying on computer screens, large interline spacing of 140% of the type size and 60-75 characters per line are generally recommended (Carter et al. 2014). The size of displaying screen is, indeed, an important parameter for onscreen reading and online tasks. Small screen users, for example, seem to be 50% less effective than users of large screens in completing the same tasks (Jones et al. 1999). In fact, users of small screens have to make substantial numbers of horizontal and vertical scrolling movements compared with large screens where much less movements are required to complete the same tasks.
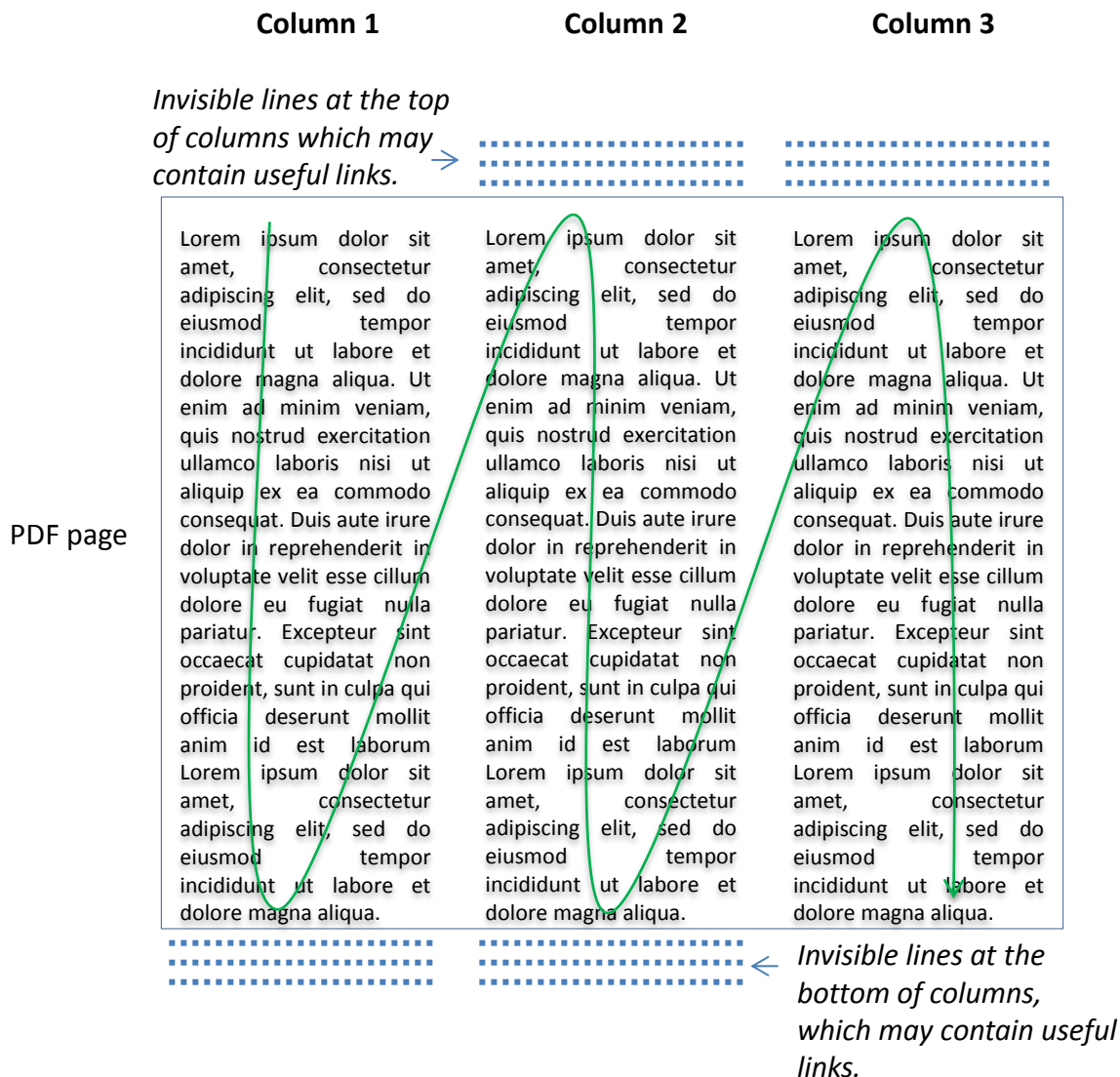
---

[1] https://acrobat.adobe.com/us/en/products/about-adobe-pdf.html, Accessed 23-November-2015

Most of us have experienced such difficulties to read digitalized texts in PDF-columned formats on computer small screens that make the reading process sometimes daunting and unfriendly because some lines at the top or the bottom of PDF pages are invisible so that readers need to scroll up-down repeatedly in the same page to read the hidden upper and lower lines in each PDF column. Scrolling horizontally or zooming in or out also are sometimes required when the reading frame on the displaying screen is not optimal, making the reading process annoying and time consuming. If an article is composed, for example, of more than one column per page, the reader will need to scroll down-up-down many times within the same page (Fig 1); the first column first then the second column second and so on to get the whole content read. If an article is composed of 5 pages of two columns per page, the reader will need to scroll down-up-down 3 times per page (or 15 times per five pages). This will be troublesome for longer articles of 3 or 4 columns per page, introducing distraction risks as readers should move up/down and right/left repeatedly.

Although PDF displaying options already offer a reading mode or a full screen option, these features, however, fit much more with large screen sizes than with small or medium-sized screens; large screens keep text sizes clear enough without forcing readers to scroll up/down or to zoom in/out compared with small screens. In other words, on large screens (i.e. ≥ 24 inches), the contents of columns in standard A4 pages are readable relatively easily since the text is visible in its entirety, page-per-page, so that readers are not obliged to scroll frequently within the same page. In small screens, on the other hand, the number of non-displayed lines at the top and/or the bottom of pages imposes on readers to make zigzagging repetitive movements within each page (Fig. 1) because the column heights are larger than the visible reading areas. Moreover, the orientation of the medium (paper vs. screen) is not always compatible or optimal for reading long texts. Although a page orientation in landscape format is sometimes available, but this format is rarely used in PDF documents. Landscape formats are mostly used in PowerPoint presentations but much less in PDF or text documents. As a result, the standard A4 page format (210 mm x 297 mm) seems to be inappropriate for a smooth and fluid reading of text organized in columns on small and medium-sized displaying monitors.

To fix these issues, genuine solutions should be developed to improve human/computer interactions and to enhance the reading process of PDF columned-texts. In this regard, three options might be considered, each of them with its advantages/disadvantages. The first option is to write texts as one column only. The second is to replace the A4 page format with smaller formats, for example, A5 (148 mm x 210 mm) or A6 (105 mm x 148 mm). The third option, and probably the most effective one, is to make the PDF viewers as responsive tools (with responsive layouts) to adapt the viewable text areas to the displaying screen dimensions to display texts as readable chunks at once. In other words, PDF pages should display the number of lines (or paragraphs) that readers can read at once without the need to scroll up or down

Column 1 **Column 2** **Column 3**

Invisible lines at the top of columns which may contain useful links.

PDF page

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

Invisible lines at the bottom of columns, which may contain useful links.

Figure 1. An example of three-columned PDF page where readers have to scroll vertically many times to read the whole content of each PDF colum. In most current vizualizing PDF viewers, some lines are usually invisible on digital screens unless to scrol down and up. When scolling down, however, lines at the top of the subsequent columns (column 2, 3, etc) become in turn invisible. Readers are obliged to scroll up again to get them read and so on, making the onscreen reading process disturbing and long, particularyl for people with disability when for example the invisibile lines at the top or the bottom of pages contain links for further information (readers should scroll up and/or down to find them).

No invisible lines at the top of PDF columns

Column 1	**Column 2**	**Column 3**
Lorem ipsum dolor sit amet, fed uop reaipsom consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore.	Lorem ipsum dolor sit amet, hera na consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	Lorem ipsum dolor sit amet, herea ola fop consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore .

PDF page

No invisible lines at the bottom of PDF columns

Figure 2. An example of the output of potential solution to optimize reading of PDF columned-texts would be to design responsive PDF viewers that adapt the amount of text proportionally to the dimensions of displaying screens, so no hidden lines would remain at the top or the bottom of each column. As such, readers will not be obliged to scroll vertically several times in the same page. Such a feature would allow reading PDF-columned texts smoothly from up-to-down and left-to-right without wasting time in repetitive vertical and horizontal scrolling movements.

repeatedly within the same page. A displaying format limited to one paragraph at a time was previously reported to improve the accuracy and reading performance (Creed et al. 1988), which reinforces the importance of the proposal note discussed here.

For the first option (writing text in one column only), this is already the case for HTML files, so this option would not offer many advantages over HTML formats currently used, except for PDF annotation tools that are currently available in PDF but not for HTML files (Laarni et al. 2004). The second option consisting of potential replacement of the standard A4 page format with smaller formats (i.e. A6 or A5) would be a good alternative choice, at least for online reading. However, for print or paper versions, A6 and A5 formats may present some inconveniences, relating chiefly to print layouts, which are much more adapted to the A4 format than for A5 or A6 formats in the scientific world. Such an option may, thus, need deep changes at paper manufacturing industry which is highly conservative and reluctant to transformation. Moreover, since the A4 format was standardized since a long time ago (in 1922 by the German Standardization Institute, DIN: http://www.din.de), and as it is still the dominant format for office work, the readership would also be reluctant to change. Alternatively, printing more than one text page per paper to reduce the print cost could also be considered. Nonetheless, when printing is not required (i.e. reading online only), adopting smaller page sizes for PDF files would be a worthy option to consider. However, for long texts, A5 or A6 formats would increase the number of pages per file and, therefore, increase the print costs, but since it concerns the displaying on screens only, this issue should not matter much given that printing is usually not required.

The third, and maybe the best, option in my view to improve reading on digitalized screens and to enhance the interaction human/machine would be the development of responsive PDF viewers. That is, making the PDF viewers as adaptable tools or layouts that should automatically adapt the viewable text area to the size of the viewing digital environment (i.e. computer screens). As such, only the amount of text that could be read in one block is displayed at a time without the need to scroll vertically or horizontally within the same page. In other words, the content of PDF pages should fit automatically with the dimensions of the displaying screen. In such a case, the PDF functions should be able to recognize the screen resolution (vertical and horizontal dimensions) and then displaying only the amount of text that is readable within the visible reading frame. Doing so, the reader will not need to waste time in scrolling down-up-down in each page (Fig 2), but s/he will need to scroll only between pages in a comparable way to the reading process on paper. Such an option is already available for reading texts on many tablets or smartphones models, but it should also be of high interest to be available for PDF files opened on desktop and laptop screens to ensure smoother navigation and reading with minimal scrolling or zooming movements. Toward such an objective, the National Center for Biotechnology Information (NCBI) has already developed a new presentation style called PubReader to facilitate reading of articles archived in PubMed Central

repository (Lipman 2012). PubReader breaks the text into multiple columns and pages to improve the readability and navigation in web browsers in a similar way to printed papers, whatever the displaying platform (desktop, laptop, tablets etc.). Similar options to the online PubReader functions should also be developed for offline PDF viewers to make the reading process as friendly as possible while saving important scrolling time. The PubReader view mode in browsers is, in fact, a tangible example that PDF viewers could also offer as responsive tools adaptable to multiple reading platforms. Another important aspect of displaying text proportionally to the screen dimensions through responsive layouts is to improve the aesthetic effects that seem to enhance the reading process (Strebe 2015).

Whatever the screen sizes, however, the process of reading on printed papers remain more efficient than reading on digitalized screen because some features of digitized texts may impede some aspects of the reading comprehension (Mangen et al. 2013). Both speed and accuracy of proof-reading texts on screen seem to be impaired compared with reading on printed papers (Wright and Lickorish 1983) (Picking 1997) (Wilkinson and Robinshaw 1987), although reading performance on screen has been substantially improved, for example, by increasing screen resolution or font sizes and using new fonts (i.e. Clear Type). However, under time pressures (for e.g. exam conditions), the reading comprehension still score higher on paper than on screen (Ackerman and Lauterman 2012). Learning on paper also seems to be more efficient than on screen, and self-assessments of knowledge on paper are better as well (Ackerman and Lauterman 2012).

Despite these advantages of reading and learning on paper, the trend for onscreen reading is progressively dominant, particularly with mobile devices and web services offered exclusively online. For some reading tasks, however, reading from a hard copy will be preferred, and for others, reading from an electronic support would be preferred (Hendry 1995). Onscreen reading tools should thus be optimized to maximize the reading output for people with or without disabilities. To improve accessibility and writing on screens for disabled people, many methods have already been suggested, such as the EasyWrite system (Godinho et al. 2014, Godinho et al. 2015), the trackball text entry (Wobbrock and Myers 2006), the reading by colors (Irlen 2005), and the No-Look Notes as an eyes-free text entry system that uses multi-touch input and audio output (Bonner et al. 2010). Companies that do not take into account the needs of users with disabilities may incur some risks to their legal position and reduce their business success (Curran et al. 2007). The opinion piece discussed here to introduce responsive PDF text layouts would facilitate the access to web links in webpages because it implies that the links on pages would be easily visible in the viewable reading areas so that disabled or impaired people would view them straightforwardly in the visible text area and spend less time and movements to access them.

In addition to automatic visual and aesthetic adjustments, automatic acoustical adjustments are also important in the digital era to improve the interaction between human

and digital devices. Recent applications, for example, were proposed to automatically adjust smartphones' ringtones to improve the automation of digital devices and prevent interruptions and performance impairment for people exercising regular tasks (Moustafa 2014).

In summary, a good responsive design of PDF tools that adjust the text layout to the displaying environment would provide an optimal viewing and reading practice. Since the visual representation of information is an effective communicative practice (Snyder 2014) (Strebe 2015), using smooth and proportion-based grids and resizable text and images would make the reading of PDF files as a pleasurable experience. Present and future technological advancements in responsive web design would allow specialized developers of PDF programs to incorporate such a desirable customer/reader back-end service.

Disclosure: None

References

Ackerman, R. and T. Lauterman (2012). "Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure." Computers in Human Behavior **28**(5): 1816-1828.

Bernard, J. B., G. Kumar, J. Junge and S. T. Chung (2013). "The effect of letter-stroke boldness on reading speed in central and peripheral vision." Vision Res **84**: 33-42.

Bonner, M., J. Brudvik, G. Abowd and W. K. Edwards (2010). No-Look Notes: Accessible Eyes-Free Multi-touch Text Entry. Pervasive Computing. P. Floréen, A. Krüger and M. Spasojevic, Springer Berlin Heidelberg. **6030:** 409-426.

Carter, R., P. B. Meggs, B. Day, S. Maxa and M. Sanders (2014). Typographic Design: Form and Communication, Wiley.

Chung, S. T. (2004). "Reading speed benefits from increased vertical word spacing in normal peripheral vision." Optom Vis Sci **81**(7): 525-535.

Creed, A., I. Dennis and S. Newstead (1988). "Effects of display format on proof-reading with VDUs." Behaviour & Information Technology **7**(4): 467-478.

Curran, K., N. Walters and D. Robinson (2007). "Investigating the problems faced by older adults and people with disabilities in online environments." Behaviour & Information Technology **26**(6): 447-453.

Godinho, R., P. A. Condado, M. Zacarias and F. G. Lobo (2014). "Improving accessibility of mobile devices with EasyWrite." Behaviour & Information Technology **34**(2): 135-150.

Godinho, R., P. A. Condado, M. Zacarias and F. G. Lobo (2015). "Improving accessibility of mobile devices with EasyWrite." Behaviour & Information Technology **34**(2): 135-150.

Hendry, D. (1995). "Breakdowns in writing intentions when simultaneously deploying SGML-marked texts in hard copy and electronic copy." Behaviour & Information Technology **14**(2): 80-92.

Hsu, S. H. and K. C. Huang (2000). "Interword spacing in Chinese text layout." Percept Mot Skills **91**(2): 355-365.

Irlen, H. (2005). Reading by the colors: Overcoming dyslexia and other reading disabilities through the Irlen method, Penguin.

Jones, M., G. Marsden, N. Mohd-Nasir, K. Boone and G. Buchanan (1999). "Improving Web interaction on small displays." Computer Networks **31**(11–16): 1129-1137.

Kwon, B. C., J. S. Yi and Y. Zhu (2013). "ReadingMate: the effect of the content stabilizing technique, font size, and interline spacing on the letter-counting task performance of treadmill runners." Hum Factors **55**(6): 1101-1111.

Laarni, J., J. Simola, I. Kojo and N. Risto (2004). "Reading vertical text from a computer screen." Behaviour & information technology **23**(2): 75-82.

Legge, G. E. and C. A. Bigelow (2011). "Does print size matter for reading? A review of findings from vision science and typography." J Vis **11**(5).

Lipman, D. (2012). "The PubReader View: A New Way to Read Articles in PMC." NLM Tech Bull. 2012 Nov-Dec;(389):e7.

Mangen, A., B. R. Walgermo and K. Brønnick (2013). "Reading linear texts on paper versus computer screen: Effects on reading comprehension." International Journal of Educational Research **58**: 61-68.

Moustafa, K. (2014). "A smart application for smartphone: a proposal to reduce noise pollution for people having regular tasks." Front Public Health **2**: 122. DOI: 10.3389/fpubh.2014.00122.

Nes, F. V. (1984). "Limits of visual perception in the technology of visual display terminals." Behaviour & Information Technology **3**(4): 371-377.

Pelli, D. G., K. A. Tillman, J. Freeman, M. Su, T. D. Berger and N. J. Majaj (2007). "Crowding and eccentricity determine reading rate." J Vis **7**(2): 20 21-36.

Picking, R. (1997). "Reading music from screens vs paper." Behaviour & Information Technology **16**(2): 72-78.

Risse, S. (2014). "Effects of visual span on reading speed and parafoveal processing in eye movements during sentence reading." J Vis **14**(8): 11.

Slattery, T. J. and K. Rayner (2013). "Effects of intraword and interword spacing on eye movements during reading: exploring the optimal use of space in a line of text." Atten Percept Psychophys **75**(6): 1275-1292.

Snyder, J. (2014). "Visual representation of information as communicative practice." Journal of the Association for Information Science and Technology **65**(11): 2233-2247.

Strebe, R. (2015). "Aesthetics on the web: effects on approach and avoidance behaviour." <u>Behaviour & Information Technology</u>: 1-17.

Wilkinson, R. and H. M. Robinshaw (1987). "Proof-reading: VDU and paper text compared for speed, accuracy and fatigue." <u>Behaviour & Information Technology</u> **6**(2): 125-133.

Wobbrock, J. and B. Myers (2006). <u>Trackball text entry for people with motor impairments</u>. Proceedings of the SIGCHI conference on Human Factors in computing systems, ACM.

Wright, P. and A. Lickorish (1983). "Proof-reading texts on screen and paper." <u>Behaviour & Information Technology</u> **2**(3): 227-235.

---

[i] Email: khaled.moustafa@gmail.com