



**FFI** Forsvarets  
forskningsinstitutt

23/00824

**FFI-RAPPORT**

# Påvirkning og robusthet på sosiale medier

– resultater og erfaringer fra piloteksperiment

Anne Lise Bjørnstad



# **Påvirkning og robusthet på sosiale medier – resultater og erfaringer fra piloteksperiment**

Anne Lise Bjørnstad

---

**Emneord**

Påvirkningsoperasjoner  
Sosiale medier  
Psykologiske sårbarheter  
Eksperimenter  
Robusthet  
Priming

**FFI-rapport**

23/00824

**Prosjektnummer**

1582

**Elektronisk ISBN**

978-82-464-3490-2

**Engelsk tittel**

Social media influence and robustness: results and learning points from a pilot experiment

**Godkjennerne**

Stig Rune Sellevåg, *forskningsleder*  
Janet Martha Blatny, *forskningsdirektør*

*Dokumentet er elektronisk godkjent og har derfor ikke håndskreven signatur.*

**Opphavsrett**

© Forsvarets forskningsinstitutt (FFI). Publikasjonen kan siteres fritt med kildehenvisning.

---

---

## Sammenheng

Uønsket og fiendtlig påvirkning gjennom digitale medier representerer en sentral trussel mot norsk sikkerhet ifølge EOS (etterretnings-, overvåkings- og sikkerhets)-tjenestene. Mer kunnskap om de psykologiske faktorene som inngår i bruken av sosiale medier vil kunne bidra i arbeidet mot slik påvirkning. Dette gjelder i forhold til effektene av påvirkning på sosiale medier, og hvordan man kan benytte virkemidler identifisert i psykologisk forskning til å øke robustheten mot slike påvirkningsforsøk. Rapporten tar både for seg sentral forskning på temaene og beskriver et utforskende eksperiment (piloteksperiment) som bringer denne forskningen videre.

Rapporten presenterer og diskuterer forskning på holdninger og oppfatninger, og effektene påvirkning kan ha på adferd. Videre tar den for seg sentral forskning på generelle psykologiske sårbarheter for påvirkning, individuelle forskjeller i sårbarhet og tiltak og teknikker som kan styrke robusthet mot forsøk på påvirkning.

Piloteksperimentet hadde som mål å undersøke hvordan innhold på sosiale medier påvirker menneskers oppfatninger av saker. Eksperimentet skulle også utforske tiltak for å øke robustheten i befolkningen mot uønsket påvirkning i et kontrollert miljø.

Deltagerne ble tilfeldig inndelt i fire forskjellige grupper som fikk ulike betingelser. De fikk enten forskningsbasert muntlig, skriftlig, spill-type eller irrelevant (kontrollgruppe) informasjon relatert til hvordan man kan bli påvirket på sosiale medier. Alle fikk deretter lese om fiktive saker på et simulert sosialt medium (Twitter-klonen *Mastodon*).

Det ble funnet målbare endringer i oppfatning av saker etter at deltagerne hadde lest meldingene på Mastodon. Til tross for at utvalget var lite ( $N = 20$ ) var halvparten av disse endringene i oppfatning signifikante. Endringene tenderte til å være størst i kontrollgruppen og minst i muntlig- og spill-gruppene. Den samlede tendensen syntes altså å være lovende med hensyn til typen betingelser testet ut her. Det vil si at forskningsbasert muntlig, skriftlig og spill-type informasjon om påvirkning så ut til å kunne gi en gunstig robusthetseffekt mot uønsket påvirkning gjennom sosiale medier. Slike tiltak vil derfor være formålstjenlig å undersøke nærmere i fremtidige studier. Som med endringene i oppfatning, syntes deltagerne i muntlig og spill-gruppene også å vise tendens til noe mer robusthet mot å tro på usanne tweets. Her var det kun tendenser og ingen signifikante funn. Kvalitative data fra gruppesamtaler med deltagerne understøttet de kvantitative tendensene og funnene.

Begrensninger i studien inkluderte et lite utvalg ( $N = 20$ ) og at meldingene benyttet i piloteksperimentet kunne ha vært mer sofistikerte. Erfaringene med piloteksperimentet resulterte også i anbefalinger for fremtidige eksperimenter. Litteraturen på forskningsområdet indikerte at man må påregne oppfriskning og gjentakelser av budskap og informasjon til alle deler av befolkningen man ønsker å gjøre mer motstandsdyktig mot påvirkning. I dette ligger også en tilrettelegging av budskap, slik at det blir tilgjengelig for ulike grupper. Her er det gjort lite og det vil derfor være behov for mer forskning for å forstå hvordan dette best kan gjøres.

Samlet ble resultatene og erfaringene tolket som lovende med hensyn til å gjennomføre fremtidige eksperimenter i større målestokk med utvalg fra flere populasjoner, inkludert Forsvaret og andre aktører i totalforsvaret.

---

---

## Summary

Unwanted and hostile influence through digital media represents a threat to Norwegian security according to Norwegian Intelligence, Surveillance and Security (EOS) Services. Enhancing our understanding of the psychological factors associated with the use of social media contributes to the work against unwanted influence. This includes both effects of social media influence and techniques to enhance resilience against influence attempts on social media. This report presents key research on the above-mentioned themes and provides insights from an exploratory (pilot) experiment that serves to advance this research.

This report explores research on attitudes and perceptions, and the effects of influence on behavior. It also presents central research on general psychological vulnerabilities to influence, individual differences in vulnerability, as well as measures and techniques that can strengthen the societal resilience against influence attempts.

The goal of this pilot experiment was twofold; first, to explore how social media content influence the perception of issues, and second, to explore measures to increase resilience against unwanted influence within a controlled environment. The participants were randomly divided into four different groups that were subjected to different conditions. They were either given research-based verbal, written, game-type, or irrelevant (control group) information related to how one can be influenced on social media. The participants were subsequently presented with fictitious posts on a simulated social platform (Twitter clone *Mastodon*).

There were found measurable changes in views on issues after the participants had read the messages on Mastodon. Despite the small sample size ( $N = 20$ ), half of the changes in views were significant. The changes in views tended to be the largest in the control group and the least in the verbal and game groups. The overall tendency thus appeared promising with regard to the types of conditions tested here. The findings indicate that research-based verbal, written and game-type information about influence may increase societal resilience against social media influence, suggesting such measures may be further investigated in future studies. The participants in the verbal and game-priming groups seemed also to show a tendency towards somewhat more robustness with regard to believing false tweets. These results were only tendencies (non-significant). Qualitative data from group discussions with the participants supported the trends and findings in the quantitative data.

Limitations included a small sample size and that the social media posts used in the pilot experiment may have been more sophisticated. The analyses also yielded recommendations for future experiments. The existing literature in this area further indicate a time decay of the effects of influence counter measures suggesting a need to refresh and repeat. There is also a need to adapt messages to different groups, which is another area in need of more research.

Overall, the results and experiences were promising for conducting future experiments on a larger scale with samples from other populations, including the Norwegian Armed Forces and Total Defense contributors.

---

---

# Innhold

<b>Sammendrag</b>	<b>3</b>
<b>Summary</b>	<b>4</b>
<b>Forord</b>	<b>8</b>
<b>1 Innledning</b>	<b>9</b>
1.1 Bakgrunn	9
1.2 Sikkerhetsutfordringer, menneskelig sårbarhet, psykologiske faktorer og sosiale medier	10
1.3 Situasjonsutløsende faktorer	11
1.4 Avgrensning	11
1.5 Sum – mål, problemstilling og rapportinnhold	12
<b>2 Teori og tidligere forskning</b>	<b>12</b>
2.1 Holdninger, oppfatninger, endringer og adferd	13
2.2 Psykologiske sårbarheter for påvirkning	13
2.2.1 Persuasjonsprinsippene	15
2.2.2 Feilkilder i menneskelig oppfattelse	15
2.2.3 Oppfattelser av sant og usant	16
2.3 Individuelle forskjeller i sårbarhet for des- og misinformasjon	17
2.3.1 Need for cognition (NFC) og kritisk tenkning (“Critical thinking”)	17
2.4 Kjennetegn på påvirkningsforsøk og teknikker for å øke robusthet	18
2.4.1 Kjennetegn på forsøk på påvirkning	18
2.4.2 «Priming», «framing» og «prebunking»	19
2.4.3 Medietreningsstudier («media-literacy»-studier)	19
2.4.4 «Accuracy-nudge» intervensjon	20
2.4.5 Vaksine («inoculation») – strategi	20
2.5 Forskningsspørsmål	22
<b>3 Metode</b>	<b>23</b>
3.1 Forskningsdesign og operasjonalisering	23
3.2 Verktøy	23
3.2.1 Simuleringsverktøyet «Mastodon»	23
3.2.2 Spørreskjema	24

3.3	Eksperimentelt oppsett	24
3.3.1	Primerne (de eksperimentelle betingelsene)	25
3.3.2	Mastodon-meldingene	25
3.4	Gjennomføring	25
3.5	Utvalg	26
3.6	Etiske betraktninger	27
3.7	Målene på oppfatning av saker med reliabilitetsanalyser	28
3.7.1	Tillit til vaksine/myndigheter	28
3.7.2	Oppfatning av konflikt by-land	28
3.7.3	Oppfatning av Forsvaret (negativ)	28
3.7.4	Oppfatning av likestilling (negativ)	29
3.7.5	Oppfatning av at det er mye rasisme i Norge	29
3.7.6	Oppfatning av president Biden (positiv)	29
3.8	Mål på evne til å gjenkjenne sanne, og robusthet mot å tro på usanne og gråsome utsagn/tweets	29
3.8.1	Mål på svarsikkerhet (konfidens)	29
3.9	Kontrollvariabler	30
3.10	Need for Cognition (NFC)	30
3.11	Kvalitative data	32
3.12	Statistiske analyser	32
3.12.1	Innledende analyser	32
3.12.2	Hovedanalyser	33
<b>4</b>	<b>Resultater</b>	<b>34</b>
4.1	Kontrollvariablene	34
4.1.1	Korrelasjonsanalyser: oppfatningsendringer og kontrollvariablene	35
4.1.2	Korrelasjonsanalyser: evne til å gjenkjenne sanne, gråsome og usanne nyheter og kontrollvariablene	36
4.2	Endring av oppfatning	38
4.2.1	Holdning 1: Tillit til vaksine/myndigheter	38
4.2.2	Holdning 2: Oppfatning av konflikt by-land	39
4.2.3	Holdning 3: Oppfatning av Forsvaret (negativ)	41
4.2.4	Holdning 4: Oppfatning av likestilling (negativ)	42
4.2.5	Holdning 5: Oppfatning av at det er mye rasisme i Norge	43
4.2.6	Holdning 6: Oppfatning av president Biden (positiv)	44
4.3	Analysen av gruppeforskjeller i evne til å gjenkjenne sanne tweets, og robusthet mot å tro på usanne og gråsome tweets	46
4.3.1	ANOVA (enveis variansanalyse)	46
4.3.2	Analyser av mål på konfidens (svarsikkerhet)	50
4.4	Kvalitative dataanalyser: fra observasjoner, gruppe- og fellesdiskusjonene	51



---

---

<b>5</b>	<b>Diskusjon</b>	<b>52</b>
5.1	NFC og de andre kontrollvariablene	52
5.2	Endring i oppfatninger av saker og videre forskning	53
5.3	Robusthet mot endring i oppfatninger av saker og videre forskning	54
5.4	Forskjeller i evne til å skille sanne, usanne og gråsone tweets og videre forskning	55
5.5	Konfidens (svarsikkerhet) og videre forskning	55
5.6	Resultatene fra oppfatninger av saker og sant/usant-analysene samlet og videre forskning	56
5.7	Kvalitative analyser og videre forskning	56
5.8	Begrensninger og videre forskning	58
5.8.1	Utvalgsstørrelsen	58
5.8.2	Effekter over tid	58
5.8.3	Tweetsene	58
5.8.4	Primerne (de eksperimentelle betingelsene)	59
5.8.5	Somulatoren (teknologien for å simulere sosiale medier)	59
5.8.6	Etiske utfordringer	60
<b>6</b>	<b>Konklusjon, veien videre og noen anbefalinger</b>	<b>60</b>
	<b>Referanser</b>	<b>63</b>
<b>A</b>	<b>Vedlegg</b>	<b>71</b>
A.1	URL spørreskjema 1 & 2, Forhåndsinformasjon 1, 2, 3 og Twitterklone/Mastodon	71
A.2	Eksperiment-manual for FFI (C-SPI) 16.02.2022	72
A.3	Spørreskjema	75
A.3.1	Pre-eksperiment spørreskjema	75
A.3.2	Post-eksperiment spørreskjema	79

---

---

## Forord

Dette forskningsarbeidet er gjort som en del av FFI-prosjekt *Cyber-Social Propaganda and Influence – Påvirkning i Informasjonsdomenet (C-SPI)*.

Jeg vil først få takke Arild Bergh (seniorforsker ved FFI) for arbeidet med Somulatoren, Mastodon og ikke minst for hans tekniske støtte og innsats ved gjennomføringen av eksperimentet beskrevet i denne rapporten.

Jeg vil deretter få takke elevene og deres lærere på Medie-linjen ved Lillestrøm Videregående for svært godt og hyggelig samarbeid om oppretting og utsendelse av fiktive tweets til eksperimentet. En spesiell takk går til: Maria Tenstad Damm, Christian Alsaker, Kamilla Bakstad, Jonatan Meke og deres lærer Gunnulv Hellesylt for deres deltagelse på forberedelser, selve eksperimentet og generelt gode samarbeid. I tillegg utviste elevene en entusiasme som vi alle hadde stor glede av.

Jeg vil også få takke deltagerne som stilte opp på dette eksperimentet og bidro til at vi har data å analysere. En stor takk til dere!

Og sist, men ikke minst, en takk til Nina Hellum, Lea Kristina Petronella Bjørgul, Julie Celine Berghaust, Betina Slagnes, Eskil Grendahl Sivertsen, Rune Lausund (forskere ved FFI) og forskningsleder Stig Rune Sellevåg – som alle har bidratt til gjennomføringen av eksperimentet.

Kjeller, 27.januar 2023  
Anne Lise Bjørnstad

---

---

# 1 Innledning

## 1.1 Bakgrunn

Digitale medier er sentrale informasjonskilder for folk i hele verden, og bruken øker for hvert år (f.eks., Grizăne, Isupova & Vorteil, 2022). 58 prosent av verdens befolkning bruker sosiale medier (pr. januar 2022), og fra 2021 til 2022 var det en 10 prosent økning (Kemp, 2022). Også journalister tilknyttet sentrale nyhetsorganisasjoner i Europa er funnet å ofte bruke innhold fra sosiale medier som sin primære nyhetskilde (Brandtzæg, Luders, Spangenberg, Rath-Wiggins & Følstad, 2016). Medietilsynet rapporterte i 2022 at så å si alle mellom 13 og 18 år i Norge bruker sosiale medier (Medietilsynet, 2022), og at det er på sosiale medier de unge oftest har kommet over nyheter de mistenkte ikke var sanne. Av hele befolkningen over 16 år, rapporterte 68 prosent at de har kommet over informasjon de var i tvil om var sann på nett (Medietilsynet, 2021). To befolkningsundersøkelser om kritisk medieforståelse gjennomført av Medietilsynet i Norge (Medietilsynet, 2019, 2021), indikerer at mange mennesker i Norge også har problemer med å gjenkjenne en falsk nyhet. I 2019 fant Medietilsynet at 44 prosent av de under 30 ikke avslørte en falsk nyhet, mens tallet var 49 prosent for de over 60. I 2021-rapporten indikerte at tallet på de som ikke avslørte en falsk nyhet var 23 prosent for befolkningen generelt, 30 prosent for de mellom 60 og 79 år, og 49 prosent for de over 80 år. Selv om tallene er litt forskjellige i de to undersøkelsene, avhengig av målemetode og utvalg, viser begge at det er behov og potensiale for å øke evnen til å gjenkjenne falske nyheter og forsøk på påvirkning i den norske befolkningen.

Oppsummert kan man si det slik at mennesker generelt har problemer med å skille sant fra usant på nett og at problemet med usant synes å være størst på sosiale medier (Medietilsynet, 2021, 2022; Sentio Research Group, 2022; Sivertsen et al., 2022). Samtidig er det en sterkt økende bruk av sosiale medier. Stortingsmelding 10 (Forsvarsdepartementet, 2021–2022) beskriver desinformasjon og fiendtlig påvirkning gjennom digitale medier som en sentral og økende trussel mot norsk sikkerhet. Forsvarskommisjonen av 2021 (NOU 14, 2023) underbygger dette og indikerer at både Forsvaret og samfunnet har behov for å utvikle sin forsvarsevne og robusthet mot påvirkning i informasjonsmiljøet («Informasjonsmiljøet omfatter totaliteten av enkeltpersoner, organisasjoner og systemer som samler inn, behandler, sprer eller handler basert på informasjon»; NOU 14, 2023, s. 208). Det er derfor behov for økt kunnskap om hvordan falske nyheter, desinformasjon (intendert feilinformasjon) og misinformasjon (uintendert feilinformasjon) på sosiale medier påvirker oss og hvilke tiltak som kan øke robustheten i befolkningen mot slik uønsket og/eller fiendtlig påvirkning.

Psykologisk forskning indikerer at alle i en eller annen grad er påvirkelige – det er rett og slett en del karakteristiske trekk ved det å være menneske som gjør oss sårbare for påvirkning (for en oversikt, se f.eks., Bjørnstad, 2019). FFI-rapport 19/01224 (Bjørnstad, 2019) presenterer forskning fra psykologisk fagfelt som kan belyse noen av disse sårbarhetene. Som rapporten reflekterer, har mye av psykologisk forskning på området tradisjonelt hatt hovedfokus på hvordan man kan oppnå påvirkning, det har vært mindre forskning direkte på hvordan man kan

---

øke robustheten mot påvirkning. Dette har endret seg siden 2019, og det er nå fremvekst av mye ny forskningsaktivitet med fokus på hvordan man kan gjøre mennesker mer motstandsdyktige (se f.eks., Pennycook, McPhetres, Zhang, Lu & Rand, 2020; Roozenbeek, van der Linden, Goldberg, Rathje & Lewandowsky, 2022). Denne forskningen er en god og viktig start, men indikerer også at det er fortsatt er mange uløste problemstillinger.

Som påpekt, representerer uønsket påvirkning gjennom digitale medier en sentral trussel mot norsk sikkerhet. Det forstås derfor som formålstjenlig å øke kunnskapen om de psykologiske faktorene involvert i bruken av sosiale medier i forsvarsøyemed. Dette gjelder både med hensyn til effektene av påvirkning på sosiale medier og hvordan man kan benytte virkemidler identifisert i psykologisk forskning til å øke robustheten mot påvirkningsforsøk på sosiale medier. Dette er i tråd med Forsvarskommisjonen (NOU 14, 2023) som påpeker betydningen av psykologisk motstandskraft i befolkningen i møte med sammensatte trusler som desinformasjon.

Målsettingen med denne studien er både å ta for seg sentral forskning på temaene og rapportere fra et piloteksperiment. Sistnevnte bidrar til å teste hvordan innhold på sosiale medier påvirker oppfatninger av saker og å utforske tiltak for å øke robustheten i befolkningen mot uønsket og/eller fiendtlig påvirkning i et kontrollert miljø. Som indikert ved benevnelsen piloteksperiment planlegges det videre eksperiment i større målestokk, med justeringer basert på erfaringene beskrevet i denne rapporten.

## **1.2 Sikkerhetsutfordringer, menneskelig sårbarhet, psykologiske faktorer og sosiale medier**

Omfanget og økningen i bruk av sosiale medier gjør dette til en sannsynlig vei inn for en fiendtlig aktør til å forsøke å påvirke hele eller deler av en befolkning til å endre oppfatning, holdning og/eller adferd. Som indikert av Etterretningstjenesten, Nasjonal sikkerhetsmyndighet (NSM) og Politiets Sikkerhetstjeneste (PST) (Etterretningstjenesten, 2021, 2022; Nasjonal sikkerhetsmyndighet, 2021; Politiets Sikkerhetstjeneste; 2021), har for eksempel Russland demonstrert både i Norge (f.eks., hacking av norske stortingsrepresentanters e-poster høsten 2020) og andre land (f.eks., propaganda og desinformasjon ifm. invasjonen/krigføringen i Ukraina og innblanding i presidentvalget i 2016 i USA; Mueller, 2019) at de bedriver cyberangrep og påvirkningsoperasjoner. Kommersielle selskaper kan også frivillig eller ufrivillig bli brukt av nasjonalstater i påvirkningsoperasjoner, for eksempel gjennom å skaffe til veie og/eller selge informasjon, eller distribuere misinformasjon eller desinformasjon. FFI-rapport 21/02746 (Sivertsen et al., 2022) indikerte at utenlandske, ikke-statlige aktører er involvert i å spre desinformasjon gjennom sosiale medier i Norge. Mange forskjellige aktører kan altså ha interesse av å påvirke hele eller deler av den norske befolkningen.

Kognitive, emosjonelle, sosiale og kulturelle faktorer spiller inn og gjør mennesker sårbare for forsøk på påvirkning (for en oversikt se f.eks., Bjørnstad, 2019; Fiske & Taylor, 2017). I tillegg til slike generelle menneskelige sårbarheter, er det individuelle forskjeller som påvirker graden av sårbarhet. (f.eks., Baron, Isler, Yilmaz, 2022; Pennycook & Rand, 2019; Roozenbeek et al., 2022b).

---

---

Videre kan sosiale medier gjennom deling av i økende grad ekstremt innhold med likesinnede, fungere som et verktøy for polarisering og dannelse av ekstremistiske holdninger (f.eks., Van Swol, Lee & Hutchins, 2022). Selve gruppekonteksten leder mennesker til å søke belønning fra de andre gruppe medlemmene, senker oppfattelsen av fare og fører til disinhibisjon blant gruppe medlemmene (dvs. at adferd som ikke er sosialt akseptabel ikke blir undertrykt innen gruppene; se f.eks., Hinsz & Jackson, 2022). Slike gruppeprosesser på sosiale medier kan forklare hendelser som stormingen av den amerikanske kongressen 6. januar 2021 (Hinsz & Jackson, 2022; Winget & Park, 2022).

I tillegg kommer algoritmene på sosiale medier som forsterker og sprer det som skaper størst engasjement slik at konstruerte nyheter som spiller på emosjoner, ytterliggående standpunkter, kontroversielle temaer og lignende får mest spredning (f.eks., Bergh, 2020; Roozenbeek & van der Linden, 2019; Sivertsen, Hellum, Bergh & Bjørnstad, 2021). Sosiale medier kan på denne måten oppfattes å være «direkte faktafiendtlige» (Hofoss, 2022). Videre tenderer fiktive nyheter å benytte seg av enklere språk (Carrasco-Farré, 2022), og noe som er enklere å lese tenderer mer til å bli oppfattet som sant enn det som er vanskeligere å lese (Reber & Schwartz, 1999).

### **1.3 Situasjon utløsende faktorer**

Desinformasjon, misinformasjon og konspirasjonsteorier på sosiale (og andre) medier i kombinasjon med dramatiske verdensomspennende begivenheter kan gi verden store utfordringer i årene fremover. Covid 19-pandemien og Russlands invasjon og krigføring i Ukraina er eksempler på hendelser som gir mange land i verden store utfordringer. En situasjon med ressursknapphet og uro destabiliserer og kan åpne opp for mange typer aktører som vil kunne ønske å utnytte seg av situasjonen, både lokalt, nasjonalt og internasjonalt. Maktbalanser kan flyttes og polariseringen mellom ulike grupper øke og mange kan få behov for å søke noen å legge ansvaret på. Sosiale medier muliggjør og forenkler samlinger av mennesker med sammenfallende synspunkter. Digitale plattformer brukes blant annet til å kontrollere narrativene, spre propaganda, konspirasjonsteorier og løgner satt i system, og kan mane til både destruktive og voldelige handlinger (f.eks., stormingen av Kongressen i USA i 2021; Hinsz & Jackson, 2022; Winget & Park, 2022).

### **1.4 Avgrensning**

Sårbarheter for påvirkning omhandler mange nivå, fra det internasjonale og politiske, til det samfunnsmessige/kulturelle, gruppe, organisasjon, til psykologiske sårbarheter på individnivå. Man kan også snakke om sårbarheter i Forsvaret, i økonomien, og så videre. Denne rapporten har fokus på psykologiske sårbarheter på individnivå, og hva man kan gjøre for å motvirke disse – nærmere bestemt uttesting av metoder for å øke den individuelle robustheten. Som indikert over er dette begrunnet i at psykologisk forskning kan forklare hvordan og hvorfor mennesker er sårbare for påvirkning, og gir indikasjoner på virkemidler som kan brukes til å øke robustheten. Samtidig ble det påpekt behov for forskning for å forstå dette bedre – spesielt i konteksten av en digitalisert sosial virkelighet. Behovet for forståelse av psykologiske faktorer og økning av

---

---

psykologisk motstandsdyktighet ble også løftet frem i Forsvarskommisjonens rapport (NOU 14, 2023).

Teknologimediert påvirkning kan videre foregå gjennom mange kanaler. Denne rapporten begrenser seg til å omhandle sosiale medier. Bakgrunnen for dette er, som beskrevet innledningsvis, den svært økende og omfattende bruken av sosiale medier (Medietilsynet, 2022; Kemp, 2022; Sivertsen et al., 2021), funnet av en stor mengde mis- og desinformasjon og falske nyheter på sosiale medier (Medietilsynet, 2021; Sentio Research Group, 2022; Sivertsen et al., 2022), erfaringen med at sosiale medier brukes av både Kina og Russland i påvirkningsoperasjoner i vestlige demokratier (Etterretningstjenesten, 2021, 2022; Nasjonal sikkerhetsmyndighet, 2021; Politiets Sikkerhetstjeneste; 2021; Sivertsen et al., 2021) og at algoritmene på sosiale medier fremmer spredningen av des- og misinformasjon (Bergh, 2020; Hofoss, 2022). Alt dette gjør sosiale medier til en egnet kanal for påvirkningskampanjer, mens forskningen ikke helt har økt i samme takt som bruken og misbruken av sosiale medier.

## 1.5 Sum – mål, problemstilling og rapportinnhold

Et viktig mål med arbeidet for å minimere effektene av uønsket påvirkning fra fremmede nasjoner er å forstå våre sårbarheter og kunne igangsette tiltak for å øke motstandskraften mot påvirkning. I tråd med dette er målsettingen med forskningen beskrevet i denne rapporten både å utvide kunnskapen om menneskelige sårbarheter for påvirkning, samt å starte en uttesting av teknikker for å øke robustheten hos det enkelte menneske mot påvirkningsforsøk, med spesielt fokus på sosiale medier.

Rapporten starter med en presentasjon av teori og tidligere forskning på temaet og tar deretter for seg utviklingen og gjennomføringen av et piloteksperiment. Dette er det første av det som planlegges å kunne bli en serie eksperimenter og forskning på psykologiske faktorer ved forsøk på påvirkning gjennom sosiale medier. Planen er på sikt å utvide både utvalg og teknikker og utvikle simuleringsplattformen (Somulator; Bergh, 2023) for forskningsbruk.

## 2 Teori og tidligere forskning

Med målsetting om å bedre forstå sårbarhetene og hvordan man kan gjøre mennesker mer robuste mot uønsket påvirkning presenteres i det følgende relevant forskning fra psykologisk fagfelt. Presentasjonen er ikke ment å være uttømmende, men forklarer grunnlaget for de hypotesene forskningen som her blir rapportert søker å svare på. Først presenteres relevant forskning som tar for seg *hva* som påvirkes – det vil si, holdninger og effektene som endrede holdninger og oppfatninger kan ha på adferd. Deretter følger en presentasjon av forskning på psykologiske sårbarheter for påvirkning, fulgt av individuelle forskjeller i sårbarhet for des- og

---

---

misinformasjon og til slutt forskning på tiltak og teknikker som kan motvirke og bygge motstandskraft mot forsøk på påvirkning.

## **2.1 Holdninger, oppfatninger, endringer og adferd**

Holdninger handler om hvordan vi evaluerer en sak eller andre mennesker, og kan defineres som “the psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor” (Eagly & Chaiken, 1993; p. 1). Holdninger påvirker hvordan mennesker selekterer, oppfatter og evaluerer informasjon (f.eks., Stahlberg & Frey, 1996) – og holdninger og oppfatninger påvirker adferd (f.eks., Ajzen, 1991; Fiske & Taylor, 2017). Samtidig ser man at adferd er med på å forme holdninger gjennom for eksempel kognitiv dissonans (Festinger, 1957). Det betyr at adferd som ikke er i tråd med ens holdninger gjør at man justerer holdningen på grunn av ubehaget med dissonans mellom holdning og adferd.

Påvirkningsoperasjoner og forsøk på påvirkning gjennom bruk av desinformasjon og falske nyheter kan gi endrede oppfatninger av saker, personer, grupper, institusjoner, eller land. Ved for eksempel over tid å sverte en myndighetsperson gjennom en desinformasjonskampanje, kan dennes myndighet og troverdighet brytes ned slik at myndighetspersonen kan miste gjennomslagskraft. Det samme kan skje på nasjonalt nivå, slik at for eksempel våre partnere i NATO kan bli mindre motivert til å hjelpe Norge i en konfliktsituasjon hvis Norge blir oppfattet som for eksempel lite troverdig og/eller uetisk. Russlands portrettering av Ukraina som naziregime og lignende var med på å berede grunnen for invasjonen ved å prøve å bygge opp forståelse og aksept for en invasjon (f.eks., Kumankov, 2023; Roache et al., 2023). Påvirkning av oppfatninger og holdninger kan dermed forstås som en sentral del av Russlands forberedelser til invasjonen. En metaanalyse av Bierwiazzonec og kolleger (Bierwiazzonec, Gundersen & Kunst, 2022) knyttet troen på konspirasjonsteorier til uønsket adferd under Covid-19-pandemien. Forskerne fant lavere vaksinetall og mindre sosial distansering blant de som trodde på konspirasjonsteorier, hvilket underbygger sammenhengen mellom endring i oppfattelse og adferdsendring. Fiktive opphetede debatter på sosiale medier kan også bygge en feilaktig oppfatning av konflikter mellom interesser og grupper. Dette er bare noen eksempler. Poenget som gjøres her er at påvirkning av holdninger og oppfatninger både brukes og har mange ulike effekter, inkludert på adferd.

## **2.2 Psykologiske sårbarheter for påvirkning**

Kognitiv elaborering (dybdetenkning) er basis for dannelsen av robuste holdninger og oppfatninger (Chaiken, 1980; Petty & Cacioppo, 1986), og er også indikert å være sentralt for å undersøke påstander og nyheter på sosiale medier og dermed kunne oppdage feil og forsøk på påvirkning (Bago, Rand & Pennycook, 2020; Maertens, Roozenbeek, Basol & van der Linden, 2021). Manglende elaborering og bruk av heuristikker (dvs., kognitive snarveier basert på tidligere erfaringer, oppfatninger og følelser; Chaiken, 1980) kan derimot gjøre oss lettere å lure og dermed mer påvirkelige av falske nyheter og desinformasjon (Bago, et al., 2020; for en oversikt se også, Bjørnstad, 2019). Appellering til følelser er for eksempel en mye brukt teknikk

---

---

for å gjøre et budskap troverdig og gi en påvirkningseffekt (se f.eks., Carrasco-Farré, 2022; Roozenbeek & van der Linden, 2019). I tråd med dette, fant Filkukova & Langguth (2021) at emosjonell respons økte sannsynligheten for at deltagerne trodde på fiktive nyheter.

Generelt er autonomi, ansvarsfølelse og personlig involvering funnet å ha en positiv sammenheng med kognitiv elaborering (Fiske & Taylor, 2017). Dette kan forstås som motivasjonsfaktorer for å gå mer i dybden på en sak, nyhet eller melding. Uten motiverende faktorer henfaller mennesker ofte til mer overfladisk tenkning og heuristikker, det vil si at mennesker ikke bruker mer innsats enn det som synes å være strengt tatt nødvendig (se «cognitive misers»; Schumann, Kotowski, Ahn & Haugtvedt, 2012). Hvis kommunikator oppfattes å være troverdig og man har tillit til hen, at det som blir kommunisert er i tråd med ens eksisterende holdninger og oppfatninger, at saken ikke er så viktig for en selv, eller at en ikke opplever å ha ansvar, øker sannsynligheten for bruk av heuristikker og mer overfladisk tenkning (f.eks., Priester & Petty, 2003; Fiske & Taylor, 2017). I grupper på sosiale medier kan man få en slik effekt, da man gjerne har større tillit til de som tilhører gruppen enn de som er utenfor og dessuten gjerne har et meningsfelleskap innad i gruppen (for en gjennomgang se Bjørnstad, 2019; Sivertsen et al., 2021).

Videre viser forskning på kognisjon at mer kompliserte budskap krever mer motivasjon for at mottakeren skal ta det inn (Fiske & Taylor, 2017). Falske nyheter og desinformasjon er funnet å være mindre kognitivt krevende å prosessere (Carrasco-Farré, 2022). Resultatet er da gjerne at mer kognitivt krevende faktainformasjon, forskning, motargumenter og såkalt «debunking» (f.eks., Lewandowsky et al., 2020) av falske nyheter og desinformasjon, risikerer ikke en gang å bli lest eller lyttet til fordi det oppfattes som for kognitivt krevende.

Grupper på sosiale medier fungerer ofte som en kapasitetsbesparende sil; man slipper å forholde seg til det store og hele kompliserte bildet ved å la gruppene man tilhører gjøre jobben for seg. Jo mer komplisert verden og nyhetsbildet blir, jo større behov er det sannsynlig at individene får for å velge slike kapasitetsbesparende løsninger. Distraksjoner og stress bidrar også til lavere kognitiv kapasitet og øker behovet for å bruke heuristikker og kapasitetsbesparende løsninger (Fiske & Taylor, 2017). Høy arbeidsbelastning, krig, uro, pandemier, økonomiske nedgangstider og lignende er situasjoner som skaper stress, som igjen er med på å øke betydningen av grupper på sosiale medier.

Problemet er at slike prosesser gjør oss sårbare for påvirkning innenfra gruppene på sosiale medier. Aktører som ønsker å påvirke oss kan infiltrere gruppene på sosiale medier, opprette grupper på falske premisser og lignende. Algoritmene på sosiale medier og søkemotorer for øvrig, som er innstilt på å gi deg mer av det samme, vil være med på å understøtte forestillingen om sannhetene i ditt eget «gruppe-univers» (for mer om dette, se også, Sivertsen et al., 2021). Dette kan gjøre det vanskelig å snu feilaktige oppfatninger og holdninger hvis de har fått fotfeste i en befolkning eller i undergrupper av en befolkning.



---

---

## 2.2.1 Persuasjonsprinsippene

Robert Cialdini (f.eks., Cialdini, 2001) er kjent for å ha definert seks persuasjonsprinsipper eller teknikker for overtalelse. Disse teknikkene er godt kjent og brukes i alt fra reklame til politikk til påvirkningskampanjer. Forskningen på persuasjon viser at noen teknikker gjør det lettere å påvirke både individer og grupper. Hvis du liker kommunikator, gjerne basert på likhet, skryt/positiv tilbakemelding, attraktivitet, felles tilhørighet og lignende, er det større sannsynlighet for at du blir overtalt eller påvirket. Dette kan for eksempel utnyttes i sosiale grupper på nett og i kommunikasjon. Videre er det slik at vi mennesker ønsker å speile andre menneskers adferd – såkalt resiprositet. For eksempel hvis du får noe, vil du gjerne gi noe tilbake. Vi mennesker vil også gjerne følge sosiale normer, og mene, utrykke og gjøre det som blir forventet av oss – for å bli godtatt – spesielt av mennesker som ligner oss selv og som vi identifiserer oss med. Et slikt behov for sosial bekreftelse kan lett utnyttes i sosiale medier og grupper på nett. Det neste prinsippet handler om at mennesker tenderer til å ville være konsistente i sin adferd, holdninger og verdier (dvs., unngå kognitiv dissonans). Det betyr blant annet at det kan være vanskelig å snu hvis man først har godtatt en usannhet. Dernest utnyttes ofte autoritet i form av ekspertise eller makt/posisjon for å gjøre et budskap troverdig og skape en påvirkningseffekt. Til slutt har det vist seg at mennesker liker følelsen av eksklusivitet, slik at man gjerne ønsker seg mer av noe det er lite av. Dette er svært mye brukt i reklame: «kun få eksemplarer», «snart tomt» – løp og kjøp!

Oppsummert dreier persuasjonsprinsippene seg altså om: 1) å like («liking»), 2) resiprositet («reciprocity»), 3) sosial bekreftelse/konsensus («social proof»), 4) konsistens («consistency /commitment»), 5) autoritet («authority») og 6) knapphet («scarcity»). Som indikert i Bjørnstad (2019) er det noe usikkerhet rundt effektiviteten til disse prinsippene i konteksten av sosiale medier, men Muscanell og kolleger fant at prinsippene ofte er blitt benyttet i forbindelse med nettsvindler (Muscanell, Guadagno & Murphy, 2014).

## 2.2.2 Feilkilder i menneskelig oppfattelse

Feilkilder i menneskelig oppfattelse («biases») utnyttes ofte også i påvirkningsøyemed. Rent kognitivt er det slik at vi mennesker lett kan la oss manipulere gjennom noen ganske enkle teknikker. Her er noen eksempler.

### 2.2.2.1 «Base-rate fallacy»

Mennesker tenderer til å overdrive betydningen av, og dermed lettere la seg påvirke av, sterke enkelthistorier fremfor statistikk, forskning og generell informasjon (f.eks., Fiske & Taylor, 2017). Dette forklares av enkelthistorienes livaktighet og evne til å vekke følelser, noe som gjør at de lettere blir lagret i minnet og husket. Hvis man derimot oppfatter den generelle informasjonen som viktig for en selv («salient») og den blir presentert før man har gjort seg opp en mening, blir den tillagt mer vekt (Ginossar & Trope, 1980; Chun & Kruglanski, 2006). Dette er noe av basisen for «prebunking»-teoriene og forskningen som presenteres under (kap. 2.4.2).

---

---

### 2.2.2.2 «Priming» og «framing»

Budskapets kontekst kan ha stor betydning for hvordan mennesker oppfatter en sak. Hva man får vite om en person eller sak rett i forkant av en hendelse vil farge ens oppfattelse – såkalt «priming» (f.eks., Fiske & Taylor, 2017). Dette brukes for eksempel mye under en valgkamp – de fleste politiske partier vil prøve å få oppmerksomhet rundt sine saker. Dette aktiverer og påvirker tilgjengeligheten til et tema i minnet. Man ser det også relatert til påvirkningskampanjer, at det er vanskelig å rette opp feilinformasjon i etterkant av at den har fått «slå rot» (f.eks., Lewandowsky et al., 2020).

Tilsvarende som priming påvirker rammen en sak settes i, eller selve fremstillingen, hvordan vi mennesker oppfatter den («framing»; f.eks., Kahneman & Tversky, 1984; Tversky & Kahneman, 1981). Dette er av typen, om glasset er halvfullt eller halvtomt. Bevisst bruk av ulike nyhetssjangre kan fungere som framing. For eksempel, mange nettbaserte ikke-redaktørstyrte nyhetsmedier som gjerne produserer mye falske nyheter benytter seg av navn som etterligner mer seriøse aktører med tradisjonelle redaksjoner. Falske nyheter kan også ikles en seriøs nyhetsdrakt for å lure oss (f.eks., Alme, 2019). Dette gjør det selvfølgelig vanskeligere å kjenne igjen falske nyheter og medfører at man ikke kan nøye oss med å sjekke stilen nyheten blir presentert i.

### 2.2.2.3 Attribusjon

En annen type feilkilde i vår oppfattelse dreier seg om feil-attribusjoner eller såkalte «biased» attribusjoner. Attribusjoner handler om sosial kognisjon; hvordan mennesker forstår den sosiale verden vi lever i ved å utlede kausale forklaringer på hendelser eller adferd (f.eks., Fiske & Taylor, 2017). Den mest kjente feil-attribusjonen er kanskje den *fundamentale attribusjonsfeilen* (Heider, 1958; Ross, 1977). Den består i at mennesker tenderer til å legge for mye vekt på en persons karakter fremfor situasjonen for å forklare hendelser. Videre er attribusjonene våre ofte selvsentrerte («self-centered») og selvforherligende («self-serving») (Miller & Ross, 1975; Ross & Sicoly, 1979). Attribusjonsfeilene gjør at mennesker lett kan forledes til å forklare negative hendelser med negative fortolkninger av andre personer eller grupper, noe som også kan forsterkes av ulik gruppetilhørighet på sosiale medier.

## 2.2.3 Oppfattelser av sant og usant

På bakgrunn av de mange feilkildene i menneskelig oppfattelsesevne (som beskrevet over), er det ikke overraskende at vi har problemer med å skille sant fra usant i det vi møter av nyheter og påstander på digitale plattformer og sosiale medier (som beskrevet innledningsvis: Medietilsynet, 2019, 2021). Forskning fra psykologisk fagfelt indikerer at mennesker generelt har problemer med å skille sant fra usant (f.eks., Piksa et al., 2022; Roozenbeek et al., 2022b). Forskningen til Piksa og kolleger (2022) indikerte at de fleste mennesker faller inn i to omtrent like store grupper: de som tror på, og de som tviler på det meste av nyheter på sosiale medier – uavhengig av om nyhetene er sanne eller usanne. De fant at et mindretall havnet i to andre grupper: de som tenderer til nesten alltid, og de som tenderer til nesten aldri, å kunne skille sant fra usant. Dette var imidlertid et relativt lite utvalg og forskningen hadde også flere andre

---

---

metodiske svakheter, men resultatene tolkes å kunne gi noen perspektiver på funnene fra Medietilsynet.

I studier av misinformasjon og falske nyheter er oppfattelser av hva som er sant og usant målt på mange ulike måter. Roozenbeek og kolleger (2022b) gjorde en sammenfattende studie av åtte ulike måter å måle dette på, fra nøyaktighet («accuracy»), manipulativitet, reliabilitet, troverdighet til sant/usant («real/fake»). Konklusjonen fra denne studien var at forskningsresultatene er sammenlignbare til tross for noe varierende skalaer. Den største forskjellen mellom skalaene var at de som målte manipulativitet oppnådde høyere konfidensskårer enn de andre. Det er med andre ord enklere for mennesker å gjenkjenne bruk av manipulasjonsteknikker, enn å avgjøre hva som er sant og usant.

## **2.3 Individuelle forskjeller i sårbarhet for des- og misinformasjon**

Som indikert over gir lavere kognitiv kapasitet mindre kognitiv elaborering og økt behov for å bruke heuristikker og andre kapasitetsbesparende løsninger (Fiske & Taylor, 2017), hvilket kan gjøre mennesker mindre kritiske og lettere å påvirke – gjennom blant annet falske nyheter og desinformasjon. En annen del av forskningen på kognisjon ser på sammenhengen mellom kognitiv kapasitet og mindre kognitiv elaborering ut fra individuelle forskjeller i kognitive evner og hvordan det kan forklare ulikheter i sårbarhet for å tro på des- og misinformasjon.

For eksempel skårer mennesker som tror mest på falske nyheter skårer lavere på kognitive tester som måler analytisk tenkning («Cognitive Reflection Test» – CRT; f.eks., Pennycook & Rand, 2019), åpenhet i tenkningen («Actively open-minded hinking» – AOT; f.eks., Baron et al., 2022; Roozenbeek et al., 2022b), og nysgjerrighet og det å like å tenke i dybden («Need for Cognition» – NFC; Cacioppo, Petty & Kao, 1984; Cacioppo, Petty & Morris, 1983; Borah, 2022). Samtidig er lav NFC blitt knyttet til overdreven og problematisk internett-bruk (Shi, Chen & Tian, 2011). Dette kan bety at de menneskene som er minst kritiske til det de leser på sosiale medier og minst evner å skille sant fra usant er den gruppen som bruker sosiale medier mest, noe som videre kan tolkes å kunne øke sannsynligheten for spredning av des- og misinformasjon. Det kan altså synes å være gode muligheter for aktører som ønsker hjelp til å spre usann informasjon på nett, hvis brukerne som minst evner å skille sant fra usant er overrepresentert blant de mest aktive.

### **2.3.1 Need for cognition (NFC) og kritisk tenkning (“Critical thinking”)**

“Need for cognition”, eller behov for kognisjon (NFC), refererer til individuelle forskjeller i tendensen til å ønske og like å bedrive krevende tenkning (Cacioppo et al., 1984). Forskjeller i NFC er blitt positivt knyttet til et stort antall andre kognitive forskjeller som for eksempel kunnskapsrikhet, intelligens og motivasjon for å søke og elaborere informasjon, og negativt knyttet til for eksempel autoritarianisme, «closed-mindedness» og forenkling ved unngåelse, forvriddning og ignorering av ny informasjon (for en oversikt se, Cacioppo, Petty, Feinstein & Jarvis, 1996). NFC har også blitt knyttet til hvor påvirkelig man er (f.eks., Borah, 2022; Cacioppo et al., 1983; Haugtvedt & Petty, 1992) og analytisk kognisjon eller metakognitive

---

---

aktiviteter<sup>1</sup> (CRT; Maloney & Retanal, 2020). Verplanken, Hazenberg og Palenewen (1992) fant en positiv sammenheng mellom NFC og informasjonssøkning («external information seeking effort»). På lignende vis fant Mokhtari, Davarpanah, Dayyani & Ahanchian (2013) en sammenheng mellom NFC og informasjonssøkende adferd (“information seeking behavior” – ISB). Informasjonssøking gir et bredere og dypere informasjonsgrunnlag og forstås derfor å bidra til lavere sannsynlighet for påvirkning. Lav NFC kan med andre ord sees som en sårbarhet for påvirkning, mens høy NFC gir robusthet (se også Bjørnstad, 2019).

NFC har videre blitt sett på som en predisposisjon for, eller som en del av kritisk tenkning («Critical thinking»; f.eks., Fischer, Spiker & Riedel, 2009; Heijltjes, van Gog, Leppink & Paas, 2014; Klaczynski, Fauth & Swanger, 1998). Kritisk tenkning er ansett å være en viktig evne for militært personell, spesielt i forhold til deres evne til tolkning av informasjon og beslutningstaking (Fischer, et al., 2009).

Det er også funnet en positiv sammenheng mellom NFC og utdanning og akademiske resultater (f.eks., Cacioppo et al., 1996; Liu & Eveland, 2005), og utdanning ser også ut til å beskytte mot tro på mis- og desinformasjon (Roozenbeek & van der Linden, 2019).

## **2.4 Kjennetegn på påvirkningsforsøk og teknikker for å øke robusthet**

Det foregår mye lovende forskning også på mottiltak mot påvirkningsforsøk og spredning av des- og misinformasjon (for en oversikt over psykologisk forskning se Abrams, 2021). Som tidligere indikert er det å gå ut med informasjon i etterkant av feilinformasjon vanskelig å nå igjennom med på en uforberedt befolkning (Lewandowsky et al., 2020). Et hovedmål blir derfor å finne måter å gjøre mennesker mer robuste mot feilinformasjon og påvirkningsforsøk.

### **2.4.1 Kjennetegn på forsøk på påvirkning**

Forskerne Roozenbeek & van der Linden (2019) ved Cambridge University tok utgangspunkt i seks teknikker som de fant at ofte brukes i desinformasjon i forsøk på påvirkning. De seks teknikkene, polarisering, skape følelser, spredning av konspirasjonsteorier, trolling, ansvarsfraskrivelse og falske kontoer («impersonation»), kan forstås som kjennetegn på desinformasjon og forsøk på påvirkning. I videreføringen av denne forskningen, har Roozenbeek, van der Linden, Basol og kolleger fokusert på: polarisering, følelser, konspirasjonsteorier, trolling og forsterkning («amplification», dvs. bruk av bots, falske kontoer og lignende for å spre meldinger/nyheter og få de til å se populære ut) (Roozenbeek & van der Linden, 2020); moralsk-emosjonelt språk, konspiratorisk argumentasjon og falske eksperter (Basol et al., 2021). I den nyeste publikasjonen (Roozenbeek et al., 2022), fokuserte forskerne på påvirkningsteknikkene: emosjonelt språk, inkoherens, falske dikotomier, uthengning («scapegoating») og personangrep («ad hominem attacks»).

---

<sup>1</sup> CRT er beskrevet både som et mål på analytisk tenkning og metakognitiv aktivitet (Maloney & Retanal, 2020; Pennycook & Rand, 2019).

---

---

Teknikkene studert i denne rekken av eksperimenter kan tolkes som variasjoner av de initielt studerte. For eksempel kan uthengning og personangrep sees som eksempler på trolling og ansvarsfraskrivelse, og falske kontoer et eksempel på en type teknikk brukt til forsterkning. Polarisering kan også forstås som en type eskalering.

Fra norsk kontekst, har Medietilsynet laget en sjekklister (Medietilsynet, 2022b) for å hjelpe befolkningen med å bli bedre til å oppdage falske nyheter og desinformasjon. De ber folk være skeptiske til sterke følelser, sensasjonelle/utrolige saker/nyheter, fengende overskrifter, bruk av store bokstaver og utropstegn, skjev fremstilling og bilder (fordi de kan manipuleres). De oppfordrer også til å sjekke saken i flere medier (fortrinnsvis noen man vet er til å stole på), samt adresselinjen og kilden.

Oppsummert synes appellering til og bruk av sterke følelser samt polarisering og skjev fremstilling å være sentralt i påvirkningsøyemed, poengtert både i forskning og som praktisk råd for gjenkjenning av Medietilsynet.

#### **2.4.2 «Priming», «framing» og «prebunking»**

Priming og framing, som presentert over, brukes av aktører som ønsker å påvirke, men man kan også bruke denne forskningen til å lage tiltak for å beskytte seg mot påvirkningsforsøk. For eksempel kan man se for seg at man ved å informere og utdanne befolkningen i kjennetegn på mis- og desinformasjon og hvordan dette typisk spres, kan øke robustheten i befolkningen mot slik påvirkning. Med faktainformasjon i forkant kan man farge tolkningen av etterfølgende påvirkningsforsøk – det vil si, lære å se den i et mer korrekt lys. Slike tiltak baserer seg dermed på de samme psykologiske prosessene som man finner i priming.

Flere forskere argumenterer for en lignende strategi som de kaller «prebunking» (f.eks., Basol et al., 2021; Lewandowsky et al., 2020). Dette går ut på å være i forkant med informasjon om saker/hendelser man forventer kan bli utsatt for påvirkningsforsøk gjennom desinformasjon og falske nyheter. Eksempler på slike saker/hendelser kan være lokale og nasjonale valg og vaksiner og vaksiner under Covid-19 pandemien. Ved å gå bredt ut med forskningsbasert informasjon tidlig i forbindelse med sistnevnte, kunne man demme opp for mye des- og misinformasjon, konspirasjonsteorier og falske nyheter.

#### **2.4.3 Medietreningsstudier («media-literacy»-studier)**

Medietrening kan forstås som en måte å prime mennesker til å lettere kjenne igjen mis- og desinformasjon og påvirkningsforsøk. Som indikert over, gir Medietilsynet i Norge råd i form av en sjekklister om hva man bør se etter for å avdekke feilinformasjon og påvirkningsforsøk (Medietilsynet, 2022b). Andre eksempler er amerikanske The News Literacy Project («The News Literacy Project», 2022) som også har online kurs, blant annet for elever i skolen. Slike tiltak er funnet å kunne ha positive effekter på evnen til å skille mellom falske og faktiske nyheter og å evaluere innhold på digitale medier (f.eks., Guess et al., 2020; McGrew, 2019).

---

---

#### 2.4.4 «Accuracy-nudge» intervensjon

En annen måte å gjøre en befolkning mer robust mot påvirkningsoperasjoner er å be de reflektere over sannheten i påstander og nyheter. Blant annet fant Pennycook og kolleger (Pennycook et al., 2020) at en slik såkalt «accuracy-nudge» intervensjon hadde en positiv effekt på deltageres selv-rapporterte tilbakeholdenhet med å dele det de trodde var usanne nyheter med andre på Twitter. Slike påminnelser kan forståes som en slags primer; påminnelsen aktiverer fokuset på sannhet/usannhet før lesningen starter og gjør at mennesker i større grad legger oppfattelse av korrekthet til grunn for beslutningen om eventuelt å dele en nyhet (andre hensyn kan ellers være kongruens med egne holdninger og oppfatninger, politisk overbevisning, gruppetilhørighet, ønske om «likes», osv.).

#### 2.4.5 Vaksine («inoculation») – strategi

Forskningsgruppen, *Cambridge Social Decision-Making Lab* (CSDML), ved Cambridge University med blant andre Roozenbeek, Van der Linden, Basol og Maertens har siden 2019 publisert mange artikler rundt hvordan man kan høyne robustheten mot falske nyheter og desinformasjon ved det de omtaler som vaksine-teknikk («inoculation technique»; benevnes ofte også som en type «prebunking»). Vaksinen består av en opplæring i å kjenne igjen vanlige manipulasjonsteknikker i falske nyheter og desinformasjon som benyttes for å oppnå en påvirkning (som presentert over i kapittel 2.4.1). Dette ble gjort gjennom å trene deltagerne i å bruke disse manipulasjonsteknikkene selv i on-line spill. Forskningsgruppen har utviklet flere slike nettbaserte spill, blant annet, «Bad News» (Roozenbeek & Van der Linden, 2019), «Harmony Square» (Roozenbeek & Van der Linden, 2020) og «Go Viral!» (Basol et al., 2021).

Forskerne Roozenbeek & Van der Linden (2019) utviklet spillet «Bad news» med målsetting om å trene deltagerne i å kjenne igjen manipulasjonsteknikkene ved at de lærte å bruke dem selv. Post-testresultatene indikerte en positiv effekt av dette, i form av at deltagerne ble bedre til å kjenne igjen disse teknikkene. Men studien hadde noen svakheter som manglende kontrollgruppe, at deltagerne kunne bli negativt influert av opplæringen i påvirkningsteknikker, og usikkerhet rundt generaliserbarhet og langtidseffektene. Basol og kolleger (Basol, Roozenbeek & van der Linden, 2020) tok denne forskningen videre ved å ta i bruk kontrollgruppe og å inkludere et mål på svarsikkerhet (konfidens). Resultatene støttet funnene fra Roozenbeek & Van der Linden og viste i tillegg en positiv effekt av spillet på konfidensskårene. Maertens og kolleger (Maertens et al., 2021) tok for seg effekten av spillet over tid, og kunne vise at effekten avtok over tid slik at den ikke lenger var signifikant etter to måneder, eller tre måneder med en påminnelse (såkalt «boost») inkludert.

I videreføringen av denne forskningen, med bruk av spillet «Go Viral!», viste Basol og kolleger (Basol et al., 2021) at funnene var generaliserbare til flere land og språklige kontekster – det vil si, til Storbritannia, Frankrike og Tyskland. Gjennom bruk av spillet «Harmony Square» tok Roozenbeek & Van der Linden (2020) forskningen videre ved å ha mer fokus på politisk des- og misinformasjon. Resultatene støttet opp om de tidligere funnene i form av at deltagerne fikk økt evne til å kjenne igjen manipulasjonsteknikkene, økt konfidens i gjenkjennelsen av manipulasjonsteknikkene og lavere motivasjon til å dele misinformasjon (selv-rapportert).

---

---

I den nyeste publikasjonen (Roozenbeek et al., 2022), tok forskerne eksperimentene ut i en sosial medie-setting (YouTube) og inkluderte kontrollgrupper. Her konstruerte Roozenbeek og kolleger en forenklet form av de tidligere eksperimentene. Forskerne benyttet fem korte videoer (uavhengig variabel, istedenfor spillene) som hver for seg eksemplifiserte én påvirkningsteknikk, og testet deretter om deltagerne kjente igjen de ulike teknikkene i en påstand/nyhet de fikk lese (avhengig variabel). Også her fant de en positiv effekt av intervensjonen (dvs., videoene i disse eksperimentene), på evne til gjenkjennelse av påvirkningsteknikker i innhold vist på et sosialt medie. I tillegg konkluderte Roozenbeek og kolleger med at deltagerne ble sikrere på sine evalueringer, bedre til å skille troverdig fra ikke-troverdig innhold, og at de ble mindre villige til å dele feilinformasjon videre. En mangel i denne forskningen i forhold til naturlige situasjoner, var at påstandene/nyhetene deltagerne skulle evaluere ikke var tilknyttet noen avsender eller konto. Når forskningen på området indikerer at kilde er noe av den viktigste informasjonen for å evaluere troverdighet (f.eks., Fiske & Taylor, 2017; Priester & Petty, 2003), kan dette tolkes som en mangel. Villighet til å dele feilinformasjon var også basert på selv-rapportert intensjon, ikke faktisk adferd.

Pilot-eksperimentet presentert i denne rapporten søker å bringe denne type forskning videre ved å lage en enda mer realistisk eksperimentell setting som involverer en simulert sosial medie-kontekst. Dette inkluderer informasjon om avsenderens profil på det sosiale mediet. Videre hersker det usikkerhet rundt den relative effektiviteten til de ulike teknikkene som er brukt i alt fra vaksine-strategistudiene til medietreningsstudiene («media-literacy») til informasjonskampanjer basert på for eksempel forskningsbasert informasjon. Roozenbeek & van der Linden (2019) hevdet at deres spill-metode var mer effektiv fordi den innebar mer involvering av deltagerne og derfor oppnådde mer kognitiv bearbeiding eller elaborering. Dette er ikke blitt testet i en mer naturalistisk sosial medie-kontekst hvor deltagerne ikke bare blir testet i å kjenne igjen enkelte teknikker, men hvor de også evaluerer sannhetsgehalten i meldinger som kan inneholde både disse og andre manipulasjonsteknikker. Dessuten kan også vaksine/spill-metoden forstås å ha noen bivirkninger; opplæringen i å bruke manipulasjonsteknikkene aktivt kan rett og slett lede noen deltagere (spesielt de sårbare) til å bruke teknikkene selv snarere enn bare å kjenne dem igjen hos andre. Pilot-eksperimentet rapportert her vil derfor også teste ut flere teknikker, som overordnet kan forstås å være ulike typer «primer»-teknikker, opp mot hverandre i en simulert sosial medie-setting.

#### **2.4.5.1 Konfidens (svarsikkerhet)**

Basol og kolleger (Basol et al., 2020; Basol et al., 2021) fant at deltageres rapporterte konfidens (dvs. hvor sikre de var på svarene sine) i å kjenne igjen påvirkningsteknikker også ble høyere etter gjennomgang av spillene «Bad news» og «Go Viral». Spillene syntes med andre ord, å gi deltagerne både bedre forutsetninger for å kjenne igjen påvirkningsteknikkene som ble brukt i spillene og gi dem følelsen av å være sikrere på sine svar. Den nyeste forskningen av Roozenbeek et al. (2022) støtter disse funnene; konfidensen økte i takt med evnen til å gjenkjenne påvirkningsteknikkene.

---

---

Samtidig viser forskning fra kognitiv psykologi at de som har dårligst evner gjerne har for stor tro på egne evner – såkalt overkonfidens («overconfidence»; se f.eks., Miller & Geraci, 2011). Med andre ord er det ikke sikkert at det er en fordel med høy konfidens.

## 2.5 Forskningsspørsmål

På bakgrunn av forskningen presentert i dette kapittelet vil studien søke å svare på følgende forskningsspørsmål:

- 1) Kan man få til en målbar påvirkning av oppfatninger/holdninger gjennom en kortvarig eksperimentell påvirkning gjennom et sosialt medie?
- 2) Kan man høyne robustheten/motstandsdyktigheten mot påvirkning av oppfatninger/holdninger på sosiale medier gjennom tre forskjellige «primer»-teknikker; det vil si:
  - a) forskningsbasert informasjon (muntlig, høy troverdighet)?
  - b) forskningsbasert informasjon (skriftlig, høy troverdighet)?
  - c) et vaksine/ «inoculation»-spill?
  - d) Og i så fall, er disse «primer»-teknikkene likeverdige, eller er det forskjeller?
- 3) Kan man høyne evnen til kjenne igjen sanne nyheter og utsagn og robustheten mot å tro på usanne på sosiale medier gjennom tre forskjellige «primer»-teknikker; det vil si:
  - a) ved bruk av forskningsbasert informasjon (muntlig, høy troverdighet)?
  - b) ved bruk av forskningsbasert informasjon (skriftlig, høy troverdighet)?
  - c) ved bruk av et «inoculation»-spill?
  - d) Og i så fall, er disse «primer»-teknikkene likeverdige, eller er det forskjeller?
- 4) Er det sammenheng mellom evne til å kjenne igjen sanne og usanne nyheter og utsagn på sosiale medier og robusthet mot påvirkning av oppfatninger/holdninger?
- 5) Er evne til gjenkjennelse av falske nyheter forbundet med høyere eller lavere konfidens i en sosial medie kontekst?
- 6) Gir lav NFC høyere sårbarhet og høyere NFC robusthet mot påvirkning av oppfatninger/holdninger og evne til gjenkjennelse av falske nyheter?



---

---

## 3 Metode

I dette kapittelet følger en presentasjon av forskningsdesign, operasjonalisering, verktøy, eksperimentelt oppsett, gjennomføring, utvalg, etiske betraktninger, mål, variabler og en overordnet beskrivelse av typen analyser som følger i kapittel 4.

### 3.1 Forskningsdesign og operasjonalisering

For å belyse problemstillingene og hypotesene beskrevet innledningsvis, er det valgt et eksperimentelt design. Eksperimentet vil teste ut bruk av informasjon gitt i ulike former som metode for å øke robustheten mot påvirkningsforsøk på sosiale medier. Eksperimentet er utforskende (eksplorerende) og defineres som et piloteksperiment. Det var viktig å finne ut om metode og oppsett fungerte med hensyn til å kunne fremskaffe resultater og svar på forskningsspørsmålene. Flere gjennomføringer vil være nødvendig for å oppnå tilstrekkelig statistisk styrke i eksperimentet og generaliserbarhet til flere populasjoner enn den som deltagerne i piloteksperimentet representerer.

For å undersøke forskningsspørsmålene (kapittel 2.5) i et eksperimentelt oppsett, ble påvirkning operasjonalisert både som målbar endring i oppfatning av eller holdning til saker og tro på usanne og gråsome<sup>2</sup> «tweets» (*tweets* indikerer Twitter-type meldinger – med og uten lenker til nyheter). Robusthet mot påvirkning forstås dermed som mindre eller manglende oppfatningsendring og mindre eller manglende tro på usanne og gråsome tweets. Selve påvirkningen i dette eksperimentet besto av å lese «tweets» på nett (presentert i det følgende, kapittel 3.2.1).

### 3.2 Verktøy

#### 3.2.1 Simuleringsverktøyet «Mastodon»

Eksperimentet ble gjennomført ved bruk av «Mastodon» (et åpent kildekode-verktøy) for realistisk å simulere Twitter – en av de mest brukte sosiale medietjenestene. Mastodon muliggjør å undersøke oppfatning av saker og påvirkning av disse gjennom ulike typer budskap på et sosialt medie med ulike grupper, samt å teste ut ulike teknikker for å fremme robusthet mot påvirkning på sosiale medier. Mastodon er en del av «Somulator», som nå også inkluderer fire andre simulerte sosiale medietjenester. Somulator og Mastodon er nærmere beskrevet i Bergh (2023).

---

<sup>2</sup> *Gråsome tweets* indikerer tweets som er en blanding av sant og usant. Det kan f.eks. være overdrivelser, meninger og saker tatt ut av kontekst.

---

---

### 3.2.2 Spørreskjema

Deltagerne svarte på to spørreskjemaer, ett før eksperimentet («pre-eksperiment») og ett etter eksperimentet («post-eksperiment»). Det pre-eksperimentelle spørreskjemaet inneholdt spørsmål om bakgrunn (dvs., kjønn, alder, utdanning og etat), Need for Cognition (NFC; Cacioppo et al., 1984) og oppfatning/holdning til seks ulike saker. Det post-eksperimentelle spørreskjemaet var todelt: den første delen inneholdt de samme oppfatning/holdnings-spørsmålene som pre-eksperimentelt spørreskjema, den andre delen skulle måle deltagerens evne til å gjenkjenne sanne, usanne og gråsones utsagn/tweets. Spørreskjemaene og målene er nærmere beskrevet i kapittel 3.7-3.9 nedenfor og gjengitt i vedlegg (A.3).

### 3.3 Eksperimentelt oppsett

Vi hadde et eksperimentelt oppsett med tre eksperimentelle betingelser og én kontroll. De tre eksperimentelle gruppene fikk forhåndsinformasjon, forstått som en type positiv priming. Denne primingen var altså den uavhengige variabelen ( $UV$ ) som studien søkte å undersøke effekten av. Gruppene var som følger:

1. Den første eksperimentelle gruppen fikk informasjon i form av forskning formidlet skriftlig om påvirkning på sosiale medier og hva man bør være bevisst for ikke å bli påvirket av fiendtlige aktører (innholdet er identisk med den muntlige informasjonen gitt til gruppe to). For deltagerne så informasjonen ut til å være skrevet av FFI-forskere og ligge på FFIs hjemmesider.
2. Den andre eksperimentelle gruppen fikk informasjon i form av forskning formidlet muntlig (innholdet var identisk med den skriftlige informasjonen gitt til gruppe én) om påvirkning på sosiale medier og hva man bør være bevisst for ikke å bli påvirket av fiendtlige aktører. For deltagerne så informasjonen ut til å være presentert av FFI-forskere og ligge på FFIs hjemmesider.
3. Den tredje eksperimentelle gruppen fikk informasjon i form av nett-spillet «Harmony Square», som har vært benyttet til å øke evnen til å kjenne igjen manipulasjonsteknikker ofte brukt i falske nyheter og utsagn (Roizenbeek & van der Linden, 2020; se kapittel 2.4.5).
4. Den fjerde gruppen, kontrollgruppen, fikk ingen relevant informasjon; de fikk høre podcast og lese rapporter om et annet urelatert forskningsprosjekt (menneskelig yteevne i kaldt vær), også på FFIs hjemmesider.

De avhengige variablene ( $AV$ ) var:

- Oppfatnings/holdningsendringer (differansen mellom de pre- og post-eksperimentelle målingene presentert i kapittel 3.7)
- Robusthet mot å tro på usanne og gråsones utsagn/tweets (målene er presentert i kapittel 3.8)
- Svaresikkerhet/konfidens relatert til bedømmelsen av sant, usant og gråsones (kapittel 3.8.1)

---

---

### 3.3.1 Primerne (de eksperimentelle betingelsene)

To av primerne bestod av forskningsbasert informasjon om påvirkning, presentert enten skriftlig eller muntlig – begge digitalt fremstilt. Her fikk deltagerne informasjon om temaer som:

- Hva som menes med påvirkning og hvordan påvirkningsoperasjoner oppnår effekt
- Påvirkning og desinformasjon i sosiale medier: egenskaper, algoritmer, narrativ og målgrupper
- Hvordan fabrikasjon, identitet, retorikk, symbolikk, og teknologi brukes for å oppnå påvirkning
- Psykologiske sårbarheter som gjør at mennesker blir påvirket: fra feilkilder i menneskelig oppfattelse, til følelser, persuasjon og den sosiale konteksten
- Vanlige teknikker som benyttes og sjekklister for å kjenne igjen forsøk på påvirkning

Den tredje primeren var spillet «Harmony Square» (Roozenbeek & van der Linden, 2020). Spillet går som beskrevet i kapittel 2.4.1 og 2.4.5 ut på å lære opp deltagerne i fem vanlige teknikker som brukes i påvirkningsøyemed: polarisering, følelser, konspirasjonsteorier, trolling og forsterkning. Deltagerne får selv prøve teknikkene i spillet.

### 3.3.2 Mastodon-meldingene

Mastodon-meldingene, heretter omtalt som «tweetsene» omhandlet seks ulike saker. FFI hadde inngått et samarbeid med Medie-linjen på Lillestrøm Videregående for hjelp til å opprette fiktive profiler og tweets på Mastodon og å sende disse ut til deltagerne på eksperimentdagen. De fikk være med oss som forskere og lære av oss på området påvirkning og bidro med på å lage tweets med målsetting om å kunne påvirke deltagerne på eksperimentet.

## 3.4 Gjennomføring<sup>3</sup>

Ved ankomst ble deltagerne tatt imot av FFI-forskerne (heretter omtalt som bare FFI). Deltagerne ble tilfeldig delt inn i fire grupper (randomisert inndeling); hver deltager trakk en lapp hvor det sto 1, 2, 3 eller 4 – hvilket betegnet det rommet og den gruppen de skulle gå til. Deltagerne fikk deretter ta opp egen laptop og logge seg på et felles Microsoft Teams-møte hvor FFI ga en kort innledning (5-10 min) om hva dagen skulle inneholde (grunnen til at innledningen ble holdt digitalt og ikke i et fellesrom skyldtes smittevernstiltak i forbindelse med den da pågående Covid-19-pandemien). Det ble poengtert at deltagelse var frivillig og at deltagerne kunne trekke seg når og om de skulle ville det.

Deretter ble deltagerne bedt om å gå inn på det nettbaserte pre-eksperimentelle anonymiserte spørreskjemaet (bakgrunn/kontrollvariabler, NFC og oppfatning/holdninger til utvalgte enkelttema; gjengitt i A.3.1).

---

<sup>3</sup> FFI hadde før selve gjennomføringen utført en test av det eksperimentelle oppsettet internt i FFI-prosjektet (*Cyber-Social Propaganda and Influence – Påvirkning i Informasjonsdomenet: C-SPI*).

---

---

Etter å ha besvart spørreskjemaene gikk deltagerne gjennom de ulike eksperimentelle betingelsene, altså primerne (beskrevet i kapittel 3.3.1). Dette innebar en aktivitet av en halv times varighet innen hver betingelse/gruppe, hvoretter de fikk en kort pause.

I hvert rom var det en FFI-forsker som kunne svare på spørsmål og geleidet deltagerne gjennom eksperimentet, som beskrevet i eksperiment-manualen gjengitt i Vedlegg (A.2). Manualen ble benyttet for å sikre at deltagerne i hver av gruppene fikk helt identisk informasjon med unntak av de eksperimentelle betingelsene.

Dernest fikk deltagerne vite at de hadde fått en bruker på Twitter-klonen Mastodon hvor det var lagt ut en del innhold (innlegg, lenker til nyheter, osv.) som de kunne få lese og forholde seg til på samme måte som de vanligvis ville gjøre hvis de var vant til å bruke Twitter. Deltagerne ble så bedt om å logge seg på Mastodon. Alle deltagerne fikk tilgang til de samme tweetsene og nyhetslinkene som ble sendt ut i fire omganger i løpet av én time. På Mastodon var det lagt inn fiktive tweets, med linker til nyhets saker fra både tradisjonelle redaktørstyrte medier og ikke-redaktørstyrte nettsteder som ofte har blitt forbundet med publisering av falske nyheter. Nyhetssakene var en blanding av sant, usant og gråsone – det vil si reelle nyheter, falske nyheter og vinklede/delvis sanne nyheter. Kildene eller opphavspersonene bak tweetsene, var også fiktive, selv om deres tilsynelatende tilhørighet kunne være til reelle organisasjoner. Kildene fremsto som mer eller mindre troverdige.

Deltagerne ble deretter bedt om å gå inn på nettsiden med det post-eksperimentelle anonymiserte spørreskjemaet (oppfatnings/holdningsspørsmål og sannhetsgehalt; A.3.2). Dette ble fulgt av en kort pause før FFI-forskerne ledet en oppsummerende diskusjon i hver gruppe om hvordan deltagerne hadde opplevd dagen. Deltagerne ble spurt om hvordan de synes oppgavene og dagen hadde vært så langt, hvordan de opplevde sakene på Mastodon og om (og hvordan) de syntes informasjonen de hadde fått i forkant påvirket hvordan de oppfattet sakene på Mastodon. Spørsmålsguiden til disse diskusjonene er gjengitt i eksperimentmanualen i Vedlegg (A.2). Svarene ble notert av FFI-forskerne på gruppenivå behandlet i en felles kvalitativ analyse.

Avslutningsvis gikk deltagerne igjen inn på et felles Microsoft Teams-møte hvor FFI gjennomførte en debrief med presentasjon av de ulike eksperimentelle gruppene, hypotesene, forskningsprosjektet og hvordan dataene ville bli behandlet videre (anonymisert, aggregert og kun til dette forskningsformålet). Etter dette ble det invitert til en fellesdiskusjon og alle deltagerne ble til slutt takket for innsatsen.

Deltagerne jobbet individuelt under hele eksperimentet.

### **3.5 Utvalg**

Utvalget ( $N = 20$ ) besto av en gruppe frivillige fra en organisasjon, hvor 19 var i alderen 21-30 år og én i alderen 31-40 år. De var alle sivile. Utdannelsesnivået varierte fra lavere grad/annet til masternivå, med et gjennomsnitt på ca. bachelorgradsnivå ( $M = 2.9$  år), hvor 75% hadde

---

---

utdannelse på bachelor- eller mastergradsnivå. Det var 11 kvinner og 9 menn i utvalget. Det var ingen samvariasjon mellom kjønn og utdannelse ( $r = .01, p = .97$ ).

### 3.6 Etiske betraktninger

Eksperimentet er gjennomført i henhold til FFI sine føringer med hensyn til etiske retningslinjer og GDPR (General Data Protection Regulation). Instituttets prosedyrer for slik godkjenning er fulgt. Det ble ikke samlet inn personopplysninger under eksperimentet; alle data var anonymisert. Det var heller ingen innsamling av sensitive opplysninger. Deltagerne fikk lese oppdiktede tweets og det innsamlede datamaterialet besto i hovedsak av deltageres syn på og oppfatninger av disse – i form av svar på spørsmål i et spørreskjema. Demografiske spørsmål var begrenset til kun overordnede ikke-identifiserbare spørsmål.

Alle deltagerne ble ved oppstart informert om hva de skulle være med på i løpet av dagen og det overordnede fokuset i dette FFI-forskningsprosjektet. De ble også informert i forkant og underveis i eksperimentet om at deres deltagelse var på frivillig basis og at de sto fritt til å kunne trekke seg underveis. Det var 34 påmeldte deltagere og 20 som møtte opp på eksperimentdagen. Det var én person som valgte å gå før siste spørreskjema, ellers valgte de resterende 19 å gjennomføre. Deltagerne fikk til slutt en debrief hvor de ble nærmere informert om forskningsaktiviteten og hvilke forskningsspørsmål som ble undersøkt gjennom eksperimentet de hadde vært med på.

Et etisk dilemma i denne forskningen er at man må indusere en viss form for påvirkning for å kunne studere effektene av mottiltak. Men tweetsene i forskningsoppsettet her beskrevet utgjør ingen forskjell fra det deltagerne, og alle andre som leser tweets, utsetter seg for i det daglige. Det er derfor vurdert at den påvirkningen deltagerne ble utsatt for i eksperimentet, er en betydelig mildere form enn det man utsettes for gjennom regulær bruk av nett og sosiale medier. Tweetsene benyttet i denne studien var kjennetegnet av ganske lav troverdighet både i form og innhold, og pågikk altså kun over svært kort tid (1 time). Hvis man kan påvise noen form for påvirkning av dette, kan det forstås som en indikasjon på at påvirkningen befolkningen i store deler av verden utsettes for i det daglige kan være betydelig større enn mange hittil har antatt. Ikke minst fra aktører som bevisst ønsker å påvirke, og som systematisk og over lang tid legger ut massive mengder og et nettverk av, ikke bare tweets, men også falske nyheter, falske profiler, desinformasjon og så videre. Dette bakteppet peker på at denne studien er relevant og viktig; det er behov for å bedre forstå hvor påvirkelige vi mennesker faktisk er, og ikke minst, hva som kan gjøres for å øke vår robusthet. Dette anses som en svært viktig del av arbeidet med å trygge det norske samfunnet og menneskene der. Alle aktørene i totalforsvaret inngår i dette. Ved hjelp av en grundig debrief i etterkant, er det vår vurdering at deltagerne også fikk mulighet til å justere sine oppfatninger tilbake og vil kunne ha oppnådd litt høyere robusthet med hensyn til risiko for å bli påvirket i det daglige på sosiale medier etter deltagelsen.

Tilbakemeldingene fra deltagerne i studien ga indikasjoner på at de hadde opplevd deltagelsen som berikende. De fikk to presentasjoner om FFI i tillegg til innsyn i FFI-prosjektet. Deltagerne ble tilbudt forfriskninger og lunsj.

---

---

### 3.7 Målene på oppfatning av saker med reliabilitetsanalyser

Vi hadde seks mål på oppfatninger av ulike temaer som deltagerne hadde fått lese om på de simulerte tweetsene med tilhørende nyhetssaker, kalt «Holdning 1-6». Oppfatningene ble alle målt på en 5-punkts Likert-skala med svaralternativer fra «helt uenig» (1) til «helt enig» (5), administrert pre- og post-eksperimentelt. Endring (differansen) fra pre- til post-eksperiment på disse seks oppfatningene eller holdningene representerte det ene settet med avhengige variabler (AV). Benevnelsene holdninger og oppfatninger sidestilles i det følgende.

#### 3.7.1 Tillit til vaksine/myndigheter

Holdning 1, Tillit til vaksine/myndigheter, ble målt med seks spørsmål: 1. Jeg føler meg trygg på vaksinene mot Korona/Covid-19, 2. Generelt føler jeg meg trygg på myndighetenes anbefalinger, 3. Jeg mener vaksinene mot Korona/Covid-19 er utrygge, 4. Generelt føler jeg meg ikke trygg på myndighetenes anbefalinger, 5. Jeg har tillit til myndighetenes pandemiltak, 6. Jeg har ikke tillit til myndighetenes pandemiltak. Spørsmål 3, 4 og 6 ble rekodet før utregning av gjennomsnitt-skåre, slik at høy skåre betyr høy tillit til vaksine/myndigheter. Målet viste høy reliabilitet med en Cronbachs alpha,  $\alpha = .85$  basert på pre-eksperimentell måling, og akseptabel reliabilitet med  $\alpha = .70$  basert på post-eksperiment måling.

#### 3.7.2 Oppfatning av konflikt by-land

Holdning 2, Oppfatning av konflikt by-land, ble målt med spørsmålet, «Det er mye splittelse mellom by og land i Norge». Høy skåre betyr oppfatning av at det er mye konflikt by-land i Norge.

#### 3.7.3 Oppfatning av Forsvaret (negativ)

Holdning 3, Oppfatning av Forsvaret (Negativ), ble målt med fire spørsmål: 1. Jeg har et positivt inntrykk av det norske Forsvaret, 2. Norge bruker for mye penger på Forsvaret, 3. Jeg har et negativt inntrykk av det norske Forsvaret, 4. Forsvaret tar ikke miljøhensyn. Spørsmål 1 ble rekodet slik at høy skåre indikerer negativ oppfatning av Forsvaret. Målet viste en noe lav reliabilitet med en  $\alpha = .51$  basert på pre-eksperimentell måling, og  $\alpha = .59$  basert på post-eksperiment måling. Dette er et mål som innholdsmessig dekker ulike sider av Forsvaret, og derfor kan forstås som en heterogen skala hvor det er naturlig med lavere alfa-verdi enn den konvensjonelle grensen på  $0.70^4$ .

---

<sup>4</sup> Selv om alfa-verdiene ideelt sett burde vært høyere, er standarder for akseptabel reliabilitet, som  $.70$  (Nunnally, 1978, p. 245), konvensjoner og ikke klare grenseverdier. For en diskusjon, se Pedhazur & Schmelkin, 1991 (s. 109–110). Alfa-koeffisienten avhenger videre i stor grad av antallet indikatorer i målet (skalaen) slik at det med et lavt antall indikatorer må forventes lavere alfa-nivåer, spesielt i tilfeller der indikatorene fanger opp ulike, men relaterte aspekter ved et fenomen.

---

---

### 3.7.4 Oppfatning av likestilling (negativ)

Holdning 4, Oppfatning av likestilling (negativ), ble målt med to spørsmål: 1. Likestillingen har gått for langt – kvinner har fått fordeler på bekostning av menn på mange områder, 2. Likestillingen har fortsatt langt å gå for at kvinner skal stå på likefot med menn. Spørsmål 2 ble rekodet slik at høy skåre indikerer en negativ oppfatning av likestilling. Målet viste en noe lav reliabilitet med en Cronbachs alpha,  $\alpha = .63$  basert på pre-eksperimentell måling, og  $\alpha = .45$  basert på post-eksperiment måling. Lav N gir mer tilfeldigheter/ustabilitet i målingene og kan forklare forskjellen i reliabilitetsestimater mellom de pre- og post-eksperimentelle målingene. Som nevnt tidligere gjør et lavt antall indikatorer det generelt vanskeligere å oppnå høy alfa-verdi.

### 3.7.5 Oppfatning av at det er mye rasisme i Norge

Holdning 5, Oppfatning av at det er mye rasisme i Norge ble målt med spørsmålet, «Det er mye rasisme i Norge». Høy skåre indikerer tro på at det er mye rasisme i Norge.

### 3.7.6 Oppfatning av president Biden (positiv)

Holdning 6, Oppfatning av president Biden (positiv), ble målt med spørsmålet, «Jeg har godt inntrykk av president Biden». Høy skåre indikerer en positiv oppfatning.

## 3.8 Mål på evne til å gjenkjenne sanne, og robusthet mot å tro på usanne og gråsome utsagn/tweets

Det var videre spørsmål som målte hvordan deltagerne oppfattet sannhetsgehalten i utsagn hentet fra tweetsene de hadde fått lese. Deltagerne fikk evaluere åtte usanne, én gråsome og tretten sanne påstander på en 5-punkts skala fra «helt usant» (1) til «helt sant» (5). Dette ble målt post-eksperiment. Et eksempel på usanne utsagn er, «Vaksinene mot Korona/Covid-19 er farlige» og et eksempel på sanne utsagn er, «Etniske nordmenn flytter fra Groruddalen». Svarene på disse spørsmålene om oppfattelse av sannhet representerte det andre settet med avhengige variabler (*AV*).

### 3.8.1 Mål på svarsikkerhet (konfidens)

Det ble også målt hvor sikre deltagerne var på sine oppfattelser av sannhet/usannhet. Denne svarsikkerheten (konfidensen) ble målt med spørsmålet «Hvor sikker er du på dette svaret?» med svaralternativ på en 5-punkts skala fra «helt usikker» (1) til «helt sikker» (5). Dette målet var inkludert for å svare på et mer eksplorerende forskningsspørsmål; om deltagerne som var mer robuste mot å tro på usanne tweets tenderte til også å være mer sikre på sine svar sammenlignet med de som trodde mer på usanne tweets, eller motsatt.

---

---

### 3.9 Kontrollvariabler

Alder, utdanning, kjønn og NFC ble sett som mulige «confoundere» og ble derfor evaluert inkludert som kontrollvariabler i analysene. Alder var ikke nødvendig å kontrollere for videre da nesten alle i utvalget var fra samme aldersgruppe (som beskrevet over). Av kontrollvariablene var NFC ansett å være den viktigste for våre analyser, basert på teori og tidligere forskning (se kapittel 2.3.1). NFC var også ansett å være viktig i seg selv da det var en målsetting å undersøke forholdet mellom NFC og robusthet. Utdannelse er tatt med fordi det er funnet å ha sammenheng med NFC (f.eks., Cacioppo et al., 1996; Liu & Eveland, 2005) og påvirkning (Roozenbeek & van der Linden, 2019), og kjønn er med fordi det er tidligere funnet at kvinner og menn kan være i ulik grad påvirkelige ut fra hva slags medium de får informasjonen gjennom (Guadagno & Cialdini, 2007). Spørreskjemaene er gjengitt i vedlegg (A.3).

### 3.10 Need for Cognition (NFC)

Basert på hypotesen om at lav NFC gir sårbarhet og høy NFC gir robusthet mot påvirkning, ble NFC hos deltagerne målt. Dette ble gjort både for å teste denne hypotesen i seg selv, men også for å kontrollere for at NFC-nivået hos de enkelte deltagerne ikke påvirket resultatene av eksperimentet.

NFC ble målt med 18 spørsmål/indikatorer (Cacioppo et al., 1984) på en 5-punkts Likert-skala («strongly disagree – strongly agree»); 9 av spørsmålene var snudd og ble rekodet før det samlede målet ble kalkulert (dvs., gjennomsnittsverdien av de 18 spørsmålene; i tråd med bl.a. Kuvaas & Kaufmann, 2004). Utvalget hadde et samlet gjennomsnitt på  $M = 3.89$  og et standardavvik<sup>5</sup> på  $SD = 0.26$ .

Målet viste en lav reliabilitet med en Cronbachs alpha på  $\alpha = .41$ . «Corrected Item-Total Correlation» og «Cronbach's Alpha if Item Deleted» for målet er vist i tabellen under sammen med gjennomsnitt ( $M$ ) og standardavvik ( $SD$ ).

---

<sup>5</sup> Standardavvik ( $SD$ ) er et mål på graden av variasjon i skårer mellom respondentene.



Tabell 3.1 Gjennomsnitt ( $M$ ), standardavvik ( $SD$ ) = 0.26, «Corrected Item-Total Correlation» og «Cronbach's Alpha if Item Deleted» for NFC (Need for Cognition).  $N = 20$ .

NFC indikatorer	$M$	$SD$	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
NFC 1	3.50	1.19	.21	.37
NFC 2	4.25	.72	.34	.36
NFC 3*	4.10	.64	-.02	.43
NFC 4*	3.90	1.02	-.03	.45
NFC 5*	4.55	.51	.36	.37
NFC 6	3.55	.76	.21	.38
NFC 7*	3.65	1.04	-.34	.53
NFC 8*	3.60	.99	.35	.33
NFC 9*	3.25	1.02	.28	.35
NFC 10	4.20	.62	.05	.41
NFC 11	4.10	.79	-.01	.43
NFC 12*	4.30	.92	-.08	.45
NFC 13	3.70	.66	.39	.35
NFC 14	3.65	.75	.30	.36
NFC 15	4.15	.82	.48	.31
NFC 16*	3.85	.74	.12	.40
NFC 17*	4.10	.72	.03	.42
NFC 18	3.65	1.18	.03	.44

Notat. \* = Rekodet.  $N = 20$ .

NFC er et veletablert mål, men med en lav alfa var det nødvendig å se nærmere på om spørsmålene målte det samme<sup>6</sup>. Spørsmålene som ladet negativt (spørsmål 3, 4, 7, 11, 12 - «Corrected Item-Total Correlation») ble først fjernet, da negativ korrelasjon indikerer at de ikke måler det samme. Dette ga en alfa på  $\alpha = .61$ . Deretter ble også spørsmål 18 fjernet, da spørsmålet også hadde en negativ korrelasjon med de andre måleindikatorene etter at de fem første spørsmålene var tatt ut. Dette ga en alfa på  $\alpha = .67$  – en reliabilitet som er akseptabel for bruk i forskningssammenheng (Pedhazur & Schmelkin, 1991). Utvalgsstørrelsen gjør også at en må regne med noe lavere reliabilitetsmål. NFC basert på 12 indikatorer hadde da et gjennomsnitt på  $M = 3.86$  og et standardavvik på  $SD = 0.38$ .

<sup>6</sup> Reliabiliteten ved bruk av 6 indikatorer (dvs. NFC-indikator 1, 2, 3, 4, 11 og 15) som foreslått av Lins de Holanda Coelho, Hanel & Wolf (2020) ble også undersøkt. Dette ga en reliabilitet på  $\alpha = .21$ , en indikasjon på at denne kortformen ikke var brukbar på vårt utvalg.

---

---

### **3.11 Kvalitative data**

Det ble samlet inn kvalitative data fra gruppediskusjonene, den avsluttende fellesdiskusjonen, og observasjoner gjort underveis i eksperimentene. Dette var et pilot-eksperiment og kvalitative data ble ansett som viktig tilleggsinformasjon – både for å bedre forstå de kvantitative dataene (som var begrenset av utvalgsstørrelse) og for å kunne gjøre forbedringer i oppsettet til neste gjennomføring av eksperimentet.

Etter eksperimentene fasiliterte FFI-forskerne en kort diskusjon (10 min) i hver gruppe og tok notater. Det var tre hovedspørsmål vi søkte å få tilbakemeldinger på: hvordan deltagerne syntes oppgavene og dagen hadde vært, hvordan deltagerne opplevde sakene på Mastodon og hvordan deltagerne syntes informasjonen de fikk i forkant (primeren) hadde påvirket hvordan de oppfattet sakene på Mastodon.

I den avsluttende diskusjonen etter debrief på Microsoft Teams ble det også bedt om spørsmål og tilbakemeldinger fra deltagerne.

### **3.12 Statistiske analyser**

#### **3.12.1 Innledende analyser**

Første del av resultatene (kapittel 4) tar for seg kontrollvariablene (NFC, kjønn og utdanning) og deres relasjon til endring i oppfatning av saker, evne til å gjenkjenne sanne, gråsoner og usanne nyheter og utsagn på sosiale medier og hvor sikre deltagerne var på svarene på sistnevnte (konfidens). Endring for hvert av målene på oppfatning ble kalkulert (differansen) på basis av de pre- og post-eksperimentelle målingene (kapittel 4.1). Korrelasjonsanalyse ble benyttet for å undersøke om disse endringene i oppfatning hadde sammenheng med noen av kontrollvariablene. Kontrollvariablene som viste seg å ha sammenheng med endring, ble tatt med i de videre analysene der det var indikert et behov for det (dvs., signifikant/nær signifikant korrelasjon). Deretter ble det gjennomført korrelasjonsanalyser av kontrollvariablene og evnen deltagerne demonstrerte til å gjenkjenne sanne, gråsoner og usanne nyheter og utsagn på sosiale medier. Kontrollvariablene som demonstrerte en sammenheng med evne til å gjenkjenne sanne, gråsoner og usanne utsagn, ble inkludert i de relevante analysene. Til slutt ble det gjennomført en korrelasjonsanalyse for å avdekke eventuelle sammenhenger mellom kontrollvariablene og konfidens relatert til oppfattelsen av om tweetsene var sanne/usanne.

I tillegg til å indikere behov for å ta med kontrollvariablene i videre analyser, hadde de innledende korrelasjonsanalysene en egenverdi. I tråd med forskningsspørsmålene presentert i kapittel 2.8, var det en målsetting å avdekke om NFC og utdanning hadde sammenheng med robusthet mot å la seg påvirke (påvirkning operasjonalisert og målt som endring i oppfatning av saker) og evnen til å gjenkjenne falske nyheter på sosiale medier og forsøk på påvirkning.

---

---

### 3.12.2 Hovedanalyser

For å svare på forskningsspørsmålene og hypotesene (kapittel 2.8), hadde hovedanalysene som mål å undersøke om:

- 1) Det var endring i oppfatning som følge av den eksperimentelle påvirkningen (operasjonalisert som tweets på Mastodon).
- 2) Man kan høyne robusthet mot endring av oppfatninger ved tre ulike primer-teknikker. Det vil si, viste de tre eksperimentelle betingelsene noen effekt sammenlignet med kontrollgruppen; var det en eller flere primer-teknikker som viste seg effektive mot påvirkning, og var det eventuelt noen forskjell i hvilke som var mest effektive?
- 3) Man kan høyne evne til å kjenne igjen sanne nyheter og utsagn og robusthet mot å tro på usanne meldinger på sosiale medier ved tre ulike primer-teknikker. Og hvordan relaterte dette til svarkonfidens (dvs., hvor sikre deltagerne var i sine evalueringer av sannhet/usannhet)?
- 4) Det er sammenheng mellom evnen til å kjenne igjen sanne og usanne nyheter og utsagn på sosiale medier og robusthet mot påvirkningen av oppfatninger?

For å undersøke oppfatningsendringer (punkt 1), ble det utført en statistisk sammenligning av gjennomsnittsskåren på de seks oppfatningsmålene («holdning 1-6») for hele utvalget før og etter eksperimentell påvirkning ved hjelp av en t-test<sup>7</sup>.

Forskjell i robusthet mot påvirkning av oppfatning (punkt 2) mellom de tre eksperimentelle gruppene og kontrollgruppen ble gjennomført gjennom en Mixed Design ANOVA (analysis of variance)<sup>8</sup> separat for hvert av oppfatningsmålene.

Det ble deretter gjort en samlet evaluering av resultatene på de ulike oppfatningsmålene for å se om det kunne være en gjennomgående trend i effekten av de ulike primerne. Dette ble gjort gjennom en kvalitativ evaluering og en ren opptelling av hvilke grupper som syntes å være mest robuste ut fra tendensen i dataene. Disse analysene må tolkes med forsiktighet og kun som tendenser.

Det ble så undersøkt om det var en forskjell mellom de tre eksperimentelle gruppene og kontrollgruppen i forhold til evnen de demonstrerte til å gjenkjenne sanne og usanne nyheter og utsagn på sosiale medier (punkt 3). Dette ble gjort gjennom ANOVAer (en-veis)<sup>9</sup> som sammenlignet de tre eksperimentelle gruppene og kontrollgruppen med hensyn til korrekt oppfattelse av sannhet/usannhet. Det ble også undersøkt om det var forskjeller i konfidens.

---

<sup>7</sup> En t-test, eller nærmere bestemt en *paired samples t-test*, er en analyse som tester om differansen i gjennomsnittsverdi mellom to grupper er signifikant, det vil si, om det er mindre enn 5 % sannsynlighet for at resultatet skyldes tilfeldigheter ( $p < .05$ ).

<sup>8</sup> En *Mixed Design ANOVA (analysis of variance)* er en statistisk analyse der man undersøker om det er en signifikant endring i skårer mellom to eller flere tidspunkt, og om en eventuell endring er forskjellig for ulike grupper.

<sup>9</sup> En en-veis ANOVA er en statistisk analyse som undersøker om det er en signifikant forskjell i gjennomsnittsverdi mellom flere grupper med ett målepunkt.

---

---

Gjennomgående i alle analysene ble det kontrollert for NFC, kjønn og utdanning der det var relevant ved å inkludere disse som kovariater i ANCOVA (analysis of covariance) analyser.

Avslutningsvis ble det evaluert om det syntes å være en sammenheng mellom evnen til å kjenne igjen sanne og usanne nyheter og utsagn på sosiale medier og robusthet mot påvirkningen av ens oppfatninger gjennom en kvalitativ sammenligning av gruppene (punkt 4).

## 4 Resultater

Dette kapitlet presenterer hovedanalysene med tilhørende resultater. Det starter med en analyse av kontrollvariablene, fulgt av analysene av oppfatningsendring, og deretter analysene av evne til å gjenkjenne sanne, usanne og gråsone tweets. Kapitlet avsluttes med de kvalitative dataanalysene fra observasjoner, gruppe- og fellesdiskusjonene. Dette kapitlet inneholder kun en presentasjon av resultatene, oppsummering av de viktigste resultatene samt tolkninger og diskusjonen av disse kommer i neste kapittel (kapittel 5).

### 4.1 Kontrollvariablene

Det ble først utført korrelasjonsanalyser mellom kontrollvariablene NFC, kjønn og utdanning for å undersøke om det var noen interne sammenhenger mellom kontrollvariablene. Dette er vist i Tabell 4.1.

Tabell 4.1 Korrelasjonsanalyser<sup>10</sup>, gjennomsnitt (*M*) og standardavvik (*SD*) for kontrollvariablene NFC, kjønn og utdanning. *N* = 20.

Kontrollvariablene	<i>M</i>	<i>SD</i>	<i>r/p</i>	NFC	Kjønn
NFC (høy)	3.86	0.38			
Kjønn (mann = 1 / kvinne = 2)	1.55	0.51	<i>r</i> <i>p</i>	-.34 .15	
Utdanning (høy)	2.90	1.02	<i>r</i> <i>p</i>	.44 .05	.01 .97

Notat. *r* = Pearson Correlation (korrelasjonskoeffisienten). *p* = sannsynlig tilfeldighet.<sup>11</sup>

---

<sup>10</sup> Disse korrelasjonsanalysene kalles *zero-order correlation* analyser og undersøker relasjonene mellom to og to av variablene. Koeffisientene er standardiserte (mellom -1 og +1).

<sup>11</sup> *p* er sannsynligheten, eller probabiliteten, for at korrelasjonen egentlig er 0 i populasjonene og at man ved en tilfeldighet får en korrelasjon som er forskjellig fra null i utvalget i undersøkelsen. En *p*-verdi på under .05 ( $p < .05$ ) er ansett "signifikant" da det tilsier at sannsynligheten for at resultatet skyldes en tilfeldighet er mindre enn 5%.

Som forventet indikerte korrelasjonsanalysene ingen sammenheng mellom NFC og kjønn, men det var en marginalt signifikant sammenheng mellom NFC og utdanning.

#### 4.1.1 Korrelasjonsanalyser: oppfatningsendringer og kontrollvariablene

Korrelasjonsanalyser av kontrollvariablene, NFC, kjønn og utdanning opp mot endring på oppfatningsmålene fra pre- til post-eksperimentell påvirkning ble gjennomført for å undersøke om det var grunnlag for å inkludere kontrollvariablene i analysene av oppfatningsendring. Først ble endring i oppfatning av de seks ulike sakene fra pre- til post-eksperimentell måling kalkulert. En korrelasjonsanalyse undersøkte så om kontrollvariablene påvirket tilbøyeligheten til oppfatningsendring (Tabell 4.2).

Tabell 4.2 Korrelasjonsanalyser: oppfatningsendring og kontrollvariablene.  $N = 20$ .

		NFC	Kjønn (mann=1/kvinne=2)	Utdannelse (høy)
Endring pre-post.eksp. holdning 1	<i>r</i>	-.21	.20	-.16
	<i>p</i>	.37	.41	.50
Endring pre-post.eksp. holdning 2	<i>r</i>	-.19	-.04	-.50*
	<i>P</i>	.43	.88	.02
Endring pre-post.eksp. holdning 3	<i>r</i>	.25	-.18	.18
	<i>P</i>	.29	.46	.46
Endring pre-post.eksp. holdning 4	<i>r</i>	-.16	.36	.13
	<i>P</i>	.51	.12	.57
Endring pre-post.eksp. holdning 5	<i>r</i>	-.09	.26	.09
	<i>P</i>	.72	.28	.71
Endring pre-post.eksp. holdning 6	<i>r</i>	-.07	.00	-.48*
	<i>P</i>	.77	1.00	.03

Notat.  $r$  = Pearson Correlation. \*  $p < .05$ .

Tabell 4.2 indikerte to signifikante sammenhenger mellom oppfatningsendring og én av kontrollvariablene. Endring i holdning 2 og 6 fra pre- til post-eksperiment var negativt korrelert med utdanning. Det vil si at høyere utdanning var forbundet med mindre endring av oppfatning. Korrelasjonsanalysene indikerte altså at det var grunnlag for å kontrollere for utdanning i de videre analysene av endring av holdning 2 og 6.

#### 4.1.2 Korrelasjonsanalyser: evne til å gjenkjenne sanne, gråsome og usanne nyheter og kontrollvariablene

Det ble deretter utført korrelasjonsanalyser av kontrollvariablene NFC, kjønn og utdanning opp mot evnen deltagerne demonstrerte til å gjenkjenne sanne, gråsome og usanne nyheter og utsagn på sosiale medier. Analysene undersøker om utdanning og NFC hadde sammenheng med evne til å gjenkjenne falske nyheter og påstander på sosiale medier. Signifikante sammenhenger indikerer i tillegg at det er grunnlag for å inkludere den aktuelle kontrollvariablen i de ANOVA-analysene av gruppeforskjeller i kapittel 4.3. Korrelasjonsanalysene er presentert i Tabell 4.3, 4.4 og 4.5, én tabell for de usanne, én for gråsome/meninger og én for de sanne nyhetene og utsagnene deltagerne hadde fått lese på Mastodon.

Tabell 4.3 Korrelasjonsanalyser av kontrollvariablene NFC, kjønn og utdanning opp mot evnen deltagerne demonstrerte til å gjenkjenne usanne nyheter og utsagn på sosiale medier. Høy gjennomsnittsskåre (*M*) indikerer lav evne til å gjenkjenne usannhet. *N* = 19.

Usanne tweets – spørsmål nr.	<i>M</i>	<i>SD</i>		NFC	Kjønn (mann=1/kvinne=2)	Utdanninge (høy)
# 1.	1.42	.51	<i>r</i>	-.16	-.35	-.70**
			<i>p</i>	.51	.14	.00
# 2.	3.58	1.12	<i>r</i>	.06	.16	.11
			<i>p</i>	.80	.52	.67
# 6.	2.37	.90	<i>r</i>	.34	-.25	.33
			<i>p</i>	.16	.30	.17
# 8.	2.63	1.21	<i>r</i>	-.18	.01	.39
			<i>p</i>	.46	.99	.10
# 11.	1.79	.98	<i>r</i>	-.29	-.41	-.24
			<i>p</i>	.23	.08	.32
# 12.	3.89	.46	<i>r</i>	-.01	.28	.39
			<i>p</i>	.98	.25	.10
# 17.	1.68	.89	<i>r</i>	-.43	.18	.13
			<i>p</i>	.07	.46	.59
Usanne samlet (gjennomsnitt)	2.48	.27	<i>r</i>	-.28	-.16	.28
			<i>p</i>	.25	.52	.25

Notater. *r* = Pearson Correlation. \* *p* < .05. \*\* *p* < .01. *SD* = standardavvik.

Tabell 4.3 viser en negativ sammenheng mellom utdanning og klassifisering av én usann tweet. Den negative sammenhengen betyr at de med lavere utdanning tenderte til å klassifisere denne usanne tweeten mer som sann enn de som hadde høyere utdanning. Resultatene indikerer at utdanning bør kontrolleres for i de videre analysene av usann tweet #1 (kapittel 4.3).

Tabell 4.4 Korrelasjonsanalyser av kontrollvariablene NFC, kjønn og utdanning opp mot evnen deltagerne demonstrerte til å gjenkjenne gråsoner/meningsbaserte nyheter og utsagn på sosiale medier. Høy gjennomsnittsskåre (M) indikerer lav evne til å gjenkjenne gråsoner/mening. N = 19.

Gråsoner tweets – spørsmål nr.	M	SD	NFC	Kjønn (mann=1/kvinne=2)	Utdanning (høy)
# 5.	2.58	1.07	<i>r</i> .31 <i>p</i> .20	-.35 .15	.44 .06

Notater. *r* = Pearson Correlation. *SD* = standardavvik.

Det ble funnet en nær signifikant positiv sammenheng mellom kontrollvariabelen utdanning og oppfattelse av den ene gråsoner-tweeten. Resultatet betyr at flere med høyere utdanning tenderte til å tro at denne tweeten var sann. Selv om sammenhengene ikke var helt signifikante er det nært nok til at det forstås som at utdanning bør kontrolleres for i ANOVA-analysene (kapittel 4.3).

Tabell 4.5 Korrelasjonsanalyser av kontrollvariablene NFC, kjønn og utdanning opp mot evnen deltagerne demonstrerte til å gjenkjenne sanne nyheter og utsagn på sosiale medier. Høy gjennomsnittsskåre (M) indikerer høy evne til å gjenkjenne sannhet. N = 19.

Sanne tweets – spørsmål nr.	M	SD	NFC	Kjønn (mann=1/kvinne=2)	Utdanning (høy)
# 3.	4.05	1.03	<i>r</i> .44 <i>p</i> .06	-.17 .49	.17 .48
# 4. –	2.74	.73	<i>r</i> .42 <i>p</i> .07	.13 .59	.56* .01
# 7.	3.58	.84	<i>r</i> .43 <i>p</i> .07	-.05 .85	.07 .78
# 9.	2.84	.83	<i>r</i> -.34 <i>p</i> .15	-.17 .50	.00 1.00
# 10.	3.26	.93	<i>r</i> -.05 <i>p</i> .84	-.11 .67	-.13 .61
# 13.	2.21	1.03	<i>r</i> -.05 <i>p</i> .83	-.03 .89	.23 .35
# 14.	3.37	.68	<i>r</i> .37 <i>p</i> .12	-.49* .03	.09 .73
# 15.	3.37	.83	<i>r</i> .26 <i>p</i> .28	-.14 .57	.07 .77
# 16.	3.11	.74	<i>r</i> .23 <i>p</i> .34	-.02 .92	.16 .51
Sanne samlet (gjennomsnitt)	3.17	.387	<i>r</i> .39 <i>p</i> .10	-.25 .31	.29 .23

Notater. *r* = Pearson Correlation. \* *p* < 0.05. *SD* = standardavvik.

Tabell 4.5 indikerer at utdanning er signifikant positivt korrelert med å riktig klassifisere tweet #4 som sann, og at kjønn er negativt korrelert med å riktig klassifisere sann tweet #14 som sann.

---

---

Dette betyr at de som hadde høy utdannelse hadde høyere sannsynlighet for å klassifisere tweet #4 riktig, mens menn hadde høyere sannsynlighet for å klassifisere tweet #14 riktig. Samlet sett var hverken kjønn eller utdannelse i nærheten av å være signifikant korrelert med riktig oppfattelse av sannhet. Men en nær signifikant positiv sammenheng mellom NFC og sanne påstander samlet kan indikere at de med høyere NFC har hatt noe lettere for å kjenne igjen sanne tweets enn de med lavere NFC. Resultatene i Tabell 4.5 indikerer behov for å kontrollere for utdannelse i forhold til tweet #4 og kjønn i forhold til tweet #14 i de videre analysene (kapittel 4.3). I ANOVAen av samlevariabelen for de sanne tweetsene er det indikert mest grunnlag for å ta med NFC som kontrollvariabel.

## 4.2 Endring av oppfatning

### 4.2.1 Holdning 1: Tillit til vaksine/myndigheter

#### 4.2.1.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget en gjennomsnitt på  $M = 4.61$  før eksperimentell påvirkning og gjennomsnitt på  $M = 4.54$  etter eksperimentell påvirkning. Endringen kan defineres som liten ( $d = .21$ )<sup>12</sup> og var i retning mer negativ oppfatning av vaksiner og myndigheter etter eksperimentell påvirkning. Endringen var i forventet retning når man ser på innholdet i påvirkningen (tweetsene) deltagerne fikk lese. *En Paired Samples T-test* viste at denne effekten ikke var statistisk signifikant ( $p = .36$ ,  $t = 0.95$ ), og kan dermed ha vært tilfeldig.

#### 4.2.1.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

Deretter ble det gjennom en *Mixed Design ANOVA* som sjekket om det var noen forskjeller mellom de tre eksperimentelle gruppene og kontrollgruppen. Forskjellen mellom gruppene var litt under middels (endring  $\times$  gruppe,  $F^{13} = .39$ , *Partial Eta Squared*<sup>14</sup> = .07) og ikke signifikant ( $p = .76$ ). Det er med andre ord betydelig sannsynlighet for at forskjellen kan ha vært tilfeldig – altså ingen effekt av primerne de eksperimentelle gruppene ble presentert for. Figuren under (Figur 4.1) er likevel tatt med for å vise tendensen i dataene; gruppene med muntlig- og spill-primere var de som endret seg minst mellom før og etter eksperimentell påvirkning.

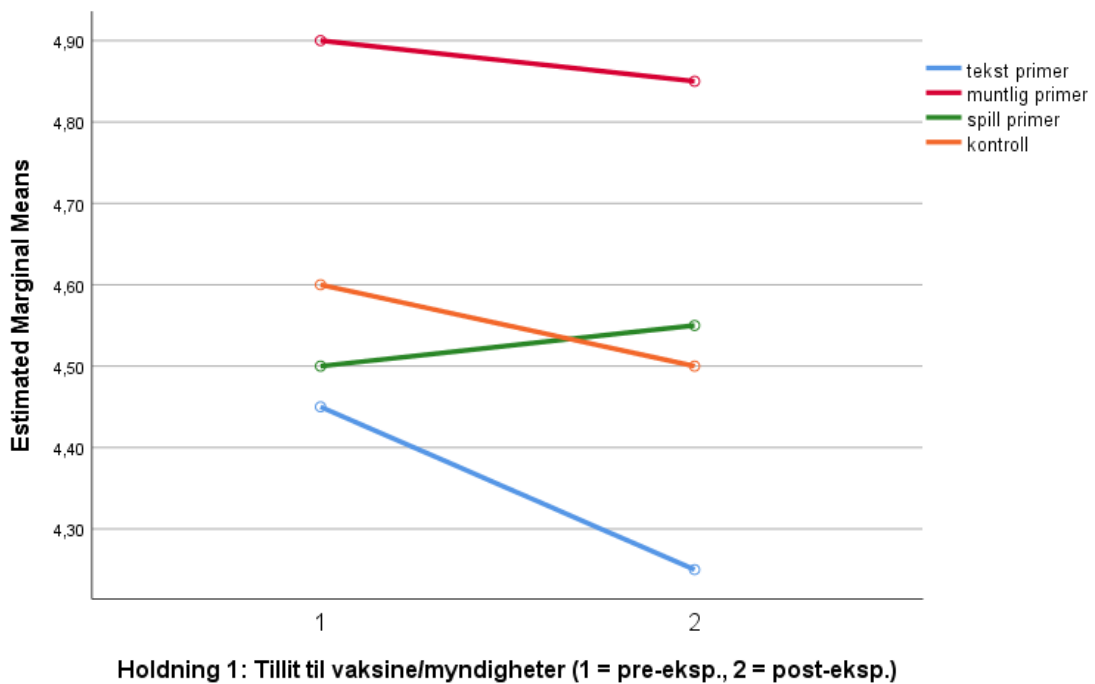
---

<sup>12</sup> *Cohen's d-value* ( $d$ ) er et mål på størrelsen på forskjell mellom grupper;  $d = .20$  er ansett å være en liten forskjell,  $d = .50$  en medium stor forskjell og  $d = .80$  en stor forskjell (Cohen, 1988).

<sup>13</sup>  $F$  er en såkalt testobservator beregnet gjennom å dele variansen i skårer mellom gruppene på variansen i skårer innad i gruppene. Jo større  $F$ -verdien er, jo lavere er sannsynligheten for at en forskjell i gjennomsnittsverdi mellom gruppene er oppstått ved en tilfeldighet (se f.eks. Pedhazur Scmelkin, 1991).

<sup>14</sup> *Partial Eta Squared* er et mål på effektstørrelse, i denne sammenheng indikerer det størrelsen på forskjellene mellom gruppene (*Partial Eta*<sup>2</sup> < .01 er ansett å være en liten forskjell, *Partial Eta*<sup>2</sup> = .09 en medium stor forskjell og *Partial Eta*<sup>2</sup> = .25 en stor forskjell (Cohen, 1988).





Figur 4.1 Gruppeforskjeller i robusthet mot å bli påvirket, holdning 1, tillit til vaksine/myndigheter.  $N = 20$ .

## 4.2.2 Holdning 2: Oppfatning av konflikt by-land

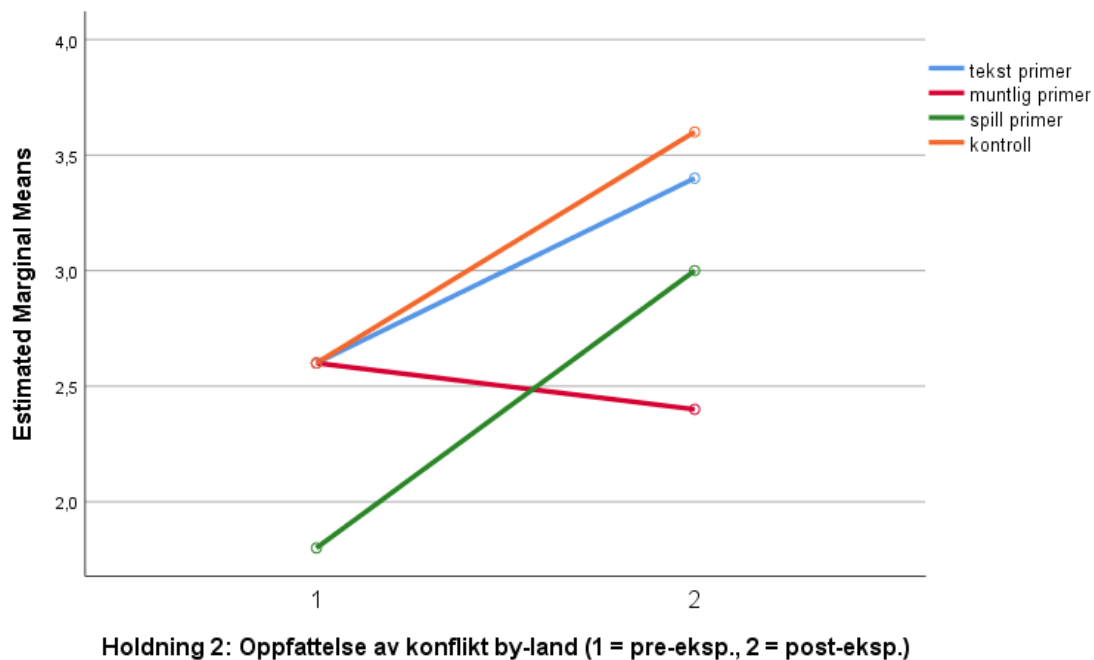
### 4.2.2.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget et gjennomsnitt på  $M = 2.40$  før eksperimentell påvirkning og et gjennomsnitt på  $M = 3.10$  etter eksperimentell påvirkning. Endringen kan defineres som stor ( $d = .81$ ) og var i retning oppfatning av at det var mer konflikt mellom by og land etter eksperimentell påvirkning. Endringen var i forventet retning i forhold til påvirkningen deltagerne ble utsatt for. En *Paired Samples T-test* viste at denne effekten var statistisk signifikant ( $p < .01$ ,  $t = -3.62$ ).

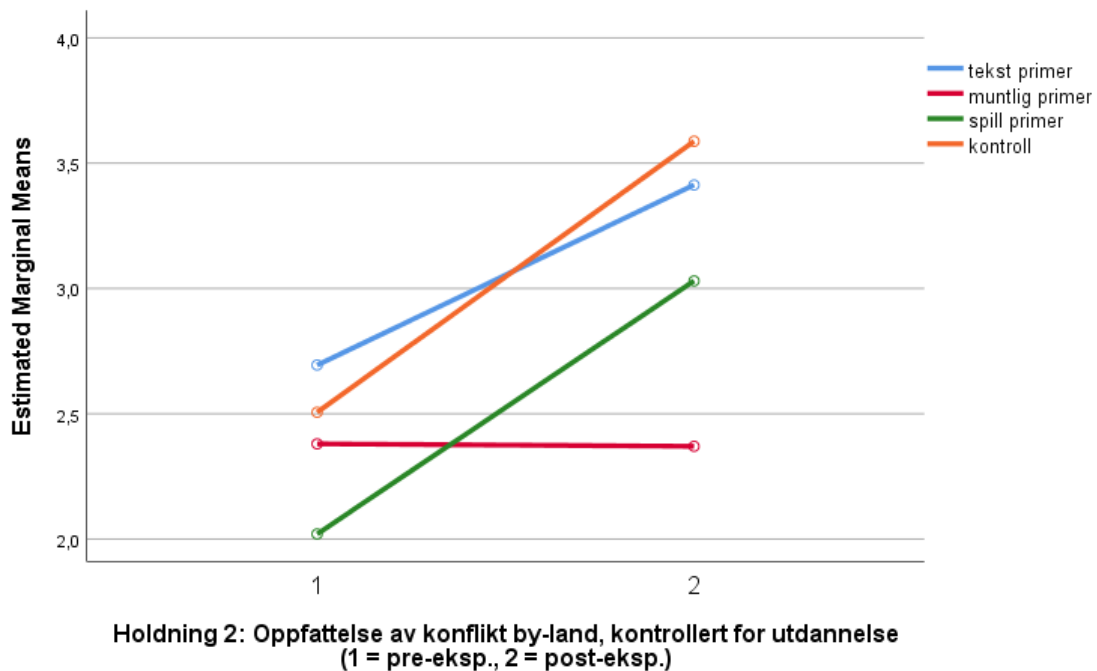
### 4.2.2.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

Deretter ble det gjennom en *Mixed Design ANOVA* undersøkt om det var noen signifikante forskjeller mellom de tre eksperimentelle gruppene og kontrollgruppen. Det ble funnet en signifikant forskjell i endring mellom gruppene (endring  $\times$  gruppe,  $F = 3.68$ ,  $p = .03$ , *Partial Eta*<sup>2</sup> = .41) – det vil si en stor effekt. Som vist i Figur 4.2 under var det gruppen som fikk muntlig primer som var signifikant forskjellig fra de andre gruppene. Denne gruppen demonstrerte ikke den samme endringen i oppfatning som de andre gruppene og kan forstås som å ha hatt større robusthet mot påvirkningen enn de andre gruppene. Det ble videre

kontrollert for utdanning (Figur 4.3) da utdanning var korrelert med denne variabelen (se Tabell 4.2), men utdanning påvirket ikke resultatene ( $p = .17$ ).



Figur 4.2 Gruppeskjeller i robusthet mot å bli påvirket, holdning 2, oppfatning av konflikt by-land.  $N = 20$ .



Figur 4.3 Gruffeforskjeller i robusthet mot å bli påvirket, holdning 2, oppfatning av konflikt by-land, kontrollert for utdannelse.  $N = 20$ .

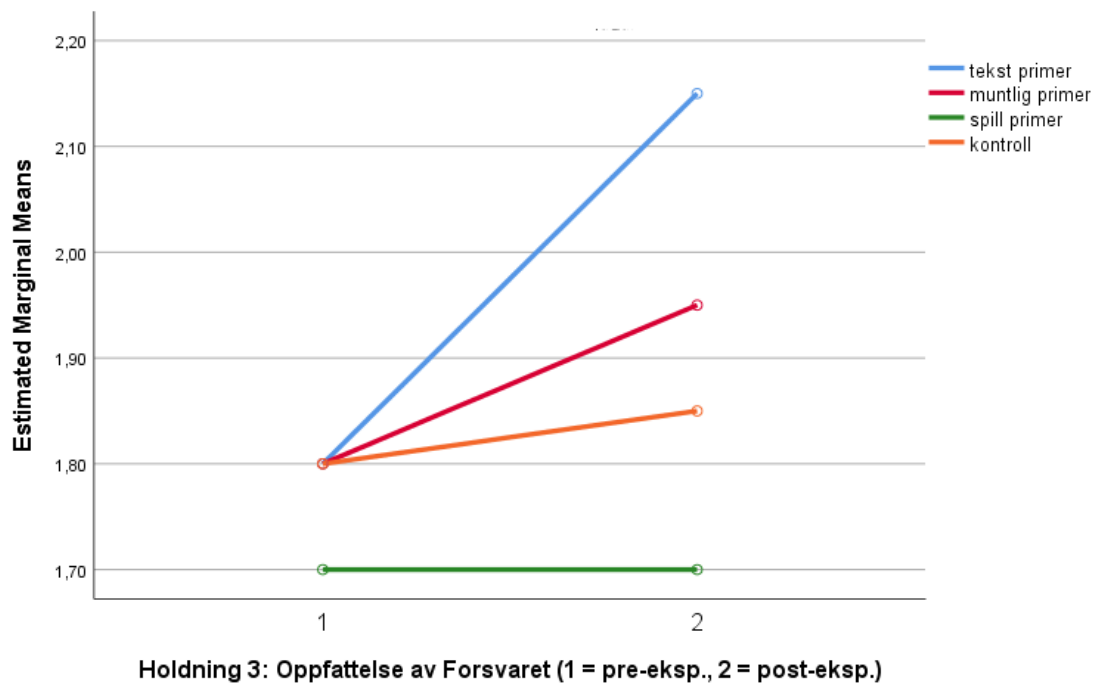
#### 4.2.3 Holdning 3: Oppfatning av Forsvaret (negativ)

##### 4.2.3.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget et gjennomsnitt på  $M = 1.78$  før eksperimentell påvirkning og et gjennomsnitt på  $M = 1.91$  etter eksperimentell påvirkning. Endringen kan defineres som middels ( $d = .48$ ) og var i retning mer negativ oppfatning av Forsvaret etter eksperimentell påvirkning. Endringen var i forventet retning når man ser på innholdet i påvirkningen deltagerne ble utsatt for (alle «tweets» var negative). En *Paired Samples T-test* viste at denne effekten var signifikant ( $p < .05$ ,  $t = -2.15$ ).

##### 4.2.3.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

Deretter ble det utført en *Mixed Design ANOVA* for å undersøke forskjellene mellom de tre eksperimentelle gruppene og kontrollgruppen. Det var til dels store forskjeller mellom gruppene (endring  $\times$  gruppe,  $F = 1.60$ ,  $Partial\ Eta^2 = .24$ ), men ingen statistisk signifikante ( $p = .23$ ). Det kan altså ikke konkluderes med noen effekt av primerne de eksperimentelle gruppene ble presentert for. Figuren under (Figur 4.4) viser at tendensen i gruppeforskjellene gikk i retning av at spill-primer gruppen var mest robust mot påvirkning.



Figur 4.4 Gruppeforskjeller i robusthet mot å bli påvirket, holdning 3, oppfatning av Forsvaret (negativ).  $N = 20$ .

#### 4.2.4 Holdning 4: Oppfatning av likestilling (negativ)

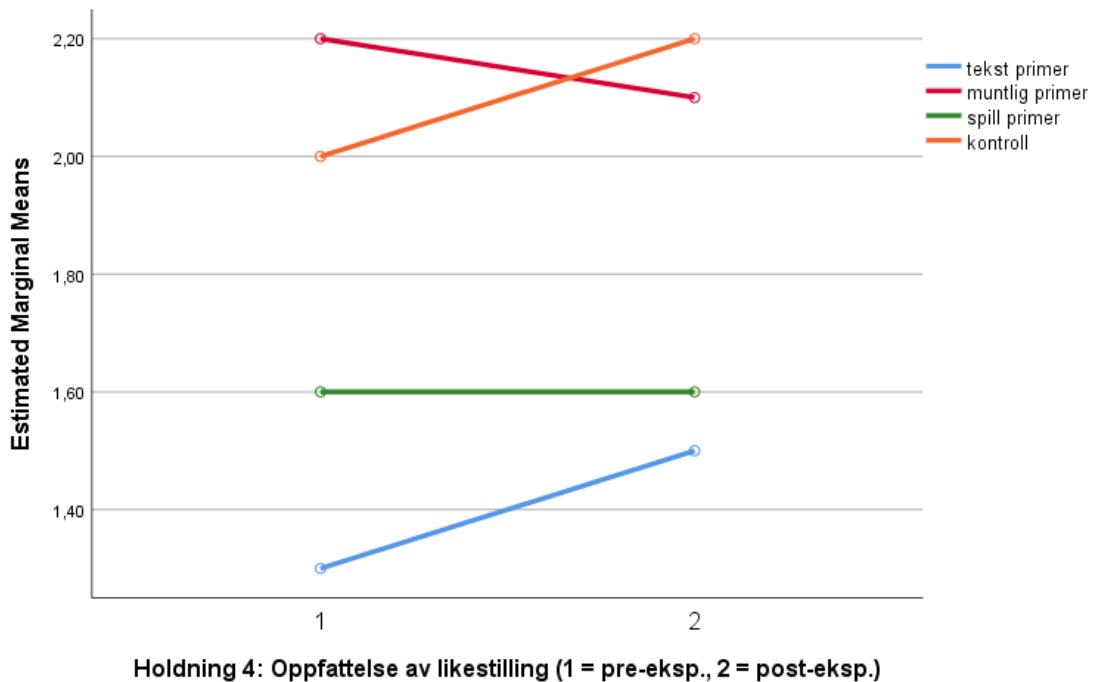
##### 4.2.4.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget et gjennomsnitt på  $M = 1.78$  før eksperimentell påvirkning og et gjennomsnitt på  $M = 1.85$  etter eksperimentell påvirkning (sprikende standpunkter i tweetsene, men primært usant og meninger). Endringen kan defineres som svært liten ( $d = .14$ ) og var i retning mer negativ oppfatning av likestilling etter eksperimentell påvirkning. En *Paired Samples T-test* viste at denne effekten ikke var signifikant ( $p = .53$ ,  $t = -0.65$ ), og kan dermed ha vært tilfeldig.

##### 4.2.4.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

Deretter ble det gjennom en *Mixed Design ANOVA* undersøkt om det var forskjeller mellom de tre eksperimentelle gruppene og kontrollgruppen. Det var ingen signifikante forskjeller mellom gruppene (endring  $\times$  gruppe,  $p = .77$ ,  $F = 0.77$ ,  $Partial \eta^2 = .07$ ) – altså ingen effekt av primerne de eksperimentelle gruppene ble presentert for. Figuren under (Figur 4.5) viser at tendensen i gruppeforskjellene gikk i retning av at muntlig- og spill-primer-gruppene var mest

robuste mot påvirkning, men her er forskjellene litt under middels og som nevnt ikke-signifikante.



Figur 4.5 Gruffeforskjeller i robusthet mot å bli påvirket, holdning 4, Oppfatning av likestilling (negativ).  $N = 20$ .

#### 4.2.5 Holdning 5: Oppfatning av at det er mye rasisme i Norge

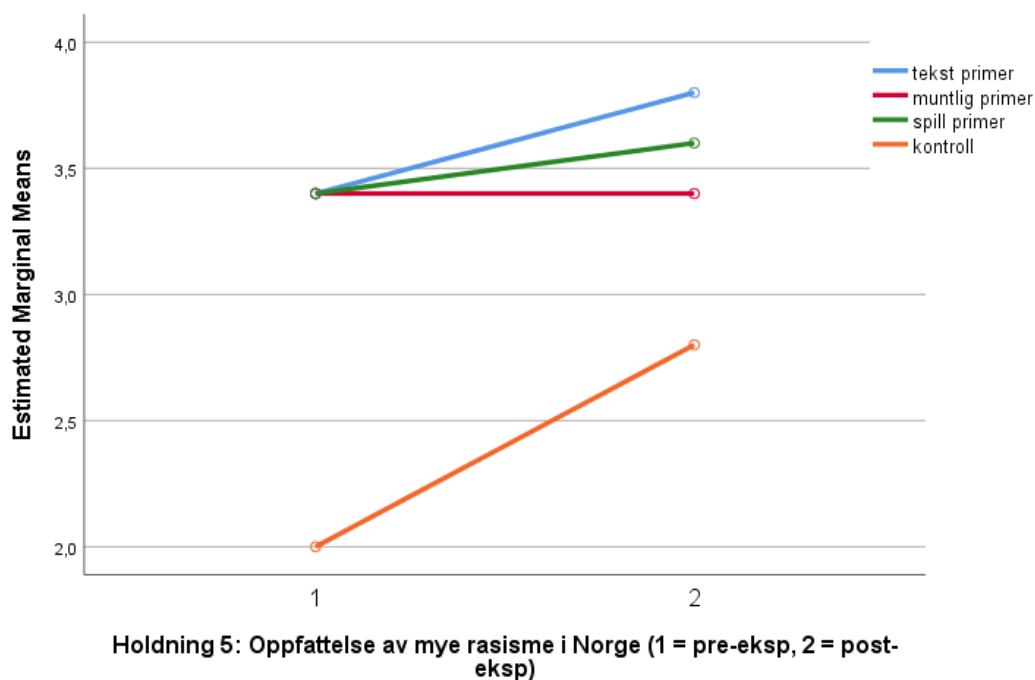
##### 4.2.5.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget et gjennomsnitt på  $M = 3.05$  før eksperimentell påvirkning og et gjennomsnitt på  $M = 3.40$  etter eksperimentell påvirkning. Endringen kan defineres som stor ( $d = .47$ ) og var i retning oppfatning av mer rasisme i Norge etter eksperimentell påvirkning. Innholdet i påvirkningen deltagerne ble utsatt for var både rasistisk og antirasistisk (dvs., blanding av sanne, usanne, menings- og gråsome tweets), med en liten overvekt av rasistisk (og usant) innhold. En *Paired Samples T-test* viste at denne endringen var signifikant ( $p < .05$ ,  $t = -2.10$ ).

##### 4.2.5.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

Gjennom en *Mixed Design ANOVA* ble det undersøkt om det var forskjeller mellom de tre eksperimentelle gruppene og kontrollgruppen. Det var en middels til stor forskjell mellom gruppene (endring  $\times$  gruppe,  $F = 1.06$ ,  $Partial \eta^2 = .17$ ), men den var ikke signifikant ( $p =$

.39). Det betyr at det ikke var noen påviselig effekt av primerne de eksperimentelle gruppene ble presentert for. Men som vist i Figur 4.6 var det en tendens til at oppfatningen til kontrollgruppen endret seg mest (dvs., oppfattet at det var mer rasisme i Norge) etter påvirkning (dvs., av tweets med både rasistisk og antirasistisk innhold), mens gruppen med muntlig primer ikke demonstrerte noen endringer (dvs., viste mest robusthet).



Figur 4.6 Gruppeforskjeller i robusthet mot å bli påvirket, holdning 5, Oppfatning av at det er mye rasisme i Norge.  $N = 20$ .

#### 4.2.6 Holdning 6: Oppfatning av president Biden (positiv)

##### 4.2.6.1 Endring fra pre- til post-eksperimentell påvirkning

Samlet hadde hele utvalget et gjennomsnitt på  $M = 3.65$  både før og etter eksperimentell påvirkning, altså ingen endring i alle gruppene samlet.

##### 4.2.6.2 Robusthet mot å bli påvirket (sammenligning av gruppene)

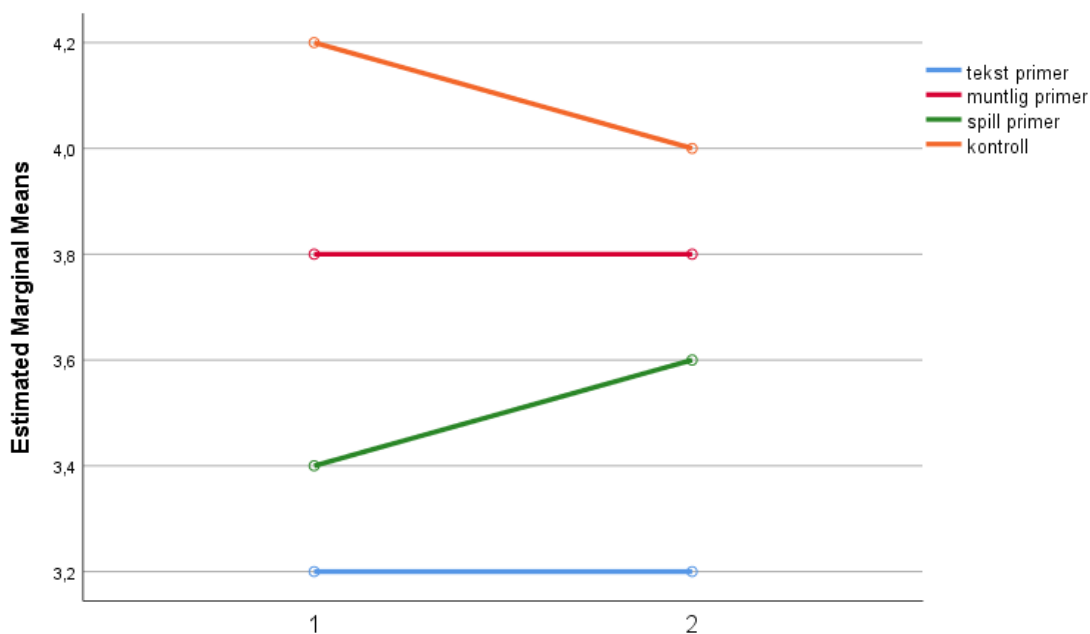
Deretter ble det gjennom en *Mixed Design ANOVA* sjekket om det var noen forskjeller mellom de tre eksperimentelle gruppene og kontrollgruppen. Her kom det frem at det var en medium til stor effekt ( $Partial\ \eta^2 = .20$ ), men ikke en signifikant forskjell mellom gruppene (endring  $\times$  gruppe,  $p = .30$ ,  $F = 1.33$ ). Det ble også kontrollert for utdanning (da denne kontrollvariabelen hadde vist en sammenheng med denne endringsvariabelen i korrelasjonsanalysen, Tabell 4.2);

---

---

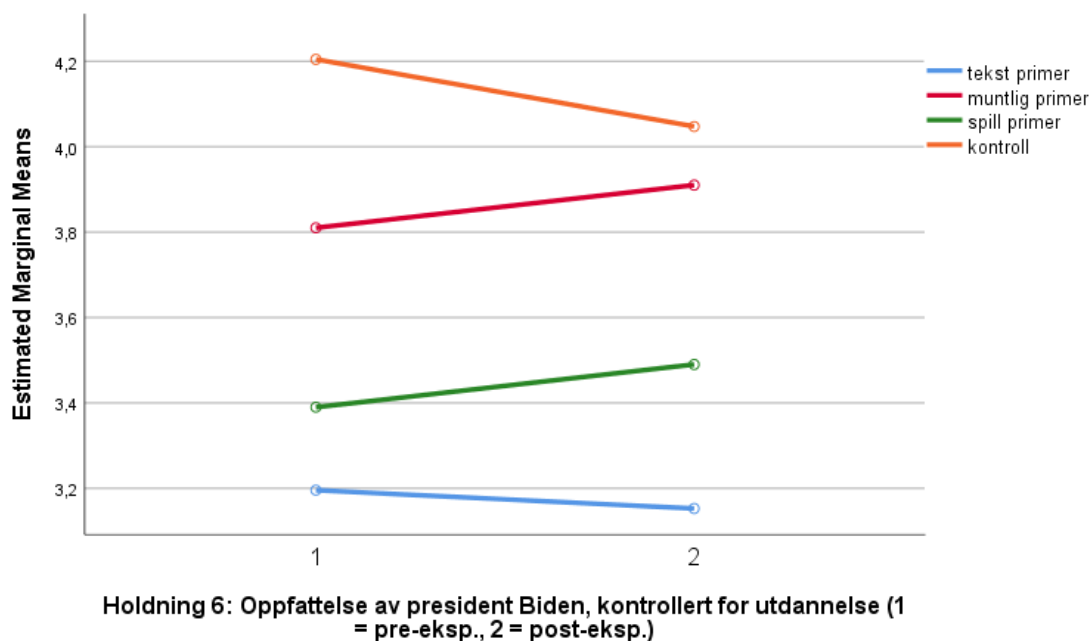
analysen viste en nær signifikant effekt av utdanning på endringen i oppfatning (endring  $\times$  utdanning,  $p = .09$ ,  $F = 3.26$ ,  $Partial\ Eta^2 = .18$ ). Ved å inkludere kontrollvariabelen, ble forskjellen mellom gruppene mindre og lengre fra signifikant (endring  $\times$  utdanning,  $p = .47$ ,  $F = 0.88$ ,  $Partial\ Eta^2 = .15$ ).

Figur 6 og 7 viser at tendensen i dataene var ingen eller en liten positiv endring i spill-, muntlig- og skriftlig primer gruppene; de eksperimentelle gruppene tenderte til å vise litt mer robusthet mot påvirkning sammenlignet med kontrollgruppen.



Holdning 6: Oppfattelse av president Biden (1 = pre-eksp, 2 = post-eksp)

Figur 4.7 Gruppeforskjeller i robusthet mot å bli påvirket, holdning 6, Oppfatning av president Biden (positiv).  $N = 20$ .



Figur 4.8 Gruffeforskjeller i robusthet mot å bli påvirket, holdning 6, Oppfatning av president Biden (positiv), kontrollert for utdannelse.  $N = 20$ .

### 4.3 Analyser av gruppeforskjeller i evne til å gjenkjenne sanne tweets, og robusthet mot å tro på usanne og gråsoner tweets

Det var 21 som målte hvordan deltagerne oppfattet sannhetsgehalten i utsagn hentet fra tweetsene de hadde fått lese: syv usanne, én gråsoner og tretten sanne tweets/påstander.

#### 4.3.1 ANOVA (enveis variansanalyse)

For å teste om det var en forskjell mellom de tre eksperimentelle gruppene og kontrollgruppen i forhold til evnen de demonstrerte til å gjenkjenne sanne, usanne og gråsoner nyheter og utsagn i form av tweets på sosiale medier ble det gjennomført enveis ANOVA som sammenlignet de tre eksperimentelle gruppene og kontrollgruppen. Disse analysene er presentert i Tabell 4.6, 4.7 og 4.8 nedenfor (høy skåre betyr at deltagerne i større grad trodde på usannheter, gråsoner og sanne tweets). Det ble i tillegg kontrollert for NFC, kjønn og utdannelse gjennom en ANCOVA, der korrelasjonsanalysene over (Tabell 4.3, 4.4 og 4.5 i kapittel 4.1.2) hadde indikert sammenhenger mellom kontrollvariablene og de ulike tweetsene. Detaljene i disse analysene er rapportert etter hver tabell i det følgende.



Tabell 4.6 Forskjeller i gjennomsnittsverdi mellom de tre eksperimentelle gruppene og kontrollgruppen med hensyn til oppfattelse av usanne tweets. En-veis ANOVA.  $N = 19$ .

Usanne tweets – spørsmål nr.	<i>F</i>	<i>p</i>	Primer	<i>M</i>	<i>SD</i>
# 1.	1.52	.25	tekst	1.20	.45
			muntlig	1.60	.55
			spill	1.20	.45
			kontroll	1.75	.50
# 2.	1.05	.40	tekst	3.20	1.10
			muntlig	3.80	1.30
			spill	3.40	1.34
			kontroll	4.00	.82
# 6.	.44	.73	tekst	2.40	.55
			muntlig	2.20	1.10
			spill	2.40	1.14
			kontroll	2.50	1.00
# 8.	.08	.97	tekst	3.00	.71
			muntlig	1.60	.89
			spill	3.20	1.10
			kontroll	2.75	1.71
# 11.	2.05	.15	tekst	1.60	1.34
			muntlig	2.00	1.23
			spill	1.80	.84
			kontroll	1.75	.50
# 12.	1.05	.40	tekst	4.20	.45
			muntlig	3.80	.45
			spill	3.80	.45
			kontroll	3.75	.50
# 17.	.06	.98	tekst	1.80	1.30
			muntlig	1.60	.89
			spill	1.60	.89
			kontroll	1.75	.50
Usant samlet (gjennomsnitt)	.41	.75	tekst	2.49	.40
			muntlig	2.37	.26
			spill	2.49	.33
			kontroll	2.61	.24

Notater. Høy skåre på gjennomsnitt (*M*) indikerer høy tro på at tweeten er sann. *SD* = standardavvik.

Tabell 4.6 indikerte ingen signifikante forskjeller mellom gruppene i oppfattelsen av enkelt-tweetsene eller samlet sett.

Det ble kontrollert for utdanning i forhold til usann tweet # 1 (fordi utdanning var signifikant korrelert med usann tweet #1, Tabell 4.3, og dermed kunne ha påvirket resultatene i ANOVA-

analysen), gjennom en ANCOVA med utdanning som kontrollvariabel. Det å kontrollere for utdanning endret ikke resultatene, det vil si at det ikke ble funnet signifikante forskjeller mellom gruppene ( $F = 1.66, p = .22$ ) med hensyn til usann tweet #1. Det betyr at forskjellene mellom gruppene ikke ble vesentlig endret ved å ta med denne kontrollvariabelen.

Det ble videre kontrollert for NFC i forhold til samlevARIABLEN for de usanne tweetsene (da NFC var nær å ha en signifikant korrelasjon med denne, Tabell 4.3). Det ble kjørt en ANCOVA med NFC som kontrollvariabel, hvor det ikke ble funnet signifikante forskjeller mellom gruppene ( $F = 1.00, p = .42$ ). Det vil si, kontroll for NFC endret ikke resultatene for samlevARIABLEN og forskjellene mellom gruppene ble ikke vesentlig endret ved å ta med denne kontrollvariabelen.

*Tabell 4.7* Forskjeller i gjennomsnittsverdi mellom de tre eksperimentelle gruppene og kontrollgruppen med hensyn til oppfattelse av gråsome tweets. En-veis ANOVA.  $N = 19$ .

Gråsome tweet – spørsmål nr.	<i>F</i>	<i>p</i>	Primer	<i>M</i>	<i>SD</i>
# 5.	1.02	.41	tekst	2.20	.89
			muntlig	2.20	.84
			spill	3.20	.84
			kontroll	2.75	1.71

*Notater.* Høy skåre på gjennomsnitt (*M*) indikerer høy tro på at tweeten er sann. *SD* = standardavvik.

Tabell 4.7 indikerte ingen signifikante forskjeller mellom gruppene i oppfattelsen av den ene gråsome-tweeten. Det ble videre kontrollert for utdanning i forhold til gråsome tweet # 5 (fordi korrelasjonsanalysene indikerte en nær signifikant korrelasjon; Tabell 4.4). Dette ble gjort gjennom en ANCOVA med utdanning som kontrollvariabel. Det å kontrollere for utdanning endret ikke resultatene; det ble ikke funnet signifikante forskjeller mellom gruppene ( $F = .57, p = .65$ ). Det betyr at forskjellene mellom gruppene ble ikke vesentlig endret ved å ta med denne kontrollvariabelen.

Tabell 4.8 Forskjeller i gjennomsnittsverdi mellom de tre eksperimentelle gruppene og kontrollgruppen med hensyn til oppfattelse av sanne tweets. En-veis ANOVA.  $N = 19$ .

Sanne tweets – spørsmål nr.	<i>F</i>	<i>p</i>	Primer	<i>M</i>	<i>SD</i>
# 3.	1.00	.42	tekst	4.60	.55
			muntlig	3.80	1.30
			spill	4.20	.84
			kontroll	3.50	1.29
# 4.	1.09	.38	tekst	2.40	.55
			muntlig	2.60	1.14
			spill	3.20	.45
# 7.	2.22	.13	kontroll	2.75	.50
			tekst	3.80	.45
			muntlig	3.40	1.14
# 9.	.31	.82	spill	3.00	.71
			kontroll	4.25	.50
			tekst	3.00	.71
# 10.	3.17	.055	muntlig	3.00	1.00
			spill	2.80	1.10
			kontroll	2.50	.58
			tekst	2.40	1.14
# 13.	1.32	.30	muntlig	3.40	.55
			spill	3.40	.55
			kontroll	4.00	.82
			tekst	2.60	1.14
# 14.	2.02	.16	muntlig	2.00	.71
			spill	1.60	.89
			kontroll	2.75	1.26
			tekst	3.00	.000
# 15.	1.09	.38	muntlig	3.40	1.14
			spill	3.20	.45
			kontroll	4.00	.00
			tekst	3.60	.55
# 16.	.56	.65	muntlig	3.60	1.34
			spill	2.80	.45
			kontroll	3.50	.58
			tekst	3.40	1.14
Sanne samlet (gjennomsnitt)	.50	.69	muntlig	2.80	.45
			spill	3.20	.45
			kontroll	3.00	.82
			tekst	3.20	.15
			muntlig	3.11	.73
			spill	3.04	.23
			kontroll	3.36	.06

Notater. Høy skåre på gjennomsnitt (*M*) indikerer høy tro på at tweeten er sann. *SD* = standardavvik.

---

---

Tabell 4.8 indikerte ingen signifikante forskjeller mellom gruppene i oppfattelsen av enkelt-tweetsene eller samlet sett.

Vi kontrollerte i tillegg for utdanning i forhold til sann tweet #4 og for kjønn i forhold til sann tweet #7 gjennom to ANCOVAer, da korrelasjonsanalysene hadde indikert at disse var signifikant korrelert (Tabell 4.5). Det å kontrollere for utdanning og kjønn endret ikke resultatene; det ble ikke funnet signifikante forskjeller mellom gruppene hverken relatert til tweet #4 ( $F = 1.22, p = .34$ ) eller #7 ( $F = 1.47, p = .27$ ). Med andre ord, forskjellene mellom gruppene ble ikke vesentlig endret ved å ta med disse kontrollvariablene.

#### **4.3.1.1 Resultatene fra ANOVA og ANCOVA oppsummert**

ANOVAene indikerte ingen signifikante forskjeller mellom gruppene knyttet til evne til å gjenkjenne usanne, gråsome og sanne tweets.

Men det var en tendens til at kontrollgruppen hadde noe høyere skårer enn de andre gruppene relatert til usanne tweets – det vil si at kontrollgruppen tenderte til å være dårligst til å kjenne igjen usanne tweets. Muntlig primer-gruppen hadde de laveste skårene – hvilket betyr at de tenderte til å være best til å kjenne igjen usanne tweets.

Spillprimer-gruppen hadde den høyeste skåren i forhold til å gjenkjenne en gråsome tweet. Deltagerne i spillprimer-gruppen tenderte altså til å tro mer på gråsome tweeten enn deltagerne i de andre gruppene. Deltagerne i tekst- og muntligprimer-gruppene skåret lavest – de tenderte altså til å tro minst på gråsome tweeten.

Kontrollgruppen tenderte til å ha de høyeste skårene også i forhold til å tro på sanne tweets. Spillprimer-gruppen skåret lavest – det vil si de trodde minst på de sanne tweetsene.

ANCOVA-analysene med kontrollvariablene inkludert endret ikke resultatene fra ANOVAene.

#### **4.3.2 Analyser av mål på konfidens (svarsikkerhet)**

Gjennomsnittsskåren på konfidens for usanne tweets samlet var  $M = 3.58$ , for sanne  $M = 2.66$  og for gråsome  $M = 3.11$ . Det betyr at deltagerne tenderte til å være mer usikre på evalueringene sine av de sanne tweetsene enn de usanne.

Det ble innledningsvis gjennomført en korrelasjonsanalyse for å undersøke eventuelle sammenhenger mellom kontrollvariablene og konfidens relatert til evalueringene av om tweetsene var sanne/usanne. Tendensen var negativ men ikke signifikant for NFC og konfidens til evalueringene av sanne tweets ( $r = -.17, p = .48$ ) og usanne tweets ( $r = -.25, p = .30$ ), men signifikant negativ med hensyn til evalueringen av gråsome-tweeten ( $r = -.46, p < .05$ ). Det betyr at tendensen var i retning av at de som skåret høy på NFC også følte seg mer usikre på sine evalueringer av sannhet. Det var ingen gjennomgående tendens eller signifikans i korrelasjonene mellom utdanning og konfidens, og kjønn og konfidens.

---

---

Det ble deretter gjennomført fire enveis ANOVAer for å undersøke om det var noen forskjeller mellom hvor sikre deltagerne var på sine evalueringer av om tweetsene var sanne/usanne og de ulike eksperimentelle gruppene. Det ble først gjort tre separate ANOVAer for gråsone, sanne og usanne tweets og til slutt en samlet for alle tre typer tweets. Det var ingen signifikante forskjeller. Men samlet hadde kontrollgruppen høyest skåre ( $M = 3.31$ ), tekst-primer gruppen hadde nest høyeste skåre ( $M = 3.13$ ) fulgt av spill-primer gruppen ( $M = 3.10$ ) og muntlig-primer gruppen ( $M = 2.96$ ) ( $F = .15, p = .93$ ). Selv om det understrekes at forskjellene er svært små og langt fra signifikante, kan det synes som at tendensen er i retning av at deltagerne som fikk informasjon gjennom en primer var mer usikre i sine evalueringer av sannhet.

En korrelasjonsanalyse indikerte ingen sammenheng mellom konfidens og tro på at tweetsene var sanne ( $r = -.23, p = .35$ ), men tendensen var i retning av at deltagerne tenderte til å være mer sikre i sin sak når de mente en tweet var usann.

#### **4.4 Kvalitative dataanalyser: fra observasjoner, gruppe- og fellesdiskusjonene**

Deltagerne uttrykte at de hadde fått nok tid på de ulike oppgavene og at de generelt var godt forståelige.

Noen bemerket imidlertid at det var vanskelig å forstå spørsmålene i sant/usant-delen før de fikk det forklart av FFI-representanten. De fleste deltagerne følte seg usikre på svarene. Noen indikerte at det var vanskelig å vite om de skulle evaluere egen eller generell oppfattelse av sannhet/usannhet og om det var på generelt grunnlag eller bare ut fra tweetsene de hadde lest.

Enkelte deltagere uttrykte at de opplevde det som uklart hvordan de kunne oppføre seg på Mastodon (var det OK å være nettroll, engasjere fullt og helt som normalt eller skulle de kun lese?). Noen sa også at det var «gøy å være nettroll når man kan være anonym», mens andre sa at de hadde lyst til å oppføre seg som nettroll, men at de lot være på grunn av instruksjonen om å «oppføre seg som de pleide».

Som svar på spørsmålet om deltagerne opplevde noe som rart og hvordan de opplevde sakene på Mastodon, var det mange som indikerte at meldingene på Mastodon var lite troverdige. Mange meldte videre at «feeden» var full av lite troverdige meldinger, konspirasjonsteorier, polariserende meldinger, emosjoner, over-tydelige meldinger, radikale og betente temaer og aggresjon. Men deltagerne uttrykte likevel at meldingene godt kunne vært reelle til tross for at de ikke var troverdige. Noen sa at de ville blokkert mange av brukerne i virkeligheten. Likevel syntes kommentarene og resultatene å indikere at det kunne være vanskelig for deltagerne å skille sant/usant etterpå. Andre fortalte også at de bare skrollet forbi saker.

Bruk av noen kjente memes og gamle saker, gjorde også at deltagerne sa at ingen ville dele eller tro på det. Ellers indikerte deltagerne at de syntes det var intenst, med mye info på kort tid, men at det også var morsomt.

---

---

Deltagerne informerte videre om at det oppsto en situasjon med en deltager som opprettet en fiktiv bruker der hen utga seg for å være en kjent FFI-forsker med lite normative meninger. Denne var laget av en deltager og ble tatt ned av «spillkommandoen» da den ble oppdaget.

Deltagerne sa også at kontoer med navn som lignet troverdige kontoer (f.eks., «NKR»), ble oppfattet som litt vanskelige å bestemme troverdigheten av. De ble i tvil om det var skrivefeil eller bevisst villedning.

På spørsmål om deltagerne syntes informasjonen de fikk i forkant påvirket hvordan de oppfattet sakene på Mastodon, svarte deltagerne i de eksperimentelle gruppene som hadde fått informasjon i forkant, enten det var muntlig, skriftlig eller gjennom spill (dvs., «primer»-gruppene), konfirmerende. De forklarte at de ble litt mer bevisst. Men én deltager mente også at hen var ganske kritisk fra før, slik at effekten ikke var så stor. Deltagerne fra spillprimergruppen kommenterte også at de ble mer observante på trolling, overskrifter og nøkkelord for å få likes og oppmerksomhet, «character assassination» for lettere å «kaste skit» og reaksjon på kritikk. Deltagerne fra kontrollgruppen utrykte at de ikke opplevde noen effekt av saken de fikk lese og høre om, og utrykte at de opplevde den som irrelevant.

I den avsluttende diskusjonen etter debrief på Microsoft Teams ble det også bedt om spørsmål og tilbakemeldinger fra deltagerne. Her kom det ikke frem noe særlig nytt utover det som kom frem i gruppediskusjonene. Det ble også tydelig at det var langt vanskeligere å få deltagerne i tale på Microsoft Teams enn direkte i gruppene.

## 5 Diskusjon

Dette kapitlet gir en oppsummering og diskusjon av resultatene presentert i kapittel 4, og inkluderer begrensninger, videre forskning, etiske utfordringer og konklusjon. Den første delen tar for seg en oppsummering av resultatene fra analysene og hvordan de svarer på hypotesene/forskningsspørsmålene.

### 5.1 NFC og de andre kontrollvariablene

Basert på hypotesen om at lav NFC gir sårbarhet og høy NFC gir robusthet mot påvirkning, ble NFC hos deltagerne målt. Dette ble gjort både for å teste denne hypotesen i seg selv, men også for å kontrollere for at NFC-nivået hos de enkelte deltagerne ikke påvirket resultatene av eksperimentet.

NFC er et veletablert mål, men med ikke-akseptable reliabilitetsmål var det nødvendig å justere målet ved kutte de indikatorene som ikke syntes å måle det samme. Det kan være at en

---

---

kombinasjon av utvalget<sup>15</sup> og tiden som har gått siden målet ble utviklet, har redusert spørsmålenes egnethet og dermed reliabiliteten. Resultatet ble et NFC-mål med 12 indikatorer (istedenfor de originale 18) og en akseptabel reliabilitet. Gjennomsnittet på NFC for utvalget var relativt høyt ( $M = 3.86$ ; for en sammenligning se f.eks., Kuvaas & Kaufmann, 2004). Det viste seg også at det var lite varians i NFC mellom deltagerne, noe som gjorde det lite sannsynlig at forskjeller i NFC ville påvirke hovedanalysene.

Som forventet indikerte de innledende korrelasjonsanalysene ingen sammenheng mellom NFC og kjønn, men det ble funnet en marginalt signifikant sammenheng mellom NFC og utdanning. Sammenhengen indikert i sistnevnte er i tråd med funn fra tidligere forskning (f.eks., Cacioppo et al., 1996; Liu & Eveland, 2005). Lav varians i NFC-skårene forklarer at det ikke ble funnet noen sammenheng mellom NFC og endring i oppfatning av saker eller mellom NFC og mistro til usanne tweets.

De innledende korrelasjonsanalysene indikerte at det i tillegg var grunnlag for å kontrollere for utdanning i analysene av endring i oppfatning av saker (dvs. holdning 2 og 6) og i oppfattelse av én sann, én gråsone og én usann tweet, og kjønn i forhold til én sann tweet. Disse analysene indikerte imidlertid ingen effekt av kontrollvariablene på resultatene.

## 5.2 Endring i oppfatninger av saker og videre forskning

Holdning 2 (oppfatning av konflikt by-land), 3 (negativ oppfatning av Forsvaret) og 5 (oppfatning av at det er mye rasisme) viste signifikante endringer i størrelsesorden medium til stor fra pre- til post-eksperimentelle målinger. Det vil si, etter å ha lest alle de fiktive tweetsene tenderte deltagerne til å mene at det var mer konflikt mellom by og land i Norge, ha en mer negativ oppfatning av Forsvaret og mene at det er mer rasisme i Norge sammenlignet med før de leste tweetsene. For holdning 2 og 3 gikk endringen i samme retning som tweetsene de fikk lese om disse sakene. Hva gjelder holdning 5 var påvirkningen fra tweetsene mindre samstemt, med både rasistiske utsagn og antirasistiske utsagn, men hvor fokuset på temaet kunne få deltagerne til å mene at det var mer rasisme enn de hadde tenkt før påvirkningen.

Tendensen (ikke signifikant) til holdning 1 var en liten endring i retning av hva alle tweetsene om temaet uttrykte, det vil si mindre tillit til vaksiner og myndigheter. Tendensen (ikke signifikant) til holdning 4, oppfatning av likestilling, var en liten endring i negativ retning, selv om tweetsene om temaet uttrykte meninger/informasjon i både positiv og negativ retning (primært meninger og usanne utsagn). Samlet var det ingen endring på holdning 6 (positiv oppfatning av president Biden), men her indikerte Mixed Design ANOVAene at to av gruppene tenderte til å endre seg i to ulike retninger; kontrollgruppen i negativ retning i tråd med tweets-innholdet, spill-gruppen i motsatt retning. Den lille positive endringen i spill-gruppen kan tolkes som en motreaksjon på et påvirkningsforsøk i negativ retning. Fenomenet er beskrevet i

---

<sup>15</sup> Kuvaas & Kaufmann (2004) fikk oppnådde dog en akseptabel reliabilitet med et norsk utvalg og med samme skala og utregningsmetode.

---

tidligere forskning (f.eks., Schlosser & Shavitt, 2009), hvor et gjennomskuet forsøk på påvirkning kan gi motsatt effekt.

*Oppsummert tenderte fem av seks oppfatninger av saker å ha endret seg i størrelsesorden litt til mye fra pre- til post-eksperimentelle målinger. Tre av disse endringene i oppfatning var signifikante, hvilket indikerer at den eksperimentelle påvirkningen i form av tweets syntes å ha hatt en viss effekt.*

For videre forskning og eksperimenter anbefales å kun bruke tweets som går i én retning relatert til en gitt sak – for å lettere kunne måle og tolke resultatene med hensyn til om det var noen påvirkning.

### **5.3 Robusthet mot endring i oppfatninger av saker og videre forskning**

Den andre delen av resultatene relatert til oppfatningsendring skulle svare på hypotesen om motstandsdyktigheten mot påvirkning av oppfatninger av saker på sosiale medier kunne økes ved bruk av muntlig eller skriftlig forskningsbasert informasjon eller et vaksine- («inoculation») spill som primer; var det en eksperimentell betingelse (primer) som var målbart mer effektiv mot påvirkning enn de andre? Og/eller var det observert noen gjennomgående trend i effekten av de ulike primerne samlet for de ulike målene på oppfatning?

Det ble gjennomført en samlet kvalitativ evaluering av figurene fra Mixed Design ANOVAene og en ren optelling av hvilke eksperimentelle grupper som syntes å ha hatt minst påvirkning, for å se om det var noen gjennomgående trend i dataene som kunne indikerte om noen grupper syntes å være mer robuste enn andre.

Gruppene med enten muntlig eller spill-primer viste tendenser til robusthet på 5 av 6 oppfatningsmål, gruppen som fikk lese skriftlig primer viste tendenser til robusthet i 2 av 6 oppfatningsmål, mens kontrollgruppen ikke viste noen tendenser til robusthet i forhold til noen av oppfatningsmålene. Den samlede tendensen syntes å være at kontrollgruppen ble noe mer påvirket i sine oppfatninger enn de eksperimentelle gruppene som fikk høynet sitt kompetansenivå med hensyn til påvirkning, falske nyheter og desinformasjon. Det understrekes at disse resultatene må tolkes med forsiktighet og kun som tendenser. Tendensene tolkes først og fremst som en indikasjon på at det er grunnlag for å gjøre flere eksperimenter med nye og større utvalg. Ett av disse resultatene var dog signifikant og kan festes mer lit til; gruppen som fikk muntlig informasjon viste seg mer robust fra pre- til post-eksperiment enn de andre gruppene med hensyn til endring av holdning 2. Men andre ord syntes denne gruppen å demonstrere mest robusthet knyttet til oppfatning av konflikt by-land.



---

---

Tweetsene deltagerne ble presentert for var gjennomgående lite troverdige<sup>16</sup>. Det anses som interessant at man til tross for dette kunne se tendenser i dataene til at de som fikk informasjon<sup>17</sup> i forkant, syntes å vise tendenser til mer robusthet knyttet til endring i oppfatning.

#### **5.4 Forskjeller i evne til å skille sanne, usanne og gråsoner tweets og videre forskning**

ANOVAene indikerte ingen signifikante forskjeller mellom gruppene relatert til evne til å gjenkjenne hva som var usanne, gråsoner og sanne tweets. Det var altså ingen klare svar på hypotesen om at man kan høyne evne til kjenne igjen sanne nyheter og utsagn og robusthet mot å tro på usanne på sosiale medier ved bruk av forskningsbasert informasjon eller ved et vaksinespill.

Tendensene var i retning av at deltagerne i kontrollgruppen syntes å tro mest på alle tweetsene – både de sanne og usanne. De i muntlig- og spill-primergruppene derimot, syntes å være mest skeptiske – både til de sanne og usanne tweetsene. Men forskjellene her er små og langt fra signifikante. Som tidligere nevnt var utvalget så lite at det er vanskelig å oppnå signifikante resultater. Tendensene kan dog tolkes som at det er grunnlag for å videreføre forskningen med justeringer i nye eksperimenter med andre og større utvalg.

Majoriteten av tweetsene under denne pilotkjøringen fremsto generelt som lite troverdige – både sanne, gråsoner og usanne (mer om begrensninger i kapittel 5.7). Dette kunne derfor også være en årsak til, i tillegg til størrelsen på utvalget, at analysene viste liten forskjell på gruppene. På bakgrunn av at tweetsene fremsto som lite troverdige kan det tolkes som et positivt resultat at kontrollgruppen var de som tenderte til å tro mest på alt – både sant og usant. Det kan synes som at primerne gjorde deltagerne mer skeptiske generelt – og da spesielt den muntlige- og spillprimeren.

#### **5.5 Konfidens (svarsikkerhet) og videre forskning**

Forskjellene i svarene på konfidens var svært små og ikke signifikante, men det kunne synes som at tendensen var i retning av at deltagerne som fikk informasjon gjennom en primer var mer usikre i sine evalueringer av sanne, usanne og gråsoner tweets. Denne tendensen var motsatt av hva Roozenbeek og van der Linden (Roozenbeek & van der Linden, 2019) og Basol og kolleger (Basol et al., 2020; Basol et al., 2021) fant i sine analyser. En mulig grunn til denne forskjellen kan være at det i måleinstrumentet under eksperimentene av Roozenbeek og kolleger kun ble brukt utsagn/nyheter som brukte de spesifikke teknikkene for desinformasjon som de ble trent på i spillet. Basol og kolleger fant heller ingen endring i konfidens relatert til faktiske nyheter. Som indikert innledningsvis (kap. 2.4.5.1) viser også forskning på overkonfidens at de som har

---

<sup>16</sup> Dvs. dårlig staving/ortografi, mange utropstegn, store bokstaver, lite godt språk generelt og tydelige meninger. Som tidligere beskrevet ble de mer troverdige og gråsoner-tweetsene ved en feil i gjennomføringen, ikke sendt ut.

<sup>17</sup> Dvs. enten skriftlig, muntlig eller informasjon gjennom spill.

---

---

dårligst evner gjerne har for stor tro på egne evner (f.eks., Miller & Geraci, 2011), og at konfidens derfor ikke nødvendigvis er udelt positivt.

En annen mulighet til forskjellen i denne studiens funn sammenlignet med Roozenbeek og kolleger, er at tweetsene benyttet i denne studien generelt ble oppfattet å være mindre troverdige. Det var også slik at denne studien testet troen på tweets deltagerne hadde vært eksponert for i en simulert twitter-situasjon med kontrollgruppe, mens det i de initiale eksperimentene av Roozenbeek og van der Linden kun ble gjort før- og ettermålinger uten kontrollgruppe. Dette ble imidlertid forbedret i oppfølgerstudiene av Basol og kolleger. En tilleggsforskjell er at Basol og kolleger målte hvor sikre deltagerne var på sin «rating» av hvor manipulative (Basol et al., 2021) og hvor reliable (Basol et al., 2020) utsagnene var; ikke hvor sanne deltagerne mente utsagnene var. Denne diskrepansen med tidligere forskning blir viktig å følge opp i fremtidige full-skala eksperimenter.

Det var også en tendens i dataene til at deltagerne var mer sikre i sin sak når de mente en tweet var usann. Dette kan skyldes at tweetsene generelt fremsto som lite troverdige.

## **5.6 Resultatene fra oppfatninger av saker og sant/usant-analysene samlet og videre forskning**

Tendensene i analysene av forskjeller mellom de eksperimentelle betingelsene i forhold til oppfatningsendring og gjenkjenning av sant/usant syntes å peke i samme retning; deltagerne i den muntlige (og til dels spill-) primergruppen tenderte til å vise mest robusthet mot påvirkning av oppfatning av saker og skepsis mot usanne tweets, mens deltagerne i kontrollgruppen tenderte til å vise minst robusthet. Det synes meningsfylt at de gruppene som stolte minst på informasjonen i tweetsene også var de som lot seg påvirke minst i sine oppfatninger.

Dette gir grunnlag for å følge opp forskningen med utvidede eksperimenter med andre og større utvalg, men med oppdaterte tweets (dvs., inkludere tweets med høyere troverdighet, inkludert bedre oppbyggede profiler og nettverk for å oppnå dette, og flere gråsone-tweets), kontoer/profiler og innlegg – også på andre sosiale medie-kloner i Somulatoren.

## **5.7 Kvalitative analyser og videre forskning**

Noen deltagere uttrykte under gruppe- og fellesdiskusjonene at det var vanskelig å forstå spørsmålene i sant/usant-delen av post-eksperiment spørreskjemaet. På bakgrunn av dette bør det i fremtidige eksperimenter presiseres bedre i spørreskjemaet at deltagerne skulle svare ut fra egne oppfatninger og at det var greit å være usikker da det ikke var noen riktige eller gale svar. Spørreskjemaet inneholdt dog denne informasjonen (Se vedlegg A.3.2): «Dette spørreskjemaet har som formål å fange opp dine oppfatninger og evalueringer av sakene som du fikk lese på Mastodon (twitterklonen). Det er frivillig å svare. Vi understreker at det er ingen riktige eller gale svar. [...]», og videre «Klikk på det svaret som stemmer best med din oppfatning av følgende saker – på en skala fra helt uenig til helt enig. Det er ingen riktige eller gale svar. Kun

---

---

ett svar-klikk for hvert spørsmål.»). I tillegg ble dette kommunisert muntlig av FFI-forskeren (se vedlegg A.2, eksperimentmanualen: «klikk på det svaret som passer best for deg», «det er ingen riktige eller gale svar» og «er det vanskelig: helt ok å gjette»). Målsettingen må derfor være å gjøre veiledningene enda klarere, ved for eksempel å fremheve skriften, slik at alle deltagerne får med seg denne informasjonen i introduksjonen til spørreskjemaet.

Noen deltagere opplevde det som uklart hva de kunne gjøre på Mastodon. Derfor bør det i fremtidige eksperimenter presiseres at deltagerne kan gjøre det de vanligvis ville gjort på Twitter. Likevel kan det være en utfordring at deltagerne er anonyme, fordi anonymitet kan invitere til mer grenseoverskridende adferd. Dette er et kjent funn i sosialpsykologisk forskning (f.eks., Hewstone, Stroebe, & Stephenson, 1996; Matheson & Zanna, 1989); med anonymitet faller hemningene og etterfølgelse av sosiale normer. Dette er et interessant funn i seg selv. Også gruppene som ikke hadde spilt Harmony Square (som til dels er en opplæring i trolling og andre påvirkningsteknikker) rapporterte dette.

På bakgrunn av noe problemer med innlogging for noen av deltagerne bør man i fremtidig bruk av Mastodon i eksperimenter gjøre innloggingen enklere – det vil si, mer automatisk.

Kommentarene om at sakene på Mastodon opplevdes som lite troverdige, indikerer at det i fremtidige eksperimenter også bør inkluderes mer seriøse, troverdige og mindre tydelige meldinger. Det understrekes at deltageres oppfattelse av meldingene var i tråd med hvordan vi som forskere fra FFI også oppfattet dem. For ordens skyld var det også planlagt utsendelse av mer troverdige og mer gråsome type tweets. Disse tweetsene ble, som tidligere beskrevet, ikke sendt ut, grunnet en menneskelig feil under gjennomføringen av eksperimentet.

De kvalitative dataene illustrerte også et vanlig problem på sosiale medier og på nett generelt; kontoer med navn som ligner troverdige kontoer oppleves som villedende. Å låne troverdighet fra seriøse kontoer, benyttes ofte av useriøse aktører som ønsker å bli oppfattet som seriøse.

Deltagerne uttrykte at de syntes informasjonen de fikk i forkant påvirket hvordan de oppfattet sakene på Mastodon. På denne måten støttet de kvalitative tilbakemeldingene tolkningen av resultatene fra de kvantitative analysene; primerne (den eksperimentelle betingelsen) med forhåndsinformasjon syntes å ha hatt en preventiv effekt på påvirkning fra tweetsene. De kvalitative tilbakemeldingene støttet også hypotesen om at informasjonen kontrollgruppen fikk var irrelevant og ikke hadde noen effekt på oppfatningen av tweetsene de fikk lese i etterkant.

Det var tydelig at det var langt vanskeligere å få deltagerne i tale på Microsoft Teams enn direkte i gruppene, noe som understøtter nytten av å samle inn kvalitative data i mindre grupper og ikke gjennom en digital samarbeidsplattform.

De kvalitative dataene understøttet de kvantitative med hensyn til at deltagerne selv også oppfattet at primerne hadde en positiv effekt på deres evne til å oppfatte sant/usant og ikke la seg påvirke ved generelt å gjøre de mer skeptiske. Konklusjonen om å benytte mer troverdige profiler og meldinger i fremtidige eksperimenter er også i tråd de kvalitative tilbakemeldingene. Til slutt understøtter de kvalitative dataene planene om å gjøre innlogging og lignende mer

---

---

automatisert, gjøre presiseringer relatert til at deltagerne skal oppføre seg som de vanligvis gjør når de er på Twitter/de sosiale mediene, og hvordan spørsmålene i siste del av spørreskjemaet skal forstås (sant/usant-delen).

## **5.8 Begrensninger og videre forskning**

### **5.8.1 Utvalgsstørrelsen**

Denne studien besto av et piloteksperiment med relativt få deltagere ( $N = 20$ ). Det var likevel ansett som formålstjenlig å gjennomføre statistiske analyser med de kvantitative dataene. Fordi selv om man måtte påregne utfordringer med å finne statistisk signifikante resultater, er resultatene tolket å gi viktige indikasjoner på verdien av å forfølge og videreutvikle denne eksperimentaktiviteten. Funnene var overordnet ikke-signifikante, men lovende, og inkluderte også noen signifikante resultater med middels til store effekter.

### **5.8.2 Effekter over tid**

Pilot-eksperimentet presentert i denne rapporten tok ikke for seg effekter av primerne over tid. Roozenbeek, van der Linden og kolleger referert til i denne rapporten (f.eks., kap. 2.4.5), har derimot studert foreldelseeffekten relatert til deres bruk av spill. Ut fra deres funn, må man gå ut fra at tiltak for å øke robustheten mot påvirkningsoperasjoner også gjennom informasjon, opplæring og lignende er «ferskvare». Man må altså påregne oppfriskning og gjentakelser av budskap og informasjon til alle deler av befolkningen som man ønsker å gjøre mer motstandsdyktig. I dette ligger også en tilrettelegging av budskap slik at det blir tilgjengelig for ulike grupper. Her er det lite forskning og det er behov for en bedre forståelse av hvordan en slik tilrettelegging best kan gjøres.

### **5.8.3 Tweetsene**

Alle tweetsene var fiktive med varierende grad av sannhet både i selve meldingen og i eventuelle nyheter den henviste til. Klassifiseringen av noen tweets som sanne/usanne kan diskuteres. Dette er generelt en utfordring, da det kan være mye uenighet rundt en del påstander – om de er sanne eller usanne. I en del tilfeller kan selve utsagnet være basert på en falsk nyhet, men kan likevel være helt eller delvis sann, og motsatt. Dette er en utfordring videre forskning bør ta med seg.

Det kan forstås å være et problem at det kun var med én tweet av typen gråsoner (blanding av sant, usant og meninger), da resultatene knyttet til denne ene kan ha vært litt tilfeldig. Det anses å være behov for å ha med flere gråsoner tweets i fremtidige eksperimenter.

En annen utfordring med tweetsene som ble benyttet i dette eksperimentet er at de generelt kunne bli oppfattet som lite troverdige i form og tone. De kvalitative dataene presentert over (kapittel 5.5) understøtter en slik tolkning. Som også nevnt i kapittel 5.2, kan denne svakheten ha påvirket resultatene. Fordi de sanne tweetsene også var lite troverdige i sin form, ble også

---

---

disse oppfattet som lite troverdige. Dette behøver ikke være uheldig; kanskje er det formålstjenlig å mistro alle tweets med lav troverdighet, uansett om de er sanne eller usanne (dvs., bedre å bli generelt mer skeptisk enn forbli generelt naiv). Det kan være at en generell skepsis til det som ser lite troverdig ut er prisen å betale for ikke å tro på påvirkningsoperasjoner med falske nyheter, tweets og annen desinformasjon.

Det var videre slik at gruppen med tweets med en mer seriøs fremtoning (inkludert flere gråsome tweets) aldri ble sendt ut under eksperimentet. Dette var en menneskelig feil som kan usannsynliggjøres gjennom å automatisere utsendelsen av tweets, slik at alt kan settes opp på forhånd. På bakgrunn av kommentarene om at sakene på Mastodon opplevdes som lite troverdige, er det en målsetting for videre eksperimenter at det også inkluderes mer seriøse, troverdige og gråsome meldinger og mindre av de lite troverdige meldingene.

#### **5.8.4 Primerne (de eksperimentelle betingelsene)**

Resultatene tolkes som at primerne som ble brukt i piloteksperimentet har potensial til å øke robustheten mot påvirkning på sosiale medier. Likevel anbefales det en gjennomgang, evaluering og en eventuell videre forbedring av informasjonen og formidlingen av dette materialet.

Spillet Harmony Square kan kritiseres for at det ikke bare lærer opp deltagerne i hvordan de kan kjenne igjen påvirkningsteknikker og trolling, men også i hvordan de kan bruke teknikkene selv. Spillet gir også en belønning for den negative adferden, som kan inspirere spillerne til å ville prøve teknikkene selv. Det er mulig at man ikke kan sammenligne deltageradferden i Mastodon direkte med deltagerens vanlige adferd på Twitter da alt var anonymisert i eksperimentene. Likevel er dette en sannsynlig bieffekt for noen mennesker av denne typen opplæring, som ennå ikke har blitt undersøkt. Det er derfor tilrådelig med en grundig undersøkelse av slike bieffekter i et representativt utvalg før en eventuelt benytter slike spill i en opplysningskampanje eller generell opplæring rettet mot hele eller deler av befolkningen. Dette for å unngå utilsiktede effekter og negative overraskelser.

En mulighet er også å lære av hvordan krigsspill brukes i militær utdanning og øvelse. Forskning på bruk av krigsspill i opplæring peker blant annet på viktigheten av at deltagerne går gjennom en debrief med grundig refleksjon og diskusjon etter spillet for å få den forståelsen og læringseffekten man ønsker og unngå negativ læring (se f.eks., Alme, & Hvidsten, 2022; Wong, 2019).

#### **5.8.5 Simulatoren (teknologien for å simulere sosiale medier)**

Simulatoren, med bruk av den sosiale medie-klonen Mastodon til å simulere Twitter, fungerte overordnet fint og etter sin hensikt i pilot-eksperimentet rapportert her. Men den hadde også noen begrensninger. Det var flere deltagere som opplevde problemer med innlogging. Dette skyldes både at innloggingsprosedyren var litt omstendelig og at det ikke fungerte helt likt på alle systemer. Denne er nå forbedret til fremtidige eksperimenter (Bergh, 2023).

---

---

Det er også et spørsmål om deltagerne bør få se hva de andre deltagerne gjør og skriver på Mastodon, slik tilfellet var nå. Dette gir en naturlig kontekst, men gjør at påvirkningen fra meldingene deltagerne får lese ikke er kontrollert. Dette kan også gi etiske utfordringer da man ikke har kontroll på hva deltagerne skriver og legger ut. Forskerne vil dermed kunne stå ansvarlige for eventuelle overtramp som gjøres av enkelte deltagere ovenfor de andre. Dette bør derfor evalueres i fremtidige eksperimenter.

### **5.8.6 Etiske utfordringer**

Svarene på spørsmålene knyttet til oppfatning av sakene presentert på Mastodon før og etter at deltagerne hadde lest tweetsene indikerte noen endringer. Disse var ikke større enn at det er sannsynlig at de ikke representerte vedvarende endringer. Som indikert i metodekapittelet, ble det også gitt en debrief for å motvirke eventuelle påvirkninger av oppfatninger eller annet fra eksperimentet. Likevel er dette, som tidligere diskutert et etisk dilemma; behovet for å indusere en viss form for påvirkning i eksperimentsammenheng for å kunne studere, utvikle, teste og sette inn mottiltak mot påvirkning. Det er derfor ytterst viktig at man i videreføringen av denne forskningsaktiviteten hele tiden evaluerer dette nøye og at det blir gitt en grundig debrief til deltagerne. Da er det vår forståelse at ulempen dette eventuelt måtte medføre, er klart mindre enn utbyttet – også for deltagerne selv (i form av at deltagerne fikk økt forståelse og årvåkenhet ift. påvirkning de blir utsatt for i deres egen hverdag).

Som indikert over (kap. 5.7.4) kan spillet Harmony Square også representere et etisk dilemma fordi det lærer opp spillerne i å bruke påvirkningsteknikker og trolling selv. I tillegg inneholder spillet en belønning for benyttelse av disse teknikkene, hvilket kan beskrives som klassisk betingning (f.eks., Bitterman, 2006; Pavlov, 1927) av en adferd som ikke er ønskelig. Spillet gir med andre ord spillerne en god følelse ved ugrei adferd, hvilket kan anspre til mer ugrei adferd – for noen også utenfor spillets rammer.

## **6 Konklusjon, veien videre og noen anbefalinger**

Vi mennesker er «programmert» til lett å kunne påvirkes i en digitalisert hverdag og gjennom sosiale medier – noe som kan bli utnyttet av både nasjonalstater eller andre entiteter som ønsker å påvirke Norges befolkning, utvalgte grupper eller organisasjoner. De psykologiske sårbarhetene kan derimot også benyttes for å gjøre hele eller deler av befolkningen mer robust mot uønskede påvirkningsforsøk, for eksempel ved å være i forkant og gi informasjon som kan øke menneskers evne til gjenkjennelse av påvirkningsforsøk. Dette har vært hovedhypotesen pilot-eksperimentet beskrevet i denne rapporten har søkt å svare på. Forskningen ved FFI er et bidrag til arbeidet med å øke samfunnets og totalforsvarsaktørens evne til å motstå fiendtlig påvirkning gjennom sosiale medier.

---

---

Resultatene fra de kvantitative analysene var overordnet ikke-signifikante, men lovende og med noen signifikante funn med middels til store effekter. Tendensene i funnene pekte i samme retning:

1. Det ble funnet målbare endringer i oppfatninger av saker som følge av en kortvarig påvirkning gjennom en Twitter-klone (Mastodon). Til tross for lav utvalgsstørrelse var halvparten av disse oppfatningsendringene signifikante.

2. Endringene i oppfatninger av saker tenderte til å være størst for kontrollgruppen og minst for muntlig og spill-primergруппene. Og selv om forskjellene mellom gruppene var overveiende ikke-signifikante (én signifikant), tolkes det som at den samlede tendensen er lovende med hensyn til typen primere testet ut her. Det vil si at forskningsbasert muntlig, skriftlig og spill-type informasjon, kan gi en gunstig robusthetseffekt mot påvirkning på sosiale medier, og at slike tiltak derfor vil være gunstig å undersøke nærmere i fremtidige studier.

3. Som med oppfatningsendringene syntes deltagerne i muntlig og spill-primergруппene å vise tendens til noe mer robusthet mot å tro på usanne tweets. Her skal det poengteres at den samlede tendensen var at deltagerne i de eksperimentelle gruppene generelt syntes å være mer skeptiske til alle tweets – ikke bare de usanne. Dette kan tolkes som en indikasjon på at den generelle skepsisen var forhøyet, snarere enn evnen til å skille sant fra usant. Likevel kan effekten synes å være konstruktiv i denne konteksten; tweetsene bar preg av å være litt lite troverdige (dårlig språk, skrivefeil, store bokstaver, utropstegn, tydelige meningsutsagn, selvmotsigelser, osv.).

Deltagerne i muntlig og spill-primergруппene syntes å stole minst på informasjonen i tweetsene samtidig som de også tenderte til å vise mest robusthet mot å bli påvirket i sine oppfatninger av saker. Dette ble tolket som meningsfullt. Deltagerne i kontrollgruppen syntes å tro mest på tweetsene og viste samtidig minst robusthet mot påvirkning av sine oppfatninger. Kvalitative data fra gruppesamtaler med deltagerne understøttet de kvantitative tendensene og funnene. Begrensninger inkluderte lite utvalg og at meldingene benyttet i piloteksperimentet generelt var i en lite troverdig form og tone.

Samlet tolkes resultatene og erfaringene som lovende for fremtidige eksperimenter i større målestokk med utvalg fra flere populasjoner (inkludert Forsvaret og andre aktører i totalforsvaret). Anbefalinger for slike gjennomføringer inkluderer fokus på færre, større og bedre oppbyggede saker, flere troverdige profiler, nettverk og saker, mer i retning av reelle påvirkningskampanjer, ikke påvirkning i alle retninger (som gjør det vanskelig å måle påvirkning), profiler, nettverk, saker som er mer aktuelle i forhold til dagens politiske situasjon og bruk av flere sosiale medie-kloner (f.eks., Friendica/FB og Pixelfed/Instagram).

Litteraturen på området indikerte videre at man må påregne oppfriskning og gjentakelser av budskap og informasjon til alle deler av befolkningen som man ønsker å gjøre mer motstandsdyktig mot påvirkning. I dette ligger også en tilrettelegging av budskap slik at det blir tilgjengelig for ulike grupper. Her er det behov for mer forskning for å bedre forstå hvordan man best kan gjøre denne tilretteleggingen. Men med truslene som ble påpekt innledningsvis med hensyn til for eksempel valg og våre relasjoner til allierte vi er avhengig av, anbefales det å

---

ikke vente med informasjon og utdanning i påvente av mer forskning – dette synes uansett å ha en gunstig effekt. Her er det også gode eksempler fra våre naboland, for eksempel Finland (Gross, 2023).



---

---

## Referanser

- Abrams, Z. (2021). Controlling the spread of misinformation. *Monitor on Psychology*, 52 (2). Retrieved from <https://www.apa.org/monitor/2021/03/controlling-misinformation>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50, 179–211.
- Alme, V. (2019). *Falske nyheter som sjanger*. FFI-rapport 19/00660. Forsvarets forskningsinstitutt.
- Alme, V., & Hvidsten, A. (2022). To Learn or not to Learn: On the Importance of Mode Switching in Educational Wargames. *Scandinavian Journal of Military Studies*, 5(1), 269–281. DOI: <https://doi.org/10.31374/sjms.123>
- Bago, B., Rand, D. G. & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608–1613. doi: 10.1037/xge0000729
- Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P. & van der Linden, S. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society*, 8(1), 1-18. Doi: 10.1177/205395172111013868.
- Basol, M., Roozenbeek, J. & van der Linden, S. (2020). Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News. *Journal of Cognition*, 3(1), 1-9. Doi: 10.5334/joc.91.
- Baron, J, Isler, O. & Yilmaz (2022). Actively open-minded thinking and the political effects of its absence. In V. Ottati & C. Stern (Eds.), *Divided: Open-Mindedness and Dogmatism in a Polarized World*. Oxford: Oxford University Press. Preprint doi: 10.31234/osf.io/g5jhp
- Bergh, A. (2023). Guide to social media training with Somulator (FFI-rapport 2023/00493). Kjeller: Forsvarets forskningsinstitutt.
- Bergh, A. (2020). *Påvirkningsoperasjoner i sosiale medier – oversikt og utfordringer* (FFI-rapport 20/01694). Kjeller: Forsvarets forskningsinstitutt.
- Bierwaczzonec, K., Gundersen, A. B. & Kunst, J. R. (2022). The role of conspiracy beliefs for COVID-19 health responses: A meta-analysis. *Current Opinion in Psychology*, 46. Doi: 10.1016/j.copsyc.2022.101346
- Bitterman, M. E. (2006). Classical conditioning since Pavlov. *Review of general psychology*, 10(4), 365-376. Doi: 10.1037/1089-2680.10.4.365

---

---

Bjørnstad, A. L. (2019). *Understanding communication and influence in a defense context: A review of relevant research from the field of psychology* (FFI-rapport 19/01224). Kjeller: Forsvarets forskningsinstitutt.

Borah, P. (2022). The Moderating Role of Political Ideology: Need for Cognition, Media Locus of Control, Misinformation Efficacy, and Misperceptions About COVID-19. *International Journal of Communication*, 16, 26. Retrieved from <https://ijoc.org/index.php/ijoc/article/view/18261/3845>

Brandtzæg, P. B., Luders, M., Spangenberg, J., Rath-Wiggins, L. & Følstad, A. (2016). Emerging journalistic verification practices concerning social media. *Journalism Practice*, 10 (3) 323–342. doi:10.1080/17512786.2015.1020331

Cacioppo, J. T., Petty, R. E., Feinstein, J. A. & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119(2), 197–253. Doi:10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E. & Kao, C. F. (1984). The efficient assessment of Need for Cognition. *Journal of Personality Assessment*, 48 (3), 306–307.

Cacioppo, J. T., Petty, R. E. & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, 45(4), 805–818.

Carrasco-Farré, C. (2022). The fingerprints of misinformation: how deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. *Humanities & Social Sciences Communication*, 9, 162 doi: 10.1057/s41599-022-01174-9

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752–766.

Chun, W. & Kruglanski, A. W. (2006). The role of task demands and processing resources in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*, 91, 205–217.

Cialdini, R. B. (2001). Harnessing the science of persuasion. *Harvard Business Review*, October (R0109D), 72–79.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. San Diego, CA: Harcourt Brace Jovanovich.

- 
- Etterretningstjenesten, Forsvaret. (2022). *Fokus 2022: Etterretningstjenestens vurdering av aktuelle sikkerhetsutfordringer*. Retrieved from [https://www.forsvaret.no/aktuelt-og-presse/publikasjoner/fokus/rapporter/Fokus-2022-til-web.pdf/\\_attachment/inline/ec6bec00-d2d3-41c0-af08-02b3b494e8b7:e4014ab4d0e3bd8b2509e7974430fe121e0473ba/Fokus-2022-til-web.pdf](https://www.forsvaret.no/aktuelt-og-presse/publikasjoner/fokus/rapporter/Fokus-2022-til-web.pdf/_attachment/inline/ec6bec00-d2d3-41c0-af08-02b3b494e8b7:e4014ab4d0e3bd8b2509e7974430fe121e0473ba/Fokus-2022-til-web.pdf)
- Etterretningstjenesten, Forsvaret. (2021). *Fokus 2021: Etterretningstjenestens vurdering av aktuelle sikkerhetsutfordringer*.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Forsvarsdepartementet (2021–2022). *Meld. St.10: Prioriterte endringer, status og tiltak i forsvarssektoren*. (Tilråding fra Forsvarsdepartementet). Retrieved from <https://www.regjeringen.no/no/dokumenter/meld.-st.-10-20212022/id2908167/>
- Fiske, S. T. & Taylor, S. E. (2017). *Social cognition: From brains to culture (3<sup>rd</sup> Ed)*. London: Sage.
- Fischer, S. C., Spiker, V. A. & Riedel, S. L. (2009). *Critical thinking training for army officers volume two: A model of critical thinking*. U.S. Army Research Institute for the Behavioral and Social Sciences: Research Report 1882.
- Filkukova, P. & Langguth, (2021). Blinded by emotions: The association between emotional reactivity and trust in fictitious news stories on crime. *Studia Psychologica*, 63 (4), 404-416. Doi: 10.31577/sp.2021.04.833
- Ginossar, Z. & Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of experimental social psychology*, 16, 228–242.
- Grizāne, A., Isupova, M. & Vorteil, V. (2022). *Social media monitoring tools: An in-depth look*. Riga: NATO STRATCOM COE. ISBN: 978-9934-619-20-5.
- Gross, J. (2023, January 10). How Finland is teaching a generation to spot misinformation. *The New York Times*. Retrieved from <https://www.nytimes.com/2023/01/10/world/europe/finland-misinformation-classes.html>
- Guadagno, R. E. & Cialdini, R. B. (2007). Persuade him by e-mail, but see her in person: Online persuasion revisited. *Computers in Human Behavior*, 23, 999–1015.
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J. & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536-15545.

- 
- Haugtvedt, C. P. & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology*, Vol 63(2), 308–319.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Heijltjes, A., van Gog, T., Leppink, J. & Paas, F. (2014). Improving critical thinking: Effects of dispositions and instructions on economics students' reasoning skills. *Learning and Instruction*, 29, 31–42.
- Hewstone, M., Stroebe, W. & Stephenson, G. M. (1996). *Introduction to social psychology*. Oxford: Blackwell Publishers Ltd.
- Hinsz, V. B. & Jackson, J. W. (2022). The relevance of group dynamics for understanding the U.S. Capitol insurrection. *Group Dynamics: Theory, Research, and Practice*, 26(3), 288–308. Doi: 10.1037/gdn0000191
- Hofoss, E. (2022, 7. September). *Maner til felles innsats mot påvirkning* (Nyhetssak om Påvirkningskonferansen 2022). Forsvarets Forskningsinstitutt. Retrieved from <https://www.ffi.no/aktuelt/nyheter/maner-til-felles-innsats-mot-pavirkning>
- Kemp, S. (2022, February 1). Digital 2022: Another year of bumper growth. *We Are Social UK*. Retrieved from <https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/>
- Klaczynski, P. A., Fauth, J. M. & Swanger, A. (1998). Adolescent identity: Rational vs. experiential processing, formal operations, and critical thinking beliefs. *Journal of Youth and Adolescence*, 27(2), 185–207.
- Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341–350.
- Kumankov, A. (2023). Nazism, Genocide and the Threat of the Global West: Russian Moral Justification of War in Ukraine. *Etikk i praksis - Nordic Journal of Applied Ethics*, 1-21. Doi: 10.5324/eip.v17i1.4997
- Kuvaas, B. & Kaufmann, G. (2004). Impact of Mood, Framing, and Need for Cognition on Decision Makers' Recall and Confidence. *Journal of Behavioral Decision Making*, 17, 59–74. Doi: 10.1002/bdm.461
- Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., ... Zaragoza, M. S. (2020). *The Debunking Handbook 2020*. Doi: 10.17910/b7.1182

- 
- Lins de Holanda Coelho, G. L., Hanel, P. H. P. & Wolf, L. J. (2020). The Very Efficient Assessment of Need for Cognition: Developing a Six-Item Version. *Assessment*, 27(8) 1870–1885. Doi: 10.1177/1073191118793208
- Liu, Y.-I. & Eveland, W. P. (2005). Education, Need for Cognition, and Campaign Interest as Moderators of News Effects on Political Knowledge: An Analysis of the Knowledge Gap. *Journalism & Mass Communication Quarterly*, 82(4), 910–929. Doi: 10.1177/107769900508200410
- Maertens, R., Roozenbeek, J., Basol, M. & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied*, 27(1), 1–16. Doi: org/10.1037/xap0000315
- Maloney, E. A. & Retanal, F (2020). Higher math anxious people have a lower need for cognition and are less reflective in their thinking, *Acta Psychologica*, 202, 102939. Doi: 10.1016/j.actpsy.2019.102939
- Matheson, K. & Zanna, M. P. (1989). Persuasion as a function of self-awareness in computer-mediated communication. *Social behavior*, 4, 99–111.
- McGrew, S. (2019). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education*, 145, 103711. Doi: 10.1016/j.compedu.2019.103711
- Medietilsynet (2022). *Barn og medier 2022: Barn og unges bruk av sosiale medier*. Retrieved from [https://www.medietilsynet.no/globalassets/publikasjoner/barn-og-medier-undersokelser/2022/Barn\\_og\\_unges\\_bruk\\_av\\_sosiale\\_medier.pdf](https://www.medietilsynet.no/globalassets/publikasjoner/barn-og-medier-undersokelser/2022/Barn_og_unges_bruk_av_sosiale_medier.pdf)
- Medietilsynet (2022b). *Stopp. Tenk. Sjekk. Slik avslører du desinformasjon og falske nyheter*. Retrieved from <https://www.medietilsynet.no/digitale-medier/kritisk-medieforstaelse/stopp-tenk-sjekk>
- Medietilsynet (2021). *Kritisk medieforståelse i den norske befolkningen*. Retrieved from [https://www.medietilsynet.no/globalassets/publikasjoner/kritisk-medieforstaelse/211214-kmf\\_hovudrapport\\_med\\_engelsk\\_2021.pdf](https://www.medietilsynet.no/globalassets/publikasjoner/kritisk-medieforstaelse/211214-kmf_hovudrapport_med_engelsk_2021.pdf)
- Medietilsynet (2019). *Kritisk medieforståelse i den norske befolkningen*. Retrieved from <https://www.medietilsynet.no/globalassets/dokumenter/rapporter/kritisk-medieforstaelse-undersokelser-2019-2020/2019-Kritisk-medieforstaelse-rapport/>
- Miller, D. T. & Ross, M. (1975). Self-serving biases in attribution of causality: Fact or fiction? *Psychological Bulletin*, 82, 213–225.
- Miller, T. M. & Geraci, L. (2011). Unskilled but aware: Reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 502–506. Doi: 10.1037/a0021802

---

Mokhtari, H., Davarpanah, M.-R., Dayyani, M.-H. & Ahanchian, M.-R. (2013). Students' need for cognition affects their information seeking behavior. *New Library World*, 114 (11/12), 542-549. Doi: 10.1108/NLW-07-2013-0060

Mueller, R. (2019). *Report on the investigation into Russian interference in the 2016 Presidential Election: Volume I of II*. U.S. Washington DC: Department of Justice. Retrieved from <https://www.justice.gov/archives/sco/file/1373816/download>

Muscanell, N. L., Guadagno, R. E. & Murphy, S. (2014). Weapons of influence misused: A social influence analysis of why people fall prey to internet scams. *Social and Personality Psychology Compass*, 8 (7), 388–396.

Nasjonal sikkerhetsmyndighet (2021). *Risiko 2021: helhetlig sikring mot sammensatte trusler*. Retrieved from [https://nsm.no/getfile.php/136419-1616673370/NSM/Filer/Dokumenter/Rapporter/NSM\\_Risiko\\_2021\\_web\\_enkeltside\\_1203.pdf](https://nsm.no/getfile.php/136419-1616673370/NSM/Filer/Dokumenter/Rapporter/NSM_Risiko_2021_web_enkeltside_1203.pdf)

NOU, Norges offentlige utredninger, 14 (2023). *Forsvarskommisjonen av 2021: Forsvar for fred og frihet*. Oslo: Teknisk redaksjon, Departementenes sikkerhets- og serviceorganisasjon. ISBN: 978-82-583-1559-6.

Nunnally, J. (1978). *Psychometric theory* (2 ed). New York: McGraw-Hill.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Translated and edited by G. V. Anrep. Oxford, England: Oxford University Press.

Pedhazur, E. J. & Schmelkin, L. P. (1991). *Measurement, design and analysis. An integrated approach*. Laurence Erlbaum Associates: London.

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, 31 (7), 770-780. Doi: 10.1177/0956797620939054

Pennycook, G. & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39-50. Doi: [10.1016/j.cognition.2018.06.011](https://doi.org/10.1016/j.cognition.2018.06.011)

Petty, R. E. & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (vol. 19, pp.123–205). New York: Academic Press.

Piksa, M., Noworyta, K., Piasecki, J., Gwiazdzinski, P., Gundersen, A. B., Kunst, J. & Rygula, R. (2022). Cognitive Processes and Personality Traits Underlying Four Phenotypes of Susceptibility to (Mis)Information. *Frontiers in Psychiatry*, 13. doi: 10.3389/fpsy.2022.912397

---

---

Politiets Sikkerhetstjeneste (2021). *Nasjonal Trusselvurdering 2021*. Retrieved from <https://www.pst.no/alle-artikler/trusselvurderinger/nasjonal-trusselvurdering-2021/>

Priester, J. R. & Petty, R. E. (2003). The influence of spokesperson trustworthiness on message elaboration, attitude strength, and advertising effectiveness. *Journal of Consumer Psychology*, 13, 408–421.

Reber, R. & Schwartz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth. *Consciousness and Cognition*, 8, 338–342.

Roache, M., Tewa, S., Cadier, A., Labbe, C., Padovese, V., Schmid, R., ... Palmer, K. (2023, June 6). Russia-Ukraine Disinformation Tracking Center: 378 Websites Spreading War Disinformation And The Top Myths They Publish. *NewsGuard*. Retrieved from: <https://www.newsguardtech.com/special-reports/russian-disinformation-tracking-center/>

Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R., Sultan, M. & van der Linden, S. (2022b). Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgment and Decision Making*, 17 (3), 547-573.

Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S. & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation in social media. *Science Advances*, 8, 1-11.

Roozenbeek, J. & van der Linden, S. (2020). Breaking Harmony Square: A game that “inoculates” against political misinformation. *The Harvard Kennedy School (HKS) Misinformation Review*, 1(8). <https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37366465>

Roozenbeek, J. & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, 5 (65). Doi: 10.1057/s41599-019-0279-9

Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in Experimental Psychology* (vol.10, pp.174–221). New York: Academic Press.

Ross, M. & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, 37, 322–337.

Schlosser, A. E. & Shavitt, S. (2009). The effect of perceived message choice on persuasion. *Journal of Consumer Psychology*, 19(3), 290-301.

Schumann, D. W., Kotowski, M. R., Ahn, H. A. & Haugtvedt, C. P. (2012). The Elaboration Likelihood Model: A 30-Year Review. In S. Rodgers, and E. Thorson, (Eds.), *Advertising Theory* (pp. 51–68). New York: Routledge.

---

---

Sentio Research Group (2022). *Meningsmåling for medietilsynet om medievaner og falske nyheter*. Retrieved from <https://www.medietilsynet.no/globalassets/publikasjoner/kritisk-medieforstaelse/220404-undersokelse-nyhetsvaner-og-falske-nyheter-i-forbindelse-med-krigen-i-ukrania.-sentio-pa-vegne-av-medietilsynet.pdf>

Shi, J., Chen, Z. & Tian, M. (2011). Internet self-efficacy, the need for cognition, and sensation seeking as predictors of problematic use of the Internet. *Cyberpsychology, Behavior, and Social Networking*, 14 (4), 231–234. Doi: 10.1089/cyber.2009.0462.

Sivertsen, E., Bjørgul, L., Lundberg, H., Endestad, I., Bornakke, T., Kristensen, J. B. & Christensen, N. M. (2022). *Uønsket utenlandsk påvirkning? – kartlegging og analyse av stortings-valget 2021* (FFI-rapport 21/02746). Kjeller: Forsvarets forskningsinstitutt.

Sivertsen, E. G., Hellum, N., Bergh, A. & Bjørnstad, A. L. (2021). *Hvordan gjøre samfunnet mer robust mot uønsket påvirkning i sosiale medier* (FFI-rapport 21/01237). Kjeller: Forsvarets forskningsinstitutt.

Stahlberg, D. & Frey, D. (1996). Attitudes: structure, measurement and functions. In M. Hewstone, W., Stroebe, & G. M. Stephenson (Eds.) *Introduction to social psychology* (pp 205–239). Oxford: Blackwell Publishers Ltd.

Van Swol, L., Lee, S. & Hutchins, R. (2022). The banality of extremism: The role of group dynamics and communication of norms in polarization on January 6. *Group Dynamics: Theory, Research, and Practice*, 26(3), 239–251. Doi: 10.1037/gdn0000180.

“The News literacy project” (2022). Retrieved from <https://newslit.org/>

Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Verplanken, B., Hazenberg, P. T. & Palenewen, G. R. (1992). Need for cognition and external information search effort. *Journal of Research in Personality*, 26(2), 128-136.

Winget, J. R. & Park, E. S. (2022). Discharged in D.C.: The role of disinhibition in the behavior of insurrection group members. *Group Dynamics: Theory, Research, and Practice*, 26(3), 252–262. Doi: 10.1037/gdn0000182

Wong, Y. (2019). Developing an academic discipline of wargaming [Public Lecture, King’s Wargaming Network]. <https://www.youtube.com/channel/UCgHWLM5I32fRKgoclCDaNhg>



---

---

## A Vedlegg

### A.1 URL spørreskjema 1 & 2, Forhåndsinformasjon 1, 2, 3 og Twitterklone/Mastodon

#### Spørreskjema

ca.KI.09:20 Spørreskjema 1/ Innledende spørreskjema:

<https://survey.euro.confirmit.com/wix/p145422769916.aspx>

ca.kl.11:20 Spørreskjema 2/ Post-eksperiment Spørreskjema:

<https://survey.euro.confirmit.com/wix/p610389561096.aspx>

#### ca.kl.09:35 Forhåndsinformasjon

Info gruppe 1: <https://www.ffi.no/forskning/eksperiment-tekstinnhold>

Info gruppe 2: <https://www.ffi.no/forskning/eksperiment-video/>

Info gruppe 3: <https://harmonysquare.game/en>

Info gruppe 4: <https://www.ffi.no/aktuelt/podkaster/menneskelig-yteevne-i-kaldt-vaer>

ca.KI.10:20 <b>Twitter kloner («Mastodon»)</b>	
1. Gå inn på dette nettstedet i nettleseren:	<a href="https://tw.demo.c-spi.net">https://tw.demo.c-spi.net</a>
2. En dialogboks vil spørre deg om brukernavn og passord for å komme inn på denne nettsiden. Logg in med disse detaljene som er felles for alle:  (NB: på noen nettlesere må dette gjøres to ganger)	<b>Bruker:</b> somulator <b>Passord:</b> PRRu87TVw5bEYkMU_

3. Deretter vil du se en vanlig nettside med et innloggingsskjema. Logg inn med ditt spesifikke brukernavn & passord:

**Bruker:** XyZ1-34  
**Passord:** AmUrask1-34

## A.2 Eksperiment-manual for FFI (C-SPI) 16.02.2022

0900-0910	<b>Ankomst med Kaffe/Te: Tilfeldig inndeling i grupper (1-4) ved å trekke en lapp</b>
0910-0920	<b>Velkomst/innledning på Microsoft Teams</b>  - Én fra FFI følger og veileder hver gruppe fra 4 forskjellige rom (AL-gr1-rom2, Nina-gr2-rom 1, Eskil/Betina-gr3-rom 5/6, Julie-gr4-rom7/8)  - FFI-representanten ber deltagerne i sin gruppe om å:  Ta frem Laptop og headset, gå inn på Microsoft Teams hvor det blir felles velkomst/innledning  Konferansesenterets nett: <b>ISS_GUEST</b> , Passord: <b>GoodService2019</b>
0920-1135	<b>Eksperiment/hoveddel [2 timer, 15 min]</b>  <b>VIKTIG:</b> Ingen må røpe hypotesen/ hva vi tester i eksperimentet før ETTER vi er ferdige. Da får alle en debrief med all informasjon.
0920-0935	<b>Spørreskjema 1 [15 min]</b> <a href="https://survey.euro.confirmit.com/wix/p145422769916.aspx">https://survey.euro.confirmit.com/wix/p145422769916.aspx</a>  FFI-representanten ber alle deltagerne i gruppen hen har ansvaret for om å:  1) Gå inn på nettadressen/URL over og svare på spørreskjemaet der (Spørreskjema 1/Innledende: NFC, demografi, oppfatninger)  2) <b>Viktig – først:</b> Fyll ut din <b>individuelle Mastodon-bru</b> ker på <b>forsiden</b> (dvs., den andre innloggingsbrukeren som starter med <b>XyZ</b> etterfulgt av et tall fra <b>1</b> til <b>34</b> )  3) Si videre at:  - klikk på det svaret som passer best for deg: det er ingen riktige eller gale svar  - spørreskjemasvarene er anonymisert (dvs., at ingen kan knytte det til deltagerens navn/person)  - dataene vil kun bli analysert på aggregert nivå  - det er frivillig å svare

	<p>- dataene vil behandles konfidensielt av forskere på FFI og skal brukes som kontrollvariabler i analysene</p> <p>- det er lov til å trekke seg undervegs hvis noen skulle ønske det</p>
<b>0935-1015</b>	<p><b>Forhåndsinformasjon [35 min]</b></p> <p>FFI-representanten ber alle deltagerne i gruppen hen har ansvaret for om å:</p> <p>- gå inn på «Informasjon1, «Informasjon2», «Informasjon3», eller «Informasjon4» i henhold til hvilket gruppenummer de har fått</p> <p>- <b>gr1 (Anne Lise):</b> deltagerne blir bedt om å lese det de finner på:  <a href="https://www.ffi.no/forskning/eksperiment-tekstinnhold">https://www.ffi.no/forskning/eksperiment-tekstinnhold</a></p> <p>- <b>gr2 (Nina):</b> deltagerne blir bedt om å lytte til det de finner på:  <a href="https://www.ffi.no/forskning/eksperiment-video/">https://www.ffi.no/forskning/eksperiment-video/</a></p> <p>- <b>gr3 (Eskil/Betina):</b> deltagerne blir bedt om å gå inn på spillet <i>Harmony Square</i>:  <a href="https://harmonysquare.game/en">https://harmonysquare.game/en</a></p> <p>og spille dette; si at deltagerne IKKE skal fylle ut spørsmålene som blir stilt avslutningsvis.</p> <p>- <b>gr4 (Julie):</b> deltagerne blir bedt om å høre pod kast/lese rapporter om menneskelig yteevne i kaldt vær ved å gå inn på FFIs nettside:  <a href="https://www.ffi.no/aktuelt/podkaster/menneskelig-yteevne-i-kaldt-vaer">https://www.ffi.no/aktuelt/podkaster/menneskelig-yteevne-i-kaldt-vaer</a></p>
<b>1015-1025</b>	<p><b>Pause: Kaffe/toalett [10 min]</b></p>
<b>1025-1125</b>	<p><b>Mastodon/Twitter-klone [1 time]</b></p> <p>FFI-representanten ber alle deltagerne i gruppen hen har ansvaret for om å:</p> <p>- Logge seg inn på Mastodon:</p> <p>1) Be deltagerne gå inn på: <a href="https://tw.demo.c-spi.net">https://tw.demo.c-spi.net</a></p> <p>2) En dialogboks vil spørre deltagerne om brukernavn og passord for å komme inn på denne nettsiden. Logg in med disse detaljene som er felles for alle (NB: på noen nettlesere må dette gjøres to ganger): Bruker: <b>somulator</b>, Passord: <b>PRRu87TVw5bEYkMU_</b></p> <p>3) Deretter vil deltagerne se en vanlig nettside med et innloggings skjema. Her logger deltagerne inn med sitt spesifikke brukernavn &amp; passord: Bruker: <b>XyZ1-34</b>, Passord: <b>AmUrask1-34</b> (dvs., de individuelle brukernavnene starter på XyZ1 og går til XyZ34 og passordene fra AmUrask1 til AmUrask34)</p>

	<p>- Be deltagerne deretter om å lese (gr1)/se&amp;høre (gr2)/spille (gr3)/se&amp;høre (gr4) det de finner der:</p> <ul style="list-style-type: none"> <li>• Si at de står fritt til å velge å lese bare meldinger, eller også å klikke seg inn på linkene og lese der</li> <li>• Si også at de kan retvitte, osv. – hvis de har lyst til det</li> <li>• Mao – de kan gjøre det de pleier/har lyst til på Twitterklonen</li> <li>• Viktig: advare om at det kan være innhold som kan oppleves støtende</li> <li>• Og minne om at det er lov å trekke seg undervegs..</li> </ul>
1125-1140	<p><b>Spørreskjema 2 [15 min]</b> <a href="https://survey.euro.confirmit.com/wix/p610389561096.aspx">https://survey.euro.confirmit.com/wix/p610389561096.aspx</a></p> <p>FFI-representanten ber alle deltagerne i gruppen hen har ansvaret for om å:</p> <ol style="list-style-type: none"> <li>1) Avslutte det de holder på med på Mastodon</li> <li>2) Gå inn på nettadressen/URL over og svare på spørreskjemaet der (Spørreskjema 2/ Post-eksperiment: oppfatninger og evaluering av saker fra Twitterklonen/Mastodon)</li> <li>3) <b>Viktig – først:</b> Fyll ut din <b>individuelle Mastodon-bruker på forsiden</b> (dvs., den samme som du skrev inn på det første spørreskjemaet)</li> <li>4) Si videre at: <ul style="list-style-type: none"> <li>- klikk på det svaret som passer best for deg: det er ingen riktige eller gale svar</li> <li>- er det vanskelig: helt ok å gjette</li> <li>- svarene er anonymisert, frivillig å svare</li> <li>- dataene analyseres på aggregert nivå og behandles konfidensielt av forskere på FFI</li> </ul> </li> </ol>
1140-1145	<b>Pause: Kaffe/toalett [5 min]</b>
1145-1215	<b>Avsluttende del [30 min]</b>
1145-1155	<p><b>Gruppediskusjon</b> (FFI representanten i hver gruppe fasiliterer og tar notater) <b>[10 min]</b></p> <p>FFI-representanten ber alle deltagerne i gruppen hen har ansvaret for om å:</p> <ol style="list-style-type: none"> <li>1) dele sine opplevelser om hvordan de synes oppgavene og dagen har vært så langt: <ul style="list-style-type: none"> <li>- fikk de nok tid på de ulike delene?</li> <li>- var oppgavene forståelige?</li> <li>- var det noe som ikke virket (teknisk)?</li> <li>- var det noe de opplevde som rart ellers?</li> </ul> </li> </ol>

1155-1215	<p>2) dele hvordan de opplevde sakene på Mastodon/Twitterklonen</p> <p>- var sakene underholdende, realistiske, rare, eller noe annet?</p> <p>3) dele om de syntes informasjonen (gr1 og 2)/spillet (gr3)/nyhetene (gr4) de fikk i forkant påvirket hvordan de oppfattet sakene på Mastodon</p> <p>– hvis ja: hvordan?</p> <p><b>Avslutning/debrief</b> på Microsoft Teams: [20 min]</p> <p>FFI-representanten ber til slutt deltagerne i sin gruppe om å gå inn på Teams, hvor det blir avslutning med debrief/bakenforliggende informasjon om hva de har vært med på og fellesdiskusjon</p>
1215-1300	Lunsj [45 min] (på gruppe-rommene)
1300-1330	Presentasjon FFI/ Avd. Totalforsvar/ C-SPI ved Janet og Lea på Microsoft Teams (bruker samme møteinnkalling som før lunsj): [30 min]

### A.3 Spørreskjema

#### A.3.1 Pre-eksperiment spørreskjema

*Dette spørreskjemaet har tre deler: bakgrunnsinformasjon, mål på noen oppfatninger og et mål på kognitive preferanser. Besvarelsen er anonym og vil bli analysert kun på aggregert nivå. Svarene på spørsmålene i dette skjemaet skal brukes som kontrollvariabler i analysene.*

*Det er frivillig å svare. Vi understreker at det er ingen riktige eller gale svar.*

*Dataene vil behandles konfidensielt av forskere på FFI.*

*Tusen takk for din deltagelse!*

#### Bakgrunn

Klikk på det svaret som passer best for deg.

Er du?

Mann	<input type="checkbox"/>
Kvinne	<input type="checkbox"/>

Alder?

21-30 <input type="checkbox"/>	41-50 <input type="checkbox"/>	61-70 <input type="checkbox"/>
31-40 <input type="checkbox"/>	51-60 <input type="checkbox"/>	

Er du?

Militær <input type="checkbox"/>
Sivil <input type="checkbox"/>

Hva er din høyeste utdanning?

Befalsskole <input type="checkbox"/>	Krigsskole <input type="checkbox"/>	Stabsskole <input type="checkbox"/>	Dr.grad <input type="checkbox"/>
Sivil lavere grad <input type="checkbox"/>	Universitet/Høyskole (Bachelor-nivå) <input type="checkbox"/>	Universitet/Høyskole (Master-nivå) <input type="checkbox"/>	Annet <input type="checkbox"/>

Hvor har du ditt virke?

Forsvaret <input type="checkbox"/>	Offentlig sektor <input type="checkbox"/>	Student <input type="checkbox"/>
Departement <input type="checkbox"/>	Privat næringsliv <input type="checkbox"/>	Annet <input type="checkbox"/>

### Oppfatninger av saker generelt

Klikk på det svaret som stemmer best med din oppfatning av følgende saker – på en skala fra helt uenig til helt enig. Det er ingen riktige eller gale svar. Sett kun ett svar-kryss for hvert spørsmål.

Jeg føler meg trygg på vaksinene mot Korona/Covid-19				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Generelt føler jeg meg trygg på myndighetenes anbefalinger				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg mener vaksinene mot Korona/Covid-19 er utrygge				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Generelt føler jeg meg ikke trygg på myndighetenes anbefalinger				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Det er mye konflikter og splittelse mellom by og land i Norge				

Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har et positivt inntrykk av det norske Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Norge bruker for mye penger på Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har et negativt inntrykk av det norske Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Forsvaret tar ikke miljøhensyn				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Likestillingen har gått for langt – kvinner har fått fordeler på bekostning av menn på mange områder				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Likestillingen har fortsatt langt å gå for at kvinner skal stå på likefot med menn				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har tillit til myndighetenes pandemiltak				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har ikke tillit til myndighetenes pandemiltak				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Det er mye rasisme i Norge				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har godt inntrykk av USAs president Biden				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>

### Kognitive preferanser (NFC; Cacioppo, Petty & Cao, 1984)

Kryss av for hvor enig/uenig du er i de følgende utsagnene. Sett kun ett svar-kryss for hvert spørsmål.

1. I would prefer complex to simple problems.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>

2. I like to have the responsibility of handling a situation that requires a lot of thinking.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
3. Thinking is not my idea of fun.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
5. I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
6. I find satisfaction in deliberating hard and for long hours.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
7. I only think as hard as I have to.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
8. I prefer to think about small, daily projects to long-term ones.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
9. I like tasks that require little thought once I've learned them.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
10. The idea of relying on thought to make my way to the top appeals to me.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
11. I really enjoy a task that involves coming up with new solutions to problems.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
12. Learning new ways to think doesn't excite me very much.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
13. I prefer my life to be filled with puzzles that I must solve.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
14. The notion of thinking abstractly is appealing to me.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
15. I would prefer a task that is intellectual, difficult and important to one that is somewhat important but does not require much thought.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>



16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
17. It's enough for me that something gets the job done; I don't care how or why it works.*				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>
18. I usually end up deliberating about issues even when they do not affect me personally.				
Strongly disagree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Neutral <input type="checkbox"/>	Agree <input type="checkbox"/>	Strongly Agree <input type="checkbox"/>

\* = reverse coding.

### A.3.2 Post-eksperiment spørreskjema

*Dette spørreskjemaet har som formål å fange opp dine oppfatninger og evalueringer av sakene som du fikk lese på Mastodon (Twitterklonen). Det er frivillig å svare.*

*Vi understreker at det er ingen riktige eller gale svar.*

*Besvarelsen er anonym og vil bli analysert kun på aggregert nivå. Dataene vil behandles konfidensielt av forskere på FFI.*

*Takk for dine svar – de er svært viktige for oss.*

### Oppfatninger av saker generelt

Klikk på det svaret som stemmer best med din oppfatning av følgende saker – på en skala fra helt uenig til helt enig. Det er ingen riktige eller gale svar. Kun ett svar-klikk for hvert spørsmål.

Jeg føler meg trygg på vaksinene mot Korona/Covid-19				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Generelt føler jeg meg trygg på myndighetenes anbefalinger				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg mener vaksinene mot Korona/Covid-19 er utrygge				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Generelt føler jeg meg ikke trygg på myndighetenes anbefalinger				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Det er mye konflikter og splittelse mellom by og land i Norge				

Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har et positivt inntrykk av det norske Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Norge bruker for mye penger på Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har et negativt inntrykk av det norske Forsvaret				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Forsvaret tar ikke miljøhensyn				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Likestillingen har gått for langt – kvinner har fått fordeler på bekostning av menn på mange områder				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Likestillingen har fortsatt langt å gå for at kvinner skal stå på likefot med menn				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har tillit til myndighetenes pandemiltak				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har ikke tillit til myndighetenes pandemiltak				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Det er mye rasisme i Norge				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>
Jeg har godt inntrykk av USAs president Biden				
Helt uenig <input type="checkbox"/>	Delvis uenig <input type="checkbox"/>	Verken eller <input type="checkbox"/>	Delvis enig <input type="checkbox"/>	Helt enig <input type="checkbox"/>

### Evaluering av sakene på Mastodon (Twitterklonen)

Klikk av for hva som best reflekterer din evaluering av de følgende sakene i forhold til sannhetsgehalt – på en skala fra helt usant til helt sant. Kun ett svarklikk for hvert spørsmål.

Du vil kanskje synes at noen er vanskelige å bestemme – da ber vi deg å gi din beste gjetning. Klikk også av for hvor sikker du er på hvert svar du gir på en svarskala fra 1 («helt usikker») - 5 («helt sikker»).

1. Vaksinerne mot Korona/Covid-19 er farlige				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
2. Det er mange som mener vaksinerne mot Korona/Covid-19 er farlige				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
3. Ikke alle vaksinerte oppnår full beskyttelse. Det betyr at etter hvert som en større andel av den voksne befolkningen er vaksinert, vil naturlig nok også en større andel av de som smittes, blir syke og som behøver sykehusinnleggelse eller dør, være i gruppen av vaksinerte.				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
4. Langs Vestlandets fjord – og innlandsstrøk går én hovedvei, som liksom skal knytte forbindelsen mellom tettstedene: E39. Med unntak av noen veldig få bra strekninger, er E39 mange steder like kronglete og smal som en hvilken som helst bygdevei.				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
5. Når Oslo skal få nye sykkel veier verdt 361 millioner kroner og vi kanskje ikke får ny vei, er det problemer i ledelsen.				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
6. Når til og med en professor ved Institutt for samfunnsøkonomi sier at urbanisering er farlig må vi høre!! Hør på Victor D. Norman ved NHH				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
7. Norges planer for reduksjon i klimagassutslipp tar ikke hensyn til olje og gass som selges til utlandet				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
8. Innen 2055 så har temperaturen i havet gått opp med 5 grader				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
9. En frisk 20-åring ble feildiagnostisert med en psykiatrisk diagnose og måtte dra hjem etter noen uker i forsvaret. Hun informeres ikke engang om diagnosen, men dimitteres umiddelbart.				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
10. I 2020 ble det behandlet rundt 500 søknader om abort på sykehuset. Det har aldri vært så lave tall!				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
11. 1 av 5 kvinner bruker graviditet til å lokke en mann til å bli værende i forholdet, ifølge tall fra NRK.				

Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
12. 4 av 10 personer med minoritetsbakgrunn i Norge har opplevd diskriminering i arbeidslivet				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
13. Ikke en eneste nordmann i første klasse på Mortensrud				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
14. Politiet bruker tåregass mot antirasistiske demonstranter				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
15. Etniske nordmenn flytter fra Groruddalen				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
16. Little Mix tar pause etter at Perrie og Leigh ble gravide på samme år				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>
17. Flere kvinner enn menn tar selvmord				
Helt usant <input type="checkbox"/>	Mest usant <input type="checkbox"/>	Blandet sant/usant <input type="checkbox"/>	Mest sant <input type="checkbox"/>	Helt sant <input type="checkbox"/>

**Kvalifiserings spørsmål til hver påstand over (dvs., alle sant/usant-spørsmålene):**

Hvor sikker er du på dette svaret?				
1 (Helt usikker) <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 (Helt sikker) <input type="checkbox"/>

## Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan, med særskilte fullmakter underlagt Forsvarsdepartementet.

## FFIs formål

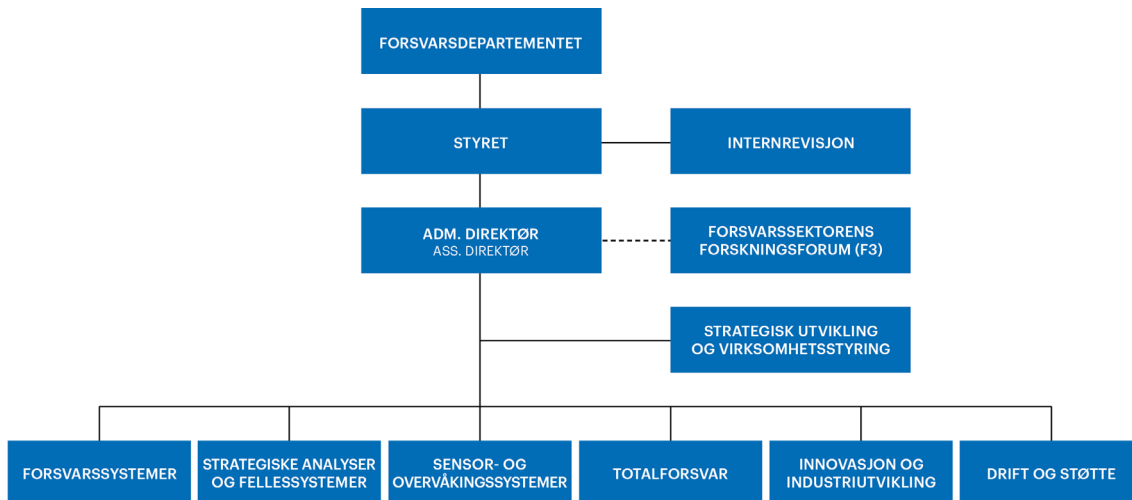
Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

## FFIs visjon

FFI gjør kunnskap og ideer til et effektivt forsvar.

## FFIs verdier

Skapende, drivende, vidsynt og ansvarlig.



Forsvarets forskningsinstitutt (FFI)  
Postboks 25  
2027 Kjeller

Besøksadresse:  
Kjeller: Instituttveien 20, Kjeller  
Horten: Nedre vei 16, Karljohansvern, Horten

Telefon: 91 50 30 03  
E-post: [post@ffi.no](mailto:post@ffi.no)  
[ffi.no](http://ffi.no)

Norwegian Defence Research Establishment (FFI)  
PO box 25  
NO-2027 Kjeller  
NORWAY

Visitor address:  
Kjeller: Instituttveien 20, Kjeller  
Horten: Nedre vei 16, Karljohansvern, Horten

Telephone: +47 91 50 30 03  
E-mail: [post@ffi.no](mailto:post@ffi.no)  
[ffi.no/en](http://ffi.no/en)