

Rapportnr.

Forfattere

Dato

Måling av kontraster på skjerm

Teknologi- og metodeanbefaling



DART/14/2017

Till Halbach, Kristin Skeide Fuglerud

2017-12-30

Forfatterne

Till Halbach er seniorforsker ved Norsk Regnesentral og del av E-inkluderingsgruppen.

Kristin Skeide Fuglerud er sjefsforsker ved Norsk Regnesentral og leder av E-inkluderingsgruppen som arbeider med digital inkludering, universell utforming og brukeropplevelse.

Norsk Regnesentral

Norsk Regnesentral (NR) er en privat, uavhengig stiftelse som utfører oppdragsforskning for bedrifter og det offentlige i det norske og internasjonale markedet.

NR ble etablert i 1952 og har kontorer i Kristen Nygaards hus ved Universitetet i Oslo. NR er et av Europas største miljøer innen anvendt statistisk-matematisk modellering og har et senter for forskningsdrevet innovasjon, Big Insight, med finansiering fra Norges forskningsråd, bedrifter og offentlige partnere. Innen statistikk jobbes det med et bredt spekter av problemstillinger, for eksempel finansiell risiko, jordobservasjon, estimering av fiskebestander, helse og beskrivelse av geologien i petroleumsreservoarer. NR er ledende i Norge innen utvalgte deler av informasjons- og kommunikasjonsteknologi. Innen IKT-området har NR innsatsområdene e-inkludering, informasjonssikkerhet og smarte informasjonssystemer.

NRs visjon er forskningsresultater som brukes og synes.

Tittel **Måling av kontraster på skjerm****Forfattere** **Till Halbach, Kristin Skeide Fuglerud****Dato** 2017-12-30**Publikasjonsnummer** DART/14/2017**Gradering** Offentlig**Sammendrag**

Denne rapporten diskuterer ulike teknologier og metoder for å måle skjermkontraster på en pålitelig og effektiv måte. Dokumentet er utarbeidet ved å se på relevante standarder og relatert forskning, og ved å hente inn uttalelser av eksperter, samt brukerinnspill.

Vår anbefaling er at en kontrastmåling bør være tredelt og bestå av to objektive og en subjektiv verdi. De to objektive verdiene er forhold av høy og lav luminans, målt med luminansmeter en gang i et mørkt rom og en gang under diffus belysning. I tillegg anbefaler vi å supplere labmålingene med en brukertesting. Denne fremgangsmåten skal fange opp de tre hovedfaktorene uavhengige skjermegenskaper, skjermens robusthet, det vi si hvor bra den klarer å beholde den maksimale kontrasten under ekstern lyssetting, og hvordan kontrasten faktisk oppleves sett med menneskelige øyne.

Emneord	Kontrast, skjerm, objektiv, subjektiv,
Målgrupper	Interessegrupper, forskere, innkjøpere, tilsynsmyndigheter
Gradering	Offentlig
Prosjektnummer	DART/14/2017
Satsningsfelt	E-inkludering
Antall sider	25
© Copyright	Norsk Regnesentral

Innholdsfortegnelse

Introduksjon.....	7
Målet med rapporten.....	8
Utførende og metode.....	8
Takk.....	8
Involvering av sosiale medier.....	9
Begrepsavklaring.....	9
Relevante standarder.....	9
CEN/TS 15291:2006.....	10
Vurdering.....	10
NS-EN ISO 9241-20:2009.....	10
Vurdering.....	10
ISO/TR 22411:2008.....	11
Vurdering.....	11
NS-EN 12414:2000.....	11
Vurdering.....	12
NS-EN ISO 9241-303:2011.....	12
Vurdering.....	12
NS-EN ISO 9241-305:2008.....	13
Vurdering.....	13
NS-EN ISO 9241-306:2008.....	13
Vurdering.....	14
NS-EN ISO 9241-307:2008.....	14
Vurdering.....	15
W3C WCAG 2.....	15
Vurdering.....	15
W3C CSS 2.1.....	15
Vurdering.....	15
ITU-T P.800.1 og ITU-T P.800.2.....	16
Vurdering.....	16
ITU-T P.910.....	16
Vurdering.....	17
ITU-R BT.500-13.....	17
Vurdering.....	17
Relevant forskning.....	18
Rapport: Lesbarhet av trykt tekst for svaksynte.....	18
Vurdering.....	18
Rapport: Undersøkelse av synsforhold ved billettautomater og informasjonstavler.....	18
Vurdering.....	19
Annen forskning om polaritet og subjektive vurderinger.....	19
Vurdering.....	19

<u>Ekspertanbefalinger.....</u>	<u>19</u>
<u>Vurdering.....</u>	<u>21</u>
<u>Brukerinnspill.....</u>	<u>21</u>
<u>Vurdering.....</u>	<u>21</u>
<u>Anbefaling og begrunnelser.....</u>	<u>22</u>
<u>Anbefaling for utregning av minimumskontrast, $L_{h,min}$ og terskelverdier.....</u>	<u>24</u>
<u>Konklusjon.....</u>	<u>24</u>

Figuroversikt

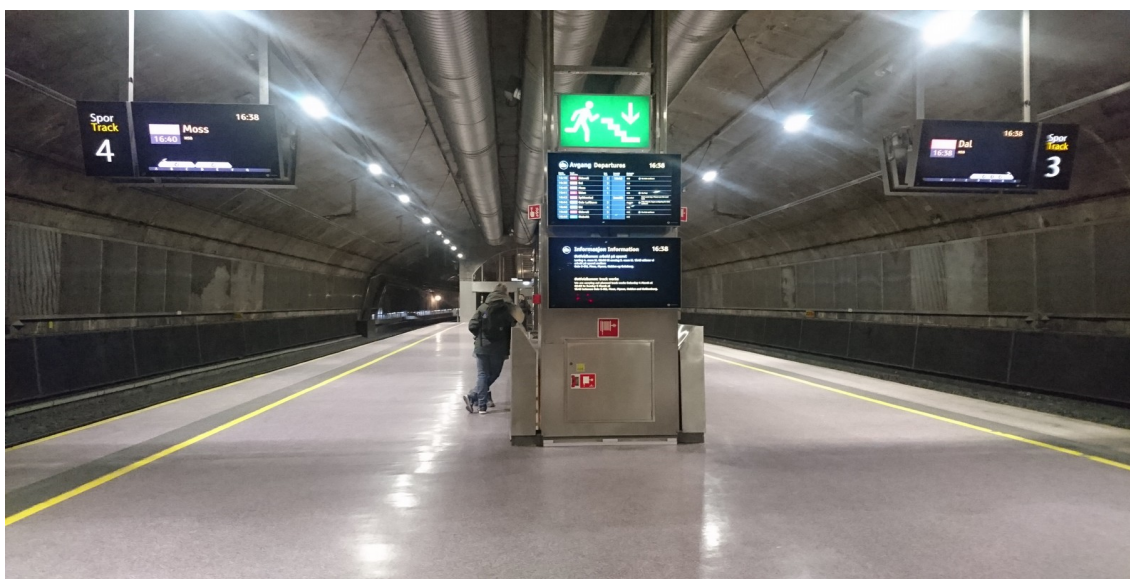
Figur 1: Illustrasjonsbilde for ulike informasjonstavler under kunstig lyssetting på Nasjonalteater stasjon i Oslo.....	7
---	---

1 Introduksjon

Meningen med elektroniske skjermer er å formidle informasjon på en visuell måte og legge til rette for dialog/interaksjon med brukeren.

Det å ha tilstrekkelige kontraster på skjerm er derfor ofte avgjørende for hvordan brukeropplevelsen blir. Dette er spesielt viktig for personer med nedsatt syn, som er antatt å utgjøre ca. 6% av den norske befolkningen¹, men det er mange med vanlig syn som også har utfordringer med å oppfatte det som står på skjermen i tilfelle dårlige kontraster eller vanskelig lyssetting. I de aller fleste tilfeller består informasjonen av tekst. Kontrastmålinger er derfor også knyttet til lesbarhet av tekst på skjerm.

Problematikken gjelder i utgangspunkt alle skjermer/display² med egen lyskilde (selvlysende), både analoge og digitale, og ordinære og berøringsskjermer. Eksempler er billett- og parkeringsautomater, ofte omtalt med paraplybetegnelsen selvbetjeningsautomater, samt informasjonstavler, det vil si skjermer som ikke krever brukerinteraksjon. Forekomsten av ulike typer skjermer har økt betraktelig i den senere tid, forårsaket av progresjonen av digitaliseringen av produkter, tjenester og tekniske systemer, og ledsaget av inntoget av flere og flere mobile enheter. Det forventes at skjermforekomster vil øke ytterligere fremover. Derfor er det behov for gode skjermkontraster ikke bare på datamaskiner, men også nettbrett, telefoner og for eksempel smartklokker. Og, ettersom mange skjermer blir brukt ute i det fri, er gode kontraster spesielt viktig i dagslys og med solinnstråling.



Figur 1: Illustrasjonsbilde for ulike informasjonstavler under kunstig lyssetting på Nasjonalteater stasjon i Oslo

Det må videre nevnes at krav om gode kontraster er lovpålagt som en del av forskriften om universell utforming av IKT-løsninger³. Denne rapporten er derfor relevant for ulike interesseforeninger, innkjøpere, tilsynsorganer og utviklende bedrifter, med flere.

1 SSB, Levekårsundersøkelse om helse, 2012

2 På engelsk brukes forkortelsen VDT, for Visual Display Terminal

3 Kommunal- og moderniseringsdep., Forskrift om universell utforming av informasjons- og kommunikasjonsteknologiske løsninger, 2013

Det er derimot ikke trivielt å måle skjermkontraster på en riktig måte og kvalitetssikre at en skjerm har tilstrekkelig eller god kontrast. For hva er «tilstrekkelig» eller «god»? Og for hvem? Andre faktorer som også spiller inn er skjermens innebygde lysstyrke og -kontrast, fargegjengivelse, samt skjermens refleksjonsegenskaper. Man bør, som utdypet tidligere, også ta hensyn til hvordan skjermen er lyssatt (kunstig og/eller gjennom sollys), og om lyssettingen varierer med tiden. Det er viktig å presisere at det er den endelige/effektive (netto-)kontrasten som er avgjørende; det hjelper altså svært lite å etterleve f.eks. WCAG⁴-kravene for en webside (som legger føringer for bruttokontrasten) om skjermens lysstyrke er altfor mørkt innstilt, eller om solens refleksjoner på skjermen gjør det umulig å se noe.

2 Målet med rapporten

Målet med denne rapporten er å diskutere ulike teknologier, verktøy og metoder, og å gi en konkret, begrunnet anbefaling for den mest pålitelige og effektive måten å måle skjermkontraster på.

3 Utførende og metode

Det er Norsk Regnesentral (NR) som har utarbeidet rapporten, med seniorforsker Till Halbach som ansvarlig prosjektleder. Sjefsforsker Kristin Skeide Fuglerud har vært fagressurs.

Det ble intervjuet en rekke fagpersoner i forbindelse med denne rapporten. Disse er:

- Bjørn Nygård, tidligere rådgiver ved Hjelpemiddelsentralen i Sør-Trøndelag, avdeling Syn
- Arne Tømte, synspedagog i Blindforbundet
- Ruedi à Porta, tidligere grafisk designer og industridesigner
- Jonny Nersveen, førsteamanuensis ved Institutt for vareproduksjon og byggteknikk (og Norsk forskningslaboratorium for universell utforming), NTNU,
- Ole Lund, førsteamanuensis ved Institutt for design, NTNU
- Magne Helland, dosent ved Institutt for optometri, radiografi og lysdesign (og Nasjonalt senter for optikk syn og øyehelse), HSN

Tilnærmingen til problemstillingen er i hovedsak basert på en litteraturstudie og intervjuer av og drøftinger med eksperter vedrørende synsfaglige aspekter. Intervjuene ble gjennomført på telefon og/eller e-post. I konklusjonen inngår vurderingene av både ekspertuttalelsene, forskningsresultater, standarder og annen litteratur, samt innspill fra en kampanje på sosiale medier, se nede.

Rapporten fikk tilskuddsmidler fra Norges Blindforbunds forskningsfond.

3.1 Takk

Forfatterne retter en stor takk til alle ekspertene for gode uttalelser og nyttige innspill. Rigmor Baraas ved Nasjonalt senter for optikk syn og øyehelse formidlet kontakt med Magne Helland. Weiqin Chen ved Forskningsgruppen for universell utforming, HiOA, sendte forespørselen videre til The American Federation for the Blind.

Takk også til Mia Jacobsen og Tonje Bjerkemoen i Blindforbundet for hjelp med å spre budskapet på sosiale medier, samt alle som lot seg engasjere i saken og kom med kommentarer, eksempler og bilder.

4 WC3, Web Content Accessibility Guidelines, 2008

3.2 Involvering av sosiale medier

For å belyse problemstillingen fra brukersiden ble det besluttet å kjøre en kampanje på Facebook bestående av to innlegg^{5,6} med noen få eksempler og oppfordring om å sende inn egne opplevelser og eksempler. Dette ble gjort på Facebook-gruppen uleselig.no, som styres av Blindeforbundet.

Gruppen har i skrivende stund i overkant av 21 000 følgere. Blant disse kan det antas er det mange personer med nedsatt syn. Aktivitetsnivå på kanalen varierer en del: I 2017 ble 15 saker lagt ut på Uleselig, med alt mellom under 10 og over 100 reaksjoner i form av *likes* og lignende, med en gjennomsnitt på rundt 30. For å øke engasjement i saken ble det lovet ut en premie i andre innlegg for beste innsendte eksempelet.

4 Begrepsavklaring

Med *kontrast* menes her forskjellen mellom to lyssignaler. En eksempelverdi er 1500:1, også skrevet 1500.

Luminans er betegnelsen på styrken (amplituden) av et lyssignal. Noe forenklet sier luminans altså noe om hvor sterkt et lysinntrykk er.

Bruttokontrasten er kontrasten som gjenspeiler intensjonen av informasjonstilbyder, f.eks. en webredaktør som spesifiserer fremstillingen av en nettside i en grafisk nettleser.

Bruttokontrasten vil alltid være lik på tvers av hardware uansett skjerminnstilling, ekstern lyssetting og refleksjoner.

Begrepet *nettokontrast*⁷ er brukt her for kontrasten som kan oppleves/måles i situasjoner der lys fra skjerm, der også skjermens egenskaper og innstillinger inngår, blandes med lys fra andre lyskilder, inkludert dagslys. Ettersom dette gjelder i de fleste anvendelsesområdene, vil ordene kontrast og nettokontrast her brukes om hverandre med samme betydning. Som en ytterligere avgrensning menes med kontrast her bokstavkontrast, eller forholdet mellom lysstyrke av linjene i bokstavene og bakgrunnen. Dette fordi informasjonsformidling i de fleste tilfeller vil dreie seg om tekst.

5 Relevante standarder

Skjermer er i de fleste tilfeller en vesentlig bestanddel av selvbetjeningsautomater. Det er derfor naturlig å se på Difis liste av tekniske dokumenter for universell utforming av automater⁸ for hva som er relevant for måling av skjermkontraster. Disse dokumentene er diskutert lenger nede.

- NS-EN 1332-1:2009,
- NS-EN 1332-2:1998,
- NS-EN 1332-3:2008,
- NS-EN 1332-4:2007,
- NS-EN 1332-5:2006,
- ISO 20282-1:2006 og

5 <https://www.facebook.com/uleselig.no/photos/a.1422603374634662.1073741828.1415372415357758/2073574236204236/?type=3&theater>

6 <https://www.facebook.com/uleselig.no/photos/a.1422603374634662.1073741828.1415372415357758/2085521315009528/?type=3&theater>

7 På engelsk: perceptual contrast

8 <https://uu.difi.no/krav-og-regelverk/automater/krav-til-automater>

- ISO/TS 20282-2:2013

ble vurdert til ikke å være relevant i denne sammenhengen. En annen standard, NS-EN 12414:2000, er ikke med på Difis liste, men allikevel i høyeste grad relevant for dette prosjektet av grunner beskrevet lenger nede.

I tillegg ble det vurdert standarder fra ISOs 9241-300-serie. I denne serien er NS-EN ISO 9241-300:2008 og NS-EN ISO 9241-302:2008 kun indirekte relevant her; den førstnevnte introduserer 300-serien, og 302-dokumentet forklarer terminologi og definerer viktige begrep brukt i de andre standardene.

5.1 CEN/TS 15291:2006

Teknisk spesifisering CEN/TS 15291:2006 med tittelen «Identitetskortsystemer – Retningslinjer for utforming for tilgjengelig kort-aktivert utstyr»⁹ gjelder utforming av automater som av ulike grunner håndterer maskinlesbare kort, f.eks. bankterminaler.

Hva gjelder skjermkontraster skal automaten være lysatt i tilfelle mangel på dagslys (avsnitt 5.4.1). Videre skal skjermen ikke være vendt mot direkte innstråling fra solen. Refleksjoner fra skjermen må være ubetydelige (5.4.2). Selve skjermen skal ha tilstrekkelig lysstyrke og kontrast, være i stand til å gjengi passende farger og må ikke vise flakringseffekter. Informasjon på skjermen må kunne leses med opptil 40 graders innfallsvinkel (6.5.2). Spesifikasjonen nevner videre at lesbarhet avhenger av en del andre faktorer, sånn som tekststørrelse, skrifttype, og andre typografiske parametre (6.3.2). Hva angår fargekombinasjoner er lys skrift på mørk bakgrunn å foretrekke framfor mørk skrift på lys bakgrunn (D.2.1).

5.1.1 Vurdering

Det legges til grunn at CEN/TS 15291:2006 kun er en teknisk spesifisering. Som sådan stiller anbefalingene noe svakere enn for eksempel de som er gitt i en teknisk standard eller norm. Selve anbefalingene bærer også preg av dette. Det er brukt en rekke ulne begrep som kan tolkes på ulike måter: «bra synlighet», «tilstrekkelig kontrast» og lignende, uten at det blir definert nærmere hvordan dette kan omsettes i praksis. Dokumentet er derfor nyttig for å vurdere hvilke faktorer som påvirker måling av kontraster, men kan ikke brukes for å si noe om hvordan dette kan løses.

5.2 NS-EN ISO 9241-20:2009

NS-EN ISO 9241-20:2009 har tittelen «Ergonomi for samhandling mellom menneske og system - Del 20: Veiledning om tilgjengelighet for informasjons-/og kommunikasjonsteknologiske (IKT) innretninger og tjenester»¹⁰.

Det er ganske begrenset hva som er relevant for måling av skjermkontraster i denne standarden: Det skal være mulig å endre kontrasten av informasjon på skjermen (avsnitt 7.2.7) og størrelsen på det som blir vist (7.2.8/7.2.9). Ellers blir det vist til allerede nevnte ISO 9241-300 og -302.

5.2.1 Vurdering

For en standard å være er ISO 9241-20 overraskende vag med formuleringer som «tilstrekkelig» lysstyrke og «tilstrekkelig» kontrast. Det er derimot meningen at andre, mer detaljerte standarder benyttes der disse er tilgjengelige. Standarden kan derfor uten problemer ignoreres i denne sammenhengen her.

⁹ Identification Card Systems - Guidance on design for accessible card-activated devices

¹⁰ Ergonomics of human-system interaction - Part 20: Accessibility guidelines for information/communication technology (ICT) equipment and services

5.3 ISO/TR 22411:2008

Den internasjonale tekniske anbefalingen ISO/TR 22411:2008 med tittelen «Ergonomics data and guidelines for the application of ISO/IEC Guide 71 to products and services to address the needs of older persons and persons with disabilities» er ikke direkte relevant for utforming av produkter, tjenester og omgivelser. Formålet med anbefalingen er å være bindeledd mellom en løsning og ISO/IEC Guide 71¹¹ og veilede med tanke på ergonomiske saksforhold.

En rekke faktorer er listet opp som påvirker god lesbarhet av det som vises på skjermen: Lyssetting; Automatenes utforming bør ta hensyn til de mest sannsynlige lysforholdene i felten (avsnitt 8.4.2). Ettersom kunstig belysning for automater er anbefalt vurdert, kan det i nesten alle tilfeller gåes ut fra at lys fra andre lyskilder enn skjermen vil påvirke lesbarheten (8.4.2). Refleksjoner; Polerte overflater og overflater med høy refleksjonskoeffisient kan føre til blinding, dette gjelder spesielt for selve skjermen (8.4.4). Skjermens tekniske egenskaper; lysstyrke og kontrast spiller selvsagt en viktig rolle, og her spesielt luminanskontrast, som må være tilstrekkelig på den ene siden, men som samtidig ikke må blende brukeren med en for høy kontrast heller (8.5.4). Skjermens fysiske egenskaper; Skjermstørrelse og avstand til betrakteren påvirker brukskvaliteten vesentlig (8.2.3.1, B.2). Egenskapene til synssansen; Kontrasten oppleves forskjellig avhengig av fargene som er brukt og fargenes intensitet (8.5.2). Betrakterens heterogenitet; Synsopplevelsen påvirkes også av betrakterens alder og eventuelle synsnedsettelse (8.5.1). Dessuten kan en persons eventuelle funksjonsnedsettelse også påvirke synsopplevelsen. Anbefalingen nevner her at individer med nedsatt syn som oftest foretrekker negativ polaritet (lyst på mørkt) framfor positiv polaritet (mørkt på lyst) (8.6.1). Tekst og symboler; Størrelse av bokstaver og ikoner samt skrifttype påvirker lesbarheten (8.6). Skjermbildet; Måten innhold er fordelt på utover skjermen, for eksempel såkalt romslig frekvens påvirker også lesbarheten og helhetsinntrykket (9.2.1.7). I Annex B.2 spesifiserer TR 22411 en metode for å regne ut luminanskontrasten som en funksjon av spektral stråling av to lys signaler, lysets bølgelengde og betrakterens alder.

5.3.1 Vurdering

Det trekker ned at dokumentet kun er en teknisk anbefaling, og at den som sådan er litt vag i formuleringene. Det som gjør den verdifull i denne konteksten er at den nevner «utallige» faktorer som påvirker betrakterens opplevelse, og at dokumentets hovedtyngde ligger på å synliggjøre behovene til spesielt eldre og mennesker med redusert synsfunksjon. Anbefalingene er derimot for lite konkret til å hente inspirasjon til måling av skjermkontraster, bortsett fra metoden nevnt i Annex B.2.

5.4 NS-EN 12414:2000

Den europeiske standarden EN 12414:1999 er også norsk standard med betegnelsen NS-EN 12414:2000 og gjelder utforming av parkeringsautomater. Denne standarden er ikke på Difis liste om relevante standarder for universelt utformede automater, men er tatt med her fordi arbeidsgruppen CEN/TC 226/ WG 9¹² i skrivende stund arbeider med en revidering av standarden, hvis nyvinning også vil ta opp temaet måling av skjermkontraster.

EN 12414:1999 nevner lesbarhet av all vesentlig informasjon, som skal være lesbar både på dagtid og nattetid, med vanlig lesehjelp¹³ og uten, på en distanse av maksimalt 1 m (avsnitt 4.19.1).

11 http://www.iso.org/iso/catalogue_detail?csnumber=57385

12 European Committee for Standardization/Technical Committee 226/Working Group 9

13 Det antas å referere til brille, kontaktlinser, o.l.

5.4.1 Vurdering

Til en standard å være, er NS-EN 12414:2000 skuffende tynn og vag. Hva gjelder måling av skjermkontraster, er det lite som er relevant. Det er derimot knyttet større forventninger til 2017-revisjonen av standarden, hvis publisering skal være like rundt hjørnet.

5.5 NS-EN ISO 9241-303:2011

Denne spesifikasjonen er både norsk og internasjonal standard med tittelen «Ergonomi for samhandling mellom menneske og system - Del 303: Krav til elektronisk visuelle skjermer»¹⁴. For å evaluere overholdelse av standarden må den derimot tolkes sammen med Del 307 (se lenger nede).

I følge ISO 9241-303 er det en rekke flere parametre enn luminans som utgjør brukerens skjermopplevelse, som omstendigheter, fysiske omgivelser, lesbarhet, m.fl. (avsnitt 4, Annex B). Anbefalt avstand mellom bruker og skjerm er mellom 400 og 750 mm i de fleste anvendelsesområdene, men i noen tilfeller kan det være nødvendig med større avstand (5.1.2). Standarden krever at skjerm luminans overstiger en minimumsluminansverdi for ulike avstander, typisk lysinnstråling fra omgivelsene, og for ulike aldre til brukerne (5.2.3, Annex D). Sistnevnte fordi kravene til kontrast blir høyere med brukerens økende alder. Skjermens egenskaper kan også endre seg med alderen (5.2.3), så det kan være lurt å ta hensyn til dette med en ekstra margin for minimumskontrast for helt nye skjermer. Skjermstørrelsen spiller også inn (Annex D). På nattetid bør skjerm luminansen derimot være dempet for å unngå blanding og hjelpe til at betrakterens øyne lett kan tilpasse seg. Luminanskontrast mellom skjerm og dens nærmere omgivelse bør ikke overstige en størrelsesorden av skjerm luminansen (5.2.4). Skjermens egenskaper kan variere noe på ulike punkter på skjermflaten, derfor anbefaler standarden å gjennomføre opptil tre luminans- og fargemålinger (5.4.2, 5.4.3). Som alternativ målemetode er det mulig med brukertesting, der man måler måloppnåelse, effektivitet og brukerfornøydhet (Annex C). I Annex D er det foreslått flere metoder for utregning av minimumskontrast. Det er nevnt at ISO 9241-3 anbefaler et forhold på minst 3:1 mellom høy og lav luminans, mens ISO 13406-2 har definisjonen $1 + 10 \cdot L_l^{-0.55}$, der L_l er den lave luminansen. ISO 9241-303 prøver så å bygge bro over motsetningene ved å foreslå en aldersavhengig utregning, og en utregning der minimumskontrast ikke konvergerer mot 1 med stigende L_l . Definisjonen på minimumskontrast blir da $k_a \cdot (2,2 + 4,84 \cdot L_l^{-0.65})$, der k_a er en aldersavhengig korrekturfaktor, se tabell D.1. Standarden utviser derimot forsiktighet med å anvende denne korrekturfaktoren, det vil si at den nevnes, men den er tydeligvis ikke i bruk¹⁵.

5.5.1 Vurdering

ISO 9241-303 tar opp mange nyttige aspekter og kan vise til detaljert bakgrunnsinformasjon og relevant arbeid hva gjelder sammenspill mellom mennesker og skjermer. Men standarden sier ikke noe om hvordan overholdelse av kravene kan måles i praksis. Det er heller ikke dens oppgave, og her vises det til delene 305, 306 og 307. Noen av kravene i ISO 9241-303 er derimot relevant for målemetode, for eksempel terskelverdier, og bør derfor kombineres med det som er spesifisert i de andre delene. Annex D er vurdert til å være særdeles relevant for eventuelle krav hva gjelder minimumskontrast, se spesielt ligning D.11.

Til slutt må det nevnes at det i Del 303 er eksplisitt formulert (avsnitt 1) at standarden ikke tar høyde for funksjonsnedsettelse utover en redusert synsfunksjon pga. høyere alder¹⁶. Dette

14 ISO 9241-303:2011. «Ergonomics of human-system interaction -- Part 303: Requirements for electronic visual displays»

15 Det er uvisst hvorfor korrekturfaktoren ikke er anbefalt brukt. Kanskje det menes at datagrunnlaget er for tynt, og at nevnte studie bør verifiseres av annen forskning først før korrekturen kan bli normativ standard.

16 På engelsk brukes vendingen «with normal or corrected-to-normal vision», noe som inkluderer bruk av brille.

betyr igjen at de delene som bygger på Del 303, vil heller ikke funksjonsnedsettelse spille noen rolle i. Denne aspekten utgjør en svakhet av hele 300-serien i ISO 9241.

5.6 NS-EN ISO 9241-305:2008

Også denne spesifikasjonen er både norsk og europeisk standard og har fått tittelen «Ergonomi for samhandling mellom menneske og system - Del 305: Optiske testmetoder for elektronisk visuelle skjermer»¹⁷. Standarden må ses i sammenheng med delene 302, 303 og 307 og gjelder altså for testing i standardiserte labomgivelser.

Dette slår seg blant annet ned i at det er anbefalt å unngå eller i det minste kontrollere direkte eller indirekte lys, f.eks. fra måleutstyret eller i form av refleksjoner fra vegger eller strølys forårsaket av optisk utstyr, for å forhindre målefeil (avsnitt 5.5). Luminansmeteret bør ha mulighet for å regulere åpningen og fokus for å foreta spotmålinger (5.7). Det er beskrevet detaljerte ulike fremgangsmåter og metodikk, f.eks. til spotmålinger og tallfesting av refleksjonskoeffisienter og lignende (6), samt nødvendig måleutstyr og andre hjelpemidler. Spesielt interessant i denne sammenhengen er måling av kontrast ved hjelp av mørkerom, et gitt testmønster og to spotmålinger, i et tilfelle uten ekstern lyssetting (6.7.1), og i tre tilfeller med ekstern lyssetting og uønskede refleksjoner (6.7.2, 6.7.3, 6.7.4). I den sammenhengen er det også relevant hvordan luminans måles ved hjelp av en lignende fremgangsmåte, men med fem/ni/elleve spotmålinger (6.6.1, 6.6.7). Det er også verdt å merke seg at standarden skiller mellom spredte/diffuse refleksjoner, f.eks. forårsaket av indirekte, eksterne lyskilder som himmelen, og speilingseffekter, som ofte kan føres tilbake til direkte eksterne lyskilder som solen; begge er komponenter som bidrar til nettokontrasten.

Andre detaljer fra standarden som er relevant for måling av skjermkontraster er: Før en måling skal skjermen ha fått anledning til å stabilisere seg i minst 20 minutter (5.1). Det er anbefalt å måle et større område på minst 500 piksler for å komme fram til riktig resultat (5.6.1). Der det er brukt bokstaver i testmønstrene er disse utformet på en ikke *anti-aliased* måte, det vil si med skarpest mulig kant mellom lyst og mørkt (5.3). Det anbefales videre å foreta usikkerhetsberegninger for alle utførte målinger for å gi en øvre grense for feil med målingene (Annex E).

5.6.1 Vurdering

ISO 9241-305 er meget relevant i denne sammenhengen, og den kan sies å gi svært detaljerte retningslinjer for å gjennomføre målinger av skjermkontraster. Som sådan kan den brukes av både skjermleverandører og testinstitusjoner. Ulempen med den er derimot at alle målinger må utføres under nøye kontrollerte forhold, som begrenser dens bruk i praksis. De beskrevne måleinstrumentene og -metodene er videre ganske omfattende og ressurskrevende, både hva gjelder kostnader og tidsbruk. For målinger ute i felten vises til ISO 9241-306.

5.7 NS-EN ISO 9241-306:2008

Del 306 i 300-serien av ISO 9241 har fått tittelen «Ergonomi for samhandling mellom menneske og system - Del 306: Feltvurderingsmetoder for elektronisk visuelle skjermer» på norsk¹⁸. I motsetning til Del 305 gjelder den altså testing i felten, men det skal nevnes at den har et særfokus på ergonomi på skjermarbeidsplasser. Standarden må ses i sammenheng med delene 302, 303, 305 og 307 i ISO 9241-serien.

Ekstern lyssetting kan måles ved hjelp av en luxmeter (avsnitt 5.2.1), mens skjermens egenstråling kan måles med luminansmeter (5.2.2). Det er foreskrevet med visse

17 ISO 9241-305:2011. «Ergonomics of human-system interaction -- Part 305: Optical laboratory test methods for electronic visual displays»

18 ISO 9241-306:2008. «Ergonomics of human-system interaction -- Part 306: Field assesment methods for electronic visual displays»

målingsområder på skjermen og en bestemt orientering av målingsverktøyet i forhold til skjermen. Skjermer må få anledning til å stabilisere seg før en måling, uansett skjermt teknologi (4.1.1, 4.2.1). Avstand mellom betrakter og skjerm bør være mellom 400 og 750 mm (5.1.1), men avstanden mellom skjerm og luminansmeter avhenger av målerens blendejusteringsevne (5.5.1.1). Til målinger av glansgraden av overflater anbefales det en glossmeter (5.2.3). Standarden anbefaler videre å måle kontrasten mellom skjermen og omgivelsene, f.eks. en vegg i nærheten, med anbefalingen om at den bør være så liten som mulig (5.2.3). Måling av bokstavekontraster stiller krav om justeringsevne av blendeåpning og fokuspunkt på luminansmeteret (5.5.1.1, 5.5.1.2). Måling av bokstavtykkelsen kan gjennomføres gjennom direkte sammenligning med en referanseskala og eventuelt bruk av forstørrelsesglass/lupe (5.5.5). Ellers kan andre parametre også evalueres etter den visuelle-inspeksjonsmetoden (5.6.1, 5.6.4, 5.7.2, m.fl.).

Standarden lister videre opp en hel rekke faktorer som påvirker skjermergonomien (Annex B): avstand til betrakteren, skjermens orientering i forhold til øyet, lyssetting, lysstyrke, eventuelle vibrasjoner, vind og vær, visuelle artefakter grunnet aldringsprosesser i materialet eller produksjonsavvik og -feil, designfeil, kontrast, bildepolaritet, tekststørrelse og -utforming, med flere.

5.7.1 Vurdering

ISO 9241-306 er i høyeste grad relevant for måling av skjermkontraster. Standarden spesifiserer nøyaktig en rekke testmetoder, viser til gyldige verdier for målingsresultater og attpåtil beskriver format for rapportering. Etter vår mening er den til og med mer relevant enn Del 305 fordi den er beregnet på målinger i felten og på den måten tar bedre hensyn til realistiske forhold sammenlignet med labomgivelser.

Men, med tanke på alle faktorene som påvirker brukeropplevelsen (Annex B) mener vi at standarden ikke klarer å fange opp helhetsinntrykket ettersom det gjennomføres en rekke enkelttester uten å ta hensyn til helhetsopplevelsen til brukeren, eller med andre ord hvordan alle disse endimensjonale aspektene påvirker hverandre. Det er heller ikke tatt hensyn til hvordan eventuelle funksjonsnedsettelse påvirker bruksopplevelsen, og hvordan disse kan fanges opp av endrede eller flere terskelverdier. I den forbindelsen er det relevant å spørre om man kunne lagd en skjerm som overholder alle krav i spesifikasjonen, men som likevel gir et ubrukelig resultat for en person med nedsatt funksjonsevne.

5.8 NS-EN ISO 9241-307:2008

Del 307 i 300-serien av ISO 9241 bærer tittelen «Ergonomi for samhandling mellom menneske og system - Del 307: Metoder for analyse og samsvarsprøving for elektronisk visuelle skjermer»¹⁹. Standarden konkretiserer, med utgangspunkt i allerede omtalte Del 303, testmetodene fra Del 305 for ulike skjermt teknologier.

På et overordnet nivå skiller ISO 9241-307 på de ulike skjermtypene CRT²⁰ (avsnitt 5.1), LCD²¹ (5.2), PDP²² (5.3), projiserings skjermer (5.4) og mobil-LCD (5.5), hvis egenskaper krever ulike målemetoder. Standarden gir en rekke terskel- og referanseverdier som går utover de verdiene som er nevnt i de andre delene (f.eks. C.2.3). Det er eksplisitt nevnt at samtlige anbefalinger kun gjelder for innendørs bruk. Videre anbefaler standarden å dokumentere omstendighetene for alle målinger i form av spesifikasjon av brukerne, brukskontekst, oppgave og beskrivelsen av selve innretningen, f.eks. om den er stasjonær eller mobil, størrelse på skjermen osv. (Annex C).

19 ISO 9241-307:2008. «Ergonomics of human-system interaction -- Part 307: Analysis and compliance test methods for electronic visual displays»

20 Engelsk: cathode ray tube

21 Engelsk: liquid crystal display

22 Engelsk: plasma display panel

5.8.1 Vurdering

Del 307 i ISO 9241 er svært relevant for en teknisk måling av skjermkontraster ettersom den utgjør en ytterligere spissing av kravene nevnt i Del 305 for å ta hensyn til ulike skjermtyper. Med sine over 200 sider er den svært omfattende; noen av seksjonen for måling av kun en skjermtype er på over 50 sider! Samtidig begrenses standardens betydning av det faktum at målingene gjelder kun innendørs, og fordi eventuelle funksjonsnedsettelse eller helhetsinntrykket ikke er tatt høyde for.

5.9 W3C WCAG 2

Web Content Accessibility Guidelines (WCAG, i skrivende stund med versjonsnummer 2) er W3Cs anbefaling for tilgjengelighet på Web²³.

I avsnitt 1.4.3 defineres det et minimums kontrastforhold mellom to (relative) luminansverdier. Noen tilfeller – som dekorative elementer, logoer og merkenavn – er unntatt dette kravet. Ellers er det verdt å merke seg at WCAG skiller mellom vanlig tekststørrelse (under 14 pt²⁴ om tykke bokstaver («bold»), ellers under 18 pt) og stor tekststørrelse (over dette) med et krav om minimums kontrast på 4,5:1 i første tilfellet og 3:1 i andre tilfellet på nivå AA. På nivå AAA må tilsvarende kontrast bli minst henholdsvis 7:1 og 4,5:1.

5.9.1 Vurdering

WCAG 2 er en allment kjent, innarbeidet og ofte brukt anbefaling. Krav 1.4.3 har vist seg å være godt egnet til å spesifisere og måle *bruttokontrasten* i mange ulike kontekster, og det finnes en rekke verktøy for å måle at kontrastkravet blir overholdt.

Hovedsvakheten med WCAG 2-metoden er at den dreier seg om intensjonen til en forfatter eller designer etc., men ikke om netto- eller den opplevde kontrasten²⁵. Verdiene i anbefalingen er tatt fra ISO 9241-303²⁶, og det er også tatt høyde²⁷ for en noe redusert synsfunksjon gjennom alder. Det vil igjen si at verdiene ignorerer andre faktorer, som skjerminnstillinger (eksempelvis lysstyrke), aldring av skjermen, produksjonsfeil og lignende avvik. Denne listen inneholder også eventuelle funksjonsnedsettelse. Det må videre nevnes at det ikke er klart hvor mye eventuelt endrede terskelverdier måtte bety i form av inkludering og ekskludering²⁸, men denne diskusjonen gjelder ikke bare WCAG 2.

W3C jobber i skrivende stund med en revisjon av WCAG 2, som skal få versjonsnummeret 2.1, men det er uvisst når dette arbeidet avsluttes, hva som oppdateres og hvordan dette vil slå ut.

5.10 W3C CSS 2.1

Cascading Style Sheets (CSS, i skrivende stund med versjonsnummer 2.1) er W3Cs anbefaling for stilsett på Web²⁹.

Størrelser i CSS relateres til brukerens leseavstand til skjermen, som er satt til gjennomsnittlig armlengde eller 71 cm (avsnitt 4.3.2).

5.10.1 Vurdering

I CSS 2.1 er det kun leseavstanden som er relevant i denne sammenhengen, men dette aspektet dekkes også av f.eks. 300-serien i ISO 9241.

23 <https://www.w3.org/TR/WCAG20/>

24 1 pt er 1/7200 cm

25 Det er dog tatt hensyn til en generell strølys-effekt i formelen for kontrast.

26 WCAG 2 siterer forgjengeren til Del 303, ISO 9241-3:1992.

27 <https://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-contrast.html#visual-audio-contrast-contrast-73-head>

28 <https://github.com/w3c/wcag/issues/200>

29 <https://www.w3.org/TR/CSS21/>

W3C jobber for tiden med å revidere CSS, som skal få versjonsnummeret 2.2, men det er uvisst når dette arbeidet avsluttes, og hva som kommer til å endres. Utkastet på versjon 3 bruker samme definisjon på leseavstand³⁰.

5.11 ITU-T P.800.1 og ITU-T P.800.2

P.800.1³¹ og P.800.2³² er deler av ITU-T³³s P-serie med tekniske anbefalinger som omhandler ulike typer terminaler og metoder for både objektive og subjektive kvalitetsmålinger. Helt konkret dreier begge seg om MOS³⁴, en verdi som kvantifiserer den opplevde kvaliteten av stimuli og/eller tekniske løsninger. Begge standardene er påvirket av arbeidet VQEG³⁵ har gjort i forbindelse med kvalitetsvurderinger av video. VQEG er en i utgangspunkt uavhengig organisasjon, men flere av deres studier har resultert i tekniske ITU-anbefalinger.

MOS er definert som aritmetisk middelvei av en rekke subjektive enkelskåre³⁶. P.800.2 anbefaler en fem trinns skala med kategoriene «meget dårlig», «dårlig», «middels», «bra», og «meget bra» (avsnitt 7). MOS påvirkes av en rekke subjektive faktorer med den konsekvensen at den må suppleres av tilleggsinformasjon bestående av en statistisk analyse med målingsantall, standardavvik og konfidensintervaller som minimumskrav for at MOS-sammenligninger skal være gyldige (7, 10). Dette gjelder blant annet hva som blir vist på skjermen, varighet av synsinntrykket, med eller uten brukerinteraksjon, oppgavebeskrivelsen, og en rekke andre parametre (12). Hverken P.800.1 eller P.800.2 sier noe om hvordan de enkelte MOS-skårene skal måles. Her vises det til P.910, som er omtalt lenger nede.

5.11.1 Vurdering

P.800.1 blir ofte nevnt i samme åndedrag som P.800.2, men inneholder ikke mye relevant informasjon for eventuelle subjektive kontrastmålinger. P.800.2 anses til å være relevant for subjektive kvalitetsbedømmelser av visuelle inntrykk, selv om den først og fremst gjelder video. Hovedforskjellen til konteksten her er at P.800.2 ønsker å skape så like forhold som mulig for alle informanter i en utprøving, mens det som skal fanges opp med kontrastmålingen er nettopp variasjonen grunnet ulike omgivelser og lysforhold. Likevel mener forfatterne at en beregning av MOS teoretisk og praktisk kan brukes for å bedømme kontraster på skjerm. MOS er en av de mest brukte metodene for kvantifisering av subjektive bedømmelser, selv om den også har noen svakheter, som diskutert i P.800.2. Derfor tilkommer en statistisk analyse av MOS nøkkelrollen, og det er forventet en ganske stor spredning av de individuelle skårene. Angivelse av MOS for en gitt automat bør alltid følges av en beskrivelse for beregningen av MOS, som beskrevet i P.800.2, og en detaljert statistisk analyse.

5.12 ITU-T P.910

ITU-T P.910³⁷ er en teknisk anbefaling for gjennomføring av subjektive kvalitetsmålinger. Den er i samme serien som ITU-Ts P.800.1 og P.800.2.1 og gjelder i utgangspunkt for multimedia og video. De fleste delene i standarden er derimot så generelle at den også kan anvendes på andre områder, noe som også eksplisitt er kommentert.

ITU-Ts P.910 beskriver flere ulike testmetoder, men felles for alle er at det forutsettes full kontroll over skjermbildet (avsnitt 5-7) og omgivelsene (5.1, 7.1). Standarden foreskriver

30 <https://www.w3.org/TR/css3-values/>

31 ITU-T, Recommendation P.800.1 – Mean opinion score (MOS) terminology, 07/2016

32 ITU-T, Recommendation P.800.2 – Mean opinion score interpretation and reporting, 07/2016

33 International Telecommunication Union – Telecommunication Sector

34 Engelsk forkortelse for Mean Opinion Score

35 Video Quality Expert Group

36 https://www.wikiwand.com/en/Mean_opinion_score

37 ITU-T, Recommendation P.910 – Subjective video quality assessment methods for multimedia applications, 04/2008

mellom fire og 40 informanter, der 15 er anbefalt. Ingen av disse bør være profesjonelle for å unngå en skjevrappotering (bias). Mindre grupper kan gi meningsfulle pekepinn allerede i tidligere utviklingsstadier av tekniske løsninger (7.3). I utgangspunkt bør en fem trinns skala brukes, som den definert i P.800.2, men det kan også tenkes flere trinn der dette er hensiktsmessig, eller eventuelt en kontinuerlig skala (B.1). Dette for å unngå for mange ensformede vurderinger av totalopplevelsen. Det kan også være aktuelt å stille flere spørsmål enn det etter totalopplevelsen, for eksempel vedrørende lysstyrke, kontrast, fargeskjema, lesbarhet, refleksjoner, eventuelle animasjoner eller grafikk, med flere (B.2). Svarene til disse kan så brukes til å bekrefte eller svekke helhetsinntrykket. Alternativt kan enkeltskårene settes sammen til en ny, vektet hovedskåre.

5.12.1 Vurdering

Noen områder i P.910 er spesifikke for kvalitetsmålinger av video, men det aller meste kan også brukes for måling av kontrast på skjerm og subjektiv opplevelse. Det som ikke er relevant er de punktene som er til for å skape mest mulig like forhold for alle informanter, noe som ikke er ønskelig i denne konteksten. Også tallene 4/15/40 må ses i samme lys, og en subjektiv måling av kontrast i felten bør nok involvere et betydelig høyere antall informanter for å fange opp hele spektret av variasjoner i omgivelsene og lysforhold. Det er derimot en god idé å gjennomføre målinger med færre personer under hele utviklingsløpet, ikke kun på slutten. En annen god idé er ikke bare å måle helhetsinntrykket, men også å spørre etter relaterte faktorer som refleksjoner og lysstyrke.

5.13 ITU-R BT.500-13

BT.500-13³⁸ inneholder anbefalinger vedrørende subjektiv evaluering av bildekvaliteten i tv-overføringer. Den er publisert i BT³⁹-serien av ITU-R⁴⁰, det vil si til kringkastingsformål, men det som ikke er tv-spesifikt er også relevant for subjektive kontrastmålinger. Standarden må ses i sammenheng med andre tekniske ITU-R-anbefalinger for nøyaktig spesifisering av enkelte testmetoder (avsnitt 3). En kombinasjon av testmetoder kan vurderes (6.5).

BT.500-13 anbefaler minst 15 testdeltagere (2.5). Den enkeltes bakgrunn burde dokumenteres for å kunne forklare eventuelle forskjeller i skårer. Selv om ikke eksplisitt nevnt vil dette også omfatte en eventuell funksjonsnedsettelse. En skåre bør alltid følges av dokumentasjon for testoppsett og en statistisk analyse av enkeltmålinger, bestående i det minste av antall deltagere, aritmetisk middelværdi av enkeltmålingene og 95%-konfidensintervall (2.8). Beregningsformlene er definert i Annex 2. For konteksten av denne rapporten blir en kategorisk 5-trinns skala som i ITU-T P.800.2 anbefalt (6.1.4.1), men for høyere nøyaktighet kan skalaen også være kontinuerlig (6.1.4.3). Det kan være relevant å måle flere delaspekter ved siden av helhetsinntrykket, sånn som lesbarhet og andre faktorer (6.1.4.1). Det kan også defineres objektive målekriterier ved å gi en testdeltager en oppgave og måle oppnåelse, tidsbruk, antall feil, og så videre (6.1.4.4). Den statistiske analysen bør også inneholde et steg for å filtrere bort eventuelle uteliggere (outliers) blant svarene (Annex 2).

5.13.1 Vurdering

BT.500-13 vurderes som relevant for måling av skjermkontrast ved hjelp av en subjektiv evaluering, men det er mye av det samme i forhold til ITU-T P.910 og P.800.2. Denne anbefalingen er derimot mer detaljert og definerer nøyaktig beregningsmetoder og terskelverdier, for eksempel for konfidensintervall, og spesifiser også filtreringmetoden for å

38 ITU-R, Recommendation BT.500-13 – Methodology for the subjective assessment of the quality of television pictures, 01/2012

39 Broadcasting service (television)

40 International Telecommunication Union – Radiocommunication Sector

fjerne uteliggere. Også anbefalingen å benytte seg av objektive målekriterier og oppgaver under en subjektiv evaluering er nyttig.

6 Relevant forskning

Her omtales et utvalg av relevant forskning på området.

6.1 Rapport: Lesbarhet av trykt tekst for svaksynte

En NTNU-rapport⁴¹ fra 2016 har sett på en rekke parametre for god lesbarhet for svaksynte, inkludert kontrast.

Studien viser at mørkt/lyst-polaritet⁴² gir signifikant bedre lesbarhet enn lyst/mørkt-polaritet på papir⁴³. Kontrasten bør enten være større enn 100/60 for mørkt/lyst-polaritet eller større enn 0/40 for lyst/mørkt-polaritet. Rapporten gjør også oppmerksom på at skriftstørrelse har noe mer å si for lesbarheten enn kontrast. Studien benytter seg av en metode der leseren selv vurderer lesbarheten på en fire punkts skala. Verdiene regnes så om til en prosentuell verdi for akseptans/lesbarhet.

6.1.1 Vurdering

Rapporten er interessant i henhold til metode og resultat, men ikke direkte anvendbar her siden den kun omtaler trykksaker, og det vises til at opplevelse av kontrast med luminans (som med selvlysende skjermer) oppfattes forskjellig sammenlignet med illuminans (altså flater med eksterne lyskilder)⁴⁴.

Arbeidet setter kontrast i en direkte sammenheng med lesbarhet av tekst der også flere andre aspekter spiller inn, som polaritet, skriftstørrelse og skrifttype. Fremgangsmåten er særdeles interessant i og med at det kun brukes (subjektive) brukervurderinger der hver informant har sin egen forståelse av hva det vil si å ha god/tilstrekkelig kontrast.

6.2 Rapport: Undersøkelse av synsforhold ved billettautomater og informasjonstavler

I en annen NTNU-studie⁴⁵ fra 2013 ble det blant annet testet billettautomater, storskjermer og skilting ute i felten med den hensikt å måle lesbarhet for personer med nedsatt syn. Nersveen viser til at målinger av skjermkontrast i felten ofte vil være støy-befengt på grunn av variasjoner i strølyset gjennom skjerming og reflekterende klær (avsnitt 3.2). Videre vil det lett kunne oppstå variasjon i luminansmålinger av små områder/felter fordi dette krever at luminansmeterets minste blendeåpning brukes, som igjen fører til at apparatet blir svært følsom for bevegelse (3.2).

I utprøvingen av billettautomatene ble det kombinert luminansmeter-målinger med spørreskjema (3.2). Det er uklart om Michelsons eller Webers kontrastdefinisjon ble brukt, men kontrastberegningen gir svært lite informasjon i de fleste tilfellene uansett fordi svart skrift på en forholdsvis lys bakgrunn ga som oftest resultatet -1. I spørreskjema ble det spurt om hvor lett forsøkspersonen trodde det var å bruke skjermbildet.

41 Nersveen og Johansen. Lesbarhet av trykt tekst for svaksynte. NTNU, 2016

42 I rapporten defineres positiv kontrast-som mindre prosentuell sorthet i forgrunn enn i bakgrunn, dvs. ensbetydende med *negativ* polaritet ellers brukt i forskningen og i ISO 9241; derfor brukes her lyst/mørkt for å unngå forvirring. Negativ kontrast defineres i rapporten som større prosentuell sorthet i forgrunn enn i bakgrunn, dvs. ensbetydende med *positiv* polaritet; derfor brukes det her mørkt/lyst.

43 Dette er en motsetning til det som er nevnt i ISO/TR 22411:2008.

44 ISO/TR 22411:2008, avsnitt 8.6.4

45 Nersveen. Undersøkelse av synsforhold ved billettautomater og informasjonstavler ved NSB Oslo S og Jernbaneverket. NTNU, 2013

Evalueringen av skilting ble gjennomført ved hjelp av et spørreskjema som spurte etter faktorer som skriftstørrelse, kontrast, fargekombinasjoner o.l. med svarmulighetene «dårlig», «middels» og «god» (3.3).

Storskjermbildene ble testet i en labsetting med oppgaven å lese så mye som mulig på skjermen der tekst ble vist i mange ulike skriftstørrelser (3.4). Testen ble gjennomført med ulike horisontale belyningsstyrker i rommet målt 0,75 m over gulvet, samt ulike leseavstander, med registrering av den minste størrelsen av skrift forsøkspersonen kunne lese.

6.2.1 Vurdering

Studien er dessverre noe begrenset på grunn av en liten testpopulasjon, men er interessant fordi objektiv kontrastmåling blir kombinert med subjektiv brukerutprøving i form av spørreskjema. Som Nersveen poengterer forventes en større variasjon av måleverdier, både objektive og subjektive, på grunn av sammenspill av mange faktorer man stort sett ikke har kontroll over ute i felten. Derfor bør testpopulasjonen økes. Der det er mulig bør oppgaver benyttes hvis suksess lett kan måles (f.eks.: les ordet), framfor å spørre om eksempelvis tilfredshet eller forventning.

6.3 Annen forskning om polaritet og subjektive vurderinger

I en studie fra 2009 ble det påvist at polaritet ikke har noe å si for lesbarhet på skjerm⁴⁶. I følge studien er det snarere slik at effekten av god lesbarhet med positiv polaritet mest sannsynlig kan føres tilbake til høy skjermlyuminans og altså ikke polaritet.

Et annet arbeid fra samme år har undersøkt kontrast i bilder, der ulike algoritmer for kontrastberegning ble sammenstilt med subjektive evalueringer⁴⁷. Disse foregår i en standardisert labsetting der betrakterne bruker sin egen forståelse av hva kontrast er og gir en skår på en skala fra 0 (lavest) til 100. Målet var å komme frem til egnede algoritmer som etterligner subjektive vurderinger.

Beslektet med dette er en studie med temaet subjektiv kontrast i bilder i ukontrollerte omgivelser⁴⁸. I studien er ovenfor nevnte skala redusert til 10. Videre slår studien fast at forskjellen på den subjektive evalueringen mellom labsetting og ukontrollerte omgivelser ikke er signifikant, men viser samtidig til at de subjektive rapporteringene er mindre korrelerte med ulike objektive metrikker.

6.3.1 Vurdering

At ulike studier kommer til ulike resultater når det gjelder polaritet tyder på at temaet er ganske komplekst, og at minst én av studiene har svakheter. Så her må mer metodisk god forskning til før man eventuelt kan konkludere.

Hovedbudskap av studiene vedrørende subjektive vurderinger er at man kan få gyldige resultater selv om disse gjennomføres utenom labben.

7 Ekspertanbefalinger

Ekspertuttalelsene er gjengitt i vilkårlig rekkefølge.

For tidligere rådgiver i Hjelpemiddelsentralen i Sør-Trøndelag, avdeling Syn, **Bjørn Nygård**, må gode skjermkontraster først og fremst ses i kontekst dataskjerm og arbeidsplass. Et eksempel

46 Buchner, Mayr, Brandt. «The advantage of positive text-background polarity is due to high display luminance». *Ergonomics* Vol. 52, No. 7, July 2009, 882–886

47 Simone, Pedersen, Hardeberg, Rizzi. «Measuring perceptual contrast in a multi-level framework». *IS&T/SPIE Electronic Imaging*, 72400Q-72400Q, 2009

48 Simone, Pedersen, Hardeberg. «Measuring perceptual contrast in uncontrolled environments». *IEEE Visual Information Processing (EUVIP)*, 2010

her er baklys når dataskjermen er vendt mot vinduet. Hva som er tilstrekkelig kontrast er høyst individuelt, slik at det ikke kan gis et fasitsvar. Nygård mener dog at det er viktig ikke å henge seg opp i tall og prinsipper; skjermen bør fortrinnsvis prøves ut individuelt. «God oppløsning» er i følge ham det viktigste. Kontrast og kontrastmålinger har sine begrensninger, mener han basert på over 20 år med erfaring og over 100 ulike utprøvinger. For eksempel er god kontrast ikke mye til hjelp for personer med kikkertsyn. Et prinsipp gir han allikevel: Jo mindre skjerm, desto bedre bør kontrasten være.

Synspedagog i Blindeforbundet, **Arne Tømta**, tar forbehold om at hans erfaring med måling av skjermkontraster er begrenset. Han mener det er tekniske retningslinjer som er den mest objektive måten å sikre gode skjermkontraster på, både når det gjelder luminanskontrast og fargekontrast. Det er derimot usikkerhet omkring hvorvidt de objektive metodene klarer å estimere den subjektive opplevelsen. Tømta nevner følgende eksempel: Ved en utprøving av skjermer med tanke på best lesbarhet falt valget for påfallende mange med nedsatt syn på den skjermen som ikke hadde best kontrast teknisk sett. Derfor er det brukerens opplevelse av skjermen som er det viktigste. Subjektive utprøvinger bør alltid gjennomføres i realistiske omgivelser. I samsvar med Nygård mener også Tømta at denne opplevelsen er meget individuell. Eventuelle diagnoser spiller inn, og det samme gjør vaner. Tømta gir følgende eksempel: Høy kontrast vil redusere utholdenheten i lesingen for en med høy kontrastfølsomhet / lysømfintlighet, mens høy kontrast for andre uten denne diagnosen vil øke utholdenheten fordi den reduserer behovet for forstørring.

For tidligere grafisk designer og industridesigner **Ruedi á Porta** henger måling av kontraster på skjerm og lesbarhet tett sammen. En annen viktig parameter for lesbarhet er skjermens oppløsning, noe som gir seg utslag i netto-strektykkelse, som er den vist på skjermen. Denne kan derfor være en del tynnere enn strektykkelsen designeren har valgt. Det er et viktig poeng at skalering av skrift i kombinasjon med kantutjevning, såkalt *font smoothing* og *anti-aliasing*, gjør at skriften blir gråere og dermed dårligere lesbar med høyere oppløsning. Med liten skrift stiger derfor kravene til kontrast.

Førsteamanuensis ved NTNUs Institutt for vareproduksjon og byggteknikk, **Jonny Nersveen**, anbefaler en to trinns prosess for måling av kontrasten: En gang uten forstyrrelser fra omgivelsen, og en gang for å måle refleksjonene fra omgivelsene, men begge i en labsetting, det vil si et mørkt rom⁴⁹. Nersveen er klar på at denne prosedyren ikke fanger opp alle persepsjonelle reaksjoner en seende måtte ha, men han understreker viktigheten av at innkjøpere av utstyr er i stand til å sammenlikne kvaliteten på ulikt utstyr både med hensyn til egenkontrasten og evnen til å bevare kontrasten i utplassert stand. For utregning av kontrasten anbefaler han å bruke Webers definisjon av skjermens egenkontrast $C_S = (L_F - L_B) / L_B$, der L_F og L_B er henholdsvis tekst-/forgrunns luminans og bakgrunns luminans. Luminansen måles ved hjelp av luminansmeter rettet mot relativt store flater. Effekten av strølys fanges så opp gjennom en tilleggs målinger av sløringsluminansen L_S ved avslått skjerm i et rom med gitt diffus belysning på eksempelvis 1000 lux. Helhetskontrasten blir da $C = C_S \cdot D$ med blendingsgrad $D = L_B / (L_B + L_S(1000 \text{ lux}))$. Sløringsluminans i et rom med annen lysstyrke vil i følge Nersveen være proporsjonal med den diffuse belysningen, slik at forventet kontrast kan regnes ut ved å beregne L_S (gitt lysstyrke) først og så sette det inn i formelen for blendingsgrad fremfor $L_S(1000 \text{ lux})$. Sløringsluminansen og gitt horisontal lysstyrke (her: 1000 lux) bør derfor være en del av skjermens dokumentasjon.

Førsteamanuensis **Ole Lund** ved Institutt for design, NTNU, har spesialisert seg på typografi og informasjonsdesign. Han påpeker at kontrast bare er én av mange faktorer når det gjelder lesbarhet. Han har videre vist til en del svakheter og feil med relatert forskning i sin avhandling

49 På engelsk: darkroom

og mener at mange av forskningsresultatene blir feilaktig brukt i ukritiske anbefalinger⁵⁰. Han etterlyser både forsiktighet i tolkning av resultatene, mindre rigiditet i anvendelse og mer fleksibilitet.

Dosent **Magne Helland** ved HSNs Institutt for optometri, radiografi og lysdesign arbeider først og fremst med synskrevende arbeid ved dataskjermer og tar forbehold om begrenset erfaring med fargekontrast. Han peker derimot til en rekke studier som har undersøkt relaterte mikrovariabler som synskomfort, synsergonomi og produktivitet. Noen faktorer er ganske perifere til lesbarhet, som luftfuktighet i omgivelsene, blunkfrekvens, eventuelle reflekser i skjermen, hvordan skjermen er plassert med hensyn på blikkvinkel, stress på arbeidsplassen, og kontorlandskap versus enkeltmannskontor. Utover det anbefaler han en fremgangsmåte for kontrastmålinger som er lik den Nersveen beskriver.

7.1 Vurdering

Ekspertene kommer fra ulike fagfelt og har ulike innfallsvinkler, og derfor er det ingen overraskelse at det er stor variasjon i svarene. Nersveen har et poeng i at innkjøpere (og tilsynsmyndighetene) bør kunne sammenligne kontrasten på tvers av skjermene, og her vil en labmåling som den han skisserer, være sentral. En teknisk løsning bør spesifiseres i henhold til teoretisk kontrast og det vi kan omtale som sløringskontrast. Begge kan måles som beskrevet ovenfor. I tillegg bør en brukertesting belyse om kontrasten fortsatt er tilstrekkelig når løsningen er utplassert i felten og fange opp ikke-tekniske parametre. Evalueringen bør inkludere personer med nedsatt syn.

8 Brukerinnspill

Sammenlignet med kanalens popularitet var det få reaksjoner på de to innleggene, men begge er i skrivende stund med henholdsvis 48 og 119 godt over gjennomsnittet på 30 for 2017, så saken engasjerer mer enn ellers. De følgende eksemplene ble gitt (noen uten mulighet for verifisering av forfatterne):

- Hvit skrift på rød bakgrunn i det grafiske grensesnittet av elektroniske programguider på tv (Altibox, Canal Digital)
- Hvit tekst på hvit bakgrunn i en vignett på NRK 1
- Grå, tynn skrift på hvit bakgrunn på noen nettsider (Sanguine Brasserie, Argent Restaurant)
- Grå/svart skrift på grå bakgrunn på elektroniske prislapper på butikker

I forbindelse med første kulepunkt er kommentaren fra en grensesnittdesigner oppsiktsvekkende. Han forklarer dårlige grensesnittvalg med at mye av teknologien er patentert, og at det gir færre gode valgmuligheter.

8.1 Vurdering

Det var ikke overveldende reaksjon på innleggene våre, og eksemplene som kom inn var stort sett kjent fra før av. Kommentaren vedrørende teknologipatenter som hindrer gode løsninger var derimot meget interessant. Om denne påstanden viser seg til å stemme, er dette noe som bør behandles av tilsynsmyndighetene og politikerne.

⁵⁰ Lund. Evidence-based typography or easy-going operationalism? International Conference on Typography and Visual Communication, 2002

9 Anbefaling og begrunnelser

Basert på vurderingene av relevante standarder, relatert forskning, uttalelser av eksperter, samt brukerinnspill gis følgende anbefaling for måling av skjermkontraster. Begrunnelsene for anbefalingene er skrevet med innrykk.

Det er den opplevde, subjektive kontrasten som er fasitsvar på spørsmålet om skjermkontrasten til en teknisk løsning er tilstrekkelig / god nok for «folk flest». Denne nettokontrasten påvirkes av en rekke faktorer som materialvalg, skjerminnstillinger, design- og utformingsvalg, lyssetting og -skjerming og andre tiltak for å unngå uønskede refleksjoner, og ikke minst eventuelle variasjoner av synsfunksjonen hos det enkelte individ. Dette brede spekteret av faktorer kan kun delvis estimeres ved hjelp av objektive kriterier og bør derfor suppleres med subjektive målinger, det vil si brukerevalueringer.

En kontrastmåling bør derfor være tredelt og bestå av to objektive og en subjektiv verdi.

Verdiene skal dokumentere både skjermens egenskaper fra fabrikken og forholdene der skjermen er oppstilt. Førstnevnte skal gi innkjøpere (og tilsynsmyndighetene) anledning å sammenligne flere skjermer på en uavhengig måte, og målet med sistnevnte er å fange opp aspekter forårsaket av utplassering, lyssetting og refleksjoner. Tilnærmingen imøtekommer videre kravene til alle ekspertene som ble intervjuet.

Den objektive delen bør gjennomføres i en labsetting under kontrollerte forhold.

Dette for pålitelige, repeterbare resultat.

En objektiv kontrastmåling bør bestå av to luminansmålinger, en for høy luminans L_h og en for lav luminans L_l . Luminans bør måles ved hjelp av et luminansmeter⁵¹ plassert 75 cm fra skjermen. Luminansmeteret bør være sentrert over skjermen og rettet mot den slik at normalen gjennom skjermens midtpunkt går i rett linje gjennom luminansmeterets linse.

Skjermavstanden på 75 cm følger anbefalingen fra ISO 9241-303 og -306 og er ikke langt fra W3Cs anbefaling på 71 cm. Den er særlig beregnet på selvbetjeningsautomater der brukeren interagerer med skjermen, men kan også observeres i forbindelse med svaksynte i møte med informasjonstavler, se Nersveens studie av tavler på Oslo S. Oppsettet følger ellers anbefalingene fra ISO 9241-305.

Første objektive måling bør foregå i et mørkt rom (illuminasjon/belysningsstyrke mindre enn 1 lux; mørke, absorberende overflater osv.). Illuminasjon bør måles med et luxmeter⁵² horisontalt i rommet, i en høyde av 0,75 m over gulvet.

Dette følger anbefalingene til Nersveen og Helland.

En luminansverdi bør gjelde en flate med homogen, statisk luminans (og homogen farge). Det bør velges flater som er store nok til at luminansmeterets målingsfelt ligger fullstendig innenfor flaten slik at apparatet viser en konstant verdi.

Denne anbefalingen er til for å unngå variasjoner i målingene.

Kontrasten bør defineres som kontrastforhold⁵³ L_h/L_l , det vil si forhold mellom høy luminans og lav luminans.

Dette er definisjonen brukt i ISO 9241, ANSI og W3C.

Resultat for første objektive måling blir da den maksimale kontrasten C_{max} .

Andre objektive måling bør gjennomføres som første måling, men i et rom med diffus belysning på 1000 lux.

51 Eksempler er Konica Minolta LS-150 og LS-160

52 Eksempler er Konica Minolta T-10A og T-10MA

53 På engelsk: contrast ratio

Dette følger igjen anbefalingene til Nersveen.

Resultat blir da den noe reduserte kontrasten C_{1000} .

Den subjektive målingen bør gjennomføres ved hjelp av en brukertesting.

Dette er i samsvar med ITU-T P.800.2 og P.910, samt ITU-R BT.500-13, og videre inspirert av Nygård, Tømta og Nersveen.

Evalueringen bør foregå i felten, der en gitt automat er utplassert, med relevante skjermbilder. Mellom 15 og 20 informanter bør delta i testen, og halvparten av disse bør ha en eller annen moderat form for redusert synsfunksjon.

Brukerantallet følger anbefalingen P.910. I utgangspunkt gjelder «jo flere, desto bedre», men grensen bør dras ved ca. 20 for å holde kostnadene nede. Involvering av brukere med redusert synsfunksjon anbefales for å øke variasjonen i svarene.

Den enkeltes bakgrunn og egenskaper bør dokumenteres. Hver informant bør gi skårer for faktorene kontrast (k), lesbarhet (l), refleksjoner (r) og blanding (b). Følgende spørsmål bør stilles, ett om gangen, og uten mulighet for å se tidligere svar, i denne rekkefølgen:

1. Hva er din vurdering av kontrastene på skjermen på følgende skala?
2. Hva er din vurdering av lesbarheten av det som vises på skjermen på følgende skala?
3. Hva er din vurdering av refleksjonene på skjermen på følgende skala?
4. Blir du blendet av skjermbildet eller deler av skjermbildet?

For svaret bør det benyttes en kontinuerlig glidebryter med synlig inndeling jevnt fordelt over skalaen. For Spørsmålene 1 og 2 bør inndelingen være:

- «meget dårlig»
- «dårlig»
- «middels»
- «bra»
- «meget bra»

For spørsmål 3 bør inndelingen være:

- «meget forstyrrende»
- «ganske forstyrrende»
- «synlige, men ikke forstyrrende»
- «nesten ignorerbare»
- «helt ignorerbare»

For spørsmål 4 bør inndelingen være:

- «ja, i meget høy grad»
- «ja, i moderat grad»
- «vet ikke / ingen formening»
- «nei, i liten grad»
- «nei, overhodet ikke»

Den underliggende skalaen bør gå fra 0 (første alternativ) til 10 (siste alternativ).

Alle disse anbefalingene, med unntak av formuleringen av spørsmålene, er i henhold til P.910. Den kontinuerlige skalaen fra 0 til 10 kopierer fremgangsmåten i 2010-artikkelen

til Simone. Spørsmål vedrørende blanding er inspirert av ISO/TR 22411:2008 og ISO 9241-303.

Alle enkelskårer bør inngå i beregningen av henholdsvis MOS_k , MOS_l , MOS_r og MOS_b , altså én for hver faktor nevnt tidligere. Det bør legges ved fullstendig dokumentasjon for testoppsett og en statistisk analyse bestående av antall deltagere, aritmetisk middelværdi og 95%-konfidensintervall.

MOS er anbefalt i P.800.2, og dokumentasjonen er omtalt i BT.500-13.

Helhetsinntrykket for kontrasten, og samtidig resultatet for den subjektive målingen, C_{sub} , bør så regnes ut som gjennomsnitt av MOS_k , MOS_l , MOS_r og MOS_b . Resultatet er altså et tall mellom 0 og 10, der 10 betyr utmerket kontrast (og ikke blanding heller).

Denne fremgangsmåten er valgt fordi de enkelte MOS-verdiene har en verifiserende relasjon til hverandre. MOS_b er til for å hindre at skåren blir høy i tilfellet kontrasten er for stor.

9.1 Anbefaling for utregning av minimumskontrast, $L_{h,min}$ og terskelverdier

Det vises til ISO 9241-303, Annex D for en god anbefaling når det gjelder utregning av minimumskontrast og minste luminans $L_{h,min}$ som gir en god/tilstrekkelig leseopplevelse i forhold til L_l . Ved å kombinere ligningene D.9 og D.11 blir minimumsluminansen:

$$L_{h,min} = k_a \cdot (2,2 + 4,84 \cdot L_l^{-0,65}) \cdot (L_l + L_d + L_s) - L_d - L_s$$

med de følgende parametre:

- L_h : høy luminans
- L_l : lav luminans
- k_a : aldersavhengig korreksjonsfaktor (se tabell D.1)
- L_d : luminans forårsaket av diffus belysning
- L_s : luminans forårsaket av direkte lyskilder som spotter og solen⁵⁴

Dette er en ganske kompleks formel, men til gjengjeld tar den hensyn til en generell reduksjon i synsfunksjonen med alderen, og til både diffuse og direkte refleksjoner. Med andre ord kan det ikke gis ett tall som terskelverdi for tilstrekkelig kontrast sånn som WCAG gjør med sitt krav om én konstant minimumskontrast på 4,5 (Level AA). Den er snarere avhengig av alle nevnte parametre og vil dermed variere betraktelig.

Det må nevnes at formelen også har noen begrensninger. For eksempel fanger den ikke opp eventuelle nedsettelse av synsfunksjonen, og ISO 9241-303 nevner videre at $L_{h,min}$ også er avhengig av skjermstørrelse.

Det må også sies til slutt at det bør finnes en maksimal kontrast og maksimal luminans $L_{h,max}$ for en god leseopplevelse uten blanding, men her har forfatterne ikke funnet relevant forskning eller anbefalinger.

10 Konklusjon

Denne rapporten har tatt for seg temaet hvordan måle kontraster på skjerm på en pålitelig og effektiv måte. Dette er en høyst aktuell problemstilling ettersom flere og flere skjermer blir brukt for informasjonstavler, selvbetjeningsautomater og personlige enheter.

⁵⁴ På engelsk: specular reflection og specular luminance

I grunn og bunn dreier alt seg om problemstillingen hvordan man mest mulig objektivt kan måle kontrastinntrykket «folk flest» har med en gitt skjerm. Som denne rapporten viser, er dette en forholdsvis innviklet affære på grunn av kompleksiteten av synsfunksjonen, brukernes diversitet, bredden på anvendelsesområdene og spennvidden av de ulike faktorene som påvirker sluttresultatet.

Vi mener skjermens egenskaper og motstandsdyktighet mot refleksjoner bør inngå i en kontrastmåling. Dette for å tilrettelegge for rettferdige sammenligninger av skjermer. I tillegg bør målingen fange opp faktorene forårsaket av utplassering, omgivelsene og andre, overveiende subjektive faktorer. Vi har beskrevet en metode bestående av to objektive målinger og en brukerevaluering. Alt i alt er denne tilnærmingen forholdsvis ressurskrevende og kostbar med dagens teknologi. Til gjengjeld vil den gi et bedre resultat enn én enkel måling. Det er også en fordel at brukeropplevelsen gjenspeiles i høyere grad i målingene enn det som hittil har vært tilfellet.