

Lasse Lindekilde og Jesper Rasmussen

# Had, humor og bystander-reaktioner: danske unges reaktioner på politisk hadefulde memes<sup>1</sup>

En udfordring med hadefuldt indhold på sociale medier er, at det ofte er pakket ind i en tilsyneladende uskyldig humor. Det gør det hadefulde materiale sværere at genkende og afkode. De fleste kan genkende hadefuldt propaganda, når det følges af et hagekors eller et IS-flag. Men hvordan ser det ud, når det hadefulde budskab er præsenteret som et humoristisk meme? Samtidig rejser eksponering for hadefuldt materiale på sociale medier et vigtigt spørgsmål om, hvordan vi som *bystanders* reagerer. Hvornår intervenserer bystanders prosocialt og for eksempel anmelder hadefuldt indhold, og hvornår reagerer de antisocialt og deler og liker sådant indhold? Artiklen bidrager med en første kortlægning af danske unges vurdering af hadefulde memes, og hvordan disse vurderinger medfører de unges bystander-reaktioner i forhold til deres villighed til at like, dele og anmelde indholdet. Desuden undersøges det, hvad der kendetegner dem, som er tilbøjelige til at reagere henholdsvis prosocialt og antisocialt på hadefulde budskaber på sociale medier med hensyn til køn, politisk interesse samt aktivistiske og politisk voldelige intentioner. Artiklen trækker på en elektronisk spørgeskemaundersøgelse blandt 814 danske folkeskole- og gymnasieelever i alderen 13 til 18 år.

Nøgleord: politisk had, humor, bystander-reaktioner

80 pct. af danskerne rapporterer, at politiske diskussioner på sociale medier er domineret af ”meget ekstreme synspunkter”, mens 57 pct. mener, at debatter er så ”aggressive, at man ikke ved, hvad folk kan finde på” (Bor og Petersen, 2021). Disse forhold har betydet et stigende fokus på at klæde ikke mindst unge mennesker bedre på til at navigere i diskussioner på sociale medier. Elever i folkeskolen og på ungdomsuddannelser gennemgår således forskellige undervisningsforløb, som handler om tonen på sociale medier, grænserne for det, som må siges, hvordan man genkender hadefuldt og ekstremistisk indhold, og hvordan man bør reagere (McDougall et al., 2018).

En voksende bekymring er i denne forbindelse, at hadefulde budskaber på sociale medier i praksis kan være vanskelige at genkende og afkode. En årsag er, at hadefuldt indhold kan være pakket ind i tilsyneladende uskyldige og humoristiske memes og benytte sig af subtile og implicite former for hadefulde

ytringer og trusler. De fleste kan genkende hadefulde propaganda, når det følges af et hagekors eller et IS-flag. Men hvordan ser det ud, når det hadefulde budskab er flettet ind i et humoristisk meme? Forskning peger på, at ekstremistiske grupper benytter sig af *dog whistling* i deres kommunikation, altså en form for skjulte budskaber som kun lader sig afkode af de indviede (Askanius, 2021). Et eksempel er brugen af ”Pepe the frog”-memes blandt højreekstreme grupper i USA. ”Pepe” var oprindeligt en harmløs internettegneserie, men begyndte fra omkring 2009 at blive brugt som en skjult markør på sociale medier af højreekstreme overbevisninger.

Eksponering for hadfuldt materiale på sociale medier rejser også spørgsmålet, hvordan vi reagerer herpå. Forskning fra USA har vist, at 66 pct. af amerikanerne rapporterer at have været vidner til hadfuldt indhold på sociale medier, men kun godt en tredjedel af dem rapporterer at have reageret på dette indhold på den ene eller anden måde (Cheng et al., 2017; Duggan, 2017a, 2017b). To tredjedele forbliver altså passive. I lyset af socialpsykologisk forskning om den såkaldte ”bystander-effekt” er denne passivitet forventelig (Darley og Latane, 1968; Latane og Nida, 1981). Kort fortalt viser forskningen, at jo flere vidner der er til en form for kritisk begivenhed (et hjertetilfælde, en ulykke, chikane m.m.), jo mindre er sandsynligheden for, at hvert enkelt vidne reagerer prosocialt ved at intervenere og hjælpe. Nyere forskning har vist, at denne effekt også findes online blandt ”cyber-bystanders” i forbindelse med eksempelvis mobning på sociale medier (Bastiaensens et al., 2014; Dillon og Bushman, 2015; Freis og Gurung, 2013). Forklaringen på bystander-effekten er typisk, at der sker en form for ”diffusion af ansvar”, når der er mange vidner til en begivenhed (Fischer et al., 2011), hvilket typisk er tilfældet på sociale medier. Intervention kræver således, at man registrerer begivenheden/hændelsen, opfatter den som kritisk, føler en form for ansvar og ved, hvordan man skal reagere (Darley og Latane, 1968). Bystander-effekten reduceres således i takt med, at den hændelse, som man er vidne til, bliver mindre tvetydig og mere tydeligt ”kritisk” og dermed kræver intervention. I den sammenhæng har det altså stor betydning, om man oplever et budskab som hadfuldt eller som humoristisk.

Disse indsigter rejser spørgsmålet, hvilke faktorer som har betydning for, hvordan danske unge reagerer på hadefulde budskaber på sociale medier, herunder om der er en sammenhæng mellem, hvor hadfuldt man vurderer, et meme er, og sandsynligheden for at reagere prosocialt ved at anmelde memet til platformens moderator. Fører budskaber, som opleves som mere hadefulde eller uacceptable, til flere prosociale reaktioner og færre antisociale reaktioner (fx dele eller like indholdet)? Og hvad karakteriserer dem, som er tilbøjelige

til at reagere henholdsvis prosocialt og antisocialt på hadefulde budskaber på sociale medier?

Formålet med denne artikel er at give et første foreløbigt svar på disse spørgsmål i en dansk kontekst. Artiklen undersøger danske skole- og gymnasieelevers vurdering af og reaktioner på en række politisk hadefulde memes. Artiklen bidrager således med en første kortlægning af danske unges vurdering af hadefulde memes, og hvordan disse vurderinger og andre faktorer (køn, politisk interesse samt aktivistiske og politisk voldelige intentioner) medformer de unges bystander-reaktioner i forhold til deres villighed til at like, dele og anmelde indholdet. Mens der i en dansk kontekst er lavet undersøgelser af bystander-reaktioner i forbindelse med vold i nattelivet og offentlig transport (Liebst et al., 2018; Liebst, Heinskou og Ejbye-Ernst, 2018; Liebst et al., 2019), så er dette det første studie, som undersøger bystander-reaktioner på sociale medier i Danmark. Resultaterne viser, at både anti- og prosocial bystander-adfærd forudsiges af køn og opfattelsen af memes som hadefulde, mens voldelige politiske intentioner og opfattelsen af hadefulde memes som acceptable alene forudsiger antisocial bystander-adfærd. Sluttelig forudsiger lovlig politisk aktivisme prosocial adfærd.

I det følgende trækker vi på eksisterende teori og viden om bystander-reaktioner til at teoretisere, for det første, hvilke forhold ved et budskabs indhold som medformer bystander-reaktioner, herunder i særdeleshed betydningen af accept af budskabets indhold og vurderinger af dets hadefuldhed. For det andet teoretiserer vi, hvilke karakteristika ved bystanders som har betydning for, om man tenderer imod mere prosociale eller antisociale reaktioner, når man eksponeres for hadefuldt indhold på sociale medier. Vi benytter disse indsigter til at formulere konkrete hypoteser om, hvordan danske unge kan forventes at reagere på hadefulde memes på sociale medier. Før præsentationen af de empiriske resultater redegør vi kort for undersøgelsens metodiske og datamæssige grundlag. Artiklen afsluttes med en diskussion af de empiriske resultater, undersøgelsens begrænsninger og behovet for fremtidig forskning.

## Politisk had, humor og cyber-bystander-reaktioner

En afgrænsning af politisk had, humor og hvornår en person kan siges at være en bystander online er vanskelig og nødvendigvis delvist et empirisk spørgsmål. Det er uden for denne artikels formåen at give en fyldestgørende diskussion heraf. Her begrænser vi os til at indkredse online politisk had som ”nedværdigende, chikanerende og intimiderende onlineindhold, som er funderet i politiske forskelle eller diskriminerende overbevisninger” (Bor og Petersen, 2021). Spørgsmålet er så, hvorledes lægfolk, herunder danske unge, applicerer en sådan

definition i praksis. Det er vigtigt at slå fast, at politisk had ikke er karakteriseret ved politisk uenighed og konflikt som sådan, hvilket er en definerende egen-skab ved politik generelt (Mutz, 2015), men snarere ved fjendtlige og truende måder at udtrykke en sådan uenighed på.

Humor kan defineres som ”muligheden for mennesker, objekter eller situationer for at fremkalde latter eller vække en følelse af morskab” (Den Store Danske Encyklopædi). Men det slås også med det samme fast, at ”hvad der er sjovt, afhænger af modtagerens kultur, alder, uddannelse, intelligens og ikke mindst af den sammenhæng, det sjove præsenteres i”. Om et meme opfattes som hadefuldt eller blot sjovt og ballade, afhænger altså af konteksten og øjnene, som ser. Med andre ord kan vi altså forvente en vis variation heri.

I den klassiske socialpsykologiske litteratur om bystander-effekten defineres en bystander typisk som et tilfældigt vidne til en kritisk begivenhed, som opfatter hændelsen og derfor – i hvert fald teoretisk – har muligheden for at reagere (Darley og Latane, 1970). Hvad dette betyder i en onlinekontekst, hvor hadefulde ytringer sjældent opleves live eller er rumligt afgrænset, er mere uklart (Price et al., 2014).

Bystander-reaktioner forstås i den klassiske socialpsykologiske litteratur typisk som enten-eller: Enten reagerer man ved at hjælpe, og ellers forholder man sig passiv (Fischer et al., 2011). En række nyere studier med fokus på cyberbystanders og deres reaktioner på had på sociale medier har forsøgt at nuancere dette ved at pege på, at passivitet kun er en form for antisocial bystander-reaktion: Cyberbystanders kan også reagere på hadefulde budskaber ved aktivt at understøtte dem ved fx at like eller dele indholdet (Moxey og Bussey, 2020; Reijntjes et al., 2016). Vi betragter altså både ikkereaktion (passivitet) og proaktiv billigelse af had (dele eller like) som former for antisocial adfærd og trækker i vores teoretisering således både på indsigter herom fra forskning i bystander-adfærd specifikt samt fra forskning i hadefuld onlineadfærd generelt.

Ligeledes findes der flere former for prosociale ”interventioner”. Man kan således sondre mellem reaktioner, som har til hensigt at artikulere misbilligelse ved at appellere til afsenderen af det hadefulde indhold om at stoppe med at sprede had eller engagere sig i en form for *counter-speech*, og reaktioner, som henvender sig til offeret og tilbyder støtte og hjælp (Obermaier, Fawzi og Koch, 2016; Porten-Cheé, Kunst og Emmer, 2020). Endelig tilbyder sociale medier en række mere tekniske prosociale reaktioner til bystanders som for eksempel at anmelde hadefuldt indhold til platformens moderator (Naab, Kalch og Meitz, 2018). Mens den sidste form for prosocial bystander-reaktioner er mindre omkostningsfuld, så kræver direkte artikulation af misbilligelse en vis grad af civiltourage (Iganski, 2020; Greitemeyer et al., 2006; Greitemeyer et al., 2007),

idet sådanne reaktioner kan være forbundet med negative sociale konsekvenser i form af had mod en selv. Der findes en begyndende forskning i konsekvenserne af prosociale bystander-reaktioner på politisk motiveret had på sociale medier, som viser, at counter-speech, der benytter sig af appel til empati med offeret (Hangartner et al., 2021) og understøttelse af budskaber fra personer med høj status, som man identificerer sig med (Siegel og Badaan, 2020; Munger, 2017), er effektiv i forhold til at reducere mængden af had i den efterfølgende interaktion.

I denne artikel fokuserer vi på likes og delinger som eksempler på antisociale bystander-reaktioner og anmeldelser af hadefuldt indhold som eksempler på prosociale reaktioner. Vi begrænser os til at undersøge sammenhængen mellem vurderinger af hadefulde budskaber og bystander-reaktioner, samt hvad der karakteriserer bystanders, som reagerer henholdsvis pro- og antisocialt.

### *Vurdering af hadefulde budskaber på sociale medier og bystander-reaktioner*

Som antydnet vurderes det at være en voksende udfordring for især unge at genkende og afkode politisk hadefulde budskaber på sociale medier, fordi de ofte er pakket ind i humor, billeder og symboler, som tilfældet er med memes, eller præsenteres i en hverdagskontekst, fx i YouTube-videoer eller blogs fra andre unge (Ndoen, 2018). Teoretisk kan denne indpakning af hadefuldt indhold være med til at sløre hadet og skabe en uklarehed om, hvorvidt budskabet er ment for sjov eller faktisk er intentionelt stødende og hadefuldt. Med andre ord bliver det mere uklart, om situationen er ”kritisk”, og intervention påkrævet (Fischer et al., 2006). I mødet med sådant indhold vil vi derfor forvente, at unge i lavere grad, end hvis indholdet var rensset for humor, opfatter indholdet som hadefuldt og uacceptabelt og derfor i mindre grad er villige til at engagere sig i prosociale reaktioner og i højere grad er villige til fx at dele eller like indholdet. Generelt vil vi forvente, at jo mere alvorligt et budskab er i forhold til at være direkte truende, nedsættende eller dehumaniserende, jo mere hadefuldt opfattes budskabet, og jo mere sandsynligt er det, at borgere reagerer, fx med krav om at indholdet slettes eller forbydes (Rasmussen, 2021; Leonhard et al., 2018). Dette understøttes af forskning, som viser, at en handling, som bryder med en persons normer for social adfærd, opfattes som uacceptabel og aktiverer prosociale reaktioner med henblik på at forsvare normative standarder (Wilhelm, Joeckel og Ziegler, 2020; Rains et al., 2017).

I konteksten af dette studie leder disse indsigter os til formuleringen af følgende hypoteser:

H1: Jo mere hadefuldt og uacceptabelt et meme vurderes at være, jo mere villige vil unge være til at reagere prosocialt (anmelde).

H2: Jo mere hadefuldt og uacceptabelt et meme vurderes at være, jo mindre villige vil unge være til at reagere antisocialt (dele eller like).

I litteraturen fremhæves en række andre forhold ved et budskabs karakter som havende betydning for vurderingen af, hvor hadefuldt et budskab er, og dermed for, hvordan man er parat til at reagere. En faktor er det hadefulde budskabs målgruppe. Forskning har således peget på, at borgere er mere tilbøjelige til at reagere prosocialt, når det hadefulde budskab er rettet mod enkeltpersoner og ikke sociale grupper bredt (Naab, Kalch og Meitz, 2018; Wilhelm, Joeckel og Ziegler, 2020; Citron og Norton, 2011). Ligeledes peger forskning på, at bystanders sympati for og identifikation med offeret fremmer prosociale bystander-reaktioner (Mulvey, Palmer og Adams, 2016; Brewer og Kerlake, 2015). Om end den sidste pointe er teoretisk relevant, begrænser vi os i dette studie til at undersøge den direkte sammenhæng mellem vurderingen af, hvor hadefuldt og uacceptabelt et budskab er, og villigheden til at engagere sig i pro- og antisociale bystander-reaktioner.

### *Hvem reagerer pro- og antisocialt på politisk hadefulde budskaber på sociale medier?*

Den klassiske socialpsykologiske litteratur om bystander-effekten konkluderer generelt, at individuelle forskelle betyder relativt lidt, i forhold til hvem som er tilbøjelige til at intervenere og hjælpe (Fischer et al., 2011). Dog peges der på, at specifik træning i krisehåndtering m.m. fremmer interventioner, mens der ikke kan findes en generel uddannelseseffekt (van den Bos, Müller og van Bussel, 2009; Cramer et al., 1988).

Forskning i cybermobning blandt børn viser, at køn spiller en rolle. En række studier fremhæver således, at drenge er mere tilbøjelige til at reagere antisocialt som bystanders, mens piger er mere tilbøjelige til at reagere prosocialt (Machackova, 2020; Mackay, 2019; Quirk og Campbell, 2015). Teoretisk peges der på, at denne kønsforskel kan skyldes både evolutionær udvikling og forskelle i socialisering og normdannelse blandt piger og drenge. Således kan forklaringen på denne kønsforskel både skyldes en større tilbøjelighed blandt drenge til at hævde sig i statushierarkier ved fx at optræde aggressivt og opsøge konflikter og en mere udbredt norm blandt piger om at drage omsorg om de svage og udsatte.

I konteksten af dette studie leder disse indsigter os til formuleringen af følgende hypoteser:

H3: Drengene er mere villige end piger til at reagere antisocialt (dele og like) på hadefulde memes.

H4: Piger er mere villige end drenge til at reagere prosocialt (anmelde) på hadefulde memes.

Den begrænsede eksisterende forskning med specifikt fokus på politisk motiveret onlinehad argumenterer for, at prosociale bystander-reaktioner på sådant indhold skal ses som en moderne form for politisk deltagelse eller civilcourage (Porten-Cheé, Kunst og Emmer, 2020; Iganski, 2020). I tråd med denne opfattelse har studier peget på, at prosociale bystander-reaktioner på hadefulde politiske budskaber på sociale medier drives af nogle af de samme faktorer, som driver politisk deltagelse og engagement generelt. Eksempelvis har forskning vist, at tilslutning til principper om at bidrage aktivt til demokratiet, føle et ansvar for løsning af samfundets fælles problemer, solidaritet med dem, som er dårligere stillet end en selv, og åbenhed over for andres holdninger korrelerer positivt med at anmelde hadefulde ytringer og negativt med antisociale reaktioner i form af at like (Kunst et al., 2021). Ligeledes antyder forskning en positiv sammenhæng mellem generel politisk interesse og prosociale bystander-reaktioner (Porten-Cheé, Kunst og Emmer, 2020; Citron og Norton, 2011). Det er forventeligt, at borgere, som har en høj politisk interesse, deltager mere i politiske debatter på sociale medier. Disse borgere vil alt andet lige blive eksponeret for mere politisk motiveret onlinehad, men også være mere motiverede for at interagere med dette materiale sammenlignet med borgere med lav politisk interesse. Endvidere har studier vist, at villighed til at deltage i lovlig politisk aktivisme forudsiger både anti- og prosocial onlineadfærd (Rasmussen et al., 2021). Villighed til at deltage i lovlig politisk aktivisme udtrykker en parathed til at udnytte ens demokratiske rettigheder til at mobilisere støtte til bestemte politiske synspunkter og aktivt arbejde for at bekæmpe opfattede uretfærdigheder i samfundet. Hvis man i forlængelse af dette opfatter bystander-reaktioner på politisk motiveret online had som en form for politisk aktivisme, kan man forvente, at borgere, som er mere aktivistisk motiverede, vil være mere tilbøjelige til både at like/dele og anmelde sådant materiale alt efter deres politiske tilhørsforhold. Omvendt forudsiges deling af fjendtlige politiske rygter online i høj grad af villigheden til at deltage i voldelig politisk aktivisme offline (Arceneaux et al., 2021), som omvendt korrelerer stærkt med statussøgen og risikofyldt adfærd samt ”mørke” personlighedstræk (Petersen, Osmundsen og Bor, 2021). Argumentet er her, at det er de samme psykologiske motivationer og personlighedstræk i forhold til statussøgen og dominans, som driver norma-

tivafvigende adfærd online og offline. I forlængelse af dette kan man forvente, at borgere, som udtrykker en høj grad af accept af voldelig politisk aktivisme offline, også vil være mere tilbøjelige til at reagere antisocialt (dele og like) på politisk motiveret had online.

Baseret på disse indsigter formulerer vi de følgende konkrete hypoteser:

H5: Politisk interesse korrelerer positivt med villighed til at reagere prosocialt på (anmelde) hadefulde memes.

H6: Lovlig politisk aktivisme korrelerer positivt med villighed til at reagere antisocialt på (dele og like) hadefulde memes.

H7: Lovlig politisk aktivisme korrelerer positivt med villighed til at reagere prosocialt på (anmelde) hadefulde memes.

H8: Voldelig politisk aktivisme korrelerer positivt med villighed til at reagere antisocialt på (dele og like) hadefulde memes.

## Metode og data

For at teste vores hypoteser benytter vi os af spørgeskemadata indsamlet blandt danske elever på folkeskoler og ungdomsuddannelser. Data er indsamlet i forbindelse med et evalueringsstudie af et nyt online dialogværktøj, ”Ekstremt demokratisk”, som er udviklet af C:NTACT til at understøtte klassebaseret dialog i de ældste folkeskoleklasser (7.-10. klasse) og på ungdomsuddannelser om ekstremistisk og hadefuldt indhold på sociale medier (se <http://ekstremtdemokratisk.contact.dk>). Evalueringen var tilrettelagt som et felteksperiment med randomisering på klasseniveau indenfor årgange på de skoler, som var med i undersøgelsen. Implementeringen af felteksperimentet var i praksis udfordret af både forurening mellem indsats- og kontrolgruppe (eleverne snakkede sammen på tværs af årgangen) og af problemer med fidelitet (lærerne anvendte, på trods af nedskrevne instrukser, dialogværktøjet ret forskelligt i undervisningen). Datakvaliteten blev vurderet for lav til faktisk at vurdere effekten af dialogværktøjet. En sammenligning af bystander-reaktioner mellem indsats- og kontrolgruppe viser ingen forskel (se tabel A6 i artiklens appendiks), hvorfor det er muligt at bruge hele datasættet til nærværende analyse.

Spørgeskemaet blev udfyldt elektronisk på elevernes egen computer, tablet eller mobiltelefon via et Qualtrics-link, som læreren delte med eleverne. Spørgeskemaet blev udfyldt i klasselokalet under supervision af læren. Foruden de variable, som anvendes i denne undersøgelse, indeholdt skemaet en række spørgs-



mål, som havde til hensigt at evaluere arbejdet med dialogværktøjet. Forud for udfyldelse af spørgeskemaet afgav eleverne informeret samtykke om deltagelse i undersøgelsen. Her blev eleverne informeret om, at de i forbindelse med undersøgelsen ville blive eksponeret for indhold, som kan forekomme stødende. Det er vigtigt at pointere, at spørgeskemaet blev udfyldt i forbindelse med afviklingen af workshoppen ”Ekstremt demokratisk”, som netop fokuserede på hadefuldt indhold på sociale medier, og hvordan og hvornår man skal reagere herpå. Således fik deltagerne en grundig debriefing i klasselokalet af den pågældende lærer efter workshoppen og udfyldningen af spørgeskemaet.

I alt har 814 danske elever på folkeskoler og ungdomsuddannelser i alderen 13 til 18 år besvaret spørgeskemaet. Dataindsamlingen strakte sig over en periode fra 19. november 2019 til 10. marts 2020. Udvalgelsen af skoler til deltagelse i evalueringen var ikke tilfældig, men afhæng af skolernes interesse i at deltage og afprøve dialogværktøjet i forbindelse med undervisningen i dansk eller samfundsfag. C:NTACT stod for kontakten til skolerne og lavede aftaler om, hvor og hvornår evalueringen skulle finde sted. I alt blev der indgået aftaler med 20 folkeskoler og gymnasier, og 58 klasser deltog i undersøgelsen. Skolerne var primært placeret i København (seks), Aarhus (fem), Odense (fem), men skoler fra Aalborg, Køge, Randers og Silkeborg deltog også. Selvom data for undersøgelsen har en vis geografisk spredning og bred variation i forhold til skolernes sociodemografiske elevgrundlag, kan vores stikprøve ikke siges at være repræsentativ for danske unge i folkeskoler og på ungdomsuddannelser. Hertil er datagrundlaget for spinkelt, og selvselektionen i udvalgelsen af skoler medfører en risiko for bias. Således kan man forestille sig, at skoler i de større byer, som vælger at være med i afprøvningen af dialogværktøjet, huser elever, som har flere erfaringer med hadefulde memes og bystander-reaktioner end den gennemsnitlige danske skoleelev. Hvad dette ville betyde for vores estimater for anti- og prosociale bystander-reaktioner er dog uklart.

Undersøgelsens afhængige variable udgøres af et batteri af spørgsmål om respondents villighed til at engagere sig i forskellige handlinger som reaktion på hadefulde memes på sociale medier. Konkret blev eleverne bedt om at forestille sig ”at nedenstående meme dukker op i dit feed på dit fortrukne sociale medie”. Eleverne blev bedt om at forholde sig til seks forskellige memes. De anvendte memes kombinerede et billede og en dansk tekst og blev konstrueret af designbureauet Workz på vegne af C:NTACT. De konstruerede memes var skabt, så de i indhold og stil mindede om hadefulde memes, som faktisk cirkulerede på sociale medier, og således at der var variation i det ideologiske udgangspunkt og dermed i, hvem hadet var rettet imod. Puljen af 63 konstruerede memes indeholdt memes med hadefylde højreorienterede budskaber rettet mod

muslimer/islam, hadefulde venstreorienterede budskaber rettet mod højreorienterede/systemet og misogyne budskaber rettet mod kvinder. De anvendte seks memes er udvalgt med to memes repræsenterende hver ideologisk retning (to højreorienterede, to venstreorienterede og to misogyne). De seks udvalgte memes kan ses i artiklens appendiks. Valget om at anvende konstruerede memes og ikke eksisterende memes fra sociale medier udspringer af et ønske om at have et sample af memes, som i udtryk og virkemidler mindede om hinanden, og alle var på dansk. Desuden beror valget på at undgå at præsentere de unge for memes, som de i princippet kunne have stødt på før.

Efter hvert enkelt meme blev respondenterne bedt om at svare på, hvor uenige eller enige de var i, at de kunne finde på 1) ”at dele memet”, 2) ”at like memet” og 3) ”at anmelde memet for at overtræde platformens regelsæt”. Rækkefølgen af disse items blev randomiseret, og svarene blev givet på en syvpunkts likert-skala fra (1) ”helt enig” til (7) ”helt uenig” og efterfølgende reskaleret til at gå fra 0-1 for at kunne sammenligne koefficienterne. Svarfordelingerne for hvert enkelt meme fremgår af artiklens appendiks. Svarene på disse items udgør selvrapporateret handlingsvillighed som reaktion på et hypotetisk scenarie og kan derfor ikke sidestilles med, hvad respondenterne faktisk ville gøre, hvis de blev eksponeret for hadefulde memes i virkeligheden på sociale medier. Dette er en åbenlys metodisk begrænsning, som mindsker studiets økologiske validitet. Vi støtter os dog til, at en lang række socialpsykologiske studier har vist, at handlingsintentioner stabilt forudsiger faktiske fremtidige handlinger (Brandstätter og Opp, 2014), og vi argumenterer derfor for, at vores resultater har relevans for unges adfærd på sociale medier.

Umiddelbart efter at have svaret på undersøgelsens afhængige mål blev respondenterne bedt om for hvert af de seks memes at vurdere og svare på, hvor uenige eller enige de var i, at memet 1) ”er hadefuldt” og 2) ”indenfor grænserne af, hvad man kan acceptere”. Svarskalaen var igen en syvpunkts likert-skala fra (1) ”helt enig” til (7) ”helt uenig”. Rækkefølgen af disse items var randomiseret. Svarfordelingerne for hvert enkelt meme fremgår af henholdsvis figur A4 og figur A5 i appendiks.

Forud for visningen af de seks memes svarede respondenterne på en række baggrundsspørgsmål, herunder køn, alder og klassetrin. Undersøgelsens sample består af 53 pct. kvinder og 47 pct. mænd med en gennemsnitsalder på 15,2 år. Politisk interesse blev målt med et enkelt spørgsmål: ”Hvor interesseret er du i politik” med fire svarmuligheder: (1) ”meget interesseret”, (2) ”noget interesseret”, (3) ”lidt interesseret”, og (4) ”slet ikke interesseret” og efterfølgende reskaleret til at gå fra 0-1.

Vi bruger to mål fra litteraturen om politisk mobilisering og radikalisering til at undersøge, hvordan adfærdsmæssige intentioner korrelerer med pro- og antisocial adfærd i forhold til bystander-reaktioner på hadefulde memes (Moskalenko og McCauley, 2009; Gøtzsche-Astrup, 2018). *Lovlig politisk aktivisme* blev målt gennem et indeks bestående af fire spørgsmål, der blev omdannet til et refleksivt indeks ( $\alpha = 0.77$ ) og reskaleret til at gå fra 0-1. Deltagerne blev bedt om at forestille sig en organisation, der arbejder for deres politiske rettigheder og interesser, og erklære, hvor enige eller uenige de var i følgende udsagn: 1) ”Jeg ville følge en gruppe på de sociale medier, der kæmper for mine politiske rettigheder og interesser”, 2) ”Jeg ville donere penge til en organisation, der kæmper for mine politiske rettigheder og interesser”, 3) ”Jeg ville arbejde frivilligt (underskriftsindsamlinger, omdele brochurer eller rekrutterer andre) for en organisation, der kæmper for mine politiske rettigheder og interesser”, og 4) ”Jeg ville rejse én time for at deltage i en offentlig demonstration eller protest for at forsvare mine politiske rettigheder og interesser”. *Voldelig politisk aktivisme* blev målet gennem et lignende batteri, omdannet til et refleksivt indeks ( $\alpha = 0.88$ ) og reskaleret til at gå fra 0-1, med følgende syv spørgsmål: 1) ”Jeg ville være med til at true dem, der forsøgte at begrænse mine politiske rettigheder og interesser”, 2) ”Jeg ville deltage i en offentlig protest mod undertrykkelse af mine politiske rettigheder og interesser, selv hvis jeg troede, protesten kunne blive voldelig”, 3) ”Jeg ville angribe politiet eller sikkerhedsvagter, hvis jeg så dem overfalde en, som kæmpede for mine politiske rettigheder eller interesser”, 4) ”Jeg ville være med til at planlægge ulovlige handlinger mod autoriteter, der regelmæssigt undertrykte mine politiske rettigheder og interesser”, 5) ”Jeg ville opfordre andre til at deltage i voldelige protester mod undertrykkelse af mine politiske rettigheder og interesser, selv hvis jeg vidste, det var ulovligt”, 6) ”Jeg ville deltage i en offentlig demonstration eller protest for at forsvare mine politiske rettigheder og interesser, selv hvis den ikke var lovligt anmeldt til politiet”, og 7) ”Jeg ville fortsætte med at støtte en organisation, der kæmper for mine politiske rettigheder og interesser, selv hvis organisationen nogle gange tyede til vold”.

Til at teste vores hypoteser følger vi standardpraksis fra anerkendte studier af (antisocial) deling af indhold i onlinemiljøer (Pennycook et al., 2020; Pennycook et al., 2021; Roozenbeek, Freeman og van der Linden, 2021). Den primære analyseenhed i vores studie er hver elevs villighed til at reagere (pro- eller antisocialt) for hvert meme, og vi estimerer sammenhænge ved linerær regression med klyngerobuste standardfejl på individniveau.

## Resultater

Figur A1-A3 i appendikset viser svarfordelingerne for respondenternes villighed til at dele, like og rapportere de seks forskellige memes i undersøgelsen. Figurerne viser tydeligt, at der er lille variation i villigheden til henholdsvis at dele, like og anmelde på tværs af de seks memes. For alle tre typer reaktioner gælder det, at der er tale om højreskæve fordelinger, hvor flest respondenter således erklærer sig "helt uenige" eller "noget uenige" i, at de er villige til at dele, like eller anmelde de forskellige memes. Vi ser dog, at lidt flere respondenter erklærer sig villige til at anmelde de forskellige hadefulde memes end til at dele og like dem. Overordnet set viser dette, at danske unge formår at afkode det hadefulde indhold, selvom det er pakket ind i et humoristisk meme. Endelig erklærer flere respondenter sig "hverken enig eller uenig" i, at de vil anmelde memes, end i, at de vil dele og like. Dette kan indikere, at flere unge er i tvivl om, hvorvidt de ville være villige til at reagere prosocialt på de hadefulde memes og anmelde, mens mindre tvivl ses i forhold til at undlade at dele og like.

Figur 1 viser marginale effekter af de teoretiserede uafhængige variable på villigheden til at dele, like og anmelde hadefulde på tværs af alle memes. Den fulde model er yderligere vist i tabel A2 i appendikset. Vi evaluerer hypoteserne en efter en, hvorefter vi opsummerer de empiriske resultater nedenfor i tabel 1 og giver en samlet beskrivelse af resultaterne.

*Er danske unge mere villige til at reagere prosocialt (anmelde) og mindre villige til at reagere antisocialt (like, dele), når de opfatter memes som hadefulde?*

Ja. Når memes bliver opfattet som mere hadefulde, er de unge mere villige til at anmelde dem ( $\beta = 0.44$ ,  $SE = 0.03$ ). Med andre ord er villigheden til at reagere prosocialt større, når memes opfattes som hadefulde. Omvendt er villigheden til at dele ( $\beta = -0.13$ ,  $SE = 0.03$ ) og like ( $\beta = -0.21$ ,  $SE = 0.02$ ) memes mindre, når de opfattes som mere hadefulde. Dette støtter forventningerne i hypotese 1 og 2 om, at der er en sammenhæng mellem opfattelsen af memes som hadefulde og både prosocial og antisocial adfærd.

*Er danske unge mere villige til at reagere prosocialt (anmelde) og mindre villige til at reagere antisocialt (like, dele), når de opfatter memes som acceptable?*

Ja, i forhold til antisocial adfærd. Vi finder dermed støtte til hypotese 2: Når memes opfattes som acceptable, er villigheden til at reagere antisocialt i form af at dele ( $\beta = 0.23$ ,  $SE = 0.03$ ) eller like ( $\beta = 0.3$ ,  $SE = 0.03$ ) større. Omvendt

finder vi ikke støtte til hypotese 1, idet villigheden til at anmelde memes ikke er mindre, når de opfattes som acceptable ( $\beta = -0.02$ , SE = 0.03).

*Er piger mere villige til at reagere prosocialt og mindre villige til at reagere antisocialt?*

Ja. Vi finder dermed støtte til både hypotese 3 of 4, idet piger er mindre villige til at reagere antisocialt i form af at dele ( $B = -0.11$ , SE = 0.02) og like ( $B = -0.14$ , SE = 0.02) hadefulde memes og mere villige til at reagere prosocialt ved at anmelde memes ( $B = 0.05$ , SE = 0.02).

*Er der en sammenhæng mellem politisk interesse og prosocial adfærd?*

Nej. Unge, der tilkendegiver en høj politisk interesse, er ikke mere tilbøjelige til at anmelde hadefulde memes ( $\beta = -0.026$ , SE = 0.03), og dermed findes der ikke empirisk støtte til hypotese 5. Til gengæld er unge med højere politisk interesse hverken mere eller mindre tilbøjelige til at dele ( $\beta = -0.05$ , SE = 0.03) eller like ( $\beta = 0.02$ , SE = 0.03).

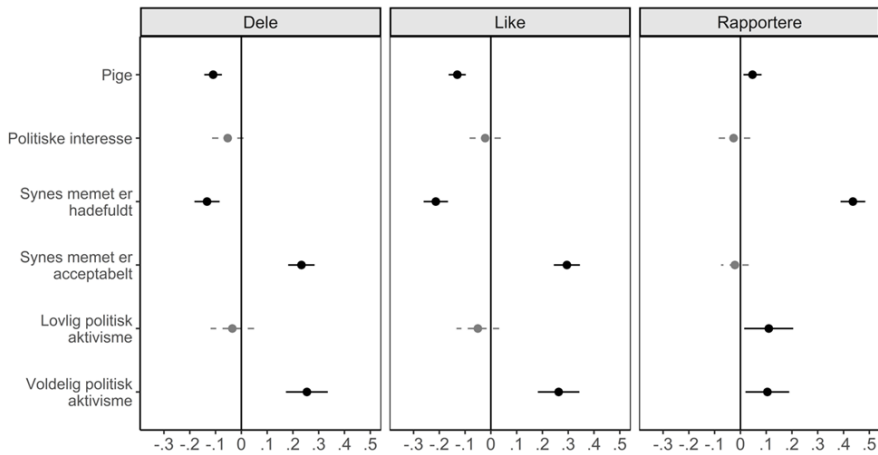
*Er der en sammenhæng mellem lovlige aktivistiske intentioner og villighed til at reagere prosocialt eller antisocialt?*

Ja og nej. Unge, der scorer højere på lovlig politisk aktivisme, er mere villige til at anmelde hadefulde memes ( $\beta = 0.1$ , SE = 0.05). Dermed finder vi støtte til hypotese 7. Til gengæld finder vi ikke støtte til hypotese 6, da de hverken er mere eller mindre villige til at reagere antisocialt i form af at dele ( $\beta = -0.04$ , SE = 0.04) eller like ( $\beta = -0.05$ , SE = 0.04).

*Er der en sammenhæng mellem voldelig politisk aktivisme og villighed til at reagere antisocialt?*

Ja. Der er en stærk positiv sammenhæng mellem voldelig politisk aktivisme og villigheden til at dele ( $\beta = 0.25$ , SE = 0.04) og like hadefulde memes ( $\beta = 0.26$ , SE = 0.04), hvilket støtter hypotese 8. Vi finder dog også en statistisk signifikant sammenhæng mellem voldelig politisk aktivisme og villigheden til at anmelde hadefulde memes ( $\beta = 0.11$ , SE = 0.04), men denne er sensitiv over for modelspecifikation. Ydermere er sammenhængen ikke statistisk signifikant i en bivariat sammenhæng mellem voldelig politisk aktivisme og villigheden til at anmelde (se tabel A5), og vi anser derfor ikke sammenhængen som robust.

Figur 1: Marginaleffekter på villighed til at dele, like og rapportere hadefulde memes



Note: Baseret på estimer fra tabel A2 med 95 pct. konfidensintervaller. Statistisk signifikante estimer er sorte, mens insignifikante estimer er grå med stiplede konfidensintervaller.

## Diskussion

Tabel 1 nedenfor opsummerer resultaterne. I det følgende diskuterer vi først de empiriske fund, herunder hvordan de relaterer sig til den eksisterende litteratur og deres praktiske implikationer. Dernæst diskuteres studiets begrænsninger og aspekter, som fremtidige studier med fordel kan undersøge.

Tabel 1: Overblik over og retning på sammenhænge

Oversigt	Antisocial		Prosocial
	Dele	Like	Anmelde
Pige	-	-	+
Politisk interesse			
Synes, det er hadefuldt	-	-	+
Synes, det er acceptabelt	+	+	
Lovlig politisk aktivisme			+
Voldelig politisk aktivisme	+	+	(+)

Note: Plus angiver en positiv sammenhæng, mens et minus angiver en negativ sammenhæng. Hvis tegnet er omsluttet af parentes, er sammenhængen sensitiv over for modelspecifikation, hvorfor estimatet skal fortolkes med varsomhed.

I denne artikel har vi vist, at der er en række faktorer, der forudsiger anti- og prosocial adfærd i form af at dele, like og anmelde hadefulde memes blandt cyber-bystanders på sociale medier. Overordnet viste det sig, at opfattelsen af hadefulde memes som acceptabel og voldelig politisk aktivisme forudsiger antisocial bystander-adfærd, mens køn og opfattelsen af memes som hadefulde forudsiger både anti- og prosocial bystander-adfærd. Lovlig politisk aktivisme forudsiger alene prosocial bystander-adfærd.

I forhold til spørgsmålet om, hvilken rolle opfattelsen af et memes budskab spiller for bystander-adfærd, viser resultaterne, at opfattelsen af, hvorvidt et meme var hadefuldt, har en sammenhæng med både pro- og antisocial adfærd i de forventede retninger. I hvilken udstrækning et meme, som opfattes som mere acceptabelt, har en sammenhæng med antisociale reaktioner, så viste resultaterne, at hadefulde memes, som vurderes som acceptable, både deles og likes mere. Tilsammen støtter disse fund op om, at danske unge i undersøgelsen vurderer memes på baggrund af deres indhold, og at der er en sammenhæng mellem denne vurdering af indholdet og de unges villighed til at interagere med hadefulde memes i form af at dele, like og anmelde dem. Disse resultater støtter op om den eksisterende litteratur om bystanders, der viser, at opfattelsen af, hvor alvorlig eller kritisk en hændelse er, har afgørende betydning for individers villighed til at handle prosocialt (Fischer et al., 2006). Ligeså har studier vist, at folks villighed til at indskrænke ytringsfriheden for hadefulde ytringer påvirkes af grovheden af udsagnet, herunder om udsagnet indeholder trusler (Rasmussen, 2021).

Samlet set tyder disse resultater på, at de unge danske brugere af sociale medier i vores undersøgelse rent faktisk forholder sig til indholdet i de opslag, de eksponeres for, og agerer i overensstemmelse hermed. Opfattelser af indhold som normativt afvigende synes at aktivere korrektive, prosociale reaktioner (Wilhelm, Joeckel og Ziegler, 2020). Sammenholdt med, at de unge i undersøgelsen formår at afkode hadefuldt indhold i humoristiske memes, efterlader dette rum for en vis optimisme. Ved aktivt at diskutere normer for online adfærd og grænserne for, hvad som kan/bør siges på sociale medier, kan unges bevidsthed om hadefuldt indhold skærpes, og prosociale reaktioner dermed understøttes.

Vi finder desuden en sammenhæng mellem køn og onlineadfærdsintentioner, der viser, at piger er mindre villige til at agere antisocialt og mere villige til at agere prosocialt. Dette resultat støtter tidligere studier, der har påvist, at drenge er mere tilbøjelige til at deltage i antisocial adfærd, herunder mobning på sociale medier (fx Bastiaensens et al., 2014). Det falder udenfor rammerne af denne undersøgelse at afgøre, om denne kønsforskel bedst lader sig forklare med re-

ference til eksempelvis evolutionær udvikling eller forskelle i socialisering og normdannelse blandt piger og drenge. Men hvis vi ønsker at hæmme antisociale cyber-bystander-reaktioner og fremme de prosociale reaktioner, er dette fund vigtigt for tilrettelæggelsen af interventioner. Dette gælder både, når det drejer sig om afgrænsning af målgrupper og udformning af interventionens indhold.

Vi mener, at bystander-reaktioner på hadefulde, politisk motiverede memes kan anses som en form for politisk deltagelse. Ved at dele, like eller anmelde indhold giver man henholdsvis sin billigelse eller misbilligelse til kende. Derfor er det også overraskende, og imod vores forventning (hypotese 5), at politisk interesse ikke synes at spille nogen rolle for online bystander-adfærd.

Disse resultater skal ses i lyset af, at lovlig politisk aktivisme forudsiger villigheden til at anmelde hadefulde memes, mens voldelig politisk aktivisme forudsiger villigheden til at like og dele hadefulde memes. Dette resultat flugter med tidligere studier, som peger på, at folk, der udviser aggressiv politisk adfærd offline, også gør det online (Bor og Petersen, 2021). Konkret ser vi, at de personer, der er villige til at gøre brug af lovlig politisk aktivisme offline, er mere tilbøjelige til at udøve prosocial adfærd online, mens de, der er villige til at gøre brug af voldelig politisk adfærd offline, er mere villige til at udøve antisocial adfærd online. Noget tyder således på, at de psykologiske motivationer, der driver forskellige former for offline politisk deltagelse, delvis spejles online og driver forskellige former for bystander-reaktioner. De, som deltager i politik for at hævde sig og dominere – ligegyldigt midlerne – gør det både offline og online, ligesom de, som primært deltager for at fremme en given sag og bekæmpe uretfærdighed, også gør det både offline og online.

### *Begrænsninger og fremtidige studier*

Vi identificerer her fire relevante områder, som fremtidige studier med fordel kan adressere. For det første er der behov at undersøge, hvilken betydning indholdsmæssige og kontekstuelle faktorer har for cyber-bystanders opfattelse af politiske budskaber og i forlængelse heraf deres reaktioner. I vores studie har vi ikke været i stand til at isolere effekten af, fx *hvem* et hadefuldt meme er rettet imod, eller variationen over politiske emner. Ud over de indholdsmæssige faktorer er det helt oplagt for fremtidige studier at undersøge effekten af kontekstuelle faktorer, eksempelvis anonymitet og størrelse af publikum i eksperimentelle designs. Selvom eksisterende studier af cybermobning har bekræftet en generel tendens til bystander-passivitet under anonymitet på sociale medier, og i takt med at publikum vokser (Bastiaensens et al., 2014; Dillon og Bushman, 2015), så er der stadig behov for at undersøge dette i forhold til *politisk* motiveret had.



For det andet bør konsekvenserne af særligt prosocial bystander-adfærd på sociale medier undersøges. Mens der er en stor litteratur, der dokumenterer de negative sociale og psykologiske konsekvenser af eksponering til hadefuldt indhold (Hyers, 2007; Brondolo, Gallo og Myers, 2009), er der en mangel på studier, der undersøger, hvordan eksponering til andres prosociale adfærd påvirker ens egen tilbøjelighed til at reagere prosocialt i fremtiden. Et centralt spørgsmål er således, om eksponering til prosociale bystander-reaktioner på sociale medier på kort sigt kan øge sandsynligheden for flere sådanne reaktioner ved at reducere tvivlen om, hvorvidt der er tale om en ”kritisk hændelse”, som kræver handling, og ved at sikre, at man ikke står alene med sin prosociale intervention. På længere sigt er spørgsmålet, om det at se andre reagere prosocialt på hadefuldt indhold på nettet, kan være med til at styrke normer om ikkediskrimination og civilitet på sociale medier (Paluck, 2011; Czopp, Monteith og Mark, 2006). Det sidste vil kræve longitudinale designs (se fx Munger, 2017; Siegel og Badaan, 2020).

For det tredje er der et behov for i fremtidige studier at inkludere mere fint graduerede mål for prosociale bystander-reaktioner. Dette studie har begrænset sig til at undersøge, hvilke faktorer prædikerer villighed til at anmelde hadefulde memes. Som antydnet indledningsvis, udgør denne form for teknisk prosocial reaktion en blandt flere, herunder mere omkostningsfulde interventioner som counter-speech rettet mod afsenderen af det hadefulde indhold og reaktioner, som henvender sig til offeret og tilbyder støtte og hjælp. Eksisterende studier af counter-speech peger på muligheden for at benytte forskellige strategier så som at henvise til korrigerende fakta, fremhæve inkonsistent argumentation, anvende humor, appellere til overordnede fælles identiteter og benævne indhold som ”hadefuldt”, ”sexistisk”, ”racistisk” m.m. (Hangartner et al., 2021; Benesch et al., 2016; Buerger, 2020). Det er en åbenlys begrænsning ved vores studie, at vi ikke indfanger disse forskellige mulige prosociale reaktioner og dermed ikke er i stand til at udtale os om, hvorvidt forskellige prosociale bystander-reaktioner prædikerer af forskellige eller de samme faktorer.

For det fjerde bør fremtidige studier eksplicit adressere, hvordan prosocial adfærd som respons på hadefuldt indhold på sociale medier kan fremmes. Mens der er en voksende litteratur om interventioner, der sigter mod at mindske *hate speech* og spredning af falske nyheder (Pennycook et al., 2020; Pennycook et al., 2021; Roozenbeek, Freeman og van der Linden, 2021; Rasmussen, Lindekilde og Petersen, 2021), mangler der i den grad viden om, hvorvidt prosocial bystander-adfærd kan fremmes via undervisning i fx kritisk tænkning, nudges eller via den måde, sociale medieplatforme designes på (for undtagelser se DiFranzo et al., 2018; Kleinsasser et al., 2015). Eksempelvis benytter en platform som

Reddit *crowd moderation* og muligheden for *down-voting* af indhold, hvilket kan fremme prosociale reaktioner på hadefuldt indhold. Imidlertid eksisterer muligheden for at *up-vote* hadefuldt indhold også, ligesom *down-voting* kan misbruges til at gemme indhold, man er politisk uenig med. I dette studie har vi begrænset os til at diskutere implikationerne af vores resultater for interventioner. Fremtidige studier bør lave effektevalueringer af sådanne indsatser.

Foruden de ovennævnte begrænsninger af vores undersøgelse er det værd at fremhæve, at selvrapporeret villighed til forskellige handlinger selvsagt ikke er det samme som faktiske reaktioner på konkrete hadefulde memes i ens sociale mediefeed. Vi studerer således handlingsintentioner og ikke adfærd. Når vi alligevel argumenterer for, at vores undersøgelse siger noget om, hvordan danske unge kan forventes at reagere på virkeligt hadefuldt indhold på sociale medier, og resultaterne dermed har praktisk relevans, så skyldes det primært, at en stor socialpsykologisk forskning viser, at intentioner udgør en robust prædikator for faktisk adfærd (Webb og Sheeran, 2006). Når det er sagt, så kan fremtidige studier med fordel benytte sig af fx simulerede sociale medieplatforme til at studere faktiske bystander-reaktioner mere direkte (se fx DiFranzo et al., 2018).

Endvidere kan der peges på en risiko for *social desirability bias* i vores data (Krumpal, 2013). Man må forvente, at adspurgt om hypotetiske intentioner om henholdsvis at dele/like og anmelde hadefulde memes vil respondenter tendere mod at underrapportere antisocial adfærd og overrapportere prosocial adfærd. Mens vi ikke kan afvise, at intentioner om at dele og like hadefuldt indhold er underestimeret i vores undersøgelse, er der ikke meget som tyder på, at prosociale reaktioner overestimeres, idet få respondenter erklærer sig enige i, at de ville anmelde de forskellige memes (se figur A3 i appendikset). Igen vil fremtidige studier med fordel kunne benytte sig af kvasifeltksp eksperimenter på simulerede sociale medieplatforme for at minimere risikoen for social desirability bias.

Endelig kan man sætte spørgsmålstegn ved undersøgelsens eksterne validitet givet dens begrænsede fokus på danske unge. Sammenlignet med en række andre lande (fx USA og Storbritannien) er Danmark karakteriseret ved en relativt lav grad af politisk og affektiv polarisering, og dermed vil vores studie i Danmark underestimere graden af både pro- og antisociale bystander-reaktioner. Vi forventer, at polarisering vil føre til stærkere antisociale reaktioner på hadefuldt indhold rettet mod ens politiske udgruppe og stærkere prosociale reaktioner på hadefuldt indhold rettet mod ens politiske indgruppe. Dog formoder vi, at disse kontekstuelle forskelle vil være mindre blandt unge (som i vores sample) end blandt voksne.

## Konklusion

Vi satte os i denne artikel for at undersøge, hvordan danske unge vurderer og reagerer på hadefulde budskaber, når de er pakket ind i et humoristisk meme på sociale medier. Specifikt var vi interesseret i at undersøge, hvilke faktorer øger sandsynligheden for henholdsvis antisociale og prosociale reaktioner, når unge er vidner til politisk had på sociale medier. Med afsæt i eksisterende forskning og teori om bystander-adfærd formulerede vi en række hypoteser om, hvordan danske unges vurdering af hadefulde memes, køn, politisk interesse og henholdsvis voldelig og lovlig politisk aktivisme medformer de unges bystander-reaktioner i forhold til deres villighed til at like, dele (antisocialt) og anmelde (prosocialt) indholdet.

Resultaterne viser, at selvom danske unge i undersøgelsen generelt udviser en lav grad af villighed til både at dele, like og anmelde hadefulde memes – hvilket kan tolkes som udtryk for en tendens til at reagere med passivitet i overensstemmelse med bystander-effekten – så spiller de unges vurdering af, hvor hadefuldt og uacceptabelt et meme er, en vigtig rolle for bystander-reaktioner. Således finder vi, at snarere end at de unge i undersøgelsen ikke formår at afkode hadefulde budskaber indeholdt i humoristiske memes, så vurderer de indholdets karakter og reagerer i overensstemmelse hermed. Jo mere hadefuldt og uacceptabelt et meme vurderes at være, jo mindre sandsynligt er det, at de unge reagerer ved at like og dele, og jo mere sandsynligt er det, at de reagerer prosocialt ved at anmelde indholdet. Dette resultat er i overensstemmelse med eksisterende studier, som fremhæver vurderingen af, hvor ”kritisk” en hændelse er for beslutningen om prosocial intervention.

Resultaterne viste ikke tegn på, at politisk interesse spiller en systematisk rolle for valget mellem forskellige bystander-reaktioner. Antisociale reaktioner (dele, like) er mere udbredte blandt drenge og personer, der i højere grad tilslutter sig voldelig politisk aktivisme. Foruden vurderingen af et budskabs indhold så afhænger prosocial anmeldelse af hadefulde budskaber af ens tilbøjelighed til at deltage i lovlig politisk aktivisme: Personer, som er mere tilbøjelige til at deltage i lovlig politisk aktivisme, er mere tilbøjelige til at anmelde hadefulde memes. Disse fund ligger i forlængelse af nyere studier, der netop viser, at folks adfærd online reflekterer deres adfærd offline. Fremtidige studier bør i den sammenhæng undersøge, hvordan man kan udnytte folks tendens til at agere prosocialt offline i onlinekontekster.

Samlet set har denne artikel bidraget ved at præsentere det første studie af danske unges vurdering af og reaktioner på politisk hadefulde memes. Artiklens resultater har bidraget med basale indsigter om, hvilke faktorer som driver og differentierer mellem prosociale og antisociale bystander-reaktioner på poli-

tisk hadefulde budskaber på sociale medier. Hermed bidrager artiklen med at forbinde hidtil adskilte empiriske indsigter i socialpsykologien om bystander-adfærd i ikkepolitiske situationer med indsigter om politisk had på sociale medier. På trods af studiets begrænsninger mener vi at have illustreret frugtbarheden af en ny forskningsdagsorden, som ikke bare analyserer, hvad som driver politisk had på nettet, men som vender sig mod bystander-figuren, og hvordan vi som almindelige borgere reagerer på sådant indhold. Set fra et demokratisk perspektiv kan en sådan forskningsdagsorden være med til at forløse det uudnyttede potentiale, der ligger i prosociale bystander-reaktioner til at modgå de negative demokratiske konsekvenser af online politisk had.

## Noter

1. Forskningen, som denne artikel bygger på, er delvist realiseret inden for STANDBY-projektet, finansieret af European Research Consortium (ERC), grant agreement number: 101002251.

## Appendiks

Appendiks findes efter litteraturlisten.

## Litteratur

- Arceneaux, Kevin, Timothy B. Gravelle, Matthias Osmundsen, Michael Bang Petersen, Jason Reifler og Thomas J. Scotto (2021). Some people just want to watch the world burn: The prevalence, psychology and politics of the "need for chaos". *Philosophical Transactions. Biological Sciences* 376 (1822): 20200147-20200147.
- Askanius, Tina (2021). On frogs, monkeys, and execution memes: Exploring the humor-hate nexus at the intersection of neo-Nazi and Alt-Right movements in Sweden. *Television & New Media* 22 (2): 147-165.
- Bastiaensens, Sara, Heidi Vandebosch, Karolien Poels, Katrien Van Cleemput, Ann DeSmet og Ilse De Bourdeaudhuij (2014). Cyberbullying on social network sites: An experimental study into bystanders' behavioural intentions to help the victim or reinforce the bully. *Computers in Human Behavior* 31: 259-271.
- Benesch, Susan, Derek Ruths, Kelly P. Dillon, Haji Mohammad Saleem og Lucas Wright (2016). *Considerations for successful counterspeech*. Dangerous Speech Project. <https://dangerousspeech.org/wp-content/uploads/2016/10/Considerations-for-Successful-Counterspeech.pdf> (25. november, 2021)
- Bor, Alexander og Michael Bang Petersen (2021). The psychology of online political hostility: A comprehensive, cross-national test of the mismatch hypothesis. *The American Political Science Review* 116 (1): 1-18.

- Brandstätter, Hermann og Karl-Dieter Opp (2014). Personality traits (“Big Five”) and the propensity to political protest: Alternative models. *Political Psychology* 35 (4): 515-537.
- Brewer, Gayle og Jade Kerslake (2015). Cyberbullying, self-esteem, empathy and loneliness. *Computers in Human Behavior* 48: 255-260.
- Brondolo, Elizabeth, Linda C. Gallo og Hector F. Myers (2009). Race, racism and health: Disparities, mechanisms, and interventions. *Journal of Behavioral Medicine* 32 (1): 1-8.
- Buerger, Catherine (2020). The anti-hate brigade: How a group of thousands responds collectively to online vitriol. *SSRN 3748803*.
- Cheng, Justin, Michael Bernstein, Cristian Danescu-Niculescu-Mizil og Jure Leskovec (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, Portland, Oregon, pp. 1217–1230.
- Citron, Danielle Keats og Helen Norton (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review* 91 (4): 1435.
- Cramer, Robert Ervin, M. Rosalie McMaster, Patricia A. Bartell og Marguerite Dragna (1988). Subject competence and minimization of the bystander effect. *Journal of Applied Social Psychology* 18 (13): 1133-1148.
- Czopp, Alexander M., Margo J. Monteith og Aimee Y. Mark (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology* 90 (5): 784-803.
- Darley, John M. og Bibb Latane (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology* 8 (4p1): 377-383.
- Darley, John M. og Bibb Latane (1970). Norms and normative behavior: Field studies of social interdependence. *Altruism and Helping Behavior*: 83-102.
- DiFranzo, Dominic, Samuel Taylor, Francesca Kazerooni, Olivia Wherry og Natalya Bazarova (2018). Upstanding by design: Bystander intervention in cyberbullying. *Conference on Human Factors in Computing Systems*, pp. 1-12.
- Dillon, Kelly P. og Brad J. Bushman (2015). Unresponsive or un-noticed?: Cyberbystander intervention in an experimental cyberbullying context. *Computers in Human Behavior* 45: 144-150.
- Duggan, Maeve (2017a). Experiencing online harassment. Pew Research Center, July 11. <https://www.pewresearch.org/internet/2017/07/11/experiencing-online-harassment/>
- Duggan, Maeve (2017b). Witnessing online harassment. Pew Research Center, July 11. <https://www.pewresearch.org/internet/2017/07/11/witnessing-online-harassment/>

- Fischer, Peter, Tobias Greitemeyer, Fabian Pollozek og Dieter Frey (2006). The unresponsive bystander: are bystanders more responsive in dangerous emergencies? *European Journal of Social Psychology* 36 (2): 267-278.
- Fischer, Peter, Joachim I. Krueger, Tobias Greitemeyer, Claudia Vogrincic, Andreas Kastenmüller, Dieter Frey, Moritz Heene, Magdalena Wicher og Martina Kainbacher (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin* 137 (4): 517-537.
- Freis, Stephanie D. og Regan A. R. Gurung (2013). A Facebook analysis of helping behavior in online bullying. *Psychology of Popular Media Culture* 2 (1): 11-19.
- Greitemeyer, Tobias, Peter Fischer, Andreas Kastenmüller og Dieter Frey (2006). Civil courage and helping behavior: Differences and similarities. *European Psychologist* 11 (2): 90-98.
- Greitemeyer, Tobias, Silvia Osswald, Peter Fischer og Dieter Frey (2007). Civil courage: Implicit theories, related concepts, and measurement. *The Journal of Positive Psychology* 2 (2): 115-119.
- Gøtzsche-Astrup, Oluf (2018). The time for causal designs: Review and evaluation of empirical support for mechanisms of political radicalisation. *Aggression and Violent Behavior* 39: 90-99.
- Hangartner, Dominik, Gloria Gennaro, Sary Alasiri, Nicholas Bahrnich, Alexandra Bornhoft, Joseph Boucher, Buket Buse Demirci, Laurenz Derksen, Aldo Hall, Matthias Jochum, Maria Murias Munoz, Marc Richter, Franziska Vogel, Salomé Wittwer, Felix Wüthrich, Fabrizio Gilardi og Karsten Donnay (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences – PNAS* 118 (50): 1.
- Hyers, Lauri L. (2007). Resisting prejudice every day: Exploring women's assertive responses to anti-black racism, anti-Semitism, heterosexism, and sexism. *Sex Roles* 56 (1): 1-12.
- Iganski, Paul (2020). Civil courage as a communicative act: Countering the harms of hate violence. *Pragmatics and Society* 11 (2): 316-335.
- Kleinsasser, Anne, Ernest N. Jouriles, Renee McDonald og David Rosenfield (2015). An online bystander intervention program for the prevention of sexual violence. *Psychology of Violence* 5 (3): 227-235.
- Krumpal, Ivar (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity* 47 (4): 2025-2047.
- Kunst, Marlene, Pablo Porten-Cheé, Martin Emmer og Christiane Eilders (2021). Do "good citizens" fight hate speech online? Effects of solidarity citizenship norms on user responses to hate comments. *Journal of Information Technology & Politics* 18 (3): 258-273.

- Latane, Bibb og Steve Nida (1981). Ten years of research on group size and helping. *Psychological Bulletin* 89 (2): 308-324.
- Leonhard, Larissa., Cristina Rueß, Magdalena Obermaier, og Carsten Reinemann. (2018). Perceiving threat and feeling responsible. How severity of hate speech, number of bystanders, and prior reactions of others affect bystanders' intention to counterargue against hate speech on Facebook. *SCM Studies in Communication and Media* 7 (4): 555-579.
- Liebst, Lasse Suonperä, Marie B. Heinskou og Peter Ejbye-Ernst (2018). On the actual risk of bystander intervention: A statistical study based on naturally occurring violent emergencies. *The Journal of Research in Crime and Delinquency* 55 (1): 27-50.
- Liebst, Lasse S., Richard Philpot, Marie B. Heinskou og Marie R. Lindegaard (2019). Bystander intervention in street violence: Current evidence and implications for practice. SocArXiv. <https://osf.io/preprints/socarxiv/7m9uv/>
- Liebst, Lasse Suonperä, Peter Ejbye-Ernst, Kasper Lykke Dausel, Peter Verbeek, Richard Philpot, Wim Bernasco, Marie Bruvik Heinskou og Marie Rosenkrantz Lindegaard (2018). Bystanders in real-life dangerous emergencies: Group relationships predict intervention. *The Danish Sociological Conference 2018*, Esbjerg, Denmark.
- Machackova, Hana (2020). Bystander reactions to cyberbullying and cyberaggression: Individual, contextual, and social factors. *Current Opinion in Psychology* 36: 130-134.
- Mackay, Yasmin (2019). *Teens who intervene: Identifying factors related to adolescent cyber-bystander intervention in cyberbullying*. Canterbury: Christ Church University.
- McDougall, Julian, Marketa Zezulcova, Barry Van Driel og Dalibor Sternadel (2018). Teaching media literacy in Europe: Evidence of effective school practices in primary and secondary education. *NESET II Report*. Luxembourg: Publications Office of the European Union.
- Moskalenko, Sophia og Clark McCauley (2009). Measuring political mobilization: The distinction between activism and radicalism. *Terrorism and Political Violence* 21 (2): 239-260.
- Moxey, Natasha og Kay Bussey (2020). Styles of bystander intervention in cyberbullying incidents. *International Journal of Bullying Prevention* 2 (1): 6-15.
- Mulvey, Kelly Lynn, Sally B. Palmer og Dominic Abrams (2016). Race-based humor and peer group dynamics in adolescence: Bystander intervention and social exclusion. *Child Development* 87 (5): 1379-1391.
- Munger, Kevin (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior* 39 (3): 629-649.
- Mutz, Diana C. (2015). *In-your-face politics: The consequences of uncivil media*. Princeton: Princeton University Press.

- Naab, Teresa K., Anja Kalch og Tino G. K. Meitz (2018). Flagging uncivil user comments: Effects of intervention information, type of victim, and response comments on bystander behavior. *New Media & Society* 20 (2): 777-795.
- Ndoen, Maria Cherry Rondang Cattleya (2018). Political meme in Instagram: How online humour became a political weapon. *Language in the Online & Offline World 6: The Fortitude*: 186.
- Obermaier, Magdalena, Nayla Fawzi og Thomas Koch (2016). Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying. *New Media & Society* 18 (8): 1491-1507.
- Paluck, Elizabeth Levy (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *Journal of Experimental Social Psychology* 47 (2): 350-358.
- Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu og David G. Rand (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science* 31 (7): 770-780.
- Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles og David G. Rand (2021). Shifting attention to accuracy can reduce misinformation online. *Nature* 592 (7855): 590-595.
- Petersen, Michael Bang, Mathias Osmundsen og Alexander Bor (2021). The psychology of status-seeking and extreme political discontent, pp. 62-80 i Joseph P. Forgas, William D. Crano og Klaus Fiedler (red.), *The psychology of populism: The tribal challenge to liberal democracy*. New York: Routledge.
- Porten-Chee, Pablo, Marlene Kunst og Martin Emmer (2020). Online civic intervention: A new form of political participation under conditions of a disruptive online discourse. *International Journal of Communication* 14: 21.
- Price, Deborah, Deborah Green, Barbara Spears, Margaret Scrimgeour, Alan Barnes, Ruth Geer og Bruce Johnson (2014). A qualitative exploration of cyber-bystanders and moral engagement. *Australian Journal of Guidance and Counselling* 24 (1): 1-17.
- Quirk, Roslynn og Marilyn Campbell (2015). On standby? A comparison of online and offline witnesses to bullying and their bystander behaviour. *Educational Psychology (Dorchester-on-Thames)* 35 (4): 430-448.
- Rains, Stephen A., Kate Kenski, Kevin Coe og Jake Harwood (2017). Incivility and political identity on the internet: Intergroup factors as predictors of incivility in discussions of news online. *Journal of Computer-Mediated Communication* 22 (4): 163-178.
- Rasmussen, Jesper (2021). The online hate speech consensus: Severity shapes Americans' and Danes' preferences for restricting online hate speech. Upubliceret materiale. [https://github.com/jesper-rasmussen/jesper-rasmussen.github.io/raw/main/jr\\_hatespeech.pdf](https://github.com/jesper-rasmussen/jesper-rasmussen.github.io/raw/main/jr_hatespeech.pdf)



- Rasmussen, Jesper, Lasse Lindekilde og Michael Bang Petersen (2021). The effectiveness of accuracy-nudge interventions on intentions to share COVID-19 misinformation on social media. <https://osf.io/akybg>
- Rasmussen, Stig Hebbelstrup Rye, Alexander Bor, Mathias Osmundsen og Michael Bang Petersen (2021). Super-unsupervised text classification for labeling online political hate. *PsyArXiv*.
- Reijntjes, Albert, Marjolijn Vermande, Tjeert Olthof, Frits A. Goossens, Liesbeth Aleva og Matty van der Meulen (2016). Defending victimized peers: Opposing the bully, supporting the victim, or both?: Defending victims of peer bullying. *Aggressive Behavior* 42 (6): 585-597.
- Roozenbeek, Jon, Alexandra L. J. Freeman og Sander van der Linden (2021). How accurate are accuracy-nudge interventions? A preregistered direct replication of Pennycook et al. (2020). *Psychological Science* 32 (7): 1169-1178.
- Siegel, Alexandra A. og Vivienne Badaan (2020). No2Sectarianism: Experimental approaches to reducing sectarian hate speech online. *The American Political Science Review* 114 (3): 837-855.
- van den Bos, Kees, Patrick A. Müller og Anke A. L. van Bussel (2009). Helping to overcome intervention inertia in bystander's dilemmas: Behavioral disinhibition can improve the greater good. *Journal of Experimental Social Psychology* 45 (4): 873-878.
- Webb, Thomas L. og Paschal Sheeran (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin* 132 (2): 249-268.
- Wilhelm, Claudia, Sven Joeckel og Isabell Ziegler (2020). Reporting hate comments: Investigating the effects of deviance characteristics, neutralization strategies, and users' moral orientation. *Communication Research* 47 (6): 921-944.

# Appendiks

*Tabel A1:* Deskriptiv statistik

Variabel	Gns.	Std.afv.	25. percentil	Median	75. percentil	Min	Max	Observationer
Dele	0.24	0.25	0.00	0.17	0.42	0.00	1.00	751
Like	0.26	0.26	0.03	0.19	0.44	0.00	1.00	751
Rapportere	0.33	0.25	0.11	0.33	0.50	0.00	1.00	751
Pige	0.53	0.50	0.00	1.00	1.00	0.00	1.00	829
Alder (% > 14)	0.54	0.50	0.00	1.00	1.00	0.00	1.00	845
Politisk interesse	0.43	0.28	0.33	0.33	0.67	0.00	1.00	848
Synes, memet er hadefuldt	0.56	0.24	0.42	0.58	0.72	0.00	1.00	751
Synes, memet er acceptabelt	0.44	0.24	0.28	0.44	0.58	0.00	1.00	751
Lovlig politisk aktivisme	0.49	0.21	0.38	0.50	0.62	0.00	1.00	736
Voldelig politisk aktivisme	0.32	0.21	0.14	0.32	0.50	0.00	1.00	730

*Table A2:* OLS Regression: resultater på tværs af villighed til at 1) dele, 2) like, 3) rapportere

	Afhængig variabel		
	Dele	Like	Rapportere
Pige	-0.109*** (0.017)	-0.130*** (0.017)	0.047** (0.018)
Politisk interesse	-0.053 (0.031)	-0.022 (0.031)	-0.026 (0.033)
Synes, memet er hadefuldt	-0.133*** (0.025)	-0.214*** (0.024)	0.436*** (0.024)
Synes, memet er acceptabelt	0.233*** (0.026)	0.295*** (0.026)	-0.021 (0.027)
Lovlig politisk aktivisme	-0.035 (0.043)	-0.050 (0.042)	0.110* (0.048)
Voldelig politisk aktivisme	0.254*** (0.041)	0.263*** (0.041)	0.105* (0.043)
Konstant	0.216*** (0.037)	0.265*** (0.036)	-0.004 (0.033)
Observationer	4192	4191	4190
R2	0.17	0.25	0.23

Note: Alder (0 = yngre end 14, 1 = 14 eller ældre) er medtaget som kontrolvariabel. Klyngerobuste standardfejl på respondent niveau. \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001.

Table A3: OLS Regression: villighed til at dele

	Afhængig variabel: villighed til at dele							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pige	-0.146*** (0.019)							-0.109*** (0.017)
Alder (> 14)		0.052** (0.018)						0.014 (0.017)
Politisk interesse			-0.096** (0.033)					-0.053 (0.031)
Synes, memet er hadfuldt				-0.187*** (0.026)				-0.133*** (0.025)
Synes, memet er acceptabelt					0.270*** (0.027)			0.233*** (0.026)
Lovlig politisk aktivisme						-0.059 (0.044)		-0.035 (0.043)
Voldelig politisk aktivisme							0.299*** (0.045)	0.254*** (0.041)
Konstant	0.316*** (0.015)	0.207*** (0.013)	0.278*** (0.017)	0.342*** (0.018)	0.118*** (0.013)	0.265*** (0.024)	0.140*** (0.016)	0.216*** (0.037)
Observationer	4344	4425	4455	4435	4432	4389	4353	4192
R2	0.05	0.01	0.01	0.04	0.07	0.00	0.04	0.17

Note: Klynge robuste standardfejl på respondentniveau. \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001. Alder (0 = yngre end 14, 1 = 14 eller ældre) er medtaget i denne model, men fungerer som kontrolvariabel i hovedmodellen.

*Table A4:* OLS Regression: villighed til at like

	Afhængig variabel: villighed til at like							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pige	-0.178*** (0.018)							-0.130*** (0.017)
Alder (> 14)		0.057** (0.019)						0.010 (0.016)
Politisk interesse			-0.076* (0.034)					-0.022 (0.031)
Synes, memet er hadfuldt				-0.272*** (0.026)				-0.214*** (0.024)
Synes, memet er acceptabelt					0.333*** (0.028)			0.295*** (0.026)
Lovlig politisk aktivisme						-0.090* (0.045)		-0.050 (0.042)
Voldelig politisk aktivisme							0.317*** (0.047)	0.263*** (0.041)
Konstant	0.359*** (0.015)	0.229*** (0.013)	0.294*** (0.018)	0.414*** (0.019)	0.115*** (0.013)	0.305*** (0.025)	0.160*** (0.016)	0.265*** (0.036)
Observationer	4340	4421	4451	4435	4430	4388	4352	4191
R2	0.07	0.01	0.00	0.07	0.11	0.00	0.04	0.25

Note: Klynge robuste standardfejl på respondentniveau. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . Alder (0 = yngre end 14, 1 = 14 eller ældre) er medtaget i denne model, men fungerer som kontrolvariabel i hovedmodellen.

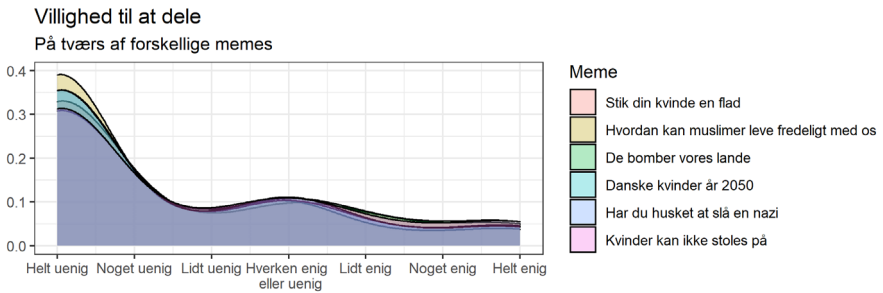
Table A5: OLS Regression: villighed til at rapportere

	Afhængig variabel: villighed til at rapportere							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pige	0.079*** (0.018)							0.047*** (0.018)
Alder (> 14)		-0.033 (0.018)						0.002 (0.017)
Politisk interesse			0.036 (0.034)					-0.026 (0.033)
Synes, memet er hadefuld				0.444*** (0.023)				0.436*** (0.024)
Synes, memet er acceptabelt					-0.046 (0.033)			-0.021 (0.027)
Lovlig politisk aktivisme						0.189*** (0.049)		0.110* (0.048)
Voldelig politisk aktivisme							0.052 (0.047)	0.105* (0.043)
Konstant	0.291*** (0.014)	0.350*** (0.013)	0.318*** (0.017)	0.085*** (0.011)	0.353*** (0.017)	0.240*** (0.025)	0.315*** (0.018)	-0.004 (0.033)
Observationer	4334	4416	4446	4432	4430	4384	4348	4190
R2	0.01	0.00	0.00	0.21	0.00	0.02	0.00	0.23

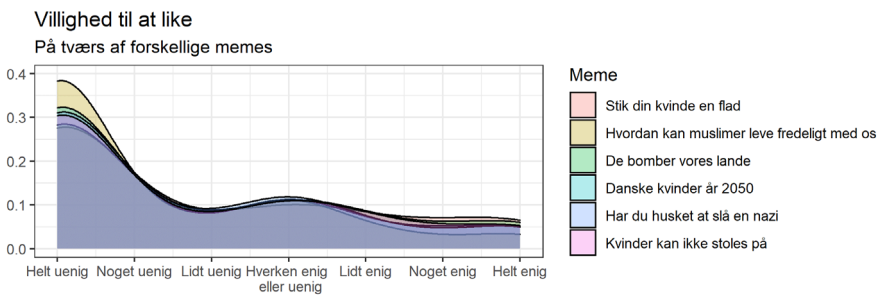
Note: Klynge robuste standardfejl på respondentniveau. \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001. Alder (0 = yngre end 14, 1 = 14 eller ældre) er medtaget i denne model, men fungerer som kontrolvariabel i hovedmodellen.

# Fordeling af villighed til at dele, like og rapportere hadefulde memes.

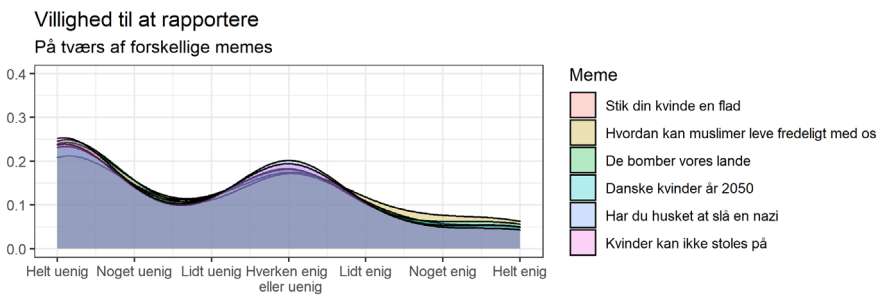
Figur A1



Figur A2

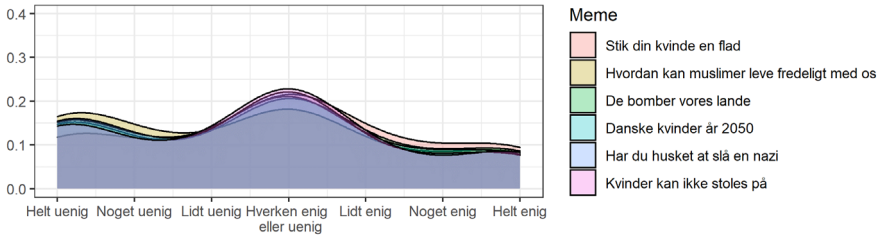


Figur A3



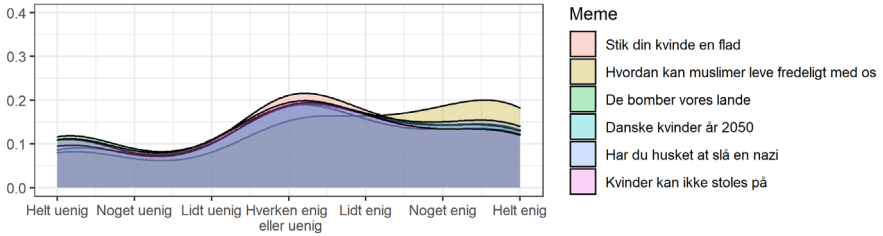
Figur A4

Memet er inden for grænserne af, hvad man kan acceptere  
På tværs af forskellige memes



Figur A5

Memet er hadefuldt  
På tværs af forskellige memes





Figur A6

”Stik din kvinde en flad”



”Hvordan kan muslimer leve fredeligt med os”



”De bomber vores lande”



”Danske kvinder år 2050”



”Har du husket at slå en nazi”



”Kvinder kan ikke stoles på”



*Table A6:* Gennemsnit på tværs af forskellige mål for eksperiment- og kontrolgruppen

Variabel	Kontrol (n = 444)	Eksperiment (n = 445)	Forskel i gennemsnit	P-værdi
Dele	0.23	0.25	0.02	0.24
Like	0.27	0.25	-0.02	0.40
Rapportere	0.32	0.35	0.02	0.18
Hadefuldt	0.56	0.56	0.01	0.60
Acceptabelt	0.45	0.43	-0.01	0.50
Alder (>14)	0.52	0.56	0.03	0.32
Pige	0.53	0.54	0.00	0.97
Politisk interesse	0.58	0.62	0.04	0.12
Lovlig politisk aktivisme	0.48	0.49	0.00	0.83
Voldelig politisk aktivisme	0.32	0.32	0.01	0.68

Note: Alle variable er omskaleret, så de går fra til 0-1. Alder og piger er har kun to værdier (0 og 1) og indikerer 1, hvis respondenter er henholdsvis over 14 år eller er en pige. Der er ingen statistisk signifikante forskelle ( $p < 0.05$ ) mellem kontrol- og eksperimentgruppen.