

Amalie Flostrand

## Manipulert eller ikke?

En kvantitativ studie av hvor flinke nordmenn er til å gjenkjenne redigeringer i bilder på sosiale medier

Masteroppgave i Medier, Kommunikasjon og Informasjonsteknologi

Veileder: Melanie Magin

Medveileder: Stefan Geiss

Juni 2022



Amalie Flostrand

## **Manipulert eller ikke?**

En kvantitativ studie av hvor flinke nordmenn er til å gjenkjenne redigeringer i bilder på sosiale medier

Masteroppgave i Medier, Kommunikasjon og Informasjonsteknologi  
Veileder: Melanie Magin  
Medveileder: Stefan Geiss  
Juni 2022

Norges teknisk-naturvitenskapelige universitet  
Fakultet for samfunns- og utdanningsvitenskap  
Institutt for sosiologi og statsvitenskap



More often than not, people think that the real images are fake and that things that are fake are real. And their confidence is very high. So people are both ignorant and confident, which is the worst combination – Henry Fraid, 2020

## **Abstract**

After the introduction of user-friendly image editing tools, editing images has become something anyone with a smartphone can do. This availability has led to users being regularly exposed to images that may have been edited in several different ways - whether one is aware of it or not.

The research question of this project is as follows: To what extent are Norwegians able to recognize edits in images, and what factors influence this? To answer this question, an experiment was conducted using an online questionnaire with 522 Norwegian participants between the age of 23 and 90 years. The experiment aimed to examine people's ability to recognize edits such as light, contrast and color changes, as well as their ability to recognize cosmetic edits aimed to make people portrayed as more "attractive". The stimulus material of the experiment is composed of four different images edited with one of three different forms of editing: light / contrast, feminine beauty stereotypical features or masculine beauty stereotypical features. In total, this amounts to 16 different images and editing combinations, including images presented without edits (experiment control group). To answer the research question, the participants were asked to assess how likely they thought it was that the picture they were presented with was edited or not. Other factors such as the participants' assessment of attractiveness, the gender depicted and the extent to which the correct ones were able to distinguish between edited and unedited images were also investigated. Self-representation, beauty stereotypes, visual literacy and signal detection theory are all theoretical concepts that forms the foundation for further exploration of the topic.

The main findings of this experiment showed that the participants had a success rate of 33% when it came to distinguishing between edited and unedited images, while the average score the participants gave when asked how likely it was that the image was manipulated was 2.72 (on a scale from 1 to 5). The form of editing that showed the greatest effect on both the participants' probability assessment and detection ability was feminine beauty stereotypical features. Furthermore, pictures depicting women had a greater effect on attractiveness assessment, detection ability and probability assessment set up against men, while age turned out to be negatively related to both attractiveness assessment, detection ability and probability assessment. The results from this experiment are considered a contribution to a small field of research, which shows that Norwegians generally do not perform well in recognizing edits in images.

## Sammendrag

Etter introduksjonen av brukervennlige bilderedigeringsverktøy har det å redigere bilder blitt noe alle med en smarttelefon kan gjøre. Denne tilgjengeligheten har ført til at man som bruker jevnlig blir eksponert for bilder som kan ha blitt redigert på en rekke ulike måter – enten man er klar over det eller ikke. Dette prosjektets problemstilling er som følger: I hvilken grad er nordmenn i stand til å gjenkjenne redigeringer i bilder, og hvilke faktorer har innflytelse på dette? For å besvare dette spørsmålet ble det gjennomført et eksperiment ved bruk av online-spørreskjema med 522 norske deltakere i alder mellom 23 til 90 år. Eksperimentet hadde som mål å undersøke folks evne til å gjenkjenne redigeringer slik som lys, kontrast og fargeendringer, samt deres evne til å gjenkjenne ulike redigeringer gjort for å gjøre personen avbildet mer «attraktiv». Eksperimentets stimulusmateriale er satt sammen av fire ulike bilder redigert med en av tre ulike redigeringsformer: lys/kontrast, feminine skjønnhetsstereotypiske trekk eller maskuline skjønnhetsstereotypiske trekk. Til sammen utgjør dette 16 ulike bilde og redigeringskombinasjoner, inkludert bilder presentert uten redigeringer (eksperimentets kontrollgruppe). For å besvare problemstillingen ble deltakernes bedt om å vurdere hvor sannsynlig de selv trodde det var at bildet de ble presentert for var redigert eller ikke. Andre faktorer slik som deltakernes bedømmelse av attraktivitet, hvilket kjønn som var avbildet og i hvilken grad de korrekt klarte å skille mellom redigerte og ikke-redigerte bilder ble også undersøkt. *Selvrepresentasjon*, *skjønnhetsstereotyper*, *visual literacy* og *signal detecting theory* er alle teoretiske begreper som legger grunnlaget for utforskning av temaet.

Eksperimentets hoved funn viser at deltakerne hadde en suksessrate på 33% når det kom til å skille mellom redigerte og ikke-redigerte bilder, mens den gjennomsnittlige skåren deltakerne ga i møte med spørsmålet om hvor sannsynlig de trodde det var at bildet var manipulert var på 2.72 (på en skala fra 1 til 5). Redigeringsformen som viste størst effekt både på deltakernes sannsynlighetsvurdering og deteksjonsevne var feminine skjønnhetsstereotypiske trekk. Videre viser resultatene at kvinner avbildet hadde større effekt på attraktivitetsvurdering, deteksjonsevne og sannsynlighetsvurdering satt opp mot menn, mens faktoren alder viser seg å ha negativ korrelasjon med både attraktivitetsvurdering, deteksjonsevne og sannsynlighetsvurdering. Resultatene fra dette eksperimentet er ansees som et tilskudd på et lite utforsket forskningsfelt, som viser at nordmenn generelt sett ikke presterer godt i å gjenkjenne redigeringer i bilder.

## Forord

Arbeidet med masteroppgaven har vært både lærerikt og krevende, og det er i den sammenhengen en god del personer som fortjener en takk.

Aller først ønsker jeg å takke min bi-veileder Stefan Geiss, som ikke bare har gitt meg veiledning og gode råd mens jeg seilte inn i metodisk ukjente farvann, men som også har bidratt med støtte og entusiasme som har holdt båten flytende. Min veileder Melanie Magin fortjener også en stor takk, som har bidratt med gjennomtenkte og detaljerte tilbakemeldinger, oppmuntrende ord og ikke minst moralsk støtte gjennom hele prosessen.

Innleveringen av denne masteroppgaven representerer både slutten av mine år som NTNU-student og starten på en helt ny epoke av livet. Så til min studiepartner gjennom både bachelor- og master år, Kristin Heggdal, vil jeg gi en takk for akademiske samtaler, latterkramper og et godt vennskap. Jeg vil også takke min personlige heiagjeng mamma Gry som aldri har vært lenger enn en telefonsamtale unna, og min evig tålmodige samboer Simen, som gjennom tykt og tynt (og eksamensperioder) har støttet og motivert meg til å gjøre mitt aller beste. Deres tilstedeværelse og kjærlighet har vært avgjørende i min tid som student i Trondheim.

Til slutt vil jeg gi en takk til vår lille master-baby som har sparket i magen og oppmuntret meg gjennom studietimene. Du har gjort mitt siste semester som student spennende, forventningsfylt og også litt krevende – ikke minst har du gitt meg en real utfordring og mine medstudenter en god latter på vei opp og ned Dragvoll sine trapper! Nå er det bare dager igjen til vi møtes, og jeg kan ikke nesten ikke vente!

Amalie Flostrand  
Juni, 2022



## Innhold

1.	Innledning .....	1
2.	Teoretisk bakgrunn og sentrale begreper .....	4
2.1	Goffmans teori om selvrepresentasjon .....	4
2.2	Kjønn og stereotypier .....	6
2.2.1	Skjønnhetsstereotypier blant kvinner og menn .....	8
2.3	Sosial sammenligning og kognitiv dissonans .....	10
2.4	Signal Detecting Theory (SDT) .....	12
2.5	Visual literacy .....	15
2.5.1	Visual literacy definert .....	15
2.5.2	Visuell kompetanse i en digital alder .....	15
3.	Metode .....	19
3.1	Valg av metode .....	19
3.2	Utvalg og rekruttering .....	20
3.3	Eksperimentets design .....	22
3.3.1	Spørreundersøkelsens struktur .....	22
3.3.2	Stimulusmateriale .....	24
3.3.3	Treatment-grupper .....	28
3.4	Variabler .....	29
3.4.1	Avhengig variabel, redigert/ikke redigert .....	29
3.4.2	Uavhengige variabler .....	29
3.5	Deskriptiv statistikk .....	33
3.6	Analysemetode .....	36
3.6.1	Balansetest .....	36
3.6.2	Estimering av treatment-effekt .....	37
3.6.3	ANOVA med gjentatte målinger og marginal prediksjon .....	38
3.6.4	Regresjonsanalyse .....	39
3.7	Eksperimentets indre og ytre validitet .....	39
3.7.1	Indre validitet .....	39
3.7.2	Ytre validitet .....	41
3.8	Utfordringer og endringer .....	42
4.	Resultater .....	44
4.1	Treatment-effekt og overordnede resultater .....	44
4.2	Subjektiv sannsynlighetsvurdering av redigering i bildet .....	45
4.2.1	Påvirkning av bilde .....	45
4.2.2	Påvirkning av redigering .....	46
4.2.3	Samspill mellom bilde og redigeringstype .....	47
4.2.4	Innflytelse av andre faktorer .....	48

4.3	Attraktivitetsvurdering .....	49
4.3.1	Påvirkning av bilde .....	51
4.3.2	Påvirkning av redigering .....	51
4.3.3	Samspill mellom bilde og redigering .....	52
4.3.4	Innflytelse av andre faktorer .....	53
4.4	Deteksjonsevne .....	54
4.4.1	Påvirkning av bilde .....	54
4.4.2	Påvirkning av redigering .....	55
4.4.3	Samspill mellom bilde og redigering .....	56
4.4.4	Innflytelse av andre faktorer .....	57
4.5	Oppsummering av sentrale funn .....	58
5.	Diskusjon .....	60
5.1	En kjønnert vurdering .....	61
5.2	Et domene forbeholdt den yngre generasjon .....	62
5.3	Lavere attraktivitetsvurdering og større skepsis .....	63
5.4	Begrensninger og videre forskning.....	64
6.	Konklusjon .....	67
	Kilder .....	69

## Figurer

<b>Figur 1:</b> Fire ulike typer av respons, SDT .....	13
<b>Figur 2:</b> Reklamebilder for kampanjen på Facebook og Instagram. ....	21
<b>Figur 3:</b> utdrag fra spørreskjema, ulike klokkesekvenser .....	23
<b>Figur 4:</b> Eksempelutdrag fra eksperimenteres 16 ulike bilde- og redigeringskombinasjoner.....	25

## Tabeller

<b>Tabell 1:</b> oversikt over redigeringer i stimulusmaterialet .....	26
<b>Tabell 2:</b> oversikt over spørreskjemaets ulike treatment-grupper .....	28
<b>Tabell 3:</b> Deskriptiv statistikk for avhengig variabel og forklaringsvariabler.....	33
<b>Tabell 4:</b> Deskriptiv statistikk for uavhengige variabler .....	33
<b>Tabell 5:</b> Deskriptiv statistikk for demografisk uavhengige variabler .....	36
<b>Tabell 6:</b> Det representerte kjønnets påvirkning på sannsynlighetsvurdering .....	45
<b>Tabell 7:</b> Redigeringstypes påvirkning på sannsynlighetsvurdering .....	46
<b>Tabell 8:</b> Post-hoc analyse av hvordan redigeringstype og kjønn av avbildet person samspiller i å påvirke attraktivitetsvurdering .....	47
<b>Tabell 9:</b> Andre faktorielle innflytelser på attraktivitetsvurdering .....	49
<b>Tabell 10:</b> Subjektiv sannsynlighetsberegning avhengig av attraktivitetsvurdering og redigeringstype .....	50
<b>Tabell 11:</b> Det representerte kjønnets påvirkning på attraktivitetsvurdering .....	51
<b>Tabell 12:</b> Redigeringstypes påvirkning på attraktivitetsvurdering .....	52
<b>Tabell 13:</b> Samspill mellom bilde og redigering, attraktivitetsvurdering .....	52
<b>Tabell 14:</b> Andre faktorielle innflytelser på attraktivitetsvurdering .....	53
<b>Tabell 15:</b> Gjennomsnittlig deteksjonsevne mellom gruppene .....	54
<b>Tabell 16:</b> Det representerte kjønnets påvirkning på deteksjonsevne .....	54
<b>Tabell 17:</b> Redigeringstypes påvirkning på deteksjonsevne.....	55
<b>Tabell 18:</b> Samspill mellom bilde og redigering, deteksjonsevne .....	56
<b>Tabell 19:</b> Logistisk regresjon, andre faktorielle innflytelser på deteksjonsevne.....	57

## 1. Innledning

Før man hadde alle dagens digitale verktøy krevde det både tid og ekspertise for å redigere bilder. Den gangen ble det gjerne praktisert manuelt i et mørkerom, mens i dag er det ikke verre enn å knipse et bilde med mobiltelefonen og åpne en redigerings-applikasjon. Etter introduksjonen av Photoshop og andre verktøy for bildemanipulering, fulgte en rekke reaksjoner fra forskere innen visuelle medier som hevdet at samfunnet hadde gått inn i en epoke der fotografier ikke lenger evnet å være trofaste gjenspeilinger av virkeligheten. Her kan man blant annet nevne titler som William J. Mitchell sin *“The Reconfigured Eye: Visual Truth in the PostPhotographic Era”* (1992), eller Fred Ritchin sin *«After Photography»* (2008). Hvor vidt dette er en tematisk overdrivelse eller ikke, er forbeholdt en annen diskusjon. Likevel er det ikke tvil om at det å ta bilder, redigere dem og dele dem med verden er noe som ikke lenger er forbeholdt erfarne fotografer eller spesielt interesserte. I dag deles ofte manipulererte bilder på ulike sosiale medier-plattformer, mens plattformene selv ikke klarer å opprettholde etiske retningslinjer for å motvirke at manipulererte bilder feilinformerer andre brukere (Tandoc et al., 2018). Selv om det i Norge i 2021 ble innført en lov om at manipulert reklame og sponsede innlegg må markeres (Hvitmyhr, 2021), er det fortsatt fritt frem for privatpersoner å skjule sine redigeringer på sosiale medier. Faktisk er det ikke uvanlig at de som deler redigerte bilder er privatpersoner som ønsker å oppnå en bedre form for selvrepresentasjon (Hancock & Toma, 2009; Slutsky, 2020).

I en artikkel fra CBC skriver Madison O'Dell (2021), en videregående-elev fra Canada, om hvordan hun selv pleide å redigere bildene hun tok av seg selv før hun la de ut på sosiale medier:

I literally start from the top and work to the bottom, whitening anything that's white, like my eyes and my teeth. Next it's my blemishes, pimples and scars. I simply erase them with the touch of a finger like they've never been there. Then my body. If I've just eaten a bunch of snacks and I'm bloated like any human would be, I edit the selfie to make me look skinnier with wider hips (O'Dell, 2021).

Slik Madison påpeker, vil man enten tenke at «dette er galskap», eller så vil man relatere seg til hennes opplevelser (O'Dell, 2021). For når det kommer til å redigere bilder av seg selv før man legger dem ut på sosiale medier, er ikke Madison alene. I 2017 var Facetune – en applikasjon som lastes ned på mobilen – Apples mest populære applikasjon (Solon, 2019). Redigeringsapplikasjonen som i utgangpunktet hadde som mål å samle inn noen hundre tusen dollar, har nå satt sitt digitale fingeravtrykk over hele Internett (Solon, 2019). Andre lignende applikasjoner har også blitt populære på grunn av hvor enkle de er å bruke, i motsetning til mer komplekse redigeringsprogrammer som Adobe Photoshop (Jennings, 2019).

Når personer redigerer bildene sine, ønsker de gjerne ikke at andre skal se at de har gjort det. «Nobody wants to be called a fake» sier Brooke Erin Duffy, en assisterende professor ved Cornell University som studerer selvrepresentasjon på Instagram (referert i Solon, 2019). Også «influencere» (personer som gjennom ønsker å påvirke holdninger blant sine følgere på sosiale medier (Pihl, 2021), ønsker å presentere seg selv autentisk, samtidig som de ønsker å poste et bilde som representerer dem selv på best mulig vis

(Duffy, referer i Solon, 2019). «If done properly, it should be hard to tell you've used it» sier Natasha Church (2017), en sminke- og skjønnhetsblogger i sin YouTube-veiledning som forklarer hvordan hun redigerer bildene sine. I videoen fortsetter skjønnhetsbloggeren å vise hvordan man kan endre på ansiktsstrukturen ved å gjøre kjeven større eller smalere, endre størrelsen på nesen og øynene, samt endre hudtonen eller jevne ut urenheter og fine linjer.

Den offentlige diskusjonen rundt det å ta og redigere bilder av selv har i stor grad vært fokusert rundt kvinner, og trekker gjerne kvinner frem som mer aktive både når det kommer til å ta og redigere bilder av seg selv (Murray, 2015; Qiu et al., 2015; Dhir et al., 2016; Fox & Vendemia, 2016). Usunne idealiseringer av kvinnekroppen i mediebildet og en hyppig formidling av at det å være attraktiv vil føre til at kvinner blir sosialt godkjent, kan ifølge Rose og kolleger (2012) være innvirkende faktorer. Samtidig har også tidligere forskning vist at økt bruk av sosiale medier kan føre til internalisering av det «tynne-idealet», som igjen kan føre til at brukere redigerer og poster bilder av seg selv oftere (Caso et al., 2020). Selv om kvinner kanskje er en mer utsatt gruppe enn menn, har en rekke studier vist at det ikke bare er kvinner som praktiserer redigering av eget utseende (Hildebrandt et al., 2008; Fox & Vendemia, 2016; Lonergan et al., 2019).

Med den økende oppmerksomheten rundt redigering av bilder har vi som brukere også blitt mer bevisste på hvor enkelt det kan være å manipulere innhold som postes på sosiale medier (Egelhofer et al., 2019). Bevisstheten later derimot ikke til å ha gjort folk flinkere til å gjenkjenne manipulasjoner i bilder. Henry Farid, en amerikansk universitetsprofessor som har spesialisert seg innen analyse av digitale bilder, beskriver menneskers evne til å identifisere redigerte bilder slik: «More often than not, people think that the real images are fake and that things that are fake are real. And their confidence is very high. So people are both ignorant and confident, which is the worst combination» (Fraid, referert i Wen, 2020). Det finnes en rekke studier som har utforsket personers evne til å identifisere redigerte bilder, blant annet Shen med kolleger (2019), Nightingale (2017), Brumberg (2011), Schetinger med kolleger (2017) og Farid og Bravo (2010) for å nevne noen. Resultatene fra disse studiene leder mot den samme konklusjonen, nemlig at folk flest ikke er spesielt flinke til å skille mellom redigerte bilder og ikke-redigerte bilder. Flere av disse studiene er likevel noe eldre, og i et medielandskap som stadig er under utvikling vil det være interessant å igjen undersøke denne tematikken, spesielt nå som sosiale-medier i stor grad er blitt en integrert del av de flestes hverdag. Det er også en god del mangler på forskningsfeltet når det kommer til den nyere formen for «kosmetisk redigering» av bilder på sosiale medier blant brukere. Med dette menes redigeringsformer ment for å endre utseendet mot en mer kosmetisk ønsket fremstilling, slik som å gjøre seg selv slankere eller andre ønskede appellerende endringer. Det er – så langt min kunnskap rekker – ikke tidligere blitt gjennomført eksperimentelle undersøkelser av menneskers evne til å gjenkjenne ulike redigeringsformer på sosiale medier i Norge. Videre er det heller ikke blitt gjennomført eksperimenter som fokuseres rundt subtile kosmetiske redigeringer man som bruker gjerne møter på sosiale medier-plattformer.

Dette eksperimentet har som mål å undersøke folks evne til å gjenkjenne redigeringer slik som lys, kontrast og fargeendringer, samt deres evne til å gjenkjenne kosmetiske redigeringer for å gjøre personen avbildet mer «attraktiv». Her vil andre underliggende faktorer slik som subjektiv antakelse av attraktivitet, kjønn avbildet og form for redigering også undersøkes opp mot eksperimentets tematikk, i et forsøk på å besvare

det overordnede forskningsspørsmålet: I hvilken grad er nordmenn i stand til å gjenkjenne redigeringer i bilder, og hvilke faktorer har innflytelse på dette?

For å besvare dette spørsmålet ble det gjennomført et eksperiment ved bruk av online-spørreskjema med 522 norske deltakere, i alderen mellom 23 til 90 år. Her ble deltakerne presentert fire ulike bilder, et original/ikke-redigert bilde og tre redigerte bilder. For å gjøre eksperimentets stimulusmateriale så «virkelighetsnært» som mulig ble bildene som presenteres for deltakerne redigert ved hjelp av applikasjonen Facetune på smarttelefon. Stimulusmaterialet består til sammen av 16 ulike sammensetninger av bilder og redigeringsformer (inkludert kontrollbildet som ikke er redigert), der hvilken sammensetning som ble presentert for deltakeren basertes på deltakerens tildeling av eksperimentgruppe.

## 2. Teoretisk bakgrunn og sentrale begreper

Gjennom de kommende sidene vil jeg presentere teoretiske perspektiver og sentrale begreper som kan bidra til å gi en dypere forståelse av prosjektets tematikk – redigering av bilder og folks evne til å gjenkjenne dem. Jeg vil trekke tidligere studier opp mot teoriene for å gi en bedre forståelse av hvordan teoriene er relevante for forskningsfeltet og dette eksperimentet. Ut ifra denne gjennomgangen vil også eksperimentets hypoteser trekkes frem. Deretter vil jeg diskutere hvilke implikasjoner manipulering av bilder kan ha på brukere av sosiale medier, samt hvordan man kan bruke eksisterende litteratur og forskning til å forstå hvordan mennesker skiller mellom manipulerede og ikke-manipulerede bilder på nett.

### 2.1 Goffmans teori om selvrepresentasjon

Goffmans teori om strategisk selvrepresentasjon har hatt en stor effekt på samfunnsvitenskapen der den har påvirket analysen av individuell atferd (Moore, 2012). Det er i boken «The Presentation of Self in Everyday Life» at Erving Goffman (1959) presenterer idéen om en strategisk selvrepresentasjon. Her beskriver han hvordan mennesker bevisst danner en definisjon av situasjonen man befinner seg i, der oppfatningen av seg selv er en viktig del når man møter andre (Goffman, 1959). Det at individet velger å presentere seg selv på en måte som best mulig gagnar sin egen agenda, og at individer vil tolke andres selvrepresentasjon gjennom de samme kriteriene, står sentralt i Goffman sin teori (Moore, 2012).

I boken fra 1959 bruker Goffman begrepet *performance* – eller opptreden – for å beskrive hvordan en person fremstiller seg selv foran en gitt mengde observatører. Han presenterer altså teorien gjennom en teatralisk metafor - eller en «dramaturgisk tilnærming» - der han snakker om individer som skuespillere med tildelte roller som opptrer på en scene (Hogan, 2010; Moore, 2012). Her bruker Goffman (1959) også begrepene *front* og *setting*. Han bruker «front» for å definere situasjonen for dem som observerer individets opptreden, mens begrepet «setting» defineres som det fysiske oppsettet som skaper en slags scene for individet som opptrer (Goffman, 1959). Her står også tanken om en «front-stage» og en «back-stage» sentralt, der «front-stage» fungerer som en arena i det sosiale liv der man blir observert av andre, mens «back-stage» er en privat arena der individet kan trekke seg tilbake uten å bli observert (Goffman, 1959). Hogan beskriver Goffmans dramaturgiske tilnærming som en: «metaphorical technique used to explain how an individual presents an “idealized” rather than authentic version of herself» (2010, s. 378).

I senere tid har det blitt vanlig å bruke Goffmans dramaturgiske tilnærming i studier av sosiale medier, der teorien blir ansett som et nyttig verktøy for å forstå presentasjonen som brukere gir av seg selv på nett (se blant annet Hewitt & Forte, 2006; Smith & Sanderson, 2015; Merunková & Šlerka, 2019; Bullingham & Vasconcelos, 2013). Bruken av Goffmans teori i disse sammenhengene har ført til en utvidelse av metaforen hans. Hogan (2010) mener selv at sosiale mediers vurderings- og søkbarhetskaraktistikk i større grad gir en følelse av «exhibition» – eller en utstilling – heller enn en «performance». Her skiller det mellom de typer nettrom hvor aktører interagerer med hverandre – det Goffman (1959) ville kalt «performance»- rom – og «utstillingsrom» der

individer administrerer artefakter for å vise til hverandre. Hogan (2010) mener at idet data knyttet til en forestilling blir lagret, blir de også omgjort til en artefakt som er tilgjengelig for andre på ulike tidspunkter og i ulike sammenhenger. På denne måten brukes data fra sosiale medier mer til asynkrone utstillinger av seg selv enn synkrone forestillinger. Denne utviklingen fører til at sosiale medier danner rom hvor individer bevisst kan velge hvilke deler av seg selv de vil skal være deres personlige virtuelle versjon (Rose et al., 2012).

En stor del av dagens nettbaserte selvrepresentasjon skjer med det som kalles «selfier» (Rettberg, 2017). I 2013 ble ordet «selfie» kåret til årets ord av Oxford English Dictionary (Qiu et al., 2015). I samme ordbok defineres begrepet på følgende måte: «A photograph that one has taken of oneself, esp. one taken with a smartphone or webcam and shared via social media» (Oxford English Dictionary, u. å). «Selfie» som form for selvrepresentasjon vil ifølge Cruz og Thornham (2015) også underbygge forestillinger om «performance», «selfbranding», og andre konsepter tilknyttet den enkeltes bruk av sosiale medier. Det er gjennom iscenesettelsen, bildetagning, deling av bildet, kommentarer og likes på digitale medier at det «fotografiske selvet» blir meningsfylt, ikke som et enkelt bilde, men som en kompleks prosess som performativt konstruerer selvet (Cruz & Thornham, 2015). Dette gjenspeiles også i tidligere forskning på feltet. I tråd med Goffmans (1959) forslag om at selvrepresentasjon er en prosess der man endrer selvet for å skape et visst inntrykk for et publikum, undersøkte Hancock og Toma (2009) virkningen av kjønn på selvrepresentasjon og sosial tiltrekningskraft på nettbaserte dating-sider. Her fant forskerne at både kvinner og menn «redigerer» profilene sine for å skape en bedre selvrepresentasjon og en form for «selektiv selvrepresentasjon»; en mer kontrollert handling der man endrer eller forvrenger egne bilder, som ofte fører til en endret fremstilling av profilen (Hancock & Toma, 2009). I tråd med dette gjennomførte Slutsky i 2020 intervjuer med regelmessige brukere av Instagram for å undersøke hvordan folk opplever selvrepresentasjon og skjønnhet knyttet opp til plattformen. Deltakerne i undersøkelsen fortalte at de redigerte kroppene sine for å nå opp til idealet om å «leve sitt beste liv» og se bra nok ut (Slutsky, 2020). Videre trekkes det frem tre sentrale faktorer som bidrar til skjønnhetsstandarder som eksisterer på plattformen, og som bidrar til det Slutsky (2020) selv kaller «the platformization of beauty». Dette handler om presset som oppstår som følger av sosial sammenligning og forventning, mulighetene for å redigere innhold, samt følelsen av å få belønning i form av likes, delinger, kommentarer og lignende. Disse faktorene beskriver hun som sentrale for hvordan normer for selvrepresentasjon på Instagram og andre plattformer har blitt formet, da det å presentere seg selv gjennom et «plattformisert skjønnhetsideal» er sentralt for mange brukere (Slutsky, 2020). Denne måten å redigere seg selv på sosiale medier er blitt dokumentert blant både kvinner og menn (Hildebrandt et al., 2008; Fox & Vendemia, 2016; Lonergan et al., 2019). Likevel har mye av fokuset blitt rettet mot kvinner, internalisering av idealiserte fremstillinger av kropp og redigering av selfier (Cohen et al., 2018; Chae, 2017; Lee & Lee, 2021). Dette er en tematikk som jeg videre vil diskuteres opp mot skjønnhetsstereotyper og kjønn i det kommende kapittelet.

Ser man tidligere forskning på feltet opp mot Goffmans (1959) teori om selvrepresentasjon og den senere utvidelsen av hans teori (Hogan, 2010), er det ikke vanskelig å se at det å ønske å representere seg selv på best mulig måte på sosiale medier kan føre til et økt fokus på skjønnhetsstereotyper, som videre kan føre til at flere redigerer bilder av seg selv før de deler dem med resten av verden. Det at folk føler de står på «utstilling» for andre, kan med andre ord føre til at man ønsker å presentere en idealisert fremstilling heller enn en autentisk fremstilling av seg selv. Denne preferansen



mot en idealisert fremstilling skaper grunnlag for å tro at det er stor sannsynlighet for at manipulerede bilder er noe man som bruker av sosiale medier stadig møter på – om man er klar over det eller ikke. Basert på gjennomgangen i avsnittene over, finner man også grunnlag for å tro at redigering av bilder skjer på et individ-nivå. Det betyr at bildene som legges ut på slike plattformer gjerne redigeres av brukere og for brukere. De ønsker altså å forbedre sin egen selvrepresentasjon med formål om å imponere andre brukere, noe som i stor grad gjenspeiler Goffmans (1959) sine tanker om «performance» på en scene, og Hogan (2012) sin senere utvidelse av teorien som et digitalt «utstillingsrom». Knyttet opp mot det nåværende eksperimentet, fungerer Goffmans teori om selvrepresentasjon som en indikator på fenomenet bildemanipulasjon på sosiale medier, og er på den måten et sentralt fundament for oppgavens tematikk og problemstilling. Forankret i Goffman sin teori, senere teoretiske utvidelser og tidligere forskning på feltet vil det være interessant å undersøke hvor flinke folk er til å identifisere redigeringer i bilder, da det er stor sannsynlighet for at redigeringer er noe som skjer på sosiale medier-plattformer hele tiden.

Goffmans teori er, slik man ser i avsnittene over, en teori som forankrer mange ulike aspekter ved selvrepresentasjon, det ideelle selvet, redigering av bilder og også representasjon av kjønn. I det kommende kapitlet vil jeg gå nærmere inn på hvilken rolle kjønn spiller i fremstillingen av en idealisert skjønnhet, og bruke dette til å danne grunnlaget for behandlingen av eksperimentets stimulusmateriale.

## 2.2 Kjønn og stereotyper

Maskulinitet og femininitet er beskrivelser som ofte brukes i daglige språket. Begrepene er ofte assosiert med fysiske og biologiske forskjeller mellom menn og kvinner. I de følgende avsnittene vil jeg gå igjennom hva som definerer skjønnhetsstereotyper bant kvinner og menn. Deretter vil jeg beskrive hvordan de ulike stereotypiske trekkene vil fungere som et rammeverk for hvordan stimulusmaterialet i dette eksperimentet blir behandlet. De ulike formene for stereotypiske trekk blir brukt som utgangspunkt for å redigere eksperimentets bilder.

I tråd med menneskers representasjon av selvet på sosiale medier finner man også visse kjønnede faktorer, blant annet når det kommer til hvordan kvinner og menn bruker selfier på sosiale medier som identifikasjon på eget kjønn (Katz & Crocker, 2015). Tidligere forskning har nemlig funnet noen sentrale forskjeller i hvordan kjønn blir presenter gjennom selfier, og i hvilken grad de opprettholder ulike former for kjønnsstereotyper. Blant annet så er det mer sannsynlig at kvinner tar selfier enn det motsatte kjønn (Qiu et al., 2015), samt at kvinner også har en høyere sannsynlighet for å redigere bilder av seg selv (Dhir et al., 2016). Sett i sammenheng med at det å være attraktiv ofte er et ideal som forbindes med femininitet, og at kulturelle normer gjerne formidler til kvinner at det å være attraktiv kan gi sosial godkjenning, er det ikke overraskende (Rose et al., 2012). Går man mer spesifikt inn på ulikheter mellom kjønnene og deres selvrepresentasjon på sosiale medier, ser man at menn ofte poster selfier av seg selv i atletisk utstyr og er alene i bildeutsnittet, gjerne mens de utøver en sport eller oppholder seg utendørs (Rose et al., 2012). Dette mener Rose og kolleger (2012) er noe som kan gjenspeile et ønske om å fremstille seg selv som sunn, sterk og uavhengig. I kontrast observerte forskerne at kvinner ofte la ut bilder med andre i

bilderammen, også gjerne i en omfavelse. Dette kan tolkes som en mer «avhengig» representasjonen av seg selv satt opp mot den mer maskuline «uavhengige» motparten (Rose et al., 2012). Kvinner viste også en høyere forekomst av å representere seg selv som «attraktive» i sine selfier, sammenlignet med menn i studien (Rose et al., 2012). Her er det også sentralt å nevne at forskjellene i representasjonen mellom kjønnene som Rose og kolleger (2012) fant, samsvarte med kjønnsstereotypiske representasjoner.

Før vi går videre er det behov for en definisjon av konseptet som videre skal diskuteres. Cambridge Dictionary (u. å.) definerer en *stereotypi* på følgende måte: «a set idea that people have about what someone or something is like, especially an idea that is wrong». Denne definisjonen bygger under tendensen man som menneske har til å kategorisere informasjon og ofte generalisere begreper, selv om de ikke nødvendigvis er sanne. «Advisory Committee on Equal Opportunities for Women and Men» gir begrepet følgende definisjon: «Stereotypes are beliefs adopted in advance; concerning characteristics of an individual, a group or an object and emphasizing the fact of not taking into account the individual traits» (2010, s. 5). Stereotypier er med andre ord sosiale forenklinger, som både strekker seg mellom kjønn, etnisiteter, nasjonaliteter og andre deler av det sosiale samfunnet (Advisory Committee on Equal Opportunities for Women and Men, 2010). Det kan tenkes at stereotypier både bærer med seg positive og negative aspekter. På den ene siden kan stereotypier være ganske praktiske, da de bidrar til å forenkle hverdagen vår ved å hjelpe oss med å kategorisere verden rundt oss. På den andre siden kan dette føre til at individer og/eller grupper behandles urettferdig. Når det kommer til kjønnsstereotypier, har man både stereotypier som beskriver hvor menn og kvinner «er» eller «skal være», samt skjønnhetsstereotypier som beskriver hvordan menn og kvinner «bør» se ut.

Begreper som vakker, stygg, ung eller gammel har også blitt brukt til å lage kategorier av kulturelle og sosiale stereotypier (Araújo et al., 2016). Her var blant annet Downs og Harrison (1985) tidlig ute med å nevne TV og kommersiell reklame som sentrale i dannelsen og opprettholdelsen av stereotypiske skjønnhetstrekk. I dag er medielandskapet ganske mer komplekst, og debatten rundt mediernes rolle i dannelsen av skjønnhetsstereotypier har fått stor plass i offentlige debatter og blant akademikere – spesielt når det kommer til den idealiserte fremstillingen av kvinner. Nå er både massemediene og sosiale medier ansett som en stor bidragsyter når det kommer til å forsterke usunne tynne skjønnhetsidealer for kvinner (López-Guimerà et al., 2010; Aparicio-Martinez et al., 2019). Her kan det tenkes at mediene bidrar til å skape en slags bedømming av andre og videre overfører normer om en ideell skjønnhet gjennom medieinnholdet (Yan & Bissell, 2014). Slik det ble nevnt tidligere, har kvinner både en høyere sannsynlighet for å ta og redigere bilder av seg selv (Qiu et al., 2015; Dhir et al., 2016; Fox & Vendemia, 2016), noe som samsvarer med stereotypien rundt femininitet og det å være attraktiv (Rose et al., 2012). Faktisk er det meste av forskningen rundt selfier fokusert rundt unge kvinner (Murray, 2015).

Det kan tenkes at de skjønnhetsstereotypiske fremstillingene av den kvinnelige kroppen gjennom årene nå viser effekt blant kvinners praktisering av selvrepresentasjon på sosiale medier. Caso et al. (2020) sin studie viser blant annet hvordan mediebruk var positivt assosiert med internalisering av det tynne-ideal, som igjen var positivt assosiert med selvobjektivering. Denne selvobjektiveringen var videre positivt assosiert med tid brukt på sosiale medier, frekvensen av selfie-posting og selfie-redigering (Caso et al., 2020). Chae (2017) fant også at kvinner ikke engasjerer seg i sosial sammenligning og påfølgende selfie-redigering på grunn av deres misnøye med utseendet, men på grunn av

ønsket om en mer ideell nettbasert selvpresentasjon. Hvilke implikasjoner dette har på hvordan man oppfatter bilder av kvinner sammenlignet med menn på sosiale medier er et lite utforsket tema. Vil det være slik at man i møte med bilder av kvinner har en underliggende forestilling om at kvinnen på bildet mest sannsynlig er redigert, basert på forutbestemte forestillinger fra samfunnet om at kvinner oftest redigerer sitt eget utseende? Forskning på manipulasjon av bilder på sosiale medieplattformer har – slik det ble nevnt i tidligere kapittel – indikert at sammenlignet med menn er det viktigere for kvinner å dyrke et sosialt ønskelig utseende ved å redigere bilder (Fox & Vendemia, 2016). Men hvilken effekt dette har på folks oppfatning av menn og kvinners bilder på sosiale medier som redigerte eller ikke, er hittil ikke dokumentert, og bærer derfor ingen klar forventning i retning av noen sammenheng mellom de to faktorene. På bakgrunn av dette stilles følgende forskningsspørsmål:

**F1:** *Finnes det en sammenheng mellom kjønn representert i et bilde og hvor sannsynlig deltakeren mener det er at bildet er redigert?*

### **2.2.1 Skjønnhetsstereotyper blant kvinner og menn**

Akkurat hva som gjør en mann eller en kvinne stereotypisk vakker, kan være noe vanskelig å besvare sett gjennom skjønnhetsteoriens diskusjon, som både beskriver skjønnhet som subjektivt og objektivt på samme tid (Sartwell, 2022). I dette gjeldende eksperimentet vil det imidlertid være gunstig å forsøke å definere noen konkrete og mer objektive skjønnhetsstereotyper forankret i tidligere forskning på feltet. I denne delen presenteres derfor en kortfattet gjennomgang av tidligere forskning på tematikken skjønnhetstrekk og kjønn. Denne gjennomgangen vil forme grunnlaget for behandlingen av eksperimentets stimulusmateriale og videre forskningsspørsmål.

Når det kommer til ansiktstrekk blant kvinner og menn, er det ulike egenskaper som anses som attraktive. I en studie fra 2004 viser Baudouin hvilke trekk som ansees som attraktive blant kvinner, basert på en undersøkelse som sammenlignet mannlige vurderinger opp mot metriske egenskaper i kvinnelige ansikter. Attraktive kvinnelige trekk ble gjerne beskrevet som mer barnslige, slik som å ha store øyne plassert lenger ut i ansiktet, fyldige lepper, samt en liten nese og hake (Baudouin, 2004). I tillegg til dette viser Baudouin sin studie at symmetri i ansiktet er et attraktivt trekk blant kvinner. Når det kommer til det mannlige ansiktet brukte Keating sin studie fra 1985 illustrasjoner for å undersøke forholdet mellom dominans og attraktivitet. Resultatene viste at små øyne, tynne lepper, firkantede kjever og tykkere bryn var trekk som ble vurdert med en høyere grad av dominans, og også som mer attraktive (Keating, 1985). Cunningham med kolleger (1990) konkluderte med at kvinner i stor grad var tiltrukket av menn som hadde dominante og modne karakteristikk i ansiktet, slik som brede kjever, tydelige haker og relativt tynne lepper. Likevel var det slik at de mennene kvinnene synes var aller mest attraktive også hadde barnslige trekk over seg, slik som større øyne, samt at de viste sosiale trekk som et bredere smil (Cunningham, 1990). Det er også vist gjennom forskning at menn som har hårvekst i ansiktet i form av lette stubber heller enn tett eller ingen skjeggvekst, blir ansett som mest attraktive, og ansett som foretrukket for både kort- og langsiktige forhold (Neave og Shields, 2008). Det er verdt å nevne at det eksisterer en viss uenighet i forskningen når det kommer til funn som tilsier at dominerende ansiktsuttrykk hos menn ansees som tiltrekkende blant kvinner, og det kan

også tenkes at det eksisterer kulturelle forskjeller som gjør det vanskelig å overføre denne type forskning fra andre land til Norge. Perrett og kolleger (1998) fant i en studie at gjennomsnittlige mannlige ansikter som var "femininisert" (gitt fyldigere lepper og mindre nese og hake) ble oppfattet som mer attraktive av kvinner, mens mer maskuline trekk førte til negative assosiasjoner blant kvinnene, slik som at personen virket «kald» eller uærlig.

Den kvinnelige og mannlige kroppen har også sine ulikheter i stereotypiske skjønnhetsidealer. Rose og kolleger (2012) beskriver hvordan kvinner ofte sammenligner seg selv med et tynnere ideal, og i større grad føler på et press til å endre kroppene sine for å øke opplevd seksualitet eller ungdommelighet. For menn ser gjerne den ideale kroppslige fremstillingen ut som en sterk, ungdommelig, aktiv og fysisk dominerende kropp (Rose et al., 2012). Dette kommer også frem i en undersøkelse fra 1988 der Silberstein og kolleger fant at kvinner og menn hadde motsatte ønsker når det kom til kroppsvekt. Kvinner ønsket ofte å være tynnere, mens menn heller ønsket å være tyngre enn tynnere. Dette kommer også til syne i kvinner og menn sine former for ønsket selvrepresentasjon, som er knyttet opp mot eksisterende forestillinger om de vestlige kroppsidealene: ekstrem tynnhet for kvinner og muskuløsitet for menn (Hildebrandt et al., 2008).

Denne empiriske gjennomgangen av ulike skjønnhetsstereotypiske trekk blant kvinner og menn vil fungere som grunnlaget for behandling av eksperimentets stimulusmateriale. En detaljert gjennomgang av hvilke redigeringer som ble gjennomført blir presentert i metodedelene, i kapittelet «3.3.2 Stimulusmaterialet», og i tabell 1.

For å undersøke hvilken effekt skjønnstereotypiske kosmetiske redigeringer vil ha mellom de ulike kjønnene, vil både menn og kvinner avbildet i stimulusmaterialet redigeres med både maskuline og feminine skjønnstereotypiske trekk. Her er man igjen inne på et felt i forskningen hvor det er vanskelig å trekke frem klare forventninger i form av hypoteser. Det eksisterer i dag ingen forskning som har undersøkt hvilken effekt maskuline og feminine skjønnstereotypiske redigeringer har på folks evne til å gjenkjenne manipulasjoner i bilder. Likevel, forankret i tidligere definisjoner av stereotyper som en forutsatt idé eller oppfatning folk har om hvordan noen eller noe er, kan det tenkes at det å bryte en fastsatt stereotypi kan ha effekt på eksperimentets resultater. Dersom en kvinne blir redigert med maskuline skjønnstereotypiske trekk, eller en mann med feminine skjønnstereotypiske trekk, kan det tenkes at det vil være lettere for deltakeren å oppdage manipulasjon – da representasjonen av personen som er avbildet allerede bryter med etablerte idéer om hvordan kjønnnet burde se ut. Basert på dette stilles følgende forskningsspørsmål:

**F2:** *Har det en effekt på hvor sannsynlig deltakeren mener det er at bildet er redigert når et bilde redigeres i motsetning til kjønnsstereotypen?*

Dette vil være interessant å undersøke da svarene forskningsspørsmålet gir kan gi en indikasjon på hvor normalisert skjønnstereotypiske representasjoner gjennom redigering er blitt blant den generelle befolkningen. Om man her finner at kvinner redigert med maskuline stereotypiske trekk vil ha høyere sannsynlighet for å bli stemplet som «manipulert», eller motsatt (at menn som er redigert med feminine stereotypiske trekk vil ha høyere sannsynlighet for å bli stemplet som «manipulert»), kan dette gi en indikasjon på hvilken kjønnnet fremstilling som ansees som mest «naturlig» i møte med folk flest.

## 2.3 Sosial sammenligning og kognitiv dissonans

I dette kapitlet skal jeg gå igjennom teoriene om sosial sammenligning og kognitiv dissonans, og beskrive hvordan disse kan knyttes opp til sosiale medier og eksperimentet.

Det å sammenligne seg selv med andre, enten det er bevisst eller ubevisst, er et ganske kjent sosialt fenomen. Likevel var det ikke før i 1954 at Leon Festinger presenterte sin banebrytende teori om *sosial sammenligning*, der han antok at vi mennesker har en indre driv for selvevaluering, der evalueringsprosessen er forankret i sammenligning med andre individer. Sammenligningen tenderer å foregå mellom individer som er like en selv, heller enn personer som er for ulike. Festinger (1954) beskriver det slik: «If some other person's ability is too far from his own, either above or below, it is not possible to evaluate his own ability accurately by comparison with this other person» (Festinger, 1954, s. 3). Festinger (1954) påpeker derimot videre at i en prestasjonsorientert kultur slik som vår egen, er det ikke usannsynlig at et individ vil sammenligne sine evner og egenskaper med noen han/hun selv tror har evner som er litt overlegne ens egne. Dette er noe som støttes av Wheeler (1966) som beskriver en *oppadgående sosial sammenligning*, mens Wills (1981) i senere tid også redegjør for en *nedadgående sosial sammenligning*. Oppgående sosial sammenligning oppstår ofte når man sammenligner seg selv med individer som har positive egenskaper og som man anser som overlegne en selv, mens nedadgående sosial sammenligning oppstår når man sammenligner seg selv med individer med negative egenskaper og som man selv anser som underlegne (Wills, 1981; Wheeler, 1966). I denne sammenhengen vil også sammenligning oppover tjene til å sette mål og skape fremgang, mens sammenligning nedover vil bidra til å øke selvtilliten. Etter hvert begynte Festinger å innse at teorien om sosial sammenligning ikke var kompleks nok i møte med det store antallet kontekster der individer ble konfrontert med bestemte avvik i sine omstendigheter, altså et avvik mellom seg selv og personen man sammenlignet seg selv med (Cooper, 2019). På bakgrunn av dette – sammen med innflytelse fra flere teoretikere (blant annet balanseteori) – utviklet Festinger (1957) teorien om *kognitiv dissonans* som en utvidelse av sosial sammenligning (Cooper, 2019).

Kognitiv dissonans teorien stammer fra Festinger (1957) sine observasjoner om kognitive uregelmessigheter, og har siden den gang vært en sentral pilar i sosial psykologi (Cooper, 2012). Hovedprinsippene i den originale versjonen av dissonansteorien er, ifølge Cooper (2012) velkjente og enkle: «The state of cognitive dissonance occurs when people perceive that a pair of cognitions is inconsistent». Denne dissonansen oppleves som et ubehag for individet, og må derfor reduseres. Reduksjonen skjer ved å endre kognisjonen som er minst motstandsdyktig mot endring, eller ved å legge til en kognisjon for å redusere den opplevde størrelsen på avviket (Festinger, 1962). Festinger mente selv at denne «dissonanskampen» er en kamp som utfoldes inne i hodet på individet, og er derfor i seg selv en psykologisk tilstand der to kognisjoner er dissonante kun om individet selv tror de er dissonante (Cooper, 2012; Festinger, 1962). Det er altså sinnet til individet som bestemmer eksistensen av en dissonans. For å sette teorien ut i «den virkelige verden» bruker Festinger (1962) det å røyke som forklarende eksempel. En person som røyker kan være klar over at det å røyke ikke er bra fra han/hun, men velger likevel å fortsette. Her eksisterer det en dissonans (eller inkonsistens) mellom handlingen å røyke og kunnskapen om at det å røyke ikke er bra. På denne måten vil personen mest sannsynlig gjøre det han/hun kan for å rasjonalisere valget av å fortsette å røyke, for eksempel å tenke at «hvis jeg stoppet å røyke vil vekten jeg legger på meg være like

usunn, om ikke verre!»). I dette eksempelet blir en kognisjon lagt til for å dempe ubehaget ved avviket som oppleves av røykeren (Festinger, 1962).

Det er flere teorier som har implisert «selvet» i Festingers teori om kognitiv dissonans. En av disse er Aronson (1968), som kanskje også presenterer den innledende vektlegging av selvkonseptet i Festingers teori (Cooper, 2012). Aronsons teori om selvkonsistens vektlegger nettopp behovet for å involvere «selvet» når kognitiv dissonans tar sted. Selvekonsistens-teorien hevder at mennesker har et behov for å se seg selv som gode, kompetente og moralske vesener (Cooper, 2012). Drar man inn Festingers eksempel om røykeren, vil en som fortsetter å røyke selv om den vet det er helseskadelig kanskje risikere å se seg selv som mindre kompetent til å ta valg. Dette vil føre til et avvik mellom personens oppførsel og personens integritet av et bedre selv – i dette tilfellet et kompetent selv (Cooper, 2012). Aronson plasserer denne selvkonsistensen i selve hjertet av dissonans-teorien. Han mener at dersom dissonans eksisterer, eksisterer det fordi personens oppførsel ikke er konsis med personens selvkonsept, noe som skaper en dissonans som må rettes opp (Cooper, 2012).

Det er blitt gjennomført flere studier som har testet ut innvirkningen på sosial sammenligning og kognitiv dissonans-teori. I denne sammenhengen er det blant annet verdt å nevne et eldre eksperiment forankret i sosial sammenligning, der Morse og Gergen (1970) undersøkte endring i selvtillit basert på sosial sammenligning under ulike forutsetninger. Halvparten av forsøkspersonene i studien møtte tilfeldig et individ med personlige egenskaper som kunne beskrives som sosialt ønskelige, mens de resterende møtte en person som ikke hadde like sosialt ønskelige egenskaper. Her fant forskerne at den sosialt ønskelige stimuluspersonen produserte en betydelig nedgang i selvtillit, mens det motsatte forekom i møte med stimuluspersonen som hadde sosialt uønskelige egenskaper (Morse & Gergen, 1970). Videre fant de også at personer med lav selvkonsistens ble mest påvirket av tilstedeværelsen av stimuluspersonen (Morse & Gergen, 1970).

Knytter man disse teoriene opp til sosiale medier, er det ikke vanskelig å se hvordan de kan skape en ideell plattform for sosial sammenligning. Tidligere forskning som knytter Festingers (1954) sin teori opp mot sosiale medieplattformer har blant annet foreslått at Facebook kan være en plattform folk bruker med formålet om å sosialt sammenligne seg selv opp mot andre brukere (Lee, 2014), og at det å følge flere fremmede på Instagram har en negativ effekt der folk tenderer til å gjøre en oppadgående sosial sammenligning (Lup et al., 2015). I denne sammenhengen er det sentralt å trekke frem rollen til individets selvtillit som motiv når man sammenligner seg med andre, også i møte med bilder av andre på sosiale medier. Det vil her være interessant å undersøke hvilken innvirkning faktorer som sosial sammenligning kan ha på folks oppfatning av stimulusmaterialet som manipulert eller ikke. Det kan tenkes at det å sosialt sammenligne (enten det er oppadgående eller nedadgående) med bilder av andre, har en innvirkning på individets selvtillit. For at selvtilliten ikke skal lide for mye i for eksempel en oppadgående sammenligning, kan det tenkes at folk vil dempe kraften av den oppadgående sosiale sammenligningen ved å anta en enkel erkjennelse: «Bildet er bare manipulert». Basert på Festinger (1945; 1962) sin teori om sosial sammenligning og teorien kognitiv dissonans som utvidelse av denne, kan det være slik at jo høyere en deltaker vurderer den avbildede personen på en «attraktivitets-skala», desto større sannsynlighet er det for at deltakeren vil tolke bilde som manipulert. Dette kan forankres i tanken om at en «oppadgående sosial sammenligning» (Wheeler, 1966) vil føre til en dissonans – eller et avvik – mellom deltakerens selv-konsept (Aronson, 1968, referer i

Cooper, 2012) og møtet med positive egenskaper (i dette tilfellet attraktivitet) i personen avbildet. Denne tankerekken kan ha ett av to eventuelle utfall: 1) personen opplever dissonans gjennom sammenligning av individet avbildet og eget selv-konsept, og fyller dissonansen med en antakelse av manipulasjon, eller 2) personen opplever *ikke* dissonans gjennom sammenligning av individet avbildet og eget selv-konsept, og man vil ikke se noen effekt på vurdering av manipulasjon. Da det ikke eksisterer forskning på dette område som kan gi klare antydninger til sammenheng mellom de ulike faktorene, vil det her stilles følgende forskningsspørsmål:

**F3:** *Finnes det en sammenheng mellom deltakerens vurdering av personen avbildet som attraktiv og sannsynligheten for at de vurderer bildet som redigert?*

## 2.4 Signal Detecting Theory (SDT)

I dette kapittelet vil jeg presentere Signal Detecting Theory (SDT) som teoretisk perspektiv og analytisk verktøy. Her vil jeg knytte teorien opp mot tidligere forskning innen bildemanipulering, hvor teorien kan brukes til å bedre forstå hvordan «signaler» i bilder kan påvirke folks evne til å gjenkjenne redigeringer. Til slutt vil jeg forklare hvordan teorien vil bli brukt som analytisk verktøy i dette eksperimentet.

Signal Detecting Theory (SDT) ble opprinnelig utviklet av radarforskere på begynnelsen av 1950-tallet, men ble raskt anerkjent av kognitive forskere som så verdien av teorien og tilpasset den for anvendelse i menneskelig beslutningstaking (Anderson, 2015). Anderson (2015) beskriver SDT sine generelle premisser slik:

The general premise of SDT is that decisions are made against a background of uncertainty, and the goal of the decision-maker is to tease out the decision signal from background noise. SDT can be applied to any binary decision-making situation where the response of the decision maker can be compared to the actual presence or absence of the target (Anderson, 2015).

Det er spesielt i eksperimenter med formål om å kategorisere tvetydig stimuli at SDT utmerker seg som et godt analyseverktøy (Abdi, 2007). I slike stimuli-eksperimenter blir gjerne to eller flere stimulusklasser utprøvd på en observatør/deltaker, og deltakeren må selv velge en passende respons til stimulusen som blir presentert (Macmillan 2001). På bakgrunn av dette er også teorien blitt brukt av forskere for å undersøke folks evne til å identifisere originale og manipulerte bilder. Nightingale og kolleger (2017) brukte blant annet SDT for å måle om mengden manipulasjon i bildene, eller mengden «signal», hadde en innvirkning på folks evne til å gjenkjenne bildene som manipulerte eller ikke. Resultatene fra eksperimentet viste at graden av forstyrrelser i bildenes pikselstruktur faktisk hadde en effekt, der eksperimentet viste at jo tydeligere pikslene i et bilde var «tuklet» med, desto større var sannsynligheten for at deltakerne mistenkte manipulasjon i bildet (Nightingale et al., 2017). Videre fant Nightingale og kolleger at selv om deltakerne evnet å skille manipulerte bilder fra ikke-manipulerte, slet de med å finne akkurat hvor i bildet manipulasjonen var, og at folk generelt sett ikke var spesielt flinke til å skille ikke-manipulerte bilder fra manipulerte.

Det finnes også studier innen feltet bildemanipulasjon som ikke tar i bruk STD sitt teoretiske rammeverk direkte, men som likevel kan knyttes opp mot teoriens grunnleggende tankegang ved deres bruk av ulike signaler som indikatorer for manipulasjon. Her kan blant annet Fraid og Bravo trekkes frem, som i 2010 skapte en rekke datagenererte scener med ulike geometriske former for å finne ut av hvor vellykkede folk er til å bruke signaler som ofte indikerer at bilder er blitt manipulert. Forskerne endret på tre ulike former for signaler i scenene: perspektiv, lys/skygger og speilbilder (Farid & Bravo, 2010). Resultatene viste at ulike former for endring i de ulike signalene hadde innvirkning på deltakernes suksess til å gjenkjenne feil i signalene. Som eksempel på dette viser forskerne hvordan det å vurdere lyskilden i et bilde var noe deltakerne klarte godt (med 95% nøyaktighet) når skyggene presentert i bildet strakk seg i helt ulike retninger (Farid & Bravo, 2010). På den andre siden, når skyggene var mer inkonsekvente og ble kombinert av to ulike lyskilder på samme side av «rommet», kunne bare 53 % oppdage feil i lyskilden. Totalt sett fant Fraid og Bravo (2010) at folk flest sitt visuelle system later til å være «bemerkelsesverdig ute av stand» til å gjenkjenne feil i bilder gjennom de tre ulike signalene som ofte er indikasjoner på bildemanipulering.

Selve signalet i SDT står sentralt i teoriens rammeverk, der styrken av et «signal» og mengden «støy» i stimulusmaterialet har en klar sammenheng med menneskers evne til å gjøre perseptuelle beslutninger på en nøyaktig måte. Trekker man teorien tilbake til radarforskerens felt, kan en radaroperatør måtte ta beslutningen om det han/hun observerer på radarskjermen antyder tilstedeværelsen av et fly (et signal) eller om det er en indikasjon på forstyrrelser (støy) (Abdi, 2007). Denne binære beslutningstakingen mellom signal eller støy fører til fire ulike responser til en gitt stimuli, slik figur 1 viser her:

REALITY	DECISION: (PARTICIPANT'S RESPONSE)	
	Yes	No
Signal Present	Hit	Miss
Signal Absent	False Alarm (FA)	Correct Rejection

**Figur 1:** The four possible types of responses in SDT. Fra "Signal Detection Theory (SDT)" av H. Abdi, 2007. I, B. McGaw, P.L. Peterson & E. Baker (Red.), *Encyclopedia of Education* (3. Utg., s.1-9). New York: Elsevier. Hentet fra: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.143.2749&rep=rep1&type=pdf>

I eksperimenter der deltakere har som oppgave å ta beslutninger om det er et signal tilstede eller ikke, vil deltakerens interaksjon med stimulusmaterialet ha fire ulike utfall, slik vist i figuren over: 1) Treff/Hit, det er et signal tilstede og operatøren sier «ja, dette er et reelt signal», 2) Bom/Miss, det er et signal tilstede men deltakeren sier «nei, dette er ikke et reelt signal», 3) Falsk alarm/FA, her er det ingen signal tilstede men deltakeren sier likevel «ja, dette er et reelt signal», og 4) Korrekt avvisning/Correct Rejection, hvor det ikke er noen signal tilstede og deltakeren sier «nei, dette er ikke et reelt signal» (Abdi, 2007).



Nevroforskeren David Heeger (2003) beskriver videre hvordan det eksisterer to ulike komponenter når det kommer til en observatør – eller deltaker – sin evne til å bedømme stimuli. Her snakker han om den rene visuelle stimulusen som presenteres for observatøren som en komponent, mens observatørens subjektive dømmekraft, eller bedømmingsevne, vil fungere som en sekundær komponent i prosessen, også kalt en *intern respons* (Heeger, 2003). Den interne responsen er en subjektiv respons, og er unik for individet som skal tyde stimulusen foran seg. På denne måten kan man presentere samme informasjon med samme signal til to ulike individer, og likevel få ulike utfall av observasjonen på grunn av individenes differanse i egen interne respons (Heeger, 2003).

Fordelen med å bruke SDT som et måleinstrument i beslutningstaking er at teorien gir en enhetsløs måling av sensitivitet – helt uavhengig av subjektbias – som også kan sammenlignes med andre sensitiviteter over vidt forskjellige situasjoner (Anderson, 2015). Etter at kognitive forskere tilpasset SDT for bruk i undersøkelser av menneskelig beslutningstaking, har teorien blitt brukt i en rekke ulike emner. Alt fra hukommelsesforskning (f.eks Banks, 1970), radiologi og diagnostikk (f. eks Boutis et al., 2010), helt til bedømmingen av nøyaktigheten til værmeldinger (Harvey et al., 1992).

Gitt mangelen på forskning som undersøker folks evne for å oppdage bildemanipulasjoner, kan SDT tilby et svært relevant perspektiv i gjeldende eksperiment når det kommer til hvilke prosesser som finner sted når et individ bedømmer tvetydig stimuli. Slik Anderson (2015) beskrev SDT som teoretisk rammeverk, kan den tas i bruk under en hvilken som helst binær beslutningstakingssituasjon der responsen til et individ er forankret i tilstedeværelse eller fraværet av «målet». I dette gjeldende eksperimentet vil tilstedeværelse eller fraværet av et «mål» representeres av tilstedeværelsen eller fraværet av manipulasjon i bildet som presenteres for deltakerne, mens «Hit», «Miss», «False Alarm» og «Correct Rejection» (se figur. 1) vil indikere hvor vellykkede deltakerne var til å korrekt identifisere de manipulerede bildene og avvise de ikke-manipulerte. Signalet som deltakerne her søker etter er med andre ord redigeringer i bildene. SDT som teoretisk perspektiv og analytisk verktøy vil i dette eksperiment brukes for å måle deteksjonen av «signal» som en dikotom variabel (oppdaget redigering, ikke oppdaget redigering). Det vil være interessant å undersøke om den subjektive dømmekraften Heeger (2003) snakker om kan sees opp mot deltakernes tidligere erfaringer og holdninger til manipulering av bilder, samt deres bruk av sosiale medier. Disse subjektive forutsetningene kan tenkes å påvirke den interne responsen deltakerne har i møte med den visuelle stimulusen, og vil i det følgende kapitlet diskuteres i sammenheng med perspektivet *visual literacy*.

## 2.5 Visual literacy

I dette kapitlet vil jeg gå igjennom *visual literacy* som konsept. Jeg vil først definere begrepet, før jeg knytter det nærmere opp mot eksperimentets tematikk – redigering av bilder. Til slutt legger jeg vekt på hvilken rolle konseptet spiller for dette eksperimentet, der eksperimentets første hypotese vil trekkes frem fra gjennomgangen.

### 2.5.1 Visual literacy definert

Visual literacy som konsept kan være noe vanskelig å definere, da det eksisterer en rekke ulike meninger om hva konseptet faktisk innebærer (Brumberger, 2011). Oversetter man begrepet direkte, blir «visual literacy» lydene «visuell leseferdighet», en ordlyd som ikke fatter rundt begrepets betydning særlig godt. Slik Orland-Barak og Maskit (2017) beskriver visual literacy handler det ikke om leseferdigheter direkte, men heller menneskers evne til å tolke og skape mening av informasjon i form av et bilde:

As such, visual literacy extends the meaning of literacy, which commonly signifies interpretation of a written or printed text. Visual literacy is based on the idea that pictures can be “read” and that meaning can be discerned through a process of reading (Orland-Barak & Maskit, 2017, s. 12-13).

På bakgrunn av dette vil begrepet «visuell kompetanse» videre brukes som en oversettelse av «visual literacy», i et forsøk på å forbedre begrepets betydning opp mot dette gjeldende eksperimentet.

American Library Association (2011) beskriver visuell kompetanse som en rekke evner som gjør det mulig for en person å effektivt finne, tolke, evaluere, bruke og lage bilder og andre visuelle medier. Her står det også sentralt at en person som besitter god visuell kompetanse forstår de ulike kontekstuelle, etiske, estetiske og tekniske komponentene involvert i prosessen av å produsere og bruke ulikt visuelt materiale: «A visually literate individual is both a critical consumer of visual media and a competent contributor to a body of shared knowledge and culture» (American Library Association, 2011). Denne beskrivelsen utvider betydningen av begrepet ved å legge til kompetansen til «å produsere», som samsvarer med den økende trenden på feltet der evnen til å tolke visuelt materiale blir likestilt med evnen til å også produsere visuelt materiale (Brumberger, 2011).

### 2.5.2 Visuell kompetanse i en digital alder

I dag er verden av visuell kommunikasjon en helt annen enn det den var i sine tidligere dager. I dag har Photoshop og andre programvarer for redigering av fotografier, sammen med folks evne til å skape, manipulere og også spre bilder ved hjelp av digitale nettverk, gjort at den visuelle kulturen og fortolkningen av den i stor grad har endret seg (Messaris, 2012). Messaris (2012) beskriver hvordan denne sosiale endringen i samfunnet har ført til en utvikling med to store komplementære implikasjoner for visuell kompetanse:

On the one hand, digital media have given the creators of images an ever-expanding arsenal of increasingly subtle, hard-to-detect visual tools. On the other hand, however, digital networks have made it much easier for the detection of visual manipulation to be publicized and, in that sense, for visual literacy to be disseminated and shared (Messaris, 2012, s. 106)

Hvordan den økende spredningen av manipulerte bilder på nett, sammen med den økende bruken og tilgjengeligheten av manipulasjonsverktøy har påvirket folks visuelle kompetanse, er noe vi ikke har særlig mye kunnskap om (Messaris, 2012). Men selv om mye av forskningen gjort rundt manipulering av bilder og folks evne til å identifisere slike manipulasjoner ikke direkte knyttes opp mot visuell kompetanse som perspektiv, har gjerne de fleste av disse studiene med seg digitale eller visuelle ferdigheter som måleenhet, enten indirekte eller direkte. Eksempler på dette ser man blant annet i Nightingale et al. (2017) som målte effekten av visuelle erfaringer, slik som å ha interesse for fotografering opp mot deltakernes evne til å gjenkjenne manipulasjon, eller Shen et al. (2019) som undersøkte det samme med deltakernes internettferdigheter, deres erfaringer med digitale bilder og bruk av sosiale medier. Det er likevel ikke enighet på forskningsfeltet om hvilken effekt slike faktorer har på folks evne til å gjenkjenne manipulasjoner i bilder. Flere studier (slik som Nightingale et al., 2017; Farid & Bravo, 2010; Brumberger, 2011) har vist at personers digitale ferdigheter, interesse og generell bevissthet rundt redigeringer av bilder på nett ikke gjør dem flinkere til å oppdage redigeringer i bilder. Dette gjenspeiles også i generasjonen som tidligere har vært ansett som mer teknologisk kyndige enn andre, nemlig de «digitalt innfødte». Denne yngre generasjonen har, slik Brumberg (2011) beskriver, i større grad blitt utsatt for nye teknologiske verktøy, noe som også har gitt dem økte ferdigheter i møte med dagens teknologiske samfunn. I denne sammenhengen snakker man gjerne også om generasjonen «digitalt innfødte» som en generasjon med en høyere grad av visuell kompetanse (Brumberger, 2011). Brumberger sin undersøkelse gikk gjennom deltakernes oppfatning av egne teknologiske ferdigheter, egen bruk av visuelle teknologier og hvordan de selv ville vurdere egen evne i å tolke visuell informasjon gjennom video eller bilder. Resultatene fra denne undersøkelsen viste en lavere grad av visuell kompetanse enn forventet (Brumberger, 2011). Videre reflekteres også disse funnene i deltakernes evne til å skille mellom redigert og ikke-redigerte bilder. I undersøkelsen ble deltakerne eksponert for to ulike bilder, der ett av dem var manipulert og det andre ikke var det. Her trodde 66% av de «digitalt innfødte» at det manipulerede bildet var autentisk, mens hele 80% svarte at det ikke-manipulerte bildet faktisk var manipulert (Brumberger, 2011).

Schetinger med kolleger gjennomførte i 2017 et eksperiment med mål om å finne ut hvor vanskelig det er for en gjennomsnittlig person å bedømme om et bilde har blitt redigert eller ikke. For å oppnå dette brukte de bilder fra kjente offentlige rettsmedisinske datasett, og resultatene fra denne studien viste at bare 58% av bildene ble korrekt klassifisert som enten redigerte eller ikke-redigerte, der bare 46% av de manipulerte bildene var identifisert som sådan (Schetinger et al., 2017). Med andre ord ble mer enn halvparten av alle redigerte bilder ikke oppdaget av deltakerne i eksperimentet. Når det kom til effekten av andre underliggende faktorer på eksperimentets deltakere, viste det seg at deltakernes tidligere erfaring med digitale bilder var den eneste faktoren som faktisk hadde en positiv innvirkning på evnen til å gjenkjenne manipulasjoner, mens utdanning ikke viste noen betydelig effekt (Schetinger et al., 2017).

I et nyere storskala online-eksperiment fra 2019 fant Shen og kolleger på den andre siden at faktorer som internetferdigheter, bruk av sosiale medier, tidligere erfaringer med digitale bilder og det de selv kaller en «pro-problematisk» holdning alle var sentrale egenskaper knyttet til folks evne til å korrekt vurdere bilder på nett som troverdige eller ikke. Ved å bruke 6 falske bilder som avbildet dyr, personer, landskap og en krigsscene hadde Shen og kolleger sitt eksperiment som mål å undersøke hvordan mennesker vurderte slike bilder som troverdige eller ikke, basert på hvilke nettplattformer bildene ble presentert i (2019). I tillegg til å presentere bildene gjennom ulike nettplattformer ble også bildene tildelt ulike mediekilder, eksempelvis Facebook/Twitter-kontoen til Bill Gates eller nyhetssiden New York Times for å undersøke hvilken innvirkning dette ville ha på resultatene. Hvilken nettplattform og mediekilde bildene ble presentert igjennom viste seg likevel ikke å ha noen innvirkning på deltakernes evne til å vurdere bildene, da det heller var deltakernes personlige egenskaper som sto sentralt i denne bedømmingen.

Forskningen som tar for seg manipulering av bilder og folks erfaringer/evner med visuelle medier har med andre ord motstridende funn når det kommer til hvordan disse faktorene spiller inn på deltakernes bedømmingsevne. Om det har vært en utvikling på feltet eller ikke er vanskelig å si, men det er likevel interessant å se at de nyere studiene på feltet (Schetinger et al., 2017, Shen et al., 2019) viser en økt effekt av faktorer som minner om visuell kompetanse, satt opp mot eldre studier som ser på det samme (Farid & Bravo, 2010; Brumberger, 2011; Nightingale et al., 2017). Kan det her tenkes at den økende spredningen av manipulerede bilder, sammen med den økende trenden med å publisere funn av visuell manipulasjon og på den måten spre visuell kompetanse (Messaris, 2012), har hatt en innvirkning på folks evne til å skille mellom redigerte og ikke-redigerte bilder? Hvis dette er tilfellet, vil det være sannsynlig at dette også gjenspeiles i brukere av sosiale medier, spesielt de som bruker sosiale medier jevnlig. Forankret i definisjonen på visuell kompetanse, for eksempel kontekstuell forståelse og generelle skapende visuelle ferdigheter, kan det argumenteres for at de som oppholder seg oftere på sosiale medier og/eller har et større forbruk av redigeringsverktøy, også vil ha en høyere grad av visuelle ferdigheter når det kommer til identifisering av bilder som redigerte eller ikke. Vi ser i dag at redigeringer av bilder som spres på sosiale medier ofte oppdages, og gjerne også av andre brukere av sosiale medier som – slik Messaris (2012) beskriver – sprer dette videre på digitale plattformer. Det finnes blant annet flere Instagram-kontoer dedikert til å spre bevis for hvordan kjendiser og influencere manipulerer bilder, blant annet Instagram-brukeren @Celebface<sup>1</sup> med 1,3 millioner følgere. Basert på denne kontekstuelle gjennomgangen og tidligere forskning på feltet trekkes følgende hypoteser:

**H 1:** *Mer erfaring med sosiale medier (h1 a: bruk av sosiale medier og h1 b: deling på sosiale medier) vil ha en positiv effekt på evnen til å skille mellom redigerte og ikke-redigerte bilder.*

**H 2:** *Tidligere erfaring med å selv redigere bilder før de legges ut på sosiale medier, vil ha en positiv effekt på evnen til å skille mellom redigerte og ikke-redigerte bilder.*

---

<sup>1</sup> @Celebface sin Instagrambruker: <https://www.instagram.com/celebface/>

Selv om alder i dette eksperimentet i hovedsak fungerer som en kontrollvariabel, forventes det likevel å finne en sammenheng mellom alder og deltakernes evne til å korrekt gjenkjenne manipulasjoner. Dette er på bakgrunn av at erfaringer med sosiale medier, bruk av digitale verktøy slik som bildemanipuleringsverktøy og en generell høyere kjennskap og bruk av digitale plattformer er forbeholdt samfunnets yngre generasjoner. Denne antakelsen gjenspeiles også iblant annet Schetinger og kolleger (2017) sitt forskningsprosjekt, som fant at deltakerne i aldersgruppen 24 til 30 år hadde størst suksess med å gjenkjenne manipulasjoner i bilder. Her fant også forskerne at deltakernes suksess sank med den økende alderen.

## 3. Metode

### 3.1 Valg av metode

Målet for eksperimentet er å undersøke hvor godt folk evner å identifisere redigeringer i bilder som publiseres på sosiale medier, og hvilke faktorer har innflytelse på dette. Forankret i dette ble populasjonsbasert survey-eksperiment valgt som metode. Med dette menes et eksperiment der forskeren kontrollerer en randomisert fordeling av deltakere til ulike varianter av den uavhengige variabelen, med hensikt av å observere effekten dette har på den avhengige variabelen (Mutz, 2011). I slike populasjonsbaserte eksperimenter der deltakere tilfeldig blir tildelt ulike eksperimentelle forhold, vil det ofte være mer praktisk at det representative utvalget for undersøkelsen ikke trenger å møte opp i et laboratorium eller en annen form for felles fysisk lokasjon (Mutz, 2011), noe som gjør bruk av digitale spørreskjema ideelt. I survey-eksperimenter kan Internett nå ut til et større og mer mangfoldig utvalg, noe som i utstrakt betydning vil si et bredere utvalg av den norske befolkningen (Hergueux & Jacquemet, 2015). I tillegg gjør Internett det mulig å presentere visuell stimuli enkelt til et bredt utvalg av mennesker (Mutz, 2011,).

Denne metodiske fremgangsmåten er i større grad knyttet opp mot eksperimentet enn spørreundersøkelsen: «By design, population-based experiments are experimental studies drawing on the power of random assignment to establish unbiased causal inferences» (Mutz, 2011, s. 3). På samme tid blir eksperimentet presentert til et tilfeldig representativt utvalg av interessepopulasjonen, akkurat på samme måte som man administrerer en vanlig spørreundersøkelse. På denne måten drar dette eksperimentet nytte av de beste kvalitetene fra spørreskjema som datainnsamlingsmetode og det eksperimentelle designet (Mutz, 2011). Det eksperimentelle designet lover å finne en faktisk årsakssammenheng mellom faktorer slik som type endinger gjort i bildet, eller «årsak» (X), og personers evne til å identifisere dette «effekt» (Y) (Ringdal, 2016). For å kunne fastslå en slik årsakssammenheng må for det første de to variablene samvariere, for det andre må årsaken overgå effekten av tid og for det tredje må forholdet mellom årsak og effekt ikke være mulig å forklare gjennom en tredje variabel (Mutz, 2011). Man kan sikre seg mot andre innvirkende faktorer gjennom randomisering, altså å la det være opp til tilfeldighetene om deltakere havner i kontroll eller eksperimentgruppe. Dette bidrar til å sikre at forskjellene mellom gruppene er opp til tilfeldighetene (Ringdal, 2016).

Selv om populasjonsbaserte survey-eksperimenter ikke er et helt nytt fenomen, har den teknologiske utviklingen i spørreundersøkelsesforskning, sammen med utviklingen av innovative teknikker i eksperimentelle design, gjort det mulig for flere ulike disipliner å ta i bruk mer komplekse og metodisk sofistikerte populasjonsbaserte survey-eksperimenter (Mutz, 2011). Ulike programvarer for å utvikle online-spørreskjemaer har gjort det mulig å gjennomføre eksperimenter mye raskere, enklere og billigere satt opp mot før hvor man var avhengig av å gjennomføre slike prosjekter i papirform.

### 3.2 Utvalg og rekruttering

For å undersøke hvordan personer vurderer bilder på sosiale-medieplattformer som ikke-redigerte eller redigerte, har jeg gjennomført et eksperiment ved hjelp av online spørreskjema. Dataene ble samlet inn i april 2022, og inkluderer deltakere i ulike aldersgrupper (23 til 90 år) med ulike demografiske forutsetninger i Norge. I følge Mutz (2011) bør populasjonen som representeres gjennom utvalget være representativt for den populasjonen forskningen har som intensjon å overføre sine funn til. Da eksperimentet har som intensjon å undersøke et representativt utvalg av hele Norges befolkning, ansees aldersspennet her som representativt for eksperimentets tiltenkte populasjonsutvalg. Det var ikke mulig for deltakere under 18 år å delta i undersøkelsen. Det ble søkt om tillatelse fra NSD (Norsk Senter for Forskningsdata) om å gjennomføre eksperimentet. I og med at eksperimentet ikke inneholder noen form for personopplysninger (heller ikke IP-adresse), konkluderte NSD med at en videre behandling av søknaden ikke var nødvendig og eksperimentet kunne gjennomføres slik det ble beskrevet.

For å rekruttere deltakere ble spørreundersøkelsen publisert på de sosiale medieplattformene Instagram og Facebook, gjennom en nettbasert kampanje med et budsjett på 1500 kr. Kampanjen ble kjørt med hensikt om å få flest mulig deltakere i undersøkelsen, og inneholdt derfor både bilde og en kortfattet tekst: «Hvordan opplever DU bilder på sosiale medier? Ta undersøkelsen!». Til sammen ble 18 966 Facebook- og Instagrambrukere eksponert for kampanjen. Kampanjen ble kjørt fra en nylig opprettet Facebookbruker som ikke inneholdt noen form for personlig informasjon, med informasjon om at spørreundersøkelsen var tilknyttet et NTNU-mastergradsprosjekt og kort om hva undersøkelsen dreide seg om. For å nå ut til så mange deltakere som mulig, ble det brukt ulike fire ulike reklamebilder i kampanjen (figur 2) slik at plattformenes algoritmer selv ville velge ut bildene som sterkest ville appellere til brukerne – i et håp om å tiltrekke så mange deltakere som mulig. Under utformingen og publiseringen av kampanjen ble det spesifisert at kampanjen skulle nå ut til en populasjon som oppholdt seg i Norge, og som var mellom 18 og 90 år. Videre spesifikasjoner ble ikke gitt da det var ønsket at kampanjen skulle nå ut til så mange som mulig innenfor disse begrensingene.

**Figur 2:** Reklamebilder for kampanjen på Facebook og Instagram.



Det var 823 deltakere som startet på undersøkelsen, og av disse var det 522 som gjennomførte hele undersøkelsen. Deltakerne ble ikke tvunget til å svare på alle spørsmålene i spørreskjemaet, slik at det vil være varierende nummer av deltakere i de ulike analysene gjennom eksperimentets analysedel. Mer informasjon om dette finnes under kapittelet «3.6 Analysemetode».



### 3.3 Eksperimentets design

Eksperimentets design er blitt konstruert ved å lage fire ulike versjoner av samme nettbaserte spørreundersøkelse, der forskjellen mellom dem er utvalget av bilder som ble presentert. De ulike versjonene ble slått sammen til en helhetlig undersøkelse der deltakerne tilfeldig ble presentert for en av versjonene innad i undersøkelsen, mens de tre andre versjonene forble skjult. På denne måten ville det kun være én versjon tilgjengelig for hver enkelt deltaker, basert på hva klokken til deltakeren var når de startet undersøkelsen. Innad i en av undersøkelsens versjoner blir deltakeren presentert for tre ulike manipulerede bilder som fungerer som eksperimentets stimulusmateriale, samt et umanipulert originalbilde som fungerer som kontrollgruppe. I sin helhet blir deltakerne presentert for fire ulike bilder. I klassiske eksperimentelle design blir analyseenheter enten utsatt for en eksperimentell faktor – eller en «behandling» – mens en kontrollgruppe ikke utsettes for behandlingen (Ringdal, 2016). I gjeldene eksperiment har derimot hver enkelt deltaker både vært innom en kontrollenhet, nemlig bildet i original versjon, og tre eksperimentelle faktorer presentert gjennom bildenes ulike redigerte versjoner.

For å undersøke deltakernes evne til å gjenkjenne redigeringer i bilder, ble deltakerne først presentert for utvalget av undersøkelsens stimulusmateriale og bedt om å bedømme hvor sannsynlig de selv synes det er at bildet er «redigert» eller ikke. For å undersøke tilstedeværelsen av kognitiv dissonans ble deltakerne deretter bedt om å vurdere det samme utvalget av bilder på nytt, med spørsmål om hvor attraktive de selv vurderte personene på bildene til å være.

#### 3.3.1 Spørreundersøkelsens struktur

Spørreundersøkelsen (se vedlegg 1) ble utført på nett via datamaskin, nettbrett eller smart-telefon, gjennom SurveyLegend.com. Denne nettsiden ble valgt som design og distribusjonsplattform for undersøkelsen på bakgrunn av nettsidens mulighet til å fremstille bilder i et større format satt opp mot andre designsider for spørreundersøkelser. Undersøkelsens varighet var mellom 10-15 minutter, avhengig av hvor raske deltakerne var til å svare. Før deltakerne startet på undersøkelsen, fikk de en svært kort introduksjon av hva undersøkelsen dreide seg om. Denne introduksjonen inkluderte ikke informasjon om at undersøkelsen omhandlet en test av deltakernes evne til å gjenkjenne redigeringer i bilder, for å unngå at deltakerne skulle føle de ble testet i noe. Spørreundersøkelsen var helt anonym, og ingen informasjon som kunne spores tilbake til deltakerne personlig ble lagret, heller ikke IP-adresse. Siden undersøkelsen var frivillig, hadde deltakerne mulighet til å avslutte undersøkelsen før den var fullført. For å hindre at mange hoppet av midt i kunne deltakerne se hvor langt de hadde kommet i undersøkelsen gjennom en progresjonsbar nederst på siden. Det ble også inkludert «oppmuntrende» kommentarer underveis, for eksempel: «Hei hvor det går, du er nå ferdig med første del av undersøkelsen!».

Slik Estellat et al. (2009) beskriver det, øker man sannsynligheten for å oppnå basiske likheter mellom de ulike eksperimentgruppene og unngår «seleksjonsbias» om man randomiserer deltakerne i et eksperiment. SurveyLegend.com tilbød dessverre ikke et

alternativ for å tildele deltakere tilfeldig i grupper, så dette måtte jeg finne et annet alternativ for å gjøre. For å sikre randomisering av deltakerne mellom undersøkelsens ulike versjoner, eller ulike former for eksperimentell behandling, ble derfor deltakerne fordelt basert på hva klokken var når de startet undersøkelsen. Deltakerne ble bedt om å velge det tids-alternativet som passet for dem der og da, med spørsmål om minuttene i timen var mellom «hel og kvart over», «kvart over og halv», «halv og kvart på» eller «kvart på og hel». Om deltakeren skulle falle mellom noen av disse alternativene ble han/hun bedt om å velge et alternativ selv.

**Figur 3:** utdrag fra spørreskjema, ulike klokkesekvenser



Videre herfra kan spørreundersøkelsens struktur deles opp i fire ulike deler. Den første delen er en presentasjon av undersøkelsens fire ulike bilder (stimulusmaterialet), med spørsmål til deltakerne om hvor sannsynlig de tror det er at bildet er redigert (på en skala fra 1-5). Del to omhandler hvor vidt deltakeren selv ville vurdert personene avbildet som attraktiv eller ikke (på en skala fra 1-5). Den tredje delen av undersøkelsen har fokus på deltakerens egen erfaring og bruk av sosiale medier. Denne delen tar for seg hvordan deltakerne selv bruker og poster bilder på sosiale medier, og om de selv tidligere har redigert/gjort endringer på bilder før de har blitt publisert.

Spørreundersøkelsens siste del omhandler deltakerens sosiodemografi, der det blir spurt etter informasjon slik som deltakerens alder, kjønn og utdanning. Avslutningsvis følger også et spørsmål om hvor kompetent deltakeren selv mener de er til å bedømme om bilder er redigerte/endret på eller ikke.

Det ble presentert en forklarende tekst i delen hvor deltakerne skulle bedømme bildene som redigerte eller ikke, med ulike eksempler på hvordan folk flest kan redigere bilder før de legger dem ut på sosiale medier (se vedlegg 1). Her var det sentralt å presentere en tekst som var beskrivende nok til at deltakerne forsto hva det var snakk om i kontekst med «redigeringer», samtidig som teksten ikke skulle være for ledende og påvirke deltakernes besvarelse. Det er også verdt å nevne at realismen i undersøkelsen ble påvirket av at deltakerne ble bedt om å se etter manipulasjoner i bildet, noe som fører til en oppmerksomhet rundt identifisering av manipulasjoner som kanskje ikke er tilstede i «det virkelig liv» når man møter bilder på sosiale medier. Det kan tenkes at dette har påvirket undersøkelsen ytre validitet – mer om dette i kapitlet «3.7.1 Indre validitet».

### 3.3.2 Stimulusmateriale

I denne delen vil jeg gå igjennom eksperimentets bilder som stimulusmateriale, hvordan materialet ble innhentet til eksperimentet og hvordan de ble behandlet/redigert. Her vil alle former for endringer i bildene gjennom Facetune bli listet opp i detalj.

Under beslutningen av hvor mange bilder som skulle inkluderes i undersøkelsen var det et par faktorer som ble vurdert. Her var det sentralt at bildene verken skulle gjøre undersøkelsen for langstrekkelig eller mangelfull. For mange bilder kunne påvirke deltakeres villighet til å delta eller fullføre undersøkelsen, mens for få bilder kunne føre til en mangelfull datainnsamling (Nardi, 2014). Fire bilder ble valgt som et ideelt antall. Bildene ble innhentet fra frivillige som jeg kjenner personlig. Av de fire ulike bildene er to av dem menn, og to av dem kvinner. Dette er for å nå undersøkelsens mål om å undersøke innflytelsen av den avbildete personens kjønn og deltakernes evne til å gjenkjenne redigeringer. Deltakerne fikk instruksjoner om at bildet skulle være helt originalt og umanipulert, og skulle ikke portrettere andre personer enn dem selv. For å sørge for at bildene var en realistisk representasjon av bilder på sosiale medier var det sentralt at deltakerne ikke tok et bilde kun med formål å bruke det i eksperimentet. Deltakerne ble derfor også bedt om å bruke et bilde de har tatt fra før, og som de selv kunne ha lagt ut på sine sosiale-medieplattformer.

I sin helhet består prosjektet av 16 ulike versjoner av disse fire bildene, varierende mellom fire ulike versjoner av hvert bilde: bildets originale utgave, bildet redigert med feminine skjønnhetsstereotypiske endringer, bildet redigert med maskuline skjønnhetsstereotypiske endringer, samt bildet redigert med lys/kontrast (se tabell. 1). Dette vil si at hvert bilde ble utsatt for de ulike redigeringene uavhengig av kjønn, der også menn ble behandlet med feminine skjønnhetsstereotypiske redigeringer, og kvinner behandlet med maskuline skjønnhetsstereotypiske redigeringer, slik vist i tabell 1. Disse bilde-versjonene ble deretter satt sammen i fire ulike sammensetninger, fordelt på spørreundersøkelsens ulike versjoner, herav: versjon A, versjon B, versjon C og versjon D (se tabell. 2), slik at hver deltaker ble utsatt for et originalt bilde og tre manipulererte bilder med ulike manipulasjoner hver.

De ulike formene for redigering er basert på den teoretiske gjennomgangen presentert i kapitlet «2.2 Kjønn og stereotyper». Kvinnelige trekk vil redigeres med barnslige karakteristikk: større øyne, fyldigere lepper, mindre hake og liten nese. Her vil også kjevepartiet gjøres mindre/småere. Videre vil rynker glattes ut, urenheter vil fjernes og kroppen redigeres mindre eller slankere. Maskuline trekk vil redigeres med modne karakteristikk, som brede kjever, tydelige haker og tynne lepper, tykkere øyenbryn, mindre øyne og en bredere/sterkere kroppsbygning. Her vil ikke rynker glattes ut, men urenheter i huden vil bli fjernet. Alle redigeringer som er gjort på eksperimentets stimulusmateriale er gjennomført via Facetune applikasjonen på smarttelefon. Dette er for å gjøre redigeringene så «virkelighetsnære» som overhodet mulig, da Facetune – slik det blir beskrevet innledningsvis – er en svært populær redigeringsapplikasjon som er foretrukket av mange for å redigere på eget utseende (Solon, 2019).

Figur. 4 presenterer fire eksempler fra de 16 ulike redigerte bildene, der bilde nr. 1 er et eksempel på original fremstilling, bilde nr. 2 er et eksempel på redigering med lys/kontrast, bilde nr. 3 et eksempel på feminin skjønnhetsstereotypisk redigering og bilde nr. 4 et eksempel på maskulin skjønnhetsstereotypisk redigering. Fullstendig oversikt over alle 16 bildeversjoner finnes i vedlegg 2.

**Figur 4:** Eksempelutdrag fra eksperimenteres 16 ulike bilde- og redigeringskombinasjoner.



**Bilde nr. 1** – Feminin skjønnhetsstereotypisk eksempel



**Bilde nr. 2** – Lys/kontrast eksempel



**Bilde nr. 3** – Maskulin skjønnhetsstereotypisk eksempel



**Bilde nr. 4** – Original eksempel

Tabell 1: oversikt over redigeringer i stimulusmaterialet

Bilde	Original	Feminine skjønnhetsstereotypiske redigeringer	Maskuline skjønnhetsstereotypiske redigeringer	Lys/kontrastendringer
Bilde 1 (kvinne)	Ingen redigeringer	<ul style="list-style-type: none"> <li>- Mindre kjeve bredde (100%)</li> <li>- Større øyne (15%)</li> <li>- Nese: Mindre nese (70%) Smalere nese (40%)</li> <li>- Lepper: Større lepper (50%), bredere lepper (60%)</li> <li>- Glatte ut fine linjer for et yngre uttrykk</li> <li>- Fjerne urenheter</li> </ul> <p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda smalere kjeve</li> <li>- Mindre hake</li> <li>- Smalere midje</li> <li>- Mindre armer og skuldre</li> </ul>	<ul style="list-style-type: none"> <li>- Bredere kjeve (-100%)</li> <li>- Bredere smil (60%)</li> <li>- Mindre øyne (-25%)</li> <li>- Tykkere øyenbryn (60%)</li> <li>- Mindre lepper (100%)</li> <li>- Mørkere farge på øyenbryn (80%)</li> <li>- Fjerne urenheter</li> </ul> <p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda bredere kjeve</li> <li>- Bredere armer og skuldre</li> <li>- Bredere hakeparti</li> <li>- Bredere hals</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrast opp (40%)</li> <li>- Lysstyrke opp (35%)</li> <li>- Fargemetning opp (25%)</li> <li>- Lyse områder opp (60%)</li> <li>- Øke mørke områder (30%)</li> <li>- Fargetemperatur varmere (20%)</li> </ul>
Bilde 2 (kvinne)	Ingen redigeringer	<ul style="list-style-type: none"> <li>- Mindre kjeve bredde (100%)</li> <li>- Større øyne (20%)</li> <li>- Nese: Mindre nese (60%), Smalere nese (20%)</li> <li>- Lepper: Større lepper (55%), bredere lepper (50%)</li> <li>- Glatte ut fine linjer for et yngre uttrykk</li> <li>- Fjerne urenheter</li> </ul> <p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda smalere kjeve</li> <li>- Mindre hake</li> <li>- Mindre armer og skuldre</li> <li>- Mindre overkropp</li> <li>- Smalere hals</li> </ul>	<ul style="list-style-type: none"> <li>- kjeve bredere (-100%)</li> <li>- Bredere smil (75%)</li> <li>- Mindre øyne (-25%)</li> <li>- Tykkere øyenbryn (60%)</li> <li>- Lepper: Mindre lepper (-100%), smalere lepper (-40%)</li> <li>- Mørkere farge på øyenbryn (80%).</li> <li>- Fjerne urenheter</li> </ul> <p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda bredere kjeve.</li> <li>- Bredere armer og skuldre</li> <li>- Bredere hakeparti</li> <li>- Bredere hals</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrast opp (40%)</li> <li>- Lysstyrke opp (35%)</li> <li>- Fargemetning opp (20%)</li> <li>- Lyse områder opp (35%)</li> <li>- Øke mørke områder (-40%)</li> <li>- Fargetemperatur varmere (15%)</li> </ul>

Bilde 3 (mann)	Ingen redigeringer	<ul style="list-style-type: none"> <li>- Mindre kjeve bredde (100%)</li> <li>- Større øyne (40%)</li> <li>- Nese: Mindre nese (60%), Smalere nese (50%)</li> <li>- Større lepper (55%)</li> <li>- Glatte ut fine linjer for et yngre uttrykk</li> <li>- Fjerne urenheter</li> </ul>	<ul style="list-style-type: none"> <li>- Kjeve bredere (-100%)</li> <li>- Bredere smil (50%)</li> <li>- Mindre øyne (-75%)</li> <li>- Mindre lepper (-50%)</li> <li>- Tykkere øyenbryn (150%)</li> <li>- Mørkere farge på øyenbryn (30%)</li> <li>- Fjerne urenheter</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrast opp (70%)</li> <li>- Lysstyrke opp (50%)</li> <li>- Fargemetning opp (30%)</li> <li>- Lyse områder opp (55%)</li> <li>- Øke mørke områder (-25%)</li> <li>- Fargetemperatur kaldere (-30%)</li> </ul>
		<p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda smalere kjeve</li> <li>- Smalere midje</li> <li>- Mindre kake</li> <li>- Mindre armer og skuldre</li> <li>- Smalere hals</li> </ul>	<p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda bredere kjeve.</li> <li>- Bredere brystkasse</li> <li>- Bredere armer og skuldre</li> <li>- Litt smalere mage for å få frem et muskuløst uttrykk</li> <li>- Bredere hakeparti</li> <li>- Bredere hals</li> <li>- Enda tynnere underlepper</li> </ul>	

Bilde 4 (mann)	Ingen redigeringer	<ul style="list-style-type: none"> <li>- Mindre kjeve bredde (100%)</li> <li>- Større øyne (40%)</li> <li>- Nese: Mindre nese (60%), Smalere nese (60%).</li> <li>- Større lepper (15%)</li> <li>- Glatte ut fine linjer for et yngre uttrykk</li> </ul>	<ul style="list-style-type: none"> <li>- Kjeve bredere (-100%)</li> <li>- Bredere smil (50%)</li> <li>- Mindre øyne (-75%)</li> <li>- Mindre lepper (-50%)</li> <li>- Tykkere øyenbryn (150%)</li> <li>- Mørkere farge på øyenbryn (30%).</li> <li>- Fjerne urenheter</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrast opp (70%)</li> <li>- Lysstyrke opp (50%)</li> <li>- Fargemetning opp (30%)</li> <li>- Lyse områder opp (55%)</li> <li>- Øke mørke områder (-25%)</li> <li>- Fargetemperatur kaldere (-30%)</li> </ul>
		<p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda smalere kjeve.</li> <li>- Mindre hake</li> <li>- Smalere midje</li> <li>- Mindre armer og skuldre</li> <li>- Smalere hals</li> </ul>	<p><b>Reshape verktøy:</b></p> <ul style="list-style-type: none"> <li>- Enda bredere kjeve.</li> <li>- Bredere brystkasse</li> <li>- Bredere armer og skuldre</li> <li>- Litt smalere mage for å få frem et muskuløst uttrykk</li> <li>- Bredere hakeparti</li> <li>- Bredere hals</li> </ul>	

Da undersøkelsens mest sentrale aspekt er bildenes ulike versjoner og deltakernes evne til å gjenkjenne redigeringene, var redigeringen av bildene en stor del av utforming av undersøkelsen. Tabellen over viser de ulike redigeringene hvert enkelt bilde ble behandlet med. Under redigeringsprosessen var det sentralt å på den ene siden behandle hvert bilde som en individuell enhet, der redigeringene som ble gjort måtte tilpasses bildet og personen avbildet på lik linje det ville blitt gjort i «virkeligheten». På den andre siden var det også viktig å utsette bildene for så lik behandling som overhodet mulig, slik at ikke variansen av eksperimentell behandling på de ulike bildene ble for stor. Redigeringene er basert på stereotypiske skjønnhetstrekk beskrevet tidligere.

### 3.3.3 Treatment-grupper

Slik det blir beskrevet tidligere, ble deltakerne fordelt mellom spørreundersøkelsens ulike versjoner på bakgrunn av hva klokken var når deltakerne startet undersøkelsen. På denne måten ble deltakerne plassert i undersøkelsens ulike «treatment-grupper» og henholdsvis presentert for følgende eksperimentell behandling:

**Tabell 2:** oversikt over spørreskjemaets ulike treatment-grupper

<b>Gruppe A</b> (spørreskjemaversjon 1 – kl var mellom «hel og kvart over»)	Bilde 1, original	Bilde 2, lys/kontrast	Bilde 3, feminin skjønnhetsstereotypisk	Bilde 4, maskulin skjønnhetsstereotypisk
<b>Gruppe B</b> (spørreskjemaversjon 2 – kl var mellom «kvart over og halv»)	Bilde 1, maskulin skjønnhetsstereotypisk	Bilde 2, original	Bilde 3, lys/kontrast	Bilde 4, feminin skjønnhetsstereotypisk
<b>Gruppe C</b> (spørreskjemaversjon 3 – kl var mellom «halv og kvart på»)	Bilde 1, feminin skjønnhetsstereotypisk	Bilde 2, maskulin skjønnhetsstereotypisk	Bilde 3, original	Bilde 4, lys/kontrast
<b>Gruppe D</b> (spørreskjemaversjon 4 – kl var mellom «kvart på og hel»)	Bilde 1, lys/kontrast	Bilde 2, feminin skjønnhetsstereotypisk	Bilde 3, maskulin skjønnhetsstereotypisk	Bilde 4, original

I eksperimenter er randomiseringen av deltakerne mellom eksperimentets ulike grupper essensiell for å motvirke seleksjons bias (Estellat et al., 2009). Når det kommer til denne randomiseringen, er det to sentrale elementer man må ta til ettertanke: 1) generering av en sekvens og 2) implementering av den (Estellat et al., 2009). Med dette mener forfatteren at en tilstrekkelig randomisering skjer om man som forsker lager en tilfeldig talltabell eller en digital tallgenerator som inndeleder deltakerne i ulike grupper, mens det man kaller for «kvasi-tilfeldige» metoder, slik som å bruke fødselsdato eller forbokstav

på navn, ikke ansees som tilstrekkelige (Estellat et al., 2009). Dette er fordi de bærer med seg en forutsigbarhet, noe som kan åpne opp for en korrelasjon mellom sekvensen (fødselsdato eller første forbokstav) og utfallet av eksperimentet (Estellat et al., 2009). På grunn av manglende mulighet til å lage en totalt randomisert sekvens innad i SurveyLegend.com sine rammer, ble klokkeslett valgt som en alternativ kvasitilfeldig metode. Dette vil, slik Estellat og kolleger (2009) beskriver, ikke være en ideell form for randomisering av gruppene, men vil fungere som en gylden middelvei mellom en tilstrekkelig tilfeldighet og kvasi-tilfeldighet metode.

Deltakerne er fordelt mellom de ulike treatment-gruppene på følgende måte:

**Gruppe A:** 24% (201 deltakere)

**Gruppe B:** 27% (223 deltakere)

**Gruppe C:** 22% (185 deltakere)

**Gruppe D:** 26% (214 deltakere)

### 3.4 Variabler

I følgende avsnitt vil de ulike variablene i eksperimentet oppsummeres.

#### 4.5.1 Avhengig variabel, redigert/ikke redigert

Etter at de fire ulike grupperingene av deltakere ble eksponert for ulike redigerte bilder, ble de stilt spørsmålet: «Hva tror du, hvor sannsynlig er det at bildet er blitt redigert?». Deltakernes svar ble målt på en skala fra 1, «svært lite sannsynlig» til 5, «helt sikkert redigert». Denne variabelen fungerer som undersøkelsen hoved-avhengige variabel.

Å måle deltakernes antakelse av bildet som redigerte/ikke-redigerte gjennom en skala gir en høyere andel av informasjon i dataene, samt at den graderte skalaen representerer fenomenet godt gjennom å også ta innover seg muligheten for usikkerhet blant deltakerne.

#### 3.4.1 Uavhengige variabler

Med unntak av variabelen «attraktivitet» hvor bildet varierte, fikk alle deltakerne samme spørsmål for å måle de uavhengige variablene, også uavhengig av hvilken treatment-gruppe de ble plassert inn i. Under følger en beskrivelse av eksperimentets ulike uavhengige variabler.



## **Attraktivitet**

Variabelen «attraktivitet» måler hvor attraktiv deltakeren selv mener personene avbildet er. Her blir deltakerne eksponert for det samme utvalget av bilder de tidligere har blitt bedt om å bedømme som redigert/ikke redigerte, og blir gitt følgende instruks: «Vennligst vurder på en skala fra 1 til 5 hvor attraktiv du selv synes personen på bildet er». Denne variabelen måles ved hjelp av en skala fra 1 (ikke særlig attraktiv), 2, 3 (ganske attraktiv), 4 og 5 (veldig attraktiv). Det var viktig å knytte denne variabelen sammen med deltakerens treatment-gruppe, da det står sentralt at deltakerne bedømmer attraktivitetsnivået på det samme utvalget av bilder de tidligere ble utsatt for (under vurdering av hvor sannsynlig det er for at bildet er redigert eller ikke). Variabelen er inkludert i eksperimentet for å undersøke hvilken sammenheng deltakernes persepsjon av personenes utseende har med hvor vidt de selv bedømmer bildene som redigerte eller ikke, for å undersøke om det eksisterer en kognitiv dissonans mellom de to faktorene (oppfattet attraktivitet og bedømming av bildes autentisitet).

Det er viktig å understreke at man under utførelsen av eksperimentet ikke har noen kunnskap om deltakernes tanker om personene avbildet som «mer» eller «mindre» attraktive enn dem selv, kun deres subjektive bedømming av hvor attraktive personene avbildet er.

## **Daglig bruk av sosiale medier**

Denne variabelen er en kontinuerlig variabel som måler hvor mange minutter deltakeren selv rapporterer å bruke på sosiale medier daglig ved hjelp av et åpent tekstfelt.

## **Bruk av sosiale medier**

Denne variabelen er en additiv indeks. Variabelen er slått sammen av 5 ulike variabler som alle måler hvor mye deltakeren bruker de ulike sosiale medieplattformene Facebook, Twitter, Instagram, LinkedIn og SnapChat på en skala fra 0 til 6 (0=aldri og 6=hele tiden). Disse variablene er summert for å finne ut hvor mye en person bruker sosiale medier totalt. Blant annet brukes denne indeksen fra variabelen til å teste hypotesen om at høyere bruk av sosiale medier fører til bedre kompetanse til å gjenkjenne redigeringer, noe som krever én variabel som oppsummerer den totale bruken av sosiale medier – kontra 5 separate variabler. Indeksen varierer teoretisk fra 0 (0+0+0+0+0) til 30 (6+6+6+6+6), mens 0 er lavest mulig bruk av sosiale medier og 30 er høyest mulig mediebruk. Empirisk observeres derimot kun verdier fra 3 til 25, der verdier mellom 0-2 og 26-30 ikke blir observert i det sammenslåtte utvalget i variabelen. Basert på dette: Jo høyere indeksverdi, jo mer tid bruker noen på sosiale medier.

## **Deling av bilder på sosiale medier**

På lik linje med variabelen «bruk av sosiale medier», er denne variabelen også en additiv indeks slått sammen av 5 ulike variabler som måler hvor ofte deltakerne deler bilder på de ulike sosiale medieplattformene Facebook, Twitter, Instagram, LinkedIn og SnapChat. Hyppigheten av deling ble målt på en skala fra 0 til 6 (0=aldri og 6=hele tiden). Her vil også den teoretiske måleenheten opprinnelig ligge på 0 til 30, men empirisk observeres kun 0 til 14, siden 15-30 ikke blir observert i det sammenslåtte utvalget i variabelen.

Basert på dette: Jo høyere indeksverdi, jo oftere deler deltakeren bilder på sosiale medier.

### **Erfaring med sosiale medier**

For å kunne måle effekten av H1: *Mer erfaring med sosiale medier (h1 a: bruk av sosiale medier og h1 b: deling på sosiale medier) vil ha en positiv effekt på evnen til å skille mellom redigerte og ikke-redigerte bilder*, ble frekvensen av variablene deling og bruk av sosiale medier slått sammen for å danne «sosiale medier erfaring». Dette ble gjort basert på en korrelasjonsanalyse mellom de tre ulike variablene «*hvor ofte deltakeren deler bilder på sosiale medier-plattformer*», «*hvor ofte de oppholder seg på sosiale medier-plattformer*» og «*om de selv har redigert bilder av seg selv eller andre før de har lagt det ut på sosiale medier*». Her viste analysen en høy korrelasjon mellom bruk og deling (0.6), men lav korrelasjon mellom deling og redigering (0.16) og mellom bruk og redigering (0.24). På denne måten vil det ikke være hensiktsmessig å måle deltakernes deling og bruk i to ulike hypoteser, da disse ikke viser en empirisk uavhengighet fra hverandre. Sammenslåingen av disse variablene gir en additiv variabel med en indeks som går fra 4-32, der indeksverdiene 0-3 og 30 ikke blir observert. Jo høyere summen av «erfaring med sosiale medier», jo større er erfaringen til deltakeren.

### **Selv redigere bilder**

Variabelen er slått sammen av to ulike variabler: deltakerens rapportering av hvor ofte han/hun selv redigerer lys/kontrast i bilder og hvor ofte han/hun selv redigerer eget eller andres utseende i bilder før det legges ut på sosiale medier. Disse to variablene består hver av skalaer fra 0 til 4 (0=aldri/jeg legger ikke ut bilder på sosiale medier, 1=sjeldent, 2=ofte, 3=svært ofte og 4=hver gang jeg deler bilder). Under sammenslåingen av disse to skalaene er verdiene blitt omkodet til en additiv indeks som går fra 0 til 6. Her er den teoretiske måleenheten 0 til 8, men empirisk observeres bare 0 til 6 da det ikke er registrert noen tilfeller av deltakere som «alltid» både redigerer bilder med lys/kontrast og endrer eget/andres utseendet før de legges ut på sosiale medier. Basert på dette: Jo høyere indeksverdi, jo mer erfaring har deltakeren med å selv redigere bilder.

### **Kompetansevurdering**

Denne variabelen er en skala fra 0 til 5. Her vurderte deltakerne selv sin egen kompetanse når det kommer til å vurdere bilder på sosiale medier som redigerte eller ikke (0=ikke kompetent i det hele tatt, 1= svært lite kompetent, 2=lite kompetent, 3=ganske kompetent, 4=kompetent og 5=svært kompetent).

### **Riktig gjettet**

For å kunne undersøke hvor vidt deltakerne korrekt vurderte bildene som redigert eller ikke, ble det lagt en egen variabel for å gjette riktig, en dikotom variabel som ble utformet ved å omgjøre variabelen «*sannsynligheten for at bildet er manipulert*» til en dikotom variabel og «*type redigering*» til en dikotom variabel. Her ble skalaen (1 til 5) i «*sannsynligheten for at bildet var manipulert*» omgjort til verdiene 1=0 (svært lite

sannsynlig), 2=0, 3=0 (ganske sannsynlig), 4=1, 5=1 (helt sikkert redigert), mens type redigering ble omgjort til verdiene 1=0 (originalbildet), 2=1 (redigert med lys/kontrast), 3=1 (feminine skjønnsstereotypiske redigeringer) og 4=1 (maskuline skjønnsstereotypiske redigeringer). Videre ble disse erstattet med den nye variabelen «gjett riktig» om deltakeren svarer 4 eller 5 (verdi 1) på skalaen for sannsynlighetsberegning av redigering, samtidig som bildet deltakeren vurderte enten hadde redigeringstype 2, 3 eller 4 (verdi 1). Dette gjaldt også om deltakeren gjettet 1, 2 eller 3 (med verdi 0) på skalaen om hvor sannsynlig de mente det var at bildet var redigert, og bildet de vurderte hadde redigeringsformen «1», altså originalversjonen (verdi 0).

## **Demografiske variabler**

### **Alder**

Kontinuerlig variabel som måler deltakernes alder i år ved hjelp av et åpent tekstfelt. Empirisk går variabelen fra 23 til 90.

### **Kjønn**

Variabelen måler deltakerens kjønn, og er omkodet til en dikotom variabel der verdien 0 representerer «mann» og 1 representerer «kvinne».

### **Utdanning**

Variabelen er utformet som en skala fra 1 til 6, som måler deltakerens høyeste fullførte utdanning. Variabelen er originalt en kategorisk variabel som er omkodet med verdiene 1 (ungdomsskole), 2 (videregående), 3 (bachelorgrad), 4 (mastergrad), 5 (profesjonsstudium), 6 (Doktor-grad/Phd), mens «annet» er kodet som «missing», eller manglende verdi. Dette er fordi «annet» kan både være enten høy eller lav utdanning, og man kan derfor ikke gi dette alternativet en tallverdi.

### 3.5 Deskriptiv statistikk

**Tabell 3:** Deskriptiv statistikk for avhengig variabel og forklaringsvariabler

<b>Avhengig variabel: sannsynlighet for at bildet er manipulert</b>		
	<b>N</b>	<b>Prosent</b>
<b>1</b> - svært lite sannsynlig	510	20,08
<b>2</b>	641	25,24
<b>3</b> - ganske sannsynlig	746	29,37
<b>4</b>	330	12,99
<b>5</b> - helt sikkert redigert	313	12,32
<hr/>		
Totale observasjoner	2540	100
<b>Forklaringsvariabel: treatment-grupper</b>		
	<b>N</b>	<b>Prosent</b>
<b>1</b>	804	24.42
<b>2</b>	892	27.10
<b>3</b>	740	22.48
<b>4</b>	856	26.00
<hr/>		
Totale observasjoner	3292	100

**Tabell 4:** Deskriptiv statistikk for uavhengige variabler

<b>Attraktivitet</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>		
<b>1</b> – ikke særlig attraktiv	206	9.33			
<b>2</b>	402	18.21			
<b>3</b> – ganske attraktiv	883	40.01			
<b>4</b>	526	23.83			
<b>5</b> – veldig attraktiv	190	8.61			
Totalt	2207	100			
<hr/>					
<b>Daglig bruk av sosiale medier</b>	<b>N</b>	<b>Gjennomsnitt</b>	<b>Min</b>	<b>Max</b>	<b>Missing</b>
Totale observasjoner	2136	60	0	500	
<hr/>					
<b>Deling av bilder på sosiale medier</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>		
<b>0</b>	52	2.51			
<b>1</b>	232	11.20			
<b>2</b>	292	14.09			
<b>3</b>	308	14.86			
<b>4</b>	244	11.78			
<b>5</b>	276	13.32			
<b>6</b>	220	10.62			
<b>7</b>	160	7.72			
<b>8</b>	108	5.21			

<b>9</b>	100	4.83
<b>10</b>	36	1.74
<b>11</b>	16	0.77
<b>12</b>	8	0.39
<b>13</b>	12	0.58
<b>14</b>	8	0.39
Totale observasjoner	2072	100

<b>Bruk av bilder på sosiale medier</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>
<b>3</b>	16	0.76	
<b>4</b>	48	2.28	
<b>5</b>	132	6.27	
<b>6</b>	72	3.42	
<b>7</b>	116	5.51	
<b>8</b>	132	6.27	
<b>9</b>	152	7.22	
<b>10</b>	176	8.37	
<b>11</b>	156	7.41	
<b>12</b>	104	4.94	
<b>13</b>	176	8.37	
<b>14</b>	248	11.79	
<b>15</b>	248	11.79	
<b>16</b>	104	4.94	
<b>17</b>	60	2.85	
<b>18</b>	48	2.28	
<b>19</b>	48	2.28	
<b>20</b>	32	1.52	
<b>21</b>	24	1.14	
<b>22</b>	8	0.38	
<b>25</b>	4	0.19	
Totale observasjoner	2104	100	

<b>Erfaring med sosiale medier</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>
<b>4</b>	24	1.16	
<b>5</b>	32	1.54	
<b>6</b>	92	4.44	
<b>7</b>	52	2.51	
<b>8</b>	88	4.25	
<b>9</b>	76	3.67	
<b>10</b>	76	3.67	
<b>11</b>	100	4.83	
<b>12</b>	92	4.44	
<b>13</b>	80	3.86	
<b>14</b>	112	5.41	
<b>15</b>	108	5.21	
<b>16</b>	108	5.21	
<b>17</b>	128	6.18	
<b>18</b>	128	6.18	
<b>19</b>	148	7.14	
<b>20</b>	92	4.44	

<b>21</b>	124	5.98
<b>22</b>	100	4.83
<b>23</b>	76	3.67
<b>24</b>	48	2.32
<b>25</b>	44	2.12
<b>26</b>	64	3.09
<b>27</b>	36	1.74
<b>28</b>	16	0.77
<b>29</b>	12	0.58
<b>31</b>	8	0.39
<b>32</b>	8	0.39

Totale observasjoner 2072 100

<b>Selv redigere bilder</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>
<b>0</b>	544	26.05	
<b>1</b>	696	33.33	
<b>2</b>	396	18.97	
<b>3</b>	228	10.92	
<b>4</b>	164	7.85	
<b>5</b>	36	1.72	
<b>6</b>	24	1.15	

Totale observasjoner 2088 100

<b>Riktig gjettet</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>
	1,078	32.75	

Totale observasjoner 3292 100

<b>Kompetansevurdering</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>
<b>1</b>	468	22.72	
<b>2</b>	876	42.52	
<b>3</b>	536	26.02	
<b>4</b>	144	6.99	
<b>5</b>	36	1.75	
Totale observasjoner	2060	100	

**Tabell 5:** Deskriptiv statistikk for demografisk uavhengige variabler

<b>Kjønn</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>		
	1764	82.43			
Totale observasjoner	2540	100			

  

<b>Alder</b>	<b>N</b>	<b>Gjennomsnitt</b>	<b>Min</b>	<b>Max</b>	<b>Missing</b>
			23	90	
Totale observasjoner	2120	57,5			

  

<b>Utdanning</b>	<b>N</b>	<b>Prosent</b>	<b>Missing</b>		
<b>1</b> – Ungdomsskole	44	2.92			
<b>2</b> - Videregående	468	31.03			
<b>3</b> - Bachelorgrad	232	15.38			
<b>4</b> - Mastergrad	704	46.68			
<b>5</b> – Profesjonsstudium	0	0			
<b>6</b> – Dr. grad (ph.d)	60	3.98			
Totale observasjoner	1508	100	1784		

### 3.6 Analysemetode

Analysen er basert på 635 besvarelser, som utgjør 2540 registrerte observasjoner av eksperimentets hoved-avhengige variabel «sannsynlighet for at bildet er redigert». Med andre ord var det kun 635 av de 823 deltakerne som startet undersøkelsen som besvarte alle spørsmålene i eksperimentets hoved-avhengige variabel. Det er 2540 registrerte observasjoner fordi hver deltaker i eksperimentet gir fire ulike vurderinger av en enhet, noe som vil si at det totale antallet deltakere ganges med det totale antallet enheter, som gir et høyere antall registrerte observasjoner ( $635 \cdot 4 = 2540$ ) enn antall deltakere. Eventuelle frafall fra deltakere – altså deltakere som ikke gjennomførte hele undersøkelsen – kan utgjøre et annet antall registrerte variasjoner i visse variabler. Det sentrale er likevel at alle deltakerne som genererer data besvarte eksperimentets hoved-avhengige variabel, altså at de vurderte hvor sannsynlig det er at de fire ulike bildene var redigerte eller ikke.

#### 3.6.1 Balansetest

Det er vanlig praksis i mange samfunnsvitenskapelige felt å styrke den interne validiteten til et gitt eksperiment ved å gjennomføre en balansetest, der man sammenligner ulike eksperimentelle grupper opp mot variabler som ikke er en del av det sentrale teoretisk rammeverk for eksperimentet (Mutz, 2011). Det er likevel viktig å understreke at selv om randomisering er riktig gjennomført, kan visse ulikheter mellom gruppene oppstå rent tilfeldig (Mutz, 2011; Estellat et al., 2009). Det er derfor viktig å beskrive eventuelle ulikheter mellom gruppene slik at leseren selv kan vurdere hvor sammenlignbare gruppene er, da størrelsen og retningen på forskjellene kan være viktigere enn statistisk signifikante tester (Estellat et al., 2009). For å sikre at det ikke er store ulikheter i

eksperimentets eksperimentgrupper, vil det gjennomføres en balansetest ved å sammenligne gjennomsnittet fra hver demografiske variabel (kjønn, alder og utdanning) opp mot hver eksperimentgruppe (vedlegg 3).

Kjønn – altså antall kvinner – i hver eksperimentgruppe blir målt i prosentandel da de ulike eksperimentgruppene ikke alle er like store. Dette vil vise om det eksisterer store variasjoner i eksperimentgruppens utvalg, og null-hypotesen her vil være at det ikke er ulikheter i eksperimentgruppene.

Resultatene fra denne testen viser at det ikke eksisterer store forskjeller i gruppene. Likevel er det verdt å nevne at det er en overordnet skjevhet mellom menn og kvinner som har deltatt i undersøkelsen, men at denne skjevheten ikke viser utslag mellom selve gruppene. På bakgrunn av dette kan det konkluderes med at eksperimentets randomiseringsprosess har vært vellykket. Eksperimentets indre validitet vil videre diskuteres i kapitlet «3.7 Eksperimentets indre og ytre validitet».

### 3.6.2 Estimering av treatment-effekt

Når man skal gjennomføre estimeringer av hvilken effekt den eksperimentelle behandlingen har på hver gruppe, må man først rapportere et sammendrag av hver gruppes utfall. Med dette menes andelen deltakere som ble utsatt for eksperimentell behandling, eller en måling av gjennomsnitt og standardavvik av målingene gjort (Estellat et al., 2009). Videre må to tilleggsdata rapporteres: 1) kontrasten mellom de to gruppene, altså selve estimeringen av behandlingseffekten, og 2) presisjonen av denne estimeringen, altså den statistiske signifikansen av behandlingseffekten (f.eks P-verdien) (Estellat et al., 2009).

Når man måler en behandlingseffekt, i dette tilfellet hvor sannsynlig deltakerne tror det er at bildet er manipulert eller ikke, vil behandlingseffekten være den gjennomsnittlige forskjellen mellom treatment-gruppen og kontroll-gruppen, mens standardavvik her blir brukt for å gjenspeile spredningen rundt dette gjennomsnittet (Estellat et al., 2009). Estimeringen av behandlingseffekt vil dermed måles på følgende måte:

$$ES = (\text{Gjennomsnitt behandlingsgruppe} - \text{Gjennomsnitt kontrollgruppe}) / SD_{\text{samlet}}$$

I denne estimeringen vil ratioen ES (effekt størrelse) finnes ved å dele forskjellene i gjennomsnittet fra de to gruppene med SD (standardavvik), noe som muliggjør en enklere sammenligning av effekten (Estellat et al., 2009). Når det kommer til hvilke resultater som utgjør en stor eller liten effekt, er det noen uenigheter om dette (Estellat et al., 2009). Likevel kan man basert på Cohen (1988, referert i Becker, 2000) sin definisjon av effektstørrelse si at en verdi på 0.2 representerer en liten behandlingseffekt, 0.5 representerer en middels behandlingseffekt og 0.8 vil vise til en større behandlingseffekt (Becker, 2000). Man snakker her om en *klinisk signifikans*, heller enn en statistisk signifikans. Klinisk signifikans tar for seg størrelsen av behandlingseffekten, eller hvor meningsfull et utslag faktisk er, mens statistisk signifikans måler troverdigheten av effekten – med andre ord at sammenhengen man finner ikke stammer fra tilfeldigheter (Estellat et al., 2009).

Statistisk signifikans representeres gjerne som p-verdi, mens effektstørrelse kan utregnes på ulike måter. Selv om det finnes mange måter å beregne en P-verdi på, er hensikten bak testen alltid den samme: P-verdien skal vise om resultatene ville ha oppstått ved en tilfeldighet hvis det ikke eksisterte noen ulikheter mellom de to



eksperimentelle gruppene treatment-gruppe og kontrollgruppe (Estellat et al., 2009). På denne måten vil P-verdien beskrive risikoen for å gjøre en falsk-positiv konklusjon om at det eksisterer en forskjell når det i virkeligheten ikke er noen forskjell i det hele tatt. Med andre ord viser P-verdien om den observerte forskjellen sannsynligvis er sann ( $P < 0,05$ ) eller bare et resultat av tilfeldigheter ( $P > 0,05$ ), det vil si statistisk ikke signifikant (Estellat et al., 2009).

### 3.6.3 ANOVA med gjentatte målinger og marginal prediksjon

ANOVA- testen, eller «varianseanalysen», ble introdusert som en teknikk for å analysere eksperimentelle data, der eksperimentvariabelen (eller faktoren) ofte er utsatt for en eksperimentell behandling (Ringdal, 2013). Det er flere måter man kan gjennomføre en ANOVA-test, men i de kommende avsnittene vil det redegjøres for analysemetoden «Repeated Measures ANOVA», da det er denne formen for analyse av data som vil bli brukt i eksperimentets analysedel. Videre vil det redegjøres for marginal prediksjon, og hvorfor dette gjerne blir brukt for å tolke resultater fra eksperimenter.

Variansanalyse er et statistisk verktøy for å undersøke ulikheten mellom gjennomsnittet i variabler av distinktive grupper av subjekter (Hand & Taylor, 1987). Man kan anse ANOVA-analysen som en utvidelse av to-gruppe t-testen, der testen blir overført til flere grupper, slik som grupper av eksperimentell behandlingstype (Hand & Taylor, 1987). Her blir en F-verdi kalkulert, som – om den er signifikant – indikerer om det eksisterer en ulikhet mellom de ulike gjennomsnittene som blir målt (Nardi, 2014). F-testen forteller på denne måten om den avhengige variabelen faktisk har en effekt. I tilfeller der et individ blir gitt en rekke ulike eksperimentelle behandlinger, og på den måten også fungerer som sin egen kontrollgruppe, brukes gjerne en «repeated measure» i varianseanalysen (Sthåle & Wold, 1989). Med dette menes at et eksperiment (på lik linje med dette eksperimentet) kan ha ulike kilder av variasjon i datasettet, da det blir presentert ulike former for stimuli for hvert enkelt individ. Dette betyr at for det gjeldende eksperimentet, er det behov for denne typen spesifisering av ANOVA satt opp mot andre eksperimenter der for eksempel påvirkningen av individet ikke vil påvirke flere enheter/målinger i eksperimentet. Det eksisterer altså variasjon innad i individet i dette eksperimentet. «Repeated measures» betyr at faktorene blir gjentatt, for eksempel at samme forsøksperson fikk mer enn én behandling (Graph Pad). Videre i analysen vil «repeated measures» oversettes til «gjentatte målinger». I dette gjeldende eksperimentet, vil de ulike behandlingene bestå av de fire ulike bildeversjonene: originalbilde (ikke-redigert), redigert med lys/kontrast og redigert med feminine- og maskuline skjønnhetsstereotypiske trekk. I og med at hvert individ både blir eksponert for kontrollgruppen (originalbildet) og flere eksperimentfaktorer, vil ANOVA-test med gjentatte målinger være en gunstig analytisk metode for å analysere eksperimentets data. Her kan man også være interessert i å undersøke eventuelle interaksjons-effekter, der man undersøker om visse faktorer i ANOVA-analysen har en kombinert innvirkning på hverandre (Nardi, 2014).

For å videre undersøke effekten av den eksperimentelle behandlingen, kan man bruke en marginal prediksjon etter ANOVA-analysen. Et marginalt gjennomsnitt er gjennomsnittsresponsen for hver kategori av en faktor, justert for eventuelle andre variabler i modellen. Marginaleffekter kan fortelle oss hvordan en avhengig variabel (eller et utfall) endres når en spesifikk uavhengig variabel (eller en forklaringsvariabel) endres (Glen, 2018).

### 3.6.4 Regresjonsanalyse

Regresjonsanalyse som analytisk metode kan både brukes på eksperimentelle og ikke-eksperimentelle data, og kan ansees som en generalisering av variansanalysen (Ringdal, 2016). I en regresjonsanalyse forutsettes det at det er en lineær sammenheng mellom den avhengige og uavhengige variabelen, noe som gjerne understrekes av en hypotese. I tilfeller der det er snakk om en eventuell sammenheng mellom den avhengige variabelen og én uavhengig variabel, tas ofte en *enkel* regresjonsanalyse i bruk. I tilfeller der man ønsker å undersøke sammenheng mellom flere tilleggsfaktorer og estimere effekten av hver enkelt av disse, brukes gjerne en *multippel* regresjonsanalyse (Sykes, 1993). I analyser der den avhengige variabelen er en dikotom variabel, brukes logistisk regresjon (Ringdal, 2016). I denne formen for regresjonsanalyse vil den avhengige variabelens predikerte verdier tolkes som sannsynligheter, mens effekter av de uavhengige variablene tolkes gjennom regresjonskoeffisientene eller odds ratio (Ringdal, 2016).

For å analysere dette eksperimentets data vil både multippel og logistisk regresjonsanalyse bli brukt, basert på hypotesen som undersøkes og den avhengige variabelens form.

## 3.7 Eksperimentets indre og ytre validitet

Man snakker ofte om både indre og ytre validitet i forskningsprosjektet, også spesielt i eksperimenter. Her er gjerne den konvensjonelle meningen at eksperimenter presterer høyt i indre validitet, men lavt i ekstern validitet (Mutz, 2011). I dette kapitlet skal jeg gå gjennom forskningsprosjektets kvalitet ved å vurdere prosjektets indre og ytre validitet.

### 3.7.1 Indre validitet

Dahlum (2021) beskriver validitet som «i hvilken grad man ut fra resultatene av et forsøk eller en studie kan trekke gyldige slutninger om det man har satt seg som formål å undersøke». I denne sammenhengen vil *indre* validitet være eksperimentets mulighet for å forklare funnene gjennom den antatte hypotesen, mens en høy indre validitet oppnås når mulige biaser i eksperimentet har blitt kontrollert for (Dahlum, 2021). I følgende avsnitt vil eksperimentets indre validitet diskuteres.

Den beste måten å evaluere effekten av en eksperimentbehandling er å sammenligne de to eksperimentgruppene – gruppen som har gjennomgått eksperimentell behandling og kontrollgruppen – for å se hvem av disse som viser størst effekt (Estellat et al., 2009). For at denne sammenligningen skal være gyldig er det i midlertidig sentralt at eksperimentgruppene er så like som overhodet mulig, gjennomgår den samme behandlingen (foruten den eksperimentelle behandlingen) og blir vurdert på samme måte på slutten av studiet (Estellat et al., 2009). For å forsikre seg om at de ulike gruppene i dette eksperimentet er så like som overhodet mulig, er deltakerne blitt randomisert til de ulike eksperimentgruppene uten innblanding for å unngå seleksjonsbias. Når man implementerer en randomiseringssekvens er det sentralt at tildelingen i ulike grupper er skjult for deltakeren, om ikke forsvinner fordelene med randomisering (Estellat et al., 2009). Randomiseringssekvensen i dette eksperimentet ble delvis skjult for deltakerne,

da det ikke ble indikert at besvarelsen av klokke-spørsmålet var den sentrale faktoren for deres tildeling av eksperimentell behandling. I tillegg til dette ble ikke deltakerne informert om at de var del av et eksperiment, kun at de skulle besvare spørsmål om deres oppfattelse av bilder på sosiale medier og egen bruk av sosiale medier-plattformer.

Likevel kan det diskuteres om metoden for randomisering basert på klokken er tilstrekkelig eller ikke. Ifølge Estellat og kolleger (2009) er riktig form for randomisering både forankret i implementeringen av sekvensen og i genereringen av den. En tilstrekkelig sekvens vil for eksempel være en tilfeldig talltabell eller en datastyrt tilfeldig tallgenerator, mens det man kaller en «kvasi-tilfeldig» metode, slik som å randomisere ved hjelp av vekslinger mellom fødselsdato eller første bokstav i fornavn/etternavn, vil være utilstrekkelig på grunn av metodens forutsigbarhet (Estellat et al., 2009). Med dette menes at det kan finnes mulige korrelasjoner mellom randomiseringselementet (f.eks fødselsdato) og utfallet. Da tekniske begrensninger førte til at en datagenerert randomiseringsfrekvens inne på spørreskjemaets plattform (Surveylegend.com) ikke var mulig, forble deltakernes minuttviser-plassering ved starten av undersøkelsen vurdert som et tilstrekkelig alternativ. Selv om denne metoden beveger seg mot en «kvasi-tilfeldig» metode, later det som at implementeringen av denne randomiseringsfrekvensen ikke har hatt store innvirkninger på deltakernes randomisering, da henholdsvis de ulike eksperimentgruppene har en fordeling på 24%, 27%, 22% og 26% av deltakerne. På samme tid viser også balansetesten (vedlegg 2) tilfredsstillende resultater.

I sammenhengen med validitet, er det også sentralt å nevne begrepsvaliditet. Denne formen for validitet betegnes også som *begrepsvaliditet* (Ringdal, 2013, s. 96). Når man ønsker å forsikre seg om en høy begrepsvaliditet er det sentralt å forsikre seg om at det teoretiske begrepet som benyttes i måleenheten faktisk måler det man ønsker, og ikke noe annet (Ringdal, 2013, s. 97). Her vil det også være sentralt å forsikre seg om at begrepene som blir brukt ikke bare tolkes slik man ønsker av mottakeren, men også at man ikke skaper spørsmål som er «ladet» i en retning, og som dytter mottakeren i en spesiell retning ved å lene spørsmålet mot ladde ord (Nardi, 2014). Eksempelvis ble en korleksjon gjennomført forankret i dette, der bruken av ordet «manipulasjon» av bilder ble endret til «redigering». Dette diskuteres nærmere i kapitlet «3.8 Utfordringer og endringer». Under undersøkelsens utvikling ble de derimot vurdert hvordan begrepet «manipulasjoner» kunne bære med seg en negativ konnotasjon for deltakeren, og ble derfor endret til et mer nøytralt begrep – redigering.

Videre var det også sentralt å finne en balanse mellom å forklare til deltakeren hva som menes med redigering av bilder, uten at dette skulle føre til at de ble ekstra oppmerksomme på noen former for redigering. Her ble det derfor tatt et valg om å gi en kort beskrivelse av ulike måter man kan redigere et bilde på, før man spurte deltakerne om hvor sannsynlig de selv trodde det var at bildene som ble presenter var redigert eller ikke. Her ble det ikke nevnt spesifikke redigeringsformer (slik som å gjøre midjen slankere, øyne større eller lignende) bevisst for å ikke trekke deltakernes oppmerksomhet mot dette i større grad enn de kanskje ville gjort i møte med slike bilder på sosiale medier.

### 3.7.2 Ytre validitet

I det Store Norske Leksikon beskrives *ytre* validitet som oppnådd om eksperimentets resultater kan generaliseres til en større andel av befolkningen (Dahlum, 2021). Altså at den begrensede mengden data fra et eksperiment kan tenkes å gjelde en større del av populasjonen enn kun den som ble undersøkt. I de følgende avsnitt vil eksperimentets ytre validitet videre diskuteres.

I publikasjonen «The External Validity of Experiments» beskriver Bracht og Glass (1968) to trusler når det kommer til ytre validitet i eksperimenter; «populasjonsvaliditet» og «økologisk validitet». Populasjonsvaliditeten går ut på at en gyldig generalisering fra eksperimentets populasjon over til den generelle populasjonen krever grundig kunnskap om begge populasjoner. Resultatene som genereres fra et eksperiment kan tenkes å kun gjelde det eksperimentelle utvalget og deres individuelle karakteristikk, og ikke resten av befolkningen (Bracht & Glass, 1968). Under vurderingen av populasjonsvaliditeten skal det også vurderes om det eksisterer en interaksjon mellom personlige variabler blant deltakerne og behandlingseffekten. Med dette menes at om en eksperimentell behandling som gjennom eksperimentet har vært sterkere enn en annen, hadde blitt reversert om deltakere med andre variasjoner av de personlige variablene ble utsatt for samme eksperimentelle behandling (Bracht & Glass, 1968). En av de beste metodene for å oppnå generalisering er ved å gjennomføre en tilfeldig seleksjon av befolkningen, der deltakerne i eksperimentet blir valgt ut ved tilfeldighet for å representere populasjonen (Campbell et al., 2002). Selv om eksperimentets utvalg er randomisert mellom eksperimentgruppene, er ikke utvalget selektert på en slik måte at man kan hevde at utvalget/resultatene er generaliserbare til resten av Norges befolkning. I og med at deltakelse i eksperimentet også var frivillig, kan det her være andre individuelle faktorer som ikke måles i utvalget – slik som personlighetstrekk basert på deltakernes eget ønske om å frivillig delta i undersøkelsen. Likevel ansees sjansen for at interaksjoner mellom personlige og avhengige variabler er tilstede som lav, der eksperimentets spesielt dominerende behandlinger også dominerer på tvers av eksperimentgruppene, på tross av ulike personlige karakteristikk innad i gruppene (dette utforskes nærmere i analysen). De eksperimentelle gruppene, slik balansetesten viser, later til å være godt randomisert med demografiske faktorer. Andre personlige egenskaper som ikke ble målt i under eksperimentet kan her ikke tas forhold til.

Økologisk validitet i eksperiment handler blant annet om at forskeren gir eksplisitt beskrivelse av den uavhengige variabelen, der det å gi tilstrekkelig kunnskap om alle aspekter ved den eksperimentelle behandlingen gir mulighet for å senere replisere eksperimentet (Bracht & Glass, 1968). Andre faktorer under økologisk validitet som kan trekkes frem for dette eksperimentet er blant annet «Hawthorne-effekten», der deltakerens oppførsel kan påvirkes av hans/hennes oppfatning av eksperimentet og hvordan han/hun bør reagere i møte med den eksperimentelle stimulusen, og den «eksperimentelle effekten», der oppførselen til forsøkspersonene kan være u- bevisst påvirket oppførselen til eksperimentatoren (Bracht & Glass, 1968). I og med at eksperimentet er gjennomført som et survey-eksperiment er det liten sannsynlighet for at noen av disse effektene har hatt innvirkning på eksperimentets ytre validitet, da deltakerne verken var informert om at de deltok i et eksperiment med ulike eksperimentelle behandlinger eller hadde noen form for interaksjon med eksperimentatoren.

### 3.8 Utfordringer og endringer

Både gjennom planleggingsfasen og utførelsen av eksperimentet møtte jeg på ulike utfordringer som førte til at jeg måtte ta stilling til endringer i eksperimentet. På bakgrunn av dette har det skjedd noen endringer fra eksperimentets første versjon til den ferdige versjonen. I de følgende avsnittene vil jeg beskrive hvordan disse utfordringene ble møtt, og hvordan de skapte grunnlag for forbedring av det ferdige eksperimentet.

Slik det blir beskrevet i diskusjonen om begrepsvaliditet, ble begrepet «manipulert» valgt bort da det kunne bære med seg eventuelle negative konnotasjoner for deltakerne. «Manipulering» defineres av det Store Norske Leksikon som å «(...) lede eller påvirke noen i en bestemt og ønsket retning, ofte på en skjult måte» (Nilstun, 2021). Før dette begrepet ble erstattet med «redigering», så ble «endring» brukt for å erstatte manipulasjonsuttrykket i spørreskjemaets første versjon. Rekkefølgen av spørreskjemaets ulike deler var også annerledes i eksperimentets i denne versjonen. I den første versjonen ble deltakerne først presentert deres treatment-gruppens fire ulike bilder med spørsmål om å vurdere deres attraktivitet. Deretter ble deltakerne igjen presentert med samme utvalget av bilder, med spørsmålet om å vurdere hvor sannsynlig de selv trodde det var at bildet var «endret». Etter at denne versjonen gikk gjennom en testrunde, ble den publisert for et mindre utvalg av deltakere. Her viste det seg at kombinasjonen av ordbruken «endring» sammen med rekkefølgen de ulike delene ble presentert i, førte til misforståelser blant flere av deltakerne. Siden de samme bildene ble presentert to ulike ganger på hver sin side, og side nr. 2 brukte ordet «endre», var det flere av deltakerne som trodde de så etter endringer i bildene fra første til andre side – noe som absolutt ikke var hensikten. Men andre ord endte spørreskjemaet opp med å måle en effekt som det ikke var hensikten å måle. Spørreskjemaet ble derfor raskt avpublisert og reorganisert, og dataene fra deltakerne som deltok i denne runden slettet. I den nye og forbedrede versjonen ble begrepet «endring» endret til «redigering». Videre ble rekkefølgen endret til å først presentere bildeutvalget med spørsmål om sannsynlighetsberegning for redigering, deretter det samme utvalget bilder med spørsmål om attraktivitetsvurdering. Denne versjonen ble så igjen testet grundig av friville testpersoner, før det igjen ble publisert.

Selv om endringene som er gjort i spørreskjemaet ansees som nødvendige for å oppnå et tilfredsstillende forskningsresultat, er det sentralt å nevne hvorfor spørreskjemaets originale versjon ble tenkt som foretrukket da det ble publisert for første gang. For å undersøke kognitiv dissonans gjennom deltakernes attraktivitetsvurdering, var det foretrukket å stille dette spørsmålet først, før deltakerne måtte vurdere om personen sannsynligvis var redigert eller ikke. Det var tenkt at dette ville føre til en mer upåvirket vurdering. I den nye versjonen, kan det diskuteres om deltakerne vurderer personene på bildet som attraktive eller ikke gjennom et perspektiv som er påvirket av deres tidligere vurdering av personene avbildet som redigerte eller ikke. Selv om dette kan ansees som en svakhet ved målingen av variabelen «attraktivitet», var det en nødvendig endring for å forsikre at eksperimentet målte det det faktisk skulle måle. Det samme gjelder begrepsendringen, som kan åpne opp for en begrepsdiskusjon i og for seg selv. Det kan her tenkes at begrepene bærer med seg to ulike konnotasjoner, der «manipulasjon» kanskje i større grad knyttes opp mot endringer i bilder slik som å endre, fjerne eller legge til sentrale elementer i et bilde, men redigering i større grad knyttes opp mot mindre justeringer i fotografier slik som endringer av lys, fargenyanser eller lignende.

Begrepet «endring» ble derfor ansett som en nøytral middelvei, men på bakgrunn av deltakernes misforståelse og manipulasjonsbegrepets eventuelle negative konnotasjoner, argumenteres det sterkt for at «redigering» gir en tilstrekkelig beskrivelse av fenomenet som måles.

Den første versjonen av spørreskjemaet var også fire ulike og separate versjoner. Når spørreskjemaet ble revidert for andre gang, ble det tatt et valg om å konstruere et mer komplekst spørreskjema med de fire ulike versjonene inkorporert innad i samme skjema. På grunn av begrensninger på spørreskjemaets plattform Surveylegend.com, var det ikke mulig å randomiserer deltakerne ved å bruke egne funksjoner. Slik det ble diskutert i kapitlet forskningskvalitet og etikk, ville dette vært en foretrukket måte å gjøre det på for å sikre best mulig randomisering. Likevel ble det ansett som gunstig å samle spørreskjemaet, da dette førte til en større kontroll rundt randomiseringsprosessen enn om fire ulike versjoner av samme spørreskjema ble publisert på sosiale-  
medieplattformer.

## 4. Resultater

I de kommende avsnittene vil jeg presentere resultatene fra eksperimentet og analysere dem. Forankret i eksperimentets forskningsspørsmål og hypoteser, vil resultatene først presenteres i rammene av bildenes påvirkning, påvirkning fra redigeringstype, samspill mellom redigering og bilde og innvirkning av andre faktorer, for eksempel deltakernes sosiodemografiske variabler. Under analysen av bildenes påvirkning har jeg valgt å fokusere på representasjon av kjønn i bildet heller hvert enkelt bilde, da det forventes at dette vil gi mer interessant informasjon enn bildene nummerert.

ANOVA-test med gjentatte målinger, justert for interaksjon mellom kjønn representert i bildet og type redigering, vil bli brukt som analytisk metode i gjennomgangen av resultatene fra eksperimentet. Her vil den avhengige variabelen skiftes ut basert på hvilke tematikker som undersøkes: subjektiv sannsynlighetsberegning av redigering til stede, attraktivitetsvurdering eller deltakernes deteksjonsevne. For å undersøke innflytelse av andre faktorer på den avhengige variabelen vil regresjonsanalyse utføres.

### 4.1 Treatment-effekt og overordnede resultater

For å undersøke effekten av eksperimentets eksperimentelle behandling, vil behandlingseffekten bli målt opp mot eksperimentets treatment- og kontroll-gruppe. For å gjennomføre dette må det genereres en ny gruppe-variabel som skiller mellom eksperimentets eksperiment- og kontroll-grupper – manipulert/ikke-manipulert. I analysen vil altså ikke type redigering stå i fokus, men heller hvilken effekt den eksperimentelle behandlingen hadde på deltakernes vurdering av sannsynligheten for at bildene var manipulerede eller ikke. Slik beskrevet i den metodiske delen av eksperimentet, vil dette gjennomføres ved å utføre følgende ligning:

$$ES = (\text{Gjennomsnitt } \textit{behandlingsgruppe} - \text{Gjennomsnitt } \textit{kontrollgruppe}) / \text{SD}_{\text{samlet}}$$

$$ES = (2.8 - 2.4) / 1.3 = 0.31$$

Effekten viser en ES på 0.31, noe som tilsier en liten til medium behandlingseffekt. Med andre ord kan man gjennom denne analysen si at eksponering for «treatment» (redigerte bilder) hadde en lav til moderat effekt på deltakernes bedømming av bildene som manipulerede.

For å undersøke om denne effekten er et tilfeldig utfall eller ikke, kjøres variablene i en ANOVA-test med gjentatte målinger, justert for interaksjon mellom variabelen som skiller mellom eksperimentgruppene og variabelen foto – en kategorisk variabel som inkluderer de fire ulike bildene nummerert. Det trengs her en gjentatt måling ANOVA fordi hver deltaker fikk mer enn en verdi på gruppevariabelen, da hver deltaker i eksperimentet gir fire ulike vurderinger av fire ulike enheter (bilder) som fører til fire stimulus-respons enheter per individ. Derfor inkluderes deltakernes id-variabel som en variabel i modellen – noe som vil sikre at deltakerens tendens til å gjette «manipulert» eller «ikke manipulert» ikke har innflytelse på kjernedelen av eksperimentet. Denne testen viser en P-verdi for hele modellen som er mindre enn 0.0001, noe som vil si at den kliniske signifikansen (ES) også er statistisk signifikant.

Resultatene fra eksperimentet viser gjennomsnittlig at deltakerne klarte å korrekt estimere om bildene de ble eksponert for var manipulerede eller ikke 33% av gangene,

mens den gjennomsnittlige skåren deltakerne ga i møte med spørsmålet om hvor sannsynlig de trodde det var at bildet var manipulert, er på 2.72 (på en skala fra 1-5).

## 4.2 Subjektiv sannsynlighetsvurdering av redigering i bildet

Gjennomsnittlig tenderer deltakerne i eksperimentet altså å plassere sannsynligheten for at bildet er manipulert på 2.72 på skalaen som går fra 1 (svært lite sannsynlig) til 5 (helt sikkert redigert). Dette er en vurderingsverdi som gjennomsnittlig plasserer deltakernes besvarelser nesten akkurat på midten av skalaen.

I de kommende avsnittene vil deltakernes sannsynlighetsvurdering bli videre analysert. Jeg undersøker først påvirkningen av kjønn representert i bildet, påvirkning av type redigering i bildet og til slutt hvilken innvirkning samspillet mellom kjønn representert i bildet og type redigering har på sannsynlighetsvurderingen.

### 4.2.1 Påvirkning av bilde

I denne analysedelen undersøker jeg hvilken påvirkning kjønn som ble representert på bildet hadde på deltakernes sannsynlighetsvurdering av redigering.

For å finne ut av dette ble det gjennomført en ANOVA-analyse med gjentatte målinger, hvor deltakernes sannsynlighetsvurdering fungerte som avhengig variabel. Her ble det også justert for interaksjon mellom hvilket kjønn som ble representert i bildet og type redigering. Denne testen viser at det eksisterer en signifikant effekt mellom kjønn som er avbildet, og deres vurdering av hvor sannsynlig det er at bildet er manipulert eller ikke:  $F(1, 1811)=56.09$ ;  $p<.001$ . Ut ifra dette kan man si hvilken sannsynlighetsvurdering en deltaker vil gi, er påvirket av hvorvidt det er en mann eller kvinne som er representert på bildet som vurderes. For å undersøke hvilken effekt det er snakk om gjennomføres en marginal prediksjon på variabelen «kjønn representert i bildet» (tabell. 6).

**Tabell 6:** Det representerte kjønnets påvirkning på sannsynlighetsvurdering

		Antall observasjoner = 2540	
		<b>Kjønn av personen avbildet</b>	
		Kvinne	Mann
Sannsynlighetsvurdering		2.88	2.55

Tabell. 6 viser at deltakerne vurderer det som mer sannsynlig at bildet som presenteres for dem er redigert når det er en kvinne som er avbildet ( $M=2.88$ ;  $95\%CI=[2.82; 2.94]$ ) satt opp mot når menn er avbildet ( $M=2.55$ ;  $95\%CI=[2.48; 2.61]$ ). Forskjellene mellom de to ulike kjønnene viser altså en økning på 0,33 på sannsynlighetsskalaen (1-5) i møte med kvinner avbildet.



Denne analysen bidrar til å gi svar på eksperimentets første forskningsspørsmål, F1: *Finnes det en sammenheng mellom kjønn representert i et bilde og hvor sannsynlig deltakerne mener det er at bildet er manipulert?*

Funnene fra analysen kan bekrefte at det faktisk eksisterer en statistisk signifikant forskjell mellom kjønn som representeres i bildet og deltakerens sannsynlighetsvurdering. Det er en større sannsynlighet for at deltakeren tror bildet er manipulert når det er en kvinne avbildet.

#### 4.2.2 Påvirkning av redigering

Her vil det undersøkes om type redigering (lys/kontrast, maskulin skjønnhetsstereotypisk, feminin skjønnhetsstereotypisk eller original (uredigert) hadde en påvirkning på deltakernes sannsynlighetsvurdering av redigering i bildene. Det blir gjennomført en ANOVA med gjentatte målinger, justert for interaksjon mellom kjønn som representeres i bildet og type redigering. Deltakernes vurdering av sannsynlighet for redigering i bildene er her satt som avhengig variabel. Testen viser:  $F(3, 1811)=93.34$ ;  $p<.001$ , som betyr at effekten redigering har på deltakernes sannsynlighetsvurdering er signifikant. Dette forteller oss at typen redigering som presenteres i bildet har en effekt på deltakernes vurdering av bildene som redigerte eller ikke. For å nærmere undersøke denne effekten gjennomføres det en marginal prediksjon på variabelen «redigeringstype» (Tabell. 7).

**Tabell 7:** Redigeringstypes påvirkning på sannsynlighetsvurdering

	Originalbilde	Lys/kontrast	Feminin stereotypisk	Maskulin stereotypisk
Sannsynlighetsvurdering	2.36	2.77	3.33	2.44

Tabell. 7 viser at bilder som blir redigert med feminine skjønnhetsstereotypiske trekk tydelig skiller seg fra de andre redigeringsmåtene når det kommer til deres påvirkning på sannsynlighetsvurderingen ( $M=3.33$ ;  $95\%CI=[3.24; 3.41]$ ). Her tenderer deltakerne å bedømme feminine skjønnhetsstereotypiske redigeringer 0,89 høyere på skalaen enn maskuline skjønnhetsstereotypiske redigeringer ( $M=2.44$ ;  $95\%CI=[2.35; 2.52]$ ). Originalbildet er den redigeringsformen som viser lavest effekt på sannsynlighetsvurderingen ( $M=2.36$ ;  $95\%CI=[2.27; 2.45]$ ), noe som betyr at eksponering for redigeringsformene hadde en større innvirkning på deltakernes mistanke om at bildene var redigert satt opp mot det ikke-redigert originalbildet – selv om variasjonen mellom effektene ikke er meget stor.

Det vil videre være interessant å undersøke om det eksisterer en interaksjonseffekt mellom kjønn representert i bildet og typen redigering som vises på bildet, med deltakernes sannsynlighetsvurdering av om bildet er manipulert som avhengig variabel.

### 4.2.3 Samspill mellom bilde og redigeringstype

Innvirkningen til samspillet mellom kjønn representert i bildet og type redigering på den subjektive sannsynlighetsberegningen av redigering blir målt ved å bruke ANOVA-analyse med gjentatte målinger justert for interaksjon mellom representert kjønn og redigeringstype. Avhengig variabel er her den subjektive sannsynlighetsberegningen av redigering. Ut ifra analysen ser man at samspillet mellom type kjønn som avbildes og typen redigering er statistisk signifikant, og at innflytelsen av kjønn og redigering på sannsynlighetsvurderingen er avhengige:  $F(3, 1811)=11.88$ ;  $p<.001$ . Med dette menes det at de ulike typer redigeringer har en innvirkning på de ulike gruppene av kjønn som er avbildet (kvinner og menn).

For å finne ut hvilken størrelse effekten har gjennomføres en marginal prediksjon med interaksjon mellom type redigering som vises på bildet og kjønn som er representert (tabell. 8). I og med at denne effekten bærer teoretisk relevans knyttet opp mot F2: «Har det en effekt på hvor sannsynlig deltakeren mener det er at bildet er manipulert når et bilde redigeres i motsetning til kjønnsstereotypen?», vil jeg utdype analysen med Scheffé post-hoc test. Scheffés test er gunstig etter man har kjørt ANOVA og fått signifikant F-statistikk, da teten kan brukes for å finne ut hvilke gjennomsnittsverdier som er signifikante eller ikke.

**Tabell 8:** Post-hoc analyse av hvordan redigeringstype og kjønn av avbildet person samspiller i å påvirke attraktivitetsvurdering

Antall observasjoner = 2,540

Bildet viser:	Orginalbilde	Lys/kontrast	Feminin stereotypisk	Maskulin stereotypisk
Kvinne	2.35 <sup>a</sup>	2.98 <sup>c</sup>	3.74	2.45 <sup>a</sup>
Mann	2.36 <sup>a</sup>	2.54 <sup>ab</sup>	2.86 <sup>bc</sup>	2.43 <sup>a</sup>

*Notat:* Resultat av en ANOVA med repeterte målinger. Gjennomsnittsverdier (marginale prediksjoner) i tabellen som ikke deler bokstav er signifikant forskjellig fra hverandre på  $p<.05$ , basert på Scheffé posthoc test.

Tabell. 8 viser interessante funn når det kommer til kvinner representert i bilder. Her ser man at bilder av kvinner som er redigert etter feminine skjønnhetsstereotypiske trekk skiller seg betraktelig fra resten av redigeringstypene ( $M=3.74$ ;  $95\%CI=[3.61; 3.87]$ ), noe som også reflekteres i post-hoc testen. Dette betyr at om det er en kvinne avbildet, som også er redigert med skjønnhetsstereotypisk feminine trekk, så estimeres det at deltakeren vil plassere sannsynligheten for at bildet er redigert (på en skala fra 1 til 5) på 3.74. Dette står i kontrast med originale/ikke-redigerte bilder der en kvinne er avbildet, som gjennomsnittlig plasseres på 2.35 på skalaen «hvor sannsynlig tror du det er at bildet er manipulert» av deltakerne ( $M=2.35$ ;  $95\%CI=[2.22; 2.49]$ ). Med andre ord øker deltakernes sannsynlighetsestimering med 1.36 ( $3.74-2.35$ ) poeng om bildet som presenteres til deltakeren er en kvinne redigert med skjønnhetsstereotypisk feminine trekk. Dett er en betraktelig økning sammenlignet med resten av modellens resultater. Til sammenligning ser man at menn redigert med maskuline skjønnhetsstereotypiske

trekk ( $M=2.43$ ;  $95\%CI=[2.29; 2.57]$ ) bærer liten forskjell fra menn representert i ikke-redigerte bilder ( $M=2.36$ ;  $95\%CI=[2.23; 2.50]$ ). Analysen presenterer nesten like verdier mellom kvinner og menn avbildet i et ikke-redigert bilde.

Disse funnene gir grunnlag for å gi svar på forskningsspørsmålet F2, der redigering som går i motsetning til kjønns stereotypier faktisk har liten effekt på hvor sannsynlig deltakeren mener det er at bildet er manipulert. Heller er det feminine stereotypiske trekk som har effekt på kvinner, men ikke på menn. Mennene som ble redigert med feminine skjønnhetsstereotypiske trekk viser en sannsynlighetsestimert på 2.86 ( $M=2.86$ ;  $95\%CI=[2.72; 3.00]$ ), satt opp mot det originale bilde av samme kjønn (2.36). I bildene av kvinner som ble redigert med maskuline skjønnhetsstereotypiske trekk ( $M=2.45$ ;  $95\%CI=[2.32; 2.58]$ ) var det nesten ingen forskjell satt opp mot det originale bildet av samme kjønn (2.35). Videre viser analysen at redigeringsformen «lys/kontrast» har en større innvirkning på deltakernes sannsynlighetsestimert av redigering når det er en kvinne avbildet ( $M=2.98$ ;  $95\%CI=[2.84; 3.12]$ ) satt opp mot det uredigerte bildet av samme kjønn (2.35, en økning på 0.63 poeng på skalaen), enn når det er menn avbildet ( $M=2.54$ ;  $95\%CI=[2.40; 2.67]$ ) opp mot det uredigerte bildet av samme kjønn (2.36, en økning på 0.18 poeng på skalaen).

Funnene fra denne analysen kan gi grunnlag til å tro at folk har en større mistanke om at bildene er redigert når det er en kvinne som er avbildet. Det er også grunnlag for å tro at folk ikke mistenker at kvinner ville ha redigert seg selv i motsatt retning, altså med maskuline skjønnhetsstereotypiske trekk.

#### **4.2.4 Innflytelse av andre faktorer**

Det kan også tenkes at sannsynlighetsvurderingen ikke bare er påvirket av bildene som ble vist i undersøkelsen. Vurderingen kan også reflektere deltakernes egenskaper og evner, for eksempel deres sosiodemografi, erfaring med sosiale medier, praktisering av bildedeling og redigering, samt deres egenvurderte kompetanse til å gjenkjenne redigeringer i bilder på sosiale medier. For å utforske hvilke innvirkninger dette eventuelt har på deltakernes sannsynlighetsvurdering gjennomførte jeg en regresjonsanalyse hvor deltakernes sosiodemografi og erfaring vil forklare sannsynlighetsvurderingen (modell 1). I en alternativ modell (modell 2) kontrollerte jeg for de eksperimentelle forholdene, det vil si: bilde, redigeringstype og deres samspill. Denne alternative modellen inkluderes for å undersøke om eventuelle innflytelser fra modell 1 vil forsvinne når eksperimentelle kondisjoner blir tatt hensyn til. På grunn av dette er ikke videre informasjon om koeffisienter inkludert i den alternative modellen, da dette ikke vil gi interessant informasjon.

**Tabell 9: Andre faktorielle innflytelser på attraktivitetsvurdering**

Antall observasjoner = 2,540

Avhengige variable: attraktivitetsvurdering		
	Modell 1	Modell 2
Alder	-.012***	-.012***
Kjønn (kvinne)	.041	.033
Utdanning	.073*	.085**
Selv redigere bilder	.035	.030
Erfaring med sosiale medier	-.004	-.003
Daglig bruk av sosiale medier (min)	-.0001	.00002
Kompetansevurdering	.138***	.128***
Bildet	—	3 koeffisienter
Redigering	—	3 koeffisienter
Bildet * Redigering	—	9 koeffisienter
(Constant)	2.81***	2.755***
R <sup>2</sup>	0.03***	0.23***

*Notat: \*\*\* betyr  $p < .001$ ; \*\* betyr  $p < .01$ ; \* betyr  $p < .05$ ; ingen markering betyr ikke statistisk signifikante resultater ( $p$ -verdi over 0.05). Jo flere stjerner jo mindre sjanse for type 1 feil.*

Både i modell 1 og modell 2 er det kun tre variabler som bærer statistisk signifikant innflytelse på deltakernes sannsynlighetsvurdering av bildene som redigerte eller ikke. Dette er de tre variablene «alder», «utdanning» og «kompetansevurdering», der (1) de med lavere alder vurderer det som mer sannsynlig at bildene er redigerte, (2) de med høyere utdanning vurderer det som mer sannsynlig at bildene er redigerte, og (3) jo høyere en deltaker vurderer sin egen kompetanse til å gjenkjenne redigeringer i bilder på sosiale medier, desto høyere er deres sannsynlighetsvurdering av redigering i bilder. De resterende variablene kjønn, det å selv redigere bilder, erfaring med sosiale medier og daglig bruk av sosiale medier bærer ikke statistisk signifikant innflytelse på sannsynlighetsvurdering. Sammenligner man de to modellene ser man at bakgrunnsvariablene alene (presentert i modell 1) har en liten effekt på deltakernes sannsynlighetsvurdering, der  $R^2=0.03$  kun forklarer 3% av variasjonen i sannsynlighetsvurderingen. Modell 2 som inkluderer eksperimentelle variasjoner, altså bildene som deltakerne eksponeres for, stiger  $R^2$  til 23%. Med andre ord forklarer den eksperimentelle variasjonen 20% av den totale variasjonen (23%-3%).

### 4.3 Attraktivitetsvurdering

For å besvare forskningsspørsmålet F3: «*Finnes det en sammenheng mellom deltakerens vurdering av personen avbildet som attraktiv og sannsynligheten for at de vurderer bildet som manipulert?*», blir det gjennomført en ANOVA-test med interaksjon mellom «redigeringstype» og «attraktivitet», «sannsynlighet for at bildet er manipulert» fungerer som avhengig variabel. Testen viser  $F(12, 1636)=2.79$ ;  $p=.0009$ , som forteller oss at innflytelsen av redigering og attraktivitetsvurdering er statistisk signifikant.

Vi går herifra videre til en marginal prediksjon av interaksjonen mellom den avhengige og de uavhengige variablene, og for å finne ut hvilken av verdiene i attraktivitetsskalaen som har størst innvirkning på deltakernes sannsynlighetsvurdering, blir det målt et gjennomsnitt mellom de ulike gruppene (tabell. 10).

**Tabell 10:** *Subjektiv sannsynlighetsberegning avhengig av attraktivitetsvurdering og redigeringstype*

Antall observasjoner = 2,207

Attraktivitetsmåling	Original	Lys/kontrast	Feminin stereotypisk	Feminin stereotypisk	Gjennomsnitt
1 – ikke særlig attraktiv	2.468	2.608	3.879	2.869	2.956
2	2.366	2.652	3.461	2.695	2.794
3	2.373	2.772	3.271	2.512	2.732
4	2.391	2.823	2.747	2.144	2.526
5 – veldig attraktiv	2.355	2.625	3.135	2.145	2.565

Slik tabell. 10 viser, vil deltakerne vurdere det som mer sannsynlig at bildene er redigert jo mindre attraktive de selv synes personen avbildet er. Dette er et noe uforventet funn, da det var forventet å finne en motsatt effekt – nemlig at jo mer attraktiv en deltaker synes personen avbildet er, jo større sannsynlighet er det at han/hun vurderer personen som redigert. Slik tabell. 10 derimot viser, er det verdien 1 (ikke særlig attraktiv) på attraktivitetsskalaen som har størst gjennomsnittlig effekt på tvers av de ulike redigeringstypene. Dette betyr at deltakeren vil gi en høyere verdi på sannsynlighetsestimering (2.956) på spørsmålet «*hvor sannsynlig er det at bildet er redigert?*» når de vurderer et bilde der personen avbildet ikke oppfattes som særlig attraktiv av deltakeren, på tvers av redigeringer. Når personen avbildet ble vurdert som 4 eller 5 (veldig attraktiv) på attraktivitetsskalaen, ser man at deltakeren tenderer til å vurdere det som mindre sannsynlig at personen avbildet er redigert, enn andre attraktivitetsvurderinger.

Forankret i forskningsspørsmålet F3, viser resultatene fra analysen og den gjennomsnittlige estimeringen at det finnes en sammenheng mellom deltakerens vurdering av personen avbildet som attraktiv og sannsynligheten for at de vurderer bildet som manipulert. Likevel er denne sammenhengen negativ. Det vil si at jo mindre attraktiv deltakeren vurderer personen som er avbildet, desto høyere vurderer deltakeren sannsynligheten for at bildet er redigert.

Mellom de ulike formene for attraktivitetsvurdering utmerker feminin skjønnhetsstereotypisk redigeringsform seg fra de andre. Deltakerne vurderer det generelt som mer sannsynlig at bilder presentert med denne redigeringsformen er manipulert satt opp mot andre, også med attraktivitetsvurderingen tatt til ettertanke. Det er også her man finner den største effekten mellom redigering og attraktivitetsvurdering mot sannsynlighetsberegning. Bildene som vurderes med 1 (ikke særlig attraktiv) og som er redigert med feminin skjønnhetsstereotypiske trekk fører til en sannsynlighetsvurdering på hele 3.879 ( $M=3.879$ ;  $95\%CI=[3.64; 4.12]$ ).

### 4.3.1 Påvirkning av bilde

I denne analysen vil det videre undersøkes hvilken påvirkning kjønnet som ble representert på bildet hadde på deltakernes attraktivitetsvurdering.

For å undersøke dette gjennomføres det en ANOVA-analyse med gjentatte målinger med deltakernes attraktivitetsvurdering som avhengig variabel, justert for interaksjon mellom hvilket kjønn som ble representert i bildet og type redigering (tabell. 11). Effekten av kjønnet til avbildet person har signifikant effekt på attraktivitetsvurdering:  $F(1, 1648)=25.93$ ;  $p<.001$  Det betyr at det finnes noen systematiske forskjeller mellom bildene som viser kvinner versus menn – der noen er vurdert som mer attraktive enn andre.

**Tabell 11:** Det representerte kjønnets påvirkning på attraktivitetsvurdering

Antall observasjoner = 2207		
	Kjønn av personen avbildet	
	Kvinne	Mann
Attraktivitetsvurdering	3.14	2.94

Slik tabell. 11 viser, vurderte deltakerne bildene som avbildet kvinner ( $M=3.14$ ;  $95\%CI=[3.08; 3.19]$ ) som mer attraktive enn bildene som avbildet menn ( $M=2.95$ ;  $95\%CI=[2.90; 3.00]$ ). Dette vil si at på attraktivitetsskalaen (1-5) plasserte deltakerne ofte sin vurdering av kvinner på 3.14, mens menn oftest fikk vurderinger 2.94.

### 4.3.2 Påvirkning av redigering

I denne analysen undersøker jeg hvilken påvirkning type redigering presentert i bildet hadde på deltakernes attraktivitetsvurdering.

For å undersøke dette ble deltakernes attraktivitetsvurdering brukt som avhengig variabel i en ANOVA-analyse med gjentatte målinger, justert for interaksjon mellom hvilket kjønn som ble representert i bildet og type redigering. Denne viser at effekten av redigering er statistisk signifikant:  $F(3, 1648)=67.42$ ;  $p<.001$ , noe som vil si at attraktivitetsvurdering er systematisk påvirket av de forskjellige typene bilderedigering. Etterfulgt av ANOVA-analysen kjøres en marginal prediksjon på redigeringstypen som uavhengig variabel (tabell. 12).

**Tabell 12:** Redigeringstypes påvirkning på attraktivitetsvurdering

Antall observasjoner = 2207

	<b>Orginalbilde</b>	<b>Lys/kontrast</b>	<b>Feminin stereotypisk</b>	<b>Maskulin stereotypisk</b>
Attraktivitetsvurdering	3.27	3.23	2.61	3.06

Det viser seg at attraktiviteten ble vurdert høyest i det originale bildet og i versjonen hvor bare lys og kontrast ble redigert. Deltakerne vurderte at det var signifikant mindre attraktivt når feminine skjønnhetsstereotypiske kvaliteter var fremhevet i bildet ( $M=2.61$ ;  $95\%CI=[2.53;2.68]$ ), mens maskuline skjønnhetsstereotypiske redigeringer ble vurdert som 3.06 på attraktivitetsskalaen blant deltakerne ( $M=3.06$ ;  $95\%CI=[2.99;3.14]$ )

### 4.3.3 Samspill mellom bilde og redigeringstype

For å undersøke samspillet mellom kjønn representert på bildet og type redigering, samt hvilken innvirkning dette har på deltakernes vurdering av personen avbildet som attraktiv eller ikke, ble det gjennomført en ANOVA-analyse med gjentatte målinger. Her var attraktivitetsvurderingen igjen avhengig variabel, og de uavhengige variablene ble justert for interaksjon mellom hvilket kjønn som var representert i bildet og type redigering. Denne analysen viser at det ikke eksisterer et statistisk signifikant samspill mellom kjønn representert i bildet og typen redigering som ble utført:  $F(3, 1648)=0.87$ ;  $p=.457$ , noe som betyr at innflytelsen av kjønn og redigering på attraktivitet er uavhengige. Dette viser seg også i marginal prediksjon og post-hoc analysen, der det er de samme typer redigeringer – maskuline og feminine skjønnhetsstereotypiske redigeringer – som fører til en lavere attraktivitetsvurdering blant deltakerne (tabell. 13). Det vil si at uansett hvilket kjønn personen avbildet på bildet har, blir attraktivitet vurdert lavest når utseende til personen var manipulert med feminine skjønnhetsstereotypiske trekk.

**Tabell 13:** Samspill mellom bilde og redigering, attraktivitetsvurdering

Antall observasjoner = 2207

Bildet viser:	<b>Orginalbilde</b>	<b>Lys/kontrast</b>	<b>Feminin stereotypisk</b>	<b>Maskulin stereotypisk</b>
Kvinne	3.37 <sup>d</sup>	3.26 <sup>d</sup>	2.68 <sup>ab</sup>	3.23 <sup>cd</sup>
Mann	3.17 <sup>cd</sup>	3.17 <sup>cd</sup>	2.53 <sup>a</sup>	2.91 <sup>bc</sup>

*Notat: Resultat av en ANOVA med repeterte målinger. Gjennomsnittsverdier (marginale prediksjoner) i tabellen som ikke deler bokstav er signifikant forskjellig fra hverandre på  $p<.05$ . Basert på Scheffé posthoc tester.*

#### 4.3.4 Innflytelse av andre faktorer

For å undersøke om andre faktorer slik som sosiodemografi, erfaring med sosiale medier og erfaringer/praktisering av bildedeling/redigering har en innvirkning på deltakernes attraktivitetsvurdering, gjennomføres to ulike regresjonsanalyser. Den ene presenteres i modell 1, mens modell 2 kontrollerer for de eksperimentelle forholdene bilde, redigeringstype og samspillet mellom disse.

**Tabell 14:** Andre faktorielle innflytelser på attraktivitetsvurdering

Antall observasjoner = 1408

Avhengige variable: attraktivitetsvurdering	Modell 1	Modell 2
Alder	-.007*	-.007**
Kvinne	.322***	.32***
Utdanning	.01	.007
Selv redigere bilder	.08***	.081***
Erfaring med sosiale medier	-.003	-.004
Daglig bruk av sosiale medier (min)	-.0002	-.0003
Kompetansevurdering	-.125***	-.122***
Bildet	—	3 koeffisienter
Redigering	—	3 koeffisienter
Bildet * Redigering	—	9 koeffisienter
(Constant)	3.408***	3.618***
R <sup>2</sup>	0.038***	0.18***

*Notat: \*\*\* betyr  $p < .001$ ; \*\* betyr  $p < .01$ ; \* betyr  $p < .05$ ; ingen markering betyr ikke statistisk signifikante resultater ( $p$ -verdi over 0.05). Jo flere stjerner jo mindre sjanse for type 1 feil.*

Modell 1 og 2 samsvarer i at det er fire bakgrunnsvariabler som har innflytelse på attraktivitetsvurderingen: alder, kjønn, det å selv redigere bilder og egen kompetansevurdering. Resultatene fra de to ulike modellene viser: (1) jo yngre deltakeren er, jo høyere er deres attraktivitetsvurdering. (2) Kvinner vurderer attraktiviteten høyere enn menn. (3) Jo oftere en deltakeren praktiserer redigering av bilder selv, jo høyere blir attraktivitetsvurderingen. (4) Høyere vurdering av egen kompetanse til å gjenkjenne redigeringer i bilder på sosiale medier fører til høyere attraktivitetsvurdering. Utdanning, erfaring med sosiale medier, og antall minutter bruk av sosiale medier hadde ikke statistisk signifikant innflytelse. Sammenlignet med eksperimentelle variasjoner hadde bakgrunnsvariabler bare en lav effekt på attraktivitetsvurderinger, der de kun forklarte  $R^2=3.8\%$  av variasjon i attraktivitetsvurderingen. Når man tar i hensyn bildene som deltakerne ble eksponert for stiger  $R^2$  til 18%. Det betyr at eksperimentelle variasjoner kan forklare 14.2% av variasjon (18%-3.8%).



#### 4.4 Deteksjonsevne

Her vil deltakernes evne til å korrekt estimere – eller gjette riktig – om bildet som presenteres er redigert eller ikke knyttes opp mot påvirkningen av kjønn representert i bildet, type redigering presentert i bildet og samspillet mellom disse. I møte med den avhengige variabelen «riktig gjettet» vil SDT sitt rammeverk vil kobles opp mot resultatene som analytisk perspektiv.

Den gjennomsnittlige deteksjonsevnen (eller treffsikkerheten) mellom alle de eksperimentelle gruppene er 33%, noe som vil si at kun 33% av deltakerne klarte å korrekt skille mellom manipulerede og ikke-manipulerede bilder. Den gjennomsnittlige prosentfordelingen mellom eksperimentgruppene er presentert i tabell. 15:

**Tabell 15:** Gjennomsnittlig deteksjonsevne mellom gruppene

Antall observasjoner = 3292

Deteksjonsevne	Gruppe A	Gruppe B	Gruppe C	Gruppe D	Totalt
	26.1%	33.5%	36.1%	.35.3%	33%

Deteksjonsevnen måles gjennom variabelen «riktig gjettet», og har til hensikt å undersøke hvor treffsikre deltakerne var i sin korrekte estimering av om bildet faktisk var redigert eller ikke.

##### 4.4.1 Påvirkning av bilde

For å undersøke hvilken påvirkning kjønn representert i bildet har på deltakernes deteksjonsevne, gjennomføres det en ANOVA-analyse med gjentatte målinger, hvor variabelen «riktig gjettet» fungerer som avhengig variabel. Analysen justeres for interaksjon mellom hvilket kjønn som ble representert i bildet og type redigering. Resultatene fra analysen viser:  $F(1, 2462)=43.02$ ;  $p<.001$ , noe som forteller at det finnes statistisk signifikante ulikheter mellom de to kjønnene som representeres i bildene. For å finne ut hvor stor denne forskjellen er, gjennomføres en marginal prediksjon (tabell. 16).

**Tabell 16:** Det representerte kjønnets påvirkning på deteksjonsevne

Antall observasjoner = 3292

	Kjønn av personen avbildet	
	Kvinne	Mann
Deteksjonsevne	.373	.283

Tabell. 16 presenter en markant forskjell (9%) mellom kvinner og menn representert i bildene. Hvis en kvinne er avbildet klarer deltakerne å korrekt bedømme om bildet er redigert eller ikke 37.3% av gangene ( $M=0.373$ ;  $95\%CI=[0.35; 0.39]$ ), satt opp mot en treffsikkerhet på 28.3% når menn er avbildet ( $M=0.283$ ;  $95\%CI=[0.26; 0.39]$ ). Med andre ord viser analysen at dersom deltakerne eksponeres for et bilde av en kvinne, øker sannsynligheten for at de vil klare å korrekt estimere om bildet er redigert eller ikke med 9% ( $0.373-0.283$ ).

Ut ifra denne tabellen kan man på lik linje med SDT sitt analytiske rammeverk tolke fire ulike utfall av deltakernes bedømmelse av bildene som redigerte eller ikke. Hvis man viderefører «signal» begrepet i SDT sitt rammeverk med «redigering», kan man se hvordan deltakerne enten fikk: 1) Treff, der bildet er redigert og deltakeren gjetter at det er redigert, 2) Bom, bildet er redigert men deltakeren gjetter at bilde ikke er redigert, 3) Falsk alarm, bildet er ikke redigert med deltakeren gjetter likevel at det er det, og 4) Korrekt avvisning, bildet er ikke redigert og deltakeren gjetter også at det ikke er redigert (Abdi, 2007).

Deltakerne opplever altså i større grad av det man i SDT kallen en «bom» (71.3%) i møte med menn avbildet, uavhengig av hvilken redigeringsform som presenteres sammen med kjønnen. I møte med bilder som representerer kvinner presterer deltakerne en bom 62.3% av gangene.

#### 4.4.2 Påvirkning av redigering

Her blir en ANOVA-analyse med gjentatte målinger gjennomført, hvor variabelen «riktig gjettet» fungerer som avhengig variabel og det justeres for interaksjon mellom kjønn representert i bildet og redigeringstype. Analysen forteller at deltakerens evne til å korrekt skille mellom redigerte og ikke-redigerte bilder er systematisk påvirket av typen redigering som presenteres i bildet:  $F(3, 2462)= 281.92$ ;  $p<.001$ . Tabell 17 viser hvilken i hvilken grad redigeringstype påvirker deteksjonsevnen:

**Tabell 17:** Redigeringstypes påvirkning på deteksjonsevne

Antall observasjoner = 3292

	<b>Originalbilde</b>	<b>Lys/kontrast</b>	<b>Feminin stereotypisk</b>	<b>Maskulin stereotypisk</b>
Deteksjonsevne	.649	.193	.334	.136

Ut ifra tabell. 17 kan man se at det originale bildet, altså bildet uten noen form for redigering, markant skiller seg ut fra resten av redigeringstypene når det kommer til påvirkning av deteksjonsevne ( $M=0.649$ ;  $95\%CI=[0.62; 0.68]$ ). I møte med det originale bildet vil deltakerne 64.9% av gangene klare å korrekt estimere at bildet ikke er redigert, altså en «korrekt avvisning» ifølge SDT sitt analytiske rammeverk. Dette kan videre tolkes som at deltakerne i møte med det uredigerte bildet også opplever en «falsk alarm», der 35.1% estimerte at bildet var redigert selv om et ikke var det ( $100\%-64.9\%=35.1\%$ ).

Som kontrast til dette viser tabellen at maskulin skjønnhetsstereotypisk redigering har den laveste effekten på 13.6% ( $M=0.136$ ; 95%CI=[0.11; 0.16]). Dette kan videre tolkes som at hele 86.4% av deltakerne viser det man ifølge SDT ville kalt en «bom» – eller «miss» (Abdi, 2007) – i møte med bildene som var redigert med maskulin skjønnhetsstereotypisk redigering. Her er bildet redigert – altså er et signal til stede – men deltakerne unngår i stor grad å registrere dette. Feminine skjønnhetsstereotypiske trekk viser en effekt på 33% ( $M=0.334$ ; 95%CI=[0.31; 0.36]). Sannsynligheten for at deltakerne klarer å gjenkjenne redigeringen når bildet er redigert med feminine skjønnhetsstereotypiske trekk, er med andre ord høyere enn de andre redigeringsformene. Sannsynligheten til å gjenkjenne redigeringen er derimot nesten halvparten så lav som på det originale bildet.

#### 4.4.3 Samspill mellom bilde og redigeringstype

For å måle hvilken innvirkning samspillet mellom kjønn representert i bildet og type redigering har på deltakernes evne til å suksessfullt bedømme om bildene er redigerte eller ikke gjennomføres nok en ANOVA-analyse med gjentatte målinger. Analysen er justert for interaksjon mellom representert kjønn og redigeringstype, og den avhengige variabelen er deltakernes deteksjonsevne, «riktig gjettet». Resultatene fra ANOVA-analysen viser et statistisk signifikant samspill mellom de to variablene «kjønn representert i bildet» og «type redigering»:  $F(3, 2462)=9.90$ ;  $p<.001$ . Med andre ord eksisterer det en avhengig innflytelse av kjønn og redigering på deltakernes deteksjonsevne. Det gjennomføres en marginal prediksjon for å videre undersøke hvordan denne effekten ser ut (tabell. 18).

**Tabell 18:** Samspill mellom bilde og redigering, deteksjonsevne

Antall observasjoner = 3292

Bildet viser:	Originalbilde	Lys/kontrast	Feminin stereotypisk	Maskulin stereotypisk
Kvinne	.674	.239	.452	.127
Mann	.625	.147	.216	.145

Slik det presenteres i tabell. 18 presterer deltakerne ganske likt i å korrekt estimere at det originale bildet ikke er redigert både når det er en mann eller en kvinne avbildet. Likevel ser man at sannsynligheten for at deltakerne estimerer riktig stiger med 5% (67.4%-62.5%) når det er en kvinne avbildet ( $M=0.674$ ; 95%CI=[0.63; 0.72]). Den mest markante ulikheten er mellom kvinner redigert med feminine skjønnhetsstereotypiske trekk ( $M=0.452$ ; 95%CI=[0.41; 0.50]) og menn redigert med feminine skjønnhetsstereotypiske trekk ( $M=0.216$ ; 95%CI=[0.17; 0.26]). Her øker deltakernes deteksjonsevne med 23.6% om det er en kvinne avbildet, en økning som kan betegnes som ganske stor. Overordnet viser tabell. 18 at deltakerne tenderer til å prestere bedre i å korrekt bedømme bildene som redigerte eller ikke når det er en kvinne

avbildet, bortsett fra når kvinnene er redigert med maskuline skjønnhetsstereotypiske trekk ( $M=0.127$ ;  $95\%CI=[0.08; 0.17]$ ). Selv om variasjonen mellom kjønnene er lav, viser analysen at deteksjonsevnen øker med 2% (14.5%-12.7%) når en mann er redigert maskulint satt opp mot en kvinne.

#### 4.4.4 Innflytelse av andre faktorer

Sosiodemografi, subjektiv kompetansevurdering og andre faktorer kan tenkes å ha en innflytelse på deltakernes deteksjonsevne. I denne sammenhengen står spesielt deltakernes erfaring med sosiale medier og erfaring med å selv redigere bilder sentralt, der innvirkningen disse variablene har på den avhengige variabelen kan gi svar på eksperimentets to hypoteser: H1 og H2. For å undersøke hvilken innflytelse disse faktorene har på deteksjonsevnen gjennomføres to ulike logistiske regresjonsanalyser, der modell 2 kontrollerer for de eksperimentelle forholdene bilde, redigeringstype og samspillet mellom disse (tabell. 19). Her rapporteres effekter av de uavhengige variablene gjennom odds ratio.

**Tabell 19:** Logistisk regresjon, andre faktorielle innflytelser på deteksjonsevne

Antall observasjoner = 1408

Avhengige variable: deteksjonsevne	Modell 1	Modell 2
Alder	.989*	.984*
Kvinne	1.154	1.243
Utdanning	1.008	1.047
Selv redigere bilder	1.02	1.004
Erfaring med sosiale medier	.994	.992
Daglig bruk av sosiale medier (min)	.999	.999
Kompetansevurdering	1.085	1.126
Bildet	—	3 koeffisienter
Redigering	—	3 koeffisienter
Bildet * Redigering	—	9 koeffisienter
(Constant)	1.119	5.919**
LR chi2	9.19	541.87***

*Notat: \*\*\* betyr  $p < .001$ ; \*\* betyr  $p < .01$ ; \* betyr  $p < .05$ ; ingen markering betyr ikke statistisk signifikante resultater ( $p$ -verdi over 0.05). Jo flere stjerner jo mindre sjanse for type 1 feil.*

Regresjonsanalysen presentert i tabell 19, modell 1, viser en  $Prob > chi2$  på 0.24, noe som vil si at tabellen ikke bærer forklaringskraft. Med dette menes at det ikke eksisterer noen bevis for at prediktorene i tabellen forklarer variansen i deltakernes deteksjonsevne. Modellen blir derimot signifikant når den kontrolleres for eksperimentelle kondisjoner, altså bilde, redigeringstype og deres samspill (tabell. 19, modell 2), med en  $Prob > chi2$  på 0.000. I I modell 2 er det de eksperimentelle faktorene (bilde, redigering

og bilde\* redigering) som bærer forklaringskraften, mens de resterende forklaringsvariablene forblir ikke signifikante. Dette er med unntak av alder, som i modell 2 viser en signifikant negativ sammenheng med deltakernes deteksjonsevne. Dette vil si at oddsen for å gjette riktig på om bildet er redigert eller ikke, predikeres til å synke for hvert år eldre en person blir (oddsen for at en person som er 1 år eldre predikeres å være omtrent 0,98 ganger så stor som for en person som er 1 år yngre). Altså vil dette si at jo yngre en deltaker er, jo flinkere vil de være til å korrekt skille mellom redigerte eller ikke redigerte bilder.

Både hypotese H1 og H2 motbevises i analysen presentert i tabell 19. Mer erfaring med sosiale medier viser ikke en positiv effekt på evnen til å skille mellom manipulerte og ikke-manipulerte bilder (H1), og tidligere erfaring med å selv redigere bilder viser heller ikke noen positiv effekt på evnen til å skille mellom manipulerte og ikke-manipulerte bilder (2). Alder viser derimot en signifikant negativ effekt på deteksjonsevnen, noe som var forventet på bakgrunn av at bruk av digitale verktøy, erfaringer med sosiale medier og god kjennskap til diverse digitale plattformer ofte er faktorer knyttet til de yngre i samfunnet.

## 4.5 Oppsummering av sentrale funn

I følgende avsnitt oppsummeres analysens mest sentrale funn.

Kvinner avbildet i det eksperimentelle materialet bemerker seg gjennom de fleste analysene i eksperimentets resultater. Kjønnetts viser er en større påvirkningskraft både når det kommer til den subjektive vurderingen av redigering (Kvinner: 2.88; Menn: 2.55), deltakernes attraktivitetsvurdering (Kvinner: 3.14; Menn: 2.94) og deteksjonsevne (Kvinner: 37.3%; Menn: 28.3%). Dette gjelder også når man måler for interaksjon mellom kjønnet representert og redigeringstypen, der kvinner som ble redigert med feminine skjønnhetsstereotypiske trekk utmerket seg både i sammenheng med deltakernes deteksjonsevne og sannsynlighetsvurdering satt opp mot det motsatte kjønn og andre redigeringsformer – samspill mellom kjønnet representert og redigeringstype opp mot deltakernes attraktivitetsmåling fant ingen signifikant sammenheng.

Funnene fra analysen viser også at feminine skjønnhetsstereotypiske trekk markere seg fra de andre redigeringsformene lys/kontrast og maskulin skjønnhetsstereotypisk redigering. Her utmerker redigeringsformen seg som høyest i påvirkningen av deltakernes subjektive sannsynlighetsvurdering (feminin: 3.33; maskulin: 2.44; lys/kontras: 2.77) og i påvirkning av deltakernes deteksjonsevne (feminin: 33.4%; maskulin: 13.6%; lys/kontras: 19.3%), men desidert lavest når det kommer til påvirkningen av deltakernes attraktivitetsvurdering (feminin: 2.61; maskulin: 3.06; lys/kontras: 3.23). Med andre ord så fører redigeringsformen med feminin skjønnhetsstereotypiske trekk til at deltakerne (1) vurderer sannsynligheten for at bildet er redigert som høyere satt opp mot andre redigeringsformer, (2) at deltakerne oftere klarer å korrekt estimere om bildet er redigert eller ikke, og (3) at deltakerne vurderer personen avbildet som mindre attraktiv satt opp mot bildene redigert med annen redigeringstype.

Når det kommer til deteksjonsevne viste kontrollgruppen (originalbildet) en særdeles høy

effekt, da deltakerne i møte bildene som ikke var redigerte klarte å korrekt estimere at bildene ikke inneholdt noen redigering 64.9% av gangene. Satt opp mot redigeringsformen med høyest effekt (feminine skjønnhetsstereotypiske redigeringen med 33.4%) er effekten på deteksjonsevnen målt blant eksperimentets originale bildeversjoner nesten dobbelt så høy. Redigeringsformene med lavest effekt på deltakernes evne til å korrekt estimere om bildet var redigert eller ikke, var redigeringsformene lys/kontrast (19.3%) og maskulin skjønnhetsstereotypisk redigering (13.6%). I møte med disse to redigeringsformene viste deltakerne en høy grad av feilvurdering, henholdsvis en feilvurdering på 80.7% i møte med redigeringen lys/kontrast ( $100\% - 19.3\% = 80.7\%$ ) og 86.4% ( $100\% - 13.6\% = 86.4\%$ ).

Den eneste bakgrunnsvariabelen som viser en statistisk signifikant innflytelse både på subjektiv sannsynlighetsberegning av redigering i bildet, vurdering av personen avbildet som attraktiv og deltakernes deteksjonsevne er alder. Alder viser en negativ korrelasjon med alle de tre avhengige variablene fra analysen, med sterkest innflytelse på deltakernes sannsynlighetsvurdering (-.012). Dette vil si at eksperimentets yngre deltakere både tenderer til å vurdere sannsynligheten for at bildene er redigert som høyere, vurdere personene avbildet som mer attraktiv og har også en høyere treffsikkerhet når det kommer til å skille mellom redigerte og ikke-redigerte bilder.

Deltakernes subjektive vurdering av egen kompetanse til å gjenkjenne redigeringer var også en av bakgrunnsvariablene som viste en statistisk signifikant effekt gjennom flere avhengige variabler, henholdsvis deltakernes sannsynlighetsvurdering og attraktivitetsvurdering. Deltakernes subjektive kompetansevurdering hadde en positiv korrelasjon mellom den subjektive sannsynlighetsvurderingen av redigering, der deltakere som vurderte egen kompetanse høyere, også vurderte det som mer sannsynlig at bildene var redigerte. Ved deltakernes attraktivitetsvurdering viser den subjektive kompetansevurderingen en negativ effekt, noe som vil si at jo lavere deltakerne vurderte sin egen kompetanse, jo høyere var deres sannsynlighetsvurdering for redigering.

## 5. Diskusjon

Denne formen for eksperimentell undersøkelse av menneskers evne til å gjenkjenne redigeringsformer på sosiale medier er den første av sitt slag i Norge. Så langt min egen kunnskap strekker er eksperimentet også det første av sitt slag forbi Norges landegrensar, der tidligere forskning ikke har inkludert kosmetiske redigeringsformer slik de ofte presenteres på sosiale medier som eksperimentelle faktorer. Målet med dette eksperimentet var å undersøke hvordan folk vurderer bilder på sosiale medier som redigerte eller ikke, og å finne ut hvilke faktorer som har innflytelse på evnen til å gjenkjenne redigeringer i bilder. I dette kapitlet vil eksperimentets funn bli brukt som grunnlag for videre diskusjon opp mot teori og tidligere forskning på feltet.

Innledningsvis ble uttalelsen fra den amerikanske universitetsprofessoren Henry Farid (referert i Wen, 2020) presentert, hvor han hevder at selv om man kanskje har høy tiltro til egne evner, er ikke folk flest spesielt flinke til å skille mellom bilder som er redigerte og bilder som ikke er det: «So people are both ignorant and confident, which is the worst combination» (Fraid, referert i Wen, 2020). Påstander slik som denne reflekteres i den tidligere forskningen på feltet, der det gang på gang er blitt vist hvordan folk flest ikke er spesielt flinke til å gjenkjenne redigeringer i bilder. Påstanden kan videre bekreftes i dette aktuelle eksperimentet. Overordnede resultater viser at deltakerne korrekt klarte å skille mellom manipulte og ikke-manipulte bilder 33% av gangene, et resultat som kan vurderes som mindre imponerende. Dette vil altså si at deltakerne presterer det man i SDT kaller en «bom» – eller «miss» – hvor det er et signal (en redigering) tilstede, men deltakerne tror det ikke er det 67% av gangene. Spesielt dominerende her er feilvurderingen av de to redigeringsformene «lys/kontrast» (80.3%) og «maskulin skjønnhetsstereotypisk» (86.4%).

Feilvurderingene fra deltakernes side kan tolkes som en generell usikkerhet i møte med bildene som presenteres. Dette kan reflekteres både i den gjennomsnittlige skåren deltakerne ga på spørsmålet om hvor sannsynlig de trodde det var at bildet var manipulert på 2.72 og den lave prosentandelen (33%) av korrekt deteksjon av redigering/ikke-redigering. En verdi på 2,72 ansees som ganske gjennomsnittlig på en skala fra 1 til 5, der verdi «3» opererer som verdien i midten. Det kan ikke argumenteres for at det gjennomsnittlige resultatet på 2.72 reflekterer verken en høy eller lav grad av sikkerhet fra deltakernes side. Gjennomsnittlig velger altså deltakeren å legge seg «midt på treet» heller enn å påstå for mye i den ene eller andre retningen. Tendensen til usikkerhet fra deltakernes side i møte med eksperimentets stimulusmateriale kan forklares gjennom det som kalles en «central tendency bias», eller *sentral tendens bias*. Det refererer til tendensen mennesker kan ha til å plassere sin vurdering på midten av en skala, heller enn å plassere vurderingen på skalaens mer «ekstreme» ytterpunkter (Statology, 2021). Dette er likevel en bias som gjerne finner sted i situasjoner der deltakeren har et personlig forhold til materialet som vurderes. I tilfeller hvor deltakeren ikke er personlig tilknyttet materialet, pleier ikke tendensen å være like markant (Statology, 2021). I denne undersøkelsen skal ikke deltakerne ha noen personlige forhold til materialet som vurderes, da dataene er innsamlet fra et helt tilfeldig utvalg. Likevel kan man spekulere i om deltakernes vurderinger som «midt på treet» er tilfeldig eller ikke.

De som vurderte sin egen kompetanse til å gjenkjenne redigeringer i bilder på sosiale medier som høy, tenderte derimot til å vurdere sannsynligheten for at bildet var redigert som høyere enn deltakerne som ga seg selv en lavere kompetansevurdering. Med andre

ord er disse deltakerne mer skeptiske enn andre, selv om det ikke blir funnet noen signifikant sammenheng mellom denne mistenksomheten og deltakernes evne til å gjenkjenne redigeringer. Overordnet viser faktisk resultatene – i kontrast til Henry Farids påstand – at folk flest heller tenderer til å vurdere sin egen kompetanse som lavere enn høyere.

Man står igjen med et sentralt spørsmål: hvilke faktorer er det egentlig som har en innvirkning på deltakernes evne til å gjenkjenne redigeringer? I et forsøk på å besvare dette spørsmålet vil de videre avsnittene diskutere eksperimentets sentrale funn opp mot gjeldende teori og tidligere forskning. Kjønns effekt vil bli diskutert på bakgrunn av den store innvirkningen kvinner presentert i bildene hadde på flere aspekter ved eksperimentet. Videre vil alder diskuteres opp mot funnene rundt deltakernes visuelle kompetanse, før funnene som handler om deltakernes attraktivitets- og sannsynlighetsvurdering vil bli diskutert opp mot kognitiv dissonans og sosial sammenligning.

## 5.1 En kjønn vurdering

Et av de mest sentrale funnene i eksperimentet er hvordan det avbildede kjønn har en innvirkning på deltakernes evne til å gjenkjenne redigeringer. Hvilket kjønn som er avbildet viser stor effekt gjennom hele analysen, både når det kommer til subjektive vurderinger av redigering, attraktivitetsvurdering og deltakernes deteksjonsevne. Effekten vises også i samspillet mellom kjønn representert og redigering både når det kommer til sannsynlighetsvurdering og deteksjonsevne. Analysen viser at folk er mer mistenksomme ovenfor bilder av kvinner satt opp mot bilder av menn. I tillegg viste kvinner redigert med feminine skjønnhetsstereotypiske trekk en markant høyere sannsynlighetsvurdering blant deltakerne satt opp mot alle andre kombinasjoner av kjønn og skjønnstereotypiske kosmetiske redigeringer. Dette kan blant annet sammenlignes med menn som ble redigert med maskuline skjønnhetsstereotypiske trekk, som så å si presterte likestilt med bildets original-versjon. Redigeringsformen viste altså ingen markant effekt.

Her er det verdt å nevne at redigeringen «feminine skjønnhetsstereotypiske trekk» også viser en forskjell når den brukes for å redigere bildene som representerer menn. Likevel later ikke dette til å bekrefte antakelsen om at redigeringsformer som strider imot kjønnets skjønnhetsstereotyper vil gjøre det mer tydelig for deltakerne at bildet er redigert (F2). Her viser for eksempel bilder av kvinner redigert med maskuline skjønnhetsstereotypiske trekk nesten ikke noen forskjell satt opp mot det ikke-redigerte bildene av kvinner. Analysen viser heller at kvinner redigert med feminine skjønnhetsstereotypiske trekk er faktoren som i stor grad skiller seg ut.

Dette kan indikere at feminine redigeringsformer er enklere for deltakerne å oppdage på grunn av tidligere erfaringer og større eksponering for denne formen for redigering på for eksempel sosiale medier. Det kan også indikere at deltakerne har en forventning om at det å redigere bilder forekommer oftere blant kvinner enn blant menn. Det kan også være at det eksisterer spesielle karakteristikk ved denne redigeringsformen, for eksempel at redigeringsformen ikke er spesielt godt utført, som gjør det enklere for deltakerne å oppdage. Om sistnevnte er tilfellet, ville det vært forventet å finne en lik fordeling mellom feminine skjønnhetsstereotypiske trekk og kjønn. Dette er likevel ikke



tilfellet, noe som kan indikere at effekten ligger i kjønnnet som faktor, heller enn i redigeringsformen.

Disse funnene besvarer eksperimentets forskningsspørsmål F1, der det kan virke som at folk bærer med seg forutbestemte holdninger i møte med bilder av kvinner. Slik tidligere forskning viser, tar kvinner både flere bilder av seg selv og redigerer seg selv oftere enn menn (Qiu et al., 2015; Dhir et al., 2016; Fox & Vendemia, 2016). Slike faktorer, kombinert med bruken av sosiale medier som rom for «utstilling» – eller «exhibition» (Hogan, 2010) – og selektive former for selvrepresentasjon (Rose et al., 2012) kan ha ført til at folk flest har forestillinger om at kvinner ofte redigerer bilder av seg selv. Dette kan igjen gjøre at folk er mer oppmerksomme på slike redigeringer i møte med bilder av kvinner.

## 5.2 Et domene forbeholdt den yngre generasjon

Når det kommer til faktorer som visuell kompetanse og hvilken innvirkning det har på vurderingen av bilder som redigerte eller ikke, viser analysen varierende funn. I motsetning til antakelsen som presenteres i H1 og H2 om at mer erfaring med sosiale medier eller erfaring med å redigere bilder har en positiv effekt på evnen til å skille mellom manipulerte og ikke-manipulerte bilder, viser de innsamlede dataene at dette ikke stemmer. Slik det blir presentert i analysen har ikke disse variablene noen forklaringskraft opp mot deltakernes deteksjonsevne, selv om deltakernes kunnskap om å redigere bilder viser en positiv effekt når det kommer til deres vurdering av personen avbildet som attraktiv. En variabel som viste seg å ha positiv effekt på deltakernes evne til å skille mellom manipulerte og ikke-manipulerte bilder, var alder. Til tross for manglende signifikante funn på variablene «selv redigere bilder» og «erfaring med sosiale medier», viser resultatene at jo yngre deltakeren er, jo større er sannsynligheten for at han/hun korrekt klarer å skille mellom redigerte og ikke-redigerte bilder. Det er ikke et sjokkerende funn at den yngre generasjonen, som også blir ansett å ha høy visuell kompetanse, er de som i størst grad klarer å skille mellom bilder som er redigert og ikke. Generasjonen med «digitalt innfødte» er – slik Brumberg (2011) beskriver det – ansett som mer teknologisk kyndige enn andre generasjoner. Dette er utsagn som stammer 10 år tilbake i tid, og det kan forventes at den visuelle kompetansen til yngre også har økt siden da. Dette kan forankres i kontrastene som finnes mellom Brumberg (2011) sin studie og dette eksperimentet. Brumberg (2011) fant for eksempel at den yngre generasjonen ikke var spesielt dyktige til å skille mellom redigerte og ikke-redigerte bilder, mens dette eksperimentet definitivt finner en sammenheng mellom yngre deltakere og deteksjonsevne. Siden den yngste alderen registrert i dette eksperimentet er 23 år, kan det også hende at effekten av alder hadde vært enda større dersom yngre deltakere hadde blitt inkludert. I dette eksperimentet var det derimot ikke mulig å inkludere unge under 18-år, da det ville ha krevd en mer omfattende prosess knyttet til datainnsamling.

Det at yngre skårer signifikant høyere på deteksjonsevne, til tross for at erfaring med sosiale medier og det å selv redigere bilder ikke når opp til en signifikant forskjell, skaper et resultat som er noe vanskelig å tolke teoretisk. Som beskrevet tidligere var det forventet at den yngre generasjonens høye bruk av sosiale medier og kjennskap til digitale verktøy ville ha en innvirkning på deres evne til å gjenkjenne redigeringer.

Likevel er det kun en yngre alder som her gir utslag, noe som kan bety at: (1) effekten av alder på deteksjonsevne kan forklares ut fra andre faktorer som ikke er blitt tatt stilling til i dette eksperimentet, (2) det at yngre viser en større deteksjonsevne stammer fra tilfeldigheter, eller (3) selv om den yngre generasjonens evne til å gjenkjenne redigeringer ikke kan forklares ut ifra deres økte erfaring med sosiale medier og redigeringsverktøy i dette eksperimentet, kan et større utvalg i en senere undersøkelse gi signifikante utslag. Her kan også kjennskap til redigeringer på sosiale medier spille en rolle. Det kan hende at yngre generelt sett er mye mer kjent med fenomenet om at folk redigerer bilder på sosiale medier, også om de ikke praktiserer slik redigering selv. Det kan også ha noe å gjøre med at den yngre generasjonen oppholder seg mer på sosiale medier, og i større grad oppholder seg rundt personer som redigerer bilder oftere enn dem selv. Dette kan igjen føre til en større skepsis i møte med bilder blant den yngre generasjonen.

At alder forblir signifikant gjennom eksperimentets ulike analyser, men ikke variablene ment for å måle visuell kompetanse blant deltakerne, er noe overraskende. Likevel er det viktig å understreke at analysen ikke nødvendigvis presenterer en sann refleksjon av virkeligheten. Spørsmål om alder er noe de fleste enkelt kan svare på. Når det kommer til andre faktorer, for eksempel erfaring med sosiale medier eller det å selv redigere bilder, er dette måleenheter som ikke er like enkle for deltakere å besvare. Å besvare slike spørsmål krever en form for selv-estimering fra deltakerne som ikke bærer noen garanti for å være korrekt. Det kan også hende at det eksisterer andre faktorer som har en innvirkning, men som ikke blir tatt hensyn til i dette eksperimentet. Ytterligere forskning bør gjennomføres for å finne ut om det egentlig bare er alderen som har utslag, eller om det er andre underliggende faktorer.

### **5.3 Lavere attraktivitetsvurdering og større skepsis**

Tidligere i den teoretiske gjennomgangen ble det presentert to mulige utfall når det kommer til deltakernes møte med en «tvunget» attraktivitetsvurdering og dens innvirkning på deres sannsynlighetsberegning av bildene som redigerte eller ikke. På den ene siden kunne deltakeren oppleve en dissonans gjennom sammenligning av individet avbildet og eget selv-konsept, og fylle dissonansen med en antakelse av redigering. På den andre siden ville ikke deltakeren oppleve dissonans gjennom sammenligning av individet avbildet og eget selv-konsept, og ingen sammenheng mellom attraktivitets- og sannsynlighetsvurdering ville bli oppdaget. Likevel peker eksperimentets resultater i en annen retning, nemlig at jo mer attraktiv deltakeren selv vurderer personen avbildet til å være, jo mindre sannsynlig tror deltakeren det er at bildet er redigert. Dette resultatet åpner opp for en ny diskusjon av konseptet kognitiv dissonans sett i sammenheng med redigering av bilder og sosiale medier. Det er spesielt interessant sett opp mot hvilke funn som var forventet i eksperimentets sammenheng.

Hovedprinsippet i dissonansteorien ligger i ubehaget et individ opplever når to ulike kognisjoner er inkonsistente, noe som kan føre til at individet legger til en kognisjon for å redusere det opplevde avviket (Festinger, 1962). Likevel viser resultatene at om deltakeren opplever et avvik mellom en sosial sammenligning eller ikke, er ikke kognisjonen som legges til en antakelse om at bildet er redigert – heller er det en

antakelse om at det ikke er det. Her er det igjen sentralt å nevne at eksperimentet ikke gir noen indikasjoner på om sosial sammenligning faktisk finner sted, om denne eventuelle sammenligningen er oppad eller nedadgående eller om deltakeren selv ville betegnet veien mellom deres attraktivitets- og sannsynlighetsvurdering som påvirket av noen form for kognitiv dissonans. Det er likevel spennende å se at forholdet mellom deltakernes attraktivitets- og sannsynlighetsvurdering faktisk viser en signifikant effekt, og at denne effekten strider imot tanken om at oppadgående sosiale sammenligninger har en sammenheng med nedgang i selvtillit (Morse & Gergen, 1970; Lup et al., 2015). Dette kan igjen føre til at en kognisjon blir lagt til for å beholde det Aronson ville kalt selvkonsistens (1968, referert i Cooper, 2012). En mulig forklaring kan ligge i online-spørreskjemaets struktur.

Under utformingen av spørreskjemaet var det som tidligere nevnt ønsket å først spørre deltakerne om hvor attraktiv de synes personen avbildet var, for så å spørre om en sannsynlighetsvurdering på om bildene var redigerte eller ikke. Årsaken var et ønske om at attraktivitetsvurderingen skulle bli gjort med et så nøytralt grunnlag som mulig. Likevel skapte denne strukturen forvirring blant deltakerne, og måtte endres i spørreskjemaets ferdigstilte versjon. Det at deltakerne i eksperimentet først ble bedt om å vurdere sannsynligheten for at bildet var redigert, for så å vurdere attraktivitet, kan ha hatt en innvirkning på resultatene. Det kan for eksempel tenkes at en deltaker som har vurdert det som «svært sannsynlig at bildet er redigert», ikke vil vurdere personen som spesielt attraktiv når vedkommende får spørsmål om det i etterkant – da deltakeren allerede har fastslått at personens utseende ikke er «ekte». En annen mulighet er at sammenhengen er tilfeldig, og at en senere måling vil gi andre resultater.

Resultatene fra eksperimentet gir grunnlag for å besvare forskningsspørsmålet F3: «Finnes det en sammenheng mellom deltakerens vurdering av personen avbildet som attraktiv og sannsynligheten for at de vurderer bildet som manipulert?». Om denne sammenhengen har et teoretisk grunnlag eller ikke, er derimot noe vanskelig å fastslå ut ifra informasjonen som er innhentet i dette eksperimentet.

#### **5.4 Begrensninger og videre forskning**

I de følgende avsnittene vil eksperimentets begrensninger diskuteres. Her vil også ulike temaer som videre ville vært interessant å undersøke i senere forskningsprosjekter trekkes frem. Avslutningsvis vil det bli gitt en kort introduksjon av andre måter man videre kan forske på gjenkjenning av redigeringer.

Eksperimentet hadde som mål å undersøke norske borgere mellom 18 og 90 år, og deres evne til å gjenkjenne redigeringer i bilder på sosiale medier. Likevel er ikke aldersgruppen mellom 18 og 23 år representert i eksperimentets utvalg. Dette kan ansees som en svakhet i utvalget da lavere alder var en av faktorene som viste effekt på alle analysens avhengige variabler. Det ville derfor vært interessant å se hvilken innvirkning flere observasjoner fra yngre deltakere vill hatt på disse effektene.

Om denne tematikken undersøkes på nytt, kan det være interessant å også inkludere enda yngre aldersgrupper inn i prosjektet (under 18-år). Funnene fra denne undersøkelsen som indikerer at den yngre generasjonen markerer seg som flinkere til å

skille mellom redigerte og ikke-redigerte bilder, forventes å utstrekke seg også til enda yngre aldersgrupper. Det er ingen tvil om at den sosiale mediebruken blant unge er godt utbredt, og det ville ha vært interessant å se hvilken effekt enda yngre brukere ville hatt på eksperimentets avhengige variabler.

Ulike måter å øke realismen i eksperimentet kunne også blitt utforsket. Tidligere ble det nevnt at ved å stille spørsmålet «*hvor sannsynlig er det at bildet er redigert?*», ble også mest sannsynlig oppmerksomheten til deltakerne rettet mot redigeringer i bildene i større grad enn de kanskje ville gjort i virkeligheten. Det er ikke nødvendigvis slik at deltakerne er fokusert på å vurdere om bilder er redigert eller ikke i møte med bilder på sosiale medier i hverdagen. Det er derfor forventet at resultatene i eksperimentet er påvirket av dette, og eksperimentets funn kan derfor ikke generaliseres til hvordan folk bruker sosiale medier generelt i hverdagen. Et inngrep som kunne blitt gjort for å videre øke realismen i eksperimentet, er å gi det en sterkere tilknytting opp mot de ulike sosiale medieplattformenes visuelle fremstilling slik som å presentere bilder i samme format som man gjerne ser på plattformen. Et eksempel på et tidligere eksperiment som har gjort dette er Shen og kolleger (2019) sitt prosjekt «*Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online*». Det kan tenkes at det å presentere eksperimentets stimulusmateriale i en ramme deltakerne er vant til å observere bilder i (f.eks via en Facebook-feed eller Instagram profil), kunne gjort eksperimentet mer virkelighetsnært. Her ville det også vært interessant å se om ulike innramminger av stimulusmaterialet, for eksempel en Facebook-feed vs. en Instagram profil, ville ført til ulike utfall av sannsynlighetsvurdering og deteksjonsevne. Selv om dette kunne styrket eksperimentets ytre validitet, antas det at dette likevel ikke vil føre til mulighet for generalisering til hverdagens bruk av sosiale medier. Det er vanskelig å skape en høy nok grad av realisme i et eksperiment som er ment til å måle gjenkjenning av redigeringer i bilder på sosiale medier, da deltakernes bruk av slike plattformer i hverdagen er ekstremt individuelt. Hvilke bilder og innhold de bli eksponert for på sine private profiler er vanskelig å gjenskape i eksperimentelle forhold, noe som videre kan legge grunnlaget for å utforske andre forskningsdesign i møte med tematikken.

Slik resultatene har vist, viser deltakerne en viss usikkerhet i møte med deres sannsynlighetsvurdering av stimulusmaterialet som redigert eller ikke. Om dette dreier seg om tilfældigheter, sentraltendens-bias eller en refleksjon av genuin usikkerhet blant befolkningen er vanskelig å tolke ut ifra dette eksperimentets resultater alene. På bakgrunn av dette ønsker jeg å bruke den siste delen av diskusjonen på å kort introdusere alternative måter å undersøke menneskers evne til å oppdage manipuleringer på sosiale medier, foruten et eksperimentelt design.

Selv om eksperimenter har styrken av å finne årsakssammenhenger, bærer metoden også med seg en svakhet der informasjon om fenomenet som undersøkes kan bli borte. Satt opp mot en kvalitativ studie, for eksempel dybdeintervjuer, er innsikten som genereres fra dette eksperimentet noe mangelfull. Det kunne vært spennende å innhente kunnskap om akkurat *hvorfor* deltakerne vurderte noen bilder som redigerte og andre ikke, samt *hvordan* ulike signaler på redigering i bildene tolkes. Denne typen innsikt er vanskelig å innhente fra kvantitative datainnsamlingsmetoder, noe som åpner opp for muligheten av å videre undersøke tematikken gjennom kvalitative metoder slik som dybdeintervjuer. En slik dyptgående kunnskap kunne supplert dette eksperimentet, og kanskje også oppdaget nye fenomener knyttet til tematikken som tidligere ikke er kjent.

Til tross for eksperimentets begrensinger, vil jeg hevde at eksperimentet fortsatt er relevant for forskningsfeltet deteksjon av redigeringer i bilder. Eksperimentet bidrar med observasjoner fra et tilstrekkelig utvalg av den norske befolkningen, og tilbyr på den måten en økt forståelse av et tidligere begrenset forskningsfelt: hvilke faktorer er det som påvirker menneskers evne til å gjenkjenne redigeringer i bilder. Dette prosjektet kan bidra til å legge et grunnlag for videre studier av temaet.

## 6. Konkusjon

Utviklingen av relativt enkle fotoredigeringsverktøy som FaceTune betyr at nesten hvem som helst kan redigere bilder på en overbevisende måte før de legges ut på sosiale medier. Til tross for den økende tilstedeværelsen av redigerte bilder i hverdagen vår, er det fortsatt lite forskning vedrørende i hvilken grad mennesker klarer å gjenkjenne slike redigeringer eller ikke. Enda mer pressende er kanskje den totale mangelen på forskning når det kommer til menneskers evne til å gjenkjenne kosmetiske redigeringer av mennesker på sosiale medier. I dette eksperimentet ble det funnet at folk flest har en ganske begrenset evne til å gjenkjenne redigering i bilder på sosiale medier. Dette går på tvers av ulike redigeringstyper, selv om deltakerne presterte desidert høyest i møte med feminine skjønnhetsstereotypiske redigeringer. Eksperimentet har også vist at menneskers evne til å gjenkjenne redigeringer ikke er upåvirket av samfunnets innflytelse, da kvinner avbildet viser en stor innvirkning på deltakernes deteksjonsevne, subjektive attraktivitetsvurdering og deres vurdering av hvor sannsynlig det er at bildet er redigert eller ikke.

Den mer dystre spådommen fra forskere innen visuelle medier om at man ikke lenger kan stole på fotografier som gjenspeilinger av virkeligheten etter introduksjonen av Photoshop og andre manipuleringsverktøy (Mitchell, 1992; Ritchin, 2008), bærer kanskje et snev av sannhet. Likevel er dette en holdning som heller ser ut til å være forbeholdt spesielt mediekynlige, da det ikke later til å være en mistanke som den større delen av befolkningen bærer med seg i møte med bilder på nettet. Funnene fra dette eksperimentet kan reflektere et behov for å øke vår kunnskap rundt redigeringer av bilder som deles på sosiale medier. Dette omhandler ikke alene risikoen om feilinformasjon, men også den mulige innvirkningen uvitenhet kan ha når det kommer til kroppspress og selvbilde blant brukere av sosiale medier. Når både massemedier og sosiale medier ansees som en av de store bidragsyterne når det kommer til å forsterke usunne skjønnhetsidealer (López-Guimerà et al., 2010; Aparicio-Martinez et al., 2019), er det ikke vanskelig å anta at de kosmetiske redigeringsene som finner sted blant brukere av sosiale medier forsterker dette. I hvert fall sett i forhold til folks manglende evne til å skille mellom slike redigeringer og virkeligheten. Det kan være at folks manglende skepsis i møte med bilder på disse plattformene fører til en usunn sirkel. Siden usunne skjønnhetsstereotyper har sirkulert i medielandskapet i årrekkker, kan dannelsen av sosiale medier som en plattform for «utstilling» (Hogan, 2010) ha ført til et økt press om å vise seg selv fra sin beste side og følge de gjeldende stereotypene medielandskapet reflekterer (Slutsky, 2020). Dette kan føre til en internalisering av idealiserte fremstillinger. Om man som individ ikke klarer å skille mellom hva som er redigert og hva som ikke er det, kan det igjen føre til en farlig antakelse om at det som egentlig er redigerte fremstillinger av idealiserte stereotyper er en sann refleksjon av virkeligheten. En mulig konsekvens kan være at flere velger å redigere bilder av seg selv for å passe inn i det stereotypiske rammeverket, for så å starte sirkelen på nytt igjen.

Samfunnet har behov for individer som er kompetente til å vurdere bildene som presenteres på sosiale medieplattformer, og individer som er kritiske og villige til å stille spørsmål ved autenticiteten til det som presenteres. I og med at «alle» har tilgang til å redigere bilder, samt tilgang til å dele innholdet med et stort antall mennesker, har nok redigerte bilder en større plass i vår hverdag enn de fleste kanskje tror. Selv om influencere og andre reklameaktører på sosiale medier er lovpålagt til å markere redigert innhold, er det fortsatt fritt fram for «vanlige» brukere. Det å belære folket til å bli

kompetente og kritiske mottakere av medieinnhold kan være sentralt for å unngå ubehagelige konsekvenser som internalisering av usunne skjønnhetsstereotyper, men også andre alvorlige konsekvenser som spredning av falsk informasjon. Ved å i større grad inkludere visuell kompetanse som en del av skolesystemet, kan det tenkes at om et lignende eksperiment hadde blitt gjennomført om 10 år, ville resultatene ha endret seg til å reflektere både en visuelt kompetent og kritisk befolkning.

## Kilder

Abdi, H. (2007). Signal Detection Theory (SDT). I, B. McGaw, P.L. Peterson & E. Baker (Red.), *Encyclopedia of Education* (3. Utg., s.1-9). New York: Elsevier. Hentet fra: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.143.2749&rep=rep1&type=pdf>

Advisory Committee on Equal Opportunities for Women and Men: Opinion on "Breaking gender stereotypes in the media". (2010). *European Commission Employment, Social Affairs and Equal Opportunities*. 1-23. Hentet fra: [https://eige.europa.eu/docs/3093\\_2010-12\\_Opinion\\_on\\_Breaking\\_gender\\_stereotypes\\_in\\_the\\_media\\_EN%5B1%5D.pdf](https://eige.europa.eu/docs/3093_2010-12_Opinion_on_Breaking_gender_stereotypes_in_the_media_EN%5B1%5D.pdf)

Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience. *Frontiers in Psychology*, (6). <https://doi.org/10.3389/fpsyg.2015.00762>

Aparicio-Martinez, P., Perea-Moreno, A. J., Martinez-Jimenez, M. P., Redel-Macías, M. D., Pagliari, C., & Vaquero-Abellan, M. (2019). Social Media, Thin-Ideal, Body Dissatisfaction and Disordered Eating Attitudes: An Exploratory Analysis. *International journal of environmental research and public health*, 16(21), 4177. <https://doi.org/10.3390/ijerph16214177>

Araújo, C. S., Meira, W., & Almeida, V. (2016). Identifying stereotypes in the online perception of physical attractiveness. *International Conference on Social Informatics*, 10046 (419-437). Springer, Cham. [https://doi.org/10.1007/978-3-319-47880-7\\_26](https://doi.org/10.1007/978-3-319-47880-7_26)

Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74(2), 81-99. <https://doi.org/10.1037/h0029531>

Baltagi, B. H., & Baltagi, B. H. (2008). *Econometric analysis of panel data* (Vol. 4). Chichester: John Wiley & Sons. Hentet fra: <https://link.springer.com/content/pdf/10.1007/978-3-030-53953-5.pdf>

Baudouin, J. Y., & Tiberghien, G. (2004). Symmetry, averageness, and feature size in the facial attractiveness of women. *Acta psychologica*, 117(3), 313-332. Hentet fra: <https://www.sciencedirect.com/science/article/pii/S0001691804000794>

Becker, L. A. (2000). Effect size (ES). Hentet fra: <https://www.uv.es/~friasnav/EffectSizeBecker.pdf>

Boutis, K., Pecaric, M., Seeto, B., & Pusic, M. (2010). Using signal detection theory to model changes in serial learning of radiological image interpretation. *Adv in Health Sci Educ*, 15(5), 647-658. <https://doi.org/10.1007/s10459-010-9225-8>

Bracht, G. H., & Glass, G. V. (1968). The External Validity of Experiments. *American Educational Research Journal*, 5(4), 437-474. <https://doi.org/10.2307/1161993>

Brumberger, E. (2011). Visual literacy and the digital native: An examination of the millennial learner. *Journal of visual literacy*, 30(1), 19-47. <https://doi.org/10.1080/23796529.2011.11674683>

Bullingham, L., & Vasconcelos, A. C. (2013). 'The presentation of self in the online world': Goffman and the study of online identities. *Journal of Information Science*, 39(1), 101-112. <https://doi.org/10.1177/0165551512470051>



Campbell, D. T., Cook, T. D., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin. Hentet fra: <https://www.alnap.org/system/files/content/resource/files/main/147.pdf>

Cambridge Dictionary. (u. å.). Stereotype. Hentet fra: <https://dictionary.cambridge.org/dictionary/english/stereotype>

Caso, D., Schettino, G., Fabbricatore, R., & Conner, M. (2020). "Change my selfie": Relationships between self-objectification and selfie-behavior in young Italian women. *Journal of Applied Social Psychology*, 50(9), 538-549. <https://doi.org/10.1111/jasp.12693>

Chae, J. (2017). Virtual makeover: Selfie-taking and social media use increase selfie-editing frequency through social comparison. *Computers in Human Behavior*, 66, 370-376. <https://doi.org/10.1016/j.chb.2016.10.007>

Church, N. (2017, 25. April). *Facetune 2 - REVIEW & DEMO photo editing app* [Video file]. YouTube. <https://www.youtube.com/watch?v=QQXknTifT8c&t=628s>

Cohen, R., Newton-John, T., & Slater, A. (2018). 'Selfie'-objectification: The role of selfies in self-objectification and disordered eating in young women. *Computers in Human Behavior*, 79, 68-74. <https://doi.org/10.1016/j.chb.2017.10.027>

Cooper, J. (2012). Cognitive dissonance theory. I P. A. M. Van Lange, A. W. Kruglanski & E. T. Higgins (Red.), *Handbook of theories of social psychology: Volume 1* (s. 377-398). London: SAGE Publications. <https://dx.doi.org/10.4135/9781446249215>

Cooper, J. (2019). Cognitive Dissonance: Where We've Been and Where We're Going. *International Review of Social Psychology*, 32(1). <http://doi.org/10.5334/irsp.277>

Cruz, E. G., & Thornham, H. (2015). Selfies beyond self-representation: the (theoretical) f (r) ictions of a practice. *Journal of Aesthetics & Culture*, 7(1), 28073. <https://doi.org/10.3402/jac.v7.28073>

Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of personality and social psychology*, 59(1), 61. Hentet fra: <https://psycnet.apa.org/fulltext/1991-01192-001.html>

Dahlum, S. (2021, 9. mars). Validitet. *Store norske leksikon*. Hentet fra: <https://snl.no/validitet>

Dhir, A., Pallesen, S., Torsheim, T., & Andreassen, C. S. (2016). Do age and gender differences exist in selfie-related behaviours?. *Computers in Human Behavior*, 63, 549-555. <https://doi.org/10.1016/j.chb.2016.05.053>

Downs, A. C., & Harrison, S. K. (1985). Embarrassing age spots or just plain ugly? Physical attractiveness stereotyping as an instrument of sexism on American television commercials. *Sex roles*, 13(1), 9-19. <https://doi.org/10.1007/BF00287457>

Egelhofer, J. L. & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: a framework and research agenda. *Annals of the International Communication Association*, 43(2), 97-116. <https://doi.org/10.1080/23808985.2019.1602782>

Estellat, C., Torgerson, D. J., & Ravaud, P. (2009). How to perform a critical analysis of a randomised controlled trial. *Best Practice & Research Clinical Rheumatology*, 23(2), 291-303. <https://doi.org/10.1016/j.berh.2009.03.003>

Farid, H. & Bravo, M. J. (2010). Image forensic analyses that elude the human visual system. *Media forensics and security II*, 7541(754106), 1-11. <https://doi.org/10.1117/12.837788>

Festinger L. A. (1954). Theory of Social Comparison Processes. *Human Relations*, 7(2).1-16. Hentet fra: <https://www.humanscience.org/docs/Festinger%20%281954%29%20A%20Theory%20o%20Social%20Comparison%20Processes.pdf>

Festinger, L. (1962). *A theory of cognitive dissonance* (Reissue. ed.). Stanford, California: Stanford University Press.

Fiona Moore, F. (2012). Self-Presentation. I A. J. Mills, G. Durepos & E. Wiebe (Red.), *Encyclopedia of Case Study Research* (s. 850-852). Thousand Oaks: SAGE Publications, Inc. <https://dx.doi.org/10.4135/9781412957397>

Fox, J., & Vendemia, M. A. (2016). Selective self-presentation and social comparison through photographs on social networking sites. *Cyberpsychology, behavior, and social networking*, 19(10), 593-600. DOI: 10.1089/cyber.2016.0248

Glen, S. (2018, 26. mai). Marginal Effects: Definition. Hentet fra: <https://www.statisticshowto.com/marginal-effects/>

Goffman, E. (1959). *The Presentation of self in everyday life*. New York: Anchor Books.

Graph Pad. (u. å). Analysis checklist: Repeated measures two-way ANOVA (and mixed model). Hentet fra: [https://www.graphpad.com/guides/prism/latest/statistics/stat\\_checklist\\_2wayanova\\_rm.htm](https://www.graphpad.com/guides/prism/latest/statistics/stat_checklist_2wayanova_rm.htm)

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication*, 59(2), 367-386. Hentet fra: <https://academic.oup.com/joc/article-abstract/59/2/367/4098407>

Hand, D. J., & Taylor, C. C. (1987). *Multivariate analysis of variance and repeated measures: a practical approach for behavioural scientists* (Vol. 5). London: Chapman & Hall

Harvey, L. O., Hammond, K. R., Lusk, C. M., & Mross, E. F. (1992). The Application of Signal Detection Theory to Weather Forecasting Behavior. *Monthly Weather Review*, 120(5), 863-883. [https://journals.ametsoc.org/view/journals/mwre/120/5/1520-0493\\_1992\\_120\\_0863\\_taosdt\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/120/5/1520-0493_1992_120_0863_taosdt_2_0_co_2.xml)

Heeger, D. (2003). *Signal detection theory*. [Undervisningsdokument]. Hentet fra: [http://neurosci.info/courses/vision2/Coding/Heeger\\_2003.pdf](http://neurosci.info/courses/vision2/Coding/Heeger_2003.pdf)

- Hergueux, J., & Jacquemet, N. (2015). Social preferences in the online laboratory: A randomized experiment. *Experimental Economics*, 18(2), 251–283. <https://doi.org/10.1007/s10683-014-9400-5>
- Hewitt, A., & Forte, A. (2006). Crossing boundaries: Identity management and student/faculty relationships on the Facebook. *Poster presented at CSCW, Banff, Alberta*, 1-2. Hentet fra: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.8152>
- Hildebrandt, T., Shiovitz, R., Alfano, L., & Greif, R. (2008). Defining body deception and its role in peer based social comparison theories of body dissatisfaction. *Body Image*, 5(3), 299-306. <https://doi.org/10.1016/j.bodyim.2008.04.007>
- Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30(6), 377–386. <https://doi.org/10.1177/0270467610385893>
- Hvitmyhr, L. B. (2021, 3. juni). Nå må influencere merke retusjert reklame. VG. Hentet fra: <https://www.vg.no/nyheter/innenriks/i/dlyWbj/naa-maa-influencere-merke-retusjert-reklame>
- Jennings, R. (2019, 25. juli). Facetune and the internet's endless pursuit of physical perfection: A face-perfecting app only widens the gap between our digital and real selves. *Vox*. <https://www.vox.com/the-highlight/2019/7/16/20689832/instagram-photo-editing-app-facetune>
- Jill Walker Rettberg, J. W. (2017). Self Representation in Social Media. I J. Burgess, A. Marwick & T. Poell (Red.), *Handbook of Social Media* (s. 1-30). California: SAGE. Hentet fra: [https://bora.uib.no/bora-xmlui/bitstream/handle/1956/13073/Self-Representation in Social Media.pdf?sequence=1](https://bora.uib.no/bora-xmlui/bitstream/handle/1956/13073/Self-Representation%20in%20Social%20Media.pdf?sequence=1)
- Katz, J. & Crocker, E. T. (2015). Selfies and photo messaging as visual conversation: Reports from the United States, United Kingdom and China. *International Journal of Communication*, 9, 12. Hentet fra: [file:///C:/Users/47932/Downloads/3180-15506-1-PB%20\(2\).pdf](file:///C:/Users/47932/Downloads/3180-15506-1-PB%20(2).pdf)
- Keating, C. F. (1985). Gender and the Physiognomy of Dominance and Attractiveness. *Social Psychology Quarterly*, 48(1), 61–70. <https://doi.org/10.2307/3033782>
- Lee, S. Y. (2014). How do people compare themselves with others on social network sites?: The case of Facebook. *Computers in human behavior*, 32, 253-260. <https://doi.org/10.1016/j.chb.2013.12.009>
- Lee, M., & Lee, H. H. (2021). Social media photo activity, internalization, appearance comparison, and body satisfaction: The moderating role of photo-editing behavior. *Computers in Human Behavior*, 114, 106579. <https://doi.org/10.1016/j.chb.2020.106579>
- Lonergan, A. R., Bussey, K., Mond, J., Brown, O., Griffiths, S., Murray, S. B., & Mitchison, D. (2019). Me, my selfie, and I: The relationship between editing and posting selfies and body dissatisfaction in men and women. *Body image*, 28, 39-43. <https://doi.org/10.1016/j.bodyim.2018.12.001>
- López-Guimerà, G., Levine, M. P., Sánchez-Carracedo, D., & Fauquet, J. (2010). Influence of mass media on body image and eating disordered attitudes and behaviors in

females: A review of effects and processes. *Media Psychology*, 13(4), 387-416.  
<https://doi.org/10.1080/15213269.2010.525737>

Lup, K., Trub, L., & Rosenthal, L. (2015). Instagram# instasad?: exploring associations among instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychology, Behavior, and Social Networking*, 18(5), 247-252.  
<https://doi.org/10.1089/cyber.2014.0560>

Macmillan, N. A. (2001). Signal Detection Theory. *International Encyclopedia of the Social & Behavioral Sciences*, 14075-14078. <https://doi.org/10.1016/B0-08-043076-7/00677-X>

Merunková, L., & Šlerka, J. (2019). Goffman's Theory as a Framework for Analysis of Self Presentation on Online Social Networks. *Masaryk University Journal of Law and Technology*, 13(2), 243-276. Hentet fra: <https://www.ceeol.com/search/article-detail?id=798067>

Messaris, P. (2012). Visual "literacy" in the digital age. *Review of Communication*, 12(2), 101-117. <https://doi.org/10.1080/15358593.2011.653508>

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, 16(1), 148-156. <https://doi.org/10.1037/h0029862>

Murray, C. D. (2015). Notes to self: the visual culture of selfies in the age of social media. *Consumption Markets & Culture*, 18(6), 490-516,  
<https://doi.org/10.1080/10253866.2015.1052967>

Mutz, D. C. (2011). *Population-Based Survey Experiments*. Princeton University Press. Hentet fra: <https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=370823&site=ehost-live>

Nardi, M. P. (2014). *Doing survey Research: A guide to quantitative methods* (3. Utg.). Colorado: Paradigm Publisher.

Neave, N., & Shields, K. (2008). The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces. *Personality and Individual Differences*, 45(5), 373-377. <https://doi.org/10.1016/j.paid.2008.05.007>

Nightingale, S. J., Wade, K. A. & Watson, D. G. (2017). Can people identify original and manipulated photos of real-world scenes? *Cognitive research: principles and implications*, 2(30), 1-21. <https://doi.org/10.1186/s41235-017-0067-2>

Nilstun, C. (2021, 7. november). Manipulere. *Store norske leksikon*. Hentet fra: <https://snl.no/manipulere>

O'Dell, M. (2021, 4. desember). Here's how I used to edit my selfies to look pretty — and why I won't do it anymore. *CBC*. <https://www.cbc.ca/news/canada/newfoundland-labrador/unfiltered-social-media-and-mental-health-1.6267941>

- Orland-Barak, L. & Maskit, D. (2017). *Methodologies of Mediation in Professional Learning*. Springer, Cham. <https://doi.org/10.1007/978-3-319-49906-2>
- Oxford English Dictionary. (u. å.). selfie, n. Hentet fra: <https://www.oed.com/view/Entry/390063?redirectedFrom=selfie>
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, 394(6), 884-887. Hentet fra: <https://www.nature.com/articles/29772>
- Pihl, R. (2021, 7. november). Influenser. *Store norske leksikon*. Hentet fra: <https://snl.no/influenser>
- Qiu, L., Lu, J., Yang, S., Qu, W., & Zhu, T. (2015). What does your selfie say about you?. *Computers in Human Behavior*, 52, 443-449. <https://doi.org/10.1016/j.chb.2015.06.032>
- Ringdal, K. (2013). Enhet og mangfold: Samfunnsvitenskapelig forskning og kvantitativ metode (3. utg.). Bergen: Fagbokforlaget.
- Rose, J., Mackey-Kallis, S., Shyles, L., Barry, K., Biagini, D., Hart, C., & Jack, L. (2012). Face it: The impact of gender on social media images. *Communication Quarterly*, 60(5), 588-607. <https://doi.org/10.1080/01463373.2012.725005>
- Sartwell, C. (2022). Beauty. *The Stanford Encyclopedia of Philosophy*. Hentet fra: <https://plato.stanford.edu/entries/beauty/>
- Schetinger, V., Oliveira, M. M., da Silva, R., & Carvalho, T. J. (2017). Humans are easily fooled by digital images. *Computers & Graphics*, 68, 142-151. <https://doi.org/10.48550/arXiv.1509.05301>
- Shen, C., Kasra, M., Pan, W., Bassett, G. A., Malloch, Y., & O'Brien, J. F. (2019). Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online. *New media & society*, 21(2), 438-463. <https://journals.sagepub.com/doi/10.1177/1461444818799526>
- Silberstein, L. R., Striegel-Moore, R. H., Timko, C., & Rodin, J. (1988). Behavioral and psychological implications of body dissatisfaction: Do men and women differ?. *Sex roles*, 19(3), 219-232. Hentet fra: <https://link.springer.com/article/10.1007/BF00290156>
- Smith, L. R., & Sanderson, J. (2015). I'm going to Instagram it! An analysis of athlete self-presentation on Instagram. *Journal of Broadcasting & Electronic Media*, 59(2), 342-358. <https://doi.org/10.1080/08838151.2015.1029125>
- Solon, O. (2019, 9. Mars). FaceTune is conquering Instagram – but does it take airbrushing too far? *The Guardian*. <https://www.theguardian.com/media/2018/mar/09/facetune-photoshopping-app-instagram-body-image-debate>
- Statology. (2021, 5. Februar). What is Central Tendency Bias? Hentet fra: <https://www.statology.org/central-tendency-bias/>

Sthåle, L., & Wold, S. (1989). Analysis of variance (ANOVA). *Chemometrics and intelligent laboratory systems*, 6(4), 259-272. [https://doi.org/10.1016/0169-7439\(89\)80095-4](https://doi.org/10.1016/0169-7439(89)80095-4)

Sykes, A. O. (1993). An introduction to regression analysis. *Coase-Sandor Institute for Law & Economics Working Paper*, 20, 0-33. Hentet fra: [https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1050&context=law\\_and\\_economics](https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1050&context=law_and_economics)

Yan, Y., & Bissell, K. (2014). The globalization of beauty: How is ideal beauty influenced by globally published fashion and beauty magazines?. *Journal of Intercultural Communication Research*, 43(3), 194-214. <https://doi.org/10.1080/17475759.2014.917432>

Wen, T. (2020, 10. juni). The hidden signs that can reveal a fake photo. *BBC Future*. Hentet fra: <https://www.bbc.com/future/article/20170629-the-hidden-signs-that-can-reveal-if-a-photo-is-fake>

Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, 1, 27-31. [https://doi.org/10.1016/0022-1031\(66\)90062-X](https://doi.org/10.1016/0022-1031(66)90062-X)

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90(2), 245-271. <https://doi.org/10.1037/0033-2909.90.2.245>

