The Effect of Font Type on Screen Readability by People with Dyslexia

LUZ RELLO and RICARDO BAEZA-YATES, Web Research Group, DTIC,
Universitat Pompeu Fabra, Barcelona, Spain

15

Around 10% of the people have dyslexia, a neurological disability that impairs a person's ability to read and write. There is evidence that the presentation of the text has a significant effect on a text's accessibility for people with dyslexia. However, to the best of our knowledge, there are no experiments that objectively measure the impact of the typeface (font) on screen reading performance. In this article, we present the first experiment that uses eye-tracking to measure the effect of typeface on reading speed. Using a mixed between-within subject design, 97 subjects (48 with dyslexia) read 12 texts with 12 different fonts. Font types have an impact on readability for people with and without dyslexia. For the tested fonts, *sans serif*, *monospaced,* and *roman* font styles significantly improved the reading performance over *serif*, *proportional,* and *italic* fonts. On the basis of our results, we recommend a set of more accessible fonts for people with and without dyslexia.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Screen design, style guides*; K.4.2 [**Computers and Society**]: Social Issues—*Assistive technologies for persons with disabilities*

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Dyslexia, learning disability, best practices, web accessibility, typeface, font, readability, legibility, eye-tracking

ACM Reference Format:
Luz Rello and Ricardo Baeza-Yates. 2016. The effect of font type on screen readability by people with Dyslexia. ACM Trans. Access. Comput. 8, 4, Article 15 (May 2016), 33 pages.
DOI: http://dx.doi.org/10.1145/2897736

1. INTRODUCTION

Worldwide, around 15–20% of the population has a language-based learning disability and likely 70–80% of these have dyslexia [International Dyslexia Association 2011b], a neurological disability which impairs a person's ability to read and write. Previous research has shown that text presentation can be an important factor regarding the reading performance of people with dyslexia [Gregor and Newell 2000; Gregor et al. 2003; Kurniawan and Conroy 2006; Rello 2014].

The research of the first author was partially funded by the Catalonian FI scholarship program and by Grant TIN2009-14560-C03-01 from the Ministry of Science and Innovation of Spain.

Authors' addresses: L. Rello (currently at), Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University, 216 Craig Sgt., Pittsburgh, PA 15213 USA; email: luzrello@acm.org; R. Baeza-Yates, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Catalonia, Spain; email: rbaeza@acm.org.

DOI: http://dx.doi.org/10.1145/2897736

Indeed, any digital text has to be written using one or several certain typefaces or font types.[1] Although the selection of font types is crucial in the text design process, empirical analyses of reading performance of people with dyslexia has focused more on font size [O'Brien et al. 2005; Dickinson et al. 2002; Rello et al. 2015] rather than on font type.

In this article, we present the first study that measures the impact of the typeface on the on-screen reading performance of people with dyslexia using eye tracking, as well as asking them their personal preferences. We limit our study to on-screen fonts because reading online content is the most frequent case today and also because this allows us to use standard eye-tracking techniques. We used a mixed between-within subject design where 97 subjects (48 subjects with dyslexia) read 12 texts with 12 different fonts. On the basis of our results, we present recommendations for font styles and a set of more accessible fonts for people with dyslexia. The main contributions of this study are:

—Font types have a significant impact on the on-screen readability of people with and without dyslexia.
—What is good for people with dyslexia regarding font types is also good for people without dyslexia.
—For the tested fonts, good fonts for people with dyslexia are *Helvetica*, *Courier*, *Arial*, *Verdana,* and *Computer Modern Unicode*, taking into consideration reading performance and subjective preferences. On the contrary, *Arial It.* should be avoided since it decreases on-screen readability.
—For the tested fonts, we found that *sans serif* and *roman* font types increased the on-screen reading performance of our participants, whereas *italic* fonts did the opposite.

The next section focuses on dyslexia, whereas Section 3 reviews related work. Section 4 explains the experimental methodology, and Section 5 presents the results, which are discussed in Section 6. In Section 7, we explain the limitations of our study, whereas in Section 8 we derive recommendations for dyslexic-friendly font types and mention future lines of research.

This article is an extended version of a paper originally presented at ASSETS [Rello and Baeza-Yates 2013]. Whereas the ASSETS paper focuses on the impact of font face on the readability of people with dyslexia only, this article adds results for a control group of 49 subjects without dyslexia. It also includes results for a third eye-tracking readability measure, the comparison of both groups, and a section dedicated to the limitations of the study.

2. DYSLEXIA

Dyslexia is a *hidden* disability. People with dyslexia, despite adequate intelligence and education, struggle with the act of reading in that they read more slowly and are more prone to reading errors such as word misidentifications. Dyslexia is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede the growth of vocabulary and background knowledge [International Dyslexia Association 2011a; Lyon 1995; Lyon et al. 2003].

[1]Although the correct term is typeface, we use font as a synonym as is commonly used outside the world of typography.

In some literature, dyslexia is referred to as a specific reading disability [Vellutino et al. 2004] and dysgraphia as its writing manifestation only [Romani et al. 1999].[2] However, we follow the standard definitions of the *International Statistical Classification of Diseases (ICD-10)* [World Health Organization 1993] and the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* [American Psychiatric Association 2000], where dyslexia is listed as a reading and spelling disorder *(ICD-10)* or a reading disorder and a disorder of written expression *(DSM-IV)*.

Popularly, dyslexia is identified with its superficial consequences, such as writing problems like letter reversals. However, a considerable number of studies confirm the biological foundations of dyslexia, with the exception of acquired dyslexia [Vellutino et al. 2004]. Moreover, the most frequent way to detect a child with dyslexia is by low performance at school [Carrillo et al. 2011]. In Spain, approximately four out of six cases of school failure are related to dyslexia [FEDIS 2008].[3]

Furthermore, dyslexia is frequent. Despite its universal neurocognitive basis, dyslexia manifestations are variable and culture-specific [Goulandris 2003]. Depending on the language, estimations on the prevalence of dyslexia differ. The National Academy of Sciences [Interagency Commission on Learning Disabilities 1987] states that 10–17.5% of the US population has dyslexia. The model of Shaywitz et al. [1992] predicts that 10.8% of English-speaking children have dyslexia, whereas in Katusic et al. [2001] the rates varied from 5.3% to 11.8% depending on the formula used. Brunswick [2010] estimates 10% for English and 3.5% for Italian. Data on the prevalence of dyslexia in Spanish speakers are much more scarce: Galván Gómez [2010] reports a 7.5% among school children in Perú; Carrillo et al. [2011] found that 11.8% of the school children examined in Murcia (Spain) exhibited difficulties associated with dyslexia, and Jiménez et al. [2009] report an 8.6% for a similar population in the Canary Islands (Spain).

The frequency, the universal neurocognitive basis of dyslexia, and its relationship with school failure are the main motivations of our study.

3. RELATED WORK

The relationship between fonts and dyslexia has drawn the attention of many fields, such as psychology, the arts, and accessibility. We divide related work into fonts recommended for people with dyslexia, fonts designed for this target group, and related user studies.

3.1. Font Recommendations

Most of the font recommendations come from associations for people with dyslexia, and they agree in using *sans-serif* fonts [British Dyslexia Association 2012; Evett and Brown 2005].

The British Dyslexia Association recommends using *Arial*, *Comic Sans,* or, as alternatives to these, *Verdana*, *Tahoma*, *Century Gothic*, and *Trebuchet* [British Dyslexia Association 2012]. However, the website does not disclose the basis on which these recommendations are made. For readers with low vision, as well as for readers with dyslexia, Evett and Brown [2005] also recommend using *Arial* and *Comic Sans*. Lockley

[2]Dysgraphia refers to a writing disorder associated with the motor skills involved in writing, handwriting, and sequencing, but also orthographic coding [Berninger and Wolf 2009]. It is comorbid with dyslexia [Nicolson and Fawcett 2011].

[3]The percentage of school failure is calculated by the number of students who leave school before finishing secondary education (high school). While the average of school failure in the European Union is around 15%, Spain has around 25–30% of school failure, with 31% in 2010 [Enguita et al. 2010].

[2002] recommended avoiding italics and fancy fonts, which are particularly difficult for a reader with dyslexia, and also points to *Arial* as the preferred font.

The only recommendation for *serif* fonts came from the International Dyslexia Centre [Hornsby 1986] and that was for *Times New Roman*. According to Ability Net [2013], *Courier* is easier to read by people with dyslexia because it is monospaced.

In the *Web Content Accessibility Guidelines* (WCAG) [Caldwell et al. 2008], dyslexia is treated as part of a diverse group of cognitive disabilities, and they do not propose any specific guidelines about font types for people with dyslexia.

Surprisingly, none of the font types recommended by the dyslexia organizations mentioned were ever designed specifically for readers with dyslexia.

3.2. Fonts Designed for People with Dyslexia

We found five fonts designed for people with dyslexia: *Sylexiad* [Hillier 2008], *Dyslexie* [De Leeuw 2010], *Read Regular*,[4] *Lexie Readable*,[5] and *OpenDyslexic*.[6] The four fonts have in common that letters are more differentiated compared to regular fonts. For example, the shape of the letter ‘b’ is not a mirror image of ‘d’. Although most of these fonts are free for personal or charity use, we chose to study *Open Dyslexic* (both roman and italic styles) because it is the only one open sourced and hence free for any kind of use.

3.3. User Studies

Most of the previous work with regular readers without dyslexia applies to the two most common fonts used on screen and in printed texts, *Arial* and *Times*, respectively [Chapman 2011]. Taking into account regular readers, Paterson and Tinker [1932] tested 10 different fonts, serif and sans serif typefaces, as well as the monospaced *American Typewriter* and *Cloister Black,* which is densely decorated. Only the last two fonts mentioned showed a significant decrease in reading speed. Later, Boyarski et al. [1998] compared *Times*, a serif font designed for printed text, with *Georgia*, a serif font designed for the screen. Users preferred *Georgia* as more pleasing and easier to read. In a second test, they compared *Georgia* with *Verdana*, a sans serif face designed for on-screen use. Users expressed subjective preference for *Verdana*, but they performed better reading *Georgia*. Bernard et al. [2003] compared two fonts—*Arial* and *Times*—and two font sizes—10 and 12 points—with 35 participants. The experiment used the same dependent measures: 10-point *Arial* typeface again was read more slowly than the other conditions, and the 12-point *Arial* typeface was preferred to the other typefaces.

Regarding readers with dyslexia, there are several studies about the effect of different parameters of text presentation, such as font and background colors [Rello et al. 2012], font size [O’Brien et al. 2005; Rello et al. 2013], or letter spacing [Zorzi et al. 2012].

The closest work to ours is a study with people with dyslexia [De Leeuw 2010] that compared *Arial* and *Dyslexie*. The researchers conducted a word reading test with 21 students with dyslexia (Dutch One Minute Test). *Dyslexie* did not lead to faster reading, but could help with some dyslexic-related errors in Dutch. In Sykes [2008], text design for people with dyslexia is explored with a qualitative study with 11 students using class observations, interviews, and questionnaires. In some tasks, the participants were asked to choose the font they preferred, but no analysis of the chosen fonts is presented.

[4] http://www.readregular.com/.
[5] http://www.k-type.com/fonts/lexie-readable/.
[6] http://opendyslexic.org/.

This is Arial	This is Myriad
This is Arial It.	This is OpenDyslexic
This is Computer Modern	*This is OpenDyslexic It.*
This is Courier	This is Times
This is Garamond	*This is Times It.*
This is Helvetica	This is Verdana

Fig. 1. Font types used in the experiment.

3.4. What Is Missing?

What is missing is a sound investigation into the effect of the most frequently used fonts on reading performance. Our experiment advances previous work by providing this evidence via quantitative data from eye-tracking measurements.

In addition, by testing 12 different fonts with 48 participants and a control group of 49 participants, we compare a greater number of font types with a larger number of participants than do previous studies. We selected the fonts on the basis of their popularity and frequency of use on the Web.

4. METHODOLOGY

To study the effect of font type on the readability and comprehensibility of texts on the screen, we conducted an experiment in which 97 participants (48 with dyslexia) had to read 12 comparable texts with varying font types. Readability and comprehensibility were analyzed via eye-tracking and comprehension tests, respectively, using the latter as a control variable. The participants' preference was gathered via questionnaires.

4.1. Design

In Table I, we show a summary of the experimental design of our experiment.

4.1.1. Independent Variables. In our experimental design, *Font Type* served as an *independent variable* with 12 levels: *Arial*, *Arial Italic*, *Computer Modern Unicode (CMU)*, *Courier*, *Garamond*, *Helvetica*, *Myriad*, *OpenDyslexic*, *OpenDyslexic Italic*, *Times*, *Times Italic*, and *Verdana* (See Figure 1). We use for brevity *OpenDys* for the corresponding fonts.

We chose to study *Arial* and *Times* because they are the most common fonts used on screen and in printed texts, respectively [Chapman 2011]. *OpenDyslexic* was selected because is a free font type designed specifically for people with dyslexia and *Verdana* because it is a recommended font for this target group [British Dyslexia Association 2012]. We chose *Courier* because it is the most common example of a monospaced font [Chapman 2011]. *Helvetica* and *Myriad* were chosen for being broadly used in graphic design and for being the typefaces of choice of Microsoft and Apple, respectively. We chose *Garamond* because it is claimed to have strong legibility for printed materials [Chapman 2011], and we selected *CMU* because it is widely used in scientific publishing, is the default of the typesetting program TeX, and is a free font supporting many languages [Knuth 1986].

We did not used a fully factorial design for our experiment because fonts cannot be altered across all the factors that we consider. For example, there is no monospaced version of *Arial*. Hence, we decided to test the most frequently used fonts to maximize the usefulness of our study. Therefore, we performed a constrained experiment taking

Table I. Methodological Summary for the Experiment

Experiment		
Design	Mixed between/within subject	
Independent Variables	*Font*	*Arial, Arial Italic* *Computer Modern Unicode (CMU)* *Courier* *Garamond* *Helvetica* *Myriad* *OpenDyslexic, OpenDyslexic Italic* *Times, Times Italic* *Verdana*
	[±Italic]	[+Italic] *Arial It.*, *OpenDys. It.*, *Times It.* [−Italic] *Arial*, *OpenDys.*, *Times*
	[±Serif]	[+Serif] *CMU*, *Garamond*, *Times* [−Serif] *Arial*, *Helvetica*, *Myriad*, *Verdana*
	[±Monospace]	[+Monospace] *Courier* [−Monospace] *CMU*, *Garamond*, *Times*
	[±Dyslexic]	[+ Dyslexic] *OpenDyslexic* [− Dyslexic] *Arial*, *Helvetica*, *Myriad*, *Verdana*
	[±Dyslexic It.]	[+ Dyslexic It.] *OpenDyslexic It.* [− Dyslexic It.] *Arial it*
Dependent Variables	*Reading Time* *Fixation Duration* *Number of Fixations*	(objective readability)
	Preference Rating	(subjective preferences)
Control Variable	*Comprehension Score*	(objective comprehensibility)
Participants	Group D (48 participants)	22 female, 26 male **Age:** range from 11 to 50 ($\bar{x} = 20.96$, $s = 9.98$) **Education:** high school (26), university (19), no higher education (3)
	Group N (49 participants)	28 female, 21 male **Age:** range from 11 to 54 ($\bar{x} = 29.20$, $s = 9.03$) **Education:** high school (17), university (27), no higher education (5)
Materials	Texts	12 story beginnings
	Text Presentation	
	Comprehension Quest.	12 literal items (1 item/text)
	Preferences Quest.	12 items (1 item/condition)
Equipment	Eye tracker Tobii 1750	
Procedure	**Steps:** Instructions, demographic questionnaire, reading task (× 12), comprehension questionnaire (× 12), preferences questionnaire (× 12)	

fonts that are frequently used, recommended, or designed for people with dyslexia. We also made sure that the fonts cover variations of essential font characteristics:

—**[±Italic]** served as an independent variable with two values: [+Italic] denotes the condition where the text was presented using an italic type (i.e., a cursive typeface), and [−Italic] denotes the condition where the text was presented in a roman type. We compare three [±Italic] types: *Arial*, *OpenDyslexic*, and *Times*. Even if earlier typography studies [Tinker 1965] suggest not presenting the text entirely in italics, we decided to include three italicized fonts because, to the best of our knowledge, the effect of italic text was not previously measured with participants with dyslexia using an eye-tracking methodology.

—**[±Serif]** served as independent variable with two values: [+Serif] denotes the condition where the text was presented with typefaces with serifs (i.e., small lines trailing from the edges of letters and symbols), and [−Serif] denotes the condition where the text used typefaces without serifs (*sans serif*). For the comparison, we used the three *serif* fonts of our set (*CMU*, *Garamond*, and *Times*) together with the four *sans serif* fonts (*Arial*, *Helvetica*, *Myriad*, and *Verdana*).

—**[±Monospace]** served as independent variable with two values: [+Monospace] denotes the condition where the text was presented using a monospaced type (i.e., a font whose letters and characters each occupy the same amount of horizontal space), and [−Monospace], where the text was presented using proportional fonts. We chose the most commonly used monospaced font, the *roman serif* font *Courier*, and we compare it with the rest of the *roman* and *serif* fonts that are *proportional*: *CMU*, *Garamond,* and *Times*.

—**[±Dyslexic]** served as an independent variable with two values: [+Dyslexic] denotes the condition where the text was presented using a font type which was specifically designed for people with dyslexia *OpenDyslexic*, and [−Dyslexic], where the text was presented using other fonts. Since *OpenDyslexic* is *sans serif* and *roman*, we use the rest of the *sans serif*, *roman* fonts (*Arial*, *Helvetica*, *Myriad*, and *Verdana*) for its comparison. To cover the italic variant *OpenDyslexic It.*, [±Dyslexic It.] served as independent variable with two values: [+Dyslexic It.], which correspond to *OpenDyslexic It.,* and [−Dyslexic It.], which correspond with the only *sans serif*, *italic* font we had in our set, *Arial it.*

We used a mixed between-within subject design, with one between-subjects factor with two levels (dyslexia or not dyslexia) and a within-subjects factor with 12 levels (type faces). Each participant read 12 different texts in 12 different fonts, hence, contributing to each condition (this leads to 66 pairwise comparisons). We counterbalanced texts and fonts to avoid experimental sequence effects. Therefore, the data with respect to text-font combinations were evenly distributed.

4.1.2. Dependent Variables. For quantifying readability, we used three *dependent measures*: *Reading Time*, *Fixation Duration,* and *Number of Fixations* extracted from the eye-tracking data. We used areas of interest that included the text only to determine when a participant has finished reading. That is, we only took into account data coming from the areas of interest. To control text comprehension of the texts, we use the *Comprehension Score* question as a *control variable*. To collect the participant preferences, we used subjective *Preference Ratings* through questionnaires.

—**Reading Time:** Shorter reading durations are preferred to longer ones since faster reading is related to more readable texts [Williams et al. 2003]. Therefore, we use *Reading Time*—that is, the time it takes to a participant to completely read one text—as a measure of readability.

—**Fixation Duration:** When reading a text, the eye does not move continuously over the text, but alternates saccades and visual fixations (i.e., jumps in short steps and rests over pieces of text). *Fixation Duration* denotes how long the eye rests on a single spot of the text. We use the average of the fixation durations as a metric for readability.
—**Number of Fixations:** The total number of fixations while reading a text.

Fixation Duration and *Number of Fixations* have been shown to be a valid indicator of readability because eye movement measures can be used to infer moment-to-moment cognitive processes in reading. Shorter fixations are associated with better readability, whereas a greater number of fixations and longer fixations can indicate that processing loads are greater [Rayner 1998]. For instance, people without dyslexia present more and longer fixations on low-frequency words than on high-frequency words [Inhoff and Rayner 1986; Just and Carpenter 1980; Raney and Rayner 1995; Rayner and Duffy 1986; Rayner and Raney 1996; Rello et al. 2013a]. On the other hand, it is not directly proportional to reading time because some people may fixate more often in or near the same piece of text (re-reading).

Three studies show why fixation duration and number of fixations are valid indicators also for people with dyslexia. First, Hyönä and Olson [1995] found that readers with dyslexia show the typical word frequency effect in which low-frequency words are fixated longer (fixation duration, number of fixations, and regressions) than high-frequency words.

Second, Pirozzolo and Rayner [1978] and Olson et al. [1983] found that when people with dyslexia were given a text appropriate for their reading level, their eye movements (fixations, saccades, and regressions) were much like those of normal readers at that particular age level.

Third, Rayner [1986] showed that the eye movements (fixation durations, saccade lengths, and the size of the perceptual span) of children without dyslexia shared the characteristics of readers with dyslexia when they were given a text that was too difficult for them.

—**Comprehension Score:** To check that the text was not only read but also understood, we used literal questions (i.e., questions that can be answered directly from the text). We used multiple-choice questions with three possible choices: one correct choice and two wrong choices. We use this comprehension question as a *control variable* to guarantee that the recordings analyzed in this study were valid. If the reader did not chose the correct answer, the corresponding text was discarded from the analysis.
—**Preference Ratings:** In addition, we asked the participants to provide their personal preferences. For each of the 12 text-font pairs, the participants rated on a 5-point Likert scale how much they liked the font type used in the text presentation (see example in Figure 3).

4.2. Participants

We had 48 people (22 female, 26 male) with a confirmed diagnosis of dyslexia taking part in the study (group D). Their ages ranged from 11 to 50 ($\bar{x} = 20.96$, $s = 9.98$), and they all had normal vision. All of them presented official clinical results to prove that dyslexia was diagnosed in an authorized center or hospital.[7] Except for three participants, all participants were attending school or high school (26 participants), or they were studying or had already finished university degrees (19 participants).

[7]In the Catalonian protocol of dyslexia diagnosis [Speech Therapy Association of Catalonia 2011], the different kinds of dyslexia extensively found in the literature are not considered.

El texto habla de: *'The text is about:'*
– Un sueño. *'A dream.'* – Un parque de atracciones. *'An amusement park.'* – Un helado de chocolate. *'A chocolate ice cream.'*

Fig. 2. Comprehension control question example.

We also had a control group (group N) composed of 49 people without dyslexia (28 female, 21 male) with an age ranging from 11 to 54 years old ($\bar{x} = 29.20$, $s = 9.03$). Except for five participants who had no higher education, the rest were either attending school or high school (17 participants) or were studying or had finished university degrees (27 participants).

4.3. Materials

4.3.1. Texts. All the texts used in the experiment meet the comparability requirements because they all share the parameters commonly used to compute readability [Saggion et al. 2015]. For instance, the Flesch Reading Ease score takes into consideration the number of words, sentences, and syllables [Flesch 1948], and the Coleman-Liau index uses the number of letters and number of sentences [Coleman and Liau 1975]. All the texts were extracted from the same book, *Impostores* (Impostors), by Lucas Sánchez [2012]. We chose this book because its structure (32 chapters) gave us the possibility of extracting similar texts. Each chapter of the book is an independent story, and it starts always by an introductory paragraph. Thus, we went through the book and selected the introductory paragraphs sharing the following characteristics:

(a) Same genre and same style.
(b) Same number of words (60 words). If the paragraph did not have that number of words, we slightly modified it to match the number of words.
(c) Similar word length, with an average length ranging from 4.92 to 5.87 letters.
(d) Absence of numerical expressions [Rello et al. 2013c], acronyms, and foreign words [Cuetos and Valle 1988] because people with dyslexia especially encounter problems with such words.

4.3.2. Text Presentation. Since the presentation of the text has an effect on the reading speed of people with dyslexia [Gregor and Newell 2000], we used the same layout for all the texts. They were left-justified, using a 14-point font size, and the column width did not exceeded 70 characters/column, as recommended by the British Association of Dyslexia [British Dyslexia Association 2012]. The color used was the most frequently used in the Web for text: black text on white background.

4.3.3. Comprehension Questionnaires. After each text, there was one literal comprehension control question. The order of the correct answer was counterbalanced. An example of one of these questions is given in Figure 2. The difficulty of the questions chosen was similar.

4.3.4. Preference Questionnaires. For each of the 12 text-font pairs, the participants rated each font on a 5-point Likert scale on how much they liked the font type used in the text presentation. An example of the items is given in Figure 3.

4.4. Equipment

The eye-tracker we used was the Tobii 1750 [Tobii Technology 2005], which has a 17-inch TFT monitor with a resolution of 1024×768 pixels. The time measurements

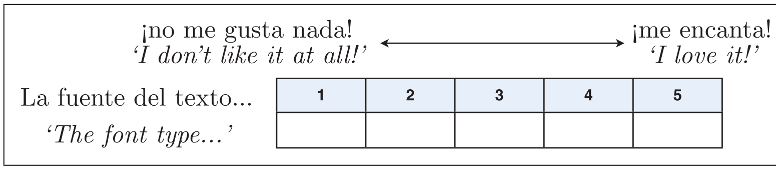

Fig. 3. Preference rating item.

of the eye-tracker have a precision of 0.02 seconds. Hence, all time values are given with an accuracy of two decimals. The eye-tracker was calibrated individually for each participant, and the light focus was always in the same position. The distance between the participant and the eye-tracker was constant (approximately 60 cm. or 24 in.) and the participant's angle to the screen (in front of it) was controlled by using a fixed chair. The light position (coming from the participant's left side) was also controlled. First, the participant was asked to sit comfortably for reading. Second, we performed the calibration of the eye-tracker.

4.5. Procedure

The sessions were conducted at the Universitat Pompeu Fabra and lasted around 20 minutes. Each session took place in a quiet room where only the interviewer (first author) was present, so that the participants could concentrate. Each participant performed the following three steps. First, we began with a questionnaire that was designed to collect demographic information. Second, the participants were given specific instructions. They were asked to read the 12 texts in silence and complete the comprehension control questions after each text. In answering the question they could not look back on the text. Their reading was recorded by the eye-tracker. Finally, each participant was asked to provide his or her preference ratings.

4.6. Statistical Analysis

We used the R Statistical Software 2.14.1 [R Development Core Team 2011] for our analysis, with the standard condition of $p < 0.05$ for significant results. We use the Shapiro-Wilk test for checking if data fit a normal distribution and the Levene test to check for homogeneity. Because our data were not normal nor homogeneous, we include the median and box plots for all our measures in addition to the mean and the standard deviation. For the same reason, to study the effects of the dependent variables, we used the two-way Friedman's nonparametric test for repeated measures plus a complete pairwise Wilcoxon rank sum post-hoc comparison test with a Bonferroni correction that includes the adjustment of the significance level. We used the same procedure to show effects of the conditions within groups, dividing the data for each group, and to study the effect of the second-level independent variables *Italics*, *Serif*, *Monospace*, and *Dyslexic*. To handle missing values in the repeated-measures statistical tests, we filled the gap with a "NA" (not available) value. In R, NA is a placeholder especially defined for this purpose.

In the post-hoc tests, we used the Bonferroni adjustment [Bonferroni 1936] because it is the most conservative approach in comparison with other adjustment methods such as the ones included by Holm [1979], Hochberg [1988], Hommel [1988], Benjamini and Hochberg [1995], and Benjamini and Yekutieli [2001]. If we remove the Bonferroni correction from the tests, we achieve much lower p-values. In R, the corresponding Bonferroni adjustment is divided by the number of tests performed. That is, the α value for each comparison is equal to α/n, where n is the number of tests.

Table II. Median, Mean, and Standard Deviation of *Reading Time* in Seconds for Groups N and D

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean ± SD*	*%*		*Median*	*Mean ± SD*	*%*
	Reading Time				*Reading Time*		
Arial	24.22	28.35 ± 12.39	100	Arial	15.63	11.83 ± 4.32	100
OpenDys	23.81	29.17 ± 15.79	103	Helvetica	16.78	12.41 ± 4.17	105
CMU	26.06	29.58 ± 12.05	104	CMU	17.21	13.03 ± 4.29	110
Courier	29.73	29.61 ± 10.87	104	OpenDys	16.34	13.09 ± 5.85	111
OpenDys It.	25.44	29.68 ± 14.44	105	Garamond	16.71	13.33 ± 4.38	113
Helvetica	27.18	31.05 ± 15.04	109	Courier	16.90	13.75 ± 5.50	116
Verdana	28.97	31.16 ± 13.03	110	Myriad	17.86	13.87 ± 5.32	117
Times	29.30	31.68 ± 11.81	112	OpenDys It.	17.63	14.34 ± 6.14	121
Times It.	28.55	32.38 ± 12.34	114	Times It.	18.50	14.55 ± 6.38	123
Myriad	26.95	32.66 ± 14.80	115	Times	18.35	14.75 ± 5.09	125
Garamond	30.53	33.30 ± 15.45	117	Verdana	18.86	15.69 ± 7.88	133
Arial It.	29.68	34.99 ± 16.60	123	Arial It.	21.25	17.24 ± 7.53	146

We include the relative percentage for *Reading Time*, with respect to the smallest average value, Arial. In all tables, fonts are sorted by the mean, and ties were resolved using the median and then the standard deviation.

Finally, we used the Spearman's rank-order correlation for nonparametric data to understand the strength of association between the main indicator of dyslexia, *Reading Time* [Serrano and Defior 2008] with the rest of our dependent variables: *Fixation Duration*, *Number of Fixations,* and *Preference Rating*.

5. RESULTS

In the first step, we cleaned up the data considering the answers to the comprehension questions. Out of the 1,164 data points (12 texts read by 97 participants), 10 data points were eliminated due to failing the comprehension test, six from people with dyslexia and four from people without dyslexia. Two of these occurred in the same text, the rest in different texts.

For group D, Shapiro-Wilk tests showed that nine, eight, and five out of the 12 datasets were not normally distributed for the *Reading Time*, *Fixation Duration*, and *Number of Fixations*, respectively. For group N, Shapiro-Wilk tests showed that three, eleven, and three out of the 12 datasets were not normally distributed for the *Reading Time*, *Fixation Duration,* and *Number of Fixations*, respectively. Shapiro-Wilk tests showed that no datasets were normally distributed for the *Preference Rating* of both groups. Also, Levene tests showed that none of the datasets had a homogeneous variance for all the measures and both groups. Then we proceeded with the analysis mentioned in Section 4.6. We only report post-hoc test results when significant effects were found.

5.1. Font Type

5.1.1. Reading Time. Table II shows the main statistical measures for the *Reading Time* for each of the *Font Type* conditions.

There was a significant effect of *Font Type* on *Reading Time* ($\chi^2(11) = 13.99$, $p < 0.001$).

The results of the post-hoc tests show that:

—**Between Groups:** Participants with dyslexia had significantly longer reading times ($\bar{x} = 24.94$, $s = 12.15$ seconds) than the participants without dyslexia ($\bar{x} = 17.69$, $s = 5.73$ seconds, $p < 0.001$). For *Reading Time* between groups, the Spearman's correlation coefficient is $\rho = 0.657$, and it is statistically significant ($p = 0.024$).

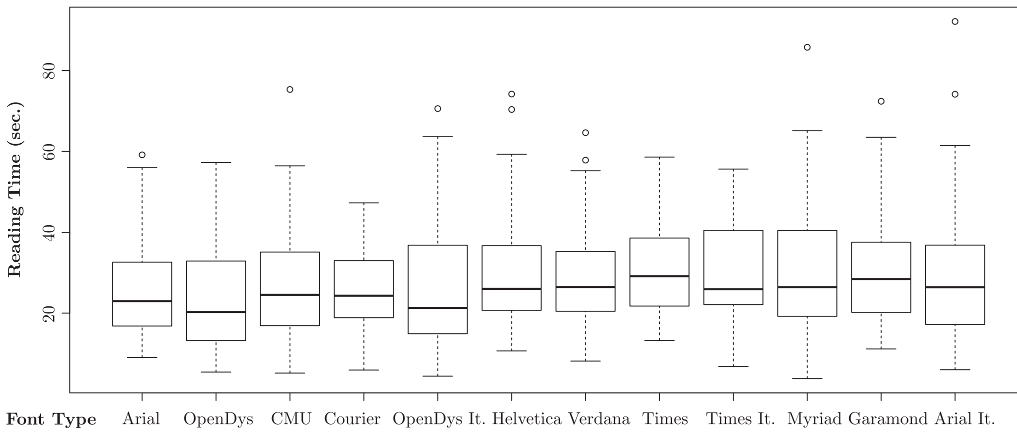

Fig. 4. *Reading Time* box plots by *Font Type* for group D (ordered by average *Reading Time* for group D).

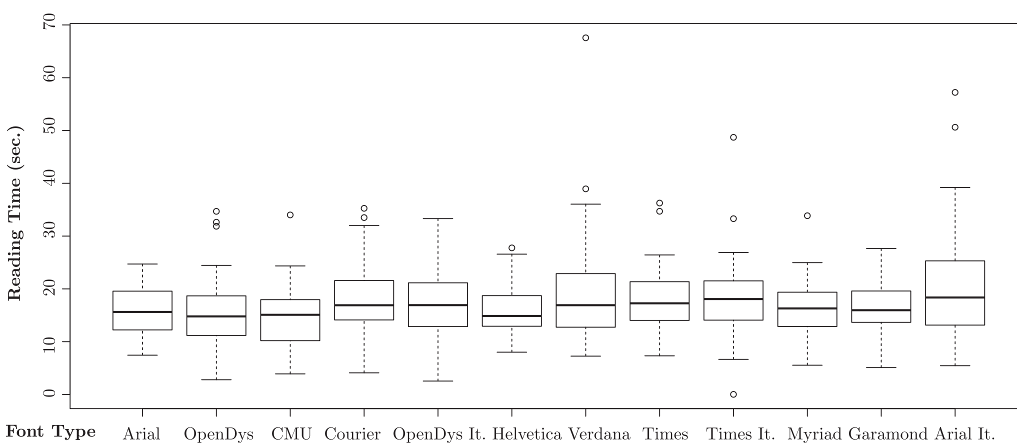

Fig. 5. *Reading Time* box plots by *Font Type* for group N (ordered by average *Reading Time* for group D).

—**Group D:** There was a significant effect of *Font Type* on *Reading Time* ($\chi^2(11) = 31.55$, $p < 0.001$) (Table II, Figure 4).
The results of the post-hoc tests show that:
—*Arial It.* had the longest reading time mean. Participants had significantly longer reading times using *Arial It.* than *Arial* ($p = 0.011$), *CMU* ($p = 0.011$), and *Helvetica* ($p = 0.034$).
—**Summary:** Participants had significantly longer reading times with *Arial It.* than with three other fonts. In this case, only 3 out of the 66 pairwise comparisons were significant.

—**Group N:** There was a significant effect of *Font Type* on *Reading Time* ($\chi^2(11) = 85.07$, $p < 0.001$) (Table II, Figure 5).
The results of the post-hoc tests show that:
—*Arial It.* had the longest reading time mean. Participants had significantly longer reading times using *Arial It.* than *Arial* ($p < 0.001$), *CMU* ($p = 0.001$), *Courier* ($p = 0.043$), *Garamond* ($p < 0.001$), *Helvetica* ($p = 0.013$), and *Times It.* ($p = 0.033$).

Table III. Median, Mean, and Standard Deviation of *Fixation Duration* in Seconds for Groups N and D

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean* ± *SD*	%		*Median*	*Mean* ± *SD*	%
	Fixation Duration				*Fixation Duration*		
Courier	0.22	0.22 ± 0.05	100	Courier	0.18	0.19 ± 0.03	100
Verdana	0.22	0.23 ± 0.07	105	Verdana	0.19	0.19 ± 0.03	100
Arial	0.23	0.24 ± 0.07	109	Arial	0.20	0.19 ± 0.03	100
Helvetica	0.24	0.24 ± 0.06	109	Helvetica	0.19	0.19 ± 0.04	100
Times	0.24	0.25 ± 0.07	114	Times It.	0.19	0.19 ± 0.04	100
Myriad	0.25	0.25 ± 0.07	114	Garamond	0.20	0.20 ± 0.03	105
OpenDys	0.24	0.26 ± 0.07	118	Myriad	0.20	0.20 ± 0.04	105
Times It.	0.25	0.26 ± 0.06	118	CMU	0.20	0.20 ± 0.04	105
OpenDys It.	0.26	0.26 ± 0.07	118	Times	0.21	0.20 ± 0.03	105
Garamond	0.25	0.27 ± 0.07	123	OpenDys	0.21	0.21 ± 0.04	111
CMU	0.25	0.27 ± 0.08	123	OpenDys It.	0.21	0.21 ± 0.04	111
Arial It.	0.28	0.28 ± 0.08	127	Arial It.	0.21	0.21 ± 0.04	111

We include the relative percentage for *Fixation Duration* with respect to the smallest average value.

- *Arial* had the shortest reading time mean. Participants had significantly shorter reading times using *Arial* than *Courier* ($p = 0.022$), *OpenDys It.* ($p = 0.031$), *Times* ($p < 0.001$), *Times It.* ($p = 0.003$), *Arial It.* ($p < 0.001$), and *Verdana* ($p < 0.001$).
- *Verdana* has the second longest reading time mean. Participants had significantly longer reading times with *Verdana* than with *Arial* ($p < 0.001$), *CMU* ($p < 0.001$), *Garamond* ($p = 0.003$), and *Helvetica* ($p = 0.021$).
- **Summary:** Participants had significantly shorter reading times with *Arial* than with six other fonts, and *Arial It.* and *Verdana* led to significantly longer reading times than five other fonts.[8] In fact, as two of the significant pairwise comparisons are repeated, 14 out of the 66 pairwise comparisons were significant.

5.1.2. Fixation Duration. Table III shows the main statistical measures for the *Fixation Duration* for each of the *Font Type* conditions. There was a significant effect of *Font Type* on *Fixation Duration* ($\chi^2(11) = 180.16$, $p < 0.001$).

For group D, *Fixation Duration* and *Reading Time* had a Spearman's correlation coefficient of $\rho = 0.987$ ($p < 0.001$). For group N, the correlation of *Fixation Duration* and *Reading Time* had a Spearman's correlation coefficient of $\rho = 0.551$ ($p = 0.063$).

The results of the post-hoc tests show that:

- **Between Groups:** Participants with dyslexia had significantly longer fixation duration means ($\bar{x} = 0.25$, $s = 0.07$ seconds) than the participants without dyslexia ($\bar{x} = 0.20$, $s = 0.04$ seconds, $p < 0.001$). For *Fixation Duration* between groups, the Spearman's correlation coefficient is $\rho = 0.717$, and it is statistically significant ($p = 0.009$).
- **Group D:** There was a significant effect of *Font Type* on *Fixation Duration* ($\chi^2(11) = 93.63$, $p < 0.001$) (Table III, Figure 6). The results of the post-hoc tests show that:
  - *Courier* had the lowest fixation duration mean. Participants had significantly shorter fixation durations reading with *Courier* than with *Arial It.* ($p < 0.001$), *CMU* ($p < 0.001$), *Garamond* ($p < 0.001$), *Times It.* ($p < 0.001$), *OpenDys It.* ($p = 0.001$), and *Arial* ($p = 0.046$).

[8]Here and in the following summaries, this should be interpreted as disjoint additive conditions.

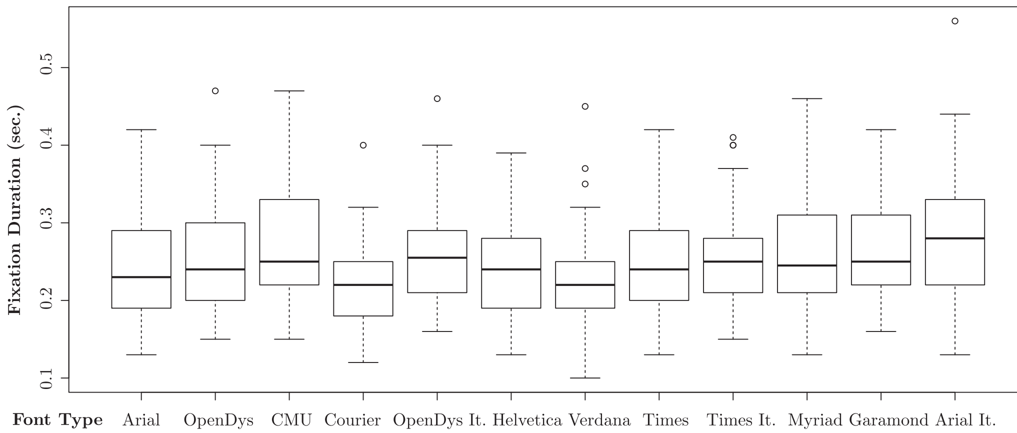

Fig. 6. *Fixation Duration* box plots by *Font Type* for group D (ordered by average *Reading Time* for group D).

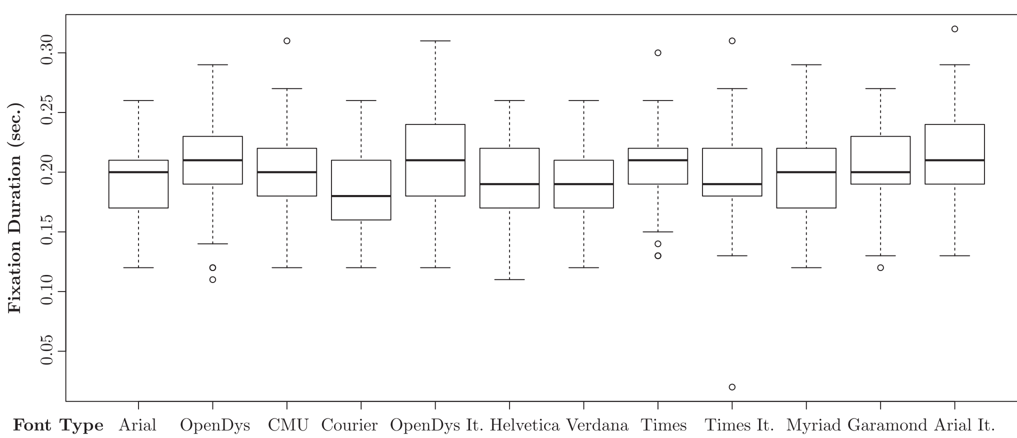

Fig. 7. *Fixation Duration* box plots by *Font Type* for group N (ordered by average *Reading Time* for group D).

—*Helvetica* had the third lowest fixation duration mean. Participants had significantly shorter fixation durations reading with *Helvetica* than with *Arial It.* ($p < 0.001$), *CMU* ($p = 0.001$), and *Garamond* ($p = 0.006$).
—Participants had significantly shorter fixation durations reading with *Arial* than with *CMU* ($p = 0.020$).
—*Arial It.* had the highest fixation duration mean. Participants had significantly longer fixation durations reading with *Arial It.* than with *Courier* ($p < 0.001$), *Helvetica* ($p < 0.001$), *Arial* ($p < 0.001$), *Times It.* ($p < 0.001$), *Times* ($p = 0.003$), *Myriad* ($p = 0.004$), *Garamond* ($p = 0.011$), and *Verdana* ($p = 0.049$).
—**Summary:** *Courier*, *Helvetica,* and *Arial* led to significantly shorter fixations durations than six other fonts, and *Arial It.* led to significantly longer fixations durations than eight other fonts. In fact, 16 out of the 66 pairwise comparisons were significant.

—**Group N:** There was a significant effect of *Font Type* on *Fixation Duration* ($\chi^2(11) = 95.99$, $p < 0.001$) (Table III, Figure 7). The results of the post-hoc tests show that:
—*Courier* had the lowest fixation duration mean. Participants had significantly shorter fixation durations reading with *Courier* than with *Arial It.* ($p < 0.001$),

Table IV. Median, Mean, and Standard Deviation of the *Number of Fixations* for Groups N and D

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean* ± *SD*	%		*Median*	*Mean* ± *SD*	%
	Number of Fixations				*Number of Fixations*		
CMU	84.0	87.60 ± 28.10	100	Arial	58.0	60.59 ± 18.67	100
OpenDys	83.0	88.05 ± 30.35	101	Helvetica	63.0	63.29 ± 17.80	104
Arial	87.5	88.76 ± 30.09	101	OpenDys	60.0	63.95 ± 21.43	106
OpenDys It.	87.0	93.31 ± 35.64	107	Times	63.0	64.14 ± 20.08	106
Times	95.5	98.24 ± 27.56	112	Courier	63.0	65.80 ± 19.80	109
Helvetica	93.0	99.23 ± 40.43	113	Myriad	66.0	69.02 ± 21.29	114
Arial It.	89.0	100.22 ± 40.68	115	OpenDys It.	65.5	69.19 ± 20.32	114
Garamond	96.5	101.02 ± 39.39	114	CMU	68.0	71.88 ± 20.46	119
Courier	93.0	101.33 ± 33.67	116	Arial It.	67.0	72.98 ± 25.42	120
Myriad	93.0	101.37 ± 38.33	116	Times It.	72.0	74.33 ± 26.81	123
Times It.	98.0	101.78 ± 29.22	116	Garamond	73.0	80.75 ± 34.02	133
Verdana	101.0	103.33 ± 34.66	118	Verdana	76.0	80.83 ± 37.21	133

We include the relative percentage for *Number of Fixations* with respect to the smallest average value.

CMU ($p = 0.001$), *Garamond* ($p = 0.001$), *Myriad* ($p = 0.004$), *OpenDys It.* ($p = 0.001$), and *Times* ($p = 0.001$).
—*Verdana* had the second lowest fixation duration mean. Participants had significantly shorter fixation durations reading with *Verdana* than with *Arial It.* ($p < 0.001$), *CMU* ($p = 0.030$), *Garamond* ($p = 0.029$), *Arial* ($p = 0.003$), and *Arial It.* ($p = 0.031$).
—*Arial It.* had the highest fixation duration mean. Participants had significantly longer fixation durations reading with *Arial It.* than with *Arial* ($p < 0.001$), *Courier* ($p < 0.001$), *Times It.* ($p = 0.001$), *Times* ($p = 0.046$), and *Verdana* ($p < 0.001$).
—**Summary:** *Courier* and *Verdana* led to significantly shorter fixations durations than 11 other fonts, and *Arial It.* led to significantly longer fixations durations than five other fonts. In fact, 14 out of the 66 pairwise comparisons were significant.

5.1.3. Number of Fixations. Table IV shows the main statistical measures for the *Number of Fixations* for each of the *Font Type* conditions.

There was a significant effect of *Font Type* on the *Number of Fixations* ($\chi^2(11) = 30.196$, $p = 0.001$).

For group D, *Reading Time* and *Number of Fixations* had a Spearman's correlation coefficient of $\rho = 0.629$ ($p = 0.032$). For group N, the correlation of *Reading Time* and *Number of Fixations* had a Spearman's correlation coefficient of $\rho = 0.608$ ($p = 0.040$). This is expected because larger ratings are better whereas smaller are better.

The results of the post-hoc tests show that:

—**Between Groups:** Participants with dyslexia had significantly more fixations ($\bar{x} = 97.07$, $s = 69.74$ seconds) than the participants without dyslexia ($\bar{x} = 34.47$, $s = 24.96$ seconds, $p < 0.001$).
—**Group D:** There was a significant effect of *Font Type* on *Number of Fixations* ($\chi^2(11) = 30.20$, $p = 0.001$) (Table IV, Figure 8). The results of the post-hoc tests show that:
  —*CMU* had the smallest mean for number of fixations. Participants had significantly fewer fixations reading with *CMU* than with *Times It.* ($p = 0.007$).
  —*OpenDys* had the second smallest mean for number of fixations. Participants had significantly fewer fixations reading with *OpenDys* than with *Verdana* ($p = 0.004$).
  —*Arial* had the third smallest mean for number of fixations. Participants had significantly fewer fixations reading with *Arial* than with *Courier* ($p = 0.023$).

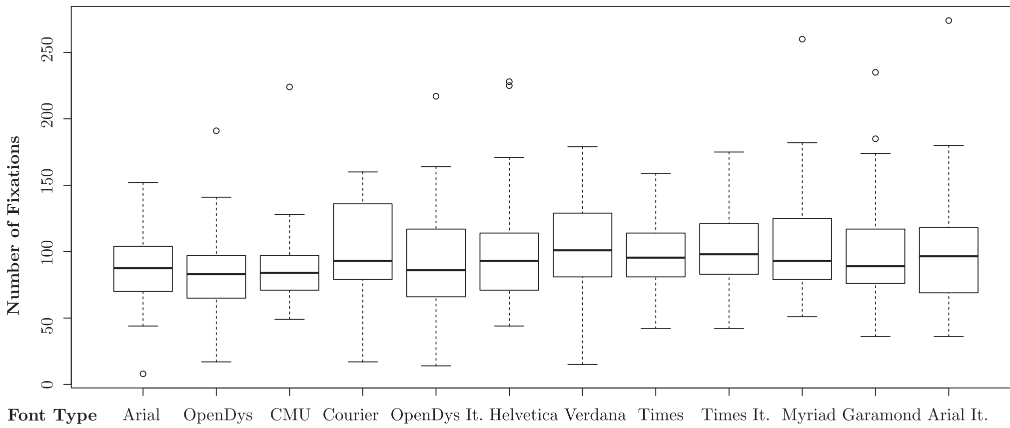

Fig. 8. *Number of Fixations* box plots by *Font Type* for group D (ordered by average *Reading Time* for group D).

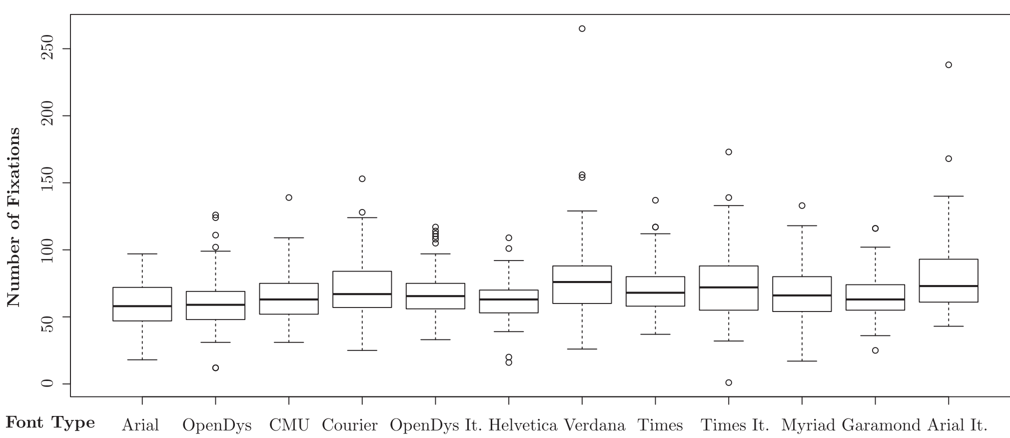

Fig. 9. *Number of Fixations* box plots by *Font Type* for group N (ordered by average *Reading Time* for group D).

- **Summary:** *CMU, OpenDys* and *Arial* led to significantly less fixations than three other fonts. In this case only 3 out of the 66 pairwise comparisons were significant.

—**Group N:** There was a significant effect of *Font Type* on *Number of Fixations* ($\chi^2(11) = 68.84$, $p < 0.001$) (Table IV, Figure 9). The results of the post-hoc tests show that:
- *Arial* had the smallest mean for number of fixations. Participants had significantly fewer fixations reading with *Arial* than with *Arial It.* ($p < 0.001$), *Courier* ($p = 0.003$), *Times* ($p < 0.001$), *Times It.* ($p = 0.003$), and *Verdana* ($p < 0.001$).
- *Verdana* had the highest mean for number of fixations. Participants had significantly more fixations reading with *Verdana* than with *Arial* ($p < 0.001$), *CMU* ($p = 0.001$), *Garamond* ($p < 0.001$), and *Helvetica* ($p = 0.014$).
- Participants had significantly fewer fixations reading with *Times* than with *CMU* ($p = 0.022$).
- Participants had significantly more fixations reading with *Arial It.* than with *CMU* ($p = 0.018$), and *Garamond* ($p = 0.006$).

Table V. Median, Mean, and Standard Deviation of the *Preference Rating* for Groups N and D

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean* ± *SD*	%		*Median*	*Mean* ± *SD*	%
	Preference Rating				*Preference Rating*		
Verdana	4	3.79 ± 0.98	100	Helvetica	4	3.97 ± 0.92	100
Helvetica	4	3.62 ± 1.08	96	Verdana	4	3.97 ± 0.95	100
Arial	4	3.60 ± 1.13	95	Arial	4	3.84 ± 0.80	97
Times	4	3.45 ± 1.15	91	CMU	4	3.79 ± 0.90	95
Myriad	3.5	3.40 ± 0.99	90	Myriad	4	3.66 ± 0.93	92
CMU	3	3.31 ± 0.98	87	Times	4	3.64 ± 0.89	92
Courier	3	3.14 ± 1.39	83	Arial It.	3	3.36 ± 1.30	85
Arial It.	3	2.90 ± 1.10	77	Garamond	3	3.33 ± 0.93	84
Times It.	3	2.86 ± 1.20	75	Times It.	3	3.18 ± 1.07	80
OpenDys	3	2.57 ± 1.15	68	Courier	3	2.85 ± 1.08	72
Garamond	2	2.57 ± 1.15	68	OpenDys	2	2.24 ± 1.09	56
OpenDys It.	2	2.43 ± 1.17	64	OpenDys It.	2	2.03 ± 1.04	51

We include the relative percentage for *Preference Rating* with respect to the largest average value, Verdana.

—**Summary:** *Arial* and *Times* led to significantly less fixations than six other fonts, and *Verdana* and *Arial It.* led to significantly more fixations than five other fonts. That is, as one comparison appears twice, 10 out of the 66 pairwise comparisons were significant.

5.1.4. Preference Rating. Table V shows the main statistical measures for the *Preference Rating* for each of the *Font Type* conditions.

There was a significant effect of *Font Type* on the *Preference Rating* ($\chi^2(11) = 120.92, p < 0.001$).

For group D, *Reading Time* and *Preference Rating* had a Spearman's correlation coefficient of $\rho = -0.998$ ($p < 0.001$). For group N, the correlation of *Reading Time* and *Preference Rating* had a Spearman's correlation coefficient of $\rho = -0.767$ ($p = 0.004$). A negative correlation was expected because larger ratings are better whereas smaller are worse.

The results of the post-hoc tests show that:

—**Between Groups:** Participants with dyslexia had significantly lower preferences ratings ($\bar{x} = 3.14$, $s = 1.21$) than the participants without dyslexia ($\bar{x} = 3.32$, $s = 1.17$, $p = 0.024$). For *Preference Rating* between groups, the Spearman's correlation coefficient is $\rho = 0.912$, and this is statistically significant ($p < 0.001$).

—**Group D:** There was a significant effect of *Font Type* on subjective preference ratings ($\chi^2(11) = 79.61, p < 0.001$) (Table V, Figure 10). Pairwise post-hoc comparisons showed significant differences between the following conditions:
  - —*Verdana* is significantly preferred over *Arial It* ($p < 0.001$), *OpenDys* ($p = 0.002$), *OpenDys It.* ($p = 0.004$), *Garamond* ($p = 0.008$), and *Times It.* ($p = 0.041$).
  - —*Helvetica* is significantly preferred over *OpenDys It.* ($p = 0.010$), *OpenDys* ($p = 0.020$), and *Arial It.* ($p = 0.031$).
  - —*Arial* was significantly more preferred than *Arial It.* ($p = 0.028$).
  - —*Garamond* was significantly less preferred than *Verdana* ($p = 0.008$), *Times* ($p = 0.023$), *Arial* ($p = 0.023$), and *CMU* ($p = 0.030$).
  - —**Summary:** The participants significantly preferred *Verdana*, *Helvetica,* and *Arial* and significantly disliked *Garamond* in comparison with other fonts.

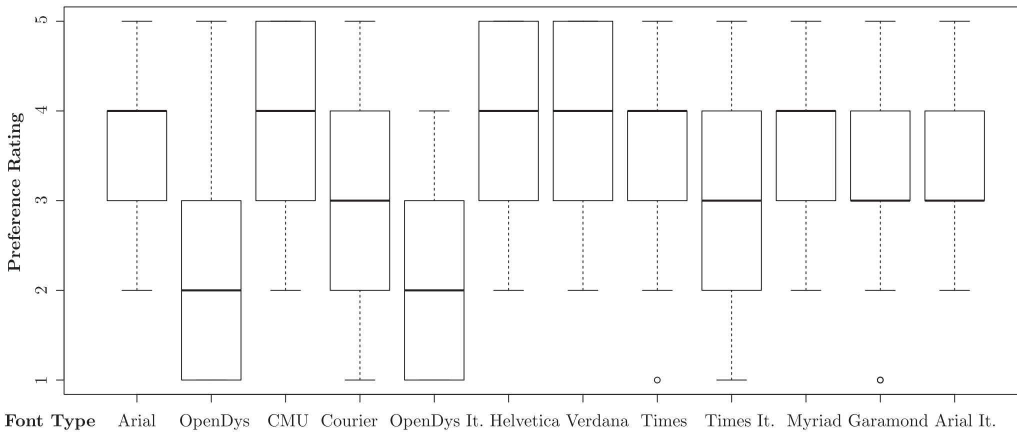

Fig. 10. *Preference Rating* box plots by *Font Type* for group D (ordered by average *Reading Time* for group D).

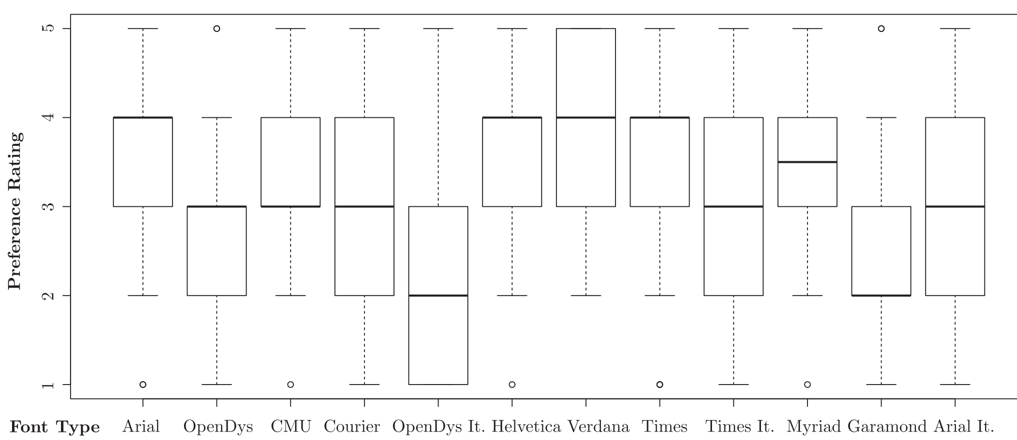

Fig. 11. *Preference Rating* box plots by *Font Type* for group N (ordered by average *Reading Time* for group D).

—**Group N:** There was a significant effect of *Font Type* on subjective preference ratings ($\chi^2(11) = 50.65$, $p < 0.001$) (Table V, Figure 11). Pairwise post-hoc comparisons showed significant differences between the following conditions:
- —*Verdana* is significantly preferred over *Courier* ($p = 0.005$), *OpenDys* ($p = 0.012$), and *OpenDys It.* ($p = 0.001$).
- —*Helvetica* is significantly preferred over *OpenDys* ($p = 0.011$), and *OpenDys It.* ($p = 0.004$).
- —*Arial* is significantly preferred over *OpenDys It.* ($p < 0.001$).
- —*OpenDys It.* was significantly less preferred than *Arial* ($p < 0.001$), *Arial It.* ($p = 0.005$), *CMU* ($p = 0.002$), *Garamond* ($p = 0.021$), *Helvetica* ($p = 0.004$), *Myriad* ($p = 0.006$), *Times* ($p = 0.005$), and *Verdana* ($p = 0.001$).
- —**Summary:** The participants without dyslexia significantly preferred *Verdana*, *Helvetica,* and *Arial*, and significantly disliked *OpenDys* and *OpenDys It.* in comparison with other fonts.

Table VI. Median, Mean, and Standard Deviation of *Reading Time*, *Fixation Duration*, *Number of Fixations*, and *Preference Rating* for [±*Italic*]

Font Type	Group D			*Font Type*	Group N		
	Median	*Mean* ± *SD*	*%*		*Median*	*Mean* ± *SD*	*%*
		Reading Time				*Reading Time*	
[−Italic]	27.04	29.74 ± 13.40	Not Sig.	[−Italic]	16.69	17.40 ± 5.62	100
[+Italic]	28.77	32.35 ± 14.62	Not Sig.	[+Italic]	18.91	20.11 ± 7.55	116
		Fixation Duration				*Fixation Duration*	
[−Italic]	0.24	0.25 ± 0.07	100	[−Italic]	0.20	0.20 ± 5.62	Not Sig.
[+Italic]	0.26	0.27 ± 0.07	108	[+Italic]	0.20	0.21 ± 7.55	Not Sig.
		Number of Fixations				*Number of Fixations*	
[−Italic]	88.5	91.74 ± 29.50	Not Sig.	[−Italic]	62.0	65.50 ± 20.62	100
[+Italic]	95	98.47 ± 35.42	Not Sig.	[+Italic]	70	74.80 ± 27.89	114
		Preference Rating				*Preference Rating*	
[−Italic]	3	3.21 ± 1.22	100	[−Italic]	3	3.24 ± 1.17	100
[+Italic]	3	2.73 ± 1.20	85	[+Italic]	3	2.86 ± 1.16	88

For the post-hoc tests with significant effects we include the relative percentage with respect to the best average value. For this table and the following tables, "Not Sig." stands for the post-hoc tests with no significant effects.

5.2. Italics

We study three [±Italic] typefaces: *Arial*, *OpenDyslexic,* and *Times* in comparison with *Arial It.*, *OpenDys It.*, and *Times It.*

Between groups, participants with dyslexia had significantly longer *Reading Time*, significantly longer *Fixation Duration*, and significantly more *Number of Fixations*, than the participants without dyslexia ($p < 0.001$) for the three measures. In Table VI, we show the medians, means, and standard deviations of each group.

Reading Time. There was a significant effect of [±Italic] on *Reading Time* ($\chi^2(1) = 27.27$, $p < 0.001$). The results of the post-hoc tests show that:

—**Group D:** We did not find a significant effect of [±Italic] on *Reading Time* ($p = 0.120$) (Table VI).
—**Group N:** There was a significant effect of [±Italic] on *Reading Time* for participants without dyslexia ($p = 0.001$). The reading time mean of fonts in [+Italic] (*Arial It.*, *OpenDys. It.*, and *Times It.*) ($\bar{x} = 20.11, s = 7.55$) was significantly larger than the reading time mean of the fonts in [−Italic] or *roman* (*Arial*, *OpenDys,* and *Times*) ($\bar{x} = 17.40, s = 5.62$) (Table VI).

Fixation Duration. There was a significant effect of [±Italic] on *Fixation Duration* ($\chi^2(1) = 8.07$, $p = 0.005$). The results of the post-hoc tests show that:

—**Group D:** There was a significant effect of [±Italic] on *Fixation Duration* for participants with dyslexia ($p = 0.040$). The fixation duration mean of the fonts [+Italic] (*Arial It.*, *OpenDys. It.*, and *Times It.*) ($\bar{x} = 0.27, s = 0.07$), was significantly larger than the fixation duration mean of the fonts [−Italic] (*Arial*, *OpenDys*, and *Times*) ($\bar{x} = 0.25, s = 0.07$) (Table VI).
—**Group N:** We did not find a significant effect of [±Italic] on *Fixation Duration* ($p = 0.280$) (Table VI).

Number of Fixations. There was a significant effect of [±Italic] on *Number of Fixations* ($\chi^2(1) = 7.418$, $p = 0.006$). The results of the post-hoc tests show that:

Table VII. Median, Mean, and Standard Deviation of *Reading Time*, *Fixation Duration*, *Number of Fixations*, and *Preference Rating* for [±*Serif*]

Font Type	Group D			*Font Type*	Group N		
	Median	*Mean* ± *SD*	%		*Median*	*Mean* ± *SD*	%
	Reading Time				*Reading Time*		
[−Serif]	27.08	30.80 ± 13.84	Not Sig.	[−Serif]	17.48	17.98 ± 6.58	Not Sig.
[+Serif]	29.06	31.53 ± 13.21	Not Sig.	[+Serif]	16.78	18.07 ± 5.09	Not Sig.
	Fixation Duration				*Fixation Duration*		
[−Serif]	0.24	0.24 ± 0.07	100	[−Serif]	0.20	0.20 ± 0.03	100
[+Serif]	0.25	0.26 ± 0.07	108	[+Serif]	0.19	0.19 ± 0.03	95
	Number of Fixations				*Number of Fixations*		
[−Serif]	93	98.15 ± 36.24	Not Sig.	[−Italic]	65	68.32 ± 25.92	Not Sig.
[+Serif]	91	95.68 ± 32.46	Not Sig.	[+Italic]	66	67.41 ± 20.30	Not Sig.
	Preference Rating				*Preference Rating*		
[−Serif]	4	3.60 ± 1.04	100	[−Serif]	4	3.84 ± 0.94	Not Sig.
[+Serif]	3	3.11 ± 1.15	86	[+Serif]	4	3.62 ± 0.95	Not Sig.

For the post-hoc tests with significant effects we include the relative percentage with respect to the best average value.

—**Group D:** We did not find a significant effect of [±Italic] on *Number of Fixations* ($p = 0.14$) (Table VI).

—**Group N:** There was a significant effect of [±Italic] on *Number of Fixations* for participants without dyslexia ($p = 0.002$). They had significantly more fixations when the texts were presented with [+Italic] fonts (Table VI).

Preference Rating. There was a significant effect of [±Italic] on *Preference Rating* ($\chi^2(1) = 40.11$, $p < 0.001$). The results of the post-hoc tests show that:

—**Group D:** There was a significant effect of [±Italic] on *Preference Rating* for participants with dyslexia ($p = 0.002$). The *Preference Rating* for the fonts [−Italic] (*Arial*, *OpenDys*, and *Times*) ($\bar{x} = 3.21, s = 1.22$) was significantly higher than for the fonts [+Italic] (*Arial It.*, *OpenDys. It.*, and *Times It.*) ($\bar{x} = 2.73, s = 1.22$) (Table VI).

—**Group N:** There was a significant effect of [±Italic] on *Preference Rating* for participants without dyslexia ($p = 0.018$). The *Preference Rating* for the fonts [−Italic] (*Arial*, *OpenDys*, and *Times*) ($\bar{x} = 3.24, s = 1.17$) was significantly higher than for the fonts [+Italic] (*Arial It.*, *OpenDys. It.*, and *Times It.*) ($\bar{x} = 2.86, s = 1.16$) (Table VI).

5.3. Serif

For the [±Serif] comparison, we used the three *serif* fonts of our set (*CMU*, *Garamond*, and *Times*) together with the four *sans serif* fonts (*Arial*, *Helvetica*, *Myriad*, and *Verdana*).

Between groups, participants with dyslexia had significantly longer *Reading Time*, significantly longer *Fixation Duration*, and significantly more *Number of Fixations*, than the participants without dyslexia ($p < 0.001$ for the three measures). In Table VII, we show the medians, means, and standard deviations of each group.

Reading Time. We did not find a significant effect of [±Serif] on *Reading Time* ($\chi^2(1) = 1.65$, $p = 0.199$).

Fixation Duration. There was a significant effect of [±Serif] on *Fixation Duration* ($\chi^2(1) = 9.31$, $p = 0.002$). The results of the post-hoc tests show that:

—**Group D:** There was a significant effect of [±Serif] on *Fixation Duration* for people with dyslexia ($p = 0.015$). Indeed, the fixation duration mean of the [+Serif] fonts

Table VIII. Median, Mean, and Standard Deviation of *Reading Time*, *Fixation Duration*, *Number of Fixations*, and *Preference Rating* for [± Monospace]

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean ± SD*	*%*		*Median*	*Mean ± SD*	*%*
	Reading Time				*Reading Time*		
[+Monospace]	29.73	29.61 ± 10.87	Not Sig.	[+Monospace]	16.90	18.57 ± 6.24	Not Sig.
[−Monospace]	29.06	31.53 ± 13.21	Not Sig.	[−Monospace]	17.46	18.04 ± 5.09	97
	Fixation Duration				*Fixation Duration*		
[+Monospace]	0.22	0.22 ± 0.05	100	[+Monospace]	0.18	0.19 ± 0.03	100
[−Monospace]	0.25	0.26 ± 0.07	118	[−Monospace]	0.20	0.20 ± 0.03	105
	Number of Fixations				*Number of Fixations*		
[+Monospace]	93	101.33 ± 33.67	Not Sig.	[−Monospace]	66	67.27 ± 20.26	Not Sig.
[−Monospace]	91	95.68 ± 32.46	Not Sig.	[+Monospace]	67	72.98 ± 25.42	Not Sig.
	Preference Rating				*Preference Rating*		
[+Monospace]	3	3.14 ± 1.39	Not Sig.	[+Monospace]	3	2.85 ± 1.30	79
[−Monospace]	3	3.11 ± 1.15	Not Sig.	[−Monospace]	4	3.59 ± 0.98	100

For the post-hoc tests with significant effects we include the relative percentage with respect to the smallest average value.

($\bar{x} = 0.26, s = 0.07$) was significantly larger than the fixation duration mean of the *sans serif* fonts [−Serif] ($\bar{x} = 0.24, s = 0.07$) (Table VII).

—**Group N:** Similarly, there was a significant effect of [±Serif] on *Fixation Duration* for people without dyslexia ($p = 0.007$). The fixation duration mean of the [+Serif] fonts ($\bar{x} = 0.20, s = 0.03$) was significantly larger than the fixation duration mean of the [−Serif] fonts ($\bar{x} = 0.19, s = 0.03$) (Table VII).

Number of Fixations. There was a significant effect of [±Serif] on *Number of Fixations* ($\chi^2(1) = 4.931$, $p = 0.026$).

However, the results of the post-hoc tests show that the effect comes from the between-groups comparison only. We did not find a significant effect of [±Serif] on *Number of Fixations* for Group D ($p = 0.64$), nor for Group N ($p = 0.81$).

Preference Rating. There was a significant effect of [±Serif] on *Preference Rating* ($\chi^2(1) = 13.88$, $p < 0.001$). The results of the post-hoc tests show that:

—**Between Groups:** Participants with dyslexia had significantly lower preference ratings than the participants without dyslexia ($p < 0.001$). In Table VII, we show the medians, means, and standard deviations of each group.

—**Group D:** There was a significant effect of [±Serif] on *Preference Rating* for people with dyslexia ($p < 0.001$). They significantly preferred [−Serif] fonts ($\tilde{x} = 3.60, s = 1.04$) over [+Serif] fonts ($\tilde{x} = 3.11, s = 1.15$) (Table VII).

—**Group N:** We did not find a significant effect of [±Serif] on *Preference Rating* for people without dyslexia ($p = 0.091$) (Table VII).

5.4. Monospace

For the [±Monospace] analysis, we compare the monospaced font *Courier*, which is *roman serif*, with the rest of the proportional fonts, which are also *roman serif* as well as *proportional*. These are *CMU*, *Garamond,* and *Times*.

Between groups, participants with dyslexia had significantly longer *Reading Time*, significantly longer *Fixation Duration*, and significantly more *Number of Fixations* than the participants without dyslexia ($p < 0.001$, for the three measures). In Table VIII, we show the medians, means, and standard deviations of each group.

Reading Time. We did not find a significant effect of [±Monospace] on *Reading Time* ($\chi^2(1) = 3.40$, $p = 0.065$) (Table VIII).

Fixation Duration. There was a significant effect of [±Monospace] on *Preference Rating* ($\chi^2(1) = 25.28$, $p < 0.001$). The results of the post-hoc tests show that:

—**Group D:** There was a significant difference of [±Monospace] on *Fixation Duration* ($p < 0.001$). We found that the fixation duration mean of the [+Monospace] font ($\bar{x} = 0.22, s = 0.05$), was significantly shorter than the fixation duration mean of the *proportional* or [−Monospace] fonts ($\bar{x} = 0.26, s = 0.07$) (Table VIII).

—**Group N:** There was a significant difference of [±Monospace] on *Fixation Duration* ($p = 0.002$). Similarly, the fixation duration mean of the [+Monospace] font ($\bar{x} = 0.19, s = 0.03$), was significantly shorter than the fixation duration mean of the [−Monospace] fonts ($\bar{x} = 0.20, s = 0.03$) (Table VIII).

Number of Fixations. There was a significant effect of [±Monospace] on *Number of Fixations* ($\chi^2(1) = 4.129$, $p = 0.042$). However, the results of the post-hoc tests show that the effect comes from the between-groups comparison only. We did not find a significant effect of [±Monospace] on *Number of Fixations* for Group D ($p = 0.22$), nor for group Group N ($p = 0.20$).

Preference Rating. There was a significant effect of [±Monospace] on *Preference Rating* ($\chi^2(1) = 6.45$, $p = 0.011$). The results of the post-hoc tests show that:

—**Group D:** We did not find a significant effect of [±Monospace] on the participants with dyslexia preferences ($p = 0.79$) (Table VIII).

—**Group N:** There was a significant effect of [±Monospace] on *Preference Rating* for participants without dyslexia ($p = 0.003$). They significantly preferred [−Monospace] fonts (*CMU*, *Garamond,* and *Times*) ($\bar{x} = 3.59, s = 0.98$) over the [+Monospace] font (*Courier*) ($\bar{x} = 2.85, s = 1.30$) (Table VIII).

5.5. Dyslexic Fonts

For the [±Dyslexic] analysis, we compare the monospaced font *OpenDyslexic*, which is *sans serif roman*, with the rest of the *sans serif roman*. These are *Arial*, *Helvetica*, *Myriad*, and *Verdana*. To cover the italic variant *OpenDyslexic It.,* we compared it with the only *sans serif italic* font we had in our set, *Arial it*.

Between groups, participants with dyslexia had significantly longer *Reading Time* and significantly longer *Fixation Duration* than the participants without dyslexia ($p < 0.001$ for the two measures). In Table IX, we show the medians, means, and standard deviations of each group.

Reading Time. We did not find a significant effect of [±Dyslexic] on *Reading Time* ($\chi^2(1) = 1.67$, $p = 0.197$), but there was an effect of [±Dyslexic It.] on *Reading Time* ($\chi^2(1) = 12.89$, $p < 0.001$) (Table IX).

—**Group N:** We found a significant effect of [±Dyslexic It.] on *Reading Time* for people without dyslexia ($p = 0.035$). The visit duration means of [+Dyslexic It.] font *OpenDys It.* ($\bar{x} = 18.73, s = 5.60$), were significantly longer than the ones from [−Dyslexic It.] ($\bar{x} = 22.38, s = 9.02$) (Table IX).

Fixation Duration. There was a significant effect of [±Dyslexic] on *Fixation Duration* ($\chi^2(1) = 7.45$, $p = 0.006$). In contrast, we did not find a significant effect of [±Dyslexic It.] on *Fixation Duration* ($\chi^2(1) = 2.91$, $p = 0.088$) (Table IX).

—**Group N:** We found a significant effect of [±Dyslexic] on *Fixation Duration* for participants without dyslexia ($p = 0.027$). The fonts which were not designed for people

Table IX. Median, Mean and Standard Deviation of *Reading Time*, *Fixation Duration*, *Number of Fixations*, and *Preference Rating* for the [± Dyslexic] and [±Dyslexic It.] Conditions

Font Type	**Group D**			*Font Type*	**Group N**		
	Median	*Mean* ± *SD*	*Sig./%*		*Median*	*Mean* ± *SD*	Sig./%
		Reading Time				*Reading Time*	
[+Dyslexic]	23.81	29.17 ± 15.79	Not Sig.	[+Dyslexic]	16.34	17.04 ± 6.08	Not Sig.
[−Dyslexic]	27.08	30.80 ± 13.84	Not Sig.	[−Dyslexic]	16.85	18.01 ± 6.59	Not Sig.
[+Dys. It.]	25.44	29.68 ± 14.44	Not Sig.	[+Dys. It.]	17.63	18.73 ± 5.60	100
[−Dys. It.]	29.68	34.99 ± 16.60	Not Sig.	[−Dys. It.]	21.25	22.38 ± 9.02	119
		Fixation Duration				*Fixation Duration*	
[+Dyslexic]	0.24	0.26 ± 0.07	Not Sig.	[+Dyslexic]	0.21	0.21 ± 0.04	111
[−Dyslexic]	0.24	0.24 ± 0.07	Not Sig.	[−Dyslexic]	0.19	0.19 ± 0.03	100
[+Dys. It.]	0.25	0.26 ± 0.07	Not Sig.	[+Dys. It.]	0.21	0.21 ± 0.04	Not Sig.
[−Dys. It.]	0.28	0.28 ± 0.08	Not Sig.	[−Dys. It.]	0.21	0.21 ± 0.04	Not Sig.
		Number of Fixations				*Number of Fixations*	
[+Dyslexic]	83.0	88.04 ± 30.35	Not Sig.	[+Dyslexic]	60.o	63.95 ± 21.43	Not Sig.
[−Dyslexic]	93.0	98.15 ± 36.24	Not Sig.	[−Dyslexic]	65.5	68.43 ± 26.02	Not Sig.
[+Dys. It.]	87.0	93.31 ± 40.68	Not Sig.	[+Dys. It.]	65.5	69.19 ± 20.31	Not Sig.
[−Dys. It.]	96.5	100.22 ± 35.64	Not Sig.	[−Dys. It.]	73.0	80.76 ± 34.02	Not Sig.
		Preference Rating				*Preference Rating*	
[+Dyslexic]	3	2.57 ± 1.15	71	[+Dyslexic]	2	2.24 ± 1.09	58
[−Dyslexic]	4	3.60 ± 1.04	100	[−Dyslexic]	4	3.86 ± 0.91	100
[+Dys. It.]	2	2.42 ± 1.27	Not Sig.	[+Dys. It.]	2	2.03 ± 1.05	60
[−Dys. It.]	3	2.90 ± 1.10	Not Sig.	[−Dys. It.]	3	3.36 ± 0.93	100

For the post-hoc tests with significant effects we include the relative percentage with respect to the smallest average value.

with dyslexia [−Dyslexic] (*Arial*, *Helvetica*, *Myriad*, and *Verdana*) ($\bar{x} = 0.19, s = 0.03$), led to significantly shorter fixations durations than *OpenDyslexic* ([+Dyslexic]) ($\bar{x} = 0.2, s = 0.04$).

Number of Fixations. We did not find a significant effect of [±Dyslexic], nor of [±Dyslexic It.], on *Number of Fixations* ($\chi^2(1) = 3.62$, $p = 0.057$ and $\chi^2(1) = 3.18$, $p = 0.075$, respectively)

Preference Rating. There was a significant effect of [±Dyslexic] on *Preference Rating* ($\chi^2(1) = 94.32$, $p < 0.001$) as well as an effect of [±Dyslexic It.] on *Preference Rating* ($\chi^2(1) = 8.1$, $p = 0.004$) (Table IX).

—**Group D:** We found a significant effect of [±Dyslexic] on *Preference Rating* for participants with dyslexia ($p < 0.001$). The fonts which were not designed for people with dyslexia (*Arial*, *Helvetica*, *Myriad*, and *Verdana*) ($\bar{x} = 3.60, s = 1.04$), were preferred to *OpenDyslexic* ($\bar{x} = 2.57, s = 1.15$). In contrast, we did not find a significant effect of [±Dyslexic It.] on *Preference Rating* for people with dyslexia ($p = 0.06$) (Table IX).
—**Group N:** We found a significant effect of [±Dyslexic] on *Preference Rating* for participants without dyslexia ($p < 0.001$). The [−Dyslexic] fonts (*Arial*, *Helvetica*, *Myriad*, and *Verdana*) ($\bar{x} = 3.86, s = 0.91$) were preferred to *OpenDyslexic* [+Dyslexic] ($\bar{x} = 2.24, s = 1.09$). There was a significant effect of [±Dyslexic It.] on *Preference Rating* for people without dyslexia ($p < 0.001$). The [−Dyslexic It.] font type ($\bar{x} = 3.36, s = 0.93$) was significantly more preferred than the [+Dyslexic It.] font *OpenDys It.* ($\bar{x} = 2.03, s = 1.05$) (Table IX).

6. DISCUSSION

The differences in reading performance between groups are consistent with the majority of eye tracking literature from experimental psychology [Adler-Grinberg and Stark 1978; Eden et al. 1994; Elterman et al. 1980; Lefton et al. 1979; Martos and Vila 1990]. The eye movements of readers with dyslexia are different from regular readers. People with dyslexia, as well as beginning readers, make longer fixations and more fixations than normal readers. However, participants with dyslexia had significantly lower preferences ratings than the participants without dyslexia. This can be explained by their reading difficulties caused by dyslexia.

The correlations between groups were significant for all the measures. The fonts that are more readable for people with dyslexia are also beneficial for people without dyslexia. People without dyslexia also prefer the fonts that people with dyslexia preferred. While there were strong positive correlations between the two measures for reading performance (*Reading Time* and *Fixation Duration*), for the tested fonts the correlations between these two measures were negative and significant for both groups. What is beneficial for the readability of people with and without dyslexia is not necessarily what they prefer.

Our results on reading performance provide evidence that font types have an impact on readability for both people with and without dyslexia. Second, these results are consistent with most of the current text design recommendations for people with dyslexia. For the tested fonts, *sans serif* and *roman* style led to shorter fixation durations in our participants with dyslexia, as recommended by Lockley [2002]. However, these styles did not lead to significantly shorter reading durations for people with dyslexia for the tested fonts. For people without dyslexia, *italic* fonts led to longer reading times and fixation durations for the tested fonts.

Overall, for the tested fonts, the reading time of the *italic* fonts was always worse than its *roman* counterpart, confirming the commonly established fact that cursive letters are harder to read for people with dyslexia. Both groups preferred *roman* fonts. Although for the tested fonts *sans serif*, *monospaced,* and *roman* fonts led to significantly shorter fixation durations for people with dyslexia, we did not find a significant difference in reading time. Hence, our conclusions on these characteristics are weaker. People without dyslexia presented the same behavior regarding *sans serif* and *monospaced* tested fonts in our study: They presented longer fixation durations for *serif* and *monospaced* tested fonts. But only people with dyslexia preferred *sans serif* fonts, and, surprisingly, people without dyslexia preferred *proportional* fonts among the fonts tested.

Although most of the recommendations for people with dyslexia suggest using *sans serif* typefaces [British Dyslexia Association 2012; Evett and Brown 2005], we found no significant differences in reading time. Similarly, Beymer et al. [2008] conducted an eye-tracking study with regular readers without finding significant differences in *serif* vs. *sans serif* fonts presented on-screen. In fact, there are plenty of studies where no significant statistical difference was found between the legibility of *serif* and *sans serif* typefaces [Paterson and Tinker 1932; Boynton et al. 1995; De Lange et al. 1993; Moriarty and Scheiner 1984; Poulton 1965]. Taking into account the preference ratings, we have calculated subgroups in the participants in which certain fonts are better than others. In Table X, we show the preferences percentages per group and font style taking into account the 5-point Likert scale ratings. *Like* merges ratings 4 and 5, *Ambivalent* equals rating 3, and *Dislike* merges ratings 1 and 2. Our study corroborates the contradiction of previous studies regarding *serif* and *sans serif* fonts. Participants have similar opinions toward fonts styles. Among the fonts tested in this study, the roman fonts were preferred to italics for both groups. People with dyslexia like monospaced

Table X. Percentages of Preferences Per Group and Font Style
(*Like* = {5, 4}, *Ambivalent* = {3}, *Dislike* = {2, 1})

Preferences (%)	**Group D**			**Group N**		
Font Style (%)	*Like*	*Ambivalent*	*Dislike*	*Like*	*Ambivalent*	*Dislike*
[−Italic]	45.24	26.19	28.57	46.46	28.28	25.25
[+Italic]	28.57	25.40	46.03	31.31	29.29	39.39
[−Serif]	55.95	28.57	15.48	62.69	30.60	6.72
[+Serif]	42.06	26.19	31.74	53.61	37.11	9.28
[−Monospace]	42.06	26.19	31.74	52.53	37.37	10.10
[+Monospace]	47.62	16.67	35.71	33.33	24.24	42.42
[−Dyslexic]	55.95	28.57	15.48	63.64	30.30	6.06
[+Dyslexic]	21.43	30.95	47.62	15.15	18.18	66.67
[−Dyslexic It.]	33.33	23.81	42.86	48.48	30.30	21.21
[+Dyslexic It.]	21.43	23.81	54.76	12.12	18.18	69.70

fonts (47.62%) and fonts that are not specifically designed for people with dyslexia (55.95%). Note that [+Serif] and [−Monospace] have the same percentages because they represent the same set of fonts: *CMU*, *Garamond,* and *Times*.

The fonts designed specifically for dyslexia, *OpenDys* and *OpenDys It.*, did not lead to better or worse readability. As De Leeuw [2010] shows, *OpenDys* did not lead to faster reading. However, we did not perform a reading out loud test, which might improve with the use of specially designed fonts [De Leeuw 2010]. For participants without dyslexia, *OpenDys It.* led to shorter reading times than *Arial It,* and the non-dyslexics fonts *Arial*, *Helvetica*, *Myriad*, and *Verdana* led to shorter fixation durations compared to *OpenDys*. In addition, both groups significantly preferred *Verdana* or *Helvetica* over *OpenDys*, and *Verdana*, *Helvetica,* and *Arial OpenDys It*. Participants without dyslexia were more extreme with their preferences and also preferred *Arial It.*, *CMU*, *Myriad*, *Times,* and *Garamond* over *OpenDys It.*, even if they objectively read faster with *OpenDys It.* compared to *Arial It.*

Although *Arial* had the shortest reading time for both groups and is highly recommended in literature for dyslexia [British Dyslexia Association 2012; Evett and Brown 2005; Lockley 2002], we cannot conclude that this font type leads to better readability because we only found significant differences with respect to *OpenDys It.* and *Arial It* in participants with dyslexia. However, for people with and without dyslexia, *Arial It.* did lead to significantly longer reading times than *Helvetica*, *Arial*, and *CMU*. It also led to significantly longer fixation durations than most of the fonts. Hence, we recommend avoiding *Arial It*. Moreover, participants with dyslexia significantly preferred *Arial* to *Arial It*.

Among the tested fonts, the two fonts that led to shorter fixation durations were *Courier* and *Helvetica*. Hence, the use of these fonts might help people with dyslexia to read faster on screen. This is consistent with the recommendation of Ability Net [2013] to use *Courier* and with Lockley [2002] to use sans serif fonts in the case of *Helvetica*. Also, *Helvetica* was the second most significantly preferred font by our participants after *Verdana*.

Regarding reading time, more significant differences were found for participants without dyslexia, whereas, regarding fixation duration, more effects were found within the participant with dyslexia. Similar to participants with dyslexia, *Arial It.* had longer reading times, *Arial* and *CMU* presented shorter reading times, and shorter fixation durations were found using *Courier*. The discordant font was *Verdana*. While *Verdana* did not lead to shorter reading times for people with dyslexia and even presented shorter fixation durations than *Arial It.*, it seems to have the opposite effect for people

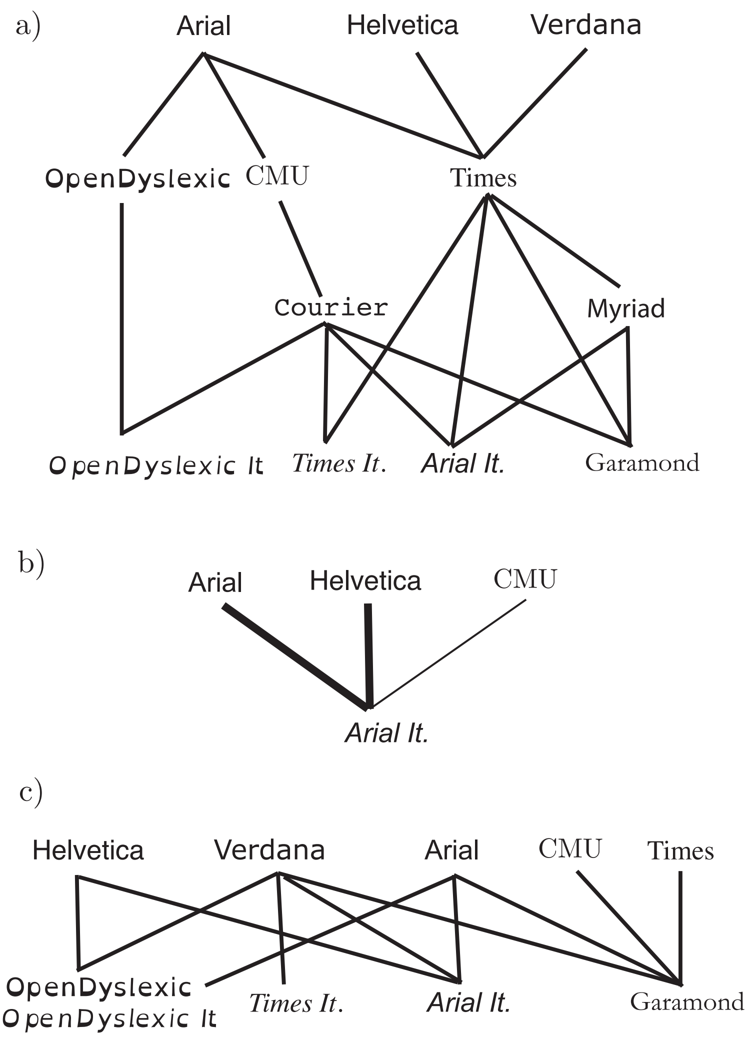

Fig. 12. Partial order (group D) obtained from the means order of *Reading Time* and *Preference Rating* (a), and the partial order for significant differences in *Reading Time* (b) and *Preference Rating* (c). Notice that in each edge of a partial order, the node above is better than the node below.

without dyslexia. For the control group, *Verdana* had the second longest reading time mean and was significantly longer compared to *Arial*, *CMU*, *Courier*, *Garamond*, *Helvetica*, and *Myriad*. Surprisingly, *Verdana* was preferred over *Courier*, which objectively led to lower fixations in people without dyslexia. However, *Verdana* also had the second lowest fixation duration mean for participants without dyslexia.

One way to understand these results is to build the partial order obtained by considering all the order relations that are valid for the average values in *Reading Time* and the *Preference Ratings*. The result is given in Figure 12(a), where the fonts can be grouped in four different levels, where higher the level implies being better. However, not all of these order relations are significant. Hence, the partial orders (b) and (c) show the significant relations for *Reading Time* and *Preference Ratings*, respectively. In the case of (b), thicker lines indicate that those relations are also significant for *Fixation*

Duration. From these partial orders, the only three fonts that are not dominated in both partial orders (b) and (c) are *Helvetica*, *CMU*, and *Arial*. These can be considered good fonts for dyslexia when we also consider the subjective preferences of the participants. The next two in importance are *Verdana* and *Times*.

7. LIMITATIONS

Our study isolates and tests certain variables of typeface and font style. Although the approach is scientific, this process has its limitations. The main limitations are (i) the participant's age; (ii) typographic variables might interact with each other; (iii) the nature of the typefaces does not allow a truly factorial design and analysis; (iv) the results only apply to our experimental conditions (i.e., they could be text dependent; (v) the effect of typeface on comprehension reminds unsolved; and (vi) the wide varieties of difficulties associated with dyslexia.

First, there is almost a 10-year difference in the average ages between the groups with and without dyslexia. Because reading ability and age are correlated, the age difference could have played a role. However, that role would have had a limited impact because reading acquisition normally finishes around nine years of age [Moats 2005], and our younger participant was 11 years old.

Second, typographic variables might interact with each other. Whereas some studies found interactions of the parameters [Tinker 1963; Bernard et al. 2003], others did not [Beymer et al. 2008]. For instance, Bernard et al. [2003] found that typeface is interdependent with font size. The authors compared two fonts—*Arial* and *Times*– and two font sizes—10 and 12 points—with 35 participants. The experiment measured reading time, preference, and errors while reading the text out loud. The 10-point *Arial* typeface was read more slowly than the other conditions, and the 12-point *Arial* typeface was preferred to the other typefaces. Moreover, font size can result in different letter sizes for different font types,[9] so parts of the observed effects might be due to the actual size of the letters. To the contrary, using eye-tracking, Beymer et al. [2008] compared font size and font type and found no significant effects. The authors studied the effect of font size and font type in online reading. The texts were presented to 82 participants using a variety of point sizes, sans serif, and serif fonts. While using smaller font size (10 points), fixation durations were significantly longer as compared to 14 points, whereas there were no significant differences in serif vs. sans serif fonts. In addition, note that if the column width is fixed, the number of characters per column depends on the font size. For instance, Tinker [1963] showed that long lines, very short lines, small type size, and combinations of these led to significantly slower readings. In contrast, Shaikh [2005] suggested that the thickest line width led to faster reading speeds. Further experiments addressing such interactions would be needed to generalize our results.

Third, comparisons between font styles such as [±Italic] or [±Serif] pose constraints because the nature of our set of typefaces does not allow a truly factorial design and analysis. For example, there is no monospaced version of *Arial*. We compared the most homogenous typeface groups we could find in our set of fonts. For instance, [+Italic] fonts were only compared to their corresponding [−Italic] fonts, whereas [+Dyslexic] fonts were only compared to *sans serif* fonts because [+Dyslexic] fonts are *sans serif*. Similarly, the [±Serif] comparison included no *italics*, *dyslexic,* nor *monospaced* fonts to try to preserve homogeneity as much as possible. Since the nonparametric tests are fairly robust to unbalanced cases, we believe the results are useful for the set of fonts tested in this study. However, generalizations to printed fonts or different screen technology should still be regarded as tentative.

[9] See Figure 1 in Boyarski et al. [1998], who compared *Times* with *Georgia* and *Verdana*.

Fourth, our experiment provided data on readings of texts of 60 words. When using eye-tracking to study reading, it has been found that the initially measured fixation durations are longer since users are still in a familiarization phase [Nielsen 2006; Nielsen and Pernice 2010]. However, fixation durations normalize when reading on, and, despite the short lengths of the texts, our findings have high validity [Buscher et al. 2010]. On the other hand, the results depend on the kind of texts used in the experiment. These findings could not be extended to texts belonging to other genres or having other degrees of complexity, since readers—with and without dyslexia—are sensitive to word length and frequency [Rello et al. 2013a].

Another limitation that would be worth study in future work is the impact of typeface on comprehension. In this work, comprehension was only used as a control variable using only literal questions. These questions had no prior validation apart from the authors criteria. Therefore, it cannot be fully guaranteed that these comprehension questions completely cover the comprehension of the test rather than simple word recall. Despite this limitation, we decided to use literal comprehension questions instead of inferential questions because the texts were short and the experiment was already long. In addition to that, the inclusion of inferential questions would have increased its complexity. Hence, we believe that future approaches will be further enriched by controlling and studying comprehension.

The final limitation comes from the fact that dyslexia hardly ever occurs in isolation. There are other visual difficulties associated with dyslexia [Evans 2001] that could be alleviated by modifications of the visual display, such as the visual stress syndrome (Meares-Irlen syndrome) [Kriss and Evans 2005]. The Meares-Irlen syndrome is characterized by symptoms of visual stress and visual perceptual distortions that are alleviated by using individually prescribed colored filters. Moreover, there are many people, such as persons with impaired color vision or with atypical color perception, who have significant problems reading fonts with sub-pixel rendering [Lee 2012]. Interrelations between these difficulties and dyslexia shall be addressed in future work, although in our study no subject had a visual difficulty diagnosed.

8. CONCLUSION AND FUTURE WORK

For this study, we tested the effect of 12 font types on objective readability and preferences on screen of people with and without dyslexia. The main conclusion is that font types have an impact on readability for people with and without dyslexia. Considering the tested fonts, for people with dyslexia, the fonts that significantly led to better objective readability were *Courier*, *Helvetica*, *Arial* (shorter fixation durations), *CMU (Computer Modern Unicode)*, and *Helvetica* (shorter reading times). For people without dyslexia, the fonts that led to better objective readability were *Arial*, *CMU* (shorter reading times), *Courier*, and *Verdana* (shorter fixation durations). The fonts that people with or without dyslexia significantly preferred were the same: *Verdana*, *Helvetica*, and *Arial*.

For the tested fonts, non-italic fonts (*roman* fonts) led to better reading performance for people with dyslexia (shorter fixations) and without dyslexia (shorter reading times). Consistently, both groups significantly preferred fonts in *roman* than fonts in *italic*. For the tested fonts, *sans serif* fonts led to better reading performance for people with dyslexia (shorter fixations) and without dyslexia (shorter fixations). Only participants with dyslexia significantly preferred *sans serif* fonts among the fonts tested. For the tested fonts, *Monospaced* fonts led to better objective readability by people with dyslexia (shorter reading times and fixations) and without dyslexia (shorter fixations). However, participants without dyslexia preferred *proportional* fonts, not monospaced. No effect regarding preferences were found for people with dyslexia.

Table XI. Summary of the Results and Typeface Recommendations for the Tested Fonts

		Values with positive effects for people	
Condition	**Measures**	**with Dyslexia**	**without Dyslexia**
Font	Obj. Readability	*Arial* *Courier* *CMU* *Helvetica*	*Arial* *Courier* *CMU*
	Preferences	*Verdana* *Helvetica* *Arial*	*Verdana* *Helvetica* *Arial*
	Recommendation: Arial, Courier, CMU, Helvetica, and Verdana		
	Obj. Readability	*roman** *sans serif** *monospaced**	*roman* *sans serif** *monospaced**
	Preferences	*roman* *sans serif* Not Sig.	*roman* Not Sig. *proportional*
	Recommendation: roman, sans serif, and monospaced		

Values marked with "*" should be considered as weak recommendations since significant effects were only found for one measure (*Fixation Duration*) out of the three measures used for objective readability.

Among the tested fonts, good on-screen fonts for people with dyslexia are *Helvetica*, *Courier*, *Arial*, *Verdana,* and *CMU*, taking into consideration both reading performance and subjective preferences. Also, for the tested fonts, *sans serif*, *monospaced*, and *roman* font types increased reading performance significantly, whereas *italic* fonts decreased reading performance. In particular, *Arial It.* should be avoided since it significantly decreases readability. What is good for people with dyslexia regarding font types is also good for people without dyslexia. See the recommendations in Table XI.

These findings can have an impact on recommendations for screen text presentation and on the text options chosen by developers, designers, or content producers when they target people with dyslexia. In fact, these findings have been integrated into three reading tools: the *IDEAL eBook Reader*[10] [Kanvinde et al. 2012] for Android, the *DysWebxia Reader* for iOS [Rello et al. 2013b], and in the web service *Text4All*[11] [Topac 2012]. These tools modify text layout for people with dyslexia. Using fonts that are good for people with dyslexia improves accessibility for a large percentage of the population and are also beneficial for those without dyslexia. Hence, the fonts we recommend should be used in practice. At the same time, most popular e-readers, such as the *Kindle reading software*[12] and *iBooks*,[13] allow users to customize most of the crucial text presentation parameters, including fonts. Similarly, popular browsers such as *Firefox*[14] and *Chrome*[15] have ways to set font styles. Moreover, there are a number of accessibility tools that allow this type of text customization, such as *SeeWord* for *Microsoft Word* [Gregor et al. 2003] and the *Firefox* extension *Firefixia* [Santana et al. 2013].

[10]https://play.google.com/store/apps/details?id=org.easyaccess.epubreader.
[11]http://www.text4all.net/dyswebxia.html.
[12]www.amazon.com/kindle.
[13]https://itunes.apple.com/en/app/ibooks/id364709193?mt=8.
[14]https://support.mozilla.org/en-US/kb/change-fonts-and-colors-websites-use.
[15]https://support.google.com/chrome/answer/95416?hl=en.

Future challenges involve studying the effect of other fonts that are specifically designed for people with dyslexia as well as the effect of font type on comprehension and in different contexts, devices, and paper-based formats. Since what is good for people with dyslexia regarding font types is also good for people without dyslexia, these recommendations should be easily applicable to educational settings with a low risk of harming any other reading groups.

ACKNOWLEDGMENTS

We thank Mari-Carmen Marcos for her assistance with the eye-tracker hardware and Martin Pielot for his help with statistics. We also thank the anonymous reviewers for their valuable comments on the article.

REFERENCES

Ability Net. 2013. NHS Choices: Help with accessibility, Changing fonts on your computer. Retrieved from http://www.nhs.uk/accessibilityhelp/index.html.

D. Adler-Grinberg and L. Stark. 1978. Eye movements, scan paths, and dyslexia. *American Journal of Optometry and Physiological Optics* 55, 8 (1978), 557–570.

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. American Psychiatric Publishing, Inc., New York.

Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 1 (1995), 289–300.

Yoav Benjamini and Daniel Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 29, 4 (2001), 1165–1188.

Michael L. Bernard, Barbara S. Chaparro, Melissa M. Mills, and Charles G. Halcomb. 2003. Comparing the effects of text size and format on the readability of computer-displayed times new roman and arial text. *International Journal of Human-Computer Studies* 59, 6 (2003), 823–835.

V. W. Berninger and B. J. Wolf. 2009. *Teaching Students with Dyslexia and Dysgraphia: Lessons from Teaching and Science*. Paul H Brookes Pub Co., Baltimore, MD.

David Beymer, Daniel Russell, and Peter Orton. 2008. An eye tracking study of how font size and type influence online reading. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 2 (BCS-HCI'08)*. British Computer Society, Swinton, UK, 15–18.

Carlo E. Bonferroni. 1936. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del Regio Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8 (1936), 3–62.

Dan Boyarski, Christine Neuwirth, Jodi Forlizzi, and Susan Harkness Regli. 1998. A study of fonts designed for screen display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, 87–94. DOI:http://dx.doi.org/10.1145/274644.274658

J. Boynton, H. Hersh, and T. Tullis. 1995. Readability of fonts in the windows environment interactive posters. In *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, Vol. 2. ACM, Denver, CO, 127–128.

British Dyslexia Association. 2012. Dyslexia Style Guide. Retrieved from http://www.bdadyslexia.org.uk/.

Nicola Brunswick. 2010. Unimpaired reading development and dyslexia across different languages. In *Reading and Dyslexia in Different Orthographies*, Sine McDougall and Paul de Mornay Davies (Eds.). Psychology Press, Hove, 131–154.

Georg Buscher, Ralf Biedert, Daniel Heinesch, and Andreas Dengel. 2010. Eye tracking analysis of preferred reading regions on the screen. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. ACM, Atlanta, GA, 3307–3312.

B. Caldwell, M. Cooper, L. G. Reid, and G. Vanderheiden. 2008. Web content accessibility guidelines (WCAG) 2.0. (2008).

M. S. Carrillo, J. Alegría, P. Miranda, and Sánchez Pérez. 2011. Evaluación de la dislexia en la escuela primaria: Prevalencia en Español (evaluation of dyslexia in primary school: The prevalence in Spanish). *Escritos de Psicología (Psychology Writings)* 4, 2 (2011), 35–44.

C. Chapman. 2011. The Most Popular Fonts Used by Designers. Retrieved from http://www.webdesignerdepot.com/2011/08/the-most-popular-fonts-used-by-designers/.

M. Coleman and T. L. Liau. 1975. A computer readability formula designed for machine scoring. *Journal of Applied Psychology* 60, 2 (1975), 283.

F. Cuetos and F. Valle. 1988. Modelos de lectura y dislexias (reading models and dyslexias). *Infancia y Aprendizaje (Infancy and Learning)* 44 (1988), 3–19.

Rudi W. De Lange, Hendry L. Esterhuizen, and Derek Beatty. 1993. Performance differences between times and helvetica in a reading task. *Electronic Publishing* 6, 3 (1993), 241–248.

R. De Leeuw. 2010. *Special Font for Dyslexia?* Master's thesis. University of Twente.

A. Dickinson, P. Gregor, and A. F. Newell. 2002. Ongoing investigation of the ways in which some of the problems encountered by some dyslexics can be alleviated using computer techniques. In *Proceedings of ASSETS'02*. ACM Press, Edinburgh, Scotland, 97–103.

G. F. Eden, J. F. Stein, H. M. Wood, and F. B. Wood. 1994. Differences in eye movements and reading problems in dyslexic and normal children. *Vision Research* 34, 10 (1994), 1345–1358.

R. D. Elterman, L. A. Abel, R. B. Daroff, L. F. Dell'Osso, and J. L. Bornstein. 1980. Eye movement patterns in dyslexic children. *Journal of Learning Disabilities* 13, 1 (1980), 16–21.

Mariano Fernández Enguita, Luis Mena Martínez, and Jaime Riviere Gómez. 2010. *Fracaso y Abandono Escolar en España (School Failure in Spain)*. Obra Social, Fundación la Caixa, Barcelona, Spain.

Bruce Evans. 2001. *Dyslexia and Vision*. Vol. 5. Wiley, New Jersey.

L. Evett and D. Brown. 2005. Text formats and web design for visually impaired and dyslexic readers-clear text for all. *Interacting with Computers* 17 (2005), 453–472. Issue 4.

FEDIS. 2008. Presidente de la FEDIS (Federación Española de Dislexia) asiste a las Jornadas de León. (Januray 2008). http://actualidadfedis.blogspot.com.es/.

R. Flesch. 1948. A new readability yardstick. *Journal of Applied Psychology* 32, 3 (1948), 221.

Rosa Galván Gómez. 2010. Dislexia va en aumento (Dyslexia is increasing). El Peruano (The Peruvian, newspaper). (January 2010). page 22.

N. E. Goulandris. 2003. *Dyslexia in Different Languages: Cross-Linguistic Comparisons*. Whurr Publishers, London.

P. Gregor, A. Dickinson, A. Macaffer, and P. Andreasen. 2003. SeeWord: A personal word processing environment for dyslexic computer users. *British Journal of Educational Technology* 34, 3 (2003), 341–355.

Peter Gregor and A. F. Newell. 2000. An empirical investigation of ways in which some of the problems encountered by some dyslexics may be alleviated using computer techniques. In *Proceedings of ASSETS'00 (ASSETS 2000)*. ACM Press, New York, NY, 85–91.

Robert Hillier. 2008. Sylexiad: A typeface for the adult dyslexic reader. *Journal of Writing in Creative Practice* 1, 3 (2008), 275–291.

Yosef Hochberg. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75, 4 (1988), 800–802.

Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6, 2 (1979), 65–70.

Gerhard Hommel. 1988. A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika* 75, 2 (1988), 383–386.

B. Hornsby. 1986. *Overcoming Dyslexia*. Martin Dunitz, London.

J. Hyönä and R. K. Olson. 1995. Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 6 (1995), 1430.

Albrecht Werner Inhoff and Keith Rayner. 1986. Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics* 40, 6 (1986), 431–439.

Interagency Commission on Learning Disabilities. 1987. *Learning Disabilities: A Report to the U.S. Congress*. Government Printing Office, Washington DC.

International Dyslexia Association. 2011a. Definition of Dyslexia: http://interdys.org/DyslexiaDefinition.htm. (2011). Based in the initial definition of the Research Committee of the Orton Dyslexia Society, former name of the IDA, done in 1994.

International Dyslexia Association. 2011b. Frequently asked questions about dyslexia. (2011). http://www.interdys.org/.

J. E. Jiménez, R. Guzmán, C. Rodríguez, and C. Artiles. 2009. Prevalencia de las dificultades específicas de aprendizaje: La dislexia en español (the prevalence of specific learning difficulties: Dyslexia in Spanish). *Anales de Psicología (Annals of Psychology)* 25, 1 (2009), 78–85.

M. A. Just and P. A. Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review* 87 (1980), 329–354.

G. Kanvinde, L. Rello, and R. Baeza-Yates. 2012. IDEAL: A Dyslexic-friendly e-book reader (poster). In *Proceedings of ASSETS'12*. ACM Press, Boulder, CO, 205–206.

S. K. Katusic, R. C. Colligan, W. J. Barbaresi, D. J. Schaid, and S. J. Jacobsen. 2001. Incidence of reading disability in a population-based birth cohort, 1976-1982, Rochester. *Mayo Clinic Proceedings* 76, 11 (2001), 1081.

Donald E. Knuth. 1986. *The Computer Modern Fonts, Volume E of Computers and Typesetting*. Addison-Wesley, Reading, MA.

Isla Kriss and Bruce J. W. Evans. 2005. The relationship between dyslexia and Meares-Irlen syndrome. *Journal of Research in Reading* 28, 3 (2005), 350–364.

S. Kurniawan and G. Conroy. 2006. *Comparing Comprehension Speed and Accuracy of Online Information in Students with and without Dyslexia*. Idea Group, Hershey, PA, 257–70.

Anthony Lee. 2012. Accessibility issues due to sub-pixel rendering. In *Text Customization for Readability, W3C WAI RDWG Online Symposium*. W3C.

Lester A. Lefton, Richard J. Nagle, Gwendolyn Johnson, and Dennis F. Fisher. 1979. Eye movement dynamics of good and poor readers: Then and now. *Journal of Literacy Research* 11, 4 (1979), 319–328.

Sharon Lockley. 2002. Dyslexia and higher education: Accessibility issues. The Higher Education Academy. (2002). https://www.yumpu.com/en/document/view/4172057/dyslexia-and-higher-education-accessibility-issues-sharon-lockley.

G. R. Lyon. 1995. Toward a definition of Dyslexia. *Annals of Dyslexia* 45, 1 (1995), 3–27.

G. R. Lyon, S. E. Shaywitz, and B. A. Shaywitz. 2003. A definition of Dyslexia. *Annals of Dyslexia* 53, 1 (2003), 1–14.

F. J. Martos and J. Vila. 1990. Differences in eye movements control among dyslexic, retarded and normal readers in the Spanish population. *Reading and Writing* 2, 2 (1990), 175–188.

Louisa Cook Moats. 2005. *The Speech Sounds of English: Phonetics, Phonology, and Phoneme Awareness (Module 2)*. Technical Report. Sopris West Educational Services.

Sandra E. Moriarty and Edward C. Scheiner. 1984. A study of close-set text type. *Journal of Applied Psychology* 69, 4 (1984), 700.

R. I. Nicolson and A. J. Fawcett. 2011. Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex* 47, 1 (2011), 117–127.

Jakob Nielsen. 2006. F-shaped pattern for reading web content. https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/. (2006). Last visited April 13, 2016.

Jakob Nielsen and Kara Pernice. 2010. *Eye Tracking Web Usability*. New Riders Pub, San Francisco, CA.

Beth A. O'Brien, J. Stephen Mansfield, and Gordon E. Legge. 2005. The effect of print size on reading speed in dyslexia. *Journal of Research in Reading* 28, 3 (2005), 332–349.

R. K. Olson, R. Kliegl, and B. J. Davidson. 1983. Dyslexic and normal readers' eye movements. *Journal of Experimental Psychology: Human Perception and Performance* 9, 5 (1983), 816.

Donald G. Paterson and Miles A. Tinker. 1932. Studies of typographical factors influencing speed of reading. X. style of type face. *Journal of Applied Psychology* 16, 6 (1932), 605.

Francis J. Pirozzolo and Keith Rayner. 1978. The neural control of eye movements in acquired and developmental reading disorders. *Studies in Neurolinguistics* 4 (1978), 1–27.

Eustace Christopher Poulton. 1965. Letter differentiation and rate of comprehension in reading. *Journal of Applied Psychology* 49, 5 (1965), 358.

R Development Core Team. 2011. A language and environment for statistical computing (Computer Software). (2011). http://www.R-project.org.

Gary E. Raney and Keith Rayner. 1995. Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology* 49, 2 (1995), 151.

Keith Rayner. 1986. Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology* 41, 2 (1986), 211–236.

K. Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124 (1998), 372–422.

K. Rayner and S. A. Duffy. 1986. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition* 14, 3 (1986), 191–201.

Keith Rayner and Gary E. Raney. 1996. Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review* 3, 2 (1996), 245–248.

L. Rello. 2014. *DysWebxia. A Text Accessibility Model for People with Dyslexia*. Ph.D. Dissertation. Universitat Pompeu Fabra.

L. Rello and R. Baeza-Yates. 2013. Good fonts for dyslexia. In *Proceedings of ASSETS'13*. ACM Press, Bellevue, WA, 14–21.

L. Rello, R. Baeza-Yates, L. Dempere, and H. Saggion. 2013a. Frequent words improve readability and short words improve understandability for people with dyslexia. In *Proceedings of INTERACT'13*. IFIP, Cape Town, South Africa, 203–219.

L. Rello, R. Baeza-Yates, H. Saggion, C. Bayarri, and S. D. J. Barbosa. 2013b. An iOS reader for people with dyslexia (demo). In *Proceedings of ASSETS '13*. ACM Press, Bellevue, Washington, WA, 37–38.

L. Rello, S. Bautista, R. Baeza-Yates, P. Gervás, R. Hervás, and H. Saggion. 2013c. One half or 50%? An eye-tracking study of number representation readability. In *Proceedings of INTERACT'13*. IFIP, Cape Town, South Africa, 229–245.

L. Rello, G. Kanvinde, and R. Baeza-Yates. 2012. Layout guidelines for web text and a web service to improve accessibility for dyslexics. In *Proceedings of W4A '12*. ACM Press, Lyon, France, 1–8.

L. Rello, M. Pielot, and M. C. Marcos. 2015. Make it big! The effect of font size and line spacing on online readability. In *Proceedings of CHI 16*. ACM, San Jose, CA.

L. Rello, M. Pielot, M. C. Marcos, and R. Carlini. 2013. Size matters (spacing not): 18 points for a dyslexic-friendly wikipedia. In *Proceedings of W4A'13*. ACM Press, Rio de Janeiro, Brazil, 1–8.

C. Romani, J. Ward, and A. Olson. 1999. Developmental surface dysgraphia: What is the underlying cognitive impairment? *Quarterly Journal of Experimental Psychology* 52, 1 (1999), 97–128.

H. Saggion, S. Štajner, S. Bott, S. Mille, L. Rello, and B. Drndarevic. 2015. Making it simplext: Implementation and evaluation of a text simplification system for Spanish. *ACM Transactions on Accessible Computing (TACCESS)* 6, 14 (2015), 1–14.

L. Sánchez Sampedro. 2012. *Impostores (Impostors)*. Crowdfunding, Madrid, Spain. Retrieved from http://www.impostores.es.

V. F. Santana, R. Oliveira, L. D. A. Almeida, and M. Ito. 2013. Firefixia: An accessibility web browser customization toolbar for people with dyslexia. In *Proceedings of W4A'13*. ACM, Rio de Janeiro, Brazil, 1–8.

Francisca Serrano and Sylvia Defior. 2008. Dyslexia speed problems in a transparent orthography. *Annals of Dyslexia* 58, 1 (2008), 81–95.

A. Dawn Shaikh. 2005. The effects of line length on reading online news. *Usability News* 7, 2 (2005), 1–4.

S. E. Shaywitz, M. D. Escobar, B. A. Shaywitz, J. M. Fletcher, and R. Makuch. 1992. Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine* 326, 3 (1992), 145–150.

Speech Therapy Association of Catalonia. 2011. *PRODISCAT: Protocol de Detecció i Actuació en la Dislèxia. Àmbit Educativo (Protocol for detection and management of dyslexia. Educational scope)*. Education Department of the Catalonian government, Barcelona, Spain.

Joy Sykes. 2008. *Dyslexia, Design and Reading: Making Print Work for College Students with Dyslexia: A Qualitative Interaction Design Study*. Ph.D. Dissertation. Carnegie Mellon School of Design.

Miles A. Tinker. 1963. Influence of simultaneous variation in size of type, width of line, and leading for newspaper type. *Journal of Applied Psychology* 47, 6 (1963), 380.

Miles Albert Tinker. 1965. *Bases for Effective Reading*. Vol. 1. University of Minnesota Press Minneapolis.

Tobii Technology. 2005. Product description of Tobii 50 Series. (2005).

V. Topac. 2012. The development of a text customization tool for existing web sites. In *Text Customization for Readability Symposium*. W3C, New York, USA.

F. R. Vellutino, J. M. Fletcher, M. J. Snowling, and D. M. Scanlon. 2004. Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry* 45, 1 (2004), 2–40.

S. Williams, E. Reiter, and L. Osman. 2003. Experiments with discourse-level choices and readability. In *Proceedings of Enlg'03*. ACL, Budapest, Hungary, 127134.

World Health Organization. 1993. *International Statistical Classification of Diseases, Injuries and Causes of Death (ICD-10)* (10th ed.). World Health Organization, Geneva, Switzerland.

M. Zorzi, C. Barbiero, A. Facoettia, I. Lonciari, M. Carrozzi, M. Montico, L. Bravar, F. George, C. Pech-Georgel, and J. Ziegler. 2012. Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences* 109 (2012), 11455–11459.

Received June 2014; revised February 2016; accepted February 2016