Citation	Procedure	Authors measured	Concurrent Measures	Native Language	Reading or Scanning?	Screen Type	Web Factors	Participants
Abubaker and Lu, 2012	Empirical study: participants read fonts in different font sizes, types and line lengths. The authors recorded the voice of the study and recorded the time with digital watch. The authors took notes about errors during reading.	Then the authors measured reading errors and reading speed.	After the study was over, students were asked to answer two questions: (1) which characters are more difficult to read? (2) which font size is easier to read?	Arabic	Reading	Desktop	Font size, line length and Font type	30 students, ages ranged from 10-12
Ali et al., 2013	Empirical study: The authors created reading passages or text blocks using those two fonts. The passages were the same size and approximately the same difficulty to read. They had each participant read both blocks of text and the participant could rest between each activity.	A research assistant recorded the time taken to read the text and the number of errors committed throughout the reading. The authors of this paper scored each of readings based on prior literature that demonstrated how many words and sentences can be detected by the reader and the clarity of vocabulary and grammar in words and verses.		Malay	Reading	Desktop	Font Type	48 undergraduates,majore in Information and Communication Technology (ICT). All participants have experience with using computers.
Ardit and Lu, 2008	No methodology, developed a new technology to support readability for individuals with low vision	N/A		English	Reading	Desktop	N/A	N/A
Banerjee et al., 2011	Empirical study: Within subject design, font conditions were compared by having participants read eighteen passages. The text of each passage comprised of a font from one of the eighteen type and font size conditions.	Reading time was recorded by using a digital stop watch. Accuracy of reading (in terms of "omission" and "misreading") were noted during the time of reading by two experimenters.	The NASA-TLX questionnaire was used to measure the overall mental workload. Although this method provides six dimensions, this experiment used only four dimensions, i.e., mental demand, performance, effort and frustration.	English	Reading	Desktop	Font type and Font Size	40 participants - All had 20/20 or better unaided or corrected vision. Everyone had experience reading documents on a computer screen for varied amount of times.
Bernard et al., 2003	Empirical study: Line conditions were compared by having participants read three passages, each with different line lengths. Both the adults' and children's passages were 12-point Arial, which was black on a white background.	Reading score was measured. The score was the time taken to read the passages divided by the percentage of accurately detected subsuinted words in the passages.	Readability questionnaire with 6 likert questions. The questionnaire focused on ease of use questions and asked participants to preference each of the three conditions	English	Reading	Desktop	Line Length	40 participants (20 children and 20 adults). All adults had experience reading on a computer screen. All participants had 20/40 or better unaided or corrected vision.
Bernard et al., 2002	Methods are the same as Bernard et al., 2001			English	Reading	Desktop	Font type and font size	60 participants with normal or corrected vision
Bernard et al., 2001	Empirical study: Within subjects design (4X2) - participants were asked to read passages with different fonts and with different sizes. The participants were asked about thier perceptions and qualitative feedback. 10 substitute words were taken out the text to make sure passages were actually read.	Prefense was measured using the Friedman C2	Readability questionnaire with 6 likert questions. The questionnaire focused on ease of use questions and asked participants to preference each of the three conditions	English	Reading	Desktop	Font Type and Font Size	27 participants (9-11 year old). All participants had 20/20 or 20/20 corrected vision. Most participants had experience reading text on computer screen.
Beymer et al., 2008	Empirical study: Between-subjects design, each participant was assigned a one-page story and asked to read the story.	Eye-tracking data was collected. Multiple choice post-test of retention after each story		English	Reading	Desktop	Font Size and Font Type	84 participants froma major computer company.
Bhatia et al., 2011	Empirical study: Font tasks were developed to measure effectiveness on the three web factors.	Completion and time taken was used to measure effectiveness and efficiency	Survey questionnaire for satisfaction, 10 questions that asked about ease of use, likeability, structure and overall experience.	English	Does not specify - participants were given multiple webpages and asked to find content on any of the pages then respond to each task. Thus, a participant could scan or read bat it was their own decision.	Desktop	Font Size, Font Style and Color Count	180 young adults (at least 19 years of age) enrolled in courses offered by the Department of Psychology at the University of Nebraska.
Boyarski et al., 1998	Empirical study: Participants were asked to read the text in different font types and styles. Texts were prepared from the Nelson-Denny Test. After each reading, participants completed the comprehension test.	Comprehension and perception was measured	Subjective perception questionnare	English	Reading	Desktop	Font Type	48 participants - university faculty, staff and graduate students
Burmistrov et al., 2016	Eye-tracking: Within-subjects study with independent variables of font weight, background color and contrast between text and background. Participants were asked to search for a target word on a text page.	The authors measured visual sarch and oculomotor indicators - fixation duration and saccade amplitude.		English	Scanning	Desktop	Font style (font weight), background color and contrast	24 participants - experienced internet and text editor users and had normal to corrected visual acuity
Chaparro et al., 2005	Empirical study: Within subject design: Participants were asked to read pages with either an enhanced page layout and a poor page layout. They read the documents and then were given 8 comprehension questions.	The authors measured comprehension and reading performance		English	Reading	Desktop	Whitespace	20 college students with normal or corrected vision. Most participants visited and read from the web daily.
Chen et al., 2014	Empirical study: Participants were asked to read four texts on different layouts (paper, tablet and desktop). Then complete a 5 multiple choice question and a short summary of the text	Reading comprehension and tablet familiarity were measured	Tablet familiarity questionnaire	Chinese	Reading	Desktop and Tablet	N/A	90 second-year college students from Beijing, China.
Darroch et al., 2005	Empirical study: Participants were asked to read different passages on the handheld device in different font sizes. To make sure participants were reading the ask, the authors used word substitute errors in their methodology.	Reading speed and reading accuracy were recorded.	After being presented with an initial set of 16 passages to read users answered questions on what they thought of the different text sizes and were asked to pick a preferred text size by browsing through the passages.	English	Reading	Handheld Device	Font size	Twenty-four participants took part in the experiment. There were young adults and older adults. All participants were fluent in English as their first language and educated to at least secondary/high school level. All participants had 20/40 vision or better. Participants had no or very minimal experience of handheld computers before the experiment.
Dyson, 2004	Literature Review	N/A	N/A	English	Reading	Desktop	Columns, Line length, Window size and Line spacing	N/A
Dyson and Haselgrove, 2001	Empirical study: Participants were asked to read texts. The independent variables were reading speed- fast and normal and line length (3 versions). After reading participants were given comprehension test in the form of multiple choice.	Comprehension, reading rate and scrolling patterns were measured.	After reading participants were given comprehension test in the form of multiple choice. The multiple choice quesitonnaire included Tite questions, Main Idea questions, structure questions, main factual questions, incidental questions and recognition questions	English	Reading and Scanning	Desktop	Line Length	36 participants, undergraduates or postgraduate students at the University
Dyson and Kipping, 1998	Empirical study: Within-subjects design, participants were asked to read texts with different line lengths.			English	Reading	Desktop	Line Length	
Flanders and Willis, 1998	Inspection Methods: Authors conducted a web analysis of different web pages based on current readability guidelines.			N/A	Reading	Desktop	N/A	N/A
Granquist et al., 2018	An online survey was sent out to participantsto arrange a text passage for typical reading and to report viewing distance, screen dimensions, and the number of characters per line.		Normally sighted participants were asked to complete a similar survey but to view it only their device (smart phone, tablet, desktop)	English	Reading	N/A	N/A	75 adult subjects (most with early-onset low vision, few with central field loss). 12 normally sighted controls reported the same information while viewing the passage on cell phones, tablets, and computers.
Hasagawa et al., 2008	Empirical study: Graphic characters were displayed in various fonts on the LCDs and read aloud by subjects.	Reading time (RT), viewing distance (VD) at the end of reading, and the number of errors (Err) were measured and recorded	Subjective evaluation (SE) was performed every time the reading of one sample was completed by selecting on a 1-5 likert scale (very easy to very easy)	Japanese/English	Reading	Desktop	N/A	Experiment 1: 23 Native speakers of English and Experiment 2: 98 Native speakers of Japanese and 55 native speakers of English
Hill and Scharff, 1997	Large scale survey and Empirical study: A large scale survey was used to choose the foregorund/background color combinations. Then participants were asked scan a screen of text and find a target shape word in a within-subjects study.			English	Scanning	Desktop	foreground/background color, Font Types and Word Styles	42 participants. . All participants were tested for normal color-vision, and all participants had 20/20 or corrected to 20/20 vision. Participants were informed of the required procedure and completed a consent form.
Hill, 1994	Empirical study: Within subject design - three independent variables leading to a 3 (background luminance levels) x 2 (text/background combinations) x 6 (luminance contrasts) mixed factorial design. Participants scanned each onscreen text paragraph for the hidden target shape word. Once they located the target word, they quickly and accurately as possible used the mouse to click on the corresponding shape at the bottom of the screen.	Search time and accuracy were measured	Subjective test: The participants rated the stimuli conditions on a Likert scale of 1 to 5 (1 = dislike and 5 = like). They clicked on the number that corresponded to their rating.	English	Scanning	Desktop	foreground-background color combinations, font styles, and font types	Sixteen participants between the ages of 18 and 511 were used in this experiment. All participants had normal color vision and at least 20/20 or corrected to 20/20 vision
Hojjati and Muniandy, 2014	Empirical study: Participants were asked to read 4 passages and respond to comprehension questions. Each passage has a different font type and different spacing.	Reading speed, ease of reading and reading comprehension		Malay		Desktop	Font-type and line-spacing	30 randomly chosen international postgraduate students from a Malaysian University
Holmqvist et al., 2003	Eye-tracking: The authors created two recordings of eye movement data from readers of two net papers and two newspapers. 12 subjects read the net papers and 15 subjects read the newspapers.	Eye movement data (fixation) was collected.	The post-questionnaire asked about their experience. Participants noted that the eye tracking device did not bother their reading in the post-questionnaire.	English	Reading and Scanning	Desktop	N/A	Not described
Hussain et al., 2011	Literature Review			N/A	Reading & Scanning	N/A	color contrast, white space, line spacing, font style, font size, text width, headings, graphics and animation	Broadly, children, teenagers and old age users
Jang et al., 2007	Empirical study: Evaluated the satisfaction frequency of current web-safe colors. Then conducted a readability evaluation to test color contrast. Participants were asked to read sentences presented on the screen with different color contrast combinations	Satisfaction and readability were measured	Subjective questionnaire: he users were asked to choose a readability level on a scale of 1 to 7, from the most clear to the most unclear	English	Reading	Desktop	Color Contrast	87 students of both genders, 10 students diagnosed with a learning disorder at elementary and junior high school. No students with color blindness or color deficiencies joined the test
Legge, 2016	Literature review			English	Reading	N/A	font size, line spacing and color contrast	N/A
Lemmerich et al., 2019	Large scale multiple-choice survey sent to readers of 14 Wikipedia languages and receiving more than 210,000 responses.	Gathered quantative data about reading behaviors through the survey	The authors additionally used quantative data found from Wikipedia logs that trace a sample of users through their usage of the platorm. Furthermore, they use country-level datasets to understnd socio-economic and cultural indicators.	N/A	Reading	N/A	N/A	Wikipedia readers from 14 different language editions

Citation	Procedure	Authors measured	Concurrent Measures	Native Language	Reading or Scanning?	Screen Type	Web Factors	Participants
Li et al., 2019	Inspection method: Li et al. tested how often Reader View finds webpages transformable. Reader View changes the standard reading to web factors that are easier to read such as font size and white space. The authors did their own evaluation on 100 websites. Usability Method: the authors conducted a 10-minute within subjects study to better understand the two conditions of Standard Web Page vs. Reader View. Participants were given the text in each condition then asked to read the webpage word by word and then respond to a few comprehension questions.	Reading Speed, perceived readability and comprehension were measured	The author also presented users with a survey that had 7 readability questions, 9 user experiences question and 1 RSD question for the last condition they read. Aesthetics and user experience were measured. Additionally, RSD (relative subjective duration).	English	Reading	Desktop	N/A	291 participants with web reading experience (including 42 who self-reported having been diagnosed with dyslexia)
Ling and van Schaik, 2007	Empirical study: Within subjects task with indepent variable of line spacing. The between subjects factor was text alignmet of left aligned and justified. Participants were presented with different readings. They had to perform a visual sarch task that required them to find a hyperlink on the screen.	The authors measured accuracy and speed of visual search as well as subjective measures relating to aesthetic appeal.	participants completed a series of questions presented by computer. The questions covered demographic details (age, sex, use of the Web), aesthetic value of pages and preference for line length and font type	English	Scanning	Desktop	Line Spacing and Text Alignment	63 undergraduate participants. All participants used the Web and had been doing so for more than a year. Frequency of using the Web varied from more than once a day to less than once a month, with a majority (76%) using the Web at least once a day. All participants had normal or corrected-to-normal visual acuity.
Liu et al., 2016	Empirical study: participants were required to search for characters in digital texts.	The search time per target character, correct response number, and correct response rate were used to measure the legibility.		Chinese	Scanning	Desktop	Font size, stroke width and character complexity	XX
Miniukovich et al., 2017	Formation of Web Usability Guidelines: Experts in usability, practitioners and researchers came together using their own knowledge to build a set of guidelines.They initially began with two researchers reviewing the entire corpus of guidelines and excluded those belonging to one of three categories. Then the experts were invited to resolve contradictions and review the validity of guidelines. Each guideline was printed on a card, they reviewed them individually and then discussed further. After coming up with 61 guidelines, they then conducted a usability test to further narrow down the guidelines. In the usability test, experts were asked to rate seven webpages on 47 guidelines.	Measured compliance rating, if applicable to a particular webpage and if understood	Automatic measurement: The compliance of each web page with the WCAG 2.0 accessibility guidelines was automatically measured.	Italian and English	Reading	Desktop	N/A	Workshops with Dyslexia experts. ; 13 of them reported having been certified with dyslexia (7 female) whereas 23 (11 female, 1 refused to specify a gender) did not. We call these two groups as dyslexics and average readers. 14 experts participated in the usability test
Moran, 2020	Eye-tracking: Large scale eye-tracking study to understand general eye movement patterns.	The authors collected eye movement data.		English	Scanning	Desktop	N/A	211 participants from Raleigh, North Carolina and Beijin, China.
Nanavati and Bias, 2005	Literature Review				Reading and Scanning	All	Line Length	N/A
Nielsen and Pernice, 2009	Eye-tracking: Large scale eyetracking study that collected 1.5 million instances of eye movement.	The authors collected eye movement data.		English	Reading and Scanning	Desktop	N/A	Conducted the study with 300 participants with a diversity of experience and demographics. The study was located in Manhattan, NY and most participants had some web readability experience.
Öquist, 2006	Empirical study: within-subject repeated-measurement, participants were given texts to read using different presentation styles on mobile phones.	Reading speed and comprehension were measured		Swedish	Reading	Mobile Devices	N/A	Participants had normal or corrected visual acuity. Each of the 5 studies had around 10-15 participants each.
Reiher-Kuijpers et al., 2021	Systematic Literature Review			Second language learning	Reading	N/A	N/A	N/A
Ricardo and Baeza-Yates, 2016	Eye-tracking: The authors measure the effect of font type and style on reading speed. Participants 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.	Eye-tracking data was collected. The authors measured reading time, fixation duration and number of fixations.	Participants were also given a prefereence questionnaire at the end of the test	No	Reading	Desktop	Font type	48 people (22 female, 26 male) with a confirmed diagnosis of dyslexia taking part in the study
Rello and Bigham, 2017	Empirical study: The independent variable was background color. There were 10 different backgrounds used. A within-subject design was used so each participant read all 10 texts on 10 different backgrounds. Then participants had to answer comprehension questions. The comprehension test was given with two literal question - questions straight from the text.	The authors measured reading time and mouse distance (the number of pixels that the mouse travelled over the text).		Spanish	Reading	Laptop or Desktop	Background color	341 participants (89 with dyslexia or at risk for dyslexia)
Rello and Marcos, 2012	Eye-tracking study: Participants had to read two stories (one in verse and the other is a fragment in prose). Participants were presented the text in different layouts (grey sale, color pairs, font size, character spacing, line spacing, paragraph spacing, and column width)	The authors measured the average fixation duration of each fragment.	Questionnaire: The participant chose what they thought was the best reading alternative between the options given for each of the parameters.	Spanish		Desktop	Color contrast, line spacing, font size,	92 native Spanish speakers between the ages of 13-43. All participants are frequent internet users and readers
Rello et al., 2016	Eye-tracking: Hybrid-measure design - Used to compare readability of different font sizes an line spacing. After reading each text, a comprehension test was given with literal and inferential questions.	Measured fixation duration, comprehension score and subjective perception rating	Participants were asked to provide their subjective perceptions. They rated their perceptions of the readability on likert scales.	English	Reading	Desktop	Font Size and Line Spacng	104 volunteers (61 female, 43 male) took part in the study. Their ages ranged from 14 to 54. All participants had normal or corrected vision
Salmerón et al., 2017	Empirical study. Eye-tracking & Think aloud: Students were given a reading comprehension test. Then asked to answer questions about specific Wikipedia articles. Then they completed a retrospective think-aloud protocol.	Comprehension and eye movement data was collected	Think-aloud protocol: After the study was over, students watched screen recording video of their learning session that included one dot representing their gaze. Students needed to remeber and say what they were learning in that video.	Spanish	Scanning and Reading	Desktop	N/A	Twenty-seven ninth and tenth grade students, native speakers of Spanish, normal or corrected visual acuity. Most had experience reading Wikipedia articles.
Shaikh and Chaparro, 2004	Large scale-survey designed to collect online reading habits and demographics. Participants listed their reading habits for 5 different document types.			English	Reading	All screen-sizes	N/A	hobbyists, the UTEST list (sponsored by Clemson University), and other professional listservs. Student participants were recruited through psychology classes for course credit. A total of 330 respondents
Shreshta et al., 2007	Eye-tracking: Participants were asked to browse a text-based page or an image-based page and search for a particular piece of information. They were given 20 seconds for each task.	Fixations recorded by ClearView were defined as a motionless gaze focused on one element lasting 100 milliseconds or longer.		English	Scanning	Desktop	N/A	Twenty undergraduate psychology students participated in the study
Singer and Alexander, 2017	Systematic Literature Review			English	Reading	Desktop	N/A	N/A
TeBlunthuis et al., 2019	Quantitative analysis: Authors collected reading time data across various language editions. They then created a stratified sample of that data and completed a regression analysis			N/A	Reading	Desktop and Mobile	N/A	Wikipedia users across language editions
Wu et al., 2020	Large-scale survey: Participants were asked to share the nature and history of their low vision, their usage of assistive technologies and then complete 3 reading activities. The participant viewed a paragraph of text from Alice in Wonderland on their chosen reading display and reported properties of viewing configuration, including the viewing distance, dimensions of the display, and the number of characters on a line.			English	Reading	N/A	N/A	158 low vision participants