An Eye-Tracking Study on Text Accessibility and Comprehension

Sergio Navas-León(✉) and Jon Andoni Duñabeitia

Universidad Nebrija, Madrid, Spain
snavas@nebrija.es

Abstract. As digital reading becomes a central component of everyday information access, the design of accessible and cognitively efficient text formats has gained increased relevance. Easy-to-Read (E2R) formatting seeks to support comprehension through simplified structure, enhanced spacing, and visual cues. This study examined whether E2R formatting benefits cognitively healthy readers, compared to a standard Control format. Fifteen Spanish university students read texts in both conditions while eye movements were recorded. Eye-tracking metrics were collected alongside comprehension accuracy. Although comprehension accuracy did not differ significantly between conditions, a small numerical advantage was observed for E2R. Total reading time was also comparable across formats. Crucially, participants in the E2R condition exhibited significantly shorter fixation durations that are indicative of more efficient visual processing. These findings suggest that accessible design can enhance processing without compromising efficiency and support the broader applicability of E2R guidelines beyond their traditional target groups.

Keywords: Accessibility · Eye-tracking · Psycholinguistics

1 Background

As digital access to essential services increases, reading on the web has become a central part of daily life, especially in healthcare and public information [1]. This growing reliance on digital texts has intensified efforts to improve accessibility and cognitive efficiency through formatting strategies [1]. In this context, Easy-to-Read (hereafter, E2R) guidelines promote clarity and accessibility through linguistic simplification, spatial segmentation, and the inclusion of supportive visual markers such as icons or lists [2]. E2R supports understanding by explicitly structuring information and guiding attention through visual and textual cues [2].

In Spain, this approach has been formalised through two complementary frameworks developed by the Spanish Association for Standardisation (UNE). The UNE 153101:2018 EX [2] provides national recommendations for producing E2R content in Spanish, incorporating input from organizations representing people with intellectual disabilities, older adults, and other stakeholders. Additionally, the UNE-EN 301549:2022 [3], aligned with the Web Content Accessibility Guidelines (WCAG 2.1),

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
J. Tornero et al. (Eds.): IBERDISCAP 2025, BIOSYSROB 36, pp. 226–233, 2026.
https://doi.org/10.1007/978-3-032-18746-8_31

defines digital accessibility requirements for websites and electronic documents and is enforced through Royal Decree 1112/2018 [4, 5].

Despite the growing interest in E2R, to the best of our knowledge, no previous studies have empirically examined whether this format influences reading behavior and comprehension compared to conventional text layouts in cognitively healthy readers. Understanding this effect is a necessary step before assessing E2R's impact on more vulnerable populations. While E2R guidelines have been primarily developed to support individuals with intellectual or developmental disabilities, their systematic evaluation in cognitively and perceptually typical populations remains limited. This is a crucial step for assessing the broader applicability of these guidelines, particularly in e-health contexts and digital cognitive assessment platforms. The shift toward remote health services and online testing accelerated by the COVID-19 pandemic has revealed structural barriers to digital inclusion, even among users without diagnosed impairments [6, 7]. Inadequate formatting and cognitive overload in standard digital interfaces can compromise comprehension and engagement, potentially skewing assessment results or impeding informed consent processes. Validating the benefits of E2R formatting in normotypical users is therefore essential before its implementation across broader public health and clinical settings.

Thus, the present study aims to explore whether the E2R format changes reading patterns and comprehension outcomes when compared to a conventional layout. Specifically, we compared: (i) a Control condition with a compact text block and no graphic aids and (ii) an adapted E2R condition, based on the Spanish accessibility standards cited above, which incorporated simplified structure, increased spacing, and visual supports. We evaluated participants' eye-tracking metrics (e.g., fixations, saccades) alongside comprehension accuracy. Given the theoretical and empirical grounding of E2R in facilitating cognitive processing, we tentatively hypothesized eye-tracking metrics would reflect lower cognitive load (fewer and shorter fixations, longer saccades) in the E2R condition.

2 Methods

2.1 Participants

Young adult volunteers from Nebrija University took part in the study, ranging in age from 18 to 30 years. Recruitment occurred between April and May 2025. Inclusion criteria required (a) normal or corrected-to-normal vision, (b) age between 18 and 35 years, and (c) correct cognitive performance. Exclusion criteria included: (a) self-reported neurological conditions or brain injury, (b) history of mental illness, (c) current or past use of psychotropic or recreational substances, (d) calibration failure or incomplete data, and (e) inability to understand Spanish. All participants met the inclusion criteria, and none were excluded during recruitment. This study was approved by the Institutional Review Board of Universidad Nebrija (code: UNNE-2020-008) and was conducted in line with the Helsinki Declaration.

2.2 Materials and Procedure

Participants read descriptions of different fictional countries each containing 15 short texts covering diverse topics (e.g., economy, education, tourism). The content was refined iteratively to enhance clarity and readability, reducing sentence complexity, and using simpler synonyms. Materials were created in Microsoft PowerPoint and exported as JPEGs for presentation. Two text presentation conditions that systematically varied in their typographic and formatting features were created (see Fig. 1 for a detailed view).

The Control condition employed a sans-serif font (Lucida Sans, size 18), with justified text and no paragraph breaks, single line spacing (1.0), and default word spacing. The background was white (RGB: 255, 255, 255), and the font color was black (RGB: 0, 0, 0), resulting in an adequate contrast ratio. On the other hand, the E2R condition featured a sans-serif font (Lucida Sans, size 18), left-aligned text, increased line spacing (1.5), and expanded word spacing. Paragraph headings were enlarged, and the layout included pictograms alongside bullet points. A high-contrast white background with dark grey text (RGB: 40, 40, 40) was used to enhance readability.

Each participant read 15 texts in each condition, resulting in a total of 30 unique texts across the two conditions. Texts were presented in a blocked fashion: participants completed all texts from one condition before proceeding to the other. A Latin square design was used to counterbalance both the order of presentation and the pairing of countries with specific content. Before beginning the main task, participants completed a brief practice session to familiarise themselves with the procedure.

The experimental session lasted approximately one hour and began with the collection of demographic information, followed by a brief cognitive screening using the Cognitive Assessment Battery (CAB) PRO (CogniFit Inc., San Francisco, US; https://www.cognifit.com/cab). Participants were then seated at a fixed distance of 98 cm from a 27″ Lenovo LED monitor. Head movements were stabilized using a chin rest.

Eye movements were recorded at a sampling rate of 500 Hz using the EyeLink Portable Duo system, which relies on corneal reflection tracking. Calibration was conducted using a randomized 9-point grid, with drift correction performed as necessary. Reading time was measured from stimulus onset until the participant pressed the response key to indicate they had finished reading, as recorded by the Experiment Builder de SR Research (Ontario, Canada).

Once calibrated, participants began the reading task. After completing both reading blocks, participants completed a flanker task (i.e., a cognitive control task in which participants indicate the direction of a central arrow while ignoring surrounding distracting arrows), which served as a distractor, and then they completed a 60-item multiple-choice questionnaire (30 items per condition) designed to assess objective comprehension accuracy, administered digitally on a tablet. Accuracy was calculated as the number of correct responses per condition. All comprehension items were presented in a single set, with both item and response option order randomised.

2.3 Statistical Analysis

Data from the reading task were aggregated by computing the average reading time, fixation count, fixation duration, and comprehension accuracy across the 15 texts per condition for each participant. All comparisons between the Control and E2R conditions were

conducted using parametric paired-sample t-tests, consistent with the within-subject design.

Prior to conducting parametric tests, data for each eye-tracking metric greater than 2.5 SDs from the mean were defined as outliers. Then, we assessed the normality of the data distributions using Shapiro–Wilk tests and visual inspection of Q–Q plots. As all dependent variables met normality assumptions, parametric tests were deemed appropriate.

To complement the frequentist analyses and assess the robustness of our findings, Bayesian paired-samples t-tests were also performed for each outcome measure. These analyses estimate the relative likelihood of the data under the alternative hypothesis (H_1) compared to the null hypothesis (H_0), expressed as Bayes Factors (BF_{10}).

Statistical significance was set at $p < .05$, and effect sizes were calculated using Cohen's d. Effect sizes were interpreted according to conventional thresholds (small = 0.2, medium = 0.5, large = 0.8).

All analyses were performed using JASP (JASP Team 2024, Version 0.19.3).

A post-hoc power analysis was conducted using G*Power 3.1 [8] to estimate the achieved statistical power for each of the paired comparisons reported. The analysis specified a two-tailed paired-samples t-test, with an alpha level of .05 and a total sample size of 15 participants. Observed effect sizes were input individually for each outcome variable. The resulting power estimates were as follows: comprehension accuracy = 10%, fixation count = 45%, fixation duration = 89%, and reading time = 11%.

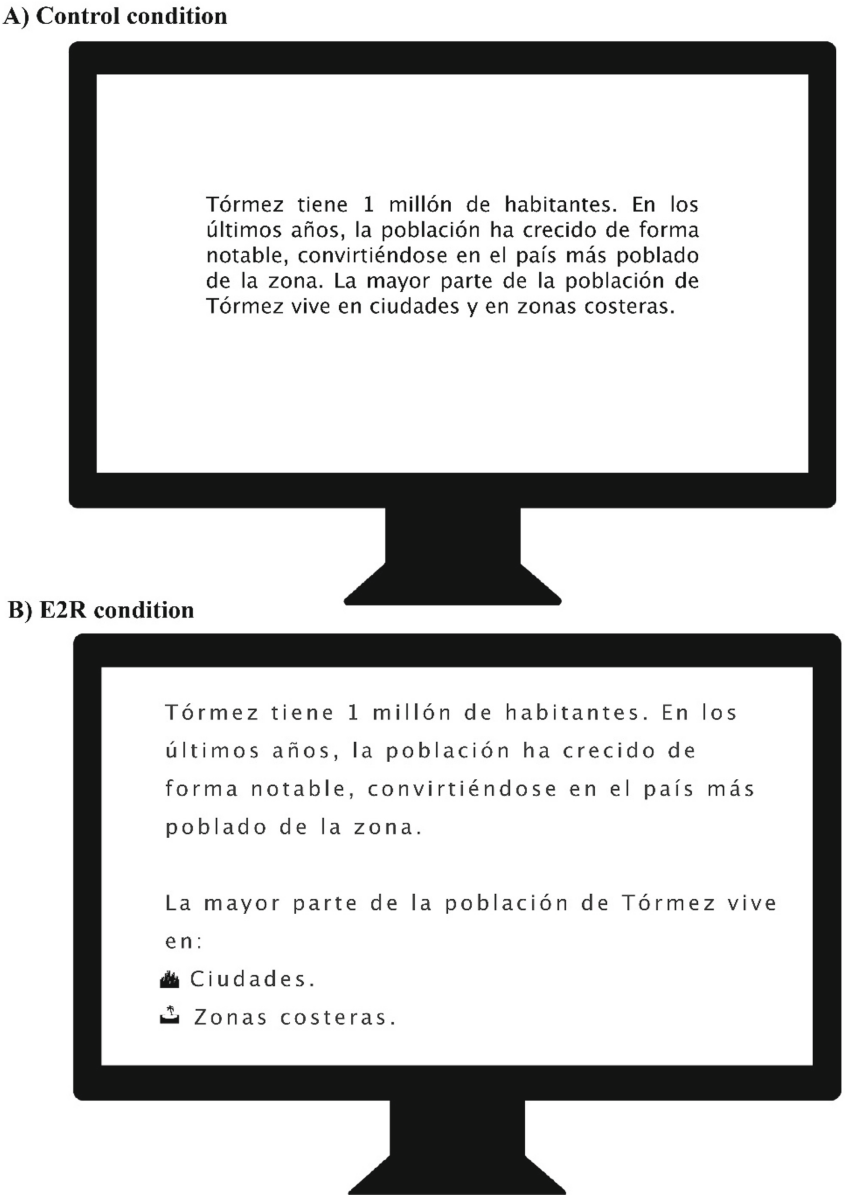


Fig. 1. Panel A shows an example of the Control condition, Panel B shows the E2R condition.

3 Results

A total of 15 participants were included in the final analyses (M = 22.61, SD = 3.71; 11 females).

Participants showed in the Control condition an average of 10,002 ms (SD = 3,565.99) reading the texts, compared to 10,462 ms (SD = 2,916.26) in the E2R condition. This difference was not statistically significant, t (14) = −0.82, p = .427, with a small effect size (d = −0.211). The Bayesian analysis supported this result, yielding a Bayes Factor of $BF_{10} = 0.345$, indicating anecdotal evidence for the null hypothesis. Regarding eye-tracking metrics, participants showed a higher fixation count in the E2R condition (M = 43.39, SD = 10.41) compared to the Control condition, with a marginal effect, t (14) = −1.99, p = .066, and a moderate effect size (d = −0.515). The Bayesian analysis yielded BF_{10}= 1.41, reflecting anecdotal to moderate evidence for the alternative hypothesis. In contrast, the average fixation duration was significantly longer in the Control condition (M = 216.716 ms, SD = 28.00) than in the E2R condition (M = 203.290 ms, SD = 22.688), t (14) = 3.45, p = .004, with a large effect size (d = 0.890). The Bayesian analysis strongly supported this finding, with a Bayes Factor of $BF_{10} = 19.1$, indicating strong evidence in favour of the alternative hypothesis. Finally, comprehension accuracy was slightly higher in the E2R condition (M = 20.29, SD = 7.03) compared to the Control condition (M = 18.79, SD = 6.72), but this difference was not statistically significant, t (13) = 0.72, p = .484, with a small effect size (d = 0.192). The corresponding Bayes Factor was $BF_{10} = 0.338$, suggesting anecdotal support for the null hypothesis.

4 Discussion

The present study aimed to examine whether the E2R format alters reading patterns and comprehension outcomes compared to a conventional text layout. Although comprehension accuracy did not differ significantly between conditions, a slight advantage was observed for the E2R format. Notably, despite including pictograms and occupying more screen space, E2R texts did not increase reading time compared to the more conventional Control format. As González-Sordé and Matamala [9] noted, prior studies found limited evidence that visual aids consistently improve comprehension. However, our results suggest that when visual elements are paired with a structured layout, they can facilitate a more efficient and guided reading process without impairing understanding.

In our study, we interpret reading efficiency as the ability to process and comprehend textual information rapidly and accurately, as reflected by specific eye movement patterns [10]. In this regard, eye-tracking data showed that participants in the E2R condition had significantly shorter fixation durations. Although the fixation count was slightly higher in E2R, this combination of more but shorter fixations may indicate that participants extracted information more rapidly per fixation, consistent with skilled reading patterns. In contrast, the Control condition required longer fixations to extract information effectively. However, it is also plausible that the structured layout and visual elements of the E2R condition prompted increased segmentation or scanning behaviour. Thus, while our results are compatible with improved processing efficiency, alternative interpretations should be considered and explored in future studies.

Taken together, these findings suggest that E2R can enhance comprehension and reduce cognitive effort without sacrificing efficiency. Despite concerns that E2R might be less appealing to cognitively healthy users, our results show potential benefits in

terms of usability and aesthetics [11]. From a practical standpoint, E2R guidelines may improve digital reading experiences, particularly in cognitively demanding contexts such as public websites. While comprehension accuracy did not reach statistical significance in this sample, it is possible that the benefits of E2R formats would be more pronounced in populations with cognitive, perceptual, or reading difficulties, for whom accessibility adaptations are most relevant.

Moreover, our findings support a broader view of accessibility that goes beyond content presentation. In line with standards like [4], we advocate for applying accessibility principles to assessment tools and procedures—such as reading tasks or self-report questionnaires—especially when used with vulnerable populations. Ensuring accessibility in both content and evaluation promotes more inclusive and equitable research practices.

Finally, some limitations must be acknowledged. The sample was small and homogeneous, and the materials consisted of short factual texts presented in a controlled setting. Moreover, comprehension accuracy differences between conditions did not reach statistical significance, which limits the strength of the conclusions in this domain. This lack of significance may be partly explained by the limited statistical power associated with the small sample size, as confirmed by a post-hoc power analysis. Future research should test E2R in more ecologically valid contexts and with more diverse populations, especially those with specific accessibility needs, and include larger samples to ensure sufficient power to detect subtle effects.

References

1. Nomura, M., Skat Nielsen, G., Tronbacke, B.: Guidelines for Easy-to-Read Materials. International Federation of Library Associations and Institutions (IFLA), The Hague (2010). https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/professional-report/120.pdf
2. Asociación Española de Normalización (UNE): Lectura fácil. Pautas y recomendaciones para la elaboración de documentos (UNE 153101:2018 EX). AENOR Internacional SAU, Madrid (2018). https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0060036
3. Asociación Española de Normalización (UNE): UNE-EN 301549:2022. Requisitos de accesibilidad para productos y servicios TIC (Norma Europea EN 301549:2021). UNE, Madrid (2022). https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=0071096
4. Gobierno de España: Real Decreto 1112/2018. BOE núm. 227 (2018). https://www.boe.es/eli/es/rd/2018/09/07/1112
5. Gobierno de España: Ley 6/2022. BOE núm. 78 (2022). https://www.boe.es/buscar/doc.php?id=BOE-A-2022-5140
6. Staccini, P., Lau, A.Y.S.: Consuming health information and vulnerable populations: factors of engagement and ongoing usage. Yearb. Med. Inform. **31**(1), 173–180 (2022). https://doi.org/10.1055/s-0042-1742549
7. Hewitt, K.C., Rodgin, S., Loring, D.W., Pritchard, A.E., Jacobson, L.A.: Transitioning to telehealth neuropsychology service: considerations across adult and pediatric care settings. Clin. Neuropsychol. **34**(7–8), 1335–1351 (2020). https://doi.org/10.1080/13854046.2020.1811891
8. Faul, F., Erdfelder, E., Lang, A.G., Buchner, A.: G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods **39**(2), 175–191 (2007). https://doi.org/10.3758/BF03193146

9. González-Sordé, M., Matamala, A.: Empirical evaluation of Easy Language recommendations: a systematic literature review from journal research in Catalan, English, and Spanish. Univ. Access Inf. Soc. **23**(3), 1369–1387 (2024). https://doi.org/10.1007/s10209-023-00975-2
10. Liu, J.C., Li, K.A., Yeh, S.L., Chien, S.Y.: Assessing perceptual load and cognitive load by fixation-related information of eye movements. Sensors **22**(3), 1187 (2022). https://doi.org/10.3390/s22031187
11. Schmutz, S., Sonderegger, A., Sauer, J.: Implementing recommendations from web accessibility guidelines: would they also provide benefits to nondisabled users. Hum. Factors **58**(4), 611–629 (2016). https://doi.org/10.1177/0018720816640962