

# Troverdigheten til barn som vitner

*Barnets uttrykte emosjoner og påvirkningen  
av disse på juryers vurderinger av  
troverdighet og skyld i saker med mistenkt  
seksuelt overgrep*

Kjell Vegard Fjeldheim Weyde



Hovedoppgave ved Psykologisk institutt

UNIVERSITETET I OSLO

Høsten 2012



## Troverdigheten til barn som vitner

*Barnets uttrykte emosjoner og påvirkningen av disse på juryers vurderinger av troverdighet og skyld i saker med mistenkt seksuelt overgrep*

© Kjell Vegard Fjeldheim Weyde

2012

Tittel: Troverdigheten til barn som vitner

*Barnets uttrykte emosjoner og påvirkningen av disse på juryers vurderinger av troverdighet og skyld i saker med mistenkt seksuelt overgrep*

Forfatter: Kjell Vegard Fjeldheim Weyde

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

# Sammendrag

I saker der det er mistanke om seksuelt overgrep av barn er barnets og tiltaltes forklaringer ofte det eneste juryen har av bevis. Dermed blir spørsmål om vitnets troverdighet ekstra viktig i slike saker. Mye forskning tyder på at et vitnes uttrykte emosjoner under vitneforklaringen kan påvirke folks vurderinger av dette vitnet. Denne studien undersøkte i hvilken grad dette også gjelder når et barn er vitne og de som vurderer sitter i en jury, noe som ikke tidligere har vært gjort. Mer spesifikt ble det undersøkt hvordan de uttrykte emosjonene til et barn som vitner om seksuelle overgrep, kan påvirke juryers vurderinger av barnets troverdighet og spørsmål om tiltaltes skyld. Filmer med fiktive avhør av en jente ble vist til 153 deltakere (hovedsakelig studenter) der jenta viste enten trist eller sint emosjon mens hun snakket om at hun var blitt utsatt for seksuelt overgrep av faren. Deltakerne ble deretter bedt om å vurdere jentas troverdighet, sjansen for at faren var skyldig, om faren var skyldig eller ikke, og hvor streng straff faren eventuelt burde få. I kontrollgruppa vurderte deltakerne hver for seg disse spørsmålene. I eksperimentgruppa ble deltakerne satt sammen i juryer på mellom fire og åtte personer og bedt om å diskutere seg frem til svar på spørsmålene. Deretter skulle de gjøre vurderinger hver for seg. Resultatene viste generelt en effekt av uttrykt emosjon. Et trist vitne ble vurdert som mer troverdig, det ble vurdert som mer sannsynlig at faren til et trist vitne var skyldig, og faren ble oftere dømt skyldig når vitnet var trist. Det var liten effekt av å diskutere i jury med tanke på å redusere virkningen av vitnets uttrykte emosjoner. Mulige forklaringer på dette diskuteres.



# Forord

Det har vært en interessant og lærerik prosess å jobbe med denne oppgaven. Flere personer fortjener en takk for å ha bidratt til at jeg kom i mål. Blant dem er veilederne mine, Ellen Wessel og Annika Melinder, som har gitt meg grundig og nyttig veiledning. Medstudent Siri Støre og min kone Helene skal også ha en stor takk for gjennomlesning av og kommentarer til oppgaven, i tillegg til at du, Helene, har støttet meg og motivert meg hele den tiden jeg har jobbet med oppgaven. Jeg vil også gi en takk til alle deltakerne i studien og til alle som bidro til å få laget filmene: Barnehuset i Oslo, for at de lot oss filme i sine lokaler og hjalp til med ferdigstilling av filmene; Kripos-avhører Tone Davik, for at hun stilte opp i filmene som seg selv; og skuespiller Tamina fra Black Box teater for at hun spilte jenta i filmene.

Kjell Vegard Fjeldheim Weyde, oktober 2012

# Innholdsfortegnelse

Sammendrag .....	V
Forord .....	VII
Innholdsfortegnelse .....	VIII
1 Innledning og presentasjon av relevant litteratur .....	1
1.1 Hvorfor denne studien er viktig .....	1
1.2 Barns uttrykte emosjoner .....	2
1.3 Emosjoner og individers vurderinger av troverdighet og skyld .....	3
1.3.1 Induksjon av emosjoner i bedømmer .....	6
1.3.2 Kognitiv belastning hos bedømmer .....	7
1.4 Juryers vurdering av troverdighet og skyld .....	8
1.4.1 Mildhetsbias .....	8
1.4.2 Argumenter <i>for</i> at jurydiskusjoner reduserer stereotypier .....	9
1.4.3 Argumenter <i>mot</i> at jurydiskusjoner reduserer stereotypier .....	10
1.4.4 Frigjøringshypotesen .....	13
1.4.5 Antall jurydeltakere .....	13
1.5 Hypoteser .....	14
2 Metode .....	16
2.1 Deltakere .....	16
2.2 Materiale .....	16
2.3 Prosedyre .....	18
2.4 Statistiske analyser .....	19
3 Resultater .....	20
3.1 Forberedende analyser .....	20
3.2 Hypotese 1 .....	20
3.2.1 Troverdighet .....	20
3.2.2 Skyld .....	21
3.2.3 Sannsynlighet for skyld .....	22
3.2.4 Måneder straff .....	23
3.3 Hypotese 2 .....	24
3.3.1 I hvilken grad predikerers skyld av betingelse og troverdighet i kontrollgruppa?	



3.3.2	I hvilken grad predikeres skyld av betingelse og troverdighet i individuellgruppa?.....	24
4	Diskusjon.....	26
4.1	Hypotese 1 .....	26
4.2	Hypotese 2 .....	30
4.3	Begrensninger.....	31
4.4	Konklusjon.....	33
	Litteraturliste .....	35
	Vedlegg 1: Spørreskjemaet for individuell vurdering.....	44
	Vedlegg 2: Spørreskjemaet for juryens vurdering .....	45
	Vedlegg 3: Informasjonsskriv gitt til deltakerne før eksperimentets start.....	47
	Vedlegg 4: Informasjon gitt til deltakerne etter endt gjennomføring av eksperimentet.....	48



# 1 Innledning og presentasjon av relevant litteratur

## 1.1 Hvorfor denne studien er viktig

I artikkelen Den perfekte drapsmann fra februar 2010 i A-Magasinet (Gulliksen, 2010) står følgende utdrag: *“Hva er det som skjer? Har hun sett dette, eller er det fantasi. Hva tror du? skriver Øglænd til etterforskeren inne på avhørsrommet. Noen minutter senere får han svar: “Hun gråter. Snakket i 30 minutter om barndom og vonde ting. Dette er sant. Jenta har samvittighet.”*

Utdraget over er et eksempel på en troverdighetsvurdering, som ofte er en sentral del av en jurs oppgave i retten. Når det er overgrepssaker retten skal ta stilling til, er vitneutsagn ofte det eneste juryen kan basere sine bedømmelser på (London, Bruck, Wright & Ceci, 2008). I tillegg oppgir dommere at fornærmedes vitneforklaring i slike saker ofte ses på som det viktigste beviset (Myers, Redlich, Goodman, Prizmich & Imwinkelried, 1999). Også i Norge er det slik at barnevitnets forklaring ofte er det eneste som foreligger av bevis i overgrepssaker (Langballe, 2007). Derfor er det ikke overraskende at vurderinger av troverdighet er funnet å kunne predikere straffebedømmelse (Cramer, Brodsky & DeCoster, 2009).

Statistikken viser at det årlig anmeldes omtrent 800 tilfeller av seksuelle overgrep mot barn i Norge (Statistisk Sentralbyrå, 2011), men at det trolig er store mørketall (Politidirektoratet, 2011). Mossige og Huang (2010) fant at omtrent ni prosent av jentene i et stort utvalg av barn under 18 år svarte at de hadde vært utsatt for voldtekt og/eller voldtektsforsøk. Blant guttene var det samme tallet på litt under én prosent. Samlet betyr dette at spørsmålet om vitnetroverdighet kan være viktig for mange barn og deres pårørende, men også for de som er mistenkt for å stå bak overgrep. Det er derfor viktig å få økt kunnskap om hvilke faktorer som på forskjellige måter kan påvirke oppfatninger av troverdigheten til barn som vitner, slik at sjansen for feilaktige vurderinger om hvorvidt overgrepet har funnet sted, kan reduseres. Flere studier har undersøkt hvordan individuelle troverdighetsvurderinger kan påvirkes av en persons holdninger, emosjoner og atferd. (Connolly, Price & Gordon, 2009). Nonverbal atferd kan være en faktor som det er vanlig å ta hensyn til når det skal vurderes i hvilken grad noen lyver (Ellison & Munro, 2009). Undersøkelser viser for eksempel en

tendens der politifolk bruker vitnets opptreden som indikasjon på hvorvidt påstandene deres er falske eller ikke (Jordan, 2004). Enkelte studier finner at det faktisk kan være en forskjell i enkelte typer nonverbal atferd mellom folk som lyver og som forteller sannheten (Vrij, Akehurst, Soukara & Bull, 2004). Men til tross for at nonverbale faktorer vektlegges i vurderinger av troverdighet, viser det seg at mange av disse nonverbale faktorene ikke har noen signifikant sammenheng med hvorvidt personen er sannferdig (Stiff & Miller, 1986; Strömwall & Granhag, 2003). Funn tyder for eksempel på at ansiktsuttrykk ikke gir noen indikasjon på hvorvidt avsender forteller sannheten eller lyver (Bond & DePaulo, 2006).

Det er blitt undersøkt hvordan et voksent vitnes og et barnevitnes uttrykte emosjoner kan påvirke individers (Heath, Grannemann & Peacock, 2004; Regan & Baker, 1998) og juryers (Dahl, Enemo, Drevland, Wessel, Eilertsen & Magnussen, 2007) vurderinger av vitnets troverdighet og tiltaltes skyld. Men denne studien er trolig den første til å undersøke hvordan juryer i en overgrepssak påvirkes av et vitnets uttrykte emosjoner når det gjelder troverdighetsvurderinger av vitnet og vurderinger av skyld for tiltalte, der vitnet er et barn (11 år gammel). Videre vil det først presenteres hva litteraturen sier om barns reaksjoner når de forteller om overgrep. Deretter vil det bli sett på individuelle vurderinger av troverdighet og skyld, og hvordan disse kan påvirkes av vitnets uttrykte emosjoner. Som det vil fremgå under, vil det kunne være viktig å vite noe om hvordan individuelle vurderinger påvirkes av vitnets uttrykte emosjoner, for bedre å kunne vurdere hvordan juryers vurderinger vil påvirkes. Det vil så gjøres rede for beslutningstaking i juryer og argumenteres for hvorfor juryers vurderinger av troverdighet og skyld ikke vil skille seg fra individuelle vurderinger. Metode for denne studien, resultater og diskusjon av resultatene vil så presenteres.

## **1.2 Barns uttrykte emosjoner**

En av grunnene til at emosjoner vektlegges i vurdering av og kommunikasjon med andre, kan være at de representerer en viktig måte å kommunisere med andre om hvem man er og hvordan man ønsker å fremstå (DePaulo, 1992). Emosjonene er nonverbale signaler som kan kommuniseres ved ansiktsuttrykk, stemme og kroppsspråk (Ekman, 2004). Hvilke emosjoner barn uttrykker, vil kunne avhenge av hvem som er til stede sammen med barnet, eller med andre ord, hvem som observerer barnet (Zeman & Garber, 1996). Zeman og Gaber undersøkte faktorer som kan påvirke barns grunner til å kontrollere eller uttrykke emosjoner (tristhet, sinne og smerte). De fant at barn oppgir frykt for negative interpersonlige konsekvenser som

en grunn for ikke å uttrykke følelser som tristhet og sinne, og at dette i større grad gjelder overfor venner enn overfor foreldre. Barna mente at det å uttrykke tristhet eller sinne ikke ville møte like stor forståelse hos venner som hos foreldrene. På den annen side mente barna at grunner for å uttrykke tristhet eller sinne kunne være forventningen om hjelp fra andre, kanskje særlig fra foreldre, eller at emosjonene ble så intense at de ville være vanskelige å kontrollere.

Når det gjelder emosjoner blant barn som forteller om overgrep de har vært utsatt for, viser forskning at folk har oppfatninger som ikke stemmer med analyser av hvordan vitnene faktisk forholder seg. I en studie av Regan og Baker (1998) forventet 81% av deltakerne at en seks år gammel jente ville reagere med gråt i det øyeblikket hun møtte overgriperen. Andre forventede reaksjoner inkluderte blant annet frykt (67%), forvirring (43%) og sinne (30%). Dette til tross for at flertallet av barn som forteller om seksuelt eller fysisk overgrep eller neglekt, fremstår som nøytrale. Sayfan, Mitchell, Goodman, Eisen og Qin (2008) kodet intervjuer med barn og fant at 75% av dem som fortalte om overgrep hadde et nøytralt uttrykk, og at 98% ikke gråt. Det var ikke sammenheng mellom alder og kjønn og negativ affekt, men det fremkom en ikke-signifikant trend der jenter var noe mer opprørte enn gutter, og eldre barn var mer opprørte enn yngre barn. De som hadde opplevd gjentatte overgrep var mindre opprørte mens de snakket om overgrep. Studien til Goodman et al. (1992) av ekte rettssaker med barn, fant også at eldre barn var signifikant mer triste enn yngre barn. Wood, Orsak, Murphy og Cross (1996) fant i likhet med Sayfan og medarbeidere at barn som ble intervjuet om misbruk (i dette tilfellet barn det var sterk mistanke om at hadde vært utsatt for overgrep), for det meste av tiden var rolige og oppmerksomme. Reaksjoner som for eksempel det å spøke og virke ubekymret (Christianson & Lindholm, 1998) og å vise sinne eller smileatferd (Bonanno, Noll, Putnam, O'Neill & Trickett, 2003) under beskrivelsen av overgrep, er også rapportert.

### **1.3 Emosjoner og individers vurderinger av troverdighet og skyld**

Resultatene fra forskningen nevnt over, kan tyde på at det er diskrepans mellom folks forventninger og det forskningen viser når det gjelder vitners emosjonelle reaksjoner. Stereotypien om at barn som forteller om overgrep gråter, ser ut til å komme til uttrykk i individuelle bedømmelser av barnevitners troverdighet. En seks år gammel jente ble vurdert

som signifikant mer ærlig, troverdig, nøyaktig og reliabel hvis det ble beskrevet at hun gråt heller enn å være rolig (Regan & Baker, 1998). Denne effekten av et emosjonelt vitne på bedømmelser av troverdighet og skyld later til å være sterk, og er funnet i en rekke studier også av voksne vitner, der lekfolk vurderer både kvinnelige vitner (Kaufmann, Drevland, Wessel, Overskeid & Magnussen, 2003) og tiltalte menn (Heath, Grannemann & Peacock, 2004; Wessel et al., 2011). Emosjonseffekten er i tillegg funnet med politifolk som bedømmere (Bollingmo, Wessel, Eilertsen & Magnussen, 2008). Det at folk påvirkes av informasjon i omgivelsene er en prosess som kan foregå uten at bedømmer er det bevisst (Ferguson & Bargh, 2004). Bargh, Chen og Burrows (1996) ga deltakerne i et eksperiment språkoppgaver som inneholdt enten ord som hadde med høflighet å gjøre, ord som hadde med frekkhet å gjøre, eller nøytrale, tilfeldig ord. De som hadde gjort språkoppgaven med ord relatert til frekkhet, avbrøt i større grad eksperimentatoren mens denne snakket med en annen person, enn de som hadde gjort de oppgavene med ord relatert til høflighet.

Vurderinger av troverdighet er funnet å kunne predikere straffebedømmelse (Cramer et al., 2009). I sammenheng med dette, viser studien til Regan og Baker (1998) at vitnets uttrykte emosjoner ikke bare ser ut til å kunne påvirke folks oppfatning av hennes troverdighet, men også hvorvidt folk dømmer tiltalte skyldig eller ikke. Flere dømte tiltalte skyldig når vitnet gråt. Denne sammenhengen mellom emosjoner og bedømt skyld finnes også hos blant andre Castelli, Goodman og Ghatti (2005). Magnussens forskningsgruppe (Kaufmann et al., 2003) fant at trist emosjonsuttrykk hos fornærmede ga økt antall skyldbedømmelser av tiltalte, sammenliknet med nøytralt eller inkongruent (positive emosjoner, smiling) emosjonsuttrykk. I tillegg predikerte økt bedømmelsene av troverdighet økt bedømmelsene av skyld. Heath et al. (2004) fant at skyld også ble predikert av troverdighet når det var tiltalte som ble vurdert. I tillegg fungerte oppfattelsen av tiltaltes ærlighet og sympati for tiltalt som mediatorer av tiltaltes uttrykte emosjoners effekt på grad av oppfattet skyld.

Enkelte studier har også funnet resultater som tyder på at folk, i tillegg til å ha en forventning om hva slags type emosjoner vitner vil vise, har en forventning om intensiteten til disse emosjonene. Det vil dermed ikke bare kunne være av betydning om vitnet er nøytral eller trist eller sint, men også hvilken grad av emosjoner som uttrykkes. Golding, Fryman, Marsil og Yozwiak (2003) fant at for lite ("rolig") eller for mye gråting ("hysterisk gråting") fra offerets side førte til at færre deltakere dømte tiltalte skyldig, i forhold til når mengden gråt var moderat. I tråd med dette fant Salekin, Ogloff, McFarland og Rogers (1995) at for

tiltalte kvinner, men ikke menn, påvirket grad av tiltaltes emosjoner oppfattelse av skyld signifikant. Tiltalte ble oppfattet som mer skyldig når hun viste flat eller høy affekt, og mindre skyldig når hun viste moderat affekt.

Ifølge Tsoudis og Smith-Lovin (1998) og Winkel og Koppelaar (1991) vil det at en fornærmet (i disse eksperimentene voksne personer) i en overgrepssak uttrykker tristhet, kunne være et signal fra den fornærmede om at han eller hun ikke fortjente eller ønsket å bli utsatt for overgrep. Dermed vil den fornærmede ses i et mer positivt lys enn om hun hadde forholdt seg nøytral. Økt grad av tristhet uttrykt av fornærmede vil også kunne være assosiert med økt positiv vurdering av fornærmede og strengere straffeutmåling for tiltaltes del. Tilsvarende vil det, ifølge disse forskerne, også kunne være fordelaktig for tiltalte å uttrykke høy grad av tristhet.

De to foregående avsnittene gir eksempler på forskjellige typer oppfatninger som kan ligge til grunn for folks vurderinger av fornærmedes uttrykte emosjoner. Den ene muligheten er at fornærmede oppfattes som mest troverdig når emosjonenes intensitet matcher alvorlighetsgraden av forbrytelse eller overgrep. Den andre muligheten kan være at fornærmede er mest troverdig når emosjonene er sterke, uavhengig av forbrytelsens alvorlighetsgrad. Rose, Nadler og Clark (2006) satte disse to hypotesene opp mot hverandre ved å undersøke vitnets grad av emosjoner (sterk eller svak grad) opp mot alvorlighetsgraden av forbrytelsen vitnet var utsatt for (alvorlig eller mindre alvorlig). Resultatene viste at det for vitnet ville kunne være noe mer fordelaktig å matche grad av emosjoner med alvorlighetsgraden av forbrytelsen. Vitner ble vurdert å ha en uvanlig reaksjon dersom de reagerte rolig på en alvorlig forbrytelse, i dette tilfellet å under et ran få en pistol mot hodet. Vitner ble også vurdert å ha en uvanlig reaksjon dersom de reagerte sterkt på en hendelse som var av langt mindre alvorlig grad, som tyveri av gjenstand til en verdi av 100 dollar. Uvanlig reaksjon var negativt korrelert med idømt straffelengde.

I samsvar med Rose et al. (2006) mener Hackett, Day og Mohr (2008) at det ser ut til å være et forventnings-bias (expectancy-violation bias), heller enn emosjonell ekspressivitet per se, som påvirker observatørens oppfatning av troverdigheten til vitner. Effekten av vitnets uttrykte emosjoner på vurdering av troverdighet og straff vil dermed være sterkeste når vitnet uttrykker emosjoner av den type og med den intensitet som bedømmerne i størst grad forventer at vitnet skal uttrykke. Fordi forventninger om hvordan vitner skal fremstå ser ut til å ha betydning for bedømmelse av troverdighet og straff, kan det antas at triste vitner vil ses på som mer troverdige enn sinte vitner, etter som det i større grad er

forventet at barnevitner skal være triste enn sinte (Sayfan et al., 2008). Det at forventnings-bias kan forklare noe av mekanismen bak virkningen av uttrykte emosjoner, støttes av at personer som hadde sterke forventninger, i motsetning til svake forventninger til hvordan typiske overgrepsofre reagerer, i større grad så emosjonelle overgrepsofre som mer troverdige enn nøytrale overgrepsofre (Hackett et al., 2008). Forventninger til et overgrepsoffers reaksjoner vil kunne være forskjellig, avhengig av offerets kjønn. En voksen kvinne som fortalte om at hun var utsatt for uønsket sex, ble mer trodd om hun gråt og var oppskaket enn om hun var rolig, men dette gjaldt kun dersom hun ble beskrevet å være en kvinne som passet inn i et kjønnsstereotypisk mønster (Schuller, McKimmie, Masser & Klippenstine, 2010). I en studie av Melville (2007) ble det funnet at når vitnene uttrykte stereotypisk emosjon, som for menn var definert som sinne, og for kvinner som frykt, fikk tiltalte strengere straff. Selv om enkelte av de foregående studiene handler om voksne vitner og ikke barn, illustrerer de betydningen av stereotyper i vurderingen av troverdighet og straff.

### **1.3.1 Induksjon av emosjoner i bedømmer**

I tillegg til forventnings-bias-effekten, vil effekten som et emosjonelt vitne har på en bedømmer, kanskje kunne forklares av emosjoner som oppstår i bedømmeren selv. I beslutningsforskningen er det et blitt forsket mye på hvordan opplevde emosjoner kan påvirke beslutninger, og en vanlig oppfatning er at opplevde emosjoner er viktige for at mennesker skal kunne være i stand til mer effektiv beslutningstaking (Damasio, 1994). Observatører kan ha en tendens til å oppleve de samme emosjonene som de observerer hos andre (Wild, Erb & Bartels, 2001), og den emosjonelle tonen i et budskap kan indusere en tilsvarende emosjon i mottakeren (Neumann & Strack, 2000). Innen vitnepsykologiforskningen er det blitt funnet at et emosjonelt vitne vil kunne vekke større medfølelse hos bedømmeren enn et ikke-emosjonelt vitne. Denne medfølelsen vil igjen kunne være assosiert med økt tiltro til vitnet (Ask & Landström, 2010). Bedømmers sympati for fornærmede er funnet å henge sammen med ønske om økt straff for anklagede (Nadler & Rose, 2003). Tsoudis (2002) fant at tristhet hos tiltalte og fornærmede hadde en positiv sammenheng med empati hos bedømmeren; jo tristere personen var, jo mer empati fikk bedømmeren med denne personen. Det var også slik at et trist vitne var assosiert med mindre empati for tiltalte. Empati for tiltalte var videre assosiert med oppfattelsen av tiltaltes identitet, som igjen var assosiert med straff. De som følte empati for tiltalte, så tiltalte i et mer positivt lys (tiltalte angret, den negative handlingen var ikke representativ for ham som person) og dømte mindre strengt.



Et vitne vil altså kunne indusere emosjoner hos bedømmerne. Ask og Granhag (2007) undersøkte effekten av induisert tristhet eller sinne hos bedømmere, og fant at tristhet var assosiert med mer nøyaktig prosessering av saken enn sinne, som var assosiert med mer overfladisk prosessering. Sinne hos bedømmeren kan ut ifra dette se ut til å gjøre bedømmeren mer utsatt for å støtte seg på tidligere oppfatninger, forventninger og bias. I en annen studie av sosial persepsjon var det induisert sinne som skilte seg fra nøytralt og induisert trist humør, som begge ga like resultater. Sinne viste sammenheng med økt bruk av stereotypier og heuristikker, i forhold til tristhet og nøytralt humør (Bodenhausen, Sheppard & Kramer, 1994). Sinte bedømmere vil kunne se den anklagedes handling som mer intensjonell enn det triste bedømmere vil, og vil dermed kunne straffe tiltalte i økt grad (Ask & Pina, 2011). Effekten av vitnets uttrykte emosjoner på bedømt troverdighet kan dermed se ut til å kunne forklares av forventninger til hvordan overgrepsofre skal fremstå, samt av følelser som oppstår hos de som vurderer troverdigheten (Ask & Landström, 2010).

### 1.3.2 Kognitiv belastning hos bedømmer

Det er flere faktorer som kan påvirke i hvilken grad bedømmere blir influert av vitnets uttrykte emosjoner. Økt vektlegging av nonverbal informasjon er vist å kunne forekomme i situasjoner der det er for lite tale å basere vurderingen sin på (Vrij, 2004), eller i saker der graden av bedømmers involvering er lav (“Studiens mål er å skaffe data til introkurs i metode”), heller enn høy (“Veldig viktig studie, din bedømmelse viktig for at vi skal få gjennomført studien”) (Reinhard & Sporer, 2008, mine oversettelser). Bias kan også forekomme dersom situasjonen legger stort beslag på kognitive ressurser (Forsyth, 2006), slik at den komplekse situasjonen forenkles (Bodenhausen & Lichtenstein, 1987). Effekten av et vitnes emosjoner på troverdighet ble for eksempel forsterket med økt kognitiv belastning hos bedømmeren (Ask & Landström, 2010). Høy kognitiv belastning (som å huske 9 siffer mens en film av en vitneforklaring blir spilt) ser ut til å føre til økt vektlegging av nonverbal og mindre vektlegging av verbal informasjon, fordi prosessering av nonverbal informasjon krever færre kognitive ressurser hos mottakeren (Reinhard & Sporer, 2008). Disse forskningsresultatene kan tenkes å være overførbare til rettssaker der tidspress og arbeidsmengde kan være betydelig (Ask, Granhag & Rebelius, 2011).

## 1.4 Juryers vurdering av troverdighet og skyld

Effekten av vitnets uttrykte emosjoner på vurderinger av troverdighet og skyld ser ut til å være sterk. Det kan få uheldige konsekvenser i overgrepssaker dersom slike irrelevante faktorer skal være avgjørende for hvorvidt barnevitnet blir trodd og tiltalte dømt. I de studiene som er nevnt over, er det gjort individuelle vurderinger av troverdighet. Et spørsmål som kan stilles, er om grupper er bedre beslutningstakere enn det individer er, og om det å diskutere i grupper vil føre til at vurderingene i mindre grad påvirkes av irrelevante faktorer, som emosjoner. Det eksisterer så langt vi vet ingen studier som til nå har undersøkt hvorvidt juryer påvirkes av barnevitners uttrykte emosjoner i overgrepssaker, men Dahl et al. (2007) gjorde en studie der de undersøkte effekten av jurydiskusjoner på bedømmelse av troverdighet og skyld i en sak med mistenkt seksuelt overgrep. Juryer på 5-7 personer så først en kort film av en voksen kvinne som forklarte at hun hadde vært utsatt for seksuelt overgrep. Kvinnen fremsto med enten kongruente emosjoner (fortvilelse og hulking), nøytral (flat, rett frem), eller inkongruent (positive og paradoksale emosjoner, hyggelig). Mens en kontrollgruppe som ikke diskuterte, anslo sannsynligheten for at tiltalte var skyldig til henholdsvis 56 prosent (inkongruent), 67 prosent (nøytral) og 80 prosent (kongruent), var de som hadde diskutert i jury tilnærmet upåvirket av vitnets emosjoner, med sannsynligheter på mellom 75 og 80 prosent. I spørsmålet om dom, dømte henholdsvis 61 prosent (inkongruent), 80 prosent (nøytral) og 95 prosent (kongruent) i kontrollgruppa den tiltalte skyldig, mens tilsvarende tall lå på 64-71 prosent blant jurymedlemmene. Diskusjon førte med andre ord til færre dommer totalt og en reduksjon av effekten av vitnets uttrykte emosjoner.

### 1.4.1 Mildhetsbias

Studien til Dahl et al. (2007) gir et eksempel på mildhetsbias-effekten (leniency bias), som går ut på at deltakere etter jurydiskusjon, sammenliknet med før diskusjon, uttrykker mindre tro på at tiltalte er skyldig (MacCoun & Kerr, 1988), eller at tiltalte bør få mildere straff (Sporer, 1986). Et slikt mildhetsbias er funnet også dersom jurymedlemmene før diskusjon ikke har noen spesiell preferanse i skyldspørsmålet (Fraser, Thompson & Gutowski, 1986) eller ser ut til å helle mot skyldig (Tanford & Penrod, 1986). Det har vært fokusert på at juryer med majoriteter på to tredjedeler eller mer tenderer mot å dømme på linje med denne majoriteten. MacCoun & Kerr (1988) fant for eksempel at i juryer der to tredjedeler i utgangspunktet var for å dømme, dømte juryen skyldig i 67 prosent av tilfellene. Der en tilsvarende majoritet var

for frifinnelse, kom juryene frem til ikke skyldig i 94 prosent av tilfellene. Dersom juryene i utgangspunktet var delt (50/50) i oppfatningen av skyld, endte de opp med frifinnelse i 81 prosent av tilfellene. Samlet fungerte dermed jurydiskusjonene til fordel for tiltalte, dersom det ikke i utgangspunktet var en klar majoritet som var for å dømme skyldig.

#### 1.4.2 **Argumenter for at jurydiskusjoner reduserer stereotypier**

Et argument for hvorfor diskusjon i en gruppe skulle føre til vurderinger som skiller seg fra individuelle vurderinger, kunne tenkes å være forventningen om å måtte forsvare sitt eget syn i en gruppediskusjon. En slik forventning kunne gjøre at deltakerne tenker nøyere gjennom egne meninger og slik sett reduserer sjansen for å bli påvirket av stereotypier (Kruglanski & Freund, 1983; Stewart, 1998). En annen mulig forklaring er at det å bli konfrontert med sine stereotypier av andre, også ville kunne være med på å redusere stereotypiske responser (Czopp, Monteith & Mark, 2006). I en gruppe som diskuterer en vitneforklaring, kan det tenkes at noen påpeker at vitnets emosjoner bør være en viktig indikasjon på hvorvidt vitnet lyver eller snakker sant. En annen i gruppa kan da tenkes å svare med at folk som forklarer seg om overgrep, kan vise forskjellige typer emosjoner, og at disse emosjonene ikke nødvendigvis har betydning for troverdigheten. Denne konfrontasjonen ville kanskje kunnet bidra til redusert stereotypisk tenkning; at emosjonene til vitnet ikke påvirker gruppas vurderinger av troverdighet og skyld. Bollingmo, Wessel, Sandvold, Eilertsen og Magnussen (2009) undersøkte effekten av å gi informasjon til individuelle bedømmere om betydningen av vitners uttrykte emosjoner. Slike instruksjoner kan for eksempel gis av dommeren i rettssaken. I en sak som dreide seg om seksuelt overgrep, fikk noen av deltakerne beskjed om at vitnets emosjoner ikke er en god faktor å basere bedømmelser på. Andre deltakere ble informert om at vitnets emosjoner kan gi informasjon om hvorvidt vitnet er troverdig eller ikke. Dessuten var det en kontrollgruppe som ikke fikk noen slik informasjon. Resultatene viste at det å gjøre bedømmerne oppmerksomme på at vitnets uttrykte emosjoner ikke gir reliabel informasjon om troverdighet, signifikant dempet virkningene av uttrykte emosjoner på vurdering av troverdighet.

Selv om ikke disse eksemplene omhandler juryer, vil den informasjonen som deltakerne fikk, kunne tenkes å være sammenliknbar med den informasjonsutvekslingen som kan forekomme i en gruppe (Czopp et al., 2006), der ett av gruppemedlemmene vil kunne dele det samme synspunktet med resten av gruppa, som det «dommeren» i studiet til Bollingmo et al. (2009) delte med individuelle bedømmere.

### 1.4.3 Argumenter *mot* at jurydiskusjoner reduserer stereotypier

På den annen side er det ikke alltid instruksjoner fra dommeren i en rettssak demper effekten av bias (Smith & Greene, 2005). Det kan til og med være slik at slike instruksjoner har motsatt effekt. Edwards og Bryan (1997) ga deltakerne i et eksperiment instruksjon om å overse emosjonelt ladet informasjon, eller bevismateriale, i en sak som omhandlet mord og ran. Dette førte til økt virkning av dette bevismaterialet. Denne effekten ble ikke demonstrert hos deltakere som fikk beskjed om å ignorere tilsvarende bevismateriale som var av ikke-emosjonell karakter. Edwards og Bryan argumenterer for at virkningen av emosjonell informasjon kan være vanskeligere å undertrykke enn virkningen av ikke-emosjonell informasjon. Videre vil instruksjoner om å overse visse typer informasjon kunne gjøre denne informasjonen mer tilgjengelig i hukommelsen. Økt tilgjengelighet av informasjon vil kunne føre til økt påvirkning av denne informasjonen på bedømmelser (Bargh, 1989). Lieberman og Arndt (2000) mener dette kan være noe av grunnen til at instruksjoner som er ment å skulle redusere faren for bias, ikke alltid har noen positiv effekt.

Når det gjelder hvordan jurydiskusjoner kan påvirke virkningen av bias på bedømmelse av troverdighet og skyld, er det ingen entydige funn fra forskningen. I Smith og Greene (2005)-studien hadde ikke diskusjon i juryen noen effekt på feilansettelse av bevis. Saken det dreide seg om i det tilfellet, var en kollisjon der en bil var påkjørt av trailer. Deltakere skulle avgjøre hvor mye erstatning bilføreren burde få. De som kun fikk informasjon om skader på bilen og bilføreren, ga mindre erstatning enn de som i tillegg fikk den i dette tilfellet irrelevante informasjonen om hvordan trailersjåføren hadde oppført seg i en kort periode før krasjet.

Det finnes også en rekke andre eksempler som tyder på at grupper kan la seg påvirke av forskjellige irrelevante faktorer når de skal foreta troverdighetsvurderinger. Flere av studiene har et voksent vitne, men hovedpoenget er å vise til indikasjoner på at juryer kan la seg påvirke av bias. En tiltalt som var fysisk attraktiv, ble oftere frifunnet enn en tiltalt som var mindre attraktiv (MacCoun, 1990). Tiltaltes etnisitet påvirket juryers beslutninger om dødsstraff (Baldus, Woodworth, Zuckerman, Weiner, & Broffitt, 1998). Dessuten viste juryer økt tendens til å dømme tiltalte når de fikk info om at tiltalte tidligere hadde begått en kriminell handling, i forhold til det juryer uten slik informasjon gjorde. Det var derimot ingen slik forskjell blant folk som vurderte skyldspørsmålet individuelt og uten diskusjon (Hans & Doob, 1976). Det er også forsket på hvordan eksponering for negativ publisitet forut for rettssaker påvirker bedømmerens vurdering av skyld, og både juryer og personer som ikke

diskuterer, har blitt påvirket av dette biaset. Juryer eksponert for slik publisitet med sterkt emosjonelt innhold, dømte tiltalte oftere skyldig enn juryer som ikke var eksponert for dette (Devine, Clayton, Dunford, Seying & Pryce, 2001). Vitnes alder kan, i tillegg til individer, også påvirke juryers vurderinger av troverdighet. I en studie av Gabora, Spanos og Joab (1993) ga individuelle troverdighetsvurderinger av et 13- og 17-årig vitne som muligens var utsatt for seksuelt overgrep, ingen forskjeller. Men etter at deltakerne hadde gått sammen i juryer på 5-8 personer og kommet frem til enighet om dom, ble 13-åringen vurdert som mer troverdig enn 17-åringen. Dessuten ble det vurdert som mer sannsynlig at 13-åringene manglet nødvendig kunnskap for å fabrikere historie om seksuelt overgrep. Dette er på linje med tidligere presentert forskning på individers vurdering av troverdigheten til barn og voksne i overgrepssaker (Bottoms & Goodman, 1994).

Lieberman og Arndt (2000) påpeker at fordi jurydiskusjon foregår etter presentasjonen av bevis, kan det være at diskusjonen kommer for sent til effektivt å redusere bias. Ifølge Porter og ten Brinckes (2009) «Dangerous Decisions Theory» foretar lekdommere og dommere i rettssaker troverdighetsvurderinger av vitnet etter meget kort tid, for deretter å legge stor vekt på informasjon som støtter dette førsteinntrykket, og lite vekt på informasjon som går imot det. Under rettssaker vil det dermed kunne være slik at jurymedlemmene individuelt foretar egne vurderinger før jurydiskusjonen begynner. Hannaford-Agor, Hans, Mott og Munsterman (2002) spurte folk som hadde fungert som jurymedlemmer i virkelige rettssaker, om når de første gang begynte å helle mot den ene siden (det vil si gjøre seg opp en mening om skyld eller ikke skyld) i rettssaken de hadde deltatt i. 80 prosent oppga at de gjorde dette før jurydiskusjonen, og 20 prosent at de gjorde det under jurydiskusjonen. Disse resultatene indikerer at et stort flertall av jurymedlemmer starter diskusjonen med en viss grad av oppfatning om hvorvidt tiltalte er skyldig eller ikke. Dessuten vil jurymedlemmenes oppfatninger før diskusjon kunne gjenspeiles i juryens vurderinger etter diskusjon. For eksempel undersøkte Schulz-Hardt, Frey, Lüthgens og Moscovici (2000) hvordan preferanser før diskusjon hos deltakere i grupper på fem påvirket gruppediskusjonen. I gruppene var det enten tre, fire eller fem av de fem personene som i utgangspunktet støttet det samme alternativet. Resultatene viste at jo mer homogen gruppa var i utgangspunktet, jo mer gjorde et bekreftelsesbias seg gjeldende i gruppa. Det vil si at tendensen til å søke informasjon som bekrefter ens slutninger, heller enn å avkrefte dem (Forsyth, 2006), ble forsterket. Bekreftelsesbiaset var mer uttalt i homogene grupper enn blant individuelle, men ble samtidig redusert jo mer heterogen gruppa i utgangspunktet var, og var ikke signifikant i de mest

heterogene gruppene (Schulz-Hardt et al., 2000). Informasjonen en gruppes medlemmer sitter inne med før diskusjon, og hvordan denne informasjonen er distribuert i gruppa, kan dermed være av betydning for hvor vidt stereotypier vil polariseres eller ikke.

Dette ble også vist i en studie av Golding, Bradshaw, Dunlap og Hodell (2007). Disse forskerne undersøkte sammenhengen mellom bedømmers kjønn og vurdering av skyld i en fiktiv sak. De fant at det var større kvinneandel enn herreandel som før diskusjon mente tiltalte var skyldig. Dette ble reflektert i juryenes beslutninger ved at juryer med kvinnemajoritet stemte for skyldig oftere enn juryer uten kvinnemajoritet, og ved at juryer uten kvinnemajoritet stemte oftere for ikke skyldig enn det juryer med kvinnemajoritet gjorde. Dessuten forandret lekdommere i juryer med kvinnelig majoritet oftere mening fra ikke skyldig til skyldig enn det lekdommere i juryer uten en slik kvinnelig majoritet gjorde. På den annen side skiftet personer som hadde diskutert i juryer uten kvinnemajoritet, oftere mening fra skyldig til ikke skyldig enn det personer i juryer med kvinnemajoritet gjorde (Golding et al., 2007).

Det kan dermed ut ifra de sistnevnte studiene se ut til at en avgjørende faktor for utfallet av en jurydiskusjon er flertallets oppfatning før diskusjon. Flere studier funnet at majoritetens oppfatning før diskusjonen, i de klart fleste tilfeller ender opp med å bli juryens dom (Devine et al., 2001; Schmidt & Brigham, 1996). Schmidt og Brigham (1996) undersøker deltakeres oppfatning av skyld før diskusjon og juryenes oppfatning av skyld (skyldig versus ikke skyldig) i en sak med mistenkt seksuelt overgrep, og fant en signifikant sammenheng, der juryens dom reflekterte flertallets mening før diskusjon. Dette er delvis en kontrast til de studiene som finner tendens til mildhets-bias hos juryer (Dahl et al., 2007; MacCoun & Kerr, 1988). Studier av ekte juryer finner heller ikke støtte for mildhetsbias (Devine et al., 2004; Sandys & Dillehay, 1995). Det ble for eksempel funnet at juryer som i utgangspunktet ikke hadde noen klar majoritet (2/3) i forhold til oppfatning om skyld, endte opp med å dømme skyldig i 75 prosent av tilfellene. I tillegg var det en sterk sammenheng mellom det jurymedlemmene mente i begynnelsen av diskusjonen, og det som endte opp med å bli juryens dom (Devine et al., 2004).

Noe av grunnen til at majoritetens mening ofte ser ut til å bli juryens endelige beslutning, kan være at utsagn som kommer fra majoriteten tenderer mot å vektlegges i større grad i diskusjonen enn det utsagn fra minoriteten gjør (Hastie, Penrod & Pennington, 1983). Dersom én i gruppa nevner at et barnevitnes uttrykte emosjoner ikke bør vektlegges i vurderingen av troverdighet og skyld, kan det dermed tenkes at dette utsagnet vektlegges i

mindre grad dersom utsagnet på forhånd kun er delt av en minoritet. Ifølge Kaplan og Miller (1987) vil saker som avgjøres ved flertallsbeslutning, ofte kunne føre til at minoriteten overses. Når saken avgjøres ved enstemmighet, må alle lyttes til i større grad, og minoriteter vil da kunne påvirke gruppas beslutning i den ene eller andre retningen. Juryene i studien vår får ikke beskjed om hvorvidt vurderingene må være enstemmige eller ikke. Implikasjonen blir da at hvis juryene tar avgjørelser på bakgrunn av det majoriteten mener, uten å måtte ta hensyn til å skulle oppnå enstemmighet, vil minoriteten bli mindre innflytelsesrik.

#### 1.4.4 Frigjøringshypotesen

Det ser ikke ut til å være noen tydelige svar på spørsmålet om det er individer eller grupper som lar seg minst påvirke av forskjellige typer bias (Kerr, MacCoun & Kramer, 1996). Det later til å være en tendens der bias, som for eksempel publisitet i media forut for en rettssak, har minimal påvirkning på juryer når det er veldig svake eller sterke beviser i saken som klart favoriserer en side. Da har biasene større påvirkning på individuelle bedømmere. Er bevisene av moderat styrke, later det derimot til at biasene påvirker gruppa i større grad (Devine et al., 2001; Kerr, Niedermeier & Kaplan, 1999). Dette er i tråd med Kalven og Zeisels (1966) frigjøringshypotese (liberation hypothesis). Denne hypotesen sier at når bevisene i en rettssak tydelig favoriserer en side, så vil juryen stort sett stemme til fordel for denne siden. Men dersom bevisene i saken er mer tvetydige, vil juryens dom oftere påvirkes av andre, uhensiktsmessige variabler, som for eksempel publisitet forut for rettsaken (pre-trial publicity). Hypotesen støttes også av nyere forskning på ekte juryer (Devine, Buddenbaum, Houp, Studebaker & Stolle, 2009). Frigjøringshypotesen kan muligens forklare noe av årsaken til tilsynelatende tvetydige funn i forskningen på grupper og individuelles troverdighetsvurderinger.

#### 1.4.5 Antall jurydeltakere

Videre må det også nevnes at det har vært en debatt om hvorvidt juryer på seks personer skiller seg fra juryer på 12 personer, som er det vanlige antallet medlemmer i USA (i Norge er det 10). Selv om små forskjeller finnes, ser det ikke ut til å være forskjeller av betydning mellom juryer på seks og 12 personer når det gjelder straffebedømmelser (Devine et al., 2001). Også når det gjelder ekte juryer ser ikke jurystørrelsen ut til å ha en sammenheng med om det dømmes skyldig eller ikke (Devine et al., 2004). Dermed blir det i dette eksperiment

naturlig å bruke juryer på seks personer av praktiske hensyn. Det er også praktisk å benytte seg av studenter som forsøkspersoner. Det å basere eksperimentet på studenter, kan forsvares med at studentpopulasjoner og potensielle jurydeltakere fra samfunnet for øvrig ikke viser noen forskjeller i juryeksperimenter (MacCoun & Kerr, 1988).

## 1.5 Hypoteser

Samlet sett ser det ut til at det ofte er en sammenheng mellom individuelle bedømmelser og de bedømmelsene som juryer foretar. Dermed vil resultatene av forskningen gjort på individuelles bedømmelser kunne si noe om hvordan juryer vil bedømme vitner. Mer spesifikt vil det kunne være som følger: Individuelle bedømmere tenderer mot å vurdere barnevitner i overgrepssaker som mer troverdig når de gråter, og dessuten dømme tiltalte oftere (Regan & Baker, 1998). Ut ifra dette kan det tenkes at et flertall av jurymedlemmene begynner diskusjonen med en oppfatning om at barnet snakker sant dersom hun gråter, og at tiltalte er skyldig. Og hvis det er slik at det majoriteten i begynnelsen av diskusjonen mener om straffebedømmelsen, i de klart fleste tilfeller ender opp med å bli juryens dom (Devine et al., 2001), ville det ikke være overraskende å finne at også juryer lar seg påvirke av et barnevitnes emosjoner. I dette eksperimentet vil det da forventes at et trist vitne vil vurderes som mer troverdig enn et sint vitne, fordi tristhet hos barn utsatt for overgrep i større grad er en forventet emosjon enn sinne. Det var naturlig at den ene av de to emosjonene var trist, ettersom det er en meget vanlig emosjon å ha med i forskning på vitnetroverdighet. Sint emosjon er med fordi det blant folk flest er en klart mindre forventet emosjon å se hos vitnet i seksuelt overgrep-saker (Regan & Baker, 1998), og på grunn av denne forventningsforskjellen mellom trist og sint emosjon, vil det kunne introduseres et forventnings-bias (Hackett et al., 2008).

Som det fremgår av litteraturgjennomgangen hittil, har mange studier undersøkt faktorer som kan påvirke skyldbømmelse og troverdigheten til vitner og tiltalte i rettssaker. Det er gjort studier av deltakere som hver for seg bedømmer troverdighet, og av deltakere som gjennom diskusjon skal komme frem til en troverdighetsvurdering. Flere av disse studiene handler om seksuelle overgrep, i likhet med dette studiet. For eksempel har deltakere hver for seg skullet vurdere troverdigheten til et vitne som enten gråt eller var rolig/nøytral mens hun fortalte om overgrep (Regan & Baker, 1998). Når det gjelder juryer, har for eksempel Dahl et al. (2007) undersøkt hvordan vitnets emosjoner påvirker juryers vurderinger



av troverdighet av et voksent vitne i en overgrepssak. Selv om det også har vært studier der juryer vurderer barns troverdighet, later det til at få, om noen, av disse jurystudiene med barnevitner har sett på hvordan vitnes emosjoner når hun forteller om seksuelt overgrep, påvirker troverdigheten hennes. Derfor vil dette studiet undersøke dette nærmere.

Med utgangspunkt i forskningen som er presentert her, er følgende hypoteser utarbeidet:

- 1) Vitnets uttrykte emosjoner vil ha effekt både på kontrollgruppas, jurygruppas og individuellgruppas (de individuelle vurderingene til jurymedlemmene etter diskusjon) vurderinger av troverdighet og skyld. Vurderingene til individuelle bedømmere, juryer og de individuelle som har diskutert i jury, vil ikke skille seg signifikant fra hverandre.**
- 2) Vurdering av troverdighet vil predikere vurdering av skyld, slik at økt troverdighet vil predikere økt sjanse for skyldbedømmelse.**

## 2 Metode

### 2.1 Deltakere

Kriterier for utvelgelse av forsøksdeltakere var at deltakerne måtte være minst 18 år gamle og potensielle deltakere i en jury. Deltakerne ble hovedsakelig rekruttert fra Universitetet i Oslo; Psykologisk institutt, Universitetsbiblioteket og Det samfunnsvitenskapelige fakultet. I tillegg til alder, var en annen forutsetning for å få delta at deltakerne ikke hadde hatt undervisning i vitnepsykologi. Slik ble sjansen redusert for at psykologistudentene skulle vite for mye om eksperimentets hensikt. Studenter ble rekruttert på universitetets område og fordelt randomisert til forskjellige betingelser. De ble informert om at jeg kom fra Psykologisk institutt og holdt på med et psykologieksperiment i bygningen vi befant oss i, og at eksperimentet handlet om vitnetroverdighet, ville ta maks 30 minutter, og gikk ut på å se en film og svare på et spørreskjema. Ingen belønning ble tilbudt for å være med.

Utvalget besto av i alt 153 deltakere. Av disse deltok 98 i jurybetingelsen, mens 55 utgjorde kontrollgruppene. Deltakerne i jurybetingelsen diskuterte seg først frem til enighet om svar på spørreskjemaet, før de fylte ut spørreskjemaet igjen, men da hver for seg. Juryene vil heretter bli omtalt som *jury*, mens ordet *individuell* vil hen vise til jurymedlemmene som etter diskusjon gjorde individuelle vurderinger. Kontrollgruppa, de som svarte individuelt på spørreskjemaet uten noen diskusjon i forkant, vil heretter bli omtalt som kontroll.

Gjennomsnittsalderen for deltakerne totalt viste seg å være 23 år ( $M=23.13$ ,  $SD=4.38$ ), og gjennomsnittsalderen i de forskjellige gruppene skilte seg ikke fra hverandre [ $F(3,149)=.284$ ,  $p=.837$  eta $_p^2$  =.006]. Av de 153 deltakerne, var 92 kvinner (60%) og 61 menn (40%). En chi-square-analyse viste ingen signifikante forskjeller i kjønnsfordeling mellom gruppene [ $\chi^2(3, N=153)=2.78$ ,  $p=.427$ ].

### 2.2 Materiale

Studien er utarbeidet i tråd med etiske retningslinjer i Helsinkideklarasjonen og godkjent av Regional etisk komité, Helse sørøst.

En ung skuespiller, en 11 år gammel jente fra et lokalt teater, ble rekruttert til filming på Barnehuset i Oslo. På forhånd hadde hun vært på Psykologisk institutt og øvd inn en vitneforklaring. Ved å filme på Barnehuset ble settingen realistisk, etter som virkelige

avhør av barn til bruk i rettssaker filmes her. I filmene spiller jenta rollen som vitne, og hun intervjues av den nevnte avhøreren fra Kripos, som spiller seg selv. Vitne og avhører ble filmet fra topp til tå, sittende i hver sin stol inntil en vegg, med et lite bord imellom. Både vitne og avhører satt vendt mot hverandre, begge omtrent 45 graders vinkel vendt bort fra kameraet, det vil si at de ble filmet litt fra siden forfra. Filmene var like, med unntak av vitnets uttrykte emosjon. I den ene filmen fremsto vitnet som trist, og i den andre som sint. I den triste versjonen er stemmen til vitnet trist og hun snakker litt forsiktig og lavere. Hun ser mye ned, og ved en anledning tørker hun en tåre. Det er ellers lite av tydelig gråting. I den sinte versjonen er vitnet sint, men sitter allikevel rolig i stolen, har et irritert/sint uttrykk i ansiktet og snakker høyere og mer aktivt.

Varigheten av filmene var på cirka to minutter. Avhøret er bygget opp i tråd med rekkommendert praksis for avhør (se for eksempel Goodman & Melinder, 2007). Det snakkes først litt om skolen jenta går på, hvilke fag hun liker, og at læreren har sett at hun har virket ukonsentrert den siste tiden. Det kommer frem at vitnet har fortalt læreren sin at hun har vært utsatt for seksuelt overgrep. Til avhøreren forteller vitnet at hun har vært utsatt for seksuelt overgrep av faren en kveld hun var på helgebekesøk hjemme hos ham. Historien er fiktiv (for hele historien, se vedlegg 1). Avhøreren fulgte anbefalte retningslinjer for avhør av barn (Goodman & Melinder, 2007), med åpne spørsmål og videre mer fokuserte spørsmål der det var nødvendig.

Blant flere opptak som ble gjort, valgte en psykolog og to studenter, i samarbeid med en erfaren barneavhører fra Kripos, ut de to filmene som fremsto som mest realistiske, og som skulle vises til deltakerne.

Det ble utarbeidet et spørreskjema som var knyttet opp mot filmen deltakerne så (se vedlegg nummer 1 og 2). Spørreskjemaet er utarbeidet på bakgrunn av et spørreskjema brukt i en tilsvarende studie som hadde et voksent vitne i stedet for et barn (Kaufmann et al., 2003). Det ble tilpasset med tanke på at vitnet var et barn og at det var både kontrollgrupper og juryer som vurderte. I spørreskjemaet ble det ved jurydiskusjon spurt om antall jurymedlemmer. De versjonene av spørreskjemaene som skulle fylles ut individuelt, inneholdt spørsmål om kjønn, alder og jobb/utdanning. Ellers var alle skjemaene like. Den første avhengige variabelen i skjemaet var spørsmål om hvor troverdig jenta fremsto. Dette skulle vurderes på en Likert-skala fra én til sju, der én indikerte veldig troverdig og sju indikerte veldig lite troverdig. Videre skulle deltakerne foreta en sannsynlighetsvurdering av hvorvidt jentas far var skyldig eller ikke, ved å rangere sannsynligheten for skyld fra null til 100

prosent, der 100 indikerte hundre prosent sjans for skyld. Vurdering av skyldspørsmål skulle besvares med enten skyldig eller ikke skyldig. Dersom deltakeren eller juryen krysset av for skyldig, ble det bedt om å angi grad av fengselsstraff, fra én måned til seks år, eller eventuelt annen form for straff. Denne typen overgrepssaker har i Norge en øvre strafferamme på tre år, derav disse svaralternativene. Selv om det var spørsmålene om troverdighet og skyld (ja/nei) som var av hovedinteresse for hypotesene, ble det vurdert som interessant å undersøke deltakernes oppfatning omkring farens skyld nærmere. Variabelen sannsynlighet for skyld inneholdt 11 svaralternativer (0 til 100 prosent, med intervaller på 10), og kunne dermed tenkes å gi et litt finere mål på skyldoppfatning enn det den dikotome variabelen skyld kunne. Straffelengde er et sentralt spørsmål i retten, og var derfor av interesse å ha med her.

## 2.3 Prosedyre

Eksperimentet foregikk på seminarrom i de samme byggene hvor deltakerne ble rekruttert. Før eksperimentet ble det muntlig informert om at deltakerne ville få se en film på cirka to minutter der en jente fortalte en avhører fra Kripas at hun var blitt utsatt for overgrep, og at starten og slutten er klippet bort av hensyn til jentas (vitnes) identitet. Det ble ikke sagt noe om jentas alder. I tillegg fikk deltakerne lese et informasjonsskriv (se vedlegg nummer 3). Filmene ble så vist på lerret, ved hjelp av projektor (Epson EBG-5100) koblet til en bærbar pc. Eksterne høyttalere ble brukt for å sikre høy og tydelig lyd.

Etter filmen skulle deltakerne svare på spørreskjemaet. I jurybetingelsene ble deltakerne bedt om å sette seg sammen i en gruppe for i fellesskap å fylle ut spørreskjemaet. Det var ønskelig at disse gruppene skulle bestå av 5-7 personer. Deretter ble disse deltakerne bedt om å sette seg hver for seg og fylle ut det samme spørreskjemaet. Deltakerne i kontrollbetingelsene (ingen jurydiskusjon) ble bare bedt om å fylle ut skjemaet hver for seg.

Etter at alle hadde sett filmen og fylt ut spørreskjemaet(-ene), fikk deltakerne utdelt et informasjonsskriv (se vedlegg nummer 4). Her ble de informert om at avhøret var fiktivt, og de fikk en forklaring på hva som var hensikten med eksperimentet. De ble bedt om ikke å avsløre hensikten for andre, for ikke å påvirke den videre datainnsamlingen. Det ble også gitt anledning til å stille spørsmål eller komme med kommentarer, dersom noen ønsket det.

## 2.4 Statistiske analyser

Statistikkprogrammet IBM SPSS Statistic 19 ble brukt til å utføre analysene. Studien representerte et faktorielt 3 (kontroll, jury, individuell) x 2 (emosjon: trist/sint) mellomgruppe-design. For å undersøke effekten av de uavhengige variablene gruppe (kontroll, jury, individuell) og emosjon på de avhengige variablene troverdighet, sannsynlighet (for skyld) og måneder straff ble toveis mellomgruppe-ANOVA utført. Chi-square-tester ble benyttet for å undersøke virkningen av emosjon på bedømmelse av straff, etter som straff var en kategorisk variabel. Logistisk regresjon ble utført for å finne ut hvorvidt troverdighet og kjønn predikerte straff.

Variabelen troverdighet ble snudd for å gjøre sammenhengen med skårer på andre variabler tydeligere. En høy skåre på troverdighet var da et uttrykk for skåring til økt fordel for vitnet (mer troverdig), på samme måte som høy skåre på sannsynlighet, skyld og måneder straff.

# 3 Resultater

## 3.1 Forberedende analyser

I analysene ble det først undersøkt hvorvidt det var noen effekt av kjønn på vurdering av troverdighet og skyld. Undersøkelser av kontrollbetingelsen viste at kvinners og menns vurderinger av troverdighet [ $F(1,53)=1.336, p=.741$ ], sannsynlighet for skyld [ $F(1,53)=.163, p=.355$ ], skyld [ $\chi^2(1, N=54)=.153, p=.695$ ] og straffeutmåling [ $F(1,29)=3.068, p=.956$ ] ikke skilte seg signifikant fra hverandre. Det ble derfor besluttet å behandle begge kjønn under ett i alle analyser.

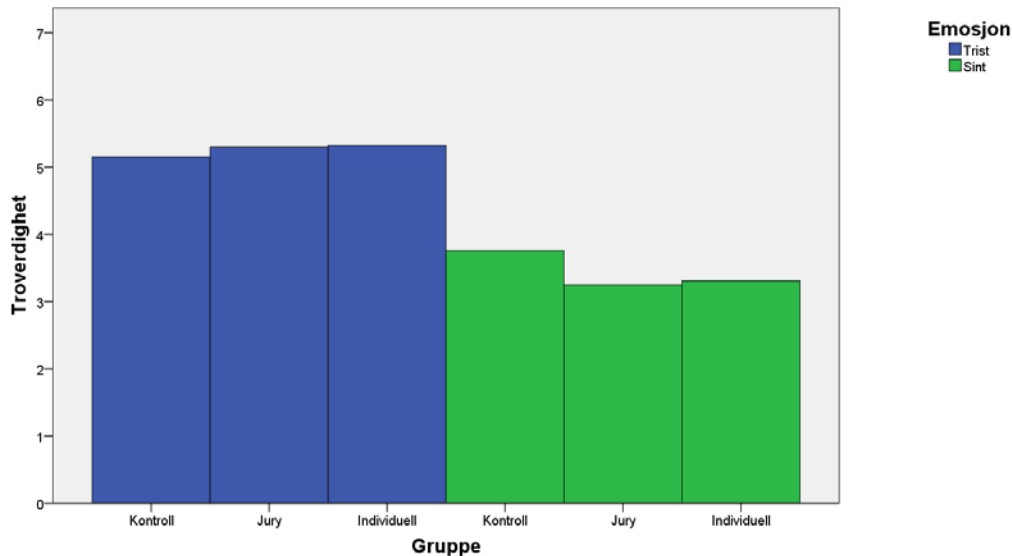
Videre ble det undersøkt i hvilken grad gruppene kontroll, jury og individuell skilte seg fra hverandre på vurderingene av henholdsvis trist og sint vitne.

## 3.2 Hypotese 1

Hypotese 1 var at vitnets uttrykte emosjoner ville ha effekt både på kontrollgruppas, jurygruppas og individuellgruppas vurderinger av troverdighet og skyld, og at vurderingene til individuelle bedømmere, juryer og de individuelle som har diskutert i jury, vil ikke skille seg signifikant fra hverandre.

### 3.2.1 Troverdighet

En toveis mellomgruppe-ANOVA med gruppe og emosjon som uavhengige variabler og troverdighet som avhengig variabel, viste en signifikant hovedeffekt av emosjon [ $F(1,165)=49.981, p=.000, \eta^2=.232$ ]. Enveis mellomgruppe-ANOVAER indikerte at effekten av emosjon på troverdighet var statistisk signifikant i både kontrollgruppa ( $[F(1,53)=12.78, p=.001]$ ; trist:  $M=5.15, SD=1.49$ ; sint:  $M=3.76, SD=1.41$ ), jurygruppa ( $[F(1,16)=12.67, p=.003]$ ; trist:  $M=5.30, SD=.82$ ; sint:  $M=3.25, SD=1.58$ ) og i individuellgruppa ( $[F(1,96)=60.09, p=.000]$ ; trist:  $M=5.32, SD=.94$ ; sint:  $M=3.31, SD=1.62$ ). Vitnet ble av alle gruppene vurdert som mest troverdig når hun fremstod som trist. Hovedeffekten av variabelen gruppe [ $F(2,165)=.235, p=.791, \eta^2=.003$ ] og interaksjonseffekten [ $F(2,165)=1.024, p=.362$ ] var ikke signifikante. (Se Figur 1.)



**Figur 1:** Grafisk fremstilling av gjennomsnittsskårene på variabelen troverdighet i gruppene med trist (blå) og sint emosjon (grønn). Høy skåre indikerer økt troverdighet, og figuren illustrerer hvordan alle gruppene i trist betingelse (blå) ga høyere troverdighetsvurdering enn gruppene i sint betingelse (grønn).

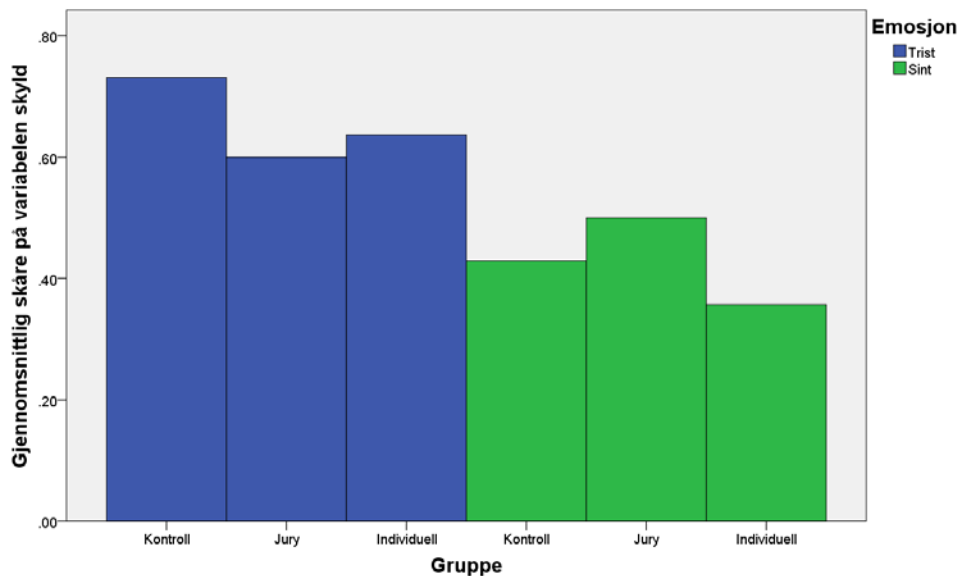
### 3.2.2 Skyld

Resultatene av en chi-square-test viste en signifikant effekt av forholdet mellom betingelse og skyld i kontrollgruppa [ $\chi^2(1, N=54)=3.875, p=.049$ ], der deltakerne i den triste betingelsen dømte skyldig oftere (19 av 26) enn deltakerne i den sinte betingelsen (12 av 28).

For å undersøke effekten av emosjon på vurdering av skyldspørsmålet i jurygruppa, ble det foretatt en Kruskal-Wallis-test (ikke-parametrisk test for uavhengige utvalg). Resultatene viste ingen signifikante forskjeller mellom trist og sint betingelse [ $\chi^2(1, N=18)=.170, p=.680$ ]. 6 av 10 juryer i trist betingelse, og 4 av 8 juryer i sint, stemte for skyldig.

Når det gjaldt individuellgruppa, var det én person som unnlot å svare på skyldspørsmålet, og som derfor er utelatt fra analysene her. Resultatene av en chi-square-test viste en signifikant effekt av forholdet mellom betingelse og straff [ $\chi^2(1, N=97)=6.36, p=.012$ ], der deltakerne i den triste betingelsen dømte skyldig oftere (35 av 55) enn deltakerne i den sinte betingelsen (15 av 42).

For å se om de tre gruppene hadde svart forskjellig når det gjaldt bedømmelse av straff versus ikke straff, ble det foretatt en Kruskal-Wallis-test. Denne viste at gruppene ikke skilte seg signifikant fra hverandre på skyldspørsmålet, verken når vitnet var trist [ $\chi^2(2, N=91)=.867, p=.648$ ] eller når hun var sint [ $\chi^2(2, N=78)=.740, p=.691$ ]. (Se Figur 2.)



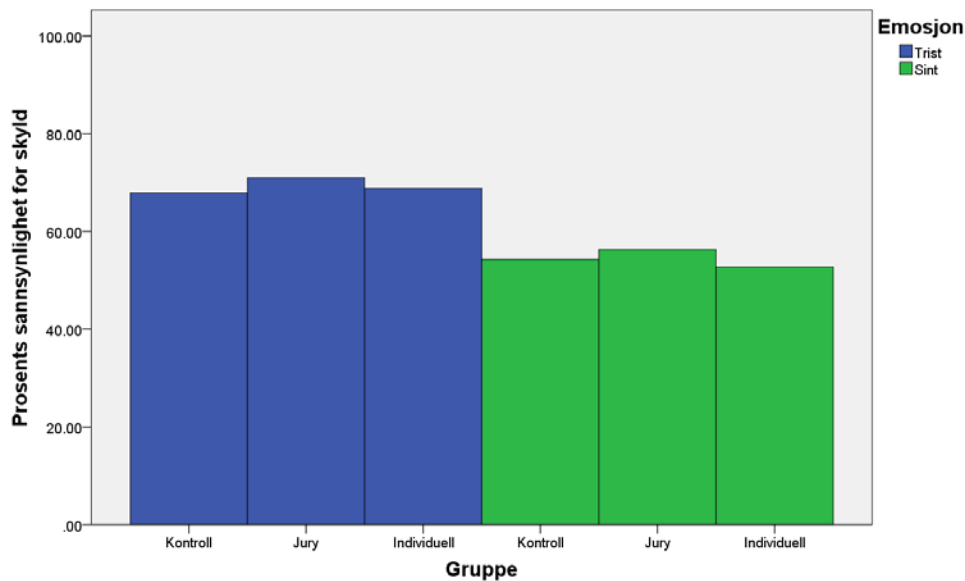
**Figur 3:** Grafisk fremstilling av skårene på variabelen skyld i gruppene med trist (blå) og sint emosjon (grønn). Økt verdi på y-aksen tilsvarer økt antall skyldbedømmelser i betingelsen.

### 3.2.3 Sannsynlighet for skyld

Spørsmålet om sannsynlighet for skyld var ikke direkte knyttet opp til hypotesene, men ble allikevel vurdert som en interessant variabel å ha med i analysene (se under punkt 2.2 Materiale).

En toveis mellomgruppe-ANOVA med gruppe og emosjon som uavhengige variabler og sannsynlighet for skyld som avhengig variabel, viste en signifikant hovedeffekt av emosjon [ $F(1,165)=15.609, p=.000, \eta^2=.086$ ]. Enveis mellomgruppe-ANOVAER indikerte at effekten av emosjon på troverdighet var statistisk signifikant i både kontrollgruppa ([ $F(1,53)=5.37, p=.024$ ]; trist:  $M=67.88, SD=19.86$ ; sint:  $M=54.31, SD=23.21$ ) og individuellgruppa ([ $F(1,96)=18.87, p=.000$ ]; trist:  $M=68.84, SD=15.28$ ; sint:  $M=52.74, SD=21.42$ ), men ikke i jurygruppa ([ $F(1,16)=3.21, p=.092$ ]; trist:  $M=71.00, SD=15.24$ ; sint  $M=56.25, SD=19.78$ ). Hovedeffekten av variabelen gruppe [ $F=(2,165)=.163, p=.850, \eta^2=.002$ ] og interaksjonseffekten [ $F(2,165)=.076, p=.927$ ] var ikke signifikant. De gruppene som så et trist vitne, vurderte det som mer sannsynlig at faren var skyldig, enn de gruppene som så et sint vitne.



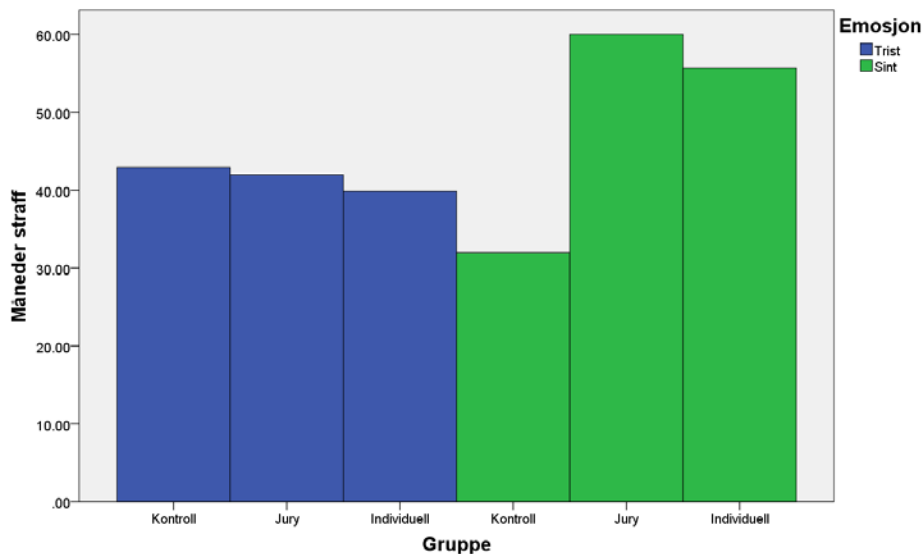


**Figur 2:** Grafisk fremstilling av gjennomsnittsskårene på variabelen sannsynlighet i gruppene med trist (blå) og sint emosjon (grønn). Høy skåre indikerer økt sannsynlighet for skyld, og figuren illustrerer hvordan alle gruppene i trist betingelse (blå) ga høyere sannsynlighetsvurdering enn gruppene i sint betingelse (grønn).

### 3.2.4 Måneder straff

Spørsmålet om måneder straff var heller ikke direkte knyttet opp til hypotesene, men ble allikevel vurdert som en interessant variabel å ha med i analysene (se under punkt 2.2 Materiale).

En toveis mellomgruppe-ANOVA med gruppe og emosjon som uavhengige variabler og måneder straff som avhengig variabel, viste en interaksjonseffekt [ $F(2,83)=4.245, p=.018, \eta^2=.093$ ]. Datafilen ble derfor splittet i to basert på emosjon. Deretter ble enveis mellomgruppe-ANOVAER utført på de to gruppene, med gruppe som uavhengig variabel og måneder straff som avhengig variabel. Analysene viste ingen signifikante forskjeller mellom gruppene som så et trist vitne [ $F(2,56)=.141, p=.869$ ], men en signifikant forskjell mellom gruppene som så et sint vitne [ $F(2,27)=6.315, p=.006$ ]. Post hoc-tester (Tukey) viste signifikante forskjeller [ $p=.040$ ] mellom kontrollgruppa [ $M=32, SD=19.37$ ] og jurygruppa [ $M=60, SD=16.97$ ] og mellom kontrollgruppa og individuellgruppa [ $M=55.71, SD=18.59$ ] [ $p=.009$ ], men ingen signifikante forskjeller [ $p=.915$ ] mellom jurygruppa og individuellgruppa. Jurygruppa og individuellgruppa ga strengere straff i sint betingelse enn det kontrollgruppa gjorde (se Figur 4).



**Figur 4:** Grafisk fremstilling av skårene på variabelen måneder straff i gruppene kontroll, jury og individuell, med trist (blå) og sint emosjon (grønn). Økt verdi på y-aksen tilsvarer økt antall skyldbedømmelser i betingelsen.

### 3.3 Hypotese 2

#### 3.3.1 I hvilken grad predikerers skyld av betingelse og troverdighet i kontrollgruppa?

En logistisk regresjonsanalyse med betingelse (trist/sint) og troverdighet som uavhengige variabler og skyld som avhengig variabel, ble foretatt for å undersøke i hvilken grad troverdighet predikerer svar på skyldspørsmålet. Av tabell 1 fremgår det at troverdighet ( $p=.007$ ) var en signifikant prediktor for skyld. Beta-verdien til troverdighet (.665) er positiv, og viser dermed at en økning på denne variabelen predikerer økning i variabelen skyld. Det vil si at høyere troverdighetsvurderinger av vitnet er assosiert med høyere sjanse for å dømme skyldig. Den totale modellen er statistisk signifikant [ $\chi^2(2, N=54)=14.269, p=.001$ ], og mål på pseudo R square (analysene Cox & Snell R Square [= .232] og Nagelkerke R Square [= .312]) tyder på at modellen forklarer mellom 23,2 og 31,2 prosent av variasjonen i variabelen skyld.

#### 3.3.2 I hvilken grad predikeres skyld av betingelse og troverdighet i individuellgruppa?

En logistisk regresjonsanalyse med betingelse (trist/sint) og troverdighet som uavhengige variabler og skyld som avhengig variabel, ble foretatt for å undersøke i hvilken grad troverdighet predikerte svar på skyldspørsmålet. Av tabell 2 fremgår det at troverdighet

( $p < .001$ ) var en signifikant prediktor for skyld. Beta-verdien til troverdighet (1.033) er positiv, og viser dermed at en økning på denne variabelen predikerte økning i variabelen skyld. Det vil si at det å skåre høyt på troverdighetsvariabelen, økte sjansen for å dømme skyldig (skyldig=1). Den totale modellen er statistisk signifikant [ $\chi^2(2, N=97)=34.115, p < .001$ ], og mål på pseudo R square (analysene Cox & Snell R Square [= .297] og Nagelkerke R Square [= .395]) tyder på at modellen forklarer mellom 29,7 og 39,5 prosent av variasjonen i variabelen skyld.

## 4 Diskusjon

Dette eksperimentet undersøkte hvordan et barnevitnes uttrykte emosjoner i en overgrepssak påvirker juryers vurderinger av barnevitnet. Disse vurderingene gjaldt troverdigheten til barnevitnet, bedømmelse av om anklagede var skyldig eller ikke, sannsynlighetsvurdering vedrørende den anklagedes skyld, og hvor lang straff anklagede eventuelt burde få. Mer spesifikt ble det sett på effekten av de to emosjonene tristhet og sinne, og om deltakere som diskuterte de overnevnte punkter i juryer, vurderte annerledes enn en kontrollgruppe som ikke diskuterte, og hvorvidt juryens vurderinger var sammenlignbare med de individuelle jurymedlemmenes vurderinger etter diskusjon.

### 4.1 Hypotese 1

Den første hypotesen var at vitnets uttrykte emosjoner ville ha effekt både på kontrollgruppas, på jurygruppas og på individuellgruppas vurderinger av troverdighet og skyld, og at disse tre gruppens vurderinger ikke ville skille seg fra hverandre. Resultatene støtter i hovedsak denne hypotesen. Det var nesten ingen signifikante forskjeller mellom kontrollene, juryene og de individuelle som hadde diskutert; alle de tre gruppene ble påvirket av emosjon, slik at de i trist betingelse ga høyere vurderinger til fordel for vitnet enn i sint betingelse (selv om ikke alle forskjellene i jurygruppa var signifikante). Ett unntak var bedømmelsen av måneders straff. Her var det en signifikant forskjell i betingelsen sint mellom kontrollgruppen på den ene siden og jury- og individuell-gruppene på den andre siden, der jury- og individuell-gruppene i gjennomsnitt ga lengre straffer enn kontrollgruppen. En mulig forklaring på dette, kan være følgende: Under observasjon av juryene ble det lagt merke til at juryen ofte lot alles meninger bli hørt, for så å la svaret bli et gjennomsnitt av dette. Det er kjent fra resultatene i kontrollgruppa at det var flere som dømte skyldig i trist betingelse (19 av 26) enn i sint (12 av 28). Etter som deltakerne dermed virket mindre tilbøyelige til å stemme skyldig i sint betingelse, kan det tenkes at det i en jury derfor ville kreves sterkere røster som argumenterte for straff i sint betingelse, i forhold til i trist, for at den samlede juryen skulle gå for skyldig. Slik kunne det tenkes at en jury som stemte skyldig, ville bestå av deltakere med i gjennomsnitt sterkere overbevisning om tiltaltes skyld, sammenlignet med deltakere i juryer i trist betingelse. Dette ville igjen kunne tenkes å komme til syne som økt skåre på straffutmålingen.

Effekten av emosjon på juryers bedømmelse av skyld var ikke signifikant, og det kan innvendes her at  $n$  var litt for lav til å undersøke effekten på skyld på en god måte. Selv om det er vanskelig å si noe om hvorvidt en signifikant effekt på skyld ville fremkommet ved høyere  $n$ , gir den signifikante effekten av emosjon på skyld i individuellgruppa en indikasjon på at jurydiskusjon ikke hadde noen effekt på deltakernes oppfatning av skyld.

Effekten av emosjon på vurdering av sannsynlighet for skyld var heller ikke signifikant i jurygruppa. Også her kan det ha hatt noe å si at  $n$  var lav. Samtidig argumenterer Stevens (1996) for at det i tilfeller der gruppestørrelsen er liten, kan være nødvendig å justere alfanivået opp for å kompensere for lav  $n$ . Ut ifra dette vil det kunne tenkes at  $p=.092$  i dette tilfellet er signifikant (men se under 4.3 Begrensninger, om diskusjon av type I- og type II-feil). Gjennomsnittsverdiene i henholdsvis trist ( $M=71$ ) og sint ( $M=56.25$ ) betingelse er på nivå med både kontroll- og individuellgruppa, og støtter dermed opp under den nevnte mulige tolkningen. I tillegg er det signifikante effekter av emosjon på sannsynlighet i individuellgruppa, slik det er i kontrollgruppa. Dette indikerer at det å diskutere i jury ikke hadde noen effekt på deltakernes oppfatning av sannsynlighet for at faren var skyldig.

Antall juryer bør i fremtidige studier uansett være høyere, slik at spekulasjoner om hvorvidt det eksisterer effekter, ikke blir nødvendig.

Cohen (1988) klassifiserer effektstørrelser ( $\eta^2$ ) som små effekter fra .01; som moderate effekter fra .06; og som store effekter fra .14 og høyere. Ut ifra denne klassifiseringen er effekten av emosjonsvariabelen stor på variabelen troverdighet, og moderat på variablene sannsynlighet for skyld og måneder straff. Forskjellen mellom emosjonsbetingelsene er dermed av betydelig praktisk signifikans.

Det at det, med ett unntak, ikke var signifikante forskjeller mellom vurderingene til kontrollgruppa, jurygruppa og individuellgruppa, er i overensstemmelse med annen forskning som finner at flertallets mening før diskusjon ser ut til å være viktig for juryens bedømmelser (Devine et al., 2001). Jurymedlemmer foretar muligens troverdighetsvurderinger av vitnet etter meget kort tid, for deretter å legge stor vekt på informasjon som støtter dette førsteinntrykket, og lite vekt på informasjon som går imot det (Porter & ten Brinke, 2009). Flertallet av jurymedlemmer gjør seg muligens opp en mening om saken før diskusjonen begynner (Hannaford-Agor et al., 2002). Som nevnt tidligere, later det til å være en vanlig oppfatning at jenter gråter når de konfronteres med overgrep, og at dette kan være noe av grunnen til at triste vitner ofte vurderes mer fordelaktig enn vitner som fremstår med en emosjon som i mindre grad er forventet (Regan & Baker, 1998), slik tilfellet var med

kontrollgruppa i dette eksperimentet. På bakgrunn av dette er det ikke urimelig å anta at jurymedlemmene før diskusjonen startet også var påvirket av vitnets emosjoner på en tilsvarende måte som kontrollgruppa. Dermed kan det tenkes at jurymedlemmene i den triste betingelsen før diskusjon vurderte vitnet som mer troverdig enn jurymedlemmene i den sinte betingelsen. Gruppene i den triste betingelsen kan i tråd med dette tenkes å før diskusjon ha bestått av et flertall jurymedlemmer som oppfattet vitnet som troverdig. Slik sett ville juryene i den triste betingelsen vært mer homogene når det gjaldt vurderingene av vitnet, slik at jurydiskusjonene kan ha bidratt til å bekrefte stereotypien om at vitnet er mest troverdig når hun er trist (Schulz-Hardt et al., 2000). Som Lieberman og Arndt (2000) påpeker, kan det være at jurydiskusjonen kommer for sent til effektivt å redusere bias, fordi jurydiskusjonen foregår etter presentasjonen av bevis.

Resultatene fra denne studien gir ikke resultatene støtte til at et mildhets-bias (MacCoun & Kerr, 1988; Sporer, 1986) gjør seg gjeldende i jurydiskusjonene. Det vises ved at svarene på skyldspørsmålet ikke skilte seg signifikant fra hverandre i kontroll-, jury- og individuellgruppene i henholdsvis trist og sint betingelse. Men resultatene fra denne studien stemmer med MacCoun og Kerrs (1988) resultater på ett punkt: MacCoun og Kerr fant at i juryer der to tredjedeler i utgangspunktet var for å dømme, dømte juryen skyldig i 67 prosent av tilfellene. I den nåværende studien var en tilsvarende andel (73 prosent) for straff i trist betingelse i kontrollgruppa. Etter å ha diskutert i jury dømte 64 prosent skyldig (individuellgruppa). På den annen side fant MacCoun og Kerr at der en tilsvarende majoritet (2/3) var for frifinnelse, kom juryene frem til ikke skyldig i 94 prosent av tilfellene. Dersom juryene i utgangspunktet var delt (50/50) i oppfatningen av skyld, endte de opp med frifinnelse i 81 prosent av tilfellene. I den nåværende studien var 43 prosent for skyld i sint betingelse i kontrollgruppa, mens 36 prosent var det etter å ha diskutert (individuellgruppa). Med andre ord var det små forandringer i svarene på skyldspørsmålet i denne studien, i forhold til det det var i MacCoun og Kerrs studie. Og effekten av diskusjon på svarene på skyldspørsmålet var omtrent den samme enten majoriteten i utgangspunktet var for skyld (kontroll, trist), eller meningene om skyld var jevnt fordelt (kontroll sint). Funnene til denne studien er dermed samlet sett mer i samsvar med funn fra studier av ekte rettssaker som ikke viser støtte til mildhetsbias (Devine et al., 2004).

I denne studien ble det undersøkt hvordan emosjonene tristhet og sinne påvirker en rekke vurderinger. En faktor som kan ha påvirket resultatene, er graden av uttrykt emosjon. I lys av forventnings-bias-effekten (Hackett et al., 2008) kan grad av uttrykt emosjon, heller

enn emosjonen i seg selv, ha vært med på å skape forskjeller i vurderingene. For eksempel kunne det tenkes at vitnet i den siste betingelsen uttrykker et nivå av sinne som passer dårligere med deltakernes forventninger enn det nivået hun uttrykker av tristhet. Men uansett hvorvidt det er emosjonen i seg selv eller graden av emosjon som spiller en rolle for vurderingene, vises det få effekter av å diskutere i jury, da gruppens vurderinger på hver av emosjonsbetingelsene ikke skilte seg signifikant fra hverandre (med unntak av måneder straff i sint betingelse). Selv om det ikke ble foretatt et objektivt mål på grad av uttrykt emosjon, ble det tilstrebet så stor grad av likhet som mulig mellom de to emosjonsbetingelsene når det gjaldt intensiteten til emosjonene. Allikevel kan det ikke med sikkerhet utelukkes at enten forskjeller i emosjonsintensitet eller andre faktorer, som vitnets talemåte (Cramer et al., 2009), kan ha hatt betydning for forskjellene i deltakernes vurderinger mellom de to emosjonsbetingelsene.

Det kan tenkes at deltakerne opplevde liten grad av involvering i prosjektet og at dette var med på å øke vektlegging av vitnets uttrykte emosjoner (Reinhard & Sporer, 2008). Selv om det ble understreket for deltakerne at de var med på viktig forskning, var alle klar over at vurderingene de gjorde, ikke hadde noen konsekvenser for verken vitne eller tiltalt. I ekte rettssaker vil juryens avgjørelse kunne få mer alvorlige konsekvenser enn i et eksperiment, slik at deltakere sannsynligvis vil oppleve en ekte rettssak som viktigere enn et kort eksperiment. Informasjonsmengden vil etter all sannsynlighet også være større i en ekte rettssak, noe som kanskje kan gjøre at emosjonene til vitnet blir mindre viktig i vurderingsprosessen, etter som deltakerne får et bredere grunnlag å ta sine beslutninger på. Det kan være at deltakerne i eksperimentet lot seg påvirke av emosjoner nettopp på grunn av at de hadde lite informasjon å basere vurderingene på (Vrij, 2004). Det ble under eksperimentet observert at enkelte av deltakerne som diskuterte, kommenterte den minimale mengden med informasjon, og at den utgjorde et altfor tynt grunnlag å dømme noen på.

Som resultatene viser, kan et vitnes uttrykte emosjoner ha effekt på individuelle bedømmere og juryers vurderinger, selv om den kognitive belastningen er minimal. Økt informasjonsmengde, for eksempel i form av flere og lengre vitneforklaringer, større bevismengde og en rekke uvante prosedyrer og regler som deltakerne må forholde seg til, kunne bety økt kognitiv belastning, og dermed økt sjanse for å basere seg på bias (Reinhard & Sporer, 2008). Men det kunne tenkes at det å diskutere i jury ville medføre at hver person i juryen ville opplevd mindre kognitiv belastning sammenlignet med enkeltpersoner som ikke satt i en jury, fordi presset og belastningen i en jury ville vært fordelt på flere personer. Det

kunne vært interessant å undersøke om det ville fremkommet forskjeller mellom juryers vurderinger og vurderingene til de som vurderer uten diskusjon, dersom den kognitive belastningen hadde vært høyere.

I denne studien lot det til at flere av juryene, etter å ha diskutert en stund, tok beslutninger ved å komme frem til et gjennomsnitt av jurymedlemmenes vurderinger eller ved å bruke det mest frekvente svaret i gruppa. En slik strategi kan reflektere det at simulerte jurydiskusjoner ofte lider av tidspress slik at diskusjonen må avsluttes unormalt tidlig, og at dette fører til høyere antall «hung juries» enn i virkeligheten (Salerno & Diamond, 2010). Det ble imidlertid i dette eksperimentet ikke krevd at juryen skulle komme frem til enstemmighet, og dermed var muligheten åpen for å løse problemet med tidspress ved å la vurderingene være et gjennomsnitt av juryens meninger, eller, som i tilfellet med skyldspørsmålet, la flertallet avgjøre. Ved eventuell bruk av slike strategier er det ikke rart at juryenes vurderinger ikke skiller seg signifikant fra kontrollenes vurderinger. Den relativt korte tiden som juryene fikk til å diskutere på, kan også tenkes å ha vært en faktor som bidro til økt vektlegging av bias, ved at det representerte et økt stressmoment i likhet med økt kognitiv belastning. Selv om dette kunne vært tilfellet, forblir det allikevel bare en spekulasjon i dette studiet. Det ville vært interessant å undersøke hvordan juryens vurderinger ville vært dersom det ble gitt bedre tid (juryene diskutere i denne studien i maks 25 minutter) til diskusjon.

## 4.2 Hypotese 2

Hypotese nummer to var at vurderingen av troverdighet ville predikere vurderingen av skyld, slik at økt troverdighet ville predikere økt sjanse for skyldbedømmelse. I tråd med dette viste logistiske regresjonsanalyser at vurdering av troverdighet predikerte svar på skyldspørsmålet i kontrollgruppa, slik at økt troverdighetsvurdering av vitnet hadde sammenheng med økt sjanse for å dømme tiltalte skyldig. Dette gjaldt også for individuellbetingelsen. Disse resultatene samsvarer med det andre studier har funnet når vurderingene gjøres uten diskusjon (Cramer et al., 2009; Kaufmann et al., 2003). Heath, Grannemann og Peacock (2004) fant, som nevnt tidligere, at skyld også ble predikert av troverdighet når det var tiltalte som ble vurdert, og at oppfattelsen av tiltaltes ærlighet og sympati også fungerte som mediatorer av tiltaltes uttrykte emosjoners effekt på grad av oppfattet skyld.



## 4.3 Begrensninger

Det er allerede nevnt noen begrensninger i denne studien, som at det kanskje er kunstig lite informasjon å basere vurderingene på, at tiden deltakerne har til rådighet er urealistisk kort, og at den kognitive belastningen på beslutningstakerne kanskje er lavere enn den i virkeligheten ville vært.

Videre er effekten av uttrykt emosjon på bedømmelse av troverdighet og skyld ikke bare funnet for fornærmede, men også for tiltalte (Wessel et al., 2011). Et spørsmål vil kunne være hvordan juryens vurdering i en ekte rettssak vil påvirkes når både vitnets og tiltaltes uttrykte emosjoner er tilstede på én gang. I et slikt tilfelle ville vi fått en mer virkelighetsnær rettssituasjon og en mer balansert fremstilling av bevis, eller forklaringer. Dette vil kunne være viktig, ettersom en del forskning tyder på at bevisstyrke (Strength of evidence, Devine et al., 2001) kan ha stor betydning for juryers vurderinger. Med både vitnets og tiltaltes forklaringer å ta hensyn til, ville ikke saken blitt like ensidig fremstilt. Det kunne også vært interessant i fremtidige lignende eksperimenter for eksempel å variere styrkegraden på bevisene til fordel for vitnet, for å se om juryen da ville vært påvirket av emosjoner i like stor grad som i dette tilfellet, og undersøke i hvilken grad forskjeller mellom kontroll og jury da ville fremkommet.

Det er også mange andre variabler som i ekte rettssaker kan påvirke vurderinger av troverdighet og skyld, og som vil kunne begrense generaliserbarheten fra dette studiet. For eksempel har eksperimentet kun basert seg på ett vitne, en 11 år gammel jente. Etter som forskning viser en betydning av kjønns- og aldersforskjeller for uttrykkelse av emosjoner og bedømmelser i retten (Fuchs & Thelen, 1988; Quas, Bottoms, Haegerich & Nysse-Carris, 2002; Zeman & Garber, 1996) og troverdighetsvurderinger kan påvirkes av vitnets alder (Leippe & Romanczyk, 1989), vil det herske usikkert omkring hvorvidt resultatene fra denne studien kan generaliseres til andre tilsvarende saker der vitnet for eksempel er en gutt og/eller har en annen alder. Også fordi blant annet kjønnsstereotyper kan interagere med emosjon når det gjelder bedømmelser av for eksempel troverdighet (Schuller et al., 2010).

Dessuten dreier denne studien seg om kvinnelig offer og mannlig overgriper. Forskning kan tyde på at forskjellige kombinasjoner av offer og overgriper når det gjelder kjønn, kan påvirke troverdigheten til offeret og overgriperen. For eksempel ble det sett på som et mer alvorlig overgrep når overgriper var mann og offeret jente, enn dersom overgriper var kvinne og offeret gutt, og vitnet ble sett på som mer troverdig når overgrepet ble vurdert som mer alvorlig (Bornstein, Kaplan & Perry, 2007). Spørsmålet er da om resultatene hadde blitt

som de ble dersom for eksempel overgriperen var kvinne. Som tidligere nevnt, er det også en rekke andre faktorer som kan påvirke troverdigheten til et vitne. For eksempel var vitnet i vår studie farget, og deltakerne fikk se en film, heller enn å lese en transkripsjon, høre et lydopptak eller se vitnet live. Det er delte meninger omkring hvorvidt presentasjonsmediet påvirker vitnets troverdighet. Bornstein (1999) finner på den ene siden at de fleste studier ikke finner noen forskjeller mellom presentasjonsmedia, i dette tilfellet video, lydopptak, skrevet tekst eller live visning. Nyere studier, som en av Landström og Granhag (2010), finner på den andre siden for eksempel at det mest fordelaktige for vitnet er å bli observert live, sammenlignet med TV-overføring.

Det kan bare spekuleres på hvorfor det var en effekt av uttrykte emosjoner i denne studien. Fremtidige studier kan med fordel filme jurydiskusjonene, slik at grunnlaget for å si noe om dette blir større.

Andre innvendinger mot studien kunne vært at antallet jurymedlemmer (4-8) ikke stemmer med slik antallet faktisk er i Norge (10). Men her viser forskningen kun små forskjeller mellom juryer med 6 og 12 deltakere (Devine et al., 2001), slik at det antallet som er brukt i denne studien trolig ikke bør anses for en svakhet. Det at deltakerne i hovedsak var studenter, bør trolig heller ikke være et problem, da de fleste studier som har undersøkt forskjellen mellom lekfolk som er studenter og lekfolk som ikke er studenter, ikke finner forskjeller (Bornstein, 1999; men se Sears, 1986).

Ved å gjøre flere sammenligninger mellom grupper, slik tilfellet er i dette studiet, øker risikoen for type I-feil, der nullhypotesen forkastes når den i virkeligheten er sann. Det vil i denne studiens tilfelle si å anta at variabelen emosjon har en effekt når den i virkeligheten ikke har det. En mulig måte å redusere sjansen for type I-feil er å velge et strengere alfanivå enn for eksempel .05 (Bordens & Abbott, 2008). Av resultatene fremgår at emosjoner har effekt på troverdighetsbedømmelser i alle tre gruppene, med  $p \leq .003$ , slik at sjansen for type I-feil virker minimal i det tilfellet. For de andre avhengige variablene er  $p$ -verdien høyere, slik at sjansen for type I-feil er noe høyere i de tilfellene. Men det at det på alle avhengige variabler, med unntak av måneder straff, er en tendens til at trist emosjon er mest fordelaktig for vitnet, styrker hypotesen om at vitnets uttrykte emosjoner påvirker folks vurderinger. Dessuten mener Bordens og Abbott at anvendt forskning muligens kan benytte mindre strenge alfaverdier, fordi det kan være mer alvorlig med type II-feil enn type I-feil. Type II-feil ville her vært å konkludere med at det ikke er en effekt av emosjoner når det i virkeligheten er det. Når det gjelder anvendelse av denne studiens resultater, er det tryggere å risikere type I-feil

enn å risikere type II-feil. Det er i situasjoner hvor troverdighet og straff vurderes tryggere å være ekstra oppmerksom på at vitnets emosjoner muligens kan påvirke ens vurderinger, enn å anta at emosjonene ikke utgjør noen trussel for gode vurderinger. I førstnevnte tilfelle vil bedømmerne bare risikere å være litt ekstra forsiktige; i sistnevnte tilfelle vil bedømmerne risikere å foreta vurderinger som er ekstra utsatt for påvirkningen av vitnets emosjon, og dermed risikere å vurdere feil.

## 4.4 Konklusjon

I denne studien ble det funnet støtte til at både de som vurderer individuelt uten å ha diskutert, juryer og de som vurderer individuelt etter å ha sittet i en jury, blir påvirket av vitnets uttrykte emosjoner når de skal vurdere troverdigheten til et jentevitne i en overgrepssak, sannsynligheten for tiltaltes skyld, ta stilling til skyldspørsmålet og vurdere straffeutmåling (selv om ikke alle resultatene i jurygruppa var statistisk signifikante). Hovedtrekkene er videre at det ikke er noen signifikante forskjeller mellom vurderingene til de tre nevnte gruppene, og at et trist vitnet gir vurderinger til fordel for vitnet, og mot tiltalte. I virkeligheten vil det ikke nødvendigvis være slik at et trist vitne alltid oppfattes som mer troverdig enn et sint vitne, selv om effekten av uttrykte emosjoner er til stede. Dette er fordi omstendighetene vitnet fremstår i kan variere. Mange andre variabler vil kunne maskere en slik effekt av emosjoner som er funnet i dette studiet. Men studien vår har vist støtte til at dersom alt annet holdes konstant, vil både de som vurderer troverdighet og skyld individuelt og i en gruppe la sine vurderinger påvirkes av vitnets emosjoner.

Studien har bidratt til ny kunnskap innen vitnepsykologien når det gjelder hvordan juryer i forhold til enkeltpersoner blir påvirket av de uttrykte emosjonene til et barn som vitner om seksuelt overgrep. Dessuten har studien tilført noe nytt ved å sammenligne de uttrykte emosjonene tristhet og sinne, der det oftest har vært vanlig å sammenligne tristhet og nøytralt uttrykk. Forhåpentlig vis vil kunnskapen fra denne studien kunne bidra til en mer rettferdig og nøyaktig behandling av spørsmål om barnevitners troverdighet.

Resultatene fra denne studien viser støtte til en hypotese om at uttrykte emosjoner har betydning for troverdighet og skyld selv etter jurydiskusjon. Mer forskning er nødvendig for å styrke funnene fra denne studien og bidra til økt generaliserbarhet.



# Litteraturliste

- Ask, K. & Granhag, P. A. (2007). Hot cognition in investigative judgments: The differential influence of anger and sadness. *Law and Human Behavior*, 31, 537-551.
- Ask, K., Granhag, P. A. & Rebelius, A. (2011). Investigators under influence: How social norms activate goal-directed processing of criminal evidence. *Applied Cognitive Psychology*, 25, 548-553.
- Ask, K. & Landström, S. (2010). Why emotions matter: Expectancy violation and affective response mediate the emotional victim effect. *Law and Human Behavior*, 34(5), 392-401.
- Ask, K. & Pina, A. (2011). On being angry and punitive: How anger alters perception of criminal intent. *Social Psychological and Personality Science*, 2(5), 494-499.
- Baldus, D. C., Woodworth, G., Zuckerman, D., Weiner, N. A. & Broffitt, B. (1998). Racial discrimination and the death penalty in the post-Furman era: An empirical and legal overview, with recent findings from Philadelphia. *Cornell Law Review*, 83(1), 638-770.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. I: J. S. Uleman og J. A. Bargh (red.), *Unintended Thought* (s. 3-51). New York, NY, USA: The Guilford Press.
- Bargh, J. A., Chen, M. & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230-244.
- Bodenhausen, G. V. & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology*, 52(5), 871-880.
- Bodenhausen, G. V., Sheppard, L. A. & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology*, 24(1), 45-62.
- Bollingmo, G. C., Wessel, E. O., Eilertsen, D. E. & Magnussen, S. (2008). Credibility of the emotional witness: A study of ratings by police investigators. *Psychology, Crime, and Law*, 14(1), 29-40.

- Bollingmo, G. C., Wessel, E., Sandvold, Y., Eilertsen, D. E. & Magnussen, S. (2009). The effect of biased and non-biased information on judgments of witness credibility. *Psychology, Crime, and Law*, 15(1), 61-71.
- Bond, C. F. & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review* 10(3), 214-234.
- Bonanno, G. A., Noll, J. G., Putnam, F. W., O'Neill, M. & Trickett, T. K. (2003). Predicting the willingness to disclose childhood sexual abuse from measures of repressive coping and dissociative tendencies. *Child Maltreatment*, 8(4), 302-318.
- Bordens, K. S. & Abbott, B. B. (2008). *Research design and methods*, (7. utg.), New York, NY, USA: McGraw-Hill.
- Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, 23(1), 75-91.
- Bornstein, B. H., Kaplan, D. L. & Perry, A. R. (2007). Child abuse in the eyes of the beholder: Lay perceptions of child sexual and physical abuse. *Child Abuse & Neglect*, 31(4), 375-391.
- Bottoms, B. L., Davis, S. L. & Epstein, M. A. (2004). Effects of victim and defendant race on jurors' decisions in child sexual abuse cases. *Journal of Applied Social Psychology*, 34(1), 1-33.
- Bottoms, B. L. & Goodman, G. S. (1994). Perceptions of children's credibility in sexual assault cases. *Journal of Applied Social Psychology*, 24(8), 702-732.
- Castelli, P., Goodman, G. S. & Ghetti, S. (2005). Effects of interview style and witness age on perceptions of children's credibility in sexual abuse cases. *Journal of Applied Social Psychology*, 35(2), 297-319.
- Christianson, S.-Å., & Lindholm, T. (1998). The fate of traumatic memories in childhood and adulthood. *Development and Psychopathology*, 10, 761-780.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ, USA: Erlbaum.
- Connolly, D. A., Price, H. L. & Gordon, H. M. (2009). Judging the credibility of historic child sexual abuse complainants: How judges describe their decisions. *Psychology, Public Policy, and Law*, 15(2), 102-123.
- Cramer, R. J., Brodsky, S. L. & DeCoster, J. (2009). Expert witness confidence and juror personality: Their impact on credibility and persuasion in the courtroom. *Journal of American Academic Psychiatry and Law*, 37, 63-74.

- Czopp, A. M., Monteith, M. J. & Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology*, 90(5), 784-803.
- Dahl, J., Enemo, I., Drevland, G. C. B., Wessel, E., Eilertsen, D. E. & Magnussen, S. (2007). Displayed emotions and witness credibility: A comparison of judgments by individuals and mock juries. *Applied Cognitive Psychology*, 21(9), 1145-1155.
- Damasio, A. (1994). *Descartes' error. emotion, reason, and the human brain*. New York, NY, USA: G. P. Putnam's Sons.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111(2), 203-243.
- Devine, D. J., Buddenbaum, J., Houp, S., Studebaker, N. & Stolle, D. P. (2009). Strength of evidence, extraevidentiary influence, and the liberation hypothesis: Data from the field. *Law and Human Behavior*, 33, 136-148.
- Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R. & Pryce, J. (2001). Jury decision making. 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, 7(3), 622-727.
- Devine, D. J., Olafson, K. M., Jarvis, L. L., Bott, J. P., Clayton, L. D. & Wolfe, J. M. T. (2004). Explaining jury verdicts: Is leniency bias for real? *Journal of Applied Social Psychology*, 34(10), 2069-2098.
- Edwards, K. & Bryan, T. S. (1997). Judgmental biases produced by instructions to disregard: The (paradoxical) case of emotional information. *Personality and Social Psychology Bulletin*, 23, 849-864.
- Ekman, P. (2004). Emotional and conversational nonverbal signals. I J. M. Larrazabal & L. A. P. Miranda (red.). *Language, Knowledge, and Representation*, 39-50. Nederland: Kluwer Academic Publishers.
- Ellison, L. & Munro, V. E. (2009). Of 'normal sex' and 'real rape': Exploring the use of socio-sexual scripts in (mock) jury deliberation. *Social and Legal Studies*, 18(3), 291-312.
- Ferguson, M. J. & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, 8(1), 33-39.
- Forsyth, D. R. (2006). *Group dynamics*, (4. utg.). Belmont, CA, USA: Thomson-Wadsworth.
- Fraser, S. C., Thompson, W. & Gutowski, J. (1986). Effects of evidence strength and admissibility on mock jury verdicts. Upublisert rådata. University of California,

Irvin, USA.

- Fuchs, D. & Thelen, M. H. (1988). Children's expected interpersonal consequences of communicating their affective state and reported likelihood of expression. *Child Development* 59(5), 1314-1322.
- Gabora, N. J., Spanos, N. P. & Joab, A. (1993). The effects of complainant age and expert psychological testimony in a simulated child sexual abuse trial. *Law and Human Behavior*, 17(1), 103-119.
- Golding, J. M., Bradshaw, G. S., Dunlap, E. E. & Hodell, E. C. (2007). The impact of mock jury gender composition on deliberations and conviction rates in a child sexual assault trial. *Child Maltreatment*, 12(2), 182-190.
- Golding, J. M., Fryman, H. M., Marsil, D. F. & Yozwiak, J. A. (2003). Big girls don't cry: The effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse and Neglect*, 27(11), 1311-1321.
- Goodman, G. S., Golding, J. M., Helgeson, V. S., Haith, M. M. & Michelli, J. (1987). Jurors' perceptions of children's eyewitness testimony. *Law and Human Behavior*, 11(1), 27-40.
- Goodman, G. S. & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology*, 12, 1-9.
- Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., Prado, L., Myers, J. E. B. & Melton, G. B. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development*, 57(5), 1-159.
- Gulliksen, Øivind (2010, 28. februar). Den perfekte drapsmann. *Aftenposten*. Hentet fra <http://www.aftenposten.no/amagasinet/article3489111.ece>
- Hackett, L., Day, A. & Mohr, P. (2008). Expectancy violation and perceptions of rape victim credibility. *Legal and Criminological Psychology*, 13(2), 323-334.
- Hannaford-Agor, P. L., Hans, V. P., Mott, N. L. & Munsterman, G. T. (2002). *Are hung juries a problem?* Williamsburg: VA, USA: National Center for State Courts, National Institute of Justice.
- Hans, V. P. & Doob, A. N. (1976). Section 12 of the Canada Evidence Act and the Deliberations of Simulated Juries. *Criminal Law Quarterly*, 18, 235-254.
- Hastie, R., Penrod, S. D. & Pennington, N. (1983). *Inside the jury*. Cambridge, MA, USA: Harvard University Press.



- Heath, W. P., Grannemann, B. D. & Peacock, M. A. (2004). How the defendant's emotion level affects mock juror's decisions when presentation mode and evidence strength are varied. *Journal of applied social Psychology, 34*(3), 624-664.
- Jordan, J. (2004). Beyond belief? Police, rape and women's credibility. *Criminology and Criminal Justice, 4*(1), 29-59.
- Kalven, H. & Zeisel, H. (1966). *The American jury*. Boston, USA: Little, Brown.
- Kaplan, M. F. & Miller, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology, 53*(2), 306-313.
- Kaufmann, G., Drevland, G. C. B., Wessel, E., Overskeid, G. & Magnussen, S. (2003). The Importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology, 17*(1), 21-34.
- Kerr, N. L., MacCoun, R. J. & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103*(4), 687-719.
- Kerr, N. L., Niedermeier, K. E. & Kaplan, M. F. (1999). Bias in jurors vs bias in juries: New evidence from the SDS perspective. *Organizational Behavior and Human Decision Processes, 80*(1), 70-86.
- Kruglanski, A. W. & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*(5), 448-468.
- Landström, S. & Granhag, P. A. (2010). In-court versus out-of-court testimonies: Children's experiences and adults' assessment. *Applied Cognitive Psychology, 24*(7), 941-955.
- Langballe, Å. (2007). Forholdet mellom frie og spontane beretninger fra barn i dommeravhør, og påliteligheten i barnets utsagn. *Tidsskrift for Norsk Psykologforening, 44*(7), 868-877.
- Leippe, M. R. & Romanczyk, A. (1989). Reactions to child (versus adult) eyewitnesses. The influence of juror's perceptions and witness behavior. *Law and Human Behavior, 13*(2), 103-132.
- Lieberman & Arndt (2000). Understanding the limits of limiting instructions. Social psychological explanations for the failures of instructions to disregard pretrial publicity. *Psychology, Public Policy and Law, 6*(3), 677-711.
- London, K., Bruck, M., Wright, D. B. & Ceci, S. J. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological

- issues, and implications for forensic interviewers. *Memory*, 16(1), 29-47.
- MacCoun, R. J. (1990). The emergence of extralegal bias during jury deliberation. *Criminal Justice and Behavior*, 17(3), 303-314.
- MacCoun, R. J. & Kerr, N. L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology*, 54(1), 21-33.
- Melville, A. (2007). The crying game: An examination of how stereotypes affect witness credibility (Honors thesis, Cornell University, USA). Hentet fra <http://hdl.handle.net/1813/7820>
- Mossige, S. & Huang, L. (2010). Omfanget av seksuelle krenkelser og overgrep i en norsk ungdomsbefolkning. *Norsk Epidemiologi*, 20(1), 53-62.
- Myers, J. E. B., Redlich, A. D., Goodman, G. S., Prizmich, L. P. & Imwinkelried, E. (1999). Jurors' perceptions of hearsay in child sexual abuse cases. *Psychology, Public Policy, and Law*, 5(2), 388-419.
- Nadler, J. & Rose, M. R. (2003). Victim impact testimony and the psychology of punishment. *Cornell Law Review*, 88, 419-456.
- Neumann, R. & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, 79(2), 211-223.
- Nightingale, N. N. (1993). Juror reactions to child victim witnesses: Factors affecting trial outcome. *Law and Human Behavior*, 17(6), 679-694.
- Politidirektoratet (2011). *Kommenterte STRASAK-tall 2010*. Hentet fra: [https://www.politi.no/vedlegg/lokale\\_vedlegg/politidirektoratet/Vedlegg\\_585.pdf](https://www.politi.no/vedlegg/lokale_vedlegg/politidirektoratet/Vedlegg_585.pdf)
- Porter, S. & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, 14(1), 119-134.
- Quas, J. A., Bottoms, B. L., Haegerich, T. M. & Nysse-Carris, K. L. (2002). Effects of victim, defendant, and juror gender on decisions in child sexual assault cases. *Journal of Applied Social Psychology*, 32(10), 1993-2021.
- Reinhard, M.-A. & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology*, 44(3), 477-488.
- Regan, P. C. & Baker, S. J. (1998). The impact of child witness demeanor on perceived credibility and trial outcome in sexual abuse cases. *Journal of Family Violence*,

- 13(2), 187-195.
- Rose, M. R., Nadler, J. & Clark, J. (2006). Appropriately upset? Emotion norms and perceptions of crime victims. *Law and Human Behavior*, 30(2), 203-219.
- Salekin, R. T., Ogloff, J. R. P., McFarland, C. & Rogers, R. (1995). Influencing juror's perceptions of guilt: Expression of emotionality during testimony. *Behavioral Sciences and the Law*, 13, 293-305.
- Salerno, J. M. & Diamond, S. S. (2010). The promise of a cognitive perspective on jury deliberation. *Psychonomic Bulletin & Review*, 17(2), 174-179.
- Sandys, M. & Dillehay, R. C. (1995). First-ballot votes, pre-deliberation dispositions, and final verdicts in jury trials. *Law and Human Behavior*, 19(2), 175-195.
- Sayfan, L., Mitchell, E. B., Goodman, G. S., Eisen, M. L. & Qin, J. (2008). Children's expressed emotions when disclosing maltreatment. *Child Abuse & Neglect*, 32(11), 1026-1036.
- Schmidt, C. W. & Brigham, J. C. (1996). Juror's perceptions of child victim-witnesses in a simulated sexual abuse trial. *Law and Human Behavior*, 20(6), 581-606.
- Schuller, R. A., McKimmie, B. M., Masser, B. M. & Klippenstine, M. A. (2010). Judgments of sexual assault: The impact of complainant emotional demeanor, gender, and victim stereotypes. *New Criminal Law Review: An International and Interdisciplinary Journal*, 13(4), 759-780.
- Schulz-Hardt, S., Frey, D., Lüthgens, C. & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, 78(4), 655-669.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51, 515-530.
- Smith, A. C. & Greene, E. G. (2005). Conduct and its consequences: Attempts at debiasing jury judgments. *Law and Human Behavior*, 29(5), 505-526.
- Sporer, S. L. (1986). Justice as fair and equal treatment before the law: The role of individual versus group decision making. I H. W. Bierhoff, R. L. Cohen og J. Greenberg (red.). *Justice in social relations*, (s. 283-298). New York, USA: Plenum Press.
- Statistisk Sentralbyrå (2011). *Anmeldte lovbrudd, 2011*. Hentet fra Statistisk Sentralbyrås nettsider: <http://www.ssb.no/emner/03/05/lovbrudda/>
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3. utg.). Mahway,

NJ, USA: Lawrence Erlbaum.

- Stewart, D. D. (1998). Stereotypes, negativity bias, and the discussion of unshared information in decision-making groups. *Small Group Research*, 29(6), 643-668.
- Stiff, J. B. & Miller, G. R. (1986). "Come to think of it..." Interrogative probes, deceptive communication, and deception detection. *Human Communication Research*, 12(3), 339-357.
- Strömwall, L. & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law*, 9(1), 19-36.
- Tanford, S. & Penrod, S. (1986). Jury deliberations: Discussion content and influence processes in jury decision making. *Journal of Applied Social Psychology*, 16(4), 322-347.
- Tsoudis, O. (2002). The influence of empathy in mock jury criminal cases: Adding to the affect control model. *Western Criminology Review*, 4(1), 55-67.
- Tsoudis, O. & Smith-Lovin, L. (1998). How bad was it? The effects of victim and perpetrator emotion on responses to criminal court vignettes. *Social Forces*, 77(2), 695-722.
- Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, 9, 159-181.
- Vrij, A., Akehurst, L., Soukara, S. & Bull, R. (2004). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Human Communication Research* 30(1), 8-41.
- Wessel, E., Bollingmo, G. C., Nielsen, L. M., Sønsteby, C., Eilertsen, D. E. & Magnussen, S. (2011). The emotional witness effect: Story content, emotional valence and credibility of a male suspect. *Psychology, Crime & Law*, 18(5), 417-430.
- Wild, B., Erb, M. & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: Quality, quantity, time course and gender differences. *Psychiatry Research*, 102(2), 109-124.
- Winkel, F. W. & Koppelaar, L. (1991). Rape victims' style of self-presentation and secondary victimization by the environment: An experiment. *Journal of Interpersonal Violence*, 6, 29-40.
- Wood, B., Orsak, C., Murphy, M. & Cross, H. J. (1996). Semistructured child abuse interviews: Interview and child characteristics related to credibility of disclosure. *Child Abuse & Neglect*, 20(1), 81-92.

Zeman, J. & Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development* 67(3), 957-973.

# Vedlegg 1: Spørreskjemaet for individuell vurdering

## VITNETROVERDIGHET

Du har nå sett et utdrag fra en vitneforklaring. Du skal nå besvare dette spørreskjemaet, hvor vi ønsker din umiddelbare vurdering.

Dette er ikke en juridisk-, men troverdighetsvurdering. Vi ønsker din personlige oppfatning. Derfor er det viktig at du svarer ut fra din vurdering.

Kvinne

Mann

Alder

Arbeid:

Du skal ringe rundt 1-7 for hvert spørsmål. Ved hvert spørsmål vil det stå beskrevet hva ytterpunktene i dimensjonen er.

## BARNETS FORKLARING

Barnet fremstår som:

1----2----3----4----5----6----7  
Troverdig                      Ikke troverdig

Barnets forklaring fremstår som:

1----2----3----4----5----6----7  
Pålitelig                      Ikke pålitelig

## HELHETLIG VURDERING

Vi er her klar over at andre momenter i saken må tas med i vurderinger av skyldspørsmål, men sett at du skulle vurdere et eventuelt skyldspørsmål i en rettsak ut fra disse vitneforklaringene, hva mener du om følgende:

Jeg tror at sannsynligheten for at den faren er skyldig er:

0 % --- 10 --- 20 --- 30 --- 40 --- 50 --- 60 --- 70 --- 80 --- 90 --- 100 %

Ved skyldspørsmål ville jeg i alt stemt for:

Ikke skyldig

Skyldig

**Utuktig handling**

Ifølge straffelovens § 212,2, har utuktig handling mot mindreårig en strafferamme på inntil 3 års fengsel.

Kun dersom du har stemt for skyldig, hvor lang fengselsstraff mener du er passende:

0 --- 1mnd --- 6mnd --- 9mnd --- 1år --- 2 --- 3 --- 4 --- 5 --- 6år

Evt. annen form for straff/behandling:

**I hvilken grad er du enig i påstandene nedenfor**

**1=enig, 4=vet ikke, 7=uenig**

Barns fortellinger om overgrep er ofte falske

1-----2-----3-----4-----5-----6-----7

Barn som ikke viser stress eller tristhet forteller trolig ikke sannheten om misbruk

1-----2-----3-----4-----5-----6-----7

Hvis et barn gir en detaljrik forklaring, er det en indikasjon på at den er sann

1-----2-----3-----4-----5-----6-----7

# Vedlegg 2: Spørreskjemaet for juryens vurdering

## VITNETROVERDIGHET

Dere skal nå diskutere spørsmålene nedenfor, slik at dere kommer frem til ett svar på spørsmålene nedenfor.

Dette er ikke en juridisk-, men en troverdighetsvurdering.

Antall jurymedlemmer:

Dere skal ringe rundt 1-7 for hvert spørsmål. Ved hvert spørsmål vil det stå beskrevet hva ytterpunktene i dimensjonen er.

## BARNETS FORKLARING

Barnet fremstår som:

1----2----3----4----5----6----7  
Troverdig                      Ikke troverdig

Barnets forklaring fremstår som:

1----2----3----4----5----6----7  
Pålitelig                      Ikke pålitelig

## HELHETLIG VURDERING

Vi er her klar over at andre momenter i saken må tas med i vurderinger av skyldspørsmål, men sett at dere skulle vurdere et eventuelt skyldspørsmål i en rettssak ut fra disse vitneforklaringene, hva mener dere om følgende:

Vi tror at sannsynligheten for at den faren er skyldig er:

0 % --- 10 --- 20 --- 30 --- 40 --- 50 --- 60 --- 70 --- 80 --- 90 --- 100 %

Ved skyldspørsmål ville vi i alt stemt for:

Ikke skyldig

Skyldig

## Utuktig handling

Ifølge straffelovens § 212,2, har utuktig handling mot mindreårig en strafferamme på inntil 3 års fengsel.

Kun dersom dere har stemt for skyldig, hvor lang fengselsstraff mener dere er passende:



0 --- 1mnd --- 6mnd --- 9mnd --- 1år --- 2 --- 3 --- 4 --- 5 --- 6år

Evt. annen form for straff/behandling:

**I hvilken grad er dere enig i påstandene nedenfor**

**1=enig, 4=vet ikke, 7=uenig**

Barns fortellinger om overgrep er ofte falske

1-----2-----3-----4-----5-----6-----7

Barn som ikke viser stress eller tristhet forteller trolig ikke sannheten om misbruk

1-----2-----3-----4-----5-----6-----7

Hvis et barn gir en detaljrik forklaring, er det en indikasjon på at den er sann

1-----2-----3-----4-----5-----6-----7

# **Vedlegg 3: Informasjonsskriv gitt til deltakerne før eksperimentets start**

## INFORMASJON OM FORSØK I VITNEPSYKOLOGI

Du vil i dette forsøket se et videoklipp hvor et barn forteller om en hendelse som mistenkes å være et overgrep. Etter at du har sett denne videoen ønsker vi at du fyller ut et spørreskjema.

Dette er en anonym undersøkelse. Spørreskjemaet vil kun inneholde kjønn og alder. Det betyr at ingen data kan spores tilbake til din person.

Du vil til enhver tid ha mulighet til å trekke deg fra forsøket om ønskelig.

Hele forsøket vil vare ca 20-30 min.

# **Vedlegg 4: Informasjon gitt til deltakerne etter endt gjennomføring av eksperimentet**

## INFORMASJON OM GJENNOMFØRT FORSØK I VITNEPSYKOLOGI

Du har i dette forsøket sett en video og besvart et spørreskjema. Barnet du har sett er en skuespiller, og hendelsen barnet fortalte om er konstruert. Det betyr at barnet du har sett ikke har noen forbindelse med hendelsene hun fortalte om.

Dette forsøket vil inkludere mange forsøkspersoner. Det er ønskelig at dere ikke forteller andre om forsøkets innhold og karakter da dette kan påvirke den videre innsamlingen av data.

Vi vil med dette takke for din deltakelse i dette forskningsprosjektet. Ditt bidrag har vært viktig for oss.