

DE NORDISKE RETTSBEVISSTHETSUNDERSØKELSENE
— NOEN METODEKRITISKE BETRÅKTNINGER

AV FØRSTEAMANUENSIS LEIF PETTER OLAUSSEN

Summary

This article discusses methodological issues within the 2009 study of attitudes towards punishment in the Nordic countries. The 2009 study used surveys to compare citizens' penal attitudes with actual levels of punishment applied in connection with six relatively serious case scenarios. Preliminary reports from 2010 conclude that the majority of citizens indicate a preference for punishments that are less severe than those actually given. In the current article, the author (who participated in the 2009 study) gives two reasons for reservations concerning this conclusion. First, the list of punishments from which respondents were asked to select did not reflect the actual list of punishments available for the case scenarios in question. A post-test of the survey questionnaire among students indicates that the inclusion of answer categories unavailable in the penal law may have significantly reduced the proportion of respondents that selected unconditional imprisonment (the actual punishment for many of the scenarios). Second, respondents' possibilities for selecting a combination of any two of the seven types of sanctions create problems of commensurability when citizens' attitudes are compared with actual levels of punishment. A comparison of three different procedures tested in the preliminary analysis to handle the commensurability problem shows that the results depend very much on which procedure is used.

*The 2009 study was partially designed to test the hypothesis that judges who meet persons awaiting sentence will show more leniency than those without contact to defendants. However, the author of the current article argues that the study design, which involves use of filmed scenarios, cannot yield a reliable test of this hypothesis because such films always plant inescapable biases in the minds those who watch them.**

1. Innledning

Foreløpige resultater fra den nordiske rettbevissthetsundersøkelsen i 2009 i Danmark, Island, Norge og Sverige ble presentert på Nordisk kriminalistmøte

* Title in English: Attitudes towards Punishment in the Nordic Countries: Remarks about Methodological Problems.

19. august 2010 i København av oss som har gjennomført undersøkelsene og ble gjengitt som artikkel (Balvig m.fl. 2010) i dette tidsskriftet. De gjentas derfor ikke her. Felles for disse presentasjoner – og for de *foreløpige* rapportene¹ som ble presentert i august 2010 – er at både det metodiske opplegget av undersøkelsene og resultatene er uten drøftende vurderinger av metodevalg og konsekvenser dette kan ha hatt. Våre tilhørere og leser kan dermed ha fått inntrykk av at alle våre metodevalg er uproblematisk, eller vi kan ha skapt berettiget undring om vi selv tror at alle våre metodevalg er uproblematisk. I denne artikkelen skal jeg forsøke å rette opp et slikt mulig inntrykk ved å gjøre oppmerksom på noen metodeproblemer som det etter min oppfatning er særlig viktig å være klar over ved vurdering av de resultater som er presentert. Jeg har gradvis blitt klar over problemer mens vi har arbeidet videre det siste halvåret med forskningsmaterialet med sikte på felles opplegg for analyser av data.

Som forsker har jeg selvsagt en etisk plikt til å gjøre åpent rede for forhold som har betydning for den tillit man kan ha til konklusjoner som trekkes. Videre er en forutsetning for at publiserte resultat ikke skal bli mistolket eller misbrukt – strukket lenger enn de rekker – at metodiske svakheter er gjort åpent rede for. Både som deltaker og initiativtaker til undersøkelsene har jeg et ansvar for å bidra til slik åpenhet. For rettsbevissthetsundersøkelsene er dette særlig viktig fordi publiserte resultater kan få samfunnsmessig betydning som bidrag til offentlig meningsdannelse om straff og straffenivå, og kanskje også danne premisser for beslutninger om straff. – De synspunkter jeg presenterer, deles ikke nødvendigvis av de andre forskerne i prosjektet.

2. Straffenivået versus folks tildeling av straff - måleproblemer

Ett av våre formål var å undersøke om det er samsvar mellom straffenivået i domstolene og den straff som vanlige kvinner og menn selv synes er passende. Dette ville vi finne ut vi ved å be både profesjonelle fagdommere og tilfældige utvalg av vanlige borgere fastsette en straff for et begrenset antall konkrete saker og så sammenligne svarene deres. I prinsippet kan dette virke som en ukomplisert fremgangsmåte, men når målingene skal gjennomføres, blir det hele mer komplisert, og resultatet blir ikke nødvendigvis så klart som ønskelig.

Profesjonelle dommere kjenner naturligvis lovgivning og rettspraksis godt nok til at de ikke har store problemer med å vite ganske sikkert hvilken straff en domstol ville gitt i hver sak som inngikk i undersøkelsene. Samtlige saker som ble brukt i undersøkelsene, ville etter norske fagdommeres vurdering endt med ubetinget fengsel med en bestemt varighet. Vanlige borgere har derimot ikke særlig godt kjennskap til straffelov og rettspraksis, så de fikk mulighet til å indikere hvilken straff de ville gi ved å krysse av én eller to reaksjoner på en oppstilling av 31² alter-

nativer, inklusive ingen straff. Svaralternativene besto av vanlige straffereaksjoner samt to reaksjonsformer som *ikke* finnes blant straffelovens reaksjoner, behandling og økonomisk kompensasjon.

Dette betyr at fagdommere og vanlige borgere *ikke valgte blant samme sett av reaksjoner*, og dette skaper måleproblemer når faktisk straffenivå, anslått av fagdommere i hver sak, skal sammenlignes med folks foretrukne reaksjoner. Hvordan skal reaksjoner som behandling (i f. eks. 1 år) eller økonomisk kompensasjon (f. eks. 80.000 kr), bøter eller samfunnsstraff kunne sammenlignes med faktisk straffenivå, ubetinget fengsel (i f. eks. 8 måneder)? Dessuten ga vi respondentene i våre undersøkelser anledning til å *kombinere* hvilke som helst av de opplistede reaksjoner. Dette øker problemene: Hvordan skal en kombinasjon av *to* reaksjoner (f. eks. samfunnsstraff i 4 måneder og 120.000 kr i økonomisk kompensasjon) kunne sammenlignes med f. eks. ubetinget fengsel i 3-5 måneder? Hvilken straff er strengest?

Slike spørsmål ble viet svært liten oppmerksomhet mens undersøkelsen ble planlagt. Det ble tatt for gitt at det metodiske design som Flemming Balvig (2006) hadde benyttet, skulle repliseres uendret, og at dette var uproblematisk.

I *ettetid* har det vist seg at vi som gjennomførte de nordiske undersøkelsene i 2009, har hatt ganske ulike oppfatninger om hvordan de nevnte spørsmål best kunne løses. I hver av de midlertidige rapportene (jfr. note 1) gjør forskeren kortfattet rede for *sin* framgangsmåte og *sine* valg, men drøftinger av konsekvenser for resultatene mangler i våre rapporter, og det faktum at vi har gjort *ulike* valg kommer ikke fram, med mindre de foreløpige rapportene leses parallelt og med stor oppmerksomhet. Våre ulike tilnærminger til nevnte problem blir tatt opp i pkt. 2.2 nedenfor. Men først skal et annet, men tilhørende problem belyses: Hvilke følger hadde introduksjonen av behandling og økonomisk kompensasjon som svaralternativ på linje med ordinære former for straff, på respondentenes tilbøyelighet til å velge *ubetinget fengsel* som straff? Dette er viktig fordi ubetinget fengsel er *straffenivået* som respondentenes eget valg av straff blir sammenlignet med i rapportene.

2.1 Ekstra svarkategorier og valg av fengselsstraff

En mulig følge av å introdusere behandling og økonomisk kompensasjon helt på linje med de vanlige straffereaksjonene i spørreskjemaet som folk skulle besvare, er at dette både kan ha *redusert* prosentandelen som valgte å krysse av for ubetinget fengsel – og at lengden på ubetinget fengsel som ble valgt av de som krysset av for denne straffen, kan ha blitt *kortere* hvis fengselsstraffen ble kombinert med behandling eller økonomisk kompensasjon. For å undersøke om dette er mer enn en hypotetisk mulighet, gjennomførte jeg en test av spørreskjemaet blant grupper av studenter ved Universitetet i Oslo medio februar 2011.

Studentene ble ikke trukket statistisk tilfeldig fra noen studentpopulasjon, men var studenter som møtte til seminarundervisning utvalgte dager i to fag. Alle fremmøtte studenter valgte å delta. Jeg valgte to ulike studentgrupper som jeg antok ville ha ganske ulike oppfatninger om straff fordi dette kan ha betydning for det som skulle testes. Den ene studentgruppen antok jeg ville være ganske tilbøyelig til å velge straffer som tradisjonelt blir brukt i strafferetten. Disse studentene kan derfor kalles *tradisjonsbærere* (N=64). De studentene som jeg antok ville føle seg mindre bundet til å benytte de tradisjonelle straffene, kan kalles *tradisjonsbrytere* (N=65).

Studentene fikk utdelt et spørreskjema der de ble spurt om hvordan de ville ”dømme” i en konkret sak. Den var identisk med saken om *kioskran* i den nordiske undersøkelsen. I noen spørreskjema var svarkategoriene nøyaktig de samme (31) som i den nordiske undersøkelsen, mens kategoriene behandling og økonomisk kompensasjon var fjernet i andre spørreskjema som studentene fikk. Uavhengig av antall svarkategorier studentene kunne velge blant, kunne alle krysse av for én eller to reaksjoner, slik det også var i de nordiske undersøkelsene. Den skriftlige orienteringen om de ulike reaksjonene var den samme som i skjema benyttet i de nordiske undersøkelsene³. Alle studenter i samme seminargruppe fikk like skjema for å unngå at de oppdaget at de ikke hadde samme svarkategorier. Siden det ikke er noen grunn til å tro at det er systematiske forskjeller mellom studentene i ulike grupper, vil det være en tilfeldig fordeling blant studentene av spørreskjema med og uten kategoriene behandling og økonomisk kompensasjon. Etter at skjema var delt ut, snakket ikke studentene med hverandre. De brukte 10-15 minutter på å besvare skjemaet.

De to studentgruppene ser ut til å skille seg nokså klart fra hverandre i sine valg av straff for kioskraneren, slik jeg hadde antatt. Tradisjonsbærerne var klart mer tilbøyelige til å velge tradisjonelle straffereaksjoner. Når studentene kunne velge blant alle 31 alternativer, valgte tradisjonsbærerne i større grad bøter, fengselsstraff og økonomisk kompensasjon til offeret. De ser derfor ut til å være mer opptatt av å straffe gjerningspersonen, og at offeret har lidd overlast, mens tradisjonsbryterne ser ut til å være mer orientert mot gjerningspersonen og muligheter for resosialisering av ham, ved at de i større grad valgte betinget fengsel, samfunnsstraff og behandling. – Her vil jeg imidlertid begrense fremstillingen til å belyse hvor vidt studentenes valg av ubetinget fengsel (eller soning med elektronisk fotlenke) var betinget av om behandling og økonomisk kompensasjon var blant alternativene i spørreskjemaet. Fokus på ubetinget fengsel skyldes at dette er straffenivået for kioskranet.

Tabell 1: Prosentandel som selv ville gitt soning med fotlenke eller ubetinget fengsel som straff for kioskran.				
	Tradisjonsbærere		Tradisjonsbrytere	
	<i>Svaralternativer inkluderte behandling og økonomisk kompensasjon?</i>		<i>Svaralternativer inkluderte behandling og økonomisk kompensasjon?</i>	
	JA	NEI	Ja	NEI
Elektronisk fotlenke	1	8	5	4
Ubetinget fengsel	37	53	19	25
Til sammen	38	61	24	29
Totalt antall reaksjoner	N= 67	N= 38	N= 59	N= 56

Det framgår av tabell 1 at prosentandelen av avkryssede reaksjoner som var ubetinget fengsel eller soning hjemme med elektronisk fotlenke er *lavere* når behandling og økonomisk kompensasjon fantes blant svaralternativene enn når de ikke fantes, men forskjellen er forholdsvis liten blant tradisjonsbryterne. I denne gruppen – der studentene er generelt lite tilbøyelig å velge ubetinget fengsel – har det begrenset betydning for utfallet om behandling og økonomisk kompensasjon er med eller ikke. Blant tradisjonsbærerne – som generelt er mer tilbøyelig til å velge ubetinget fengsel – betyr det derimot svært mye om de to svarkategoriene er med eller ikke. Blant tradisjonsbærerne som *ikke kunne* velge behandling eller økonomisk kompensasjon, var til sammen 61 % av avkryssede reaksjoner ubetinget fengsel eller elektronisk fotlenke. I gruppen av tradisjonsbærere som også *kunne* velge behandling eller økonomisk kompensasjon, var den tilsvarende prosentandelen vesentlig lavere, bare 38 %.

Dette indikerer at i en populasjon der ganske mange tenker tradisjonelt når det gjelder straff og generelt aksepterer bruk av fengselsstraff, vil introduksjon av svarkategoriene behandling og økonomisk kompensasjon kunne redusere sannsynligheten ganske betydelig for at ubetinget fengsel blir valgt. Svarene fra tradisjonsbærerne antyder dessuten at introduksjon av behandling og økonomisk kompensasjon fører til at lengden på ubetinget fengselsstraff blir noe redusert, men antall svar er for lavt til kunne konkludere sikkert når det gjelder dette.

Fordi testen er gjennomført i to forskjellige studentgrupper kommer det tydelig

fram at resultatene ikke kan anvendes til å beregne hvor stort utslag det har hatt at vi benyttet spørreskjema med behandling og økonomisk kompensasjon på linje med vanlige straffereaksjoner. Vi vet ikke en gang om den vanlig befolkningen ligner mest på tradisjonsbærerene eller på tradisjonsbryterne – fordi ingen av disse gruppene er tilfeldige utvalg av den voksne befolkningen. Jeg anser det imidlertid som høyst sannsynlig at svaralternativene behandling og økonomisk kompensasjon har bidratt til at *en lavere andel av befolkningen enn ellers valgte å krysse av for ubetinget fengsel, og til at lengden på ubetinget fengsel⁴ som er valgt, har blitt noe lavere* (kanskje 1 kategori i gjennomsnitt) enn om svaralternativene ikke hadde vært sidestilt med vanlige straffereaksjoner.

Derfor blir det etter min oppfatning ikke akseptabelt å bruke *prosentandel* som ville gitt ubetinget fengsel som *samlet mål* for folks oppfatning av hva som er passende straff i en sak, og som kan sammenlignes med faktisk straffenivå (= 100 % ubetinget fengsel). For å illustrere poenget: Hvilken prosentandel for tradisjonsbærerene i tabell 1 ville det vært riktig å sammenligne med faktisk straffenivå, 38 % eller 61 %, hvis man skulle sammenlignet denne gruppens tildelte straff med straffenivået for saken? I den nordiske undersøkelsen er opplegget slik at vi bare målte det *laveste* av to slike alternativer, og da blir det svært lite betryggende bare å bruke dette. Det kan reises tvil om det var et heldig og udiskutabelt valg å inkludere behandling og økonomisk kompensasjon blant svarkategoriene. Vi burde etter min oppfatning ikke ha benyttet dem, både fordi det ganske sikkert har redusert andelen som har valgt ubetinget fengsel og trolig også varigheten av den fengselsstraff som er valgt.

Svaralternativene behandling og økonomisk kompensasjon har dessuten økt våre kommensurabilitetsproblemer – problemer med å sammenligne de straffene folks selv ville gitt med faktisk straffenivå. Hvilken reaksjon er strengest: straffenivået på 8 måneder ubetinget fengsel eller 2 år betinget fengsel og 1 år behandling, eller 1 år betinget fengsel og 150.000 kr i økonomisk kompensasjon? – Om man bare bruker prosentandel ubetinget fengsel til å måle respondentene tildelte straff, lurar man seg stilltiende unna slike plagsomme metodiske spørsmål. For ved å bruke prosentandel ubetinget fengsel som mål på respondentenes tildelte straff, forutsettes det implisitt at alle andre reaksjoner det er krysset av for er *mildere* enn ubetinget fengsel. Men det er diskutabelt med de kombinasjonsmuligheter (doble avkryssinger) som det var adgang til i de nordiske undersøkelsene.

2.2 Doble avkryssinger og multidimensjonalitet

Når straffenivået som i alle saker er ubetinget fengsel skal sammenlignes med de straffer som vanlige borgere selv har angitt i våre spørreskjema at de ville ha gitt, oppstår det nokså komplekse sammenligningsproblemer. De skyldes at re-

spondentenes tildelte straff samlet sett er *multidimensjonale*. De har dels foreslått *andre former* for straff enn ubetinget fengsel (bot, samfunnsstraff, betinget fengsel, konfliktråd, behandling, hjemmesoning med elektronisk fotlenke eller økonomisk kompensasjon til offeret), og dels har de foreslått ulike *kombinasjoner av to* reaksjoner fordi vi hadde tillatt at respondentene kunne krysse av for to ulike reaksjoner, og svært mange gjorde det. Hvordan har vi så valgt å forholde oss til de måleproblemer som oppstår når multidimensjonale forslag til straff skal sammenlignes med ubetinget fengsel av en bestemt varighet?

De som er sterkt interessert i resultatet av undersøkelsene har kanskje allerede lest de midlertidige rapportene som ble presentert medio august 2010. I så fall framgår det meget klart at det nevnte problemet er forsøkt løst på tre ulike måter. Fordi løsningsforsøkene ikke blir sett eller drøftet i sammenheng med hverandre i våre midlertidige rapporter, kan det være vanskelig å forstå hva forskjellene består i og hvilke konsekvenser løsningene kan ha for resultatene som blir presenter for hvert land. Derfor vil jeg her se løsningsforsøkene i sammenheng med hverandre og bruke norske data for å belyse spørsmålet om hvilken betydning de tre fremgangsmåtene har for resultatene.

Flemming Balvig og jeg selv har et felles utgangspunkt ved at vi begge forsøker å lage en *straffeskala* (en indeks på ordinalt målenivå) som i prinsippet gjør det mulig å ta hensyn til alle avkryssingskombinasjoner respondentene har valgt, og som også gjør det mulig å forsøke å ”vekte” ulike reaksjoner i forhold til hverandre. Det skjer ved at en ”mild” reaksjon tilordnes en lavere tallverdi enn en ”hard” reaksjon når skalaen lages, uten at avstanden mellom tallene kan tillegges noen bestemt kvantitativ betydning. Balvig formidlet sin straffeskala til oss andre i det nordiske prosjektet i et langt notat i oktober 2009 mens datainnsamlingen pågikk i de andre landene. Det var et gunstig tidspunkt fordi det ga anledning til å tenke igjennom hans straffeskala *før* det forelå data som kunne brukes til å teste virkninger av den eller av alternative skala. De måtte eventuelt lages uten kjennskap til hvilke svar respondentene ville gi. Siden jeg intuitivt fant enkelte sider ved Balvigs straffeskala mindre heldige, laget jeg et grunnlagt alternativt forslag som jeg distribuerte til de andre for å få deres vurderinger. Etter en kort E-postdiskusjon, ble vi enige om å følge hver våre analyseopplegg, hvilket vi gjorde fram til august 2010.

Hva er så likheter og forskjeller mellom Flemming Balvigs og min straffeskala?⁵ Et viktig hensyn da jeg lagde min straffeskala høsten 2009 (med utgangspunkt i Balvigs skala) var å forsøke å opprettholde alle svarkategorier vi hadde brukt i spørreskjemaet i størst mulig grad. Sammenslåinger bryter med en slags stilltiende ”kontrakt” man har med respondentene om å ivareta de svarene de har gitt, og man kan komme til å miste viktig informasjon som følge av sammen-

slåinger. Dessuten: En forutsetning for å kunne slå sammen ulike svarkategorier i vår undersøkelse er at de kategoriene som blir slått sammen, må være noenlunde "likeverdige" som straff. Å slå sammen for eksempel kategoriene 'konfliktråd' og 'bot på 6 måneders netto lønn' forutsetter at de to reaksjonene er omtrent likeverdige som straff. Fordi vi mangler grunnlag for å vite hvilke av våre mange svaralternativer som er likeverdige⁶ som straff, forsøkte jeg i størst mulig grad å unngå sammenslåinger. For øvrig er det enklest å få oversikt over forskjeller og likheter ved de straffeskala som Flemming Balvig og jeg har benyttet, ved å gi en sidestilt presentasjon av de to skalaene, se oppstilling 1 og 2.

Oppstilling 1: Balvigs straffskala for avkryssede svar i konkrete saker 2009.	
Poeng	Reaksjoner
0	Ingen straf
1	Bøde, økonomisk kompensasjon; behandling; konfliktråd
2	Betinget fængsel
3	Samfundstjeneste
4	Ubetinget fængsel < 2mdr.; elektronisk fodlænke < 2 mdr.
5	Ubetinget fængsel fra 2 til 5 mdr.; elektronisk fodlænke fra 2 til 3 mdr.
6	Ubetinget fængsel fra 6 til 11mdr.; elektronisk fodlænke > 3 mdr.
7	Ubetinget fængsel fra 1 år til 1 år 11 mdr.
8	Ubetinget fængsel fra 2 år til 2 år 11 mdr.
9	Ubetinget fængsel fra 3 år til 4 år 11 mdr.
10	Ubetinget fængsel fra 5 år og opefter
Kilde: Balvig (2010) oppstilling 5.3 side 88.	

Oppstilling 2: Olaussens straffeskala for avkryssede svar i konkrete saker.			
	Poeng		Poeng
Ikke straff	0	Elektronisk fotlenke	
Konfliktråd	1	Under 2 måneder	10
Bot		2-5 måneder	13
Inntil 1 måneds lønn	2	Over 3 måneder	16
2-5 måneder lønn	2	Ubetinget fengsel	
6 måneder lønn eller mer	3	Under 2 måneder	10
Betinget fengsel		2-5 måneder	13
Under 2 måneder	4	6-11 måneder	16
2-5 måneder	4	1 år – 1 år og 11 måneder	19
6-11 måneder	4	2 år – 2 år og 11 måneder	22
1 år – 1 år og 11 måneder	5	3 år – 4 år og 11 måneder	25
2 år – 2 år og 11 måneder	5	5 år eller mer	28
3 år – 4 år og 11 måneder	6	Behandling	
5 år eller mer	6	Under 2 måneder	2
Samfunnsstraff		2 måneder – 1 år	3
Under 50 timer	10	Over 1 år	5
50-150 timer	13	Økonomisk kompensasjon	
Over 150 timer	16	Under 50.000 kr	2
Kilde: Olaussen 2010, side 39.		50 – 100.000 kr	3
		Over 100.000 kr	5

Flemming Balvig valgte å gi samme verdi (1) for flere grunnleggende ulike reaksjoner: behandling, alle bøter og all økonomisk kompensasjon og for konfliktråd. Den *likestilling* i straffemessig henseende som derved gjøres av bøter som tilsvarer mer enn et halvt års netto inntekt og svært høye beløp i økonomisk kompensasjon

med konfliktråd, var en viktig utløsende grunn til at jeg lagde en alternativ straffeskala. I min straffeskala har konfliktråd lavere verdi enn samtlige andre reaksjoner fordi denne reaksjonen skiller seg ut ved å være betinget av partenes samtykke, og fordi den ikke medfører anmerking i strafferegisteret. Dessuten opprettholder jeg nivåforskjeller innen behandling, for bøter og for økonomisk kompensasjon. Også når det gjelder samfundstjeneste/samfunnsstraff tenker Balvig og jeg ulikt. I min skala er verdiene tilpasset verdien for den ubetingede subsidiære fengselsstraff som tilsvarer hver kategori av samfunnsstraff. Jeg legger altså til grunn at den byrde eller belastning det er å sone samfunnsstraff ble avvendt da straffen ble innført og avspeiles i reglene om subsidiær fengselsstraff, som jeg altså følger. Balvigs skala legger derimot til grunn at all samfundstjeneste skal ha lavere verdi enn den korteste ubetingede fengselsstraff, selv om dette vel bryter med danske regler om subsidiær fengselsstraff og derved ”nedgraderer” samfundstjenestens straffemessige belastning. Det fører til at mer enn 150 timer samfundstjeneste får lavere verdi (3) på Balvigs skala enn ubetinget fengsel under 2 måneder (4). Nok en forskjell består i at Balvig tildeler all betinget fengsel samme verdi (2), mens jeg har tildelt varierende verdier (4-6) avhengig av lengden på betinget straff. Jeg er usikker på om avstanden mellom verdier for betinget og ubetinget fengsel er rimelig på den skalaen jeg lagde.

En viktig likhet mellom våre skala består i at de tradisjonelle straffereaksjonene rangeres likt, fra mildest til strengest, slik det vel er vanlig innen strafferetten: bot, betinget fengsel, samfunnsstraff og ubetinget fengsel/elektronisk fotlenke. Dessuten legges det til grunn at den ubetingete fengselsstraffens varighet gir grunnlag for entydig rangering av denne straffen, og det gjøres ingen forskjell på ubetinget fengsel og hjemmesoning med elektronisk fotlenke fordi det siste bare er en soningsform⁷ for ubetinget fengsel. Fordi Balvig og jeg forholder oss ulikt når det gjelder verditildeling til andre reaksjoner enn ubetinget fengsel, må vi også benytte ulike verdier for ubetinget fengsel – og ulik ”trinnhøyde” mellom hvert nivå av ubetinget fengselsstraff – for å unngå urimelige utslag når verdien for to avkryssinger skal summeres. – I analysene bruker både Balvig og jeg (interpolert) *medianverdi* for summerte skalaverdier for *samtlig reaksjoner* som grunnlag for å sammenligne folks tildelte straff i konkrete saker med faktisk straffenivå i de samme saker⁸. Verdi for faktisk straffenivå leses ut straffeskalaen, som derved altså både bestemmer verdi for straffenivå og verdi for respondentenes tildelte straff.

Selv om både Balvigs og mitt forslag til straffeskala var kjent da datamaterialet var klare til analyse, valgte Kristina Jerre og Henrik Tham en radikalt annen tilnærming til det måleproblemet som foreligger. Grunnen til det er ukjent. Når de sammenlignet respondentenes tildelte straff med faktisk straffenivå, beregnet Jerre og Tham (2010, s. 50-53) (interpolert) *medianverdi* for *antall måneder ubetin-*

get fengselsstraff som respondentene hadde krysset av for. Dette betyr imidlertid ikke at bare de som krysset av for ubetinget fengsel inngår i medianberegningen. Svarene fra de respondenter som ikke hadde krysset av for ubetinget fengsel eller soning med elektronisk fotlenke, men for noe annet, ble kodet om til ”null måneder fengsel”, uavhengig av om respondenten hadde krysset av for en eller to reaksjoner, for eksempel ”sammållstjenst över 150 timer” og ”över 100.000 kr i økonomisk kompensasjon” til offeret. For de respondenter som både hadde krysset av for ubetinget fengsel og en annen reaksjon, ble den andre reaksjonen kodet om til ”null måneder fengsel”. Denne fremgangsmåten er ekvivalent med at alle tilleggsavkryssinger til ubetinget fengsel ble strøket, og en meget stor andel av de svenske respondentene hadde kombinert to reaksjoner (Jerre og Tham 2010, s. 51).

Når avkryssede reaksjoner for alt annet enn ubetinget fengsel ble satt til ”null måneder fengsel” og tatt med i medianberegningen, og tilleggsreaksjoner ved avkryssing for ubetinget fengsel ble strøket ved beregning av medianen, sier det seg selv at beregningsmåten som Jerre og Tham benyttet, *med nødvendighet reduserer medianverdien* i forhold til nesten enhver annen tenkelig beregningsmåte for de straffer som respondentene hadde krysset av for. En slik fremgangsmåte er ikke begrunnet metodisk, og det finnes verken drøfting av eller begrunnelse for beregningsmåten i deres rapport. Dessuten er det viktig å være klar over at deres beregningsmåte ikke har noen betydning for verdien som representerer faktisk straffenivå som respondentenes tildelte straff skal sammenlignes med. Den ligger fast, mens beregningsmåten systematisk reduserer medianen for respondentenes tildelte straff. Dette maksimerer muligheten for å få et analyseutfall der respondentenes tildelte straff framstår som ”mildere” enn faktisk straffenivå.

Det er selvsagt ikke mulig å lage slike skalaer som Balvig og jeg har konstruert uten å anvende betydelig grad av skjønn. Derfor er det i og for seg mulig å lage et stort antall ulike skalaer, som vil kunne gi ulike utslag på analyseresultatene. Det er like åpenbart at de omkodningene som Jerre og Tham har valgt å gjøre, er basert på en skjønn. Det er neppe mulig å beregne hvilke virkninger de tre beskrevne fremgangsmåtene generelt vil ha, bortsett fra at Jerre og Thams fremgangsmåte alltid vil gi redusert medianverdi. Som regel vil nok deres beregningsmåte også gi en lavere medianverdi enn skalaverdier av den typen Balvig og jeg har benyttet. Likevel er det mulig å vise at de tre beregningsmåtene har betydning for de empiriske resultater undersøkelsen gir, og derved også for hvilke konklusjoner som kan trekkes.

I tabell 2 presenteres resultater av en analyse av spørreskjema (før film ble vist) som de norske fokusgruppedeltakerne (N=116) besvarte, og som tilsvarende skjema som ble sendt ut i posten i alle nordiske land. Jeg har benyttet samtlige tre

beregningsmåter beskrevet ovenfor på dette materialet for å beregne medianverdier som mål for respondentenes tildelte straff. De kan sammenlignes med straffnivået for å finne ut om det er samsvar (eller ikke) mellom straffenivå og den straff respondentene tildelte. Beregnede medianverdier for de seks sakene gjengis i tabell 2 sammen med angivelse av straffenivået.

Tabell 2: *Medianverdi* for straff i de seks sakene som ble bedømt av fokusgruppet deltakerne i Norge – beregnet på tre ulike måter: B = Balvigs straffskala, J+T = Jerre & Thams beregningsmåte og O = Olaussens straffeskala. * indikerer straffenivå for hver sak.

Ubetinget fengsel:	Saker					
	Partner-vold	Heroin-smugling	Kioskran	Voldtekt	Bankunderslag	Gatevold
Ikke fengsel ¹⁾	J+T					J+T
Under 2 mnd (eller fotlenke)	O B	J+T				
2-5 mnd	*					B
6-11 mnd		B				O *
1 år – 1 år og 11 mnd		O	B, J+T	J+T	J+T	
2 år – 2 år og 11 mnd			O *	O B *	B	
3 år – 4 år og 11 mnd		*			O	
5 år eller mer					*	

¹⁾ Satt til null måneder fengsel i Jerre og Thams beregning av median.

Det fremgår meget tydelig at beregningsmåtene gir *svært lite robuste resultater*. Utfallet er til dels sterkt betinget av beregningsmåten som er brukt. Medianverdiene for respondentenes valg av straff for samme sak – i forhold til straffenivået – varierer etter hvilken beregningsmåte som er valgt til tross for at nøyaktig samme datamateriale ligger til grunn for alle beregninger. Men varia-

sjonen i måleresultater er ikke like stor for alle saker. Størst variasjon er det for heroinsmugling. For denne saken gir min straffeskala en median i intervallet ”1 år – 1 år og 11 måneder” ubetinget fengsel, mens Jerre og Thams beregningsmåte gir en median på ”Under 2 måneder” fengsel. Så stor variasjon er særdeles lite betryggende, men den er ikke overraskende når man kjenner forutsetningene for beregningsmåtene. Det er heller ingen overraskelse at Jerre og Thams beregningsmåte ganske gjennomgående gir måleresultater som indikerer at de straffene som respondentene selv ville gitt, er klart ”mildere” enn det straffenivået som domstolene praktiserer. Dette skyldes ikke bare respondentenes svar, men at Jerre og Tham setter alle avkryssede straffer som ikke er ubetinget fengsel, til null måneder fengsel, og samtidig inkluderer respondentene som ikke har svart ubetinget fengsel i medianberegningene. Den straffeskala jeg har benyttet, gir minst eller ingen forskjell mellom straffenivå og respondentenes tildelte straff. Balvigs skala gir lavere medianverdier større avstand til faktisk straffenivå enn min skala, eller våre skala gir sammenfallende resultater (for partnervold og voldtekt).

Beregningsmåten betyr altså ikke like mye for alle de seks sakene, og avvikene mellom medianene er klart minst for kioskran og for voldtekt. Grunnen til dette er at prosentandelen som svarer ubetinget fengsel er høy for disse to sakene, samtidig som andelen som krysses av for samfunnsstraff, behandling, økonomisk kompensasjon eller betinget fengsel er relativt lav. Det er særlig de sistnevnte tre reaksjonene – oftest i kombinasjon med annen reaksjon – som skaper de store forskjellene som kan observeres i tabell 2. Og det finnes ingen enkel eller udiskutabel ”omregningsmåte” mellom disse reaksjonene og ubetinget fengsel. Ethvert løsningsforsøk må bygge på skjønn, og det gjelder selvsagt også den straffeskalaen jeg selv har lagd.

Det forhold at min straffeskala gir resultater som i størst grad faller sammen med (det norske) straffenivået kan selvsagt ikke brukes (alene) som belegg for at denne straffeskalaen har rimelig god validitet som gjennomsnittsmål for straffer som de norske respondentene selv ville gitt. Det er imidlertid et faktum at man også får svært godt sammenfall mellom straffenivå og respondentenes tildelte *ubetingede fengselsstraffer* i det norske fokusgruppematerialet når *modus* brukes som mål, se tabell 3. Den meget gode *interne konsistens* som det altså synes å være i det norske materialet mellom *modus* for ubetinget fengselsstraff og utfallet ved bruk av min straffeskala (jfr. tabell 2) gir grunnlag for å hevde at denne straffeskalaen ikke gir systematisk andre resultater enn *modus* om man benytter *modus* for ubetinget fengsel som alternativt mål for *samsvar* (eller ikke) mellom respondentenes tilordnede straff og faktisk straffenivå. Det forhold at lekdommere svært sjelden dissenterer i spørsmålet om straff i norske domstoler tyder også på at det neppe er betydelig avvik mellom straffer som måles ut og lekdommernes oppfatninger.

Tabell 3: Prosentandel av respondentene i norske fokusgrupper som selv ville gitt ubetinget fengsel av varierende lengde. Veide tall. [Modus med **fet** skrift. * angir straffnivå.]

Ubetenget fengsel og elektronisk fotlenke	Partner-vold	Heroin-smugling	Kioskran	Voldtekt	Bankunderslag	Gatevold
Fotlenke over 3 mnd.	6	6	3	2	5	4
Under 2 mnd.	13	2	1	1		9
2 – 5 mnd	* 29	9	6	8	5	* 30
6 – 11 mnd	19	14	17	15	7	21
1 år – 1år og 11mnd	23	18	16	14	18	23
2 år – 2år og 11mnd	8	15	* 24	* 21	22	11
3 år – 4år og 11mnd	2	* 21	14	21	24	2
5 år eller mer		15	18	15	* 20	
Totalt (n)	100 (52)	100 (66)	99 (87)	97 (84)	101 (88)	100 (47)
% som selv ville gi ubetenget fengsel eller fotlenke	45 % (N=115)	57 % (N=116)	76 % (N=115)	73 % (N=115)	77 % (N=114)	41 % (N=114)
Kilde: Olaussen (2010) tabell 3.2 side 38.						

Fordi intern konsistens i analyseresultater etter min oppfatning bør telle i en validitetsvurdering av de tre beregningsmåtene, er det rimelig klart at min straffeskala faller godt ut. Bruk av så vel Balvigs skala som Jerre og Thams beregningsmåte for å beregne gjennomsnittlig (eller “typisk”) straff tildelt av respondentene ville

gitt systematisk og til dels betydelig *for lave verdier*, sammenlignet med det man ville fått om man valgte modus for ubetinget fengselsstraff som alternativt mål. Deres beregningsmåter gir mindre konsistente resultater og åpner derved for at man kan velge hvilket resultat som skal vektlegges eller presenteres. En slik valgmulighet gjør det helt nødvendig å fortelle åpent *at* man har foretatt et valg, og om mulig, begrunne rimeligheten i det valget som er tatt.

I pkt 2.1 foran ble det vist at inkludering av svaralternativene behandling og økonomisk kompensasjon i spørreskjemaene i de nordiske undersøkelsene, sannsynligvis har redusert andelen respondenter som valgte å krysse av for ubetinget fengsel. Dette medfører at alle tre beregninger av medianverdier i tabell 2 er *underestimert* i forhold til det den ville blitt dersom behandling og økonomisk kompensasjon hadde vært utelatt fra spørreskjemaet. Denne underestimeringen kan det selvsagt ikke ”justeres for” ved å bruke modus for ubetinget fengselsstraff som alternativt mål for om det er sammenfall (eller ikke) mellom straffenivå og respondenters forslåtte straff. Poenget med å trekke inn modus for ubetinget fengselsstraff er å påpeke at det ikke spiller noen rolle om jeg bruker dette statistiske målet eller min straffeskala for å avgjøre om det er sammenfall mellom respondentenes tildelte straff og straffenivå for sakene. Disse målene gir etter min oppfatning et mer dekkende bilde av de foreliggende norske data enn *bare* å rapportere at prosentandelen av respondenter som ville gitt ubetinget fengsel i flere av sakene var betydelig lavere enn 100 % (= straffenivået), selv om dette isolert sett er korrekt⁹. Ved bare å rapportere prosentandelen som krysset av for ubetinget fengsel, ville jeg samtidig ha skjult at andre statistiske mål – som modus for ubetinget fengsel og den additive straffeskalaen som også ivaretar avkryssinger for andre reaksjoner enn ubetinget fengsel – faktisk viser at det er ganske godt sammenfall med straffenivået. Det ville vært forskningsetisk uforsvarlig rapporteringsmåte.

3. Fører nærhet til tiltalte til mildere straff?

En mulig og relevant innvending overfor bruk av *spørreskjema* til å undersøke om folks forslag til straff stemmer over ens med straffenivået, er at denne fremgangsmåten er ganske virkelighetsfjern sammenlignet med det som skjer i en domstol. Der vil dommerne både se og høre tiltalte forklare seg, og det vil ofte være både vitner og offer som kan gi sin versjon av den handlingen tiltalen gjelder. Kanskje ville vi fått litt andre svar på hva som er passende straff hvis våre respondenter hadde opplevd et umiddelbart møte i retten med tiltalte? Til grunn for dette spørsmålet ligger en hypotese om at et møte mellom dommere og tiltalte – det medmenesket som dommeren skal gi en straff – vil påvirke dommeren i retning av å gi litt mildere straff enn når straff gis uten at et slikt møte finner sted, som ved bruk av spørreskjema.

Vi ville forsøke å belyse dette spørsmålet i fokusgruppedelen av undersøkelsen. Når fokusgruppedeltakerne var samlet, ble de umiddelbart bedt om å besvare vårt spørreskjema om hvilken straff de ville gitt i konkrete saker. Deretter ble de vist en film av en fiktiv retts sak om én av de sakene som inngikk i spørreskjemaet de nettopp hadde besvart. Umiddelbart etter at filmen var slutt, ble fokusgruppedeltakerne på nytt bedt om å angi hvilken straff de ville gitt tiltalte i den filmen de nettopp hadde sett. Verken spørreskjema eller film ga noen antydninger om hva som kunne være aktuell straff.

I dette designet er det lagt til grunn at hypotesen om nærhetens betydning vil bli besvart ved å sammenligne fokusgruppedeltakernes forslag til straff *før* og *etter* at de hadde sett filmen. Men er dette en holdbar antakelse?

Når fokusgruppedeltakerne foreslo straff ved å lese en sterkt komprimert handlingsbeskrivelse (av f. eks. voldtekt) i spørreskjemaet, hadde de ikke annet enn dette og sine *egne forestillinger* om tiltalte og offer å bygge på. Hva slags person var gjerningspersonen, hvor sterk og brutal var han, angret han etterpå det han hadde gjort, var han i stand til å forstå at han hadde handlet fullstendig uakseptabelt, eller var han mer eller mindre uforstående eller likegyldig til det hele? – Og hvordan så de offeret for seg? En vever, svak og sky person, eller det motsatte, var hun kanskje en utfordrende og flørtende person som bevisst eller ubevisst kan ha påvirket tiltalte slik at han handlet som han gjorde, eller var hun en person som neppe skapte usikkerhet om hun ønsket å ha samleie eller ikke? Og hvilke forestillinger hadde fokusgruppedeltakerne om følger for kvinnen av voldtekten da de besvarte spørreskjemaet før de så filmen? Handlingsbeskrivelsen i spørreskjemaet gir ikke svar på disse spørsmålene. Men man tar neppe feil om man antar at det var betydelige individuelle variasjoner i fokusgruppedeltakernes forestillinger om alt dette, og at de ganske sikkert hadde betydning for hvilken straff de foreslo.

Når handlingsforløpet (voldtekten) skulle omsettes til film av en fiktiv retts sak mot gjerningsmannen, skulle handlingsbeskrivelsen formidles via aktors og forsvarers avhør av tiltalte og offeret. Men samtidig var det selvsagt nødvendig for regissør, instruktør og skuespillere å foreta valg om hva filmen skulle formidle om tiltalte og offer som personer, langt ut over det som sies rent verbalt. Filmatisering medfører at regissøren må ta stilling til grad av sympati, følsomhet eller brutalitet som skal prege fremstilling av gjerningsmannen, og av hvordan kvinnen opplevde handlingen og de psykiske følger den fikk. Og regissøren har store muligheter til å påvirke hvilken person de som ser filmen, skal bli sterkest berørt av, offer eller gjerningsmann. Selv om det som blir sagt i filmene lojalt følger samme handlingsbeskrivelse som i spørreskjemaet, så vil måten replikker blir utalt på, ansiktsuttrykk, blick, pust, og annet kroppsspråk hos offer, gjerningsperson, aktor og forsvarer – samt bildeutsnitt, bruk av zooming, fotograferingsvinkler etc. – gjøre

at en filmversjon alltid vil være en *regissert* fremstilling av den komprimerte handlingsbeskrivelsen i spørreskjemaet. Derfor vil dette være en *annen* fremstilling enn de fullstendig *uregisserte*, individuelle forestillingene som dannet seg ved lesingen av den komprimerte handlingsbeskrivelsen i spørreskjemaet.

Siden den regisserte filmversjonen alltid er noe annet enn den komprimerte skriftlige handlingsbeskrivelsen, kan vi ikke vite hvordan vi skal tolke og forklare en eventuell forskjell i foreslått straff mellom spørreskjema og etter å ha sett film. Hvis filmen f. eks. fremstiller tiltalte som mer usympatisk, ufølsom og brutal enn folk ”i gjennomsnitt” forestilte seg ved lesingen av handlingsforløpet i spørreskjemaet, vil dette kunne føre til at de foreslår *strengere* straff etter å ha sett filmen – selv om filmen i og for seg representerer et ”møte” mellom tiltalte og de som skal fastsette straff for ham. Hvis filmen omvendt fremstiller tiltalte som mer sympatisk og angrende enn folk ”i gjennomsnitt” forestilte seg gjerningspersonen ved lesingen av handlingsforløpet i spørreskjemaet, vil dette kunne føre til at de foreslår *mildere* straff etter å ha sett filmen. Men da er det ikke opplevelse av *nærhet* til tiltalte som gir utslag i målingene, men at filmfremstillingen har gitt en *skjev* (biased) fremstilling av tiltalte i forhold til ”gjennomsnittlige” forestillinger og tanker om ham, på bakgrunn av det skriftlig fremstilte handlingsforløpet. Et slikt resultat ville verken være uventet eller spennende, men viktigst i denne sammenheng: Det ville ikke være en følge av dommernes møte med et medmenneske – men med et *sympatisk og angrende* menneske. Vi vil følgelig ikke være i stand til å finne ut om (eventuelt) mildere tildelt straff etter å ha sett filmen skyldes at dommerne fikk et ”møte” med et medmenneske i filmen, eller om det var fordi dette medmennesket var angrende og sympatisk, og at filmen var regissert slik at dommerne ikke skulle identifisere seg sterkt med offeret.

Selv om vi ikke kan finne ut det vi ønsket, kan vi likevel sikre grunnlaget for de tolkninger vi gjør av materialet ved å få utført en faglig evaluering av regien i de fire filmene som blir brukt. En fokusgruppe av personer med faglig bakgrunn i filmregi – og ikke samfunnsvitenskap – vil kunne vurdere hvilket budskap som synes å være regissert inn i filmene. En slik evaluering vil dessuten gi våre lesere et vesentlig bedre grunnlag enn nå for å vite om filmene formidler en nøytral eller vindskjev versjon av de skriftlige handlingsforløpene. Derved blir det også lettere for andre å kontrollere om våre konklusjoner er rimelig holdbare.

4. Konklusjoner

En forutsetning for å kunne sammenligne respondentenes (den voksne befolkningens) oppfatninger av passende straff med faktisk straffenivå, er at begge målinger er gjort på en slik måte at sammenligningen gir et pålitelig resultat. Det er to grunner til å betvile at denne forutsetningen holder i den nordiske undersøkelsen

av rettsoppfatninger. Det forhold at spørreskjema hadde to ekstralegale svarkategorier som ble brukt i betydelig grad av respondentene, medfører ganske sikkert at den målte oppslutningen om fengselsstraff blir lavere enn om disse svarkategoriene ikke hadde vært brukt. Muligheten til avkryssing for to reaksjoner medfører dessuten både at avkrysset fengselsstraff trolig blir kortere enn om respondentene bare hadde krysset av for én reaksjon, og at det oppstår problemer med å beregne en samlet straff for ulike kombinasjoner av to reaksjoner som kan sammenlignes med straffenivået, ubetinget fengsel. Samlet skaper dette måleproblemer som er så store at det må tas klare forbehold om sammenligning av straffenivå og målte respondentoppfatninger gir et riktig bilde. For den norske delen av undersøkelsene er det derfor ikke grunnlag for å konkludere sterkere enn at en betydelig andel av den voksne befolkningen er *nølene til bruk av fengselsstraff*, og særlig til *lange fengselsstraffer*.

Det anvendte fokusgruppedesignet gjør det ikke mulig å finne ut om et rettslig møte mellom dommere og det medmennesket som skal dømmes, leder til at straffen blir mildere. Det skyldes at den regien som nødvendigvis er lagt inn i filmene som blir brukt i fokusgruppene, er en redigert versjon av den komprimerte, skriftlige handlingsbeskrivelsen i spørreskjemaet. Dermed vil filmregien være det som gir utslag. Om filmregissøren fremstiller en sympatisk og angrende tiltalt som innser at han handlet feil, og tilstår handlingen, eller om han virker usympatisk, likegyldig i forhold til det han har gjort, og gjør offeret delvis medansvarlig for handlingen.

Åpenhet om disse svakhetene ved undersøkelsene kan naturligvis svekke undersøkelsenes kriminalpolitiske betydning, men ikke mer enn det som er rimelig i lys av svakhetene og den forskningsetiske plikt til å være åpen om dem. Dessuten har de påpekte forhold mindre betydning for andre formål med undersøkelsene. De kan fortsatt brukes til å undersøke hvilke forskjeller det (eventuelt) er i oppfatninger om straff mellom de nordiske land. Hvor like er vi egentlig? Dessuten: Noe av det mest verdifulle materiale fra fokusgruppene er timelange diskusjoner av straff i konkrete saker blant lekfolk. Hvordan tenker og argumenterer de innfor de “minioffentligheter” som fokusgruppene er?

5. Referanser

- Balvig, Flemming (2006): *Danskernes syn på straff*. Advokatsamfundet. København.
- Balvig, Flemming (2010): *Danskernes retsfølelse og retsfornuft – et forspil*. Det Juridiske fakultet, Københavns Universitet, København.
- Jerre, Kristina og Henrik Tham (2010): *Svenskarnas syn på straff*. Rapport 2010:1. Stockholms universitet, Kriminologiska institutionen. Stockholm.

Olaussen, Leif Petter (2010): *Straffenivået og folks holdninger til straff i Norge*. Rapport. Institutt for kriminologi og retts sosiologi, Universitetet i Oslo.

Balvig, Flemming, Helgi Gunnlaugsson, Kristina Jerre, Leif Petter Olaussen og Henrik Tham (2010): ”Den nordiske rettsbevissthetsundersøgelse”. *Nordisk Tidsskrift for Kriminalvidenskab*, 97. årgang, nr. 3, s. 232-250.

- 1 Balvig (2010), Jerre og Tham (2010) og Olaussen (2010).
- 2 Antallet varierte noe mellom landene.
- 3 I skjema uten svaralternativene behandling og økonomisk kompensasjon, var disse reaksjonene heller ikke med i orienteringen om reaksjonene.
- 4 Også andre reaksjoner som bøter, samfunnsstraff og fotlenkesoning (gruppert sammen) ser ut til å ha blitt litt ”mildere” når de er kombinert med behandling eller økonomisk kompensasjon.
- 5 For dokumentasjon av skalaene: se Balvig (2010) s. 86-90 og Olaussen (2010) s. 39-41.
- 6 Likeverdigheten bør vel enten fastsettes av de som skal sone straff eller i lovgivningsprosesser eller dommer der ulike former for straff blir konkret avveid i forhold til hverandre?
- 7 Dette ble også forklart for respondentene i det norske spørreskjemaet.
- 8 Se Balvig (2010) opstilling 6.1.5 s. 109 og Olaussen (2010) diagram 3.3 s. 41.
- 9 Med unntak for det som fremholdes i pkt. 2 foran.