Ida Marie Solås

Legibility on Mobile Phones in Bright Outdoor Conditions

Master's thesis in Interaction Design
Supervisor: Frode Volden and Phil Green
June 2020

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

NTNU
Kunnskap for en bedre verden

Ida Marie Solås

Legibility on Mobile Phones in Bright Outdoor Conditions

Master's thesis in Interaction Design
Supervisor: Frode Volden and Phil Green
June 2020

Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

NTNU
Norwegian University of
Science and Technology

Preface

This master thesis is the final part of my Master in Interaction Design, completed at the department of Design at the Norwegian University of Science and Technology (NTNU) in Gjøvik. This thesis was carried out and written in the spring of 2020 with a workload corresponding to 30 ECTS. Preliminary research and planning started in the autumn of 2019 and extended some into the spring of 2020.

The interest for this topic sprung out in a project conducted in an earlier course focusing on legibility and sufficient contrast outdoors. On further investigation of the topic I realized that the basis for sufficient contrast level to ensure universal design was based on research conducted inside on outdated displays, which don't match the use case we are faced with today. Interest for this topic stuck by me and as people's usage constantly becomes more flexible, it is important that guidelines and requirements keeps up with this development supporting the use cases and ensuring accessibility regardless. I therefore wanted to use this final project to see how designers can ensure sufficient legibility on mobile phones in bright outdoor conditions by looking at different typographical factors and how they can be used to achieve good legibility.

This thesis is mainly written for those who work with design. The aim of this thesis is to add knowledge to the field of design and advocate for including use cases in future requirements and guidelines.

NTNU in Gjøvik
16-06-2020
Ida Marie Solås

Acknowledgment

I would like to thank the following persons for their great help during this thesis:

- First, I would like to thank my supervisors Frode Volden and Phil Green for their knowledge, support and helping me stay on track during this process. Their help have been invaluable and this thesis would not have resulted as is without their guidance.
- Giovanni Pignoni for additional help and advice, especially in preparation of the experiment.
- Ole Lund for sharing his knowledge of typography and spiking my interest in typography during my bachelor, and for further helpful suggestions at the start of my master thesis.
- All the participants that gave of their time to help me with my thesis.
- And a special thanks to the rest of Institute of Design in Gjøvik for the support, opportunities and knowledge during these years.
- I also want to thank my better half, Markus Nedrelid, for his unconditional support, patience and for putting up with me during this time. I would probably not have made it without you.
- Friends and family for cheering me on all the way.

I.M.S.

Abstract

Mobile phones have a growing dynamic usage, constantly expanding as its flexibility allows for a wide use wherever you are. This have led to a greater use outdoors where we are doing daily tasks, using work tools or doing other on-the-go tasks. Sunlight can make this to a demanding task, as high illumination levels and a cold colour temperature effect what we perceive and can significantly reduce legibility. To be able to read from a phone in this context, several variables can influence the difficulty level: ambient condition, phone's display characteristics, readers vision and the characteristics of the text. Of these variables design of the text is easiest to control and through this thesis a set of guidelines and good advice will be developed for designers to support sufficient legibility on mobile phones in bright outdoor conditions. Through identifying typographical factors, reviewing existing literature and research about them and a selection of existing guidelines, the premise for guidelines are made. It became clear that contrast was most crucial in this use case, closely followed by font size. Several typographic factors were reviewed to see if legibility could be optimized further where two factors was identified to be investigated further in the given use case through an online experiment. Guidelines was designed based on the review through several methods and adding results from the online experiment, finding that neither difference in stroke contrast or difference in colour had any significant effect on legibility. The online experiment did however find that pixel density (PPI) had a significant effect on legibility as the legibility of higher PPI was rated higher.

Sammendrag

Mobiltelefoner har et voksende dynamisk bruk, som stadig utvider seg siden dens fleksibilitet tillater bredere bruk hvor enn du er. Dette har ført til større bruk utendørs hvor vi gjør dagligdagse oppgaver, bruker arbeidsverktøy eller gjør andre oppgaver på veien. Sollys kan gjøre dette til en utfordrende oppgave, ettersom høye belysningsnivåer og en kald fargetemperatur påvirker det vi oppfatter og kan betydelig redusere lesbarheten. For å kunne lese fra en telefon i denne konteksten, kan flere variabler påvirke vanskelighetsnivået: omgivelsesforholdene, telefonens skjermegenskaper, leserens syn og egenskapene til teksten. Av disse variablene er design av teksten enklest å kontrollere, og gjennom denne oppgaven vil det bli utviklet et sett med retningslinjer og gode råd for designere for å støtte tilstrekkelig lesbarhet på mobiltelefoner under lyse utendørsforhold. Gjennom å identifisere typografiske faktorer, gjennomgå eksisterende litteratur og forskning om dem og et utvalg av eksisterende retningslinjer, blir forutsetningen for retningslinjene lagt. Det ble tydelig at kontrast var mest avgjørende i dette bruksscenariet, tett fulgt av skriftstørrelse. Flere typografiske faktorer ble gjennomgått for å se om lesbarheten kunne optimaliseres ytterligere der to faktorer ble identifisert for å bli undersøkt nærmere i det gitte bruksscenariet gjennom et nettbasert eksperiment. Retningslinjer ble designet basert på gjennomgangen gjennom flere metoder og resultater fra det nettbaserte eksperimentet ble lagt til, som fant at verken forskjell i strekkontrast eller forskjell i farge hadde noen betydelig effekt på lesbarheten. Det nettbaserte eksperimentet fant imidlertid at pikseltetthet (PPI) hadde en betydelig effekt på lesbarheten ettersom lesbarheten til høyere PPI ble vurdert høyere.

Contents

List of Figures

List of Tables

1 Introduction

Reading on a mobile phone outdoors on a bright sunny summer day can be a challenging task, but people expect to be able to read and access the wanted information when desired. Mobile phones and other mobile devices have opened up another level of flexible and dynamic use, making information accessible as we carry our devices with us in our pockets. The evolution of mobile devices and digitisation of information and services takes us to use our phones in numerous new situations – buying a ticket as we run to the bus, reading the newspaper while lying on the beach, navigating a map while trying to find the right location, finding a phone number while out walking, accessing a work tool to register a control conducted outdoors, and so much more. Not all that long ago these tasks would be more cumbersome to complete outdoors relying on ourselves to remember to bring it with us – buying the ticket from a controller, carry a paper version of the newspaper with us to the beach, bringing a physical map while navigating, looking for the number in a phone book or registering our control on paper before putting it in a system back at the office. Mobile phones have made this more accessible and convenient, but a phone shows us information through the display which may compromise legibility when accessed in bright conditions outdoors, unlike similar tools on paper. High levels of ambient illumination can cause glare or reflections in the display and poorer effective contrast and brightness which all reduce legibility on phones.

A mobile phone has its physical constraints, causing most of the information to rely on visual content and text as one of the significant carriers of information (Sandnes 2017). When a designer sits in an office designing an UI or a website, it might be easy to forget or hard to know how the design will be experienced when used outdoors and when it will be easy to read or not. A set of guidelines will then be useful to follow as it might not always be time or resources to test the design with a wide range of users or in a realistic environment. When designing a text there are numerous typographical factors, including different font characteristics, width, contrast, size, weight and spacing, that can increase or decrease how easy it is to read and an understanding of the balance between these are important in order to achieve sufficient legibility.

The aim of this thesis is to design a set of guidelines that will address different typographical factors and how they can be used to achieve sufficient legibility in this use case. This will be based on gaining an understanding of which typographical factors and to what extent they effect legibility, how bright outdoor condition effects reading on displays and what typographical choices that can be made to achieve good legibility in bright outdoor conditions.

1.1 Keywords

Legibility, ambient illumination, outdoor conditions, guidelines, mobile phones, typographic factors

1.2 Problem description

As mobile phones are being used in so many different ways at different locations and what we use our phones for keeps expanding, the need to address specific use cases has increased. Bright outdoor conditions effect visibility on displays, as the sun emits higher levels of luminance it weakens the experiences brightness and contrast of the display, resulting in reduced legibility. This challenging use case needs guidelines that preserve it's ecological validity by addressing its characteristics and help designers make good design decisions that supports good legibility.

The contrast of the text is the factors most affected by this use case due to the high level of illumination and the easiest way out would be to say that maximum contrast (black and white) should be used to preserve the legibility. However, this isn't feasible as we live in an aesthetic world that values design and when designed properly improves communication. The approach of maximum contrast would not be accepted as a standard or be legislated, so a compromise is needed. A legible text is influenced by several other typographical factors and is a result of how they are combined. When making these guidelines they need to preserve some of the designer's freedom while also looking at the combination of other typographical factors and how they together achieve good legibility in bright outdoor conditions, not only relying on sufficient contrast.

1.3 Justification, motivation and benefits

Over the last years there have been an increased focus on flexible design to include all types of users, often addressed as universal design also including online accessibility (Difi 2020). Accessibility on the web is legislated in Norway and are looked at as discrimination if the legislated requirement and guidelines aren't followed. This is a great step in the right direction to ensure equality to access information amongst users, but when embracing the variety of users their flexible and dynamic usage should also be included. This thesis advocate for more actively including use cases when creating guidelines as our mobile devices leads our usage to face many different environments. Different measures are needed when reading on a phone outdoors in the summer and when reading at night with only the stars as a light source, but the way most guidelines are designed today the same guidelines are used for both of these scenarios. It's not feasible to think that one set of guidelines can cover the dynamic use of mobile devices and is why the different environments also needs to be focused on – to match not only the variety of users, but also the variety of use cases that in the end will benefit us all.

The guidelines and requirements for universal design seems to often be thought on as add-ons and as just a check list to review at the end of a project (Sandnes 2017). This is one of the challenges, another one is that it's not so easy to understand why and where they come to short. In a conversation with some designers they told me about a news article they posted, with black text on a green background and how they were surprised when they got complaints about poor contrast, when it was within the legislated contrast level. This is most likely caused by the lack of ecological validity within the guidelines and that they aren't robust enough to handle this type of use.

By designing a set of guidelines that focus on the use case of reading on a mobile phone in bright outdoor conditions it can make it easier for designers to pay attention to this challenge and design for text to be legible. Another benefit from making guidelines is that it can be implemented early in the process and when making the visual profile to ensure this across the brand and all their touchpoints. It can also be a beneficial way to ensure legibility for the projects that don't have the resource or opportunity to test their user interfaces, especially when intended for outdoor use.

1.4 Research questions

The following research questions are planned to be covered and answered in this thesis:

1. How can designers ensure sufficient legibility on mobile phones in bright outdoor conditions?
    - Which variables and to what extent do they influence legibility?
        - Do difference in stroke contrast influence legibility in bright outdoor conditions?
        - Do difference in colour influence legibility in bright outdoor conditions?
    - How does bright outdoor conditions effect legibility on displays?

1.5 Planned contributions

Through this thesis, the hope is to make a contribution into the research on legibility on mobile phones in bright outdoor conditions that further open up and facilitate a wide range of use. This includes to give insight in typographical factors and how an environment effects this, which again effects usability and accessibility of its content. Further, to use this information and see how typographical choices can achieve good legibility in challenging conditions and create a set of guidelines as a contribution to the design community. To be used by designers when designing UI's and increase usability and accessibility when using mobile phones in various conditions. Hopefully this insight can also be used by others focusing on legibility, when further developing and improving criteria and guidelines by including the variety of use and increase the ecological validity by opening up for multiple and specific use cases.

2 Background and Theory

Through previous introductory research (Solås 2020), review of literature and guidelines related to the objective of this research, several topics was identified. The scope of this thesis is influences by several variables and can be summarized as "the legibility of strokes and terminations of a typeface is guided by the limitations of the human visual system, the inherent characteristics of a display technology and the environmental conditions in which reading occurs" (Dobres, Chahine, Reimer, Gould, Mehler & Coughlin 2016). 4 key variables are identified here: typographic factors, human factors, mobile phone and the reading environment. From a designer view the only one of these that can be controlled are typographic factors and this will be the main focus. The aim of this thesis is to investigate how sufficient legibility can be achieved in bright outdoor conditions through designing guidelines. Other influential topics are also discussed, such as: how bright outdoor light and human factors can influence legibility, development of displays, how different use cases have different requirements and how ambient illumination effects legibility on displays. This chapter will to a big extent answer the research questions "*which variables and to what extent do they influence legibility?*" and "*how does bright outdoor conditions effect legibility on displays?*", contributing to a deeper understanding of the topic at hand and used in the final guidelines.

2.1 Legibility and typographic factors

Visual communication have been one of the main contributors to share and gain knowledge through times. From cave paintings and writing on papyrus rolls, to introducing mechanical movable type printing that allowed for mass production of printed books starting the printing revolution, to-day's World Wide Web where accessing knowledge and new information have become a public domain. Historically, knowledge have been a sign of power and wealth only reserved those who had learned to read and could afford it, while now all are required to access information affecting each of our lives without seeing how much power lies in digitization and accessibility. This also caused a shift in legibility research towards using ergonomics as a framework and changing paradigms towards usability (Lund 1999). Accessing information online has its physical constraints as it rely on visual information and text as a foundation to communicate (Sandnes 2017). The physical constraints of the web come in the shape of laptops, smart watches, smart phones and other mobile devices (Norman 2013) which can be used anywhere, opening a whole other level of flexibility. The total of this development has made legibility more relative, depending even more on the combination of the characteristics of the text, what device used, the surroundings, the reader and the task at hand.

Designing for legibility can result in and express many things with a balance between "artistic sensibility and pragmatic concern". (Dobres, Chahine, Reimer, Gould, Mehler & Coughlin 2016).

Different purposes demand different degrees of legibility; a sign on the highway relies on being legible or not at the required distance, while reading a novel is more a scale of how comfortable it is to read or how fast. When reading on a phone in bright outdoor conditions we depend on reading the text on buttons or to read short or long texts as news article. Even though the text may be possible to read, poor legibility can still result in visual fatigue (Lin et al. 2013) and ambient illumination conditions that causes more demanding surroundings have an effect on visual fatigue and workload (Lee et al. 2011). Nielsen (2015) summarized it simply "users won't read web content unless the text is clear, the words and sentences are simple, and the information is easy to understand".

The terms readability and legibility have at times been used interchangeably (Lund 1999), when speaking of legibility, the intrinsic and extrinsic factors that constitutes the visual properties of the text is referred. While readability refers to how the reader can understand the text and the complexity of the language used. Even though legibility and readability are separated based on different qualities, they are still dependent as legibility is critical for effective readability. Based on an earlier mapping of typographical factors in an introductory literature review (Solås 2020) more knowledge is added through a more extensive review. The different factors affecting legibility will be presented further in this section, divided in intrinsic and extrinsic factors, and finally discuss how a displays characteristic can affect legibility. Intrinsic factors refers to the shape of characters including width, weight, stroke and serifs (Figure 1), while extrinsic speaks to the psychophysical variables such as size, contrast and colour (Figure 2) (Reimer et al. 2014). Even though they are divided into two groups they are not independent of each other and can be influence in a minimal or bigger degree.

2.1.1 Intrinsic factors

When choosing what typeface to use numerous factors come to play depending on both ergonomic and design aspects. Looking at a detailed level, differences in fonts can have open or closed shapes, different angel of terminals, single or double-storey a and g, length of ascenders and descenders, these can influence our ability to identify and differentiate letters (Beier 2012). Research assessing typefaces in text-rich automotive user interface highlight some characteristics to improve legibility; open shapes, ample intercharacter spacing to prevent them to blur together, unambiguous forms and varying horizontal proportions (Reimer et al. 2014). Halbach & Fuglerud (2018) summarized related work done regarding good fonts to use on displays and found that even if a highly legible font is used, it can still result in poor legibility caused by other factors. They found that it's not possible to arrive at a generalized conclusions or universally valid rules for font design, advising to choose fonts available on the user's device. The report shows that factors like linear spacing, character width and word spacing needs more knowledge, while research indicate that size, weight and contrast/colour are more important than the font and serifs.

Serif or Sans Serif

Serifs are the "*short strokes that extends from and at an angle of the upper or lower ends of the major strokes of a letterform*" (Carter et al. 2002). Examples of serif fonts are Times New Roman

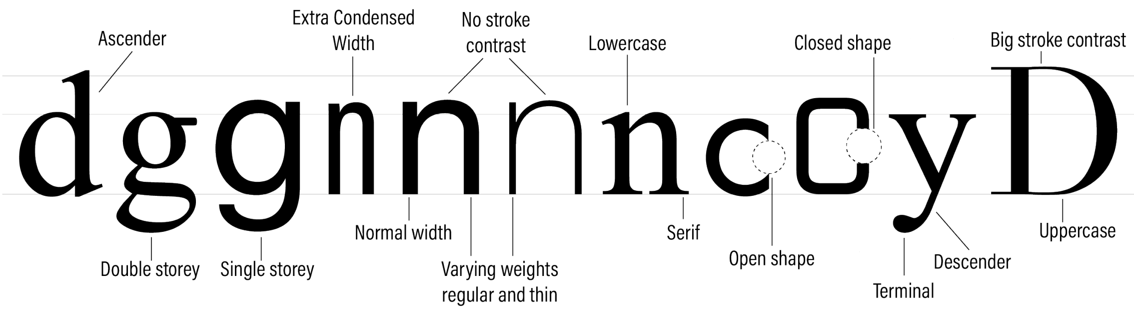


Figure 1: Illustration of intrinsic typographical factors

and Baskerville, while fonts without serifs called sans serif are Arial and Helvetica. Arguments in favor of using serifs are that they can help enhance and define the letter shape, enable to identify individual characters and keeping letters apart, and how the horizontal strokes help the reader to keep track of the line (Beier 2012). Lund (1999) did a thorough review on this subject and found that most of the research reviewed lacked internal validity, by using fonts that was very different in style in other ways than just the serifs. Bigelow (2019) summarized some newer research on serif superiority, but the reviewed studies with internal validity found little or no significant difference.

Legge & Bigelow (2011) found that in continuous text of printed newspaper and books almost all the fonts used was with serifs, while on websites nine of the ten most used fonts on Google Fonts was sans serif. This might be to achieve a cleaner look or due to the higher x-height fractions in sans serif fonts (Bigelow 2019). In the start there was a tendency leaning towards sans serif as most legible on displays, caused by the serif and other fine details causing problems and noise rather than supporting ease of reading because of a mismatch with the straight pixel pattern (Rannem 2012). Caused by the difference in resolution between paper and display fonts can look differently depending on the medium. With the technological advance supporting much higher resolution, this isn't considered such a big problem anymore as newer high-resolution displays supports a much higher level of details.

Stroke contrast

Contrast in typography refers to the contrast between the strokes in a font, the thickness of the stem and hairline (Bigelow 2019). Difference in stroke contrast are most often combined with serif or sans serif fonts, serifs usually have stroke contrast in varying degrees while sans serif usually are monolinear with optical no contrast. Typographic contrast does not seem to be a highly researched topic as a single variable (Bigelow 2019), but usually as a secluded variable when evaluating sans serif fonts as this often is a characteristic. A visual inspection of fonts on computer screens demonstrated a tendency towards low stroke contrast (Beier 2012). Dobres, Chahine, Reimer, Gould & Zhao (2016) found that bold weights are easier to read as it may increase stroke visibility, which may also speak in favor of using a font with low contrast to avoid the thinner details. Even though a high-resolution display will be able to show sufficient details, it might be that situations with high illumination levels will cause a loss in detail at such a high degree that it will reduce legibility and a font without difference in stroke is preferable.

Weight

A fonts weight speaks to the general thickness of its strokes, "*defined by the ratio between the relative width of the strokes of letterforms and their height*" Carter et al. (2002). Font families vary in number of weights, often with 4 sufficient weights – light, regular, medium and bold. The World Eide Web Consortium (W3C) defines weights from thin to thick characters with values on a scale between 0 and 1000, where 400 is equal to normal and 700 to bold. These numbers and names can vary between families and aren't specified in ratio or by coverage area (Bigelow 2019). A general guideline of weight is that a too thin or too heavy version of the font would both cause degrading in legibility, leaving a medium weight as the most legible (Carter et al. 2002).

Dobres, Chahine, Reimer, Gould & Zhao (2016) found when examining stroke weight on Chinese characters that bold weight was easier to read than medium in a glance-like context. Lightweight typography has become popular in recent years, according to Dobres, Reimer & Chahine (2016) rendering of the text can be the deciding factor. Earlier studies found superior legibility with medium weights compared to lightweight and bold, however significant findings between rendering system and font weight have been found to influence this superiority. With suboptimal rendering, light weight text degraded much more compared to heavier weights, while under the best rendering conditions lighter weight fonts had superior thresholds compared to heavier weights (Dobres, Reimer & Chahine 2016).

Width

Most commonly used are variable-width fonts which allows for each letter to have different widths, while fixed-width or monospaced have the same width independent of the letter (Bigelow 2019). Variable-width fonts leaves equal white space between the letters which can increase the reading as the gap don't disturb the reading. Character width vary from narrow (condensed) to normal to wide (expanded), a condensed font might be chosen to accommodate for little space (Carter et al. 2002). By using a condensed or expanded form the letters changes, causing extreme posture that can affect reading patterns. Condensed and expanded fonts with narrower or wider widths are also spaced differently to accommodate (Bigelow 2019). Beier (2012) research review suggest that wider forms are preferable to narrower forms as it makes room for internal space of the characters, but not too wide as this can lead to other misreading's.

Lower or uppercase

We are mostly used to read continuous text with lowercase letters and the belief is that they are more legible than uppercase. But according to the discussion by Beier (2012) research finding uppercase letters to perform badly, is because we aren't used to it and that's why lowercase is superior in continuous reading. In shorter readings like on signs, the choice is more dependent on available space and uppercase could be beneficial in some situation, even though lowercase seems best overall.

2.1.2 Extrinsic factors

Size

Size is one the most decisive factor affecting legibility – if the size is too small the reader might not be able to make out the letters, or do so comfortably, and sizes that are too big will make it hard to see the context and possibly to make out the word, especially in continuous reading. Size is defined in different units depending on the medium and purpose. Designers are used to relate to units like points for print and pixels or em when designing for the web. Even though different fonts are specified at the same point size, it doesn't mean that they are equal in size as they might vary in x-height and cap height (shown in Figure 2 . X-height is the height of the letter x and is a better measurement of the fonts size when reading lowercase letters. A font with a low x-height will in general call for bigger font size to achieve the same legibility as a smaller sized font with high x-height. A large x-height suggests enhanced letter legibility and support better performance, especially at smaller size (Beier 2012). Not accounting for different x-height when comparing fonts is one of the reasons Lund (1999) found a lack of internal validity in previous legibility studies. Legge & Bigelow (2011) review on print sizes offers several reasons why x-height should be used as a measurement in vision research, as this is the comparable variable. While Dobres et al. (2018) set the text size based on the height of the capital 'H' as defined in ISO 16673:2007 for road vehicles. Traditionally many adjusted the proportions of letters depending on its scale, called optical scaling (Bigelow 2019). Fonts designed for small scale would have greater x-height fraction, wider letters, more space between them, less contrast and thicker strokes, while fonts for larger scale would look smaller with smaller x-height fraction, more tightened spaced and narrower.

How print size effect legibility is a highly researched topic, especially in the search for a minimum size recommendation as it effect the reading experience, but is also important economically (Bigelow 2019). With a smaller size, more characters can fit on a fixed page and less ink and paper is needed to print the same information. While space on the web is "cheaper" without these concerns and more flexible, which makes it easier to adjust for individual differences and focus more on the ergonomic aspect. In the search to find a balance between the economic and ergonomic, especially related to visual fatigue (Bigelow 2019), larger text has been found to be more legible than smaller text (Dobres et al. 2018). When reviewing the extensive data Nersveen & Johansen (2016) collected when researching legibility in printed text for people with impaired vision, 12 pt was found sufficient to achieve a acceptance rate of 80%.

Both points (pt) and pixels (px) defines size and refers to the height of the available space of the font, while em is a scalable unit that is used on the web an equal to the specified size. Standard font size used is often 12 pt on print or equally 16 px or 1 em on the web for body text. Carter et al. (2002) writes that at normal reading distance 9 to 12 point is the most legible size, this range is caused by variation of x-height, but is also in a triad relationship with line length and interlinear spacing as well. Legge & Bigelow (2011) found the range to read at maximum speed looking at performance and significance of print size. Assuming the standard reading distance to be 40 cm, the fluent range is approximately 0.2° to 2°, giving a physical x-height of 1.4 mm (4 points) and 14 mm (40 points). Testing on screen, text size of 4 mm and 3 mm using capital H to set text size, Dobres et al. (2018) found the largest size to be most legible at a glance, conforming the findings in Dobres, Chahine, Reimer, Gould, Mehler & Coughlin (2016) similar study. In Halbach & Fuglerud (2018) report the literature agree on a size of at least 16 pt on print for those with reduced vision, even if reading on a display (as long as it has high quality) there shouldn't be much difference on print, indicating that approximately 21 px may be a good size. The World Wide Web Consortium don't specify a minimum font size in their Web Content Accessibility Guidelines (WCAG), but require it to be scalable up to 200% without losing functionality or content (W3C 2018*b*). In the success criteria addressing contrast they do however refer to text considering text below 14 pt bold or 18 pt normal as normal text, and text above 14 pt bold or 18 pt as large text.

The displays resolution controls how the text will be presented as its form is made out of pixels, a large size will have more pixels and small sizes fewer (Beier 2012). Technological advances have increased the resolution and pixel density of displays rapidly, resulting in displays as Apple's Retina displays giving a pixel density so high that the eye does not see the individual pixel at a normal viewing distance (Apple Inc 2020). This isn't applicable for all displays, which suggest that the resolution of the expected end users display should determine font size and font characteristics.

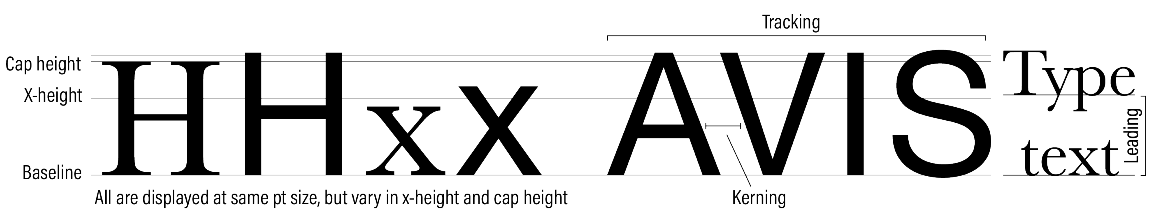


Figure 2: Illustration of extrinsic typographical factors (size and spacing)

Spacing

Spacing in typography refers to tracking, leading and kerning. Tracking or letter spacing is the general space between letters, leading is the distance between the lines and kerning is adjusting space between letters and specific combination of letters to achieve a proportional font (Rannem 2012). In typography increasing space too much is referred to as loose, while too little is tight spacing. Too loose or too tight tracking or leading will disrupt the reader and will cause difficulties in identifying the letters, reading the words or locating the next line. Adjusting tracking is usually not necessary as the font's design follow certain spacing rules, but some studies have found benefits with increased letter spacing for dyslexic readers (Bigelow 2019). Typefaces that are named "expanded" or "condensed" will have looser or tighter letter spacing and wider or narrower letters. Rannem (2012) writes how leading should be adjusted optically and optimizing legibility depends on the font's characteristics. The built in leading is adjusted to prohibit descenders and ascenders from crashing between lines, but fonts with high x-height and sans serif fonts may need some extra leading to make the lines clearer. Dobres et al. (2018) looked at both size and leading as variables and found that increased leading (tested with 0% and 33% of text size) significantly enhanced legibility, but didn't compensate for reducing text size. WCAG have detailed requirements regarding text spacing, specifying line height to at least 1.5 times the font size, tracking to at least 0.12 times the font size and word spacing to at least 0.16 times the font size. Even though W3C refers to this through research, Halbach & Fuglerud (2018) points out several weaknesses in this claim by only relying on one study with a somewhat narrow range of participants that is not verified or available to the public.

Line length

Line length, also often referred to as character per line as a more accurate measurement, gives poorer legibility by either being too long or to short (Dyson & Haselgrove 2001), by either providing to little information per fixation or when the reader are having difficulties finding the next line (Dyson 2004). Text can be designed to fill the whole page or in columns, with a fixed width on print or a recommended responsive flexible width on the web. Dyson & Haselgrove (2001) found that a medium line length of 55 characters per line support effective reading, this can be useful when designing for fixed width. When reading on a mobile phone the width of the website will vary dependent on the phone's width, and as there is no reason to shorten the length on such a small device the line length should be kept close to full width of the display.

Polarity

Positive (dark text on bright background) and negative (bright text on dark background) polarity have documented effect on legibility. Digital displays have made it easier to display text with negative polarity and are popular and often used to create day and night modes or use in night-time applications because they emit less light and interference into its surroundings (Dobres, Chahine, Reimer, Gould, Mehler & Coughlin 2016). Even so, Buchner & Baumgartner (2007) measured the effects of polarity with proofreading performance, which found positive polarity to be consistently better independent of ambient lighting and chromaticity. Positive polarity was easier to read in glance-like context on Chinese characters (Dobres, Chahine, Reimer, Gould & Zhao 2016), and Dobres, Chahine & Reimer (2017) also found that positive polarity on displays had an advantage under both dark and bright illuminated conditions under glance-like reading.

The gap that seems to appear where dark polarity sometimes is preferred and the research result showing positive polarity are more legible regardless of ambient illumination, can be explained by how "the dilation of pupils under low illumination produces optical blurring" (Dobres, Chahine, Reimer, Gould, Mehler & Coughlin 2016).

Colour and contrast

In design, colour is often used to create or accentuate an effect, mood or feeling, but "it should also be chosen with typographic legibility in mind" (Carter 2002, p. 7). The colour wheel is often used to make effective colour combinations by using different combination to make various contrasts, such as monochromatic, complementary, split complementary, analogous, neutral or incongruous. In the process of selecting colours a balance is needed between the hue of the colour (e.g. blue), the value (how bright it is) and how intense the colour is. As soon as colour is added, legibility is compromised. To optimize legibility black text on a white background will always be the safest option, but we live in a world that appreciate design and value aesthetics (Postrel 2003). Rannem (2012) writes how typography and colours is used to create associations, evoke emotions and as a design tool to create contrast and communicate. And so, a compromise is needed in order to use colour, but also make sure it's easy to read, if reading is the goal.

Humar et al. (2014) investigated how legibility was affected by colour combinations on LCD displays, by using a set of 8 basic colours different colour combinations was made and tested with different polarity. Colour combination had significant effect on the legibility and differences was found comparing these findings with research on paper and CRT displays. Subjective rating and visual performance both found black on white, black on yellow, blue on white and blue on yellow to be most legible on LCD displays. Researching the effects of screen luminance and text colour Lin (2005) found increased visual performance along with increased contrast ratio. Text colour did not significantly affect visual performance, but chromatic text was preferred over achromatic text.

Human factor's and individual differences such as colour vision influence perceived contrast and makes it hard to rely in a difference in hue. As hues are perceived differently, the contrast requirement defined in WCAG relies on a sufficient contrast in relative luminance that is independent of colour perception (W3C 2018*a*). The minimum contrast required is 4.5:1 for normal text and 3:1 for large text, indicating that as size decreases more contrast is needed. There are however some weaknesses in the contrast levels, the ratio is based on a relative luminance and don't take the effects of a dynamic ambient illumination into consideration, and the standards the minimum contrast levels are based on are standards for computer workstations in office environment. How ambient illumination and high illumination levels effect legibility on displays is further discussed in section 2.5. Pignoni (2018*a*) researched how different levels of illumination effected character recognition as a device is moved from indoors to outdoors. Based on his findings 18:1 was suggested as the minimum contrast ratio in high illumination. A tool was also made to simulate how the effective contrast will be with different contrast ratio, phones and ambient illumination.

2.1.3 Reading on displays

The factors that have been reviewed so far are relevant to text on both print and digital, but the reading experience may be different on these two mediums. Printed design is fixed and made by adding ink needing additional light to see the content, while a display consists of pixels that show content by emitting light on its own. Resolution for print have traditionally been much higher (300 DPI) than the web requires (72 PPI), but many displays now have more pixels and higher pixel density that allows for sharper and clearer rendering of web content requiring higher resolution. Resolution on displays is the number of pixels displayed on a device (width x height) while PPI is the density of pixels referring to how many that are present per inch. When lower PPI was common, anti-aliasing was often used to smooth of the edges of text and other graphic to make them appear less pixelated, but increased pixel density will in itself smooth the edges and minimize the anti-aliasing effect. The resolution, size and PPI of displays vary, and high resolution displays also have a scale factor to consider with more pixels in the same physical space that give a sharper result. Characteristics of a display is based on the expected viewing distance; a mobile phone will have higher PPI than a computer monitor as it will have a shorter viewing distance to display content sharp enough. A text set in 16 px will most likely appear as similar size from the two different viewing distances, but the text will have a different physical size on the displays due to a higher PPI on mobile phones. Compared to a printed text with a fixed output it is hard to design for the web due to the lack of control of output and viewing conditions. Even so, a designer needs to consider these display differences to design a text robust enough to be legible on a wide range of displays.

2.2 Legibility guidelines

Knowledge about how to ensure legibility comes in different forms and can be shared through good advice, best practices, design manuals or guidelines. One example is the main guidelines to ensure legibility by Nielsen (2015); "*use a reasonably large default font size and allow for the size to be changed, a clean typeface without strange shapes, use high contrast and plain background behind the text*". This is fairly general and straightforward guidelines, while others are more concrete with a clear focus or deeper explanation. A selection of different guidelines will be introduced and reviewed in this section.

2.2.1 Web Content Accessibility Guidelines

Web Content Accessibility Guidelines (WCAG) is a set of guidelines to ensure accessibility on web, developed by the international community World Wide Web Consortium (W3C) that develop web standards. The recommendations in WCAG was extended with version 2.1 in 2018 (W3C 2018*b*) and consists of success criteria that addresses how web content should be built and designed to ensure accessibility to people with disabilities, including visual, auditory, physical and cognitive, among others. In Norway, as well as in other countries, universal design is a legal requirement (*Equality and Anti-Discrimination Act* 2020) to ensure that everyone in the society can participate. The European Commission have also decided to implement standards with Web Accessibility Directive (WAD) that is in line with WCAG 2.1 and all member states need to transpose this in their national laws (European Comission 2019).

Several success criterion's addresses factors that enhances legibility (W3C 2019). SC 1.4.4 addresses resizing text which should be scalable to 200% without loss of content or functionality. Added in WCAG 2.1 is also SC 1.4.12 that includes text spacing in terms of line height/spacing at 1.5x font size, 2x spacing after paragraphs, 0.12x letter spacing and 0.16x word spacing. SC 1.4.8 speaks to visual presentation of blocks of text. Some other criteria call for text alternatives for non-text content (SC 1.1), images of text to just be used as decorations and not to convey information (SC 1.4.9), so that the same information can be accessed with reading aids which aren't possibly if displayed as images.

Success Criteria 1.4.3, 1.4.6 and 1.4.11 all address contrast in WCAG to make it accessible for people with vision deficiencies. SC 1.4.3 Contrast (Minimum) speaks to the visual representation of text, SC 1.4.6 Contrast (Enhanced) is based on the same as the minimum but with stricter requirements, while SC 1.4.11 speaks to the non-textual contrast and was added in WCAG 2.1. I will focus on SC 1.4.3 as this is the criteria legislated for ICT-systems in Norway and is what's used by designers and programmers. The minimum contrast is divided in three levels, starting with level A that mainly reflects that contrast should not be based on hue (W3C 2018*a*). Level AA and AAA differ between regular and large text, where large text is defined as text 18 pt. (24 px) or 14 pt. (18.5 px) in bold. AA defines a contrast ratio of 4.5:1 and 3:1 for large text, which is sufficient for a visual acuity of 20/40, while AAA compensates for a visual acuity of 20/80 with 7:1 and 5:1 as contrast ratio (W3C 2018*a*).

As mentioned earlier in section 2.1 there are however some weaknesses in these guidelines. Halbach & Fuglerud (2018) points at how the requirements for spacing only is based on one study that aren't verified or available to the public. Another weakness is regarding the contrast ratio. WCAG refers to a contrast in relative luminance, which refers to a combination of hue, brightness and saturation, even though the human eye is most sensitive to differences in brightness, especially important for those with colour blindness (Sandnes 2017). Relative luminance is used to reflect that web content don't emit light itself (W3C 2018*a*). The formula for contrast looks like this, where L1 is the lighter colour and L2 is the darker:

$$contrast = \frac{L1 + 0.05}{L2 + 0.05} \tag{2.1}$$

The level of contrast is based on standards for indoor workstations and the 0.05 value is included to account for ambient light, but the weakness of this is the static calculation (Sandnes 2017) when the usage is much more dynamic and likely use cases as discussed in section 2.4 would face a much higher level of illumination, which reduce the effective contrast significantly (Pignoni 2018*a*, Chen et al. 2017).

2.2.2 ISO Standards

International Organization for Standardization develop international standards covering aspects of technology and manufacturing. The different guidelines deal with specific issues like ISO 9241-304:2008 "provides guidance for assessing the visual ergonomics of display technologies with user performance test methods" to ensure that a display meets minimum requirements in the given context (International Organization Standards 2008). ISO 9241-303:2011 also deals with requirements for electronic visual displays with "generic performance specifications and recommendations that will ensure effective and comfortable viewing conditions" (International Organization Standards 2011). Breuninger (2019) illustrates how this standard can be used, with a reading distance of 30 cm on a smartphone or tablet the recommended text size is a cap height of 1.7 mm. Another example is ISO 24509:2019 that estimates the minimum legible font size for single characters for people at any age with correct vision, specifies to printed materials with fixed font size (International Organization Standards 2019). To get full access to ISO standards payment was needed, so full access and review have not been done. They are however still included as they are good examples of standards made with a specific use case in mind.

2.2.3 UX Collective

Breuninger (2019) present a few simple rules to make a text convenient to read. The main focus area was kept on three properties of the text; if the size is big enough, contrast high enough and if the typeface is feasible for its use. Font size is dependent on the cap height and viewing distance, and the recommended calculation for font size is 0,00582 x *viewing distance* for headlines and 0,00465 x *viewing distance* for text body. Or to follow ISO standard 9241-303:2011 that recommend a cap height of 1,7 mm for reading distance at 30 cm or 2.3 mm at a distance of 40 cm, both related to reading on a mobile phone.

Adequate contrast levels are referred to the contrast levels defined by W3C in WCAG, 4,5:1 for small text and 3:1 for large text, and positive polarity is preferred unless it is expected to only be used in the dark. Colour combinations of red and green, and blue and yellow should be avoided to pay attention to colour blindness, as well as read and blue/purple as this combination be hard to read due to the distance in the visible spectrum. When choosing what font to use serif is fine if the screen has sufficient resolution, light/condensed fonts should only be used when size and contrast are more than adequate, bolder weight can compensate somewhat for size and contrast and familiar fonts are usually more successful. Keep in mind that most users won't enlarge text, readers are diverse, and half probably don't have as good visual acuity as you, so spending some time on deciding text size and contrast will pay out as a better reading experience.

2.3 Colour and light

In order to experience colour, a light source, an object and a viewer is needed; as the light reflect on an object and reach the eye the energy are transmitted further to the brain and seeing colour is the result (Feisner & Reed 2014). Physiological speaking, colour is light which travel at wavelengths in the range of 400-700 nm constituting the visible spectrum. Different range of wavelengths project different colours; violet, blue-violet, blue, green, yellow, orange and red (Feisner & Reed 2014). We see colour when the light strikes a surface and wavelengths are absorbed, while the rays of the surface colour reflects to the eye. Mixing pigments as in paintings is subtractive colour mixing, while adding and mixing colored light like in displays is additive colour. Additive colour mixing and displays work with red, green and blue as primaries, but other colour models as HSL or other variations like HSB or HSV is also often used. HSL refers to hue (the colour), saturation, often also referred to as intensity or chroma (the richness or fullness of colour), and lightness (how dark or light the colour is). Warm or cold colours, or colour temperature have shown to effect people in various ways. Warm colours such as red, yellow and orange, and the cold colours blue, green and violet gives us associations to other things in nature as the sun or water (Feisner & Reed 2014). Depending on the total mix of the colour (other hues, saturation and lightness), a hue can changes its temperature.

2.3.1 Ambient conditions

Since colour is a result of light different types of illumination will influence what we see as light sources vary in colour and in strength. Feisner & Reed (2014) explains how colour temperature of different light source will influence how we see colours. Incandescent lights as candles and gas lamps emit white wavelengths that make warm surfaces hues appear brighter and cold one duller, while LEDs give a cool blue light. Sharma (2004) explains further about light sources and different colour temperatures. Artificial light sources are for example candles, tungsten lamps and fluorescent lighting, while natural light include sunlight and clear or cloudy sky. Tungsten sources emits most energy in the red part of the specter, daylight is considered neutral light with a pretty balanced emission and fluorescent light peaks in the blue prat of the spectrum. To measure colour temperature Kelvin is the unit used. When speaking of cold and warm colour temperatures this is opposite of how we speak of the colour spectrum, so the colour temperature starts with low numbers and a warm temperature and as the number increases the colour temperature turns colder. Some illuminants defined by CIE are tungsten lamp at 2856 K; different types of daylight (D50, D55, D65 and D75) at 5000–7500 K and uncalibrated monitor at 9300 K. Our visual system has a white balance mechanism called chromatic adaption, allowing the eye to compensate for colour of the light source and adapt to perceive colours "correctly".

Heiting (2017) writes about how sunlight is a combination of rays of different colours that together creates "white light" and how the rays energy varies. Different wavelengths contain different amounts of energy, the short wavelengths contain more energy while rays of long wavelengths contain less. The shortest wavelengths with highest energy, generally defined between 380 to 500 nm, constitutes HEV (high-energy visible) or blue light. We are exposed to blue light by being outdoors as sunlight is the main source, but also from other displays we use as "computer screens, smartphones and other digital devices emit significant amount of blue light" (Heiting 2017). These short wavelengths scatter more easily which makes it harder to focus on and as the eye aren't as good at blocking blue light from reaching the retina, it could appear more dazzling than longer wavelengths and cause eye strain (Heiting 2017). This is confirmed by Wolffsohn et al. (2000) finding that yellow-colored lenses that cut out blue light, less than 450 nm, increased contrast significantly and correlated with the subjective ratings.

The illumination intensity effects our environment ranging from total darkness at 0 lux, normal office environments around 500 lux, 10.000 lux on a clear day and levels up to 100.000 lux in sunlight. The various levels of intensity produce different responses in human vision as a good enough light is needed to read indoors, the levels outdoors can be dazzling, all depending on the medium. Paper reflect the ambient light as it hits the surface and from not being able to read paper in the dark it gets more legible as the illumination increases; this also applies to reflective displays that don't emit light as e-readers. On the other hand, displays that emits light can be read without ambient illumination but as the ambient illumination increases it weakens the contrast level on the display and can cause disturbing reflections and glare.

Outdoor conditions change throughout the day in both chromatic and illumination intensity. The time of day, the weather, season and location in the world (Feisner & Reed 2014). Colour temperature changes during the day starting with cool colours and in the morning and ends with warm temperature as the sun sets. The intensity of illumination will follow the sun, but also be affected by clouds causing the same effect as a diffusion box with even distribution of the light. Weather elements like snow will cause greater reflectance possibly causing dazzling. Ambient conditions can vary greatly from inside with candles in the evening to going skiing in the Easter holiday. The focus in this project is how the ambient conditions faced outdoors in bright sunlight effect legibility and are limited to these demanding viewing conditions.

2.3.2 Human factors

As light reflects on a surface and passes into the eye it comes in contact with the retina which is made up off rods, cones and other cells (Feisner & Reed 2014). Rods are more sensitive than cones and is responsible for most of our ability to see in the dark. But rods don't influence colour vision much and is why we experience less colour at lower light levels. Cones recognize red, blue-violet and green which are long, short and middle wavelengths that allows us to perceive hues (Feisner & Reed 2014). The wavelengths are recognized by the cones and pass these signals to the fovea, which further transmit them to the brain that interpret the signals into one message that tells us what colour we are seeing.

Part of the physiological process of seeing colour is based on our individual ability to process colour and leaves it vulnerable. According to WHO (2019) "at least 2,2 billion people have a vision impairment or blindness" across the world. These conditions effect the vision in various degrees and in different ways, but are usually caused by uncorrected refractive errors, cataract, age-related macular degeneration, glaucoma, diabetic retinopathy, corneal opacity or trachoma (WHO 2019). For instance, one of the symptoms of cataract is increased glare and reduced acuity caused by an unwanted light scattering in the eye lens, which normally in cases with increased levels of luminance or light with cold colour temperature will cause difficulties or discomfort (Nersveen & Blindeforbund 2009). Colour blindness makes it hard to differentiate some colour and is also globally spread, as many 1 in 12 men are colour blind (NIH 2019*a*). Seeing the difference between red and green is the most common form of colour blindness and have 4 types; deuteranomaly, protanomaly, protanopia and deuteranopia (NIH 2019*b*). Blue-yellow color blindness are less common and have two types, tritanomaly and tritanopia, and complete color blindness or monochromacy is uncommon (NIH 2019*b*).

All of these are known examples of individual differences that causes people to see and experience colour differently, some are related to each other while others aren't. Visual impairment is one of the main user groups when talking about universal design, as it effects so many it is important to facilitate and include them. This is important in conjunction with legibility, especially on the web as most communication happen through text.

2.4 Use cases and conditions of use

Affordances and constraints are two know design principles (Norman 2013) that influence how people use a design or product. Older displays were big, heavy and stationary, constrained to stay in the same location and affording it to let it stay there. This condition of use has changed, resulting in mobile phones affording to be taken and used everywhere and physical constraints allowing it to. Mobile phones open up for a big number of use cases by carrying it everywhere, having access to virtually anything. These considerations needs to be accounted for in guidelines, as Beier writes "*...reveals that typeface legibility is not a universal issue, where one feature or set of features improves legibility in all reading conditions. In other words, the level of legibility for a given typeface is not constant but varies, depending on the situation in which it is observed*" (Beier 2012, p. 11).

In a conversation with some designers working at one of Norway's news media last fall, we got to talk about contrast, the legislated requirements for universal design and how robust they are. They told about an article they recently had posted, with black text on a green background. The contrast was within the legislated requirements, but they had still gotten complaints about poor contrast, which seemed to have left them somewhat puzzled. This is an example to illustrate a challenge in how many uses these requirements as a checklist (Sandnes 2017) without giving it any further thought. Even though the legislated contrast ratio is a minimum and not best practice, it seems to be a perception that as long as you are within the limits you are good. This could be the case if these guidelines and requirements had specified use cases or designed for a more dynamic use, but most of them are not.

The following case is not directly tied to legibility, but still shows an applicable use case outdoors. A family was out on Omaha Beach in Normandy to test an AR-app that showed the events of the D-Day in 1944 (Nyre & Liestøl 2018). Naturally this happens on the beach and by using their phones they were told the story of events with Artificial Reality shifting between being on the beach in the present and seeing the events from 76 years ago. The scenery for this app is out on the beach and users will probably use it during the summer, which may pose a problem with high ambient illumination and the mobile phones. The family testing the app needed to use their t-shirts over their heads to shut the sun out in order to use the app and is a good example of the effect outdoor conditions can have when using a mobile phone.

This is an example of how different use cases and changing ambient illumination have been considered when creating a design system. The OpenBridge Design System is "developing an open platform that provides better and safer user interfaces on ships..." (OpenBridge Design System 2020). The captain of a ship and his crew is out sailing for longer periods at the time with no visible land in sight. The design system contains several aspects, but also interface design which resulted in the creation of 4 different modes to suit different times of the day. The colours and contrasts that is needed when sailing in bright daylight is at a whole other level than at night. In order to maintain the night vision when the only available light is the stars, the contrast will need to be much lower than the general levels defined in WCAG.

The mentioned examples are all real-life scenarios that reflect how users can use a user interface in different situations and some use cases will be more extreme than others. When looked at from a user-centered perspective and following a user-centered design process it is natural to include the context of use when gathering insight about the users and needs to further include this insight in methods like storyboards and scenarios to keep this in mind during the design process (Benyon 2013, Courage et al. 2015). This have become more and more common as interaction design and user-centered design thinking have emerged, but it is still often not used or maybe forgotten, either because of lacking knowledge, time or money, this is not always prioritized. The result may vary and in a exploratory stage influential factors can also be missed, but be found when user testing the interface in the environment it will be used. Usability testing would hopefully unveil potential weaknesses when tested in these environments. Guidelines could be helpful in this situation by providing an understanding for conditions with bright sunlight and how to design for this at an early stage.

2.4.1 Development of displays

Evolution in conditions of use have led to an increase in use cases, developing along with the display technology. The displays as we know them today started in the 1800's with the development of the cathode ray tube (CRT), first introduced as electronic television systems and later as computer monitors e.g. A CRT is a vacuum tube that produce images when an electron beam strikes a phosphorescent surface (Bellis 2017), by using multiple beams of electrons CRT's was able to display colours. This technology was later replaced by Liquid Crystal Display (LCD) which is commonly used in digital clocks, portable computers and appliance display. LCD displays consist of liquid crystals that align when exposed to electrical fields (Bellis 2019*a*), this flat panel display does not emit light itself and so a backlight is needed and sent through multiple layers of filters to produce images. Different variation of LCD displays used in phones are TFT LCD (thin-film-transistor liquid-crystal display) attaching each pixel to a transistor and capacitor individually improving image quality and contrast, and IPS LCD (in-plane switching liquid-crystal display) improving colour reproduction, viewing angles and energy consumption. The newer OLED "organic light-emitting diode" display allow for even thinner, brighter and crisper display than LED (Bellis 2019*b*). As an organic material, light is emitted when a current pass through an allows for individual pixels to be turned completely off or on, in opposite to the older LCD panels with backlight (McCourt 2020). This allows for a darker black and increase the contrast of the display by being brighter and having a darker black point. Variants of OLED is AMOLED which adds Active matrix like TFT and Super AMOLED integrates the touch response layer into the display.

These advances in display technology is also seen through the development in our needs and common use cases. The CRT was e.g. used to introduce television that opened up a whole other way of speaking to the public and getting entertainment and was used in computer monitors that opened up a door to digitization and automation of tasks. From there LCD technology offered a more adaptable, flexible and cheaper use, which lead us to a more flexible use that requiring even better and brighter displays that was introduced with OLED technology.

Advances in technology have resulted in brighter displays, deeper level of black and less reflection resulting in better contrast, which makes them more robust to withstand bright sunny days. Higher resolution displays allow for new ways to present content on displays like having smaller text size (Dobres, Chahine, Reimer, Gould, Mehler & Coughlin 2016). More recent functions like adaptive brightness adjustment and Apple's true tone function adjust the color temperature of the screen dependent on the ambient illumination around us. All of these advances help increasing legibility on mobile phones in bright conditions, but far from all have the newest phone and it is still only one of the variables influencing the total reading experience.

2.5 Ambient illumination effect legibility on displays

Colour temperature and intensity in ambient illumination conditions have found to significantly effect human psychophysical responses and satisfaction (Lin 2005, Lee et al. 2011, Lin & Huang 2013, Choi & Suk 2014). Colour temperature effects how we perceive colour and is known to change our perception and colour combinations effect legibility (Humar et al. 2014). Lin & Huang (2013) found that white light, normal ambient illumination and background with primal colours was the best conditions for character recognition. Choi & Suk (2014) researched user preferences of colour temperature for smartphone display under varying illuminates. Mobile phones are exposed to highly dynamic environments and chromatic adaptation helps the eye adapt to colours in various ambient illumination, but this function might not apply when the phone itself also emits light and often seem to yellow or blue. Choi & Suk (2014) findings suggest that the optimal colour temperature of the display is related to the illuminate colour temperature that enhances along with illuminate intensity, user-preferred temperatures shifted on an average towards higher colour temperatures.

Ambient illumination has showed to have an effect on visual performance, visual fatigue and effect the reading time (Pignoni 2018*a*, Lee et al. 2011, Chang et al. 2013, Lee et al. 2008). In most of the studies identified where ambient illumination is considered as a variable it is often lower levels illumination with a maximum of 1500 lux (Lee et al. 2011, Chang et al. 2013, Lee et al. 2008) which is a more accurate indoor level of illumination, and often investigating the ideal illumination level on the tested display. Pignoni (2018*a*) found that the uncertainty of text recognition grows along with the higher ambient illumination levels as the ambient contrast ratio gets affected by its surroundings, the experiment was conducted at different illumination levels up to 5500 lux. This researcher has previously also conducted an experiment (Solås 2019) on the perceived legibility in outdoor conditions, where around 9000 lux was measured as the highest illumination level. Due to weaknesses in the reliability the results can not be generalized but results still shows an indication of how the perceived legibility falls as the contrast ratio is lowered in bright conditions.

Different displays technology is a variable that can play an important role, better performance has been found on TFT-LCD screen than with CRT and TFT-LCD screens had better performance at 450 lux than 200 lux (Shieh & Lin 2000). Screen types have different characteristics and findings in performance tested on CRT screens may not apply to TFT-LCD which are frequently used on phones and computers (Lin & Huang 2006). Contrast ratio have been shown to have a significant effect on both visual recognition and subjective preference, significant findings were not found regarding preference on TFT-LCD and ambient illumination (200–700 lux), “but research on that topic is limited” (Chen & Lin 2004). Even with a lack in research on this topic, especially in higher illumination, “contrast ratio is the most important sub-factor of color combination that affects visual performance significantly” (Lin & Huang 2006).

3 Methodology

The aim of this thesis was to see how designers can ensure sufficient legibility on mobile phones in bright outdoor conditions by creating a useful set of guidelines. The research and guidelines reviewed in chapter 2 was assembled and employed into designing guidelines. Based on the background information, two typographic factors were identified to be investigated further, difference in stroke contrast as it has not been a highly researched topic and difference in colour due to how the bright sunlight can effect it. The chosen methodology was used to answer the following research questions:

1. How can designers ensure sufficient legibility on mobile phones in bright outdoor conditions?
    - Do difference in stroke contrast influence legibility in bright outdoor conditions?
    - Do difference in colour influence legibility in bright outdoor conditions?

Several methods will be used to answer the first research question in making the guidelines, including qualitative user-centered and gamestorming methods to define the use case based on context, pain-gain map, forced ranking and an affinity diagram to visualize the data and re-visit the reviewed research. To investigate if difference in stroke contrast or colour had any effect on legibility an online experiment was distributed with the aim of collecting the data when performed outdoors in bright sunlight. The outlines of the survey will further consist of description of the survey setup, design of stimuli, hypotheses to be tested and the data collected.

3.1 Guidelines

3.1.1 Purpose

The purpose of creating guidelines was to make a tool to assist designers when designing for the specified context to achieve sufficient legibility. The guidelines were designed by applying several methods helping to shape each guideline, specifying the challenging use case and prioritizing the typographical factors based on their effect.

3.1.2 Creating guidelines

Use case with pains and gains

Use case, also called scenario, is a frequently used method in user-centered design and a useful tool to include tasks and situations in the process (Courage et al. 2015). Usually a scenario is based on a persona and together they provide good insight of the user, task, goal and desire for task, functionality that is needed, etc. Benyon (2013) describe scenarios as "stories about people undertaking activities in contexts using technologies". The aim of defining a use case to include in the guideline, is to describe a framework of the contexts of reading on a phone outdoors and what challenges can be faced there. By emphasizing this the objective of the guidelines are clarified and the ecological validity are strengthened.

As a measure to convey how the bright sun can have a negative effect on legibility, how it effects different typographical factors and how it can be managed a pain-gain map was used. The purpose of this gamestorming method was to "develop an understanding of motivations and decisions" (Gray et al. 2010, p. 190) and set this in relation to the use case. With the objective that conveying this knowledge will increase the understanding and motivation to follow the guidelines.

Forced ranking

The gamestorming method forced ranking (Gray et al. 2010) was used as a method to rank the typographic factors in a prioritized list. By making this ranking the guidelines can communicate what is most crucial for the readers in bright light, so that if not all guidelines can be followed the designers can make an informed choice of where their effort should be placed. The different typographic factors were rated by their impact on legibility in the specified use case and how much each would affect the final design.

Affinity diagram

Affinity diagram is frequently used to analyse quantitative data by grouping similar concepts to identify findings and patterns that emerge (Courage et al. 2015). This was used as a method to aggregate and visualize the different books, research articles, theses, guidelines and other literature reviewed. The literature reviewed in chapter 2.1 was reviewed again and put into the affinity diagram. The affinity diagram was made in Miro (an online collaboration whiteboard software) and used to further design the different guidelines.

3.2 Online experiment

3.2.1 Purpose

From the review on typographical factors, two factors were identified to be investigated further through an online experiment. The objective of this online experiment was to see whether difference in stroke contrast or colour influenced legibility when it's being read in bright outdoor conditions. Findings from this experiment will be added to the final guidelines. As reviewed earlier in section 2.3 bright light faced outdoors especially in the summer has high illumination levels and a cold colour temperature that potentially can effect these variables. The high levels of illumination can reduce legibility by making it harder to identify sufficient details which affect our ability to read. Cold colour temperature can change the perceived colour with its blue light or these short wavelengths can make content appear more dazzling. Following is the hypothesis to be tested in the experiment, along with the independent and dependent variables:

$H_1$1: Difference in stroke contrast influence legibility in bright outdoor conditions
$H_0$1: Difference in stroke contrast have no effect on legibility in bright outdoor conditions

$H_1$2: Difference in colour influence legibility in bright outdoor conditions
$H_0$2 Difference in colour have no effect on legibility in bright outdoor conditions

Dependent variable: Legibility measured by subjective rating
Independent variables: Bright outdoor conditions (lux and temperature), difference in stroke contrast and choice of colour

3.2.2 Experiment design

The experiment was designed to account for the physical experiment first planned (see chapter 5.2.2), to conduct it in a feasible way considerations was made to accommodate for of this. It was important when choosing the software to build an experiment that allowed for customization and adjusting the design to keep it as simple as possible and remove other elements that might affect the stimuli. Another need for the survey tool provider was that it should be secure and allow for answers to be registered anonymously (not registering IP addresses) and supporting responsive design so it could be conducted by phone. SurveyMonkey satisfied the needs for the experiment and was chosen as software.

As an online experiment can be conducted anywhere at any time and so many variables can influence the individual reading experience, an effort was made to somewhat control the conditions and a framework was put in place for the participants when taking the experiment. In order to get more comparable results all participants were instructed to conduct the experiment outdoors in bright daylight with clear blue sky. To control for this the participants was also asked about their surroundings before presenting the stimuli. Following is the framework participants was instructed to follow before the experiment began:

1. Adaptive display functions are turned off on your phone (auto brightness, night mode and adaptive white balance)
    - Guide to turn off on Apple phones
    - On Android phones open "settings", tap "display" and you should see the functions
2. Use glasses or lenses if you use them on a regular basis
3. Don't use sunglasses
4. The brightness on your display is at maximum
5. You are outdoors in bright daylight and clear blue sky
6. Minimize the reflection in your screen, by keeping some distance to your surroundings
7. Don't stand in the shadow
8. Keep your phone at an arm's length distance during the questionnaire

A wide range of mobile phones would most likely be used to conduct the experiment, so by having participants report what type of phones they are using and by turning its brightness to max this could be more comparable when analysing the results. With the information the different phones contrast ratio could potentially be calculated based on its standard maximum white point and black point. The participants were also asked if they have accurate vision/corrected vision and colour vision as their ability to see correctly could influence their perceived legibility, especially choice of colour. To evaluate legibility participants was asked to look at multiple stimulus presenting text in various ways and rate how each text sample was to read.

A Likert scale was used as it is a good way to measure participant evaluation (Leedy & Ormrod 2015), with a rating scale of 1–4. When using a rating scale, a neutral option is often added in the middle to accommodate for neutral opinions, this was not deemed necessary here, as the two middle options are of a more neutral nature that leans one or the other way. Rating scale and wording is consistent with a similar study with subjective measurements conducted on printed text (Nersveen et al. 2018). The rating scale was presented like this in the experiment:

- Easy to read
- Readable, with some discomfort
- Readable, but hard
- Not readable

The general design of the experiment was also considered according to the discussion about surveys of Courage et al. (2015). When building an online experiment, some parts similar to a survey as participants do it on their own, it is important to increase the participants understanding and satisfaction. The title gave the respondents an instant sense of the purpose. On the first page instructions was laid out with detailed contact information, the purpose, time needed to complete the experiment and details about privacy and anonymity according to NSD template. Due to the mix between a survey and a controlled experiment, the time may have suffered from this and was estimated to take 25 minutes. The visual design used a sans serif font for its content, with large bold version to mark headlines. Cluster and other disturbing elements were removed to only show what was needed, especially elements that could disturb the stimuli was focused upon. Each stimulus was presented on separate pages to reduce cluster and to give a better estimate of remaining experiment in the progress bar added at the bottom of the page.

3.2.3 Stimuli

The stimuli consisted of 11 randomly chosen words presented in different ways to evaluate the independent variables. In order to customize the font, size and colour, it was made in Adobe XD and exported as svg files before implemented in the experiment. By using svg files the content stayed fully scalable in vector format, which makes it resolution independent and won't be rasterized in advance. The phones used to take the experiment would vary in width, in order to compensate for this and to avoid scaling of the image which would change the font size, the image width was kept at 260 px. This would allow the stimuli to be presented close to consistent at all phones, including the smallest phones (320 px including website padding). The text was presented over two lines with a linear spacing at 120%. As positive polarity indicates better performance and greater legibility (Buchner & Baumgartner 2007, Dobres, Chahine & Reimer 2017) the polarity of the stimuli was kept positive, with darker text on a brighter background. Stimuli with both variables was randomized before implementing it in the experiment.

Stroke contrast

As reviewed through chapter 2.1 many variables influence the reading experience. In order to account for these and preserve the validity, a font especially designed to only vary in stroke contrast while the other variables stay consistent was used (Perondi et al. 2017). The stimuli were presented at a 16 px size, more specifically an x-height of 1,9 mm and cap height of 2,8152 mm. The experiment was conducted in bright daylight where the levels of ambient illumination was high, but still can reach even brighter levels during summer, to the point where it can get even harder to read on a mobile phone. To reach similar type of conditions and to see if there was a threshold between the fonts, further simulation was needed by lowering the contrast of the text stimuli equivalent to the loss caused by high illumination levels. Achromatic colours (grey scale) was used for the stimuli, as a difference in lightness of the colour is most effective in terms of contrast, compared to hue or chroma. The different contrast levels were equally divided from a minimum at 2:1 up to 21:1. The background stayed white at maximum brightness to keep the level of reflection consistent, while the text colour changed to change the contrast. The different levels of contrast used is presented in Table 1 along with the calculated effective contrast ratio using Pignoni (2018*b*) contrast analyser with display characteristics from DisplayMate (2019) for an iPhone 11 Pro in full daylight without direct sunlight (10.000 lux) and direct sunlight (100.000 lux). In total the number of stimulus for this variable was 2 fonts, repeated 3 times at 25 contrast levels, resulting in 150 stimuli.

Choice of colour

In studies researching colour and legibility 6 colours seems to be commonly used - red, green, blue, yellow, purple/magenta and cyan (Humar et al. 2014, Lin & Huang 2013, Lin 2005). These 6 colours are also what constitutes the light wheel, which is based on additive colour mixing consisting of 3 primaries (red, green and blue) and 3 secondaries (yellow, magenta and cyan) (Feisner & Reed 2014). Therefore, these colours was used and limited to these 6 along with adding grey of equal contrast, as adding more colours would ultimately only by a further mixing of these colours. The specific colours used is based on an evenly and widely distribution in the chromaticity space (Lin & Huang 2013) and slightly adjusted to have approximately the same contrast ratio. By basing the colours on a similar chromatically contrast they are more comparable and should be influenced more equally by the higher levels of ambient illumination, as some colours are naturally brighter than others and the hue will then be the measured effect of the colour. As shown in Table 2 the chosen colours are presented along with RGB values and contrast ratio. These stimuli was presented in Times New Roman as this font is commonly used in legibility evaluations, at 16 px size, equal to a x-height of 1,9 mm and a cap height of 2,814 mm. The total number of stimuli resulted in 35 text samples, as 7 colours was repeated 5 times.

Levels of contrast ratio

sRGB text	Contrast ratio	10.000 lux	100.000 lux
0	21:1	7.28:1	1.63:1
8	20.8:1	7.12:1	1.63:1
15	19.16:1	6.91:1	1.62:1
23	17.92:1	6.7:1	1.62:1
31	16.48:1	6.44:1	1.61:1
38	15.13:1	6.19:1	1.6:1
46	13.57:1	5.92:1	1.59:1
54	12.08:1	5.58:1	1.58:1
61	10.86:1	5.29:1	1.57:1
69	9.58:1	4.95:1	1.56:1
77	8.45:1	4.61:1	1.54:1
84	7.57:1	4.33:1	1.53:1
92	6.68:1	4.03:1	1.51:1
100	5.91:1	3.73:1	1.49:1
107	5.32:1	3.5:1	1.47:1
115	4.74:1	3.24:1	1.45:1
123	4.23:1	3.01:1	1.43:1
130	3.84:1	2.81:1	1.41:1
138	3.45:1	2.61:1	1.38:1
146	3.11:1	2.43:1	1.36:1
153	2.84:1	2.28:1	1.34:1
161	2.58:1	2.12:1	1.31:1
169	2.35:1	1.97:1	1.28:1
176	2.16:1	1.86:1	1.26:1
184	1.92:1	1.73:1	1.23:1

Table 1: Reduction in contrast: sRGB values for text, contrast ratio defined in WCAG and calculated effective contrast ratio for iPhone 11 Pro at 10.000 and 100.000 lux.

Difference in colour

Colour	sRGB	Contrast ratio
Red	218, 29, 1	5.05:1
Green	1, 131, 0	4.94:1
Blue	98, 84, 255	4.96:1
Yellow	136, 108, 0	5:1
Magenta	189, 46, 182	4.96:1
Cyan	1, 125, 106	5.06:1
Grey	112, 112, 112	4.95:1

Table 2: Colour choices with sRGB values and contrast ratio defined in WCAG.

3.2.4 Pilot test

The experiment was evaluated several times while it was being designed and a pilot test was conducted with one person when the whole experiment was ready. The main purpose of this was to see how long it would take and identify any potential problems with wording, level of understanding or other potential misunderstandings, especially related to the framework of how the experiment should be conducted. During the pilot test 1 remark was made by the participant and corrected before distribution.

3.2.5 Participants and selection

The participants of this study could be anyone in the general population with accurate or corrected accurate vision and colour vision that have access to a mobile phone. The recruitment of participants was done by convenience sampling. The experiment was distributed on Facebook by sending personal messages asking people to participate in the experiment and posting in different groups. The threshold for participation is rather high with a long completion time and needing to be conducted outdoors, so it's likely that the participants would be people the researcher knows.

3.3 Ethical and legal considerations

In the planning of the online experiment it was considered if a notification to the Norwegian Center for Research Data (NSD) was needed. The personal data collected was gender and age range, and data regarding the participants vision and colour vision that could be considered as health data. By turning on anonymous answers when using SurveyMonkey no IP addresses or other electronic tracks was collected that could be traced back to the respondent. As none of the data will identify a person it was decided that a notification was not needed, after talking with NSD clarifying about the health data (appendix A.1).

Even though it wasn't considered necessary to send a notification, it is still important and ethical to inform participants. All participant was presented with information about the project and the experiment based on the guidelines from NSD about informed consent and was required to accept that they had read and understood the information before starting the experiment. No compensation or benefits was gained by participating in the experiment.

4 Results

This chapter will provide results from the methods used to create the guidelines, results from the data analysis of the online experiment and finally the final set of guidelines. The data from the online experiment was analysed using IBM SPSS.

4.1 Design of guidelines

In the process of designing the guidelines several user-centered and gamestorming methods was applied to structure and design the guidelines, like use case, pain-gain map, forced ranking and affinity diagram. The results from the used methods will be presented in this section, while the final guidelines is presented in section 4.3.

4.1.1 Creation methods

Definition of use case and pain-gain map

Use case or a scenario can be defined in several ways, here it was used to describe the context and problem areas, potentially evoking reflection about design issues with a vivid description (Benyon 2013). Many more concrete scenarios and user tasks can be performed in this context, so the objective was to outline the effects of high ambient illumination to give an understanding of its effect to be detailed further in specific projects. Following is the defined use case that are included with the guidelines:

> Mobile phones are used whereever and whenever it might suit us, more frequently happening outdoors. The physical constraints of a phone mostly rely on conveying information and navigation through text and visual content. Bright sunlight can greatly affect legibility by reducing what we are able to see on the display, effecting contrast and details. Several situations require an interface or website to be legible in these challenging conditions, so how can we design a text that will withstand this?

To further concretize the relationship between legibility and the effects of this use case, a pain-gain map was made as shown in Table 3. This pointed to specific challenges or "pains" and what "gains" can be done to improve them. The objective with a pain-gain map was to better understand motivations and decisions, an effort to increase the understanding why the guidelines should be followed.

Pain-Gain map

Pains	Gains
The bright sun causes a lower effective contrast ratio	Design with high contrast ratio, especially in body text
The bright sun can make it harder to read on a display	Using a bigger font size and supporting scalable text makes it easier to read
Reflection and glare can make it hard to differentiate details	Choose a well established font with details that don't easily merge together
Different technologies and users needs makes it hard to generalize	Design legibility to be robust enough to cope with these differences

Table 3: Pain-gain map for guidelines

Forced ranking

This method forced a ranking of the different typographical factors based on their effect on legibility in the specified use case and their effect on the final design. The effect on the final design was used to emphasize the importance of implementing this at an early stage by increasing its importance. In section 2.1 multiple factors was presented and discussed theoretically, not all of these can be handle individually but are dependent on each other. Some of the factors was there for grouped together in more meaningful groupings; sans/sans serif, stroke contrast, width and other characteristic of the font was grouped and polarity was combined with colour and contrast. Each item was ranked relative to the others (none could get the same score) resulting in a prioritized list (Gray et al. 2010) as shown in Table 4.

Forced ranking

Typographical factor	Impact on legibility	Effect on final design	Total
Font	4	2	6
Weight	6	4	10
Lower or uppercase	5	6	11
Size	2	3	5
Spacing	3	5	8
Line length	7	7	14
Polarity, colour and contrast	1	1	2

Table 4: Forced ranking of typographical factors

Affinity diagram

Based on the review of research, books, theses, collections of research, guidelines and other literature done in chapter 2 an affinity diagram was made. By re-visiting and reviewing these sources relevant findings was written on virtual post-it notes in Miro and organized by typographic factors (Figure 3). This way all the knowledge was shown in a more visual way, making it easier to aggregate further into defining the guidelines.

Figure 3: Affinity diagram

4.1.2 Background and sources for the guidelines

The final guidelines were based on previous reviews and visualized through an affinity diagram. The data constituting the affinity diagram was extracted and presented in Table 5. This is the main basis for the guidelines and findings from the online experiment (section 4.2) will be added before presenting the final guidelines in section 4.3.

Basis and sources for final guideline

Factor	Information	Source
Colour use	Minimum relative contrast ratio (3:1 and 4.5:1) or enhanced contrast ratio (4.5:1 and 7:1)	W3C (2019)
	High contrast and plain background	Nielsen (2015)
	Reference to WCAG: 3:1 and 4.5:1. Don't use red and green, blue and yellow or red and blue/purple	Breuninger (2019)
	Challenging for web due to limited control. Embedding contrast-limit constraints to design process.	Sandnes (2017)
	Positive polarity only way to achieve 90% acceptance. Highest possible. 100%-60% black and 0%-60% black maintains 80%	Nersveen & Johansen (2016)
	More important than font choice. Refers to WCAG contrast levels	Halbach & Fuglerud (2018)
	Ensure sufficient contrast in hue, value or saturation.	Carter et al. (2002)
	Consistently better with positive polarity. Independent of ambient lighting and of chromaticity.	Buchner & Baumgartner (2007)
	Black on white, black on yellow, blue on white and blue on yellow to be most legible on LCD displays	Humar et al. (2014)
	Suggest a contrast ratio of 18:1 as minimum in high illumination	Pignoni (2018*a*)

Basis and sources for final guideline

Colour use	Positive polarity advantage on displays in dark and bright illumination	Dobres, Chahine & Reimer (2017)
	Positive polarity had a strong legibility advantage	Dobres, Chahine, Reimer, Gould, Mehler & Coughlin (2016)
	Positive polarity is better than negative (max contrast)	Dobres, Chahine, Reimer, Gould & Zhao (2016)
Font size	Scalable to 200%	W3C (2019)
	Reasonably large default font size and allow for changing it	Nielsen (2015)
	Dependent on viewing distance. Calculate based on devices and relevant viewing distances. (0.00582*viewing distances for headlines and important labels and 0.00465 x viewing distance for the text body. Defined cap height	Breuninger (2019)
	Recommended cap height 20-22 arc minutes, minimum 16. 30 cm reading distance = 1.7mm rec cap height, 40 cm = 2.3 mm, 50 cm=2.9 mm	International Organization Standards (2011)
	Should be defined/compared based on x-height	Lund (1999)
	12 pt sufficient of 80% acceptance.	Nersveen & Johansen (2016)
	Large x-height suggest enhanced legibility, especially at smaller sizes. Display resolution control text presentation.	Beier (2012)
	Use cap height/h height to compare fonts	Reimer et al. (2014)
	Reduced vision at least 16 pt on print. More important than font choice.	Halbach & Fuglerud (2018)
	Generally range from 8 to 12 pt. Be aware of x-height	Carter et al. (2002)
	Economical and ergonomic balanse. Larger text is better.	Bigelow (2019)
	X-height to define size. For print. Reading distance 40 cm corresponding x-height 1,4 mm for text and 14 mm for headlines.	Legge & Bigelow (2011)
	Larger more legible than smaller, (H height at 3 and 4 m). 3 pixels separating sizes	Dobres, Chahine, Reimer, Gould, Mehler & Coughlin (2016)
	Larger more legible than smaller, (H height at 3 and 4 m)	Dobres et al. (2018)
Font choice	Extraordinary thin strokes or unusual characteristics are harder to read	W3C (2018*b*)
	A clean typeface without strange shapes	Nielsen (2015)
	Serifs ok with sufficient resolution. Condensed only with good text size and contrast.	Breuninger (2019)

Basis and sources for final guideline

Font choice	Serifs don't increase legibility. Scala Sans bold was sig better than others tested	Nersveen & Johansen (2016)
	Serifs help keep track of the line, no clear answer is found. Tendency toward low stroke contrast. Prefer wider forms over narrower	Beier (2012)
	Humanistic typeface. Open shapes, unambiguous forms and varying horizontal proportions	Reimer et al. (2014)
	Typeface and serifs don't seem that important	Halbach & Fuglerud (2018)
	Choose classical, time-tested typefaces. Medium with.	Carter et al. (2002)
	No sig difference on serif. Difference in stroke contrast aren't highly researched.	Bigelow (2019)
	Continuous text in newspaper and books almost only serif. 9 of 10 most used websites fonts are sans serif.	Legge & Bigelow (2011)
	Humanistic type had an advantage compared to square grotesque. More distinctive at smaller size. Show that instrinsic factors may also interact with extrinsic.	Dobres, Chahine, Reimer, Gould, Mehler & Coughlin (2016)
	Open letterforms, varying shapes and generous x-height are more legible.	Dobres, Chrysler, Wolfe, Chahine & Reimer (2017)
Spacing	Line spacing: 1.5 x font size. After paragraph: 2 x. Letter spacing: 0.12 x. Word spacing: 0.16 x font size.	W3C (2019) WCAG
	Ample intercharacter spacing to prevent them to blur together	Reimer et al. (2014)
	Need more knowledge and research	Halbach & Fuglerud (2018)
	Consistent, letter and word spacing. Line spacing that easily carries the eye. Indicate new paragraph.	Carter et al. (2002)
	Tracking usually fine as is, slightly increasing might help dyslexics.	Bigelow (2019)
	Too lose or tight tracking or leading will disrupt the reader	Rannem (2012)
	Wider leading enhance legibility	Dobres et al. (2018)
Font weight	Extraordinary thin strokes are harder to reader	W3C (2018*b*)
	Only light when good size and contrast. Bold improve legibility in difficult conditions.	Breuninger (2019)
	Bold weight significant increase the acceptance rate	Nersveen & Johansen (2016)
	Seem more important than font choice.	Halbach & Fuglerud (2018)
	Avoid too heavy or too light	Carter et al. (2002)

Basis and sources for final guideline

Font weight	Lightest and boldest can seem to reduce reading speed.	Bigelow (2019)
	Bold easier than medium in glance conditions	Dobres, Chahine, Reimer, Gould & Zhao (2016) Dobres, Chahine, Reimer, Gould & Zhao, 2016
	Light weight are superior in optimal rendering conditions, in suboptimal or good enough rendering it degrades more than other weights. No sig dif between regular, medium and bold.	Dobres, Chahine & Reimer (2017)
Lower- or uppercase	Prefer lowercase in continuous reading, mostly due to habit. Don't exclude uppercase	Beier (2012)
	Use upper- and lowercase for optimum readability	Carter et al. (2002)
Line length	Responsive design without scrolling horizontally. Not bigger than device with.	W3C (2019)
	Appropriate line lengths. Max 70 characters is acceeptable.	Carter et al. (2002)
	Not too long or too short. Medium at 55 characters per line	Dyson & Haselgrove (2001)

Table 5: Basis and sources for final guideline

4.2 Online experiment

A total of 24 participants completed the online experiment. As correct or corrected vision was a prerequisite one participant was excluded and one reporting to conducting the experiment inside was excluded, leaving results from 22 participants to be analysed further. One participant reported to be colour blind and one had a gap in age compared to the others, but this was considered to not influence their mean scores differently and was used in the analysis. Figure 4 shows the distribution of age range and gender of participants. Age range 18+, 45-54 and 65+ was also included in the experiment but had no participants. 55% of participants was between the age of 25-34 and 77% were female.

The online experiment was conducted outside in the spring in early May. 55% reported the weather to be bright sunlight and clear blue sky while 45% reported it to be bright sunlight with some clouds. The experiment start time ranged from 09:45 to 18:30 and was distributed throughout the day. 45% of the participants used an iPhone, while the rest was distributed on a range of other brands like Samsung and Huawei. In Table 6 a more detailed overview of the 20 different mobile phones used is presented with display characteristics. Displays characteristics was gathered from their manufacturer's websites or from reviews, some was missing PPI information and was calculated based on the displays size and resolution. An effort was also made to find the displays contrast ratio, white (maximum brightness) and black point for further analysis, but it was hard to find comparable sources or information at all for some displays, so these characteristics was excluded from the analysis.

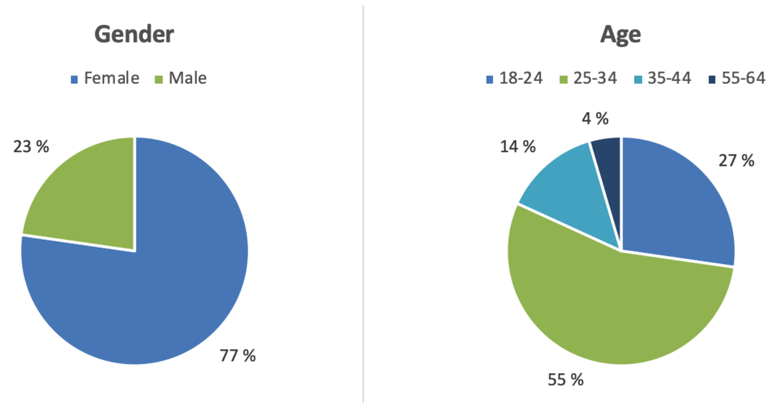


Figure 4: Age and gender distribution of participants

Phone characteristics

	Frequency	Display technology	Size	Resolution	PPI
iPhone 11 pro	1	OLED	5.8”	2436x1125	458
iPhone XS	1	OLED	5.8”	2436x1125	458
iPhone XS max	2	OLED	6.5”	2688x1242	458
iPhone X	1	OLED	5.8”	2436x1125	458
iPhone 8	1	LCD	4.7”	1334x750	326
iPhone 7 plus	1	LCD	5.5”	1920x1080	401
iPhone 7	2	LCD	4.7”	1334x750	326
iPhone 5S	1	LED	4.0”	1136x640	326
Samsung Galaxy S8+	1	AMOLED	6.2”	2960x1440	529
Samsung Galaxy S7	1	AMOLED	5.1”	2560x1440	577
Samsung Galaxy J6	1	AMOLED	5.6”	1480x720	294
Samsung Galaxy J5	1	AMOLED	5.0”	1280x720	294
OnePlus 6	1	AMOLED	6.3”	2280x1080	400
OnePlus 5	1	AMOLED	5.5”	1920x1080	401
Huawei p20	1	LCD	5.8”	2240x1080	429
Huawei mate 20 pro	1	OLED	6,39”	3120x1440	538
Huawei p30	1	OLED	6.1”	2340x1080	423
Motorola M2670	1	-	-	-	-
Google Pixel 2	1	AMOLED	5.0”	1920x1080	441
HTC10	1	LCD	5.2”	2560x1440	564

Table 6: Characteristics of mobile phones used to conduct the experiment

4.2.1 Stroke contrast

Legibility was rated using a scale from 1-4, 1 meaning "not readable" and 4 "easy to read". The acceptable level of legibility was set at scores of 3>. This data was analysed through investigating mean scores, running paired samples t-test and general linear mode on repeated measurements. The independent variable of difference in stroke contrast was divided in two levels; a font with no stroke contrast and one with stroke contrast. This analysis was performed to test the first $H_1 1$ *Difference in stroke contrast influence legibility in bright outdoor conditions*. Table 7 present descriptive analysis from the whole data set and a selection of the last 13 stimuli as this was where the ratings felled below the threshold. In Figure 5 mean rating scores are presented as legibility decrease.

On investigation of the mean scores little difference was found, only a difference of 0,0175 separated difference in stroke contrast and the difference was even smaller on the selection of the last 13 stimuli. The paired samples t-test found that there was no statistical significance between legibility and difference in stroke contrast in either pairs. Stroke contrast and no stroke contrast t=-1.256 p=.223, and selection of last 13 no stroke contrast and stroke contrast t=.804 sig=.430. These result suggests that there is no correlation between legibility and difference in stroke contrast and the $H_0 1$ is kept, "*difference in stroke contrast have no effect on legibility in bright outdoor conditions*".

Descriptive statistics

	Mean	N	Std. Deviation	Std. Error Mean
No stroke contrast	3.3867	22	.34118	.07274
Stroke contrast	3.4042	22	.34086	.07267
No stroke contrast (last 13)	2.9860	22	.54741	.11671
Stroke contrast (last 13)	2.9697	22	.52098	.11107

Table 7: Descriptive statistics: difference in stroke contrast results

Figure 5 show how the ratings gradually decreases along with the decreasing contrast ratio, as expected since reduced contrast ratio result in poorer legibility. Both variables seemed to almost simultaneously drop below the acceptable legibility level of 3, happening around stimuli 19 with text at a contrast ratio of 3.45. Even though no significant difference was found, the falling curve confirms that the simulation of brighter condition had an effect that dropped below the acceptable level of legibility.

By collecting this data with a online experiment other independent variables was added as the participants used their own phones, resulting in a wider sampling of mobile phones and characteristics; display technology, screen size, resolution and PPI (overview in Table 6). The general linear model was used to investigate if any of these variables correlated with legibility. Display technology, screen size and resolution did not result in any significant findings. However, on further investigation of the plot of display technology (Figure 6) a tendency suggested improved legibility on newer display technology. There was a significant between-subjects effect between PPI and legibility F=3.586 sig=.049. This effect is plotted in Figure 7 suggesting that higher PPI result in a general higher level of legibility. A general linear model was also performed finding PPI to be the most important variable with significance=.002.

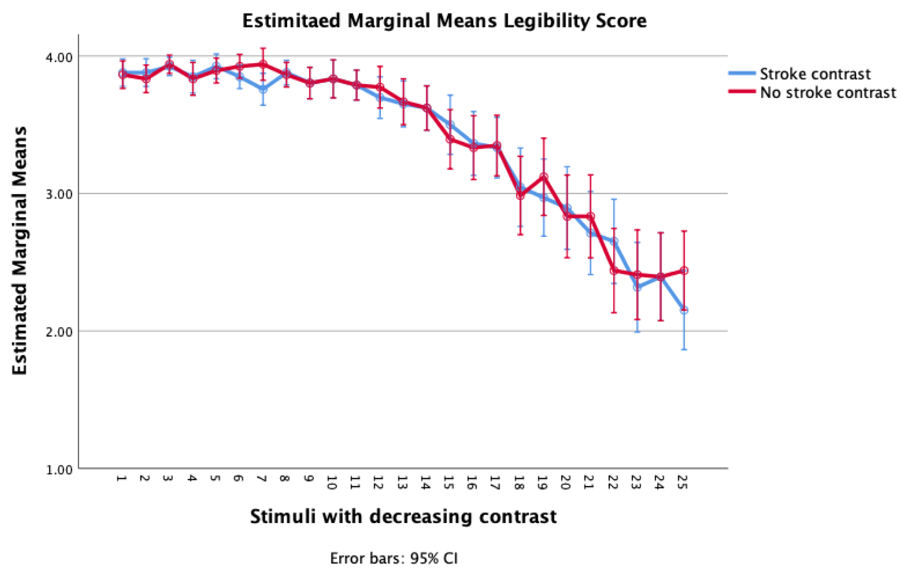


Figure 5: Mean legibility scores of difference in stroke contrast

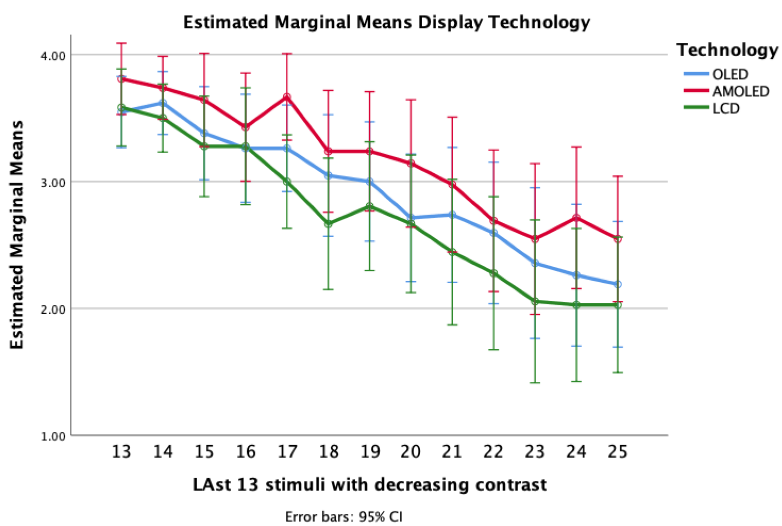


Figure 6: Effects between display technology and legibility ratings on difference in stroke contrast

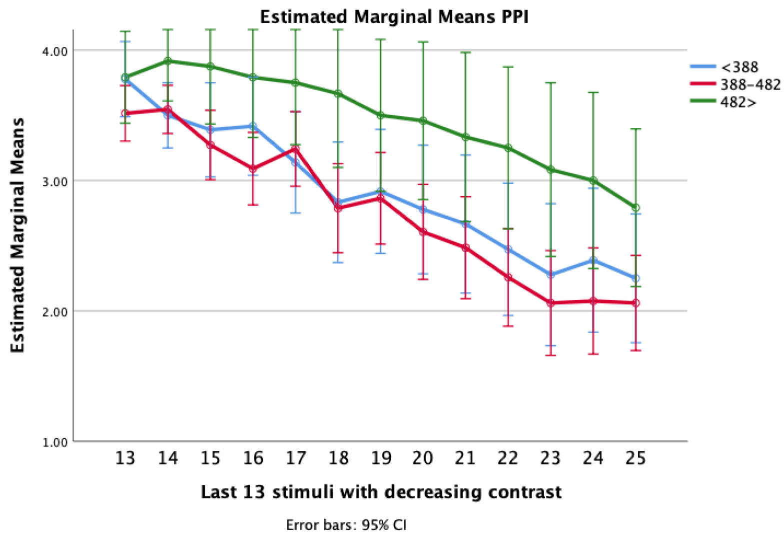


Figure 7: Effects between PPI and legibility ratings on difference in stroke contrast

4.2.2 Choice of colour

This data was analysed by investigating mean scores and running general linear model on repeated measurements. The independent variable was difference in colour, represented by 7 different colours (one grey) with similar contrast ratio. This statistics was used to test the second $H_1 2$ *difference in colour influence legibility in bright outdoor conditions*. Descriptive analysis of the colour data analysis is presented in Table 7. Mean scores showed no significant difference and stayed consistent over the acceptable legibility level of 3>. This suggest that there are no significant correlation between legibility and differences in colour and the $H_0 2$ is kept "*choice of colour have no effect on legibility in bright outdoor conditions*".

Descriptive statistics

	Mean	N	Std. Deviation
Grey	3.3273	22	.48026
Yellow	3.3606	22	.51131
Red	3.3545	22	.45327
Magenta	3.2182	22	.44469
Blue	3.3636	22	.47263
Cyan	3.4159	22	.41616
Green	3.4273	22	.51379

Table 8: Descriptive statistics: difference in colour results

General linear model was also performed to investigate effects between difference in colour and display characteristics. There was no significant effect found with either display technology, screen size, resolution or PPI. Even though PPI didn't show a significant effect with this data-set, Figure 8 shows the same tendency as the significant effect found between PPI and legibility ratings on difference in stroke contrast (see Figure 7).

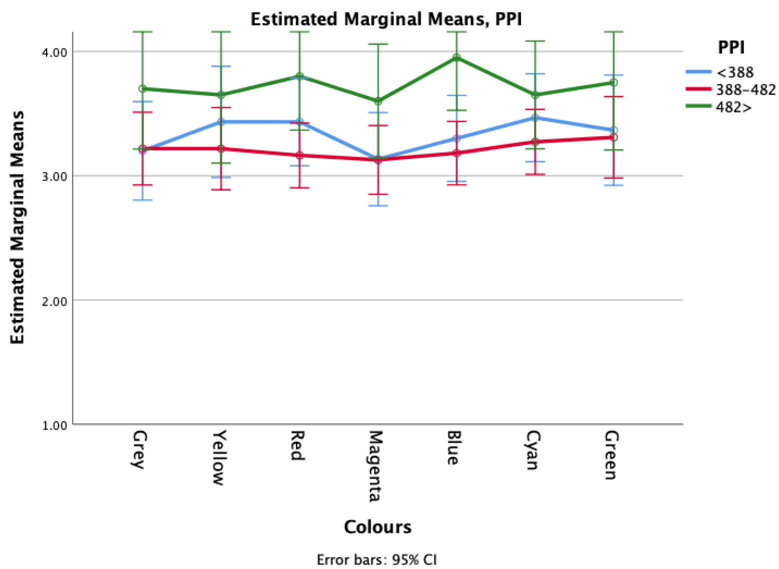


Figure 8: Effects between PPI and legibility ratings on colour differences

4.3 Final guidelines

The final guidelines are a result of combining reviewed sources consisting of 4 guidelines and 20 additional sources extracted in Table 5, and results from the online experiment added some additional findings. Final guidelines was designed combining the use case and the pain-gain map with the guidelines as presented in Figure 9.

1. Colour use

- Use a minimum contrast ratio of 18:1 for body text. Contrast ratio can be reduced with bigger size or when used on graphics
- Keep a positive polarity (dark text on light background)
- Sufficient contrast ratio is more important than what colour you use
- Don't combine red and green, blue and yellow or red and blue/purple without sufficient contrast in lightness

2. Font size

- 18-20 px is a good text size for body text
- To set body size specific to a font, calculate by using ISO 9241-303:2011

- Using a font with high x-height can allow for some reduction in size, while a low x-height might need to be increased
- Make sure text is scalable and functional, but don't depend on it to be scaled

3. Font choice

- Use a well-established serif or sans serif font without strange shapes, normal with and spacing
- A font with open letterforms, varying shapes, unambiguous forms and a generous x-height can increase legibility further

4. Spacing

- Increase the standard leading/linear spacing to 1.33-1.5 times the font size. Fonts with high x-height need some wider leading than lower x-heights
- Increasing tracking slightly by 0-0,12 times font size is helpful for dyslexics

5. Font weight

- Regular, medium or bold weight is preferred
- Use bold weight to improve legibility if other factors are compromised
- Avoid to heavy or too light weights

6. Other

- Use lower- and uppercase as normal in body text
- Uppercase can be considered in headlines or as means of creating contrast in other short phrases
- Keep the line length close to 100% of the mobile device width

How to: Design for Legibility on Mobile Phones in Bright Outdoor Conditions

Mobile phones are used where ever and whenever it might suit us, more frequently happening outdoors. The physical constraints of a phone mostly rely on conveying information and navigation through text and visual content. Bright sunlight can greatly affect legibility by reducing what we are able to see on the display, effecting contrast and details. Several situations require a interface or website to be legible in these challenging conditions, so how can we design a text that will withstand this?

- The bright sun causes a lower effective contrast ratio
- The bright sun can make it harder to read on a display
- Reflection and glare can make it hard to differentiate details
- Different technologies and users needs, makes it hard to generalize

- Design with high contrast ratio, especially in body text
- Using a bigger font size and supporting scalable text makes it easier to read
- Choose a well established font with details that don't easily merge together
- Design legibility to be robust enough to cope with these differences

1 COLOUR USE

- Use a minimum contrast ratio of 18:1 for body text. Contrast ratio can be reduced with bigger size or when used on graphics
- Keep a positive polarity (dark on light background)
- Sufficient contrast ratio is more important than what colour you use
- Don't combine red and green, blue and yellow or red and blue/purple without sufficient contrast in lightness

2 FONT SIZE

- 18–20 px is a good text size for body text
- To set body size specific to a font, calculate by using ISO 9241-303:2011
- Using a font with high x-height can allow for some reduction in size, while a low x-height might need to be increased
- Make sure text is scalable and functional, but don't depend on it to be scaled

3 FONT CHOICE

- Use a well-established serif or sans serif font without strange shapes, normal with and spacing
- A font with open letterforms, varying shapes, unambiguous forms and a generous x-height can increase legibility further

4 SPACING

- Increase the standard leading/linear spacing to 1.33-1.5 times the font size. Fonts with high x-height need some wider leading than lower x-heights
- Increasing tracking slightly by 0-0,12 times font size is helpful for dyslexics

5 FONT WEIGHT

- Regular, medium or bold weight is preferred
- Use bold weight to improve legibility if other factors are compromised
- Avoid to heavy or too light weights

6 OTHER

- Use lower- and uppercase as normal in body text
- Uppercase can be considered in headlines or as means of creating contrast in other shortphrases
- Keep the line length close to 100% of the mobile device width

Figure 9: Final guidelines

5 Discussion

This chapter will first discuss how the previous stated research questions have been answered through this thesis and then move on discussing the final guidelines, online experiment and interpretation of the results.

The overall primary research question "*how can designers ensure sufficient legibility on mobile phones in bright outdoor conditions?*" is mainly answered with the final guidelines presented in section 4.3 and through the additional sub questions. In order to answer this thoroughly a deeper understanding of "*which variables and to what extent they influence legibility*" was needed and have been answered and clearly laid out through the review in chapter 2. Difference in stroke contrast and colour was identified as factors for further research and was investigated through the methodology presented in section 3.2, answering the next two sub questions without any significant effect (section 4.2. The last research question "*how bright outdoor conditions effect legibility on displays*"*?* have been pervasive throughout the thesis, but primary discussed in section 2.5.

5.1 Guidelines

As previously stated, the purpose of the guidelines created in this thesis was to actively focus on the effects of the use case and include this when designing them. The result was guidelines addressing 6 factors effecting legibility in a prioritized order with further explanation of each. This was deemed necessary as typography is not absolute, for instance a font size aren't absolute due to the varying x-height and cap height, but at the same time most don't use these heights to determine font size. The different factors will be relative to each other and so the intention resulting in providing guidelines that takes a stand, but somewhat opens up for further consideration.

The main motivation for creating these guidelines was the effects of use case. Since there already are a lot of research on different typographical factors the focus has been on reviewing and evaluating this. Through the review some factors appeared as more critical like contrast and font size, while others didn't influence much. Some factors were also revealed as not so highly research topics, so an experiment branched out from the review to further investigate difference in stroke contrast and difference in colour.

Making content accessible for everyone have been mentioned earlier and how human factors can affect legibility have been explained. If comparing these guidelines with the criteria defined by W3C they should provide sufficient legibility within these requirements as well. However, this have not been tested and the guidelines would need further testing and evaluation with a wider user group in even brighter light if possible. Not many of the reviewed guidelines focused on a specific use case, and unlike them that is the main focus here. A specific use case might be hard to apply as many websites and other user interfaces are used in various situations. However, due to a totally different tolerance level these guidelines will provide sufficient legibility nearly regardless of the effects from its surroundings.

5.2 Online experiment

Due to the Covid-19 pandemic the methodology planned to investigate difference in stroke contrast and colour needed to change during this project. A controlled physical experiment (elaborated in section 5.2.2) was replaced with the online experiment laid out in methodology section 3.2. In the transition to a new methodology other approaches to the online experiment was considered and will be discussed before moving on to the discussion of experiment results and outlines of the planned experiment.

The focus when shifting to an online experiment was how to it could be controlled, to the best of our ability. One alternative was to instruct the participants to conduct the experiment with simulated stimuli in a completely dark room with the brightness on their phones turned to max. By having them report the make and model of their phone the results could have been closer calculated and analysed, which would have increased reliability and validity of the study and the effect of the displays could be seen. However, these conditions contradict the use case in focus and offers other challenges like the phones maximum brightness would be dazzling for participants in a dark room. Potential findings could then be caused by the dazzling effect and not the simulated change in illumination level and colour temperature. Another approached considered was to instruct the participants to take the experiment several times during different times of a day with bright sunlight. This would require them to encounter the natural changes of colour temperature during the day. By recruiting participants in the same area or city and coordinating the experiment to take place at the same times, maybe in several groups in different locations, measurements could have been made in the same area with a lux meter to get an idea of the kelvin and lux level. Several challenges were linked to this approach with recruitment of participants, proper conduct of the experiment and access to equipment and was assessed as not feasible at that point. The approach used in the online experiment instructed all participants to conduct the experiment outdoors in bright daylight and clear blue sky to achieve a "controlled and comparable environment" to the best ability.

5.2.1 Discussion of online experiment results

The results from the analysis on both difference in stroke contrast and colour showed to have no significant effect and both H_0 was kept. Ratings for stroke contrast gradually decreased along with gradually harder reading conditions forced with reduction in contrast ratio (see Figure 5). This showed that the experiment design had an effect and a level of not acceptable legibility was reached 3<, but caused by the reduction in contrast ratio and not difference in stroke contrast. If the difference in stroke contrast was even bigger with very thin hairlines a difference could potentially emerge, but fonts with these big stroke contrast would not be a good font choice to use in body text. Ratings evaluating difference in colour did all stay above the acceptable legibility level and at similar mean scores (see Table 2). Magenta had the lowest mean score and was the least legible of the colours, but no significant difference was found. The different colours were equal in contrast ratio, which suggest that sufficient level of contrast ratio is the decisive factor and not difference in colour.

On further investigation of the additional independent variables from the different mobile phones, PPI was found as the most important independent variable and to have a significant effect on legibility (see Figure 7). This variable is out of the designer's control as a digital design can be used on numerous devices, but emphasises the importance of acknowledging the effect technology can have on legibility. Interestingly, this finding was identified due to the alternative approach because it was conducted with different phones, which would not have been identified if the experiment had been performed as planned.

Collecting these data through an experiment can have added distractions or "noise" that could have influenced the results. When different phones are used with different characteristics this can have affected the font size for the different participants. As mobile phones in general have high resolution this difference is not believed to have had a big effect, compared to if it had been reviewed on different devices as a phone and a computer. Doing this experiment in a lab would allow for better control as discussed in the next section and would be a better option of especially investigating difference in colour. The participants had normal vision and colour vision, and most was in a low age range. These variables can potentially affect the results and if similar research was to be done with participants with reduced vision or colour vision, or with a higher age range (60+) findings could potentially be discovered.

5.2.2 The planned methodology: a physical experiment

The experiment was planned to be conducted physically in the Universal Design lab (UU-lab) at Mustad, which is a controlled environment that produce and control ambient illumination. The luminance level and colour temperature can be controlled, and pre-programmed if wanted, in an associated software in a control room and the luminance level can also be adjusted with a switch with multiple steps. This environment can produce a controlled colour temperature up to approx. 6500 K at 3000 lux or illumination level of approx. 10.000 lux. The room is closed off from outside light and the walls can be adjusted to different levels of reflection with curtains. As the lab allows for controlling the light intensity and chromaticity, the hypothesis in section 3.2 would be more specific as the "bright outdoor condition" variable would be split and the effect measured more detailed. The hypothesis would then be if *stroke contrast influence legibility at high levels of illumination* and if *colour on text influence legibility in cold colour temperature*.

The experiment would be conducted on a mobile phone, Motorola z3 play XT 1929-8, mounted to a Table to control the viewing angle and minimize reflections. Other variables like seating position, colour of participants clothing, and the shadow participants will cast on the screen would also be accounted for. The experimental task was planned to be a threshold visibility task by identifying characters. 4 random characters would be presented in different ways as described in section 3.2.3, and the participants would be asked to identify and type them down. To achieve the same levels of luminance and cold colour temperature that are faced outside on a bright clear blue summer day a combination of real (as long as possible) and simulated change would be used on the stimuli by changing the contrast ratio and background colour.

A physical experiment would be more consistent and make the results more comparable between participants. By using a spectroradiometer the displays white and black point would be documented and the colours used could be measured and defined in CIE 1931 coordinates which would increase the experiments reliability and increasing repetability. The spectroradiometer would also be used to measure the effective illumination level and colour temperature when planning and defining the different levels of simulated and real change. This physical experiment would also allow for screening of participants before starting the experiment, rather than relying on self-reporting. With a simple vision accuracy test and the Ishihara colour vision test the participants vision would be controlled, as accurate vision was a prerequisite for participation.

5.3 Limitations

With every study there are limitations and the spring of 2020 have been special for most of the world due to the Covid-19 pandemic, which have put its limitations on this thesis. As the breakout caused campus to close and people to stay at home the planned methodology could all of a sudden not be conducted as planned. An alternative methodology was needed and the effects of this change have been discussed earlier (see section 5.2). Using an online experiment has some limitations and relies on self-reporting which can introduce misunderstandings or other biases (Courage et al. 2015), but also gave some advantages that wouldn't have been achieved in a lab.

Resources and time limited the available user group as well. The participants all had accurate or corrected vision and colour vision, but it would be beneficial to use participants with vision deficiencies to include them and the results would be more valid from a UD perspective. As Begnum (2019) discusses in her dissertation inclusion and designing for these deficiencies will only make it better for all of us. In the end, there wasn't enough time to evaluate the guidelines and they need further evaluation and testing. Testing is needed by both designers and readers to see if they limit designers to much and if they achieve sufficient legibility in bright outdoor conditions, preferably with people with vision deficiencies.

6 Conclusion

The work done in this thesis has aimed to answer how designers can ensure sufficient legibility on mobile phones in bright outdoor conditions by designing a set of guidelines based on reviewing research, guidelines and findings from the online experiment. It is apparent that an absolute answer to this isn't feasible and that this is a complex case with many influential variables. The dynamic use of mobile phones will keep expanding and this require designs that will be legible for all users in challenging and "new" use cases. By including use case when making guidelines one of the variables are controlled to a bigger extent and allows for a better understanding to make good design choices. The contribution of this thesis is a step in the right direction, also focusing on the ecological validity of guidelines. More work is however still needed, and the guidelines needs to be tested and evaluated further, potentially moving the established thresholds.

6.1 Future work

Future work on this topic should focus on validating the proposed guidelines among users and designers, including a wider user group and measure to see if the guidelines supports accessibility for users with vision deficiencies as well. After the guidelines have been reviewed furthered supporting inclusive design, they need to be distributed and implemented in the design process. Hopefully, this thesis can also be an example advocating for inclusion of use cases when making guidelines and how to do so. Further work from a technological aspect could be to focus on a solution similar to adaptive brightness on mobile phone displays, where a designer can design for different use case that are triggered by the surrounding illumination level. Another similar suggestion could be to design for an effective contrast ratio, compared to a relative, by having the phone "reading" the ambient conditions and taking the displays characteristics into account resulting in an output for that specific scenario.

Bibliography

Apple Inc (2020), 'Using a Retina Display'.
URL: *https://support.apple.com/en-us/HT202471*

Begnum, M. E. N. (2019), Facilitating and Advancing Universal Design of ICT, PhD thesis, NTNU, Gjøvik. ISBN: 823264219X Publisher: NTNU.

Beier, S. (2012), *Reading letters: designing for legibility*, BIS, Amsterdam. OCLC: ocn751739979.

Bellis, M. (2017), 'Television History and the Cathode Ray Tube'.
URL: *https://www.thoughtco.com/television-history-cathode-ray-tube-1991459*

Bellis, M. (2019*a*), 'The History of Liquid Crystal Display'.
URL: *https://www.thoughtco.com/liquid-crystal-display-history-lcd-1992078*

Bellis, M. (2019*b*), 'OLED Technology Guidelines and History'.
URL: *https://www.thoughtco.com/who-invented-oled-technology-1992208*

Benyon, D. (2013), *Designing interactive systems: a comprehensive guide to HCI and interaction design*, third edition edn, Pearson, Boston.

Bigelow, C. (2019), 'Typeface features and legibility research', *Vision Research* **165**, 162 – 172.
URL: *http://www.sciencedirect.com/science/article/pii/S0042698919301087*

Breuninger, J. (2019), 'Legibility: how to make text convenient to read'.
URL: *https://uxdesign.cc/legibility-how-to-make-text-convenient-to-read-7f96b84bd8af*

Buchner, A. & Baumgartner, N. (2007), 'Text – background polarity affects performance irrespective of ambient illumination and colour contrast', *Ergonomics* **50**(7), 1036–1063.
URL: *http://www.tandfonline.com/doi/abs/10.1080/00140130701306413*

Carter, R. (2002), *Digital color and type*, RotoVision, Crans-Près-Céligny Hove. OCLC: 833881765.

Carter, R., Meggs, P. B. & Day, B. (2002), *Typographic design: form and communication*, 3rd ed edn, John Wiley & Sons, New York.

Chang, P.-C., Chou, S.-Y. & Shieh, K.-K. (2013), 'Reading performance and visual fatigue when using electronic paper displays in long-duration reading tasks under various lighting conditions', *Displays* **34**(3), 208 – 214.
URL: *http://www.sciencedirect.com/science/article/pii/S0141938213000425*

Chen, H., Tan, G. & Wu, S.-T. (2017), 'Ambient contrast ratio of LCDs and OLED displays', *Optics Express* **25**(26), 33643.
URL: *https://www.osapublishing.org/abstract.cfm?URI=oe-25-26-33643*

Chen, M.-T. & Lin, C.-C. (2004), 'Comparison of TFT-LCD and CRT on visual recognition and subjective preference', *International Journal of Industrial Ergonomics* **34**(3), 167 – 174.
URL: *http://www.sciencedirect.com/science/article/pii/S0169814104000654*

Choi, K. & Suk, H.-J. (2014), 'User-preferred color temperature adjustment for smartphone display under varying illuminants', *Optical Engineering* **53**(6), 061708.
URL: *http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.OE.53.6.061708*

Courage, C., Baxter, K. & Caine, K. (2015), *Understanding your users: a practical guide to user research methods*, second edition edn, Elsevier, Morgan Kaufmann, Amsterdam ; Boston. OCLC: ocn918928845.

Difi (2020), 'English'.
URL: *https://uu.difi.no/om-oss/english*

DisplayMate (2019), 'iPhone 11 Pro Max OLED Display Technology Shoot-Out'.
URL: *http://www.displaymate.com/iPhone*$_11Pro_ShootOut_1P.htm$

Dobres, J., Chahine, N. & Reimer, B. (2017), 'Effects of ambient illumination, contrast polarity, and letter size on text legibility under glance-like reading', *Applied Ergonomics* **60**, 68–73.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0003687016302459*

Dobres, J., Chahine, N., Reimer, B., Gould, D., Mehler, B. & Coughlin, J. F. (2016), 'Utilising psychophysical techniques to investigate the effects of age, typeface design, size and display polarity on glance legibility', *Ergonomics* **59**(10), 1377–1391. ISBN: 0014-0139 Publisher: Taylor & Francis.

Dobres, J., Chahine, N., Reimer, B., Gould, D. & Zhao, N. (2016), 'The effects of Chinese typeface design, stroke weight, and contrast polarity on glance based legibility', *Displays* **41**, 42–49. ISBN: 0141-9382 Publisher: Elsevier.

Dobres, J., Chrysler, S. T., Wolfe, B., Chahine, N. & Reimer, B. (2017), 'Empirical Assessment of the Legibility of the Highway Gothic and Clearview Signage Fonts', *Transportation Research Record* **2624**(1), 1–8. _eprint: https://doi.org/10.3141/2624-01.
URL: *https://doi.org/10.3141/2624-01*

Dobres, J., Reimer, B. & Chahine, N. (2016), The effect of font weight and rendering system on glance-based text legibility, *in* 'Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications', pp. 91–96.

Dobres, J., Wolfe, B., Chahine, N. & Reimer, B. (2018), 'The effects of visual crowding, text size, and positional uncertainty on text legibility at a glance', *Applied ergonomics* **70**, 240–246. ISBN: 0003-6870 Publisher: Elsevier.

Dyson, M. C. (2004), 'How physical text layout affects reading from screen', *Behaviour & Information Technology* **23**(6), 377–393. Publisher: Taylor & Francis.
URL: *https://doi.org/10.1080/01449290410001715714*

Dyson, M. C. & Haselgrove, M. (2001), 'The influence of reading speed and line length on the effectiveness of reading from screen', *International Journal of Human-Computer Studies* **54**(4), 585–612.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S1071581901904586*

Equality and Anti-Discrimination Act (2020). Act relating to equality and a prohibition against discrimination.
URL: *: https://lovdata.no/dokument/NLE/lov/2017-06-16-51*

European Comission (2019), 'Web Accessibility'.
URL: *https://ec.europa.eu/digital-single-market/en/web-accessibility*

Feisner, E. A. & Reed, R. (2014), *Color studies*, third edition edn, Fairchild Books, New York.

Gray, D., Brown, S. & Macanufo, J. (2010), *Gamestorming: a playbook for innovators, rulebreakers, and changemakers*, first edition edn, O'Reilly, Beijing Cambridge Farnham Köln Sebastopol Tokyo. OCLC: 846490200.

Halbach, T. & Fuglerud, K. S. (2018), Gode skrifttyper på skjerm, Technical Report 1041, Norsk Regnesentral, Oslo.
URL: *https://www.nr.no/directdownload/1546943674/Rapport-1041-gode-skrifttyper-paa-skjerm.pdf*

Heiting, G. (2017), 'Blue light: It's both bad and good for you'.
URL: *https://www.allaboutvision.com/cvs/blue-light.htm*

Humar, I., Gradisar, M., Turk, T. & Erjavec, J. (2014), 'The impact of color combinations on the legibility of text presented on LCDs', *Applied Ergonomics* **45**(6), 1510–1517.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0003687014000696*

International Organization Standards (2008), 'ISO 9241-304:2008 Ergonomics of human-system interaction — Part 304: User performance test methods for electronic visual displays'.
URL: *https://www.iso.org/standard/40099.html*

International Organization Standards (2011), 'ISO 9241-303:2011 Ergonomics of human-system interaction — Part 303: Requirements for electronic visual displays'.
URL: *https://www.iso.org/standard/57992.html*

International Organization Standards (2019), 'ISO 24509:2019 Ergonomics — Accessible design — A method for estimating minimum legible font size for people at any age'.
URL: *https://www.iso.org/standard/69766.html*

Lee, D.-S., Ko, Y.-H., Shen, I.-H. & Chao, C.-Y. (2011), 'Effect of light source, ambient illumination, character size and interline spacing on visual performance and visual fatigue with electronic paper displays', *Displays* **32**(1), 1 – 7.
URL: *http://www.sciencedirect.com/science/article/pii/S0141938210000727*

Lee, D.-S., Shieh, K.-K., Jeng, S.-C. & Shen, I.-H. (2008), 'Effect of character size and lighting on legibility of electronic papers', *Displays* **29**(1), 10 – 17.
URL: *http://www.sciencedirect.com/science/article/pii/S0141938207000492*

Leedy, P. D. & Ormrod, J. E. (2015), *Practical research: planning and design*, 11th edition, global edition edn, Pearson, Boston Columbus Indianapolis New York.

Legge, G. E. & Bigelow, C. A. (2011), 'Does print size matter for reading? A review of findings from vision science and typography', *Journal of Vision* **11**(5), 8–8.
URL: *http://jov.arvojournals.org/Article.aspx?doi=10.1167/11.5.8*

Lin, C.-C. (2005), 'Effects of screen luminance combination and text color on visual performance with TFT-LCD', *International Journal of Industrial Ergonomics* **35**(3), 229–235.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0169814104001593*

Lin, C.-C. & Huang, K.-C. (2006), 'Effects of ambient illumination and screen luminance combination on character identification performance of desktop TFT-LCD monitors', *International Journal of Industrial Ergonomics* **36**(3), 211 – 218.
URL: *http://www.sciencedirect.com/science/article/pii/S0169814105001691*

Lin, C.-C. & Huang, K.-C. (2013), 'Effects of ambient illumination conditions and background color on visual performance with TFT-LCD screens', *Displays* **34**(4), 276–282.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0141938213000723*

Lin, H., Wu, F.-G. & Cheng, Y.-Y. (2013), 'Legibility and visual fatigue affected by text direction, screen size and character size on color LCD e-reader', *Displays* **34**(1), 49 – 58.
URL: *http://www.sciencedirect.com/science/article/pii/S0141938212001126*

Lund, O. (1999), Knowledge construction in typography : the case of legibility research and the legibility of sans serif typefaces, Ph.D., University of Reading, Reading.

McCourt, D. (2020), 'Smartphone screens explained: display types, resolutions and refresh rates'.
URL: *https://www.androidpit.com/smartphone-displays-explained*

Nersveen, J. & Blindeforbund, N. (2009), 'kontraster–farger–belysning', *Norges Blindeforbund* pp. 1–19.

Nersveen, J. & Johansen, E. A. (2016), Lesbarhet av trykt tekst for svaksynte, Technical report, Norsk forskningslaboratorium for universell utforming, NTNU.
URL: *https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2427663*

Nersveen, J., Kvitle, A. K. & Johansen, E. A. (2018), 'Legibility in Print Text for People with Impaired Vision', *Studies in Health Technology and Informatics* pp. 862–869.
URL: *https://www.medra.org/servlet/aliasResolver?alias=iospressISBNisbn=978-1-61499-922-5spage=862doi=10.3233/978-1-61499-923-2-862*

Nielsen, J. (2015), 'Legibility, Readability, and Comprehension: Making Users Read Your Words'.
URL: *https://www.nngroup.com/articles/legibility-readability-comprehension/*

NIH (2019*a*), 'Color Blindness'.
URL: *https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness*

NIH (2019*b*), 'Types of Color Blindness'.
URL: *https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-blindness*

Norman, D. A. (2013), *The design of everyday things*, revised and expanded edition edn, Basic Books, New York, New York.

Nyre, L. & Liestøl, G. (2018), 'Folkeopplysning på stranda med mobilen i handa', *Norsk medietidsskrift* **25**(04), 1–10.
URL: *https://www.idunn.no/nmt/2018/04/folkeopplysning*$_paa_stranda_med_mobilen_{ih}anda$

OpenBridge Design System (2020), 'Next generation maritime workplaces'.
URL: *http://openbridge.no/index.html*

Perondi, L., Gerbino, W., Chia, B., Arista, R., Pignoni, G. & Gaudenzi, G. (2017), 'Tipografia parametrica e Developmental Dyslexia', *MD Journal* p. 88.
URL: *http://materialdesign.it/media/formato2/allegati*$_6230.pdf$

Pignoni, G. (2018*a*), Effects of ambient illumination on text recognition for UI development, *in* '2018 Colour and Visual Computing Symposium (CVCS)', IEEE, Gjøvik, pp. 1–6.
URL: *https://ieeexplore.ieee.org/document/8496507/*

Pignoni, G. (2018*b*), 'LCD illuminant contrast analyser'.
URL: *https://pignonig.github.io/ambient*$_contrast_analyser/$

Postrel, V. I. (2003), *The substance of style: how the rise of aesthetic value is remaking commerce, culture, and consciousness / Virginia Postrel*, 1st ed edn, HarperCollins, New York, N.Y.

Rannem, (2012), *Typografi og skrift*, Abstrakt, Oslo. OCLC: 1028449934.

Reimer, B., Mehler, B., Dobres, J., Coughlin, J. F., Matteson, S., Gould, D., Chahine, N. & Levantovsky, V. (2014), 'Assessing the impact of typeface design in a text-rich automotive user interface', *Ergonomics* **57**(11), 1643–1658. ISBN: 0014-0139 Publisher: Taylor & Francis.

Sandnes, F. E. (2017), On-screen colour contrast for visually impaired readers, *in* 'Information Design: Research and Practice', Routledge, pp. 405–416.

Sharma, A. (2004), *Understanding color management*, Thomson/Delmar Learning, Clifton Park, NY.

Shieh, K.-K. & Lin, C.-C. (2000), 'Effects of screen type, ambient illumination, and color combination on VDT visual performance and subjective preference', *International Journal of Industrial Ergonomics* **26**(5), 527 – 536.
URL: *http://www.sciencedirect.com/science/article/pii/S0169814100000251*

Solås, I. M. (2019), 'The effects outdoor conditions have on the perceived legibility'.

Solås, I. M. (2020), 'Typographical Variables and Ways to Measure Them'.

W3C (2018*a*), 'Understanding Success Criterion 1.4.3: Contrast (Minimum)'.
URL: *https://www.w3.org/WAI/WCAG21/Understanding/contrast-minimum.html*

W3C (2018*b*), 'Web Content Accessibility Guidelines (WCAG) 2.1.'.
URL: *https://www.w3.org/TR/WCAG21/*

W3C (2019), 'How to Meet WCAG (Quick Reference)'.
URL: *https://www.w3.org/WAI/WCAG21/quickref/*

WHO (2019), 'Blindness and vision impairment'.
URL: *https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment*

Wolffsohn, J. S., Cochrane, A. L., Khoo, H., Yoshimitsu, Y. & Wu, S. (2000), 'Contrast is enhanced by yellow lenses because of selective reduction of short-wavelength light', *Optometry and vision science* **77**(2), 73–81. ISBN: 1538-9235 Publisher: LWW.

A Appendices

A.1 NSD privacy notice

A.2 Survey

A.3 Stimuli presented on an iPhone 6/7/8

A.1 NSD privacy notice

Result of Notification Test: Not Subject to Notification

You have indicated that neither directly or indirectly identifiable personal data will be registered in the project.

If no personal data is to be registered, the project will not be subject to notification, and you will not have to submit a notification form.

Please note that this is a guidance based on information that you have given in the notification test and not a formal confirmation.

For your information: *In order for a project not to be subject to notification, we presuppose that all information processed using electronic equipment in the project remains anonymous.*

Anonymous information is defined as information that cannot identify individuals in the data set in any of the following ways:

- *directly, through uniquely identifiable characteristic (such as name, social security number, email address, etc.)*
- *indirectly, through a combination of background variables (such as residence/institution, gender, age, etc.)*
- *through a list of names referring to an encryption formula or code, or*
- *through recognizable faces on photographs or video recordings.*

Furthermore, we presuppose that names/consent forms are not linked to sensitive personal data.

Kind regards,
NSD Data Protection

NSD – Norsk senter for forskningsdata AS Harald Hårfagres gate 29 Tel: +47-55 58 21 17 nsd@nsd.no Org.nr. 985 321 884
NSD – Norwegian Centre for Research Data NO-5007 Bergen, NORWAY Faks: +47-55 58 96 50 www.nsd.no

A.2 Survey

Legibility on Mobile Phones in Bright Outdoor Conditions

By completing this survey, you are consenting to take part in this project to understand more about legibility on mobile phones in bright outdoor conditions. Please take some time and read each part of this consent form so you understand what will be expected of you.

Purpose of the project
This project is part of a master's thesis and the overall aim is to create design guidelines to ensure good legibility when mobile phones are used outdoors, as the bright light and colour temperature outdoors can influence the legibility. The objective of this project is to see how legible different samples with different colours and stroke contrasts (difference between thin and thick lines in a font) are when read outdoors. The results will be used as a basis for recommendations and be included in the design guidelines.

What does participation involve for you?
If you choose to participate in this project, this will involve you will fill out this online survey while being outdoors in bright daylight. The survey will mainly ask you to rate multiple samples of text from 1–4 (not readable – easy to read) and some questions about you, your vision and the phone you are taking the survey on. The survey take approximately 25 minutes to complete. Your answers will be recorded electronically.

Participation is voluntary
Participation in the project is voluntary and the collected data are anonymous. If you chose to participate, you can stop the survey at any point by simply leaving the page. No information will be stored unless you submit your answers at the end of the survey.

Confidentiality and Privacy
None of the information that you will provide will be traced back to you. All information will remain confidential and anonymous. The results of this survey will only be used for the purpose of this research. The project is scheduled to end on 31.07.20 and all data will be kept anonymous before and after this date. The only people who will have access to the data collected will be the student conducting the research and two advisors for the project.

Your rights
No personal information that can be traced back to you personally will be collected in this online survey. The information asked regarding you will be your age range, gender, what phone you are taking the survey on and if you have accurate vision and colour vision.

How to find out more:
If you have any questions or concerns, please contact Ida Marie Solås at idamsol@stud.ntnu.no a master student at NTNU in Gjøvik, or Frode Volden at frodev@ntnu.no an supervisor for the project at NTNU in Gjøvik.

This survey take approximately 25 minutes to complete and needs to be conducted outdoors in bright daylight. The survey consists of 8 check points for your environment and how you will proceed, 7 questions and following text samples that you will rate from "not readable" to "easy to read".

I consent to participate in this study

- [ ] Yes

Before you begin, please make sure you follow this:

1. Adaptive display functions are turned off on your phone (auto brightness, night mode and adaptive white balance)
    - Guide to turn off on Apple phones
    - On Android phones open "settings", tap "display" and you should see the functions
2. Use glasses or lenses if you use them on a regular basis
3. Don't use sunglasses
4. The brightness on your display is at maximum
5. You are outdoors in bright daylight and clear blue sky
6. Minimize the reflection in your screen, by keeping some distance to your surroundings
7. Don't stand in the shadow
8. Keep your phone at an arm's length distance during the questionnaire
9. Complete the survey using Chrome or Safari as browser

What date and approximately time is it?

Date / Time

Date DD/MM/YYYY

Time hh mm

AM/PM -

What is your gender?

- Male
- Female
- Other
- Prefer not to say

What is your age?

- Under 18
- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+

Do you have accurate vision? (With or without correction aids like glasses or contact lenses)

- Yes
- No
- Not sure

Do you have accurate colour vision?

- Yes
- No
- Not sure

What type of phone are you using?
Please state make and model (example: iPhone 8 Plus)

Are you outdoors?

- Yes
- No

How is the weather?
You are strongly encouraged to finish the questionnaire outdoors with strong sunlight and clear blue sky

- Bright sunlight and clear blue sky
- Bright sunlight with some clouds
- Overcast in daylight
- Other (please specify)

The remaining survey will consist of text samples that you will rate based on how they are to read.

Below is an example before you begin. Only spend a few seconds to decide on each sample. If you don't see the text sample below, please open the survey in Chrome.

newspaper painting nose apple moth
beef reading sibling exercise weathe

- Easy to read
- Readable, with some discomfort
- Readable, but hard
- Not readable

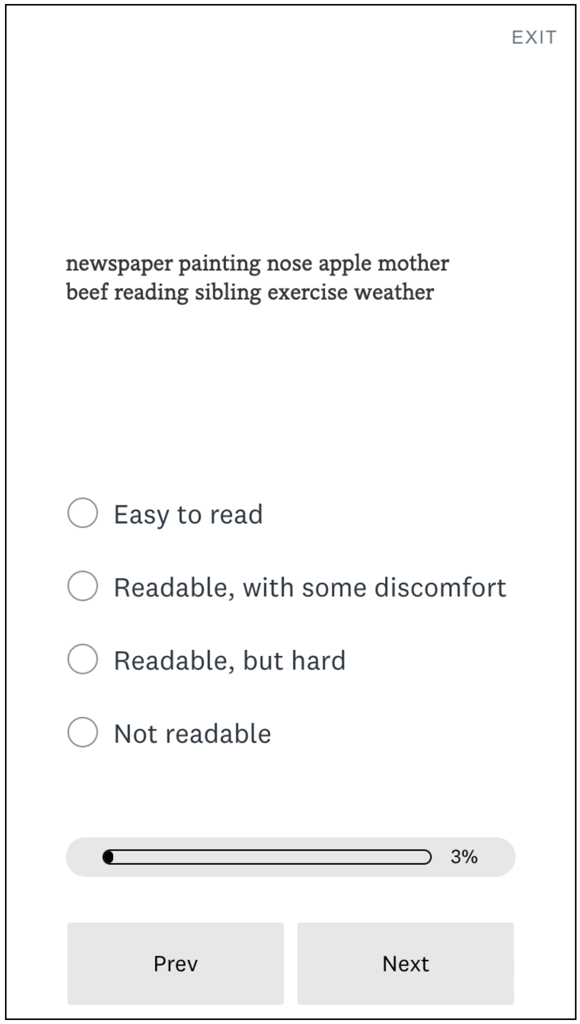
newspaper painting nose apple moth
beef reading sibling exercise weathe

- Easy to read
- Readable, with some discomfort
- Readable, but hard
- Not readable

newspaper painting nose apple moth
beef reading sibling exercise weather

- ◯ **Easy to read**
- ◯ **Readable, with some discomfort**
- ◯ **Readable, but hard**
- ◯ **Not readable**

A.3 Stimuli presented on an iPhone 6/7/8



NTNU

Kunnskap for en bedre verden