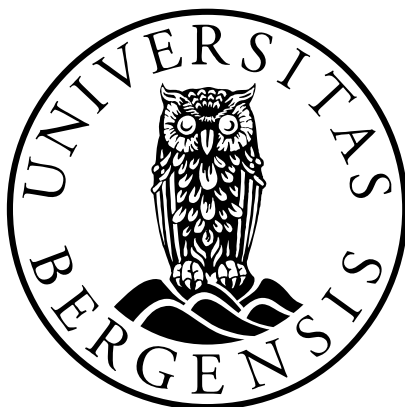


Sant nok

Kognitive feller og avhjelp i bevisvurderingen

Kandidatnummer: 103

Antall ord: 35 692



JUS396 Masteroppgave

Det juridiske fakultet

UNIVERSITETET I BERGEN

12.12.2022

«Når faktaene endrer seg, ombestemmer jeg meg. Hva gjør du?»

- John Maynard Keynes

Innholdsfortegnelse

| | |
|--|-----------|
| Innholdsfortegnelse | 3 |
| 1 Problemstilling og veien videre | 5 |
| 1.1 Tema og problemstilling | 5 |
| 1.2 Metode..... | 7 |
| 2 Sentrale begreper i bevisteori..... | 8 |
| 2.1 Hva er en bevisvurdering?..... | 8 |
| 2.2 Forholdet mellom bevisvurdering og bevistema..... | 8 |
| 2.3 Hvem er bevisbedømmer?..... | 9 |
| 2.3.1 Fagdommere og lekdommere..... | 9 |
| 2.4 Sentrale begreper innen materiell rett..... | 11 |
| 2.4.1 Bevis, bevisdata og bevissituasjon..... | 11 |
| 2.4.2 Bevisfakta og rettsfakta..... | 11 |
| 2.5 Sentrale begreper innen bevisteori..... | 12 |
| 2.5.1 Sondringen mellom den materielle retten og bevisteori..... | 12 |
| 2.5.2 Aletisk, epistemisk og psykologisk nivå..... | 12 |
| 2.6 Beviskravet..... | 13 |
| 2.7 Den frie bevisvurderingen..... | 14 |
| 2.8 Hva slags type beslutning er en bevisvurdering?..... | 15 |
| 3 Heuristikk- og skjevhetsprogrammet..... | 18 |
| 3.1 Beslutningskjeden..... | 18 |
| 3.1.1 Forholdet mellom heuristikker, skjevheter og feilslutninger..... | 18 |
| 3.1.2 Forskjellen på støy og kognitive skjevheter..... | 20 |
| 3.2 Meningsskapingsmaskinen..... | 28 |
| 3.2.1 System 1 og System 2..... | 28 |
| 3.2.2 Preferanser..... | 30 |
| 3.2.3 Forhastede slutninger..... | 34 |
| 3.2.4 Grunnleggende vurderinger..... | 36 |
| 3.2.5 Kausal mot statistisk tenkning..... | 37 |
| 3.2.6 Påliteligheten til fagdommeres intuisjon i bevisvurderingen..... | 40 |
| 4 Fortellingsperspektivet..... | 46 |
| 4.1 Samme fenomen, forskjellig språkdrakt..... | 46 |
| 4.2 Fortellinger..... | 47 |
| 4.2.1 Kjennetegn på en fortelling..... | 47 |
| 4.2.2 Hvordan fortellinger dannes i retten..... | 47 |
| 4.2.3 Generell og saksspesifikk plausibilitet..... | 48 |
| 4.3 Generell plausibilitet..... | 48 |
| 4.3.1 Hva styrker det indre signalet - koherente fortellinger..... | 48 |
| 4.3.2 Hva svekker det indre signalet..... | 56 |
| 4.4 Saksspesifikk plausibilitet..... | 59 |
| 4.4.1 Etablerte omstendigheter..... | 59 |
| 4.4.2 Heuristikker og etablerte omstendigheter..... | 60 |
| 4.5 Faren med forklaringer..... | 64 |

| | | |
|----------|---|------------|
| 4.5.1 | Etablerte omstendigheter krever en forklaring..... | 64 |
| 4.5.2 | Forståelsesillusjonen | 65 |
| 4.6 | <i>Forklaringene som bevisvurderingsmetode</i> | 68 |
| 5 | Sannsynlighetsparadigmet | 70 |
| 5.1 | <i>Sannsynlighetsparadigmet er normativt</i> | 70 |
| 5.2 | «Sannsynlighet» | 70 |
| 5.2.1 | Sannsynlighetsbegrepet er subjektivt..... | 70 |
| 5.2.2 | Objektive sannsynligheter som rettsfakta, men ikke metode..... | 72 |
| 5.2.3 | Bayes subjektive sannsynligheter og fiktive frekvenser | 73 |
| 5.3 | <i>Bayes teorem: En rasjonell måte å endre mening</i> | 74 |
| 5.3.1 | Bayesiansk analyse..... | 74 |
| 5.3.2 | Hvordan regne ut grunnfrekvensen i Bayes teorem..... | 76 |
| 5.3.3 | Hvordan anslår man beviskraften..... | 77 |
| 5.3.4 | Samlet beviskraft..... | 78 |
| 5.3.5 | Bevisverdien..... | 78 |
| 5.4 | <i>Grunnfrekvensen</i> | 79 |
| 5.4.1 | Hvorfor er det viktig å ta hensyn til grunnfrekvensen?..... | 79 |
| 5.4.2 | Neglisjering av grunnfrekvens | 81 |
| 5.4.3 | Referanseklasseproblemet..... | 82 |
| 5.5 | <i>Beviskraften</i> | 88 |
| 5.5.1 | Hva er formålet med et beviskraftsledd | 88 |
| 5.5.2 | Beviskraftsleddet flytter fokus til bevisets mengde og kvalitet | 88 |
| 5.5.3 | Feilslutninger som følge av utregningen av beviskraftsleddet..... | 89 |
| 5.5.4 | Skjevheter knyttet til anslaget av beviskraften..... | 93 |
| 5.6 | <i>Bayesiansk analyse kontra Bayes teorem</i> | 102 |
| 5.6.1 | Superanalytikere..... | 102 |
| 6 | Konklusjon | 105 |
| 6.1 | <i>En praktisk bevisvurderingsteori</i> | 105 |
| 6.2 | <i>Bedre beslutningshygiene i retten</i> | 107 |
| 6.2.1 | Gjennomgang av prosessreglene..... | 107 |
| 6.2.2 | Opplæring i fordomsfri tenkning | 107 |
| | Litteraturliste | 109 |

1 Problemstilling og veien videre

1.1 Tema og problemstilling

Tidligere førstelagmann i Borgarting lagmannsrett Nils Erik Lie har uttalt at «I praksis er den viktigste oppgaven [for den dømmende rett] å ta stilling til faktum».¹ Med dette mener han at i de aller fleste saker er rettsreglene klare. Årsaken til tvisten er at partene er uenig om hva som har skjedd, altså om sannheten. Men hvordan skal en dommer vurdere hva som er sant?

Bevisteori er fagfeltet som diskuterer hvordan man skal sikre at faktum som legges til grunn for en dom er tilnærmet hva som rent *faktisk* har skjedd. Ut fra hvor stor del av det praktiske juridiske liv som går ut på å etablere faktum, skulle man tro at bevisvurderingsmetoder var en sentral del av den juridiske utdannelsen, dommerseminarer og ble utbredt diskutert i rettsvitenskapen. Det er den imidlertid ikke.² Den litterære diskusjonen for bevisvurderingsmetoder bar preg av konsensus frem til årtusenskiftet. Konsensusen var at bevisvurderingen skulle være *fri*.³ Hvilket igjen har ført til et tilnærmet fravær av opplæring i bevisvurderingsmetoder i den juridiske utdannelsen, og til sparsomt med innslag i dommerseminarene.

Årsaken til dette påpeker Kolflaath kan være at norsk rett bygger på en forutsetning om at en normalt utrustet voksen person kan fastslå hva som er *sant* ut fra hvilke bevis som blir ført for retten ved hjelp av sin *sunne fornuft*.⁴ Diskusjonen om bevisvurderingsmetoder etter årtusenskiftet har på sin side utfordret denne forutsetningen.

Siden 80-tallet har *heuristikk og skjevhetsprogrammet* innenfor psykologisk beslutningsteori rokket ved forutsetningen om at folk flest sin sunne fornuft er egnet til å avdekke sannheten.⁵ Forskningen har avdekket at den sunne fornuft *ikke* er egnet til å vurdere hva som *sannsynligvis* har skjedd mellom partene.⁶ Dette har gitt vann på mølla i diskusjonen om

¹ (Lie, 2012) s.17 og 19

² (Kolflaath, En Metode for Bevisbedømmelsen i Straffesaker, 2015) s.509

³ (Kolflaath & Strandberg, Tre Epoker i Norsk Bevisiteori - med særlig vekt på straffeprosessen, 2015) s.42

⁴ (Kolflaath, En Metode for Bevisbedømmelsen i Straffesaker, 2015) s.509

⁵ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.163

⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.124 flg. (Kahneman, Sibony, & Sunstein, Støy, 2021) s.155

hvilke metoder bevisbedømmer rent *faktisk bruker* for å vurdere om en påstand er sann eller ikke, og hvilke metoder de *bør* anvende.

I Norsk rett har denne diskusjonen munnet ut i to paradigmer, sannsynlighetsparadigmet og fortellingsperspektivet.⁷ Sannsynlighetsparadigmet tar utgangspunkt i et *statistisk* sannsynlighetsbegrep, og legger til grunn at man skal anvende regneregler for sannsynlighet, for å sikre en rasjonell måte å vurdere om beviskravet er oppfylt. I en nordisk kontekst har sannsynlighetsparadigmet fått sin renessanse med Erling Eide og Christian Dahlman. De tar dermed utgangspunkt i den rasjonelle måten å vurdere sannsynlighet på, Bayes teorem. Og argumenterer for å anvende teoremet som en generell bevisvurderingsteori.⁸

Fortellingsperspektivet er utpreget deskriptivt. Det vil si at det har som formål å beskrive hva bevisbedømmere i realiteten gjør når de resonner og diskuterer faktum i prosessen frem mot avgjørelsen.⁹ Fellestrekk for teoriene er at sannsynlighet tolkes som hvor *plausibel* en påstand er. Man stiller dermed plausibiliteten av partenes fortellinger opp mot hverandre for å vurdere hvilken som er sann.¹⁰ Sentrale norske verker innenfor denne tradisjonen er bøkene til Eivind Kolfaath og Andre Kjus, samt en nyere doktorgrad av Line Norman Hjort.¹¹

Denne oppgaven har som formål å sammenlikne disse to paradigmene ut fra hvorvidt de i *praksis* er egnet til å etablere et faktum som er tilnærmet sannheten. En sentral forutsetning for at en bevisvurderingsmetode skal fungere i praksis, er at metoden må ta utgangspunkt i dommerens kognitive evner. Og siden dommerne, fagdommere som lekdommere er mennesker er de sårbare for kognitive skjevheter. Denne oppgaven vil drøfte hvorvidt de etablerte teoriene kommer med realistiske eller urealistiske forventninger til hvor stor grad dommerne skal greie å våke over sin begrensede rasjonalitet. For å oppsummere med Strandbakken sine ord «En bevisteori som er så komplisert at lekdommere – men også fagdommere – ikke har forutsetninger for å anvende den, er uten verdi».¹²

⁷ (Kolflaath, 2004) s.279

⁸ (Eide, 2016) (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018)

⁹ (Kolflaath, 2004) s.280

¹⁰ (Strandberg M., 2012) s.160

¹¹ (Kolflaath, Bevisbedømmelse i praksis, 2013) (Hjort, 2021) (Kjus, 2008)

¹² (Strandbakken, 2003) s.237

Problemstillingen denne oppgaven skal besvare er da følgende:

I hvilken grad åpner og avhjelper de etablerte bevisvurderingsteoriene for kognitive skjevheter.

1.2 Metode

Denne oppgaven er en teoretisk analyse av hvordan bevisbedømmer vurderer og burde vurdere bevis. Problemstillingen er tverrfaglig. Den krever at man veksler på begreper fra bevisteori, kunnskapsfilosofi og beslutningsteori. I kapittel 2 introduserer jeg begrepene fra bevisteori og kunnskapsfilosofi. Mens jeg i kapittel 3 introduserer grunnleggende beslutningsteori. På grunn av det beskjedne fokuset på beslutningsteori i norsk rett, bruker jeg en del plass på å forklare implikasjonene denne forskningen har for bevisbedømmere. Selve besvarelsen på problemstillingen starter i kapittel 4. Der drøfter jeg hvordan fortellingsperspektivet åpner og avhjelper kognitive skjevheter. Tilsvarende analyse blir gjort av sannsynlighetsparadigmet i kapittel 5. I kapittel 6 konkluderer jeg med at en praktisk anvendbar bevisteori bør ta utgangspunkt i det bevisbedømmer intuitivt gjør i bevisbedømmingen, fortellingsperspektivet. Men at man bør aktivt innarbeide metoder for å rette opp skjevheter som bayesiansk analyse.

2 Sentrale begreper i bevisteori

2.1 Hva er en bevisvurdering?

Prosessen der påstander blir til faktum kalles *bevisvurderingen*. Bevisvurderingen er hvordan bevisbedømmeren vurderer om bevistemaet er tilstrekkelig bevist av bevisene. Med andre ord, for at rettsanvender skal kunne foreta en juridisk vurdering av om noe er i tråd eller i strid med rettsreglene, må *bevisbedømmer* først konkludere med hva som rent faktisk har skjedd i saken. Dette kalles et faktum. I retten etableres faktum ved at partene fremsetter faktumpåstander om hva som har skjedd. Deretter fremlegges bevis som støtter opp under disse påstandene. Bevisvurderingen er altså vurderingen bevisbedømmer foretar av om faktumpåstanden skal legges til grunn som faktum eller ikke.

2.2 Forholdet mellom bevisvurdering og bevistema

Et bevistema defineres som «noe som skal bevises».¹³ *Hva* som skal bevises er en påstand. Påstandens innhold er at «noe er tilfelle, var tilfelle, vil bli tilfelle eller ville blitt tilfelle». Videre må påstandens innhold være relevant for en påstått rettsvirkning. Eksempelvis, en skadelidt har falt ned i en grop på et fortau og brukket armen. Skadelidte legger ned påstand om at rørleggingsarbeidet under fortauet ikke var merket, og dette var årsak til at han falt ned i gropen. Dersom påstanden blir lagt til grunn som sann blir rettsvirkningen at rørleggingsselskapet får erstatningsansvar ovenfor skadelidt.

Sammenhengen mellom bevisvurderingen og bevistema. Bevistemaet dikterer først hva som er relevante bevis, og for så å si hva beviset skal tolkes til å indikere. Bevis taler ikke for seg selv, de trenger konteksten av bevistema for å sorteres og gis meningsinnhold. Det er etter at beviset har blitt gitt relevans i konteksten av bevistemaet, at bevisvurderingen kommer inn for å vurdere om beviset kan tolkes til å indikere hvorvidt påstanden er sann. Denne oppgaven vil på sin side ikke diskutere inndelingen av bevistema

¹³ (Strandberg M. , 2012) s.411.

2.3 Hvem er bevisbedømmer?

2.3.1 Fagdommere og lekdommere

Det mest umiddelbare eksempelet på bevisbedømmelse er dommeren som vurderer bevis i en rettssak. Denne analysen vil derfor begrense seg til å vurdere bevisbedømmelsen som *dommere* i domstolssystemet foretar. Det er hovedsak er dommerens bevisbedømmelse som er omtalt i den kognitiv beslutningsteori. Det er imidlertid ikke urimelig å anta at mange av de kognitive feilslutningene som dommere begår, også gjør seg gjeldene for andre juridiske beslutningstakere som saksbehandlere, påtalemyndigheten o.l.

Det første spørsmålet er hvem som er som er dommere i rettsalene i Norge. De alminnelige domstolene i norsk rett er Høyesterett, lagmannsrettene og tingrettene jf. domstoloven §1.¹⁴ I straffesaker settes tingrettene og lagmannsrettene med en kombinasjon av fagdommere og lekdommere. I sivile saker settes retten med fagdommere, og unntaksvis med innslag av lekdommere. Hovedregelen i sivile saker er at tingretten settes med én fagdommer, mens lagmannsretten settes med tre jf. domstoloven §21 (2) 1.pkt. og §12 (1). Meddommerne skal utgjøre en del av retten når loven særskilt krever det, jf. domstoloven §14 og §21 (2) 4.pkt. I Høyesterett har man kun fagdommere.

En fagdommer er dommer av yrke, og må ha avlagt juridisk embetseksamen eller ha mastergrad i rettsvitenskap jf. domstoloven §54. Mastergrad i rettsvitenskap fokuserer i hovedsak på materiell rett og metode. Når det kommer til bevisretten og bevisteori, inneholder de obligatoriske fagene kun en innføring i de prosessuelle reglene. Opplæringen for dommerfullmektige og dommere har imidlertid et innslag i beslutningsteori, men kun en enklere innføring. Totalt sett har fagdommere lite direkte opplæring i hvordan unngå kognitive feller eller håndtere usikker informasjon.

En lekdommer har dommeroppgaven som verv. Vedkommende har enten blitt oppnevnt på bakgrunn av spesifikk fagkunnskap jf. domstoloven §72, eller folkevalgt av kommunestyret eller fylkestinget.¹⁵ Ordlyden «lek» tilsier ufaglært. Det er dermed ikke noe krav om at dommeren har juridisk bakgrunn. Dette er imidlertid ikke til hinder for at lekdommeren er

¹⁴ Lov om domsstolene av 13. august 1915 nr.5

¹⁵ (NOU 2002:11 "Dømt av likemenn") kap. 4.3

jurist, men det er i så fall en tilfeldighet. En lekdommer kalles meddommer når vedkommende utgjør en del av retten sammen med fagdommere.¹⁶

Lekdommeres tilstedeværelse i norsk rett har lange historiske røtter. Opprinnelig var det kun lekdommere som dømte i norsk rett.¹⁷ Idealet var det alminnelige prinsippet at man skal «dømmes av sine likemenn». I dag er det fremdeles et betydelig innslag av lekdommere i norsk retten, selv om jury ordningen ble avviklet i 2018.¹⁸

I sivile saker settes tingretten som regel kun med én fagdommer, men dersom retten eller partene finner det ønskelig skal det også settes to meddommere, jf. tvisteloven. §323.¹⁹ Lagmannsrettene settes med fire meddommere, dersom en av partene krever det etter tvl. §324. I straffesaker skal det *alltid* være to meddommere til stede i tingretten jf. straffeprosessloven §276 (1). I lagmannsretten varierer innslaget av meddommere ut fra hva anken gjelder. Gjelder anken bevisbedømmelsen under skyldspørsmålet, avgjørelse om straff eller straff for forbrytelse som kan medføre 6 år eller mer i fengsel, settes retten med to fagdommere og fem lekdommere jf. straffeprosessloven §332.²⁰ Sammen utgjør lekdommere og fagdommere et kollegium, der alle har individuell stemmerett.

Det vil si at etableringen av faktum i en sak ikke er en individuell oppgave, men skjer via diskusjon mellom lekfolk og utdannede jurister. For det første gjør det store innslaget av lekfolk at man har et begrenset handlingsrom når det kommer til opplæring i bevisvurderingsmetoder for å unngå kognitive skjevheter. For at man skal anvende den formelle formelen for Bayes Teorem, kreves det på sin side inngående forståelse av sannsynlighet og statistisk tenkning. En praktisk bevisvurderingsmetode bør dermed ligge tett opp mot det folk flest *gjør* når man vurderer sannsynligheter.

For det andre kompliserer diskusjonsaspektet bevisbedømmelsen. Psykologiske skjevheter påvirker tolkningsprosessen på to nivåer. Først får tolkningen slagsider når bevisene blir presentert for retten. Videre får vurderingen ytterligere slagsider når kollegiet diskuterer saken etter hovedforhandlingene. En god bevisvurderingsmetode må derfor ta høyde for hele prosessen som fører til etableringen av faktum.

¹⁶ (Øyen , 2019) s.133

¹⁷ (Øyen , 2019) s.134

¹⁸ (Domstol.no, 2021)

¹⁹ Lov om mekling og rettergang i sivile tvister av 17. juni 2005 nr. 90

²⁰ Lov om rettergangsmåten i straffesaker av 22. mai 1981 nr. 25

2.4 Sentrale begreper innen materiell rett

2.4.1 Bevis, bevisdata og bevissituasjon

Jeg har til nå gjort rede for hvem som avgjør faktum, dommerne, og rammen faktumet bestemmes innenfor, bevisvurderingen. I det følgende vil jeg fortsette med å gjøre rede for sentrale begreper som beskriver de grunnleggende bestanddelene som inngår i beslutningsprosessen, og som gjør at dommeren kan komme frem til et faktum.

En dommer som skal avgjøre en rettstvist, har ikke direkte tilgang til hva som har hendt. Hva som utgjør faktumet i saken, må utledes på bakgrunn av *bevis*. Magne Strandberg legger til grunn at «bevis er et fenomen som bærer en informasjon som tilsier at en faktumpåstand er sann eller usann».²¹ Et bevis presenteres for bevisbedømmer i form av et *bevismiddel*, som er et medium som bærer informasjonen til beviset. Strandberg kaller denne informasjonen for *bevisdata*. Dersom man legger sammen alle bevisene som bevisbedømmer kan legge vekt på, for å avgjøre om faktumpåstanden er sann eller usann, får man den gitte *bevissituasjonen*.

2.4.2 Bevisfakta og rettsfakta

Fra bevisdata kan man slutte til bevisfakta, som man igjen kan trekke en slutning til rettsfaktum. Bevisfakta er faktiske forhold som indikerer et rettsfaktum. Rettsfakta er den videre slutningen fra bevisfakta om de faktiske forhold har direkte relevans for om en rettsregel kommer til anvendelse.²² Eksempelvis er bevisdatumet at «vitnet Truls sier han så Peder Ås slå Erik Vold med et balltre i hode». Datumet leder til et bevisfaktum dersom slutningen man trekker er at «Peder Ås slo Erik Vold i hode med et balltre». Bevisbedømmer legger altså til grunn det vitnet sier han har observert som *fakta* for hva som har skjedd. Bevisdataen er altså tilstanden informasjonen er i, før den blir tolket av en bevisbedømmer. Informasjonen utgjør så et rettsfaktum. Eksempelvis dersom man ser informasjonen opp mot straffelovens §273 om kroppsskade, som sier at det er straffbart å «skade andres kropp». Tolker bevisbedømmer hendelsen, som nå har blitt lagt til grunn som bevisfakta, til å passe innunder ordlyden «skade andres kropp».

²¹ (Strandberg M. , 2012) s.49

²² (Strandberg M. , 2012) s. 51

2.5 Sentrale begreper innen bevisteori

2.5.1 Sondringen mellom den materielle retten og bevisteori

Tolkningsprosessen fra bevisfakta til rettsfakta, er også en sondering fra hva *bevisteori* diskuterer, til hva den *materielle* retten regulerer. Det som avgjør om bevisfakta utgjør rettsfaktum, er som nevnt om det kan subsumeres under ordlyden til en materiell rettsregel. Prosessen fra bevisfakta til rettsfakta reguleres altså av den materielle retten. Hva som utgjør relevant bevisfakta inngår under et større spørsmål om *hva* bevisbedømmer kan vite om hva som faktisk har hendt, i ettertid. Det vil si man beveger seg over i kunnskapsfilosofiens rike, om *hva man kan vite noe om*. Dette diskuteres i bevisteorien, som låner begreper fra kunnskapsteori. Siden denne oppgaven skal analysere den psykologiske prosessen dommeren gjennomgår i bevisvurderingen, er det relevant å utdype begreper innenfor bevisteori, og avgrense mot den materielle retten.

2.5.2 Aletisk, epistemisk og psykologisk nivå

For å kunne vurdere hva dommeren kan *vite* noe om er det ikke tilstrekkelig å skille mellom regelnivå, faktanivå og påstandsnivå, man må også skille mellom det aletiske, epistemisk og psykologiske nivået.²³ Det *aletiske* nivået betegner verden slik den er i seg selv, de faktiske forhold. Vi har imidlertid ikke direkte tilgang til verden slik den er i seg selv. Man må utlede hvordan den er gjennom informasjon. Dette skjer på det *epistemiske* nivået. Det er her vi finner bevis. De indikerer noe om verden slik den *er*, men gir ikke direkte uttrykk for den. For å oppnå kunnskap om hvordan verden er, må man tolke bevisene. Denne tolkningsprosessen skjer på det *psykologiske* nivået, og avhenger av individet som tolker bevisene.

Målet med bevisvurderingen er å avdekke hva som har skjedd på det aletiske nivået. Bevisbedømmer har på sin side kun direkte tilgang på det psykologiske nivået. Slik jeg ser det utgjør da det epistemiske og psykologiske nivået mulige *feilkilder*. Man risikerer at bevisene på det epistemiske nivået ikke er representative for det aletiske nivået. Eksempelvis kan et vitne lyve, eller et blodspor kan gi falske positivt resultat. Det er dermed en feil med selve testen/beviset som skal representere den aletiske nivået. Men bevisbedømmer kan også tolke

²³ (Strandberg M. , 2012) s.53

bevisene feil på det psykologiske nivået. Det at bevisbedømmer tolker et bevis feil kan skyldes en *kognitiv skjevhet*. Hvilket er fenomenet denne oppgaven skal drøfte.

2.6 Beviskravet

Siden dommeren ikke har direkte tilgang på det aletiske nivået, foreligger det en risiko for at man legger et usant premiss til grunn for rettsavgjørelsen. Dette kan være et resultat av en utilfredsstillende bevisførsel, altså en mangel på det epistemiske nivået. Eller at dommeren tolker representative bevis feil, herunder skjevhet på det psykologiske nivået. Ifølge Strandberg må derfor *bevisrettens* mål være «å gi dommer som er basert på en rasjonell tilnærming til denne usikkerheten».²⁴

Bevisretten har derfor etablert *beviskrav* for å sikre seg en ramme for rasjonell tilnærming til denne usikkerheten. Her må jeg understreke skillet mellom *bevisretten* og *bevisteorien*. Bevisteori er fagfeltet som diskuterer hvordan dommere rasjonelt skal la seg overbevise. Mens bevisretten er de rettslig bindene reglene som utgjør rammeverket for hvordan dette skal gjøres.

Beviskravet angir en terskel for når faktumpåstanden er sterkt nok bevist, slik at dommeren kan legge til grunn at den er sann.²⁵ Hvor sterkt faktumpåstanden må bevises varierer ut fra om man er på sivilprosessens eller straffeprosessens område. På sivilprosessens område er hovedregelen sannsynlighetsovervekt (overvektsprinsippet) jf. blant annet Rt.1992 s.64 (P-pilledom II). Det vil si at dommeren må føle seg mer enn 50 prosent overbevist om at hendelsen har inntruffet for å legge den til grunn. Innenfor strafferetten må dommeren være overbevist utenfor enhver «rimelig tvil» jf. blant annet Rt.2001s.1521.

Beviskravsmodellen er en enten-eller-modell. Det vil si at så lenge påstanden er bevist over minimumsterskelen, realiseres hele rettsvirkningen, uavhengig om faktum har sterke eller relativt svakere bevis for seg.²⁶ Beviskravet gir dermed uttrykk for hvordan retten fordeler risikoen mellom partene. Jo strengere beviskrav, jo mindre risiko aksepterer retten. I sivilretten tilsier overvektsprinsippet at man initielt fordeler usannhetsrisikoen likt mellom partene. Dette bygger på en forutsetning om at man anser partene som likestilte, og at de

²⁴ (Strandberg M. , 2012) s.17

²⁵ (Strandberg M. , 2012) s.18

²⁶ (Strandberg M. , 2012) s.19

rammes likt av usannhetsrisikoen. I strafferetten er partene på sin side ikke likestilt. Påtalemyndigheten har et større ressursapparat og risikerer ikke personlige konsekvenser med å tape saken. Tiltalte risikerer på sin side straff. Derfor styres det strafferettslige beviskravet etter at påstanden må være bevist «utenom enhver rimelig tvil». Usannhetsrisikoen bæres da av påtalemyndigheten.

Beviskravet er altså regulert av rettsregler, mens selve bevisvurderingen er fri.²⁷ Det vil si at *hvordan* dommeren lar seg overbevise ikke er regulert av bevisretten. Det vil si at jeg tar som en forutsetning at det å redusere kognitive skjevheter i bevisvurderingen er i tråd med å fremme bevisvurderingens overordnede mål om å fremme det «materielle sannhets prinsipp».²⁸ Men at jeg ikke tar stilling til hva som skjer med denne målsetningen, når usannhetsrisikoen skal fordeles. Oppgaven fokuserer da på tolkningsprosessen opp mot vurderingen om beviskravet er oppfylt.

2.7 Den frie bevisvurderingen

Hvordan dommeren lar seg overbevise er som sagt ikke regulert av rettsregler. Tvl. §21-2 sier at «retten fastsetter ved en *fri bevisvurdering* det saksforholdet avgjørelsen skal bygge på». Ordlyden «fri» tilsier at det er opp til den enkelte dommer å avgjøre hvilke bevis hun skal la seg overbevise av. Dette støttes også av forarbeidene som sier at den frie bevisvurderingen innebærer at det «ikke er regler for den vekt enkelte bevis skal tillegges».²⁹ Siden det i begrenset grad blir ført direkte bevis for Høyesterett har det i liten grad utviklet seg noen retningslinjer for den frie bevisvurderingen i praksis.³⁰ Det finnes ingen tilsvarende bestemmelse som tvl. §21-2 i straffeprosessloven. Dette kommer av at regelen ble ansett som «overflødig» å lovfeste.³¹

Prinsippet om fri bevisbedømmelse kan deles inn i to underordnede prinsipper: Fri bevisvekting og fritt metodevalg.³² Fri bevisvekting innebærer at det ikke foreligger noen rettsregler for bevisverdien til bestemte typer bevis. Eksempelvis at rettsregler tilsier at DNA-bevis veier tyngre enn vitneutsagn. Bevisbedømmer står fritt til å vekte bevisene slik hen ser

²⁷ (Strandberg M., 2012) s.79 s.309

²⁸ (Kjelby, 2015) s.84

²⁹ (Schei, et al., 2001) NOU 2001:32 A Rett på sak s. 457

³⁰ (Backer, 2015) s.310

³¹ NUT 1969:3 s.197

³² (Kolfaath, 2011) s.161

hensiktsmessig ut fra bevissituasjonen. Prinsippet om metodefrihet går på sin side ut på at fremgangsmåten i selve bevisbedømmelsen, ikke er forhåndsbestemt. Det er fremgangsmåten denne oppgaven interesserer seg for, og vil drøfte hvorvidt det er hensiktsmessig med et helt fritt metodevalg.

Prinsippet om fri bevisbedømmelse blir forsvart i norsk rettslitteratur med at forhåndsbestemt bevisvekting gir for lite fleksibilitet i møte med ulike bevissituasjoner.³³ Bevisbedømmelsen blir dermed mer treffsikker dersom man kan tilpasse bevisvektingen til situasjonen. Dette er imidlertid et forsvar av fri bevisvekting, ikke metodevalg. Hva som er begrunnelsen for fritt metodevalg, er på sin side uklart. Dersom det overordnede formålet med bevisvurderingen er å trekke en slutning på det psykologiske nivået som er tilnærmet realiteten på det aletiske, forutsetter fritt metodevalg at valgfrihet vil være det mest sannhetsfremmende.

Kolflaath argumenterer for at en forutsetning om fritt metodevalg er mest sannhetsfremmende er tvilsom da det må bygge på et uholdbart resonnement om at alle metoder er like sannhetsfremmende. Eller at man i det minste kan *vite* om en metode er mer sannhetsfremmende enn en annen. Eller at bevisbedømmer selv alltid vil lande på den mest sannhetsfremmende metoden. Denne oppgaven argumenterer videre for at enkelte bevisvurderingsmetoder er mere sannhetsfremmende enn andre. Fordi enkelte åpner for flere kognitive feller som gjør dem praktisk lite anvendelig. Dette er gitt forutsetningen om at man kommer nærmere «sannheten» med beslutninger som er fri for psykologiske skjevheter.

2.8 Hva slags type beslutning er en bevisvurdering?

Denne oppgaven skal evaluere hvilken av de to bevisvurderingsmetodene som er mest praktisk anvendbar, ut fra hensynet til kognitive skjevheter. For å kunne foreta en evaluering av metodene, må man først ha det klart for seg hva slags type beslutningsprosess bevisvurderingen er.

Som det følger av ordlyden er bevis«vurderingen», et *vurderingsspørsmål*. Et vurderingsspørsmål plasserer seg et sted imellom rene fakta og beregningsspørsmål på den ene siden og mot smak og meninger på den andre. Et vurderingsspørsmål defineres dermed av

³³ (Kolfaath , 2011) s.161

en forventning om avgrenset uenighet.³⁴ Ideelt sett skulle enhver dommer kommet frem til samme overbevisningsgrad gitt bevissituasjonen, men i realiteten vil det oppstå variasjon.

Variasjonen vil i hovedsak være et resultat av to faktorer. For det første er det forskjeller i tolkningsprosessen på de psykologiske planet, herunder psykologiske og statistiske skjevheter. For det andre oppstår det problemer på det epistemiske nivået, ved at man ikke kan få noe bekreftelse på om hypotesen bevisene skal støtte stemmer. En vurdering som blir bekreftet på det epistemiske nivået er for eksempel en værmelding. Prediksjonen om at det blir 25 grader og sol over Oslo i morgen, blir bekreftet eller avkreftet når morgendagen kommer. Men hvem som drepte Rigmor Johansen, er det bare morderen selv som vet. Og selv om morderen tilstår, kan ikke dommeren vite med sikkerhet at vedkommende snakker sant. Det er altså en prediksjon som ikke kan kontrolleres.

Hva er det da som gjør at dommeren sitter med en følelse av å ha kommet frem til «riktig» svar, når svaret egentlig ikke er noe som kan vites? Beslutningsteori påpeker at hvorvidt en vurdering er verifiserbar eller ei, endrer ikke vurderingserfaringen. Man føler seg like kompetent og tilfreds med svaret uavhengig av om det kan bekreftes.³⁵ Det man søker etter for å vurdere om man er tilstrekkelig overbevist, er et *indre signal* på om vurderingen er fullført. Hvor sterkt det indre signalet er avhenger av hvor *koherent* man opplever løsningen, eller i rettens tilfelle, fortellingen.

Men dersom fasiten på beslutningen ikke kan vites, hvordan skal man evaluere kvaliteten på beslutningen? For å sikre forutberegnelighet må bevisvurderingsmetodene ha som målsetting å minimere variasjonen mellom tilsvarende beslutninger. Siden man ikke kan gjøre noe med problemet på det epistemiske nivået, må man prøve å minimere variasjonen på det psykologiske planet. En tradisjon som er særlig utbredt innen beslutningsteori er å måle hvorvidt beslutningen overholder prinsippene for logikk og sannsynlighetsregning. Som nevnt i forrige delkapittel lar man prinsippene for logikk og sannsynlighetsregning utgjøre fasiten på «riktig» tenkemåte. *Avstanden* mellom sannsynligheten som faktisk blir anslått, og sannsynligheten som kan regnes ut ved hjelp av sannsynlighetsregneregler, kalles en *kognitiv skjevhet*.

³⁴ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.49

³⁵ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.54

Man må dermed forbedre beslutningene ved å fokusere på prosessen fremfor resultatet. Dette er en robust fremgangsmåte av to hensyn. For det første har man som nevnt ikke tilgang på resultatet. Og for det andre så vet man ikke i hvor stor grad et resultat er en konsekvens av tilfeldigheter. For å si det litt flåsete, dersom man setter mange nok sjimpanser over lang nok tid ved en skrivemaskin vil en av dem til slutt ha skrevet Hamlet.³⁶

³⁶ (Taleb, Fooled by Randomness, 2004) s.136 flg.

3 Heuristikk- og skjevhetsprogrammet

3.1 Beslutningskjeden

3.1.1 Forholdet mellom heuristikker, skjevheter og feilslutninger

I kapittel 2 drøftet jeg *hva* som skal til for at en påstand kan legges til grunn som faktum.

Videre konstaterte jeg at *hvordan* dommeren kan la seg overbevise, ikke er regulert av rettsregler. Videre kategoriserte jeg *hvilken type vurdering* bevisvurdering er, for å kunne diagnostisere hvordan den menneskelige hjernen foretar slike vurderinger. I dette kapitlet skal jeg introdusere det jeg har valgt å kalle *beslutningskjeden* og det Kahneman kaller *meningsskapingmaskinen*, den forklarer mekanismene bak heuristikkene bevisbedømmer bruker når hen skal vurdere sannsynlighet. Kapitlet dreier seg altså om hvordan hjernen intuitivt foretar sannsynlighetsvurderinger.

Den mest treffende forskningen på intuitive sannsynlighetsvurderinger er bedre kjent som *heuristikk og skjevhetsprogrammet*, som har fått to Nobels ærespriser i økonomi.³⁷

Fellesnevneren for forskningen på heuristikk og skjevhetsprogrammet er at de bruker samspillet mellom *System 1* og *System 2* til å forklare hvordan man i beslutningssituasjoner anvender *heuristikker*. Dersom beslutningen ikke er presis, utgjør heuristikken en kognitiv *skjevhet*, som igjen fører at man trekker kognitive *feilslutninger*.

Heuristikk og skjevhetsprogrammet springer ut fra ideen om at mennesket har en *begrenset rasjonalitet*, som ble introdusert av Herbert Simon på 50-tallet.³⁸ Begrepet var et motsvar på økonomenes forutsetning om det perfekte rasjonelle mennesket, Homo Economicus. I motsetning til Homo Economicus har Homo Sapiens en begrenset prosessorkraft, som ikke er i stand til å måle og vekte observasjoner som om det var en datamaskin.³⁹ Under et evolusjonært press for å sikre artens overlevelse, har hjernen i imidlertid utviklet noen sett tommelfingerregler for å operasjonalisere beslutningsprosessen.

³⁷ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.161

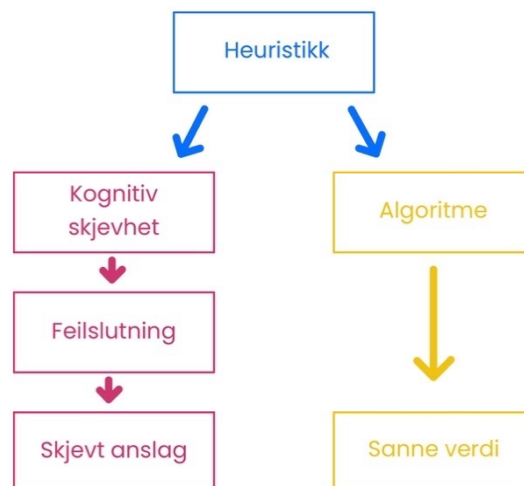
³⁸ (Simon, 1955). Siden Simon sin tid har *begrenset rasjonalitetsbegrepet* blitt omdiskutert. Dersom man tar heuristikkene i et evolusjonært perspektiv, der heuristikkene skal maksimere sjansene for artens overlevelse er heuristikker en rasjonell strategi. Dette perspektivet kalles økologisk rasjonalitet (Gigerenzer, 2012).

³⁹ (Taleb, Skin in the game, 2018) s. 216

Disse tommelfingerreglene kalles *heuristikker*. Den tekniske definisjonen av en heuristikk er «en enkel prosedyre som bidrar til å finne brukbare, men ofte ufullstendige svar på vanskelige spørsmål.»⁴⁰ Jeg mener at denne definisjonen er noe misledende. Jeg har derfor konstruert beslutningskjeden for å vise sammenhengene i begrepene heuristikker, algoritme, skjevhet og feilslutning. Definisjonen sier at heuristikken «ofte» vil føre til ufullstendige svar. Slik jeg forstår heuristikker vil de noen ganger være en god erstatning for måls spørsmålet og kalles da en *algoritme*. De utgjør da en strategi som systematisk vil returnere «riktig»-svar. Men heuristikken kan også være en dårlig erstatning for andre måls spørsmål. Heuristikken vil dermed systematisk returnere «feil»-svar på måls spørsmålet. Heuristikken er da en *kognitiv skjevhet*. Hvorvidt en heuristikk er en kognitiv skjevhet eller algoritme er da avhengig av om man bruker riktig heuristikk på riktig måls spørsmål. Ordlyden «ofte» blir da misvisende. Dersom matchingen er riktig, vil heuristikken være pålitelig til å returnere riktig svar.

Dersom man har matchet feil heuristikk med måls spørsmål, vil altså heuristikken, utgjøre en kognitiv skjevhet. Kognitive skjevheter gjør at man bryter logiske regler. Bruddet på disse reglene kalles kognitive *feilslutninger*.

Man kan oppstille begrepene i et kjede: Heuristikk → Skjevhet → Feilslutning.



Figur 1

Eksempelvis: er heuristikken det man *gjør*. Når bevisbedømmer skal vurdere hvor sannsynlig det er at tiltalte snakker sant, har man empirisk bevist at man anvender det som kalles *representativitetsheuristikker*.⁴¹ Den går ut på at i stedet for å bruke statistisk tenkning for å

⁴⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.109

⁴¹ (Guthrie, Jeffrey, & Wistrich, Blinking on the bench: how judges decide cases, 2007) s.22

anslå sannsynligheten av uttalelsens troverdighet, anvender man stereotypien. Man svarer på spørsmålet «*ser* vedkommende ut som én som *typisk* snakker sant», i stede for «hva er sannsynligheten for at vedkommende snakker sant i dette *konkrete tilfellet*». Dersom det var mulig å avdekke om mennesker lyver ut fra observasjon, kunne denne heuristikken vært en *algoritme*. Det hadde vært en tommelfingerregel som var egnet til å avdekke den sanne verdien. Beklageligvis er vitnepsykologien samstemt om at dette ikke er mulig.⁴² Dermed blir heuristikken en *kognitiv skjevhet*.

Når man anvender representativitetsheuristikker, begår man dermed en rekke *feilslutninger*. En feilslutning er når man unnlater å bruke en åpenbart logisk regel.⁴³ Eksempler på feilslutninger man trekker, når man bruker representativitetsheuristikker, er at man ignorerer grunnfrekvensen og er ufølsom for utvalgte størrelse og kvalitet. Forenklet sagt kan man si at det foreligger en beslutningskjede: Heuristikk → Skjevhet → Feilslutning.

3.1.2 Forskjellen på støy og kognitive skjevheter

Ethvert uheldig utfall er ikke en skjevhet.⁴⁴ Uheldig utfall vil være et resultat av tilfeldigheter som ikke kan forklares i kraft av beviselige skjevheter. Begrepet *skjevhet* favner dermed ikke om enhver feil som blir gjort, men påviste systematiske vurderingsfeil. Begrepet viser i seg selv til mekanismer i vurderingsprosessen, og ikke *utfallet*.

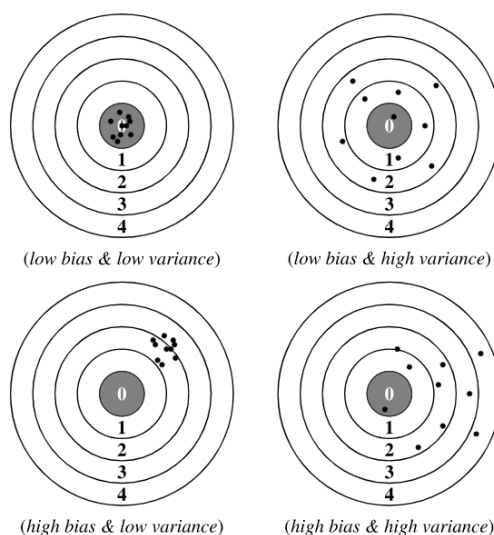
Man må skille mellom psykologiske skjevheter og statistiske skjevheter.⁴⁵ En psykologisk skjevhet er en feilslutning på individnivå. Mens en statistisk skjevhet forekommer når man på gruppenivå deler den samme psykologiske skjevheten, se figur 2. Det er disse statistiske skjevhetene denne oppgaven har som formål å diskutere hvorvidt bevisteoriene evner å korrigere. De psykologiske skjevhetene kan imidlertid slå inn i forskjellig grad eller dra i forskjellig retning for hvert enkelt individ. Differansen mellom de psykologiske skjevhetene i en gruppe skaper *støy*. Støy er altså et resultat av *variasjonen* til psykologiske kognitive skjevheter.

⁴² (Magnussen & Stridbeck, Vurdering av troverdighet: Hva sier forskningen anno 2020?, 2020) s.6

⁴³ (Kahneman, Tenke, fort og langsomt, 2013) s.173

⁴⁴ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.166

⁴⁵ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.161



Figur 2⁴⁶

Støy defineres som uønsket variasjon i vurderinger av det samme problemet.⁴⁷ Det vil si at løsningen på et problem blir forskjellig alt etter hvem som foretar vurderingen. Studier har vist at variasjonen tenderer til å være fem ganger høyere enn forutsatte akseptable verdier.⁴⁸ Selv om beslutningsteori bruker begrepet *støy* om dette fenomenet, mener jeg at det i flere sammenhenger i denne oppgaven er mer hensiktsmessig å bruke det mer beskrivende begrepet *variasjon*. Jeg vil dermed bruke begrepene om hverandre, alt etter som hva som er mest beskrivende.

Variasjon i bevisvurderingen er særlig problematisk fordi det røkter ved det grunnleggende prinsippet om forutberegnelighet.⁴⁹ Dersom resultatet er avhengig av hvem bevisbedømmer er, svekker det *tilliten* til rettssystemet. Dette har særlig blitt undersøkt i det amerikanske rettssystemet, som har skapt debatt rundt hvor stort rom det bør være for dommerskjønn.⁵⁰ Det er heller ikke urimelig å anta at de samme tendensene finnes i det norske rettssystemet. Dette er særlig med tanke på oppslutningen rundt bruken av «reelle hensyn» i rimelighetsvurderinger, og at man har en *fri* bevisvurdering. I det følgende vil jeg kommentere hvilke implikasjoner variasjon har for bevisvurderingen.

⁴⁶ (Jury analyst, 2021)

⁴⁷ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.27

⁴⁸ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.33

⁴⁹ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.53

⁵⁰ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.13 flg.

Vise masser

Når man diskuterer støyete beslutningssystemer, må man bevisst på hvordan variasjon spiller inn på en *gruppe* av beslutningstakere som skal vurdere samme sak, og variasjon i avgjørelsene *i tilsvarende saker*. Variasjon i en gruppe som skal behandle samme sak, kan være en fordel, fordi de psykologiske skjevhetene kan dra i forskjellig retning og dermed nøytraliserer hverandre. Dette er gitt at beslutningstakernes meninger er uavhengig av hverandre. Dette fenomenet kalles *massenes visdom*.⁵¹ Det matematiske fenomenet går ut på at snittet av anslaget til uavhengige beslutningstakere nærmer seg den sanne verdien ettersom utvalgets størrelse øker, også kjent som regresjon mot middelverdien. Effekten ble oppdaget av regresjonsanalysens far Francis Galton, da han spurte alle 787 landsbyboere om å gjette vekten på en okse. Det viste seg at snittet av alle gjetingene kun var ett kg unna den faktiske vekten.

Nyere forskning har også vist at man kan oppnå en tilsvarende effekt, men ikke like sterk, på individuelt plan.⁵² Man kan skape det man kaller en *indre masse*-effekt. Dette kan gjøres på to måter. Den ene måten er den gode gamle «sove på» avgjørelsen. Det vil si at man lar det gå tid mellom første og andre anslag. Man vurderer situasjonen på nytt dagen etterpå, og med mindre man har en sterk grunn til å avvike fra den første avgjørelsen, tar man snittet av de to vurderingene. Den andre metoden går ut på å forsøke å komme med to vidt forskjellige, men plausible anslag. Først gjør man et anslag, men legger så til grunn at det ikke stemmer. Videre skal man stille seg spørsmålet hvorfor anslaget eventuelt er for høyt eller for lavt. Man foretar så et nytt anslag, og tar snittet av de to. Dette kalles *dialektisk bootstrapping*.

Dette er ikke magi, men en metode for å se en problemstilling fra flest mulig vinkler. Når man står overfor en beslutning, bør man altså ha som tommelfingerregel å spørre om en fornyet vurdering fra en andrepert, eller spille djevelens advokat med seg selv. Som jeg vil komme til i neste delkapittel handler det om å aktivt koble på System 2 til å ta en grundigere evaluering av resultatet System 1 har levert. Disse metodene reduserer variasjon i vurderinger fordi de forskjellige psykologiske skjevhetene drar i hver sin retning og dermed dels kansellerer hverandre ut.

⁵¹ (Kahneman, Sibony, & Sunstein, Noise, 2021) s. 86

⁵² (Kahneman, Sibony, & Sunstein, Noise, 2021) s. 88

Gruppedynamikker kan skape variasjon *mellom* grupper som skal vurdere *samme* sak.⁵³ Sosial innflytelse på gruppen har alt å si for om man oppnår en vise massers-effekt, eller om det oppstår gruppetenkning. Sosial innflytelse på gruppedynamikken, vil si at gruppen som helhet tenderer til å holde med første taler. Dette kommer hovedsakelig av to faktorer:

Informasjonskaskader og sosialt press.

Informasjonskaskader oppstår fordi man tenderer til å være enig i første taler. Dette er gitt at man ikke har noen særlig grunn til å mistro vedkommende, eller har noen særlig sterkt belegg for å mene noe annet. Dersom majoriteten i gruppa ikke har noen sterke egne meninger og ikke har noe imot det som blir sagt, vil de være enig med første taler. Jo fler som sier seg enig i første taler, jo sterkere blir mistanken til de som egentlig holder et annet synspunkt om at majoriteten vet noe de ikke vet. De sier seg så enig med første taler.

Her spiller også *sosialt press* inn, siden man ikke ønsker å gjøre seg upopulær i gruppen med å holde en kontrær mening. Resulterer dette i at grupper tenderer til å bli *polarisert*. Det vil si at man i løpet av diskusjonen blir mer selvsikker, enig og entusiastisk til den første meningen, enn det man var da diskusjonen startet. Med andre ord så blir man sosial premiert for å støtte opinionen, og straffet for å være kritisk. På denne måten nøytraliseres ikke de psykologiske skjevhetene ut, men forsterker skjevhetene til gruppemedlemmene, og øker variasjonen *mellom* grupper som skal forta *samme* vurdering. For å oppnå en vise massers-effekt må man sikre at vurderingene til gruppemedlemmene skjer uavhengig av hverandre for å forhindre at informasjonskaskader og sosialt press virker inn og fører til gruppetenkning.

Dette har noen interessante implikasjoner for bevisvurderingssituasjonen. Resultatet av bevisvurderingen er domsteksten som gir uttrykk for faktum som dommerne legger til grunn for den rettslige vurderingen. Siden majoriteten av saker i norsk rett føres for et dommerkollegium, er domsteksten et resultat av diskusjonen fag- og meddommerne har i domskonferansen. Domskonferansen er i liten grad lovregulert.⁵⁴ Hvordan domskonferansen skal gjennomføres begrenser seg til Straffeprosesslovens §31 som sier at når retten har flere enn én dommer «skal den holde et rådslagningsmøte og stemmegivningen for stengte dører før dom blir avsagt, såfremt de ikke straks blir enige i avgjørelsen». Det er dermed opp til rettens leder, herunder fagdommeren, å legge opp til en forsvarlig diskusjon av skyldspørsmål

⁵³ (Kahneman, Sibony, & Sunstein, Noise, 2021) s. 94 flg.

⁵⁴ (Kolflaath, Domskonferanser i straffesaker, 2018) s.360

og utmåling av en eventuell straff. Med andre ord foreligger det ingen rettslig regulering eller veileder for hvordan man skal unngå gruppetenkning i bevisvurderingen under domskonferansen.

Det foreligger en konsensus blant dommere og i juridisk litteratur om at fagdommere påvirker meddommerne under domskonferansen.⁵⁵ Dette ble særlig diskutert i forbindelse med avviklingen av juryordningen i 2018. I diskusjonen ble det etablert at meddommere har en frifinnelsestilbøyelighet sammenliknet med fagdommere.⁵⁶ Det vil si at uten fagdommers påvirkning ville meddommerne stemt for frifinnelse. Diskusjonen gikk videre ut på om fagdommers påvirkning fra frifinnelse til domfellelse var ønskelig.

På den ene siden er det klart at en slik påvirkning er ønskelig dersom meddommerens frifinnelsestilbøyelighet skyldtes rettsvillfarelse. Hvor ønskelig denne påvirkningen er i bevisvurderingen, der meddommer skal vurdere hvorvidt en påstand er sann eller ikke, er mer uklart.

Svaret må bero på en tolkning av domskonferansen formål. Ordlyden i §31 sier at konferansen skal være en «rådslagning», hvilket tilsier at formålet er noe mer enn registrering av dommernes standpunkt i saken.⁵⁷ Det tilsier at konferansen har et *kvalitetssikringsformål*, som må forstås som at møtet skal sikre at bevisvurderingen får et bedre resultat enn dersom dommerne hadde vurdert hver for seg. Bedre resultat må her forstås som at man via diskusjonen kommer nærmere sannheten. Ut fra diskusjonen om massenes visdom krever dette at dommerne ytrer sine uavhengige meninger i rådslagningen, for at samtalen skal lede dem nærmere sannheten. Spørsmålet er om dette faktisk skjer under rådslagningen, eller om den utløser gruppetenkingsmekanismer.

Domskonferansen starter relativt umiddelbart etter at hovedforhandlingen er over.⁵⁸ Det vil si at dommerne har begrenset med mulighet til å gjøre seg opp en *uavhengig* mening utover det man har rukket å reflektere over under forhandlingene. Videre trekker Kolflaath frem at i en tredjedel av sakene han observerte kom rettens leder med sitt synspunkt på skyldspørsmålet innledningsvis.⁵⁹ I så fall vil meddommerne ha begrenset anledning til å danne og ytre en

⁵⁵ (Kolflaath, Domskonferanser i straffesaker, 2018) s.365

⁵⁶ (Kolflaath, Domskonferanser i straffesaker, 2018) s. 366

⁵⁷ (Kolflaath, Domskonferanser i straffesaker, 2018) s.370

⁵⁸ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.209

⁵⁹ (Kolflaath, Domskonferanser i straffesaker, 2018) s.376

uavhengig mening fra fagdommer. Gruppen vil dermed være sårbar for infomasjonskaskader, som vil vende tvilende meddommerne til å holde med første taler.

Videre har lagman Anders Bøhn påpekt virkningen av fagdommers naturlige autoritet.⁶⁰ Bevisst eller ubevisst synes meddommerne det er tryggest å innta samme standpunkt som fagdommer. Her viser han til at det klart vil foreligge et sosialt press til si seg enig med fagdommer, selv om hen ikke er første taler.

Kolflaath trekker i forbindelse med dette frem viktigheten av at rettens leder er klar over sine roller som *beslutningstaker* og *tilrettelegger*.⁶¹ Der bør man legge størst vekt på rollen som tilrettelegger, og vente så lenge som mulig med å gjøre kjent standpunktet sitt som beslutningstaker. Gjennom denne bevisstgjøringen vil man oppnå rådslagningens formål om å kvalitetssikre bevisvurderingen. En del av tilretteleggerrollen er da å fasilitere for en reell diskusjon med vekt på *bevisanalyse*. Det vil si at selv om dommerne har et klart standpunkt innledningsvis, er dette ikke noe argument for å ikke teste holdbarheten av overbevisningen ved en systematisk gjennomgang av bevismaterialet. Dette aktualiserer nødvendigheten av gode bevisteorier ytterligere, som et verktøy i bevisanalysen.

Rollen som tilrettelegger overlapper med rollen som *beslutningsobservatør*, som Kahneman m. fler anbefaler at tar del i slike prosesser for å fjerne variasjon fra beslutningsprosessen.⁶² De påpeker at det er lettere å observere skjevheter i andre enn i seg selv, og foreslår derfor at ett av gruppemedlemmene som skal treffe en beslutning har opplæring i fenomenet støy og kognitive skjevheter. Vedkommende sitter med en sjekklister over mulige kognitive feller man kan gå i, og har som oppgave å spille djevelens advokat for å motvirke fellene.

Et konkret forslag fra min side for å sikre kvaliteten på rådslagningen er å utvikle en *rådslagningsveileder*. Her må det kommenteres at i NOU'en, skrevet i forbindelse med avviklingen av juryordningen i 2018, ble det utviklet en veileder for rådslagningen.⁶³ Denne veilederen er, slik jeg leser den, basert meddommeres erfaringer under rådslagningen, men er ikke systematisk forankret i beslutningsteori.⁶⁴ Den veilederen jeg foreslår er et mer kirurgisk verktøy for å hindre gruppetenkning og kognitive skjevheter.

⁶⁰ (Bøhn, 2009) s.20-24

⁶¹ (Kolflaath, Domskonferanser i straffesaker, 2018) s.372

⁶² (Kahneman, Sibony, & Sunstein, Støy, 2021) s.242

⁶³ (NOU 2011:13 Juryutvalget: Når sant skal skrives) s.166

⁶⁴ (NOU 2011:13 Juryutvalget: Når sant skal skrives) s.161 flg.

Veilederen kunne da vært utformet som tips og obligatoriske spørsmål til hvordan man skal strukturere diskusjonen i møtet for å unngå gruppetenkning. Eksempelvis kunne den sørget for en obligatorisk tenkepause etter hovedforhandlingene. Der kan dommerne gjør opp sin uavhengige mening, som de må notere ned, og presenterer i møtet. Ytterligere kunne veilederen bestått av en liste spørsmål som dommerne kan strukturere diskusjonen rundt for å hindre gruppetenkning, kognitive feller og fremme en systematisk bevisanalyse. Fordi rådslagningen i liten grad er lovregulert, ville en slik veileder ikke kreve noen lovendring. Den ville kun vært et nyttig verktøy dommerne selv kan velge å benytte seg av. På denne måten blir rådslagningen i seg selv en måte å rette opp kognitive skjevheter.

Det finnes ikke gjennomsnittlig rettferdighet

Vise masser-effekten oppstår imidlertid bare når gruppen skal vurdere samme sak, der de psykologiske skjevhetene trekker i forskjellig retning slik at de i *snitt* nuller hverandre ut. Den samme logikken gjelder derimot ikke dersom de forskjellige individene vurderer forskjellige, men tilsvarende saker.⁶⁵ Eksempelvis dersom to personer har fått tilsvarende erstatningskrav rettet mot seg under tilsvarende forhold. Skadevolderne vurderes av hver sin dommer. Erstatningsutmålingen, i et variasjonsfritt vurderingssystem, ville vært 45 000. Den ene får imidlertid utmålt et krav på 70 000, og den andre 20 000. Rettferdighet har her i *snitt* ikke blitt utrettet, selv om snittet av 70 000 og 20 000 er 45 000. I den første saken opplever skadevolder urett i å måtte betale 25 000 for mye, mens i den andre saken opplever skadelidte at hen får 25 000 for lite. I et støyete system nuller ikke psykologiske skjevheter seg ut, de summeres opp.

Stor variasjon leder til færre sanne domspremisser over tid

Overvektsprinsippet sier at en påstand kan legges til grunn som bevisfaktum dersom dommeren er over 50% overbevist om at den er sann. Laveste overbevisningsgrad for at en påstand blir lagt til grunn er da 51%. Begrunnelsen er at dette prinsippet vil gi flest sanne domspremisser over tid.⁶⁶ Dette forutsetter på sin side at enhver dommer ville lagt til grunn den samme overbevisningen gitt at de vurderer den samme bevissituasjonen. Forskningen på *støy* viser at variasjonen ved slike vurderinger ligger på opp mot 50 prosentpoeng.⁶⁷ Selv om

⁶⁵ (Kahneman , Sibony, & Sunstein, Noise, 2021) s.29

⁶⁶ (Strandberg M. , 2012) s.471 flg.

⁶⁷ (Kahneman , Sibony, & Sunstein, Noise, 2021) s.26

denne forskningen gjelder amerikanske dommere, er det ikke urimelig å anta at den også gjelder norske dommere.⁶⁸ Årsaken til den store variasjonen i det amerikanske systemet, som også gjelder det norske, er det store tolkningsrommet den enkelte dommeren har i bevisvurderingen. I Norge er bevisvurderingen som kjent *fri*. Med så stor variasjon i avgjørelsene tenker jeg at det norske overvektsprinsippet vil føre til færre sanne domspremisser over tid.

Logikken er som følgende. Det finnes ingen fasit på hva som er sant eller usant, og man må da ta til takke med det nest beste: massenes visdoms-verdi. Massenes visdoms-verdi er den gjennomsnittlige overbevisningsgraden en representativ gruppe dommere hadde kommet frem til dersom de hadde vurdert samme sak. Men med en variasjon i vurderingen på opp mot 50% fra denne verdien, vil overvektsprinsippet lede til flest riktige domspremisser? I Tabell 1 har jeg skissert et veldig stilisert eksempel for å illustrere poenget. Her er det to dommere. Der dommer 1 er skeptisk, mens dommer 2 er lett å overbevise. Dommer 1 vil dermed slå av 50% på sin overbevisningsgrad sammenliknet med den massenes visdomsverdi, mens dommer 2 vil legge til 50%. Tabellen viser i hvilke tilfeller retten vil lande på massenes visdomsresultat, uavhengig av hvilken dommer som vurderer bevisene.

| Vise massers verdi | Dommer 1 | Dommer 2 | Matcher resultatet med klok gruppeverdien uavhengig av dommer |
|--------------------|-----------|-----------|---|
| 0% Usann | 0% Usann | 50% Usann | ✓ |
| 25% Usann | 0% Usann | 75% Sann | ✗ |
| 50% Usann | 0% Usann | 100% Sann | ✗ |
| 75% Sann | 25% Usann | 100% Sann | ✗ |
| 100% Sann | 50% Usann | 100% Sann | ✗ |

Tabell 1

⁶⁸ (Kahneman , Sibony, & Sunstein, Noise, 2021) s.69

Som det følger av eksempelet vil et beviskrav på 51%, og en variasjon i dommervurderingen på opp mot 50%, føre til at bare 1 av 5 domspremisser blir likt som den vise massen. Dette tilsier at godt over majoriteten av domspremissene ikke vil samsvare med massenes visdoms-verdi over tid. For at overvektsprinsippet skal lede til flest riktige domspremisser, må variasjonen i dommers overbevisningsgrad helt ned til 20%. For å oppnå dette kreves det klarere regulering og veiledning av bevisvurderingsmetoden.⁶⁹ Jo friere vurderingene er, jo større variasjon vil det være i avgjørelsen til enkeltindividene.

3.2 Meningsskapingmaskinen

3.2.1 System 1 og System 2

For å forstå hvordan hjernen foretar sannsynlighetsvurderinger må man ha det klart for seg at hjernen er en «meningsskapingmaskin».⁷⁰ Den leter aktivt etter meningsfylte sammenhenger også der det ikke er noen. Dette delkapittelet skal forklare mekanismene og tendensene som ligger bak denne kausale forståelsen av verden.

Meningsskapingmaskinen kan forstås som et samspill mellom det som kalles System 1 og System 2: Med ordet «system» refereres det ikke til en fysisk del av hjernen, men til eiendommeligheter ved menneskelige tolknings- og beslutningsprosesser.⁷¹ Mekanismen er utviklet for å operasjonalisere det faktum at hjernen har begrenset kapasitet og at årsakssammenhengene i verden er for kompleks til at man alltid greier å prosessere dem. I stedet for at mennesket skulle bli spist av løven mens det grublet over hva en løve *egentlig* er, lærte det seg å trekke slutningen *stor + skarpe tenner + fart i din retning = Løp!* Man utviklet tommelfingerregler, heuristikker, for å overleve i en usikker og uforutsigbar verden.

For å få en intuitiv innføring i forskjellen på System 1 og System 2, fullfør parene «salt og ...», «sukker og ...», «høyt og» og «kaffe eller». Hjernen leverer svarene «pepper», «kanel», «lavt» og «te» automatisk uten anstrengelse. Prøv derimot å regne ut $49 \cdot 57 = \dots$. Denne øvelsen krever anstrengelse, fordi man må anvende regneregler i flere operasjoner. Dette er hovedforskjellen på System 1 og System 2. System 1 virker automatisk og hurtig, med lite anstrengelse, og uten noen opplevelse av viljestyrt kontroll. System 2 på sin side

⁶⁹ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.21

⁷⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.223

⁷¹ (Kahneman, Tenke, fort og langsomt, 2013) s. 36

vekker oppmerksomheten, og retter den mot anstrengende, mentale aktiviteter som krever den.⁷² Dette er systemet man identifiserer seg med, jeget. System 1 er altså intuitivt, mens System 2 er analytisk.

Hovedfunksjonen til System 1 er å vedlikeholde og oppdatere en modell av verden slik man normalt oppfatter den, det jeg kaller *normalmodellen*.⁷³ Modellen baserer seg på assosiasjoner av vilkår, hendelser, handlinger og utfall som opptrer sammen med en viss regelmessighet. Vi erfarer at fugler har fjær, nebb, vinger, klør, legger egg og kan fly. Eller at morske øyebryn og høy stemme indikerer sinne. System 1 kan da automatisk fortelle oss at den flyvende fjærdotten er en fugl, eller sjefen er sint uten at man må anstrenge seg. Modellen er dermed en kartlegging av erfarte årsakssammenhenger. Den avgjør dermed hvordan man tolker nåtiden, og hvilke forventninger man har til fremtiden.

Overraskelse indikerer yttergrensen av hva man oppfatter som normalt. Det er det motsatte av hva man forventer. Når forventningen blir brutt blir man overrasket. Da må System 2 på banen for å oppdatere modellen.⁷⁴ Eksempelvis, ser man en pingvin for første gang vil System 1 slite med å automatisk kategorisere pingvinen siden den mangler et viktig kriterium: den kan ikke fly. Ved nærmere analyse med System 2, konkluderer man derimot med at modellen må utvides til at dette også må være en fugl. De øvrige kriteriene er tross alt oppfylt. Overraskelsen aktiverer og styrer dermed oppmerksomheten. Påkoblingen av System 2, og oppdatering av normalmodellen, er det man opplever som læring. Man erfarer og lærer en ny årsakssammenheng. Hvorvidt den erfarte årsakssammenhengen faktisk er en empirisk kausalitet, eller om det var et resultat av tilfeldigheter, er ikke System 2 nødvendigvis kritisk til.

Spørsmålet er så: hvor godt fungerer overraskelse som varselampe for System 2. Analyse fra System 2 krever anstrengelse, noe som innebærer at det er kostbart for hjernen å koble på System 2. Hva om System 1 strekker seg litt *for* langt i å konstruere et svar, for å unngå anstrengelsen med å koble på System 2? Det klassiske eksempelet, introdusert av Shane Frederick og Daniel Kahneman, er som følger: En tenniser racket og en ball koster 1,10 dollar. Racketen koster én dollar mer enn ballen. Hvor mye koster ballen?⁷⁵ I likhet med 80 prosent

⁷² (Kahneman, Tenke, fort og langsomt, 2013) s. 26 flg.

⁷³ (Kahneman, Tenke, fort og langsomt, 2013) s.81

⁷⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.82

⁷⁵ (Kahneman, Tenke, fort og langsomt, 2013) s.51

av studentene som ble spurt på universiteter i USA, vil de fleste konkludere med at ballen koster 10 cent. Selv om man til slutt regner seg frem til det riktige svaret, på 5 cent, så meldte svaret om 10 cent seg først.

Her ser man forholdet mellom System 1 og System 2 i et mikrokosmos. I stedet for at System 1 innrømmer at svaret er for komplekst til å svare på, kommer det med et intuitivt svar som det ikke har belegg for. System 2 på sin side er en relativt lat kontrollør, og vil i mange tilfeller godta System 1 sine intuitive svar uten å gå det nærmere i sømmene.⁷⁶ Dette gjelder særlig i tilfeller der man ikke vet bedre, herunder ikke kjenner det logiske prinsippet man skal følge, eller i tilfeller hvor svaret System 1 kommer med er følelsesstyrt. Da vil System 2 ha en tendens til å bekrefte svaret i stedet for å stille kontrollspørsmål.⁷⁷ System 1 vil alltid søke mot å finne årsakssammenhenger, selv i rent tilfeldige mønstre. Dette får oss til å se verden som ryddigere, enklere, mer forutsigbar og *koherent* enn den egentlig er. Den menneskelige kognisjonen er altså evolusjonært utviklet for å sikre artens overlevelse, ikke for å foreta statistiske utregninger.

3.2.2 Preferanser

Meningsskapingmaskinens mest grunnleggende bestanddel er mekanismene som danner preferanser. Preferanser dannes ved assosiativ aktivering, noe som skjer i den assosiative hukommelsen. Assosiasjon er nøkkelen for all beslutningstaking av System 1. Det er denne mekanismen som forklarer årsakene til hvordan System 1 trekker forhastede slutninger, bekreftelsestendensen, glorieeffekten og WYSIATI, og hvordan System 1 foretar bedømminger via erstatning.

For at System 1 skal kunne sy sammen et koherent verdensbilde, må det ha et startpunkt. Dette startpunktet er våre *preferanser*. Med preferanser menes at man foretrekker én ting fremfor en annen i valget mellom de to. Preferanser angår valg mellom hverdagslige ting som juice fremfor melk, til dypere spørsmål om hvilke grunnleggende verdier man styrer livet sitt etter. Preferansene blir påvirket og oppstår av to effekter som kommer av System 1-tenkning, *assosiativ koherens* og *kognitiv letthet*.

⁷⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.52

⁷⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.116

Assosiativ koherens

Assosiasjoner er et allment kjent begrep. Én tanke leder videre til en ny tanke, som leder til enda en tanke, uten at vi helt er bevisst hvorfor de henger sammen. Denne prosessen kalles assosiativ aktivering. Det sentrale med denne prosessen er at tankene er knyttet sammen, slik at det første leddet støtter og forsterker det neste. En tanke utløser minner, som utløser følelser, som utløser en kroppslig reaksjon, som utløser en handling.⁷⁸ Reaksjonen og handlingen vil igjen påvirke følelsene, som igjen vil påvirke nye tanker. Slik fortsetter det i en spiral av *tanke – følelser – handling*. Dette mønsteret kalles *assosiativ koherens*.

Dette mønsteret er sentralt for dannelsen av våre preferanser fordi det er hovedmekanismen bak hvordan System 1 lager årsakssammenhenger. Kahneman gjør et eksempel av ordene *banan* og *oppkast*. Ved å sette disse to tilfeldige ordene sammen starter System 1 automatisk å lage en årsakskjede mellom banan og oppkast. Resultatet av prosessen er at man får en midlertidig avsmak for bananer, siden man tenker at det er bananen som er *årsaken* til oppkastet. Selv om det ikke foreligger ytterligere informasjon som støtter denne påstanden. Man har dermed dannet en årsakssammenheng som ikke er der, og som gjør at vi avstår fra å spise bananer til avsmaken har gitt seg. Ta et eksempel fra retten: sett sammen ordene «Peder» og «tiltalt», og man starter automatisk å lage årsakssammenhengen Peder *er* tiltalt, hvilket trigger spørsmålet, «Hvorfor er Peder tiltalt?». Og så begynner man å lete etter *forklaringer* på hvorfor Peder er tiltalt, selv om det bare var tilfeldig at de to ordene sto ved siden av hverandre i første omgang.

Assosiativ koherens avdekker altså to ting om hvordan hjernen lagrer informasjon. For det første at den lagrer informasjonen i klynger, én tanke utløser en rekke andre tanker. For det andre at den lagrer informasjonen i årsakskjeder. For at vi skal huske ordparet *banan* og *oppkast* lager System 1 et narrativ om at banan forårsaker oppkast, eller at Peder er tiltalt. Hjernen greier ikke lagre selvstendige datapunkt, den må samle dem sammen i klynger som henger sammen i årsakskejder. Dette skaper en illusjon om at vi er i stand til å forklare alle omstendigheter rundt oss.

Assosiativ koherens er prosessen som benyttes i det som kalles *priming*. Priming handler om at man får ytre stimuli som setter i gang mønsteret *tanke – følelse – handling*. I eksperimenter har denne effekten vært illustrert ved at forsøkspersoner som blir bedt om å stokke om ord

⁷⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.59

som assosieres med eldre, ender opp med å gå saktere når de skal til neste forsøksrom.⁷⁹ Eller at forsøkspersoner som hadde en pengeseddel som skjermesparer ble mer egoistiske. Dette sier noe om hvor mektig priming-effekten er, den styrer våre lyster. Siden det er System 2 man identifiserer seg med, mens det er System 1 som blir primet, har man ingen oversikt over hvilke avgjørelser man gjør i løpet av en dag som er primet.⁸⁰ Priming kan dermed styre våre sympatier og antipatier uten at vi er klar over det.⁸¹

For å ta priming inn i en rettslig kontekst tenker jeg at effekten sier noe om implikasjonene fenomenet «dømt i pressen» har på den faktiske rettsaken. Utsagnet tilsier at man ikke får en rettfærdig rettsak, fordi mediene har gjort opp sin mening om hvorvidt tiltalte er skyldig eller ikke, allerede før saken har startet. I profilerte saker, som for eksempel Baneheia-saken, skal det godt gjøres at bevisbedømmer greier å isolere seg fra alt av medieomtale. Selv om bevisbedømmer aktivt går inn for at man ikke skal internalisere mediens meninger, er allerede assosiasjonsklyngen dannet, og påvirker ubevisst bevisbedømmers sympatier eller antipatier. Tiltalte starter dermed sitt forsvar i motbakke, hvilket svekker uskyldspresumsjonen om at tiltalte er skyldig til det motsatte er *bevist*.

Kognitiv letthet

En primet idé gir en følelse av det man kaller *kognitiv letthet*. Følelsen forklarer mekanismen som hindrer System 2 i å bli overrasket, og gir dermed ikke noe signal til System 2 om å dobbeltsjekke svaret. Dersom en tanke eller idé man blir introdusert for oppleves som kognitivt lett, omgår den med andre ord kontrollen til System 2.

Kognitiv letthet er et måleinstrument System 2 har på System 1 for å vurdere om det trengs nærmere undersøkelser av resultatene System 1 leverer.⁸² Barometeret går fra følelsen av *lett* til *anspent*. «Lett» er et tegn på at alt går bra. Det er ingen overhengende trusler, nyheter eller behov for å vekke oppmerksomheten til System 2. «Anspenthet» derimot tyder på at det foreligger et problem som krever oppmerksomhet og nærmere analyse fra System 2. Når noe føles kognitivt lett sitter vi igjen med at det er fortrolig, sant, godt og uanstrengt.⁸³ Man vil dermed ha en preferanse for det som oppleves som kognitivt lett.

⁷⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.61

⁸⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.67

⁸¹ (Kahneman, Tenke, fort og langsomt, 2013) s.63

⁸² (Kahneman, Tenke, fort og langsomt, 2013) s.68

⁸³ (Kahneman, Tenke, fort og langsomt, 2013) s.69

Hva baserer så følelsen av «lett» seg på? I grove trekk er det fire årsaker til følelsen av kognitiv letthet: primet idé, gjentakelser, klar visuell fremstilling og godt humør.⁸⁴

Gjentakelse vil si at bare i kraft av at vi har sett eller hørt noe før, får vi følelsen av fortrolighet.⁸⁵ Eksempelvis finner jeg på navnet Gerd Huseby. Bare i kraft av at navnet er nevnt nå, vil man synes dette navnet er mer fortrolig, enn et navn man ser for første gang.⁸⁶ Gjentakelse skaper en falsk følelse av erfaring. Man sitter igjen med følelsen av at man kjenner Gerd, og dermed føles det navnet mer tillitvekkende. Navnet vil vekke en letthet i den som er primet for det. Dette kan virke trivielt, men den sikreste måten å bli trodd i en løgn er å gjenta den ofte. Fordi det er vanskelig å skille følelsen av fortrolighet fra følelsen av at noe er sant.⁸⁷ Hvis studien på dette er overførbart til retten, kan partene manipulere dømmekraften til dommeren via repetisjon av påstander i løpet av en rettsak. Jo oftere en påstand blir repetert, jo mer fortrolig og sann virker den.

En annen måte å minske kognitivt anspenhet er å formidle et budskap med *klar visuell fremstilling*.⁸⁸ Det vil si at alt som gjør et budskap lettere tilgjengelig, skaper en illusjon om at det er sant. Eksperimenter har påvist at alt fra å bruke uthevet skrift, sterke kontrastfarger, skrift trykket på kvalitetspapir, tilgjengelig språk eller at budskapet blir skrevet på rim, har en effekt på hvor troverdig man synes budskapet er. Dette betyr også at man i større grad unngår kognitive feller når budskapet er mer utilgjengelig, siden det skaper kognitiv anspenhet, hvilket igjen aktiverer System 2.⁸⁹

Dersom dette er overførbart til en rettslig sammenheng, tenker jeg at dynamikken mellom kognitiv letthet og anspenhet sier noe om hvorvidt bevisbedømmer vil være mottakelig for god retorikk. Jo lettere man greier å fremstille og formidle argumentene sine, jo mindre kritisk motstand vil bevisbedømmer yte mot kvaliteten på argumentet. Dette gjør at man kan ende med å unngå å føre mer komplekse bevis, til tross for at de er av høyere kvalitet. Dette berører noe av kjernen av denne oppgavens argumentasjon: ettersom man ikke har anlegg for å

⁸⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.69

⁸⁵ (Kahneman, Tenke, fort og langsomt, 2013) s.69

⁸⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.69

⁸⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.71

⁸⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.72

⁸⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.75

intuitivt tenke statistisk, unngår man å prosedere statistiske bevis. Dette kommer av at man risikerer at bevisbedømmer vil bli lurt av kognitive illusjoner.

Til slutt kan *godt humør* påvirke hvor kognitivt lett vi føler oss. Først og fremst blir man fylt av gode følelser av at noe er kognitivt lett. Man smiler svakt når det er lett å gjenkjenne en figur på et bilde, sammenliknet med om figuren blir fremstilt uklart.⁹⁰ Dette fungerer også motsatt vei. Dersom man er i godt humør av andre årsaker, er System 2 mindre på vakt enn dersom man er i dårlig humør.⁹¹ En studie av israelske dommere viste at fengslede hadde større sjanser for å innvilge prøveløslatelse tidlig på dagen og rett etter lunsjpausen, sammenliknet med sent på dagen og rett før lunsj.⁹² Hvilken kan tilsi at når dommerne var i mette og i et godt humør var de mer tilbøyelige til å lande på et positivt utfall.

3.2.3 Forhastede slutninger

Når preferansene først har satt seg, bruker meningsskapingsmaskinen dem som et utgangspunkt for å trekke forhastede slutninger. Prosessen bak å trekke forhastede slutninger består hovedsakelig av *bekreftelsestendensen*, *glorieeffekten* og *WYSIATI*.

Bekreftelsestendensen

For å forstå en påstand må vi først tro på at den er sann.⁹³ Eller mer presist man må først vite hva ideen ville betydd hvis den var sann. Det er ikke før du har forstått påstanden at man er i stand til å ikke-tro den. Dersom aktor sier at «Viggo Kristiansen er skyldig i drapene i Baneheia», må bevisbedømmer først tro på påstanden for å forstå rekkevidden av hva denne påstanden innebærer. Den assosiative hukommelsen, System 1, vil da starte et søk etter alle bevisene som støtter denne påstanden og sy dem sammen til en koherent fortelling. Oppgaven om å tvile og ikke-tro tilhører System 2. Dersom System 2 ikke blir varslet, eksempelvis på grunn av en følelse av kognitiv letthet, eller det er distraheret av andre årsaker vil man konkludere med at påstanden er sann.

Bekreftelsestendensen stopper imidlertid ikke her. Selv om System 2 blir varslet fungerer den assosiative hukommelsen slik at System 2 får en positiv teststrategi. Det vil si at den vil sette i

⁹⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.75

⁹¹ (Kahneman, Tenke, fort og langsomt, 2013) s.79

⁹² (Danziger, Levav, & Avnaim-Pesso, 2011)

⁹³ (Kahneman, Tenke, fort og langsomt, 2013) s.91

gang et viljestyrt søk i hukommelsen etter informasjon som *bekrefter* påstanden. Eksempelvis i spørsmålet om Viggo Kristiansen er skyldig, vil System 2 sette i gang et søk i hukommelsen etter alle minnene fra rettsaken som er i overenstemmelse med at han kan være en drapsmann: han er aggressiv, er tidligere dømt for overgrep, han har ADHD, han var god venn med Jan Helge Andersen som har tilstått drapene, han fremstår ikke som en sympatisk fyr osv.

Resultatet av bekreftelsestendensen er altså at samspillet mellom System 1 og System 2 eliminerer motstridene bevis, for å skape en koherent forklaring.⁹⁴ Dette er stikk i strid med vitenskapelige prinsipper for hypotesetesting, som går ut på at man utelukker alternative forklaringer på fenomenet, fremfor å bekrefte at fenomenet stemmer. Kognisjonen tester imidlertid hypoteser med målrettede søk som bekrefter hypotesen. Det at bekreftelsestendensen er et resultat av et samspill mellom System 1 og System 2 gjør den særlig vanskelig å korrigere for.

Glorieeffekten

En effekt tett beslektet med bekreftelsestendensen er overdreven emosjonell koherens også kalt *glorieeffekten*. Dette er tendensen til å la førsteinntrykk kaste en «glorie» over alle sider ved en person eller situasjon, også de man ikke har observert.⁹⁵ Førsteintrykket tjener på samme måte som en påstand for bekreftelsestendensen, hvis man har et positivt førsteinntrykk lar man tvilen komme vedkommende til gode.⁹⁶ Eksempelvis kan en inkonsekvent forklaring unnskyldes som menneskelig feil i hukommelsen, eller tolkes som løgn. Hvorvidt bevisbedømmer tolker forklaringen i den ene eller andre retningen kan avhenge av om vitnet har gitt et førsteinntrykk av å være hyggelig og sympatisk, eller avmålt og fiendtlig. Glorieeffekten gjør at førsteinntrykk dannes raskt, og at man er treg til å nyansere dem.⁹⁷

WYSIATI

Bekreftelsestendensen og glorieeffekten sier noe om hvordan System 1 aktivt undertrykker *tvil* når man skal foreta en vurdering. En viktig medvirker i bekreftelsestendensen og glorieeffekten er WYSIATI. WYSIATI, står for “What You See I All There Is”.⁹⁸ Denne

⁹⁴ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.172

⁹⁵ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.172

⁹⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.93

⁹⁷ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.172

⁹⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.96

effekten beskriver System 1s asymmetriske behandling av informasjonen den assosiative hukommelsen har aktivert, kontra det den ikke har aktivert. System 1 konstruerer den beste historien ut fra den informasjonen den har foran seg. Informasjon som ikke hentes frem fra hukommelsen kunne like gjerne ikke eksistert. Målet for System 1 er å skape en sammenhengene fortelling. Mengden og kvaliteten på bevisene som ligger til grunn for fortellingen er for det meste irrelevant. Med andre ord har informasjonens *robusthet* ingenting å si for hvorvidt System 1 oppfatter fortellingen som troverdig. System 1 vil faktisk oppfatte det som lettere å sette sammen en sammenhengende fortelling dersom det er få bevisfakta. Da er det færre puslebrikker man må få til å passe sammen. Med begrenset informasjon vil ikke System 1 innrømme for seg selv at bevisgrunnet er for mangelfullt til å trekke en konklusjon, og heller trekke *forhastede slutninger*. Denne effekten er svært sentral for de skjevhetene man sliter med å rette opp i både forklaringsperspektivet og Bayes teorem.

3.2.4 Grunnleggende vurderinger

Erstatning

Til nå har jeg diskutert hvordan assosiativ koherens får oss til å gjøre forhastede slutninger, men den er også utgangspunktet for hvordan man foretar grunnleggende vurderinger når man blir stilt komplekse spørsmål.⁹⁹ Eksempelvis «Snakker vitnet sant?» eller «Har moren misbrukt barnet sitt?». Innenfor atferdspsykologien kaller man disse vanskelige spørsmålene for *Målspørsmål*.¹⁰⁰ Disse spørsmålene krever nitidige resonnementer som ikke nødvendigvis leder noen vei. System 1 har dermed en tendens til å komme med det som kalles *heuristiske spørsmål*. Dette er spørsmål man har bedre forutsetninger for å svare på og som relaterer seg til målspørsmålet. De heuristiske spørsmålene for disse eksemplene ville da vært «Ser vitnet ut som en løgner?» og «Synes du synd på moren?». Det at man svarer på det heuristiske spørsmålet fremfor målspørsmålet kalles *erstatning*.¹⁰¹

Det er to underliggende mekanismer bak erstatning: Det mentale *haglegeværet* og *intensitetsmatching*.¹⁰² Vi har en relativt upresis kontroll over viljestyrte beregninger. Som regel beregner System 1 mer enn det den får beskjed om, hvilket er et resultat av assosiativ

⁹⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.101

¹⁰⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.109

¹⁰¹ (Kahneman, Tenke, fort og langsomt, 2013) s. 109

¹⁰² (Kahneman, Tenke, fort og langsomt, 2013) s.105 flg.

aktivering. Det er dette som kalles det mentale haglegeværet. System 2 gir System 1 i oppdrag å svare på om «moren er skyldig?», men System 1 svarer også på om «Virker hun som en god mor?», «Er hun veltalende?», «Er hun pen?», «Kunne du sett for deg henne i fengsel?» Og «Synes du synd på henne? System 1 returnerer svarene på de spørsmålene det intuitivt kan svare på. Hvis System 1 ikke kan svare på måls spørsmålet, returnerer det en av de heuristiske alternativene som er basert på mekanismene bak *forhastede slutninger*.

Men før svaret godtas må det tilpasses måls spørsmålet. Dette gjøres ved hjelp av *intensitetsmatching*. Intensitetsmatching er evnen til å tillate komparative former.¹⁰³ System 1 har en underliggende intensitetsskala som tillater matching på tvers av dimensjoner. Eksempelvis har følelsen sorg en mørkere farge, enn glede. Eller som i vårt tilfelle, skyld og sympati er ytterpunkter på skalaene mellom skyld/ uskyld og sympati/ antipati. Man kan dermed matche følelsen av graden av sympati, med graden av skyld. Det heuristiske spørsmålet «Synes du synd på moren?» kan dermed *erstattes* med måls spørsmålet «Er moren skyldig?». Noen ganger er det heuristiske alternativet en god erstatning for måls spørsmålet, andre ganger ikke.

3.2.5 Kausal mot statistisk tenkning

Det at meningsskapingsmaskinen aktivt leter etter årsakskjeder, skaper en hel del problemer når en bevisbedømmer skal vurdere «sannsynlighet». System 1 og System 2 tenker intuitivt kausalt.¹⁰⁴ Det vil si at man utleder generelle sammenhenger fra enkelte observasjoner.¹⁰⁵ I prosessen for å bevise en hypotese, søker den assosiative hukommelsen aktivt opp minner som bekrefter påstanden, og tar kun hensyn til den informasjonen som blir hentet frem. Hva som hentes frem er informasjonen som er lettest tilgjengelig i hjernen.

Motsatsen til kausal tenkning er statistisk tenkning. Statistisk tenkning skal jeg diskutere mer inngående i kapittel 5, men hva det helt kort innebærer er at man konkluderer med årsaksforhold ved å observere grupper over tid, og registrerer hvor ofte fenomenet man studerer oppstår. Et eksempel på kausal tenkning er at: jeg leste om en mørkhudet mann som hadde slått ned en person på gaten. Jeg slutter fra dette at mørkhudede menn er mere voldelige enn hvite menn. Men statistisk sett er kanskje mørkhudede menn like lite voldelige som hvite

¹⁰³ (Kahneman, Tenke, fort og langsomt, 2013) s.105

¹⁰⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.190

¹⁰⁵ (Taleb, Fooled by Randomness, 2004) s.130

menn dersom man kontrollerer for sosioøkonomisk bakgrunn. Vold er heller relatert til hvilket samfunnslag man kommer fra, og innvandrere er overrepresentert i lavere sosioøkonomiske lag. Eller så kan man ikke lære noe om vold fra denne hendelsen, et var bare *tilfeldig*. Det som er sikkert, er at det var en person slo ned en annen. Men forklaringen på *hvorfor* volden oppsto, er enten utenfor vår forståelsesevne. Eller så kan ikke denne hendelsen si noe generelt om årsakene til hvorfor vold oppstår.

Kjernen i hvorfor man intuitivt ikke greier å tenke statistisk er at meningsskapingmaskinen ikke greier å forholde seg til *tilfeldigheter*. Man sliter med å ta høyde for at fenomenet man observerer ikke har noen meningsfylt årsakssammenheng til bevistema. For å si det med Kahneman sine egne ord: «Kausalt sett finnes ikke støy noe steder, statistisk sett finnes det overalt.»¹⁰⁶ «Støy» i denne sammenhengen må forstås som tilfeldigheter. Grunnen til dette er måten hjernen lagrer informasjon på er via i årsaksklynger. Tilfeldige observasjoner som ikke har noen forklaringskraft, blir dermed ikke lagret. Når System 2 sender ut et søk etter forklaringer på en hypotese, søker den i en database full av årsaksforklaringer, men uten tilfeldigheter. Man sitter dermed med et forvridd bilde om at de fleste fenomener i verden kan forklares, når realiteten er at det gjelder de aller færreste.

Det at meningsskapingmaskinen ikke tar høyde for tilfeldigheter gjør at System 1 er hjelpeløs ovenfor rent statistiske fakta, som kun endrer graden av sannsynligheten. Selv om tilfeldigheter ikke kan forklares, kan samlinger av tilfeldige hendelser opptre i mønster.¹⁰⁷ Ta eksempelet med et myntkast. Det er *tilfeldig* om det blir kron eller mynt. Men over tid viser statistikken at det 50/50 sjanse for om det blir mynt eller kron. Denne statistikken *forklarer* imidlertid ikke hvorfor kron oppstår halvparten av gangene. Dette gjør at meningsskapingmaskinen ikke greier skille mellom *grader* av overbevisning.¹⁰⁸

System 1 har med andre ord ikke anlegg for tvil. Maskinen henter frem den informasjonen som bekrefter påstanden, og konkluderer med to streker under svaret. Dersom System 1 blir presentert for rene statistiske fakta, som kun endrer sannsynlighetsgraden, men ikke utgjør *årsaken* til utfallet. Da vet ikke System 1 hva den skal gjøre med den informasjonen Enten blir informasjonen feilaktig lagret som en årsaksforklaring eller så blir den ignorert og glemt.

¹⁰⁶ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.220

¹⁰⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.122

¹⁰⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.126

Meningsskapingmaskinen har dermed en intuitivt kausal tilnærming til sannsynlighetsvurdering. Det vil si at man intuitivt har et binært forhold til en påstand, enten er den sann *eller* usann. Den vil overvekte eller ignorere motstridene bevis for å falle ned på den ene eller andre siden. Men den er ikke i stand til å gradere usikkerheten på en skala fra 0 til 100 som sannsynlighetsvurderingen i norsk rett krever.

Dette har mye å si for hvor effektivt en bevisvurderingsteori kan rette opp kognitive skjevheter. Sannsynlighetsparadigmet og fortellingsperspektivet har to forskjellige tolkninger av «sannsynlighet». I sannsynlighetsparadigme referer «sannsynlighet» til den statistiske forståelsen av sannsynlighet. Nærmere bestemt vurderer man hendelser som ett av flere mulige utfall. Eksempelvis slår man en mynt. Det er da to mulige utfall, kron eller mynt. *Sjansen* eller *sannsynligheten* for å få kron er én av to, eller 50%. Eller for å oversette denne tankegangen til bevisvurderingen kan man eksempelvis spørre seg: om bitemerket funnet på kvinnens bryst stammer fra Fredrik Fasting Torgersen? Bitemerker er ikke unike.

Bevisbedømmer anslår at det er tre personer som kunne gitt det bitemerket i Norge. Isolert sett er da *sannsynligheten* for at bitemerket stammet fra Torgersen 1/3. Med andre ord er det 1/3 sjanse for at det var Torgersen som hadde bitt Rigmor Johansen. Mens det var 2/3 sjanse for at bitemerket på Johansen sitt bryst stammer fra en annen person enn Torgersen.

Fortellingsparadigmet har på sin side en mer dagligdags forståelse av «sannsynlighet». Det som er sannsynlig forstås som *plausibelt*, og «sannsynligvis» betyr det samme som *trolig*.¹⁰⁹ Der hvor sannsynlig en påstand er, vurderes da ut fra hvor koherent forklaringen fremstår. Med andre ord tar sannsynlighetsparadigmet utgangspunkt i at bevisbedømmer evner å tenke statistisk, mens fortellingsparadigmet tar hensyn til at hen tenker kausalt. System 2 kan trenes til å tenke statistisk, men dette er en ressurskrevende opplæring.¹¹⁰ Flere kognitive illusjoner er så hardnakkede at man ikke slutter å se illusjonen selv om man vet bedre. Man jobber dermed i motbakke med en bevisvurderingsteori som tar utgangspunkt i statistisk tenkning.

¹⁰⁹ (Kolflaath, Bevisbedømmelse - Sannsynligheter eller fortellinger, 2004) s.283

¹¹⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.88

3.2.6 Påliteligheten til fagdommeres intuisjon i bevisvurderingen

Trening av intuisjonen via erfaring

Når bevisbedømmer skal foreta en sannsynlighetsvurdering for om beviskravet er oppfylt, er det et stort alvor som ligger over situasjonen. Man skulle tro at situasjonen i seg selv ville oppfordre til analytisk System 2 tenkning. Over tid ville bevisbedømmer opparbeide seg erfaring med å vurdere bevis, som igjen vil gjøre intuisjonen mer pålitelig. I denne oppgaven er bevisbedømmer definert som fagdommerne og lekdommere i en rettssak. Det er dermed relevant å diskutere hvorvidt fagdommere har større grunn til å stole på sin intuisjon, enn lekdommere når det kommer til bevisvurderingen.

Til nå har jeg gjort rede for meningsskapingmaskinen som ligger bak intuisjonen til System 1. Der kom det klart frem at assosiativ koherens er det grunnleggende mønsteret System 1 bruker for å hente frem minner. Intuisjon i sin kjerne er dermed «intet mer og intet mindre enn gjenkjennelse».¹¹¹ Det System 1 gjør når den setter i gang meningsskapingmaskineriet er ikke noe mer magisk enn en dagligdags hukommelsesprosess. Svarene til System 1 blir altså ikke bedre enn våre erfaringer. Det er disse erfaringene som blir avgjørende for om vi kan stole på intuisjonen eller ikke.

Det er to forutsetninger som må ligge til grunn for at System 1 skal kunne forbedre beslutningsgrunnlaget sitt:

- 1) «et miljø som er tilstrekkelig regelstyrt og følgelig forutsigbart»
- 2) «mulighet til å lære disse reglene gjennom øvelse over tid».¹¹²

Dersom et miljø er relativt forutsigbart, og man blir eksponert for den relativt samme erfaringen over tid, vil man kunne ha en opplevelse av årsak-virkning. Man lærer imidlertid ikke av erfaring i seg selv. Det kreves også tilbakemelding på om erfaringen er god eller dårlig. Tilbakemeldingen må komme nært i tid for at lærdommen skal sette seg. Kahneman sammenlikner prosessen med å lære seg å bremse i svinger når man kjører.¹¹³ Hver sving er ikke lik, men ulikheten er oversiktlig. Gradvis lærer man når man skal slippe gassen og

¹¹¹ (Kahneman, Tenke, fort og langsomt, 2013) s.256 sitat av Herbert Simons

¹¹² (Kahneman, Tenke, fort og langsomt, 2013) s.260

¹¹³ (Kahneman, Tenke, fort og langsomt, 2013) s.261

bremse, ut fra hvor krapp svingen er. Bremses man for hardt får man umiddelbart tilbakemelding i form av et ubehagelig rykk i setebeltet.

Ekspertintuisjon har blitt påvist i yrker som sjakk mestere, fysikere, forsikringsanalytikere, matematikere, brannmenn, anestesileger og sykepleiere.¹¹⁴ Selv om oppgavene de står ovenfor er komplekse, er de prinsipielt forutsigbare. Og man får relativt umiddelbar tilbakemelding på om vurderingen var god eller dårlig. Eksempelvis kan en sykepleier på fødeavdelingen stole på intuisjonen sin om spedbarnet har en infeksjon, fordi prøveresultatene vil kunne bekrefte eller avkrefte intuisjonen relativt raskt etterpå.

Det har imidlertid blitt påvist dårligere ekspertintuisjon blant aksjemeglere, kliniske psykologer, psykiatere, opptakskontor til collage, etterretningsanalytikere og *fagdommere*.¹¹⁵ Det som skilte yrkesgruppene som gjorde det bra fra de som gjorde det dårligere, var faktorer som: forutsigbarhet for utfallet til beslutningene, mengden erfaring og adgangen til å få tilbakemelding.

For å trekke paralleller til begrepene i kunnskapsteori så er det første kriteriet for opparbeidelse av ekspertintuisjon at det foreligger et regelstyrt og derav forutsigbart miljø. Problemene oppstår dermed på det epistemiske og psykologiske nivået. Man så over i det andre kriteriet om *mulighet* for å lære seg reglene til miljøet. Ta eksempelet med sykepleieren. Sykepleieren skal trene intuisjonen om et spedbarn har en infeksjon eller ikke ved å tolke symptomer. En infeksjon kan alternativt konstateres ved hjelp av en blodprøve. Både symptomene og blodprøven er bevis på det epistemiske nivået for hva som er fasiten på det aletiske nivået. Over tid er blodprøven et representativt bevis for fasiten på det aletiske nivået. Sykepleieren kan dermed trene intuisjonen sin ved å dobbeltsjekke om hen tolker symptomene riktig, fordi det foreligger bevis som er representative for fasiten. Det er dermed mulig å lage en tilbakemeldingssløyfe.

Aksjemeglere, psykologer og fagdommere har midlertid ikke tilgang på tilsvarende fasiter. De opererer i for komplekse miljøer til at man har teknologi som kan samle tilsvarende representative bevis, som blodprøven, på det epistemiske nivået. Det er ikke noe bevis å skape tilbakemeldingssløyfe på. Og gitt at man kunne laget en slik sløyfe, er det også trolig at

¹¹⁴ (Kahneman & Klein, Conditions for Intuitive Expertise: A Failure to Disagree, 2009) s.522

¹¹⁵ (Kahneman & Klein, Conditions for Intuitive Expertise: A Failure to Disagree, 2009) s.522

miljøene består av så mange variabler at hjernen ikke har tilstrekkelig kapasitet til å prosessere informasjonen korrekt på det psykologiske nivået.

Kan man bli ekspert på bevisvurdering?

Det blir imidlertid for generelt å si at disse yrkesgruppene ikke kan opparbeide seg noen form for ekspertise på fagfeltet sitt over tid. Man vil kunne opparbeide ekspertintuisjon på områdene der man har tilgang på en fasit, der det er mulig å lage en tilbakemeldingssløyfe den hjernen er kapabel til å tolke tilbakemeldingen.

Denne oppgaven har allerede poengtert at det er et skille mellom den materielle retten og bevisteori. Den materielle retten er konstruert av mennesker. Informasjonen om den materielle retten på det epistemiske nivået samsvarer da med det aletiske nivået. Man har dermed tilgang på fasiten. Siden den materielle retten er en menneskelig konstruksjon, er den også i høy grad tilgjengelig for å bli prosessert riktig på det psykologiske nivået.

Det eksisterer også en velfungerende tilbakemeldingssløyfe i form av at fagdommere blir trent i juridisk metode under studiet og i arbeidslivet. Der får man kontant tilbakemelding på om det er riktig rettsanvendelse. På studiet blir eksamen sensurert med tilbakemelding per telefon, og man blir korrigert via undervisning. I arbeidslivet blir rettsanvendelsen overprøvd av klageinstanser eller kollegaer. I retten prøves rettsanvendelsen av lagmannsrettene og Høyesterett. Jeg argumenterer dermed for at man kan anta at en fagdommer opparbeider seg ekspertintuisjon for juridiske vurderinger, på samme måte som matematikeren vet svaret før hen vet løsningen.

På bevisteoriens område stiller det seg imidlertid annerledes. Her skal dommeren avgjøre hva som er faktumet i saken. Dette er ikke en juridisk vurdering, men en vurdering om en påstand er sann eller usann. Som drøftet i kapittel 2 er bevisvurderingen ikke verifiserbar. Man kan ikke med sikkerhet på det epistemiske nivået, få bekreftet hva som har skjedd på det aletiske. Bevisbedømmer er dermed avskåret fra en fasit hen kan bygge erfaring på og trene opp intuisjonen.

Ferdighetsillusjon

For å svare på spørsmålet som ble stilt innledningsvis, så kan fagdommere stole på intuisjonen sin i juridiske vurderinger, i større grad enn lekdommerne. Men jeg argumenter

for at fagdommere ikke vil ha noen bedre intuisjon enn lekdommere i *bevisvurderingen*. Problemet oppstår dersom fagdommeren ikke kjenner grensen for sin ekspertise. Dersom fagdommeren overfører selvtilitten fra å ta juridiske vurderinger til bevisvurderinger, oppstår det man kaller en *ferdighetsillusjon*.¹¹⁶

Ferdighetsillusjonen er en overdreven *tiltro* til egne og andres evne til å predikere et utfall.¹¹⁷ Det må imidlertid poengteres at tiltro er en *følelse*. Denne subjektive følelsen er ikke nødvendigvis fundert i en evaluering av sannsynligheten for at denne vurderingen er korrekt. Følelsen gjenspeiler kun om System 1 finner påstanden *koherent*, som igjen gir en følelse av kognitiv letthet. Men det som oppleves som koherent er ikke nødvendigvis det samme som hva som er sant. WYSIATI vil pusle sammen koherente fortellinger ut fra tilfeldige observasjoner for å unngå tvil. Gitt at heller ikke norske dommere har noen tilbakemeldingssløyfe på om bevisvurderingene deres er feil, bygger de seg opp et erfaringsgrunnlag som ikke er korrigert for feil. De sitter dermed igjen med en sterk *tiltro* til at de gjør gode vurderinger, kun i kraft av at de har gjort det *mange* ganger, ikke fordi de har fått bekreftet at det har vært *riktige* vurderinger. Sterk tiltro er dermed en indikator for at man lider av denne illusjonen, med mindre man faktisk er i en oversiktlig situasjon der man kan opparbeide seg ekspertintuisjon.

Forskning på aktører i aksjemarkedet har vist at ferdighetsillusjonen fører til at prediksjonene til de som vet litt, er bedre enn de som ikke vet noen ting, mens de som vet mest har minst pålitelige prediksjoner.¹¹⁸ Årsaken til dette kommer av at desto mer kunnskap man tilegner seg, jo mindre vår blir man på usikkerhetsmomentet. Man blir da urealistisk trygg på sine egne meninger. Erfaring på usikre fagfelt som bevisvurdering og aksjeplukking, har en raskt avtakende nyttefunksjon.

Problemet med slike kognitive illusjoner er at de er hardnakkede. Hvis man sammenlikner fenomenet med visuelle illusjoner ser så hjelper det ikke å vite at det er en illusjon, man ser illusjonen likevel. Det samme gjelder for de kognitive illusjonene. Selv om aksjebransjen er fullt klar over denne forskningen, som har foreligget siden 80-tallet, stopper de ikke å gi prediksjoner.¹¹⁹ Dette kommer i hovedsak av to ting.¹²⁰ For det første er det *følelsene* som

¹¹⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.231

¹¹⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.230

¹¹⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.238

¹¹⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.231

¹²⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.236

styrer handlingene våre, ikke det intellektuelle. For det andre viser relativt entydig forskning at folk kan opprettholde en fast tro på det aller meste så lenge det foreligger en bransjekultur som utgjør et sterkt felleskap av likesinnede.

Man kunne imidlertid tenke seg at dommeryrket rekrutterte særlig analytiske personer med større anlegg for System 2 tenkning enn folk flest. En amerikansk metastudie av dommeratferd konkluderte med at fagdommere i høy grad er intuitive tenkere, slik som folk flest.¹²¹ Dette er i motsetning til ingeniører som er en yrkesgruppe som scorer særlig høyt på *kognisjon*. Altså tendensen til å tenke i System 2, fremfor System 1. I den ene studien de kommenterte, svarte majoriteten av dommerne feil på intuisjonstestene. Og de som valgte det intuitive gale svaret på testene synes at testen var enkel.¹²² Dette tyder på at dommere også lider av ferdighetsillusjoner.

Det har også blitt påvist at fagdommere oppfatter seg selv som bedre beslutningstakere enn de egentlig er.¹²³ I studien mente mer enn 80% at minst halvparten av kollegaenes avgjørelser i større grad enn dem selv ble endret ved anke. Logisk sett gir dette ikke mening, siden kun 50% av utvalget kan være bedre enn snittet.

Det at fagdommere er i overkant selvsikre på sine egne avgjørelser støttes også av feltarbeidet til Kolflaath. Han påpeker at det var én tydelig tendens til «at flere fagdommere enn meddommere hadde en klar oppfatning om skyldspørsmålet tidlig i diskusjonen. Mitt inntrykk var at meddommere oftere enn fagdommere følte behov for å diskutere seg igjennom saken under domskonferansen.»¹²⁴

Ferdighetsillusjonen og juryordningen

Ferdighetsillusjonen har en relevans til diskusjonen rundt juryordningen. Da juryordningen ble innført, uttalte Jurykommisjonen fra 1885 at begrunnelsen for å innføre en jury var å sikre en *uavhengig* motvekt til at fagdommere i kraft av sin stilling blir blendet av sin ensidige erfaring. Det ble uttalt at «Men Livet og de menneskelige Forhold ikke lader sig saaledes schematisere, og Følgen vil være at en Uskyldig kan blive straffældt».¹²⁵ Ordningen ble imidlertid avvirket igjen i 2018 til fordel for meddomsrett. Begrunnelsen var å sikre riktig

¹²¹ (Rachlinski & Wistrich, 2017) s.14

¹²² (Guthrie, Jeffrey, & Wistrich, Blinking on the bench: how judges decide cases, 2007)

¹²³ (Guthrie, Rachlinski, & Wistrich, Inside the Judicial Mind, 2001) s.813 flg.

¹²⁴ (Kolflaath , Domskonferanser i straffesaker , 2018) s.366

¹²⁵ (NOU 2011:13 Juryutvalget: Når sant skal skrives) s.33

rettsanvendelse, og begrunnelse for standpunkt i skyldspørsmålet.¹²⁶ Forutsetningen for avviklingen var at man kunne sikre det beste av begge verdener i en meddomsrett. Lekmennene representerer et mer diversifisert verdi- og erfaringsgrunnlag for bevisvurderingen og subsumsjonen. Mens fagdommere sørger for oversikt over det prosessuelle og riktig rettsanvendelse.

Det er imidlertid ikke gitt at lekdommers verdi- og erfaringsgrunnlag kommer til nytte i bevisvurderingssituasjonen i en meddomsrett dersom man tar hensyn til det jeg her har drøftet om ferdighetsillusjonen og om gruppedynamikker. Hvis fagdommer lider av en ferdighetsillusjon, vil hen lettere falle ned på en konklusjon uten å konsultere System 2. Gitt at fagdommer flagger dette selvsikre standpunktet tidlig i rådslagningen, vil dette utløse en infomasjonskaskade, som vil vende de tvilende meddommerne til å holde med første taler. Dette taler ytterligere for et behov for at fagdommer først og fremst bør innta rollen som *tilrettelegger*, og forholde seg til en oppsatt sjekkliste over mulige feilkilder i rådslagningen.

¹²⁶ (NOU 2011:13 Juryutvalget: Når sant skal skrives) s.170 og s.102

4 Fortellingsperspektivet

4.1 Samme fenomen, forskjellig språkdrakt

I kapittel 2 drøftet jeg *hva* som skal til for at en påstand skal legges til grunn som faktum og konstaterte at *hvordan* dommeren skal la seg overbevise ikke er regulert av rettsregler. Det er dermed opp til juridisk litteratur å tolke hva «sannsynlighetsovervekt» og «utover enhver rimelig tvil» innebærer. I kapittel 2 kategoriserte jeg *hvilken type vurdering* bevisvurdering er, for å kunne diagnostisere hvordan den menneskelige hjernen foretar slike vurderinger. For så å introdusere *beslutningskjeden og meningsskapingsmaskinen* i kapittel 3, som forklarer mekanismene bak heuristikkene som bevisbedømmer anvender når hen skal vurdere sannsynligheter. I dette kapittelet starter analysen av den etablerte bevisteorien, *fortellingsperspektivet* for å vurdere i hvilken grad den tar høyde for *kognitive skjevheter*.

Jeg starter analysen med fortellingsperspektivet fordi den i utgangspunktet er deskriptiv.¹²⁷ Det vil si at den i likhet med heuristikk- og skjevhetsprogrammet prøver å fastslå hva som gjør at dommere lar seg overbevise. Slik jeg forstår det beskriver de to tilnæringsmåtene i stor grad det samme fenomenene bare i forskjellig språkdrakt. Den konsentrerer seg om hva bevisbedømmer *gjør*. Bevisbedømmer skal ta stilling til sannsynligheten for om en hendelse har inntruffet. Ifølge fortellingsperspektivet vurderer bevisbedømmer sannsynligheter ved å konstruere forskjellige fortellinger for seg selv. For så vurdere hvilken av dem som er mest *plausibel*. Mens ifølge beslutningsteori *erstatte* man det vanskelige spørsmålet om hvor *sannsynlig* en hendelse er med hvor *representativ* situasjonen er for fenomenet. Det vil si man vurderer hvor mye den hendelsen som blir skildret, likner på det man har erfart før, stereotypien. Hva som avgjør hvilken fortelling bevisbedømmer blir mest overbevist av, vil være den fortellingen som ligger tettest opp mot stereotypien. Beslutningsteori supplerer dermed fortellingsperspektivets forklaring på *hvorfor* noen fortellinger er mere overbevisende enn andre.

¹²⁷ (Kolflaath, 2004) s.280

4.2 Fortellinger

4.2.1 Kjennetegn på en fortelling

Fortellingsperspektivet argumenterer for at fortellingens funksjon er å gi mening til bevisene.¹²⁸ Fortellingene er ulike *kontekster* som gir bevisene mening.¹²⁹ Bevisene taler altså ikke for seg selv, men kan ha forskjellig betydning alt etter hvilken fortelling de tolkes inn i.

Kolflaath definerer en fortelling som en *kronologi* med en *indre sammenheng*.¹³⁰ Det vil si at elementene i historien er organisert langs en tidsakse, kronologisk, og det som driver fortellingen fremover er en hendelse som leder til en annen, hvilket gir en indre sammenheng. Elementer som inngår i kronologien, er både ytre hendelser, som at A kjørte på rødt lys. Men også psykologiske elementer som motiver for en handling. Eksempelvis A kjørte på rødt lys, *fordi* kona som satt i bilen hadde hjerteinfarkt og måtte komme seg på sykehus. Kronologien kan dermed omtales som årsakskjeder, der det ene elementet utløser det neste elementet.¹³¹ Det som skjedde forut for et element, er *forklaringen* på hvorfor dette elementet inntraff.

4.2.2 Hvordan fortellinger dannes i retten

Fortellingen utvikles under hovedforhandlingene.¹³² Utgangspunktet for fortellingen er for straffesaker gjerningsbeskrivelsen i tiltalen, og påstandsgrunnlaget til partene i sivile saker. Det må imidlertid påpekes at det naturlig nok utvikles konkurrerende fortellinger.¹³³ Dersom tiltalte nekter straffeskyld, vil tiltaltes frifinnende fortelling konkurrere mot aktor sin fellende fortelling. I en sivil sak kan også partene være uenig om hva som faktisk har skjedd og deres fortellinger vil konkurrere om å bli lagt til grunn som rettslig relevante fakta. Disse fortellingene vil under forhandlingene bli konkretisert og til dels modifisert.

Sentrale bidragsyterne til fortellingene vil i straffesaker være tiltalte, mens det i sivile saker vil være partene. Vitner med selvopplevde hendelser vil også legge frem rettslige relevante fortellinger. Særlig vil advokatene bruke forhandlingene til å få en fortelling til å feste seg, med måten bevis blir fremlagt og forklart, og hvordan vitner blir forhørt. For så oppsummere

¹²⁸ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.98

¹²⁹ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.95

¹³⁰ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.102

¹³¹ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.104

¹³² (Kolflaath, Bevisbedømmelse i praksis, 2013) s.106

¹³³ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.108

fortellingen i det avsluttende innlegget. Line Normann Hjort argumenterer for at advokatenes sin utspørring av vitner ikke handler om å komme nærmere sannheten, men heller bekrefte den allerede etablerte fortellingen de prosederer på ved hjelp at det hun kaller språkspill.¹³⁴ Eller det man i beslutningsteori ville kalt etablerte kulturelle stereotyper.

Til slutt er det dommeren som skal konkludere med hva som er rettslige relevante fakta. Kolflaath påpeker to måter dommeren bidrar til å utvikle den endelige fortellingen som legges til grunn for dommen.¹³⁵ Den ene måten er at bevisbedømmer setter sammen ulike deler av bevisførselen til én sammenhengende fortelling, ut fra hvilke deler av forskjellige historier hen finner plausibel. Ofte legger dommeren til grunn hele fortellingen til én enkelt aktør. Den andre måten er at dommeren selv konstruerer en fortelling ut fra den pågående bevisførselen. Samt fyller ut verserende fortellinger under og etter hovedforhandlingene

4.2.3 Generell og saksspesifikk plausibilitet

Fortellingene vurderes da ut ifra hvilken grad dommeren tror at de er sanne. I

fortellingsperspektivet er målet på hvilken grad en fortelling ansees som sann *plausibilitet*.

Det vil si graden av bevisbedømmers overbevisning.¹³⁶ Spørsmålet er da hva som øker plausibiliteten til en fortelling. Tilbake i kapittel 2 diskuterte jeg at bevisvurderingen er en ikke-verifiserbar vurdering, bevisbedømmer må i stede vurdere overbevisningsgraden sin etter styrken på det *indre signalet* fremfor en fasit. I det følgende vil jeg redegjøre for hvordan dette indre signalet styrkes ifølge fortellingsperspektivet. Det vurderes langs to dimensjoner, *generell og saksspesifikk plausibilitet*.¹³⁷

4.3 Generell plausibilitet

4.3.1 Hva styrker det indre signalet - koherente fortellinger

Generell plausibilitet er hvor godt fortellingen fungerer i seg selv. Hvor troverdig er fortellingen som en fortelling. Mens saksspesifikk plausibilitet går ut på hvor godt fortellingen dekker alle bevisene.¹³⁸ Typisk vil den generelle plausibiliteten bero på

¹³⁴ (Hjort, 2021) s.28

¹³⁵ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.107

¹³⁶ Plausibilitet er synonymt med *troverdighet*, men siden man i dagligtalen forbinder troverdighet som en egenskap ved personer og ikke fortellinger, er det mer klargjørende å bruke ordet plausibilitet.

¹³⁷ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.119

¹³⁸ (Hastie & Pennington, The story model for jurir decision making, 1993) s.198

dommerens forventninger til fortellingens *struktur*, og *innhold*. I likhet med beslutningsteori påpeker fortellingsperspektivet at en plausibel fortelling er *koherent*. Det vil si at fortellingen er *strukturert* kronologisk, der et element forklarer *årsaken* til neste element. Fortellingens *innhold*

Generelle erfaringsgrunnsetninger og stereotypier

Når det kommer til *innhold*, siktes til det til hvilke antakelser bevisbedømmer har om hvordan hendelser *normalt* utspiller seg. Disse antakelsene blir i juridisk litteratur omtalt som *generelle erfaringsgrunnsetninger*. En fortellings troverdighet blir da vurdert etter hvor tett den ligger opp mot bevisbedømmers erfaringsgrunnsetninger. Et sentralt spørsmål er da hvordan disse erfaringsgrunnsetningene dannes, med tanke på om de er pålitelige kilder å trekke slutninger fra. Jeg mener at man her kan se til verdifulle bidrag fra det beslutningsteori kaller representativitetsheuristikker som danner og anvender det man kjenner som stereotypier.

Stereotypi

Ifølge beslutningsteori anvender man noe som heter *representativitetsheuristikker* nå man skal vurdere plausibiliteten for en hendelse. Dersom man blir spurt om hvor *sannsynlig* det er for at det var Peder Ås som drepte Marte Kirkerud, er dette et svært komplekst spørsmål. Det mentale haglegeværet vil dermed fremkalle et enklere spørsmål å svare på. Et slikt enklere spørsmål er en vurdering av *representativitet*. Man *erstatter* spørsmålet om hvor sannsynlig noe er, med hvor representativ situasjonen er for fenomenet.¹³⁹

Hvis X er representativt for Y, betyr det at de *likner*.¹⁴⁰ Ser Viggo Kristiansen ut som en drapsmann? Likner han på den forestillingen man har av hvordan drapsmenn ter seg? Representativitet handler altså om hvordan et konkret tilfelle passer med *stereotypien*.

I delkapittel 3.2.1 forklarte jeg at System 1 sin hovedoppgave er å opprettholde og vedlikeholde en stilisert modell av verden. Denne stiliseringen er det man oppfatter som normalt. I likhet med erfaringsgrunnsetninger er stereotypier altså sosiale kategorier av

¹³⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.164

¹⁴⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.164

normer.¹⁴¹ Hvordan disse sosiale normene dannes må forklares ved hjelp av to ytterligere heuristikker, tilgjengelighets- og affektheuristikker.

Tilgjengelighetsheuristikk

Hva som er normalt, er et spørsmål knyttet til hvor hyppig man forventer at et fenomen skal oppstå. Eksempelvis er det få personer som blir drept, rent statistisk sett. I 2021 var det kun 29 personer som ble drept i Norge.¹⁴² Den subjektive oppfattelsen av hva som er normalt tar imidlertid ikke utgangspunkt i empiriske data. Når man skal anslå hyppigheten av en kategori benytter vi oss av *tilgjengelighetsheuristikker*.¹⁴³ Man erstatter spørsmålet om et *objektivt anslag av hyppigheten* til en kategori, med et inntrykk av hvor *lett* og hvor *god flyt* man har for å komme på tilfeller innenfor denne kategorien. I stedet for å henvende seg til drapsstatistikken, prøver man å komme på eksempler på saker man kjenner til om folk som har blitt drept.

Hvor *lett* man kommer på tilfeller kan være betinget av statistisk frekvens, men det gjør også andre faktorer. *Iøynefallende* hendelser som tiltrekker seg oppmerksomhet, vil gjøre det lettere å komme på eksempler.¹⁴⁴ Menn i lederstillinger som misbruke sin stilling for å skaffe seg seksuell omgang fikk under metoo-skandalen mye oppmerksomhet. Rettssaker med mye medieomtale som Gaute Drevdal-saken i Norge og Virtanen-saken i Sverige, gjør det *lett* å komme på eksempler av menn som misbraker sin stilling til seksuell omgang.¹⁴⁵ Dette gjør at man anslår fenomenet som mer normalt enn det rent statistisk er. *Dramatiske* hendelser gjør kategorier midlertidig mer tilgjengelig. Man ser at under etterspillet av naturkatastrofer går kjøp av forsikringer opp, på tross av at nye katastrofer ikke er noe mer sannsynlig rett etter katastrofen.¹⁴⁶ Dramatisk er en annen side av iøynefallende, slik at sakene til Drevdal og Virtanen vil også sparke opp hvor hyppig man tenker metoo-saker er så lenge disse sakene befinner seg i den kollektive bevisstheten. Personlige erfaringer, bilder og levende eksempler er mer tilgjengelige enn andrehåndsfortellinger og statistiske fakta. En kvinnelig dommer som

¹⁴¹ (Kahneman, Tenke, fort og langsomt, 2013)s.184

¹⁴² (Politiet: Kripos , 2021)s.4

¹⁴³ (Kahneman, Tenke, fort og langsomt, 2013) s. 144

¹⁴⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.144

¹⁴⁵ (Dybvig, 2022) (Teigen, 2020)

¹⁴⁶ (Kahneman, Tenke, fort og langsomt, 2013) s.151

selv har opplevd seksuell trakassering vil lettere kunne sette seg inn i et voldtekstoffers forklaring og finne den mer plausibel, enn en dommer uten denne erfaringsgrunnsetningen.¹⁴⁷

Kategoriens størrelse anslås også av *flyt*.¹⁴⁸ System 1 anslår kategorien som stor dersom man har god flyt i å komme på eksempler. Men som oftest anvender man denne heuristikken så fort at man ikke tar seg bryet med å faktisk komme på eksempler, man går kun etter inntrykket av hvor lett og hvor god flyt man vil ha for å komme på eksempler.¹⁴⁹

Eksempelvis vet man umiddelbart at det er flere ord man kan lage med bokstavene «TAPCERHOB» enn «XUZONLCJM».

Tilgjengelighetsheuristikken er et samspill mellom System 1 og System 2 hvilket betyr at det er en *viljestyrt* problemløsningsstrategi. Dette i motsetning til en automatisk prosess som kun styres av System 1.¹⁵⁰ En automatisk prosess kan korrigeres for ved hjelp av aktivering av System 2, men i en viljestyrt prosess er System 2 en del av tommelfingerregelen og kan dermed ikke brukes til å kontrollere seg selv. Det er dermed vanskelig å rette opp kognitive skjevheter som skyldes viljestyrte heuristikker.

Affektheuristikk

Fellesfaktoren til betingelsene av hva som gjør et minne *lett* tilgjengelig er at de engasjerer *følelsene* våre. Hvor lett et eksempel dukker opp i bevisstheten, er uløselig knyttet til den følelsesmessige reaksjonen det vekker.¹⁵¹ Skremmende tanker og bilder er lett tilgjengelig i hukommelsen, som igjen gir et inntrykk av stor hyppighet, som igjen forverrer frykten for disse scenariene. Det at man konsulterer følelsene sine når man skal foreta vurderinger og valg kalles en affektheuristikk.¹⁵² Man erstatter spørsmålet «Hva *tenker* jeg», med «Hva *føler* jeg?». Dersom man liker noen, har man en tilbøyelighet til å tro på det de sier og vice versa. Dersom man har en positiv personlig opplevelse med mannlige ledere, har man mer vanskeligheter for å komme på eksempler av mannlige ledere som misbruker makt, og vil ha verre for å tro på en fornærmet som anklager sjefen sin for misbruk av sin stilling.

¹⁴⁷ (Robberstad, 2014) s.115

¹⁴⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.143

¹⁴⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.144

¹⁵⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.143

¹⁵¹ (Kahneman, Tenke, fort og langsomt, 2013) s.153

¹⁵² (Kahneman, Tenke, fort og langsomt, 2013) s.153

Affektheuristikken er altså helt avgjørende for *når* bevisbedømmer *føler* seg overbevist, altså når påstanden er sannsynliggjort. Når fortellingen som dannes i retten ligger tett opp mot stereotypien føles fortellingen som koherent og dommeren føler seg overbevist.

I motsetning til tilgjengelighetsheuristikken er affektheuristikken i utgangspunktet en automatisk prosess. Det vil si at den er et produkt av System 1. Det er System 2 sin oppgave å være kritisk til svarene System 1 avleverer. Beklageligvis tyder forskningen på at når det kommer til holdninger og personlige overbevisninger er System 2 en apologet for System 1 sine resultater.¹⁵³ Dette er antakelig en følge av bekreftelsestendensen som også er System 2 styrt.

Erfaringsgrunnsetninger er ikke rasjonelle analyser

Stereotypier blir dermed dannet i et samspill mellom tilgjengelighets- og affektheuristikken. Hva man oppfatter som *normalt* er et resultat av en prosess av hvor *lett* og hvor god *flyt* man har for å komme på eksempler innenfor den gitte kategorien. Intuitive erfaringsgrunnsetninger er altså ikke et resultat av en rasjonell analyse, og dermed ikke pålitelige.

Ta eksempelet med Viggo Kristiansen. Vitner sier han i barndommen var aggressiv, han hadde ADHD og ble dømt for overgrep på en annen jente. Dette er alle karaktersikker som er dramatiske og iøynefallende, og som vekker sterke følelser i oss. Generelt sett vil WYSIATI-effekten i slike situasjoner gjøre at man pusler sammen en koherent fortelling ut fra de bevisene man blir fremlagt. Videre tar man seg ikke en gang bryet med å komme på faktiske eksempler med personer som har disse karakteristikkene som har begått drap. Man bare «vet» at dette vil være enkelt, for det passer for godt med *stereotypien* drapsmann. Man *erstatte* da spørsmålet «Forgrep seg og drepte Viggo Kristiansen Lena Paulsen og Stine Sofie Sørstrønen 19.mai 2000 i Baneheia» med «ser Viggo Kristiansen ut som en som kunne forgrepet seg på og drept to barn». Det er imidlertid to vidt forskjellige spørsmål å vurdere sannsynligheten for at en konkret hendelse *har* inntruffet, og sannsynligheten for om at en hendelse *kan* inntreffe. Det første spørsmålet vender oppmerksomheten mot *bevisverdien* til de konkrete bevisene som er fremlagt i saken. Det andre spørsmålet vender oppmerksomheten mot hvor representativ Viggo Kristiansen generelt er for drapsmenn.

¹⁵³ (Kahneman, Tenke, fort og langsomt, 2013) s.116

Fortellinger som ligger tett opp mot stereotypien, vil dermed styrke bevisbedømmer sitt *indre signal*, og finne fortellingen mer plausibel enn en fortelling som bryter med den stereotypien. Men som jeg har drøftet, skapes ikke stereotypier ut fra statistiske analyser, men snarere på bakgrunn av ens følelser. Stereotypier eller erfaringsgrunnsetninger er dermed lite pålitelige kilder til å vurdere fortellingens generelle plausibilitet. Som jeg vil drøfte i det følgende begår man en rekke feilslutninger når man benytter seg av representativitetsheuristikker i sannsynlighetsbedømming.

Representativitetsheuristikken som skjevhet

Representativitetsheuristikken er mest berømmelig blitt påvist hos lekfolk av Kahneman og Tversky gjennom Tom W.s masterstudium-eksperimentet.¹⁵⁴ Deltakerne ble introdusert til den fiktive masterstudenten Tom W. Oppgaven var å rangere hvilket studium han fra mest til minst *sannsynlig* gikk på. For å foreta vurderingen fulgte det en personlighetsskisse av Tom fra da han gikk siste året på videregående. Grunnlaget for skissen var psykologiske tester av *ukjent validitet*. Skissen beskriver han som *intelligent, men lite kreativ. Han har behov for orden. Han skriver kjedelig og mekanisk, men kan krydre det til med nokså sprø ordspill og innslag av science fiction-aktig fantasi. Kompetanse er en stor drivkraft for ham. Han har lite sympati og finner lite glede i samvær med andre. Men han har en utpreget rettferdighetssans*. Fagene deltakerne skulle rangere var:

Bedriftsledelse

Informatikk

Ingeniørfag

Humaniora og pedagogikk

Juss

Medisin

Bibliotekfag

Fysikk og biologi

Samfunnsvitenskap

¹⁵⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.160

Sosialfag

Den *rasjonelle* måten å foreta rangeringen av fagene på er å foreta en *bayesiansk analyse*, hvilket jeg vil beskrive mer inngående i kapittelet om sannsynlighetsparadigmet. Den bayesianske analysen ville startet med å vurdere hvilke av disse studieretningene som er størst. Uten å vite noe mer om Tom W. enn at han er student ved et stort delstasuniversitet, er det rasjonelle utgangspunktet å rangere studiene etter hvor store de er. Sannsynligheten er naturlig nok større for at han går på et studium der det er mange, fremfor få. Det er dermed ikke urimelig å anta at det er flere som går på bedriftsledelse, juss, ingeniørfag enn informatikk, fysikk og biologi, bibliotekfag og humaniora og pedagogikk.

Men så har man et «bevis» som kan dra sannsynligheten i retning av et mindre studium. Beskrivelsen tyder på at sannsynligheten for at Tom W. går på et mer teknisk studium er høyere. Man kan imidlertid ikke ta innholdet av et hvert bevis for god fisk. Før man legger det til grunn må man vurdere bevisets *kvalitet*. Hvor troverdig er beviset? Det står at beskrivelsen er fra videregående. Tiden har altså gått siden beskrivelsen ble nedtegnet. Tom W. kan altså ha endret seg i løpet av bacheloren og fått andre interesser. Det står også at skissen er av *ukjent kvalitet*, det kan med andre ord være at beskrivelsen simpelt hen ikke stemmer. Personlighetstester kan være av varierende kvalitet. Man må dermed være forsiktig med å legge for stor vekt på dette beviset når man skal vurdere sannsynligheten for hvilket studium Tom W. går på. Det er også bare *ett* bevis som tilsier at Tom W. går på et mer teknisk studium. Beviset har dermed begrenset beviskraft i denne sammenheng. Men man er tvunget til å ta et standpunkt på bakgrunn av den begrensede informasjonen man har. En rimelig antakelse kan da være at blant fagene med tilsvarende størrelse er det mer sannsynlig at han går tekniske fag som ingeniørfag fremfor juss, mens et mindre fag som informatikk har fremdeles mindre sannsynlighet enn juss. Innenfor mindre fag vil likevel informatikk gå fremfor bibliotekfag.

Deltakerne i eksperimentet, satte på sin side systematisk fag som passet med stereotypien av personlighetsbeskrivelsen. Det vil si fag som informatikk og ingeniørfag høyt, mens fag som ikke passet med stereotypien som humaniora, samfunnsfag og bedriftsledelse lavt. De *erstattet* dermed spørsmålet om hvor *sannsynlig* det er at han går de forskjellige studiene, med hvor *representativ* han er for de forskjellige studiene. Deltakerne ignorerte da det som kalles *grunnfrekvensen* for hvilket studium Tom W, gikk på. I dette tilfellet var det størrelsen på de forskjellige studiene. De tilla også personbeskrivelsen for mye vekt, fordi de ikke tok stilling

til bevisets *kvalitet* og veide det heller ikke opp mot at man manglet *ytterligere* bevis som pekte i samme retning. Representativitetsheuristikken utgjør dermed en kognitiv skjevhet i den forstand at man ikke tar hensyn til grunnfrekvensen, samt bevisets sin mengde og kvalitet.

Representativitetsheuristikken har særlig blitt påvist hos lekfolk, men også fagdommere, dog i noe mildere grad.¹⁵⁵

Standardiserte fortellinger

Stereotypier dannes dermed ut fra det man oppfatter som *normalt*. Hva man oppfatter som normalt, vil i stor grad være antakelser basert på bevisbedømmers *personlige* opplevelse av hendelsers årsak-konsekvens. Men siden mennesker er sosiale dyr, som har bygget sivilisasjoner ved å dele kunnskap, er bevisbedømmer ikke avhengig av å ha erfart enhver hendelse personlig for å ta lærdom av den. Flere bevisteoretikere argumenterer for at disse erfaringsgrunnsetningene også baserer seg på det Andre Kjus kaller *nøkkelfortellinger*, Hjort kaller *standardiserte fortellinger* og det Kolflaath kaller for *forståelsesrammer*.¹⁵⁶ Det er altså ikke vanntette skott mellom personlige erfaringer og de standardiserte fortellingene. Det er de standardiserte fortellingene man bruker for å tolke og forstå egne erfaringer. De utgjør grunnlaget for hvordan man forstår seg selv og verden rundt.

Fortellingene det refereres til er gjennomgående fortellinger i vestlig kultur om hvordan verden henger sammen.¹⁵⁷ De samme temaene går igjen fra Odysseen, til Bibelen til Shakespare og Ibsen til hvordan man formidler hendelser i avisen. Nøkkelmetaforen til Kjus, referer til at fortellinger brukes av aktørene i rettsalen som nøkler til forståelse hos bevisbedømmer. De utgjør bakteppe som former ens oppfatning av hva man opplever som plausible handlingsforløp og troverdige menneskelige relasjoner. Man bruker med andre ord stereotypier/referanserammer slik at det blir lettere å fylle ut, tolke og forstå handlingene som har utspilt seg i den konkrete saken.¹⁵⁸

I Hjort sin bok bruker hun standardiserte fortellinger som tolkningsramme for hvilken fortelling dommerne fant mest plausibel i Orderud saken. Veronica blir fremstilt som den utpekulerte hovedhjernen. Hun giftet seg med Per Orderud for å klatre opp den sosiale

¹⁵⁵ (Guthrie, Jeffrey, & Wistrich, *Blinking on the bench: how judges decide cases*, 2007) s.22 flg.

¹⁵⁶ (Kjus, 2008) s.100 flg. (Hjort, 2021) s.55 flg. (Kolflaath, *Bevisbedømmelse i praksis*, 2013) s.110 flg.

¹⁵⁷ (Hjort, 2021) s.55

¹⁵⁸ (Kjus, 2008) s.100

rangstigen. Og manipulere han til å bli med på trippeldrap, for å sikre seg arven på gården. Dette drar på en arketypp Anima, en mytologisk kvinneskikkelse.¹⁵⁹ Den spiller på at kvinner har en dobbelthet, en lys og skyggeside. Bak madonnaen lurer alltid en skjøge.¹⁶⁰ Denne fortellingen forklarer hvordan hun som fremstår og fremstiller seg selv som pålitelig, kontrollert, omtenkstom og velartikulert i rettsalen kan begå et trippeldrap.

Det er ikke tilfeldig at standardiserte fortellinger har denne rollen for hva man oppfatter som plausibelt. Først og fremst finner man fortellinger plausible i kraft av at de blir repetert ofte. I kapittel 3 kommenterte jeg at repetisjon fører til at man finner det familiært og derav troverdig. Man blir med andre ord primet til å tro på disse narrative.

4.3.2 Hva svekker det indre signalet

Fortellingens struktur: Uklar årsakskjede

Spørsmålet er da hva som gjør at man oppfatter en fortelling *mindre* plausibel enn en annen, ut fra fortellingens generelle plausibilitet. Når det kommer til *struktur* blir fortellingen oppfattet som mindre plausibel, dersom elementene i fortellingen har en uklar rolle, drar i forskjellige retninger eller at elementene ikke har en klar kontinuitet. Dette resulterer i en *uklar årsakskjede*. Det *indre signalet* blir dermed svekket på grunn av mangel på koherens. Dersom man har elementer som er selvmotsigende eller logisk inkonsekvent, vet man at fortellingen ikke kan stemme. En person kan for eksempel ikke være på to steder til samme tid. Det trengs imidlertid ikke mer enn at det foreligger spenninger mellom elementene, før tvil oppstår. Eksempelvis dersom man blir stoppet i politikontroll under mistanke om å ha kjørt i ruspåvirket tilstand, fremstår det ikke som særlig rasjonelt å ta alle rusmidlene man er i besittelse av på en gang, for å unngå å bli tatt for besittelse.¹⁶¹ Et element må også tydelig forklare det neste, slik at fortellingen blir *sammenhengende*.

Brudd med stereotypien

For *innholdet* sin del vil det svekke fortellingens plausibilitet dersom aktørene har oppført seg atypisk, tatt uhensiktsmessige handlingsvalg, glemt eller ikke fått med seg åpenbare ting eller at handlinger har skjedd på urealistisk tid. I Orderud-saken prøver Veronica å fremstille seg

¹⁵⁹ (Hjort, 2021) s.246

¹⁶⁰ (Hjort, 2021) s.247

¹⁶¹ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.122

som godheten selv.¹⁶² Hun er mekleren og den omsorgsfulle i Orderudfamilien. Samtidig blir det klart at hun på lille julaften setter frem et glass med sterke piller til sin rusmisbrukene halvsøster, for å forhindre at søsteren skapte noen «scener».¹⁶³ Denne handlingen er vanskelig å få til å stemme overens med rammefortellingen/stereotypien som Veronica prøver å presentere av seg selv. Det fremstår som en atypisk handling for en omsorgsfull empatisk person å gi rusmidler til en rusmisbruker. Dette er med på å svekke det indre signalet.

Skjevhet: Ufølsom for utvalgets mengde og kvalitet

Det er en ting hva som faktisk svekker det indre signalet, en annen ting er hva som *burde* svekket det. Som nevnt i delkapittel 4.3.1 vil representativitetsheuristikken gjøre at man *kun* vurderer histories troverdighet ut fra om den er *sammenhengende*, men ignorerer *mengden* og *kvaliteten* på informasjonen som har blitt presentert.¹⁶⁴

En sammenhengende fortelling kan bestå av få bevis. Få elementer er som oftest lettere å sette sammen enn mange. Eller at få motstridene elementer får en sammenheng ved at fortellingen blir fylt ut av elementer som skaper sammenheng. Det indre signalet er dermed ukritisk til *mengden* informasjon, så lenge informasjonen skaper en sammenheng. Historien med få eller mange elementer *føles* like plausibel, så lenge den henger sammen. Videre er man også tilbøyelig til å være ukritisk til hvor informasjonen kommer fra. Det er som i eksperimentet med Tom W. der man var intuitivt ikke er kritisk til hvor personlighetsskissen kom fra. System 1 er ikke kildekritisk, det er en System 2 oppgave.

Jeg mener det at bevisbedømmer intuitivt ikke er kritisk til mengden og kvaliteten på informasjonen er et poeng som støtter Strandberg sitt argument om å innføre et robusthetskrav i bevisretten. Robusthet er «fravær av skjørhet overfor ny relevant informasjon».¹⁶⁵ Bevisbedømmer sin overbevisningsgrad er betinget av sakens bevissituasjon. Per i dag er beviskravet i norsk rett tosidig.¹⁶⁶ Hvor sannsynlig hen finner at påstand x er sann eller usann er en bevisverdi på 100% som skal fordeles mellom x og ikke-x, uavhengig av om bevissituasjonen er «god» eller «dårlig». Partenes unnlatte bevisføring eller sikring påvirker dermed *ikke* sannsynlighetsvurderingen. Strandberg argumenterer dermed for at man bør

¹⁶² (Hjort, 2021) s.126

¹⁶³ (Hjort, 2021) s.142

¹⁶⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.96

¹⁶⁵ (Strandberg M. , 2012) s.542

¹⁶⁶ (Strandberg M. , 2012) s.95

innføre et robusthetskrav, for å kvalitetssikre bevissituasjonen.¹⁶⁷ En robust overbevisningsgrad vil i liten grad påvirkes av at ny informasjon kommer til. Bevissituasjonen er da billedlig sagt «mettet». Et krav om å vurdere om bevissituasjonen er robust, ville tvunget bevisbedømmer til å koble på System 2 til å vurdere mengden og kvaliteten på bevisene. Dette vil gjøre bevisvurderingen mer *treffsikker*.

Skjevhet: Ignorerer grunnfrekvensen

Den andre skjevheten som typisk følger av representativitetsheuristikken, er at man ignorerer grunnfrekvensen til påstanden som skal bevises.¹⁶⁸ Grunnfrekvensen er hvor ofte et fenomen oppstår i et gitt utvalg. For Eksempel dersom politiet stopper 100 tilfeldige biler på veien i en rutinekontroll og det viser seg at én av dem er påvirket av alkohol, vil grunnfrekvensen for fyllekjøring i dette utvalget være 1/100 eller 1%. Grunnfrekvensen for at en tiltalt har fyllekjørt før retten har fremlagt bevis, er dermed 1%. Rasjonelt sett er dette utgangspunktet for sannsynlighetsanslaget i bevisvurderingen, før man legger frem for retten bevis som knytter seg til den konkrete saken. Bevis som en test fra et alkometer som påviser alkohol i åndedrettet, vil klart trekke opp sannsynligheten for at tiltalte har fyllekjørt gitt hvor pålitelig testen er.

Neglisjering av grunnfrekvenser har blitt påvist blant jurister i Norge og Sverige, som bygger på funnene til Mark Schweizer.¹⁶⁹ I studien ble 173 sveitsiske dommere spurt om å vurdere et hypotetisk tilfelle om en bilist som var blitt stoppet i rutinekontroll. Som en del av rutinekontrollen måtte bilisten ta en alkoholtest. Vedkommende testet positivt på denne testen. Dommerne fikk opplyst at sjansen for at en tilfeldig bilist er påvirket er 1 av 100. Det ble oppgitt at det var 95% sannsynlig at alkoholtesten ga en sann positiv, som sier noe om påliteligheten til alkoholtesten. Med andre ord sannsynligheten for at bilisten vil teste positivt *gitt* at bilisten faktisk er påvirket. Sannsynligheten for en falsk positiv var på sin side 5%. Det vil si sannsynligheten for at bilisten tester positivt *gitt* at hen ikke er påvirket. Dommerne fikk så spørsmålet hva sjansen er for at bilisten faktisk var påvirket *gitt* den positive testen. De kunne svare i oppgitte intervaller. Det riktige intervallet var 0-25%. Man kan regne ut riktig intervall ved hjelp av Bayes Teorem, som jeg inngående skal gjennomgå i neste kapittel, som

¹⁶⁷ (Strandberg M. , 2012) s.569 flg.

¹⁶⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.165 flg.

¹⁶⁹ (Eide , 2016) s. 26 flg. og (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.128 flg. (Schweizer, 2005) s.161-162

gir $1/99 * 0,95 / 0,05 = 0,19 \rightarrow 0,19 / (0,19 + 1) \approx 16\%$. Dette var det kun 9% av dommerne som svarte, majoriteten svarte 76 – 100%. De så dermed kun på påliteligheten til beviset, og tok ikke hensyn til grunnfrekvensen.

Representativitetsheuristikken vil altså føre til at man *kun* anslår plausibiliteten ut fra hvilken fortelling man kan sy sammen av *bevisene*, men vil glemme å forankre plausibiliteten i grunnfrekvensen av fenomenet. Med andre ord vil man ikke i tilstrekkelig grad vekte inn at de fleste fenomener skyldes *tilfeldigheter*. Herunder fenomenene har en forklaring som er irrelevant for saken man vurderer. Heuristikken gjør dermed at man får en overdreven vilje til å spå forekomsten av usannsynlige hendelser.

Et eksempel der man tok hensyn til grunnfrekvensen var i Landås-saken (Rt.1999 s.1742). En sykepleier hadde blitt dømt i tingretten for drap på ti pasienter. Hun ble imidlertid frikjent i lagmannsretten, etter at man hadde regnet ut frekvensen på hvor mange dødsfall som *tilfeldigvis* kunne oppstå ut fra hennes vakt plan. Utrekningen viste at hyppigheten av dødsfall som hadde skjedd på hennes vakt, ikke var signifikant forskjellig fra hva som kunne vært forårsaket av *tilfeldigheter*. Man kunne dermed ikke påvise noen korrelasjon, og hvert fall ikke kausalitet.

4.4 Saksspesifikk plausibilitet

4.4.1 Etablerte omstendigheter

Vurderingskjeden

Den generelle plausibiliteten går som sagt ut på hvor godt fortellingene fungerer som fortelling. Den saksspesifikke plausibiliteten retter oppmerksomheten mot hvor godt fortellingen *forklarer* de *etablerte omstendighetene* i saken.¹⁷⁰ En fortelling oppfattes altså som mer plausibel enn en annen dersom den i større grad dekker bevisene i saken. Den saksspesifikke plausibiliteten er derfor et spørsmål om *forholdet* mellom fortelling og bevissituasjon.¹⁷¹ For å vurdere hvor godt dette forholdet er, må man se hvorvidt det foreligger konkrete holdepunkter i bevissituasjonen som gir belegg for de aktualiserte

¹⁷⁰ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.119

¹⁷¹ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.140

fortellingene. Eller om det foreligger noen problemer i bevissituasjonen som gjør at fortellingen ikke holder vann.

De etablerte omstendighetene består av bevisdata og etablerte fakta.¹⁷² Bevisdata kommer typisk fra forklaringer fra tiltalte, vitner, sakkyndige og dokumentbevis. Noe bevisdata vil ha en mer plausibel forklaring enn andre, og legges til grunn som etablerte omstendigheter. Etablerte fakta er bevisdata som bevisbedømmer direkte legger til grunn som fakta. Det gjøres enten fordi generelle erfaringsgrunnsetninger tilsier at man skal tro på datumet, eller at bevisbedømmer ikke har noen forutsetninger for å overprøve troverdigheten til datumet.

Eksempelvis har de fleste bevisbedømmere en erfaringsgrunnsetning om at man kan stole på politimenn og legger deres forklaringer direkte til grunn som etablerte omstendigheter. Sier politimannen at A ble funnet sovende i førersetet med promille i en veiskulder, så ansees det som etablert fakta av retten. En fortelling som forklarer at A satt sovende i bilen med promille, er at A kjørte bilen i ruspåvirket tilstand. Et eksempel på bevis bevisbedømmer vanligvis ikke har forutsetninger for å overprøve er DNA-bevis.

Vurderingskjeden blir da følgende: 1) Først må bevisbedømmer vurdere troverdigheten til bevisdatumet → 2) Dersom datumet ansees som troverdig blir det ansett som en *etablert omstendighet*. → 3) Når et datum blir definert som en etablert omstendighet forutsetter fortellingsparadigmet at den må ha en *forklaring*.¹⁷³

Den mest plausible forklaringen på de etablerte omstendighetene danner dermed en årsakskjede som utgjør en sammenhengende fortelling. Dannelsen av denne fortellingen er ikke lineær, men vekselvirkninger mellom de underliggende fortellingene som prosederes av aktørene i retten. De forskjellige bevisdata som fremføres vil kaste lys over hverandre og hvor troverdig forklaringene av bevisdataene fremføres.

4.4.2 Heuristikker og etablerte omstendigheter

Kan bevisbedømmer gjennomskue løgn via observasjon?

¹⁷² (Kolflaath, Bevisbedømmelse i praksis, 2013) s.171

¹⁷³ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.172

Spørsmålet er da hva som avgjør om bevisbedømmer finner bevisdatumet troverdig. Her har jeg allerede drøftet den generelle plausibiliteten, så nå retter jeg oppmerksomheten mot hva som avgjør den saksspesifikke plausibiliteten.

Et bærende prinsipp både i straffeprosessen jf. strpl §§278, 290, 296 og 297 og sivilprosessen jf. §21-9 er bevisumiddelbarhetsprinsippet. Bevisumiddelbarhet betyr at et bevisdatum skal tas med i retten og føres for domstolen. Det innebærer blant annet at det ikke er nok å lese opp vitneforklaringer i avhør, de må forklare seg på nytt under hovedforhandlingen. Prinsippets begrunnelse er formulert i gjenopptakelsen av Fritz Moen-saken (Rt.2003 s.1389 avsnitt 102) at «ved de forklaringene som gis, og de tilføyelser, endringer eller avklaringer som skjer ved eksaminasjon i retten, og det inntrykk som formidles gjennom *opptreden* i rettssalen og under åstedsbefaringer, har den dømmende rett gjennomgående et langt bedre grunnlag for å vurdere *troverdigheten* og *holdbarheten* av forklaringene enn det som kan oppnås ved lesing av de dokumentene hvor forklaringene er nedtegnet.» (min kursiv). Det foreligger dermed en antakelse om at retten kan danne seg et bedre og sikrere inntrykk av de involvertes troverdighet og pålitelighet ved å observere dem i live, enn å få forklaringer opplest av en advokat.¹⁷⁴ Med andre ord foreligger det en presumsjon om at man kan skille sannhet fra løgn ved å *observere* den som fremfører den angivelige løgnen. Jo Hov understreker denne presumsjonen videre med følgende uttalelse i sitt læreverk «Ved umiddelbar forklaring, vil retten lettere kunne se om vitnet virker usikker – om han f.eks. stammer eller om han rødmer.»¹⁷⁵ På dette området er vitnepsykologien samstemt om at det ikke er mulig å avdekke om en person lyver ut fra hvordan et vitne fremstår i retten.¹⁷⁶

Jeg må imidlertid ta bevisumiddelbarhetsprinsippet i forsvar i den forstand at det gir dommeren en mulighet til å stille oppfølgingsspørsmål og få testet fortellingens generelle plausibilitet.

Det man oppfatter som løgn er i stor grad basert på *stereotypien* om løgnaktig atferd. Det vil si kombinasjonen av personlige erfaringer og standardiserte fortellinger om løgn, uten at stereotypen nødvendigvis er representativ for virkeligheten.¹⁷⁷ En stereotypisk oppfatning av hva som i skinnen fremstår som en sannferdig og falsk forklaring, fører til at man blir

¹⁷⁴ (Hjort, 2021) s.35

¹⁷⁵ (Hov, 2019) s.220

¹⁷⁶ (Magnussen & Stridbeck, Vurdering av troverdighet: Hva sier forskningen anno 2020?, 2020) s.6

¹⁷⁷ (Hjort, 2021) s.33

forhindret i å få en mer nyansert forståelse av hva som faktisk har hendt i det unike tilfellet.¹⁷⁸ Man risikerer at den som snakker sant ikke blir trodd, fordi vedkommende ikke passer med stereotypien av et sannferdig menneske.

Bekreftelsestendenser

Bevismiddel i form av vitner, tiltalte, fornærmede og sakkyndiges forklaringer er imidlertid kun én form for bevismiddel. Det kan også utgjøre dokumentbevis som kontoutskrifter, legejournaler, rapporter fra blodprøveanalyser, logger fra telefonmaster ol.¹⁷⁹ De fleste bevisbedømmere har ikke forutsetninger til å overprøve konklusjoner trukket av sakkyndige, eller anslå sannsynligheten for falske positive ved rettsmedisinske bevis. I disse tilfellene blir bevisdataene lagt direkte til grunn som etablerte omstendigheter og utgjør etablerte fakta. Nyere forskning har imidlertid vist at rettsmedisinske tolkninger av blant annet fingeravtrykk og DNA-bevis er sårbare for bekreftelsestendenser.¹⁸⁰ Bekreftelsestendensen tilsier at tvetydige bevis ikke vil bli tolket nøytralt, men til inntekt for påstanden man ønsker å bekrefte.¹⁸¹ Dersom bevisbedømmer ikke er kritisk til sakkyndiges uttalelser risikerer man dermed en dobbelt bekreftelseeffekt. Det vil si at først går de sakkyndige i bekreftelsesfellen, for så at bevisbedømmer tar konklusjonen til inntekt for sin egen påstand hen ønsker å bekrefte. Dette kommer jeg nærmere inn på i kapittel 5.

I tilfeller der dokumentbevis er sms'er, e-poster eller brev, er disse mediene mer åpne for tolkning av avsenders intensjon. Bevisbedømmer vil også ha anledning til å høre med forfatteren av meldingene om hva de betyr. Dette var kjernen i bevisbedømmelsen i Eirik Jensen-saken der vurderingstemaet var hva tekstmeldinger som «Det er trygt å sole seg med rett solfaktor» betydde. Eirik Jensen mener at denne meldingen betød at han hadde fint vær på ferie.¹⁸² Mens Gjermund Cappelen mente at den betød at det var trygt for han å drive med narkotikasmugling så lenge han hadde Jensen med på laget. Jensen var dermed «solfaktoren» til Cappelen i denne analogien.

Tolkningen av slike tekstmeldinger vil også være sårbar for bekreftelsestendenser. Trangen til å tolke meldingene til å passe med den underliggende fortellingen vil innvirke på

¹⁷⁸ (Hjort, 2021) s.267

¹⁷⁹ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.192

¹⁸⁰ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.252

¹⁸¹ (Magnussen & Teigen , 2020) s.27

¹⁸² (Nrk, 2017)

spørsmålsstillingen av fornærmede og vitner. Førsteintrykk som blir skapt i en bestemt kontekst kan føre til at uavhengig informasjon blir satt i sammenheng for å støtte opp under en påstand. Intuitivt vil bevisbedømmer ha problemer med å vurdere informasjon isolert og nøytralt. Man ender i stedet opp med å tolke dem i lys av fortellingen eller stereotypien man ønsker å få bekreftet.

Glorieeffekten

Videre vil *rekkefølgen* bevis blir fremlagt ha noe å si for hvilken fortelling System 1 starter å pusle sammen. Glorieeffekten og bekreftelsestendensen vil danne et førsteinntrykk for et startpunkt for hvilken retning bevisene kommer til å bli tolket i. Starter bevisfremleggelsen i tiltaltes disfavør vil dette raskt forme et førsteinntrykk som er vanskelig å endre. I den forbindelse har sorenskriver Jon Høyland påpekt at det faktisk at hovedforhandlingene starter med at aktor i sitt innledende foredrag leser opp tiltalen, og gir en kort gjennomgang av hvilke bevis som støtter hypotesen om at tiltalte er skyldig, jf. straffeprosessloven §289, prime dommeren for hva hen bemerker seg av betydning i resten av forhandlingene.¹⁸³ Satt opp mot glorieeffekten er dette vanskelig å forene med prinsippet om at tiltalte er uskyldig til det motsatte er bevist, siden prosessen systematisk primer bevisbedømmer til tro på fornærmedes forklaring.

I de mest alvorlige straffesakene der fornærmede har bistandsadvokat, skal fornærmede som hovedregel uttale seg før tiltalte, jf. §289 a).¹⁸⁴ Glorieeffekten vil da få tiltalte til å systematisk å jobbe i motbakke under forhandlingene. Denne regelen er ny av straffeprosessloven fra 2007-2008. Den ble begrunnet med at i tidligere praksis gikk forhandlingene i hovedsak mellom påtalemyndigheten og tiltalte. Fornærmede opplevde dermed å bli redusert til vitne i sin egen sak, hvilket var merbelastende. Det var derfor et ønske om å styrke fornærmedes stilling under hovedforhandlingene, ved å gi dem en større mulighet for kontradiksjon. Forarbeidene drøfter hvorvidt rekkefølgen kan lede til «psykologiske slagsider» uten å gå nærmere inn på hvilke slagsider det er snakk om. De konkluderer med at det er til tiltalte sin fordel å gi sin fremstilling *etter* fornærmede. Dersom det er glorieeffekten de sikter til vil dette være en feilaktig konklusjon, da det vil være til klar ulempe for tiltalte å uttale seg etter fornærmede. Jeg ser imidlertid ikke et problem med å bevare hensynet til fornærmedes ønske

¹⁸³ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.94

¹⁸⁴ Ot.prp.nr.11 (2007-2008) s.68-69

om kontradiksjon og å bytte om på rekkefølgen av uttalelsene. En slik endring vil gjøre at man er i tråd med uskyldspresumsjon, samtidig som man bevarer fornærmedes ønske.

En annen følge av at man vektet informasjonen man får først tyngre, enn den vi får siden, er at man ikke klarer å ignorere eller se vekk fra informasjon som har blitt gitt dersom den viser seg å være feil.¹⁸⁵ Dette er problematisk med tanke på at bevisfremleggelse handler om å bevise og tilbakevise. Beslutningsteori sier er imidlertid at «bordet fanger». Fremstår en påstand først som troverdig, hjelper det lite at påstanden blir tilbakevist som tvilsom og mangelfull i ettertid. Dette har blitt påvist som en generell tendens hos amerikanske dommere, og det er ikke urimelig å anta at det også gjelder for norske.¹⁸⁶

Et eksempel på dette er DNA-beviset i Baneheia-saken. På en pressekonferanse før rettssaken hadde startet uttalte politiet at de hadde funnet biologisk materiale som knyttet Viggo Kristiansen til åstedet.¹⁸⁷ Under rettssaken viste dette angivelige DNA-beviset seg for å være tvetydig. Det var klart at det ikke tilhørte Jan Helge Andersen. Så det tydet på at det kunne vært to gjerningsmenn. Men DNA-profilen var funnet på så sparsomt grunnlag og matchet med 54 prosent av den mannlige befolkningen i Norge. Beviset utelukket dermed ikke at Kristiansen hadde vært på åstedet, men det var langt fra å bekrefte at han hadde vært der. Isolert sett skulle tilbakevisningen i en rasjonell verden nøytralisere DNA-beviset til verken å tale til Kristiansen sin fordel eller ulempe. Men bekreftelsestendensen vil overspile det faktumet at beviset ikke kan *utelukke* at Kristiansen var på åstedet, og lagres dermed som et bevis som taler til ulempe for Kristiansen.

4.5 Faren med forklaringer

4.5.1 Etablerte omstendigheter krever en forklaring

Det siste steget i vurderingskjeden sier at etter at man har fastslått de etablerte omstendighetene, så krever de en *forklaring* på hvordan de henger sammen. På denne måten danner bevissituasjonen en sammenhengende fortelling som retten skal legge til grunn som sakens faktum.

¹⁸⁵ (Kahneman, Sibony, & Sunstein, Noise, 2021) s.173

¹⁸⁶ (Rachlinski, Wistrich, & Chris, 2005)

¹⁸⁷ (Olsbu & Kristensen, 2020)

Her må jeg imidlertid påpeke hva som menes med at enhver omstendighet har en *forklaring*, for at dette skal være i overenstemmelse med beslutningsteori. En forklaring på en etablert omstendighet kan være at omstendigheten ikke er relevant for beviset. Det vil si at satt opp mot beviset man prøver å forklare er omstendigheten *tilfeldig*. Men som jeg har diskutert under meningsskapingsmaskinen er ikke forklaringen at omstendigheten var *tilfeldig* noe som faller bevisbedømmere naturlig. Dette vil jeg utdype i det følgende.

4.5.2 Forståelsesillusjonen

Hjernen er som kjent en meningsskapingsmaskin.¹⁸⁸ Når noe dramatisk har skjedd ønsker man å finne en forklaring på hvorfor, slik at man kan leve i en illusjon om at man kan forhindre at det skjer igjen. Dette kalles *forståelsesillusjonen*. Kjernen av illusjonen er at vi tror at vi forstår fortiden, slik at vi kan forutsi fremtiden. Dersom noe uforutsett skjer, må System 1 sin *normalmodell* oppdateres.¹⁸⁹ Bekreftelsestendensen søker etter årsaksforhold som er konsistente med hva som har hendt, mens forhold som ikke er konsistente overses. Dette er et relativt effektivt system, men fører til en ubeleilig begrensing. Når modellen oppdateres lagres ikke tidligere kunnskapsnivåer eller overbevisninger.¹⁹⁰ Det vil si at med en gang man blir overbevist om noe nytt, *erstattes* den gamle overbevisningen med den nye, og mange tror de aldri har ment noe annet.¹⁹¹ Dette fører til etterpåklokkap- og utfallsskjevheter.

Etterpåklokkap og utfallsskjevhet

Etterpåklokkapsskjevhet er tendensen til å overvurdere forutsigbarheten til en hendelse etter at den har inntruffet.¹⁹² Skjevheten gjør også at bevisbedømmer vil huske mer av hendelsesinformasjonen som er konsistent med forutsigbarheten av det faktiske utfallet, enn informasjon som er inkonsistent.¹⁹³ Utfallsbedrag er at man bedømmer kvaliteten på beslutningen ut fra *resultatet* i stedet for *prosessen*, selv om forhåndssannsynligheten for utfallet var kjent for beslutningstaker.¹⁹⁴ Disse to skjevhetene rammer spesielt hardt observatører som skal vurdere godheten av andres beslutninger. Man har en tendens til å

¹⁸⁸ (Kahneman, Tenke, fort og langsomt, 2013) s.217 (Magnussen & Teigen, 2020) s.30

¹⁸⁹ (Kahneman, Tenke, fort og langsomt, 2013) s.220

¹⁹⁰ (Kahneman, Tenke, fort og langsomt, 2013) s.220

¹⁹¹ (Kahneman, Tenke, fort og langsomt, 2013) s.221

¹⁹² (American Psychological Association, 2015)

¹⁹³ (Harley, 2007) s.48-49

¹⁹⁴ (Kahneman, Tenke, fort og langsomt, 2013) s.221-220

klandre beslutningstakere for gode beslutninger som endte dårlig, og til å gi dem for lite ære for vellykkede trekk som fremstår som åpenbare i ettertid.

Forklaringer kjennetegnes med at de ofte forutsetter mål-middel-sammenhenger.¹⁹⁵

Handlinger forstås som rasjonelle ut fra klare mål og motiver. Hvis en handling fremstår uten klart motiv, ansees den som lite rasjonell og derav lite plausibel. Denne forutsetningen legger imidlertid til grunn at aktøren hadde så god oversikt over situasjonen at vedkommende var i stand til å velge det mest rasjonelle handlingsalternativet. Men i samtiden fremstår verden som regel som rotete og uoversiktlig. Man handler ut fra begrenset informasjon og magefølelse. Handlingene kan dermed fremstå som irrasjonelle når man til slutt får alle kortene på bordet. Hvor plausibel man finner en forklaring vil dermed lide av en etterpåklokskapsskjevheter. Videre havner parter som regel i retten fordi noe uheldig har skjedd. Det vil si *utfallet* av en vurdering har slått skjevt ut. Man vil dermed risikere å klandre aktøren ufortjent hardt fordi vedkommende hadde uflaks, selv om beslutningen i seg selv var god.

Et eksempel hvor denne typer feller kan oppstå er i erstatningssaker der bevisbedømmer må ta stilling til risikoens art og grad i aktsomhetsvurderingen.¹⁹⁶ Bevisbedømmer skal ta stilling til hva risikoen for utfallet var i utgangspunktet og hva en alminnelig person i tilsvarende situasjon er forventet å vite om risikoens størrelse i art og grad. Det vil si informasjon en alminnelig person ville hatt *før* risikoen realiserte seg. Bevisbedømmer vil på sin side ha problemer med å rekonstruere et tidligere kunnskapsnivå, som ikke er farget av informasjonen om at utfallet faktisk realiserte seg. Man greier da ikke å la være å tilskrive utfallet som realiserte seg en større risiko enn det den opprinnelig hadde. Jo verre faktisk konsekvens, desto større er etterpåklokskapen.¹⁹⁷ Det vil si at ved større skadeevne blir aktsomhetsnormen strengere. Etterpåklokskap vil føre til at man vurderer risikoens art og grad som større, enn den egentlig var i samtiden, i kraft av at risikoen ble realisert.

Forståelsesillusjonen fjerner altså elementet av *flaks* som bidro til den etablerte omstendigheten, noe som skadevolder ikke kan klandres for. Illusjonen hindrer

¹⁹⁵ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.161

¹⁹⁶ (Hagstrøm & Stenvik, 2015) s.93

¹⁹⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.222

bevisbedømmer å lande på forklaringer som sier at de etablerte omstendighetene skyldes *tilfeldigheter*.

Empirisk belegg for etterpåklokkenskapsskjevheter

Forskningen er samstemt når det kommer til hvorvidt lekdommere blir påvirket av etterpåklokkenskapsskjevheter.¹⁹⁸ Forskningen på om fagdommere går i den samme fellen er imidlertid mer sparsommelig og ikke like samstemt.¹⁹⁹ En rekke studier viser ingen signifikant forskjell mellom dommere som får vite utfallet, og de som ikke får vite utfallet på forhånd.²⁰⁰ På den andre siden har man flere studier som viser at fagdommere faktisk går i etterpåklokkapsfeller.²⁰¹

Rachlinski m. fler forklarer de tvetydige resultatene med at det kommer an på hvilken vurdering fagdommerne skal foreta.²⁰² I et eksperiment skulle dommerne vurdere en klagesak fra en innsatt, som hevdet at fengselet ikke hadde sørget for tilstrekkelig legehjelp.

Dommerne som viste utfallet av klagesaken falt i etterpåklokkapsfellen. Dommere greide imidlertid å manøvrere seg unna tilsvarende feller når de skulle vurdere om en politimann hadde «skjellig grunn» til å bryte seg inn i en bli for å avdekke narkotikabesittelse.²⁰³

Forskerne påpekte at årsaken til at fagdommerne ikke gikk i etterpåklokkapsfellen i dette tilfellet kunne komme av at fagdommere er opplært i reglene og rettspraksis knyttet til «skjellig grunn». Det blir da sendt et signal til System 2 om at vurderingen krever en mer systematisk vurdering. I motsetning til klagesaken som ikke hadde tilsvarende regelsett, men i større grad lente seg på en skjønnsmessig vurdering.²⁰⁴ Dette kan også forklares med at det er mulig å opparbeide seg en ekspertintuisjon på rettsanvendelse, siden det foreligger en fasit, slik som jeg diskuterte i delkapittel 3.2.6.

¹⁹⁸ (Baron & Hershey, 1988) (Guthrie, Rachlinski, & Wistrich, 2001) (Hastie, Schkade, & Paye, 1999) for oversikt se (Harley, 2007).

¹⁹⁹ (Oeberst & Goeckenjan, 2016) s. 273

²⁰⁰ (Hastie & Viscusi 1998), (Schweizer 2005), (Rachlinski, Wistrich, & Chris, 2005) s.1251 (Rachlinski, Guthrie, & Wistrich, Probable Cause, Probability, and Hindsightjels, 2011)

²⁰¹ (Anderson m.fler, 1997), (Guthrie, Jeffrey, & Wistrich, Blinking on the bench: how judges decide cases, 2007) (Oeberst & Goeckenjan, 2016)

²⁰² (Rachlinski, Guthrie, & Wistrich, Probable Cause, Probability, and Hindsightjels, 2011) s.96 flg. (Guthrie, Jeffrey, & Wistrich, Blinking on the bench: how judges decide cases, 2007) s.25 flg.

²⁰³ (Rachlinski, Wistrich, & Chris, 2005) s.1251

Dette kan tyde på at både fag og lekdommere kan være særlig sårbare for etterpåklokskapsskjevheter i bevisvurderingen, siden den i utgangspunktet er *fri* og dermed ikke regelstyrt som kan aktivere System 2.

4.6 Forklaringene som bevisvurderingsmetode

Hvis man ikke tar aktive grep vil bevisbedømmer intuitivt lage koherente fortellinger ut fra de prosederte bevisene. I stedet for å følge intuisjonen til å *finne* forklaringer, kan man også bruke forklaringer aktivt til å teste fortellingens holdbarhet. Forklaringene kan da aneeses som hypoteser.²⁰⁵ Riksadvokaten har i sitt rundskriv argumentert for at man aktivt skal bruke gjensidige utelukkende hypoteser for å komme frem til den mest robuste fortellingen av bevisituasjonen.²⁰⁶ Ideen er lånt fra empirisk forskning, der man bruker antitesen til hypotesen man ønsker å få bekreftet.

I strafferettslig kontekst vil det si at man må avkrefte at tiltalte er uskyldig for å kunne konkludere med at vedkommende er skyldig. I jakten på forklaringer skal bevisbedømmer først anta at tiltalte er *uskyldig* etter tiltalen, for så å undersøke om det finnes bevis og etablerte omstendigheter som hver for seg eller isolert blir vanskelig å forklare. Videre skal bevisbedømmer anta at tiltalte er skyldig etter hypotesen, og undersøke om det finnes bevis og etablerte omstendigheter som i så fall blir vanskelig å forklare. På denne måten avdekker man etablerte omstendigheter og bevis som kan bekrefte begge hypoteser, og derfor ikke kan tas til inntekt for domfellelse.

En slik tilnæringsmåte tvinger bevisbedømmer til å tenke alternativt, hvilket mildner bekreftelsestendenser. Dersom forklaringene man må ty til for å få omstendighetene til å passe med hypotesen fremstår som anstrengt, vil denne fortellingen fremstå som lite plausibel og ikke legges til grunn.

Hypotesetestingen kan i straffesaker formuleres opp mot beviskravet om at rimelig tvil skal komme tiltalte til gode.²⁰⁷ Dette kravet er oppfylt dersom 1) det finnes en eller flere etablerte omstendigheter (hver for seg eller i kombinasjon) som er vanskelig å forklare hvis tiltalte er uskyldig, samtidig som 2) det ikke finnes etablerte omstendigheter som er vanskelig å forklare

²⁰⁵ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.208 flg.

²⁰⁶ (Riksadvokaten, 2018) s.15

²⁰⁷ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.210

hvis tiltalte er skyldig. Herunder skal «vanskelig å forklare» forstås som at det ikke finnes en plausibel forklaring på omstendigheten.

En slik metode er i tråd med forskning på «massenes visdom» som jeg diskuterte i delkapittel 3.1.2. Mer presist er det i overensstemmelse med teknikken som kalles *dialektisk bootstrapping*.²⁰⁸ Man skaper en *indre masse* ved å først komme med én forklaring på de etablerte omstendighetene, for så aktivt finne en så annerledes forklaring på den etablerte omstendigheten man greier innenfor troverdighetens grenser. Ut fra massenes visdom vil sannheten ligge nærmere snittet av de to anslagene. Gitt at man ikke har en særlig grunn til å holde seg til én av forklaringene. På denne måten greier forklaringsparadigmet å dressere fortellingene til å ta hensyn til *tilfeldigheter* ved å tenke statistisk.

Sett bort fra majoriteten av sivile saker i tingretten, avgjøres faktum under rådslagningen etter hovedforhandlingene mellom flere dommere. Massenes visdom har ti ganger sterkere effekt dersom man greier å samle uavhengige meninger fra en gruppe, fremfor en indre masse.²⁰⁹ Ideelt sett burde man systematisert diskusjonen slik at hver og en dommer foretok en systematisk bevisvurdering ved å dressere fortellingene via forklaringene. De kunne så delt meningene sine i felleskap, og lande på et kompromiss av disse meningene. Dette ville både rettet opp de kognitive skjevhetene som jeg har diskutert over, men også kansellert ut støy i avgjørelsene.

²⁰⁸ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.88

²⁰⁹ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.87

5 Sannsynlighetsparadigmet

5.1 Sannsynlighetsparadigmet er normativt

I forrige kapittel drøftet jeg hvordan fortellingsperspektivet åpner for, men også forsøker å avhjelpe, kognitive feller. Jeg kom frem til at bevisbedømmer i hovedsak bruker representativitetsheuristikker og forståelsesillusjonen for å anslå hvor *plausibel* hen finner påstandene. Representativitetsheuristikken gjør at man blir ufølsom for utvalgs kvalitet og mengde, og at man ignorerer grunnfrekvenser. Forståelsesillusjonen fører til at man blir etterpåklok og lider av utfallsskjevhet. Fortellingsperspektivet prøver å rette opp disse skjevhetene ved å dressere måten man konstruerer fortellinger på, ved hjelp av å skape en *indre masse*.

Fortellingsperspektivet er dermed en deskriptiv teori som adresserer hva bevisbedømmer faktisk *gjør*. Og hva bevisbedømmer intuitivt *gjør* er sårbart for kognitive skjevheter. Forklaringsperspektivet prøver dermed å si at det man *burde* gjøre er å dressere disse skjevhetene med alternative forklaringer. Sannsynlighetsparadigmet har en noe annen tilnæringsmåte. Det er normativt og sier noe om hvordan bevisbedømmer *burde* tenke. Den anviser til den rasjonelle regelen for å bli overbevist på, Bayes Teorem.

5.2 «Sannsynlighet»

5.2.1 Sannsynlighetsbegrepet er subjektivt

Sannsynlighetsparadigmet har en vitenskapelig forståelse av ordet «sannsynlighet». Dette medfører at det er regneregler som representerer den riktige måten bevisbedømmer skal gå frem for å bli overbevist. Hvilket er i overenstemmelse med den dominerende oppfatningen i norsk juridisk litteratur om hvordan bevisbedømmer skal la seg overbevise, som er bedre kjent som Temametoden.²¹⁰ Strandberg påpeker at temametoden sitt sannsynlighetsbegrep er bayesiansk.²¹¹ Et baysiansk sannsynlighetsbegrep kjennetegnes ved at det er *subjektivt* og

²¹⁰ (Strandberg M. , 2012) s.91

²¹¹ (Strandberg M. , 2012) s.131

vurderingen i større eller mindre grad er inspirert av tradisjonelle *sannsynlighetsteoretiske regneregler*.

I likhet med fortellingsteorien sitt plausibilitetsbegrep er sannsynlighetsbegrepet i sannsynlighetsparadigmet subjektivt. Det vil si at det retter seg mot det psykologiske nivået til bevisbedømmer, og sier noe om i hvilken *grad* vedkommende er *overbevist*. Sannsynligheten uttrykker dermed én persons grad av overbevisning for om påstand *x* er sann.²¹² Siden folk er forskjellig vil en *subjektiv* sannsynlighet føre til at sannsynlighetsgraden vil variere fra person til person selv om påstanden og bevissituasjonen er den samme.

Sannsynlighetsgraden uttrykkes på en skala fra 0 til 1, som også kan uttrykkes i prosent. Summen av sannsynlighetsgraden til overbevisningen av påstand *x* og ikke-*x* er da til sammen 100%, hvilket følger av *negasjonsregelen*. Temametoden har et tosidig sannsynlighetsbegrep. Der høy sannsynlighet for *x*, fører til tilsvarende lavere sannsynlighet for ikke-*x*. Dette gjelder uavhengig av om bevissituasjonen er «god» eller «dårlig». Som diskutert i delkapittel 4.3.2. har ikke temametoden noe robusthetskrav.²¹³ Det vil si en terskel for hvor grundig påstand *x* er bevist før vi kan legge den til grunn som rettsfakta. Dette gjør bevisbedømmer i norsk rett særlig sårbar for ufølsomhet for mengden bevis og kvalitet på bevisene, på grunn av representativitetsheuristikken.

Problemet med et subjektivt sannsynlighetsbegrep som kun retter seg mot det psykologiske nivået er at bevisbedømmers overbevisning i seg selv er en dårlig målestokk for om en bevisvurdering er rasjonelt begrunnet.²¹⁴ Begrepet mister sin normative kraft fordi man ikke kan si at en vurdering er rett eller gal, så lenge bevisbedømmer *føler* seg overbevist. For å løse dette problemet er bayes subjektive sannsynlighet formulert som en subjektiv sannsynlighet som er *betinget* av bevissituasjonen.²¹⁵

Bayes sitt sannsynlighetsbegrep graderer dermed informasjon på det epistemiske nivået, samt bevisbedømmer sin overbevisning på det psykologiske nivået.²¹⁶ Det vil si at sannsynligheten gir uttrykk for hvor stor grad bevisbedømmer *bør* la seg overbeviste ut fra den *gitte*

²¹² (Strandberg M. , 2012) s.355

²¹³ (Strandberg M. , 2012) s. 95

²¹⁴ (Strandberg M. , 2012) s.357

²¹⁵ (Strandberg M. , 2012) s.357

²¹⁶ (Spiegelhalter, 2019) s.306

bevissituasjonen. Det kan eksempelvis uttrykkes som hva sannsynligheten er for at liket som er funnet er forvunnede X, gitt bevisene.

I sannsynligetsparadigmet skiller man på rene subjektive sannsynligheter, som vil gi forskjellige sannsynlighetsgrader ut fra hvilken person som vurderer bevissituasjonen. Og på epistemisk subjektive sannsynligheter som gir en *rasjonell fremgangsmåte* for hvordan man skal la seg overbevise av den *gitte* bevissituasjonen.

Spørsmålet er da hva denne rasjonelle fremgangsmåten er. Den rasjonelle fremgangsmåten er sannsynlighetsregneregler for hvordan man skal anslå og kombinere sannsynligheter. Siden denne oppgaven begrenser seg til å diskutere sannsynligheten for om beviskravet er oppfylt, der bevistema er gitt på forhånd vil jeg kun gå i dybden på regnereglene knyttet til Bayes teorem.

5.2.2 Objektive sannsynligheter som rettsfakta, men ikke metode

Objektive sannsynligheter finner man i empirisk vitenskap, som regel i form av statistikk.

Sannsynlighet brukes da som et mål på hvor *ofte* et fenomen oppstår i en populasjon ut fra tilfeldigheter over *tid*. Dette kalles objektive frekvenser. Man kan bruke de objektive frekvensene til å regne ut korrelasjoner og kausale sammenhenger mellom fenomen x og y. Denne typen sannsynlighet regner ut frekvenser over *tid*, og kan dermed ikke brukes til å uttrykke sannsynligheten for en enkelt hendelse.²¹⁷ I retten kan objektive frekvenser brukes som rettsfakta, slik som i Landås-saken (Rt.1999 s.1742).

En slik forståelse av sannsynlighet har imidlertid begrenset potensial som begrep i en bevisvurderingsteori. Siden denne typen sannsynlighet forutsetter at man har et fenomen som er mulig å klassifisere og observere over tid. I tillegg sier sannsynligheten bare noe om *snittet*. Den har begrenset verdi for om sannsynligheten for at det *unike* tilfellet har inntruffet. Den objektive sannsynligheten kan bare brukes dersom bevistema er å konstatere hyppighet av et fenomen. Noe som egner seg til å bevise systemiske fenomener som korrupsjon, forfølgelse og overgrep.²¹⁸ Da utgjør statistikken rettsfakta, men ikke en metode for bevisbedømmelse.²¹⁹

²¹⁷ (Strandberg M. , 2012) s.349

²¹⁸ HRDAGs arbeid med å dokumentere og analysere statistikk til å holde myndigheter ansvarlig for brudd på menneskerettigheter, <https://hrdag.org/projects/>

²¹⁹ (Strandberg M. , 2012) s. 351

Statistikk kan dermed i liten grad si noe om sannsynligheten for at et unikt tilfelle har inntruffet.

5.2.3 Bayes subjektive sannsynligheter og fiktive frekvenser

Dette er imidlertid bevisbedømmers oppgave. Hen er forpliktet til å ta stilling til hva som er sannsynligheten for at påstanden er sann i det unike tilfellet. Siden hvert tilfelle for retten er unikt, herunder at man ikke kan klassifisere eller observere over tid, må man late som. Man later som at bevistemaet var trukket ut fra et tilfeldig utvalg like bevissituasjoner, spør seg selv hvor sannsynlig det er at den er sann *gitt* bevissituasjonen.²²⁰ Ta eksempelet med bitemerket i Torgersen-saken. Bittmerker er ikke unike. Bevisbedømmer anslår at det er tre personer som kunne gitt det bitemerket i Norge. Isolert sett er da sannsynligheten for at bitemerket stammet fra Torgersen 1/3.

Her ser man at *subjektiviteten* i sannsynligheten kommer inn. For hvor henter bevisbedømmer at det er tre personer som kunne gitt det bitemerket fra? Dette problemet er bedre kjent som *referanseklasseproblemet*, som jeg kommer tilbake til i delkapittel 5.4.3.²²¹ Dette anslaget baseres på kunnskapen bevisbedømmer sitter på i vurderingsøyeblikket. Og hvilke avgrensinger av utvalget hen synes er fornuftig. Denne kunnskapen består av bevisene som partene har ført for retten, som vitneforklaringer, statistikk, rettsmedisinske bevis og sakkyndige uttalelser. Men den baserer seg også på *erfaringsgrunnsetninger* og *rimelige antakelser*.

Hvilke erfaringsgrunnsetninger bevisbedømmer har og hva hen synes er rimelige antakelser vil naturlig nok variere ut fra hvem bevisbedømmer er i den aktuelle saken. Det vil si at selv om man betinger overbevisningsgraden til bevissituasjonen på det epistemiske nivået ved å anvende regneregler. Så er det likevel sterke subjektive elementer i å anslå tallene som skal legges til grunn for å anvende regnereglene. Konsekvensen av dette, er at *enhver prosess som inneholder subjektive elementer vil være sårbare for støy og kognitive skjevheter*. Dette gjør at man risikerer forskjellige sannsynlighetsgrader fra forskjellige bevisbedømmere selv om man vurder samme påstand.

²²⁰ (Strandberg M. , 2012) s. 93

²²¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.116 flg.

Hovedformålet med å innføre begrepet om epistemiske subjektive sannsynligheter er å rasjonalisere *fremgangsmåten* for å kunne ta *praktiske* beslutninger.²²² I dette ligger det en erkjennelse om at man ikke får forutberegneligheten til objektive frekvenser, men at det i en vurderingssituasjon uunngåelig er dommeren som konstaterer hvorvidt hen er overbevist om at påstanden er sann eller ikke.²²³ Nedsiden ved at sannsynligheten er subjektiv er at resultatet blir avhengig av hvem som vurderer bevissituasjonen.

For å avhjelpe dette, prøver bayesiansk beviskravslære å formalisere *fremgangsmåten* ved hjelp av regneregler. Tilhengerne av denne metoden hevder at ved å formalisere fremgangsmåten ved hjelp av regneregler, håper man vurderingen blir mer transparent og forutberegnelig. Temametoden sverger til addisjons, multiplikasjons og negasjonsreglen. Når det kommer til samordningen av *bevisfakta* er imidlertid temametoden like opptatt av helhet i bevismaterialet som regnereglene.²²⁴ Et unntak fra denne hovedregelen er at noen anbefaler å anvende Bayes teorem til samordning av bevisdata. Lenge var dette kun anbefalt som en regneregul for samordning av bevistema og bevissituasjoner, men i senere tid har juridiske teoretikere som Erling Eide og Christian Dahlman tatt til orde for å anvende denne regneregelen som en generell bevisvurderingsmetode.²²⁵ Og sett i sammenheng med at det kun er et tidsspørsmål før man kan bruke maskinlæring til saksbehandling i norsk rett har sannsynlighetsparadigme igjen blitt aktualisert innenfor nordisk bevisteoridiskusjon.

5.3 Bayes teorem: En rasjonell måte å endre mening

5.3.1 Bayesiansk analyse

Bayes teorem er den rasjonelle måten bevisbedømmer kan oppdatere sin grad av overbevisning på etter hvert som hen får ny kunnskap om bevistema.²²⁶ Desto mer informasjon man får til å mate inn i teoremet, desto mer robust blir sannsynligheten. Det vil si

²²² (Strandberg M. , 2012) s. 355

²²³ (Strandberg M. , 2012) s. 356

²²⁴ (Strandberg M. , 2012) s.96

²²⁵ (Strandberg M. , 2012) s.96, (Eide , 2016) og (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018)

²²⁶ (Spiegelhalter, 2019) s.307

at jo mer informasjon man får, jo mindre blir sjansen for at overbevisningsgraden endrer seg vesentlig. Teoremet gir altså en rasjonell fremgangsmåte for å endre mening.²²⁷

Teoremet er en formalisering av bayesiansk analyse. Jeg vil dermed først vurdere eksempelet ved hjelp bayesiansk analyse for å gi en intuitiv forståelse av teoremet, før jeg viser hvordan man regner ut sannsynligheten ved hjelp av Bayes teorem.

Gitt scenarioet at man er dommer i en rettsak. Der saken handler om en person som har blitt påkjørt av en bil på natten. Et vitne mener å ha sett en rød bil forlate åstedet. Bevistemaet som skal vurderes er hvorvidt fargen på gjerningsbilen faktisk var rød.

Spørsmålet er da hva som er den rasjonelle fremgangsmåten for å vurdere sannsynligheten for at gjerningsbilen var rød. Hvis man husker tilbake til delkapittel 4.3.1 om Tom W. er første steg i en bayesiansk analyse anslå ut fra *tilfeldigheter* hva sannsynligheten er for at en bil som kjører på norske veier er rød. Man kan eksempelvis bruke et røft overslag, og si at er 15 av 100 biler er røde i Norge. Grunnfrekvensen for at bilen er rød er dermed 15%. Man starter analysen statistisk, ved å se til hvor ofte røde biler generelt forekommer i populasjonen. Konklusjonen er at det er større sannsynlighet for at bilen *ikke* er rød enn for at den er rød.

Men i dette tilfellet har man ytterligere informasjon, man har et vitne som hevder å ha sett at bilen var rød. For å vurdere om denne informasjonen skal tale for eller imot sannsynligheten for at bilen er rød, må man vurdere *troverdigheten* til beviset. Det vil si at selv om vitnet sier at bilen er rød, kan vitnet lyve, ha sett feil eller huske feil. Man må altså vurdere om vitnet har større grunn til å ha rett enn å ta feil. Er det grunn til å tro på vitnet, har beviset stor *beviskraft*. Dette vil trekke opp sannsynligheten på for at bilen faktisk var rød.

Beviskraften til de forskjellige bevisene vil altså trekke opp eller ned sannsynligheten fra grunnfrekvensen for at påstanden er sann, avhengig av hvor sterk beviskraften er for de enkelte bevisene.

Formelt sett sier da Bayes teorem at den endelige sannsynligheten for om bevistema er sann eller usann, er *grunnfrekvensen* multiplisert med den samlede *beviskraften* til de forskjellige bevisene, hvilket utgjør bevissituasjonen sin *bevisverdi*. Den *samlede bevisverdien* er

²²⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.168

sannsynligheten for at bevistemaet inntreffer gitt at bevisene stemmer. I dette tilfelle *sannsynligheten for at bilen er rød gitt at det stemmer at vitnet sier at den er rød*.

Bayes teorem kan regnes ut ved hjelp av to versjoner av det samme teoremet, oddsversjonen eller sannsynlighetsversjonen. For å lettere få oversikt over hvilke elementer som inngår i anslaget og oppdateringen av sannsynligheten, velger jeg å anvende oddsversjonen av teoremet, hvilket er som følger:²²⁸

$$\text{Grunnfrekvens} \times \text{Samlet beviskraft} = \text{Bevisverdi}$$

$$\frac{P(B)}{P(\bar{B})} \times \frac{P(E|B)}{P(E|\bar{B})} = \frac{P(B|E)}{P(\bar{B}|E)}$$

5.3.2 Hvordan regne ut grunnfrekvensen i Bayes teorem

En odds er sannsynlighetsraten for at en hendelse vil inntreffe, mens en rate er et forholdstall mellom to størrelser. Oddsene regnes ut ved å dividere sannsynligheten for at hendelsen inntreffer $P(B)$, på sannsynligheten på at hendelsen ikke inntreffer, $P(\bar{B})$. «Hendelsen» er det man kjenner som *bevistema* i rettslig kontekst. I eksempelet er bevistema at bilen *faktisk* er rød.

$$\frac{P(B)}{P(\bar{B})} = \frac{\text{Sannsynligheten for at bilen er rød}}{\text{Sannsynligheten for at bilen ikke er rød}} = \text{Initiell odds for at bevistema inntreffer}$$

Siden 15 av 100 biler på veien var røde, er 85 ikke-røde. Sannsynligheten for rød bil er da 15%, mens ikke-rød bil er 85%. Den initielle oddsene for at bilen *er* rød er da:

$$\frac{0,15}{0,85} = 0.1764$$

Siden man får 15 ca. seks ganger i 85, er tolkingen av at det initielt er 6 ganger mer sannsynlig at bilen *ikke* er rød fremfor at den er rød. Den initielle oddsene sier altså noe om hvordan verden *er* på det aletiske nivået. Denne oddsene kalles *grunnfrekvensen* for hvor hyppig et fenomen oppstår ut fra tilfeldigheter i en populasjon. Et utgangspunkt for å regne ut den initielle oddsene kan være å bruke empirisk data, eksempelvis tall fra SSB over hvor mange biler det er av de forskjellige typene i Norge.

²²⁸ (Spiegelhalter, 2019) s.315 flg.

5.3.3 Hvordan anslår man beviskraften

Men formålet med bevisvurderingen er tross alt å avdekke sannsynligheten for at bevistemaet har skjedd på det aletiske nivået *gitt* bevisfakta på det epistemiske nivået. Det er dette man regner ut ved hjelp av Bayes teorem, sannsynligheten for bevistemaet (B) gitt at bevisfakta (E) stemmer, $P(B|E)$. For å anslå beviskraften må man altså snu på flisa og regne ut sannsynligheten for at bevisfaktaene stemmer *gitt* at bevistemaet faktisk har inntruffet, $P(E|B)$.

Beviskrafts-leddet sammenlikner sannsynligheten for to konkurrerende hypoteser.²²⁹ I rettslig kontekst vil det være sannsynligheten for at bevistemaet har inntruffet, bilen er rød, mot sannsynligheten for at bevistemaet ikke har inntruffet, bilen er ikke rød. Sannsynligheten for at vitnet sier at bilen er rød gitt at bilen er rød kalles en *sann positiv*. Sannsynligheten for at vitnet sier at bilen er rød gitt at den ikke er rød kalles en *falsk positiv*.

Man regner dermed ut beviskraften med følgende formel:

$$\frac{P(E|B)}{P(\bar{E}|B)} = \frac{\text{Sannsynligheteten for at vitnet sier at bilen er rød gitt at bilen er rød}}{\text{Sannsynligheten for at vitnet sier at bilen er rød gitt at bilen ikke er rød}} = \text{Beviskraften}$$

Det vil si at jo vanligere det er med et positivt resultat på testen for bevistema som faktisk foreligger, desto større er sannsynligheten for et sant positivt utfall.²³⁰ Jo vanligere det er at testen gir et positivt resultat når bevistema ikke forekommer, jo høyere er sannsynligheten for et falskt positivt svar. Jo større forholdet mellom sannsynligheten for sanne og falske positiver er, jo sterkere er beviskraften. Et bevis med sterk beviskraft har høy sannsynlighet for sann positiv, og lav sannsynlighet for falsk positiv.

Eksempelvis viser tester retten har gjort at vitnet identifiserer de røde bilene riktig 80% av tiden, men tar feil 20% av gangene. Det vil si at sannsynligheten for at vitnet sier at bilen er rød *gitt* at bilen er rød 80%, er den sanne positive. Mens sannsynligheten for at vitnet sier at bilen er rød *gitt* at den ikke er rød er 20%, den falske positive.

$$\frac{0,8}{0,2} = 4$$

²²⁹ (Spiegelhalter, 2019) s.315

²³⁰ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.93

Tolkningen av beviskraften er at vitnets utsagn om at bilen er rød er 4 ganger mer sannsynlig dersom bilen faktisk er rød, sammenliknet med at den ikke er rød.

Beviskraften sier altså noe om hvor relevant et bevisfaktum er, og hvor stor vekt man skal tillegge det i bevisvurderingen. Dersom beviskraften er 1, er beviset verdiløst.²³¹ Ut fra eksempelet ovenfor betyr 1 at det er like stor sannsynlighet for at bilen var rød som ikke rød, gitt at vitnet sa at den var rød. I eksempelet kom jeg frem til at beviskraften til vitneutsagnet var 4, hvilket er større enn 1. Det betyr at beviskraften drar bevisverdien opp. En beviskraft mindre enn 1 trekker ned bevisverdien.

5.3.4 Samlet beviskraft

I bevisvurderingssituasjon er det ofte flere bevisfakta som støtter eller taler imot bevistema. Bayes teorem tillater også sammenstilling av flere bevisfakta. Det vil si at man kan legge til flere beviskraftledd, i formelen, slik at den ser slik ut:

$$\text{Beviskraft}_1 \times \text{Beviskraft}_2 \times (\dots) \times \text{Beviskraft}_n = \text{Samlet beviskraft}$$

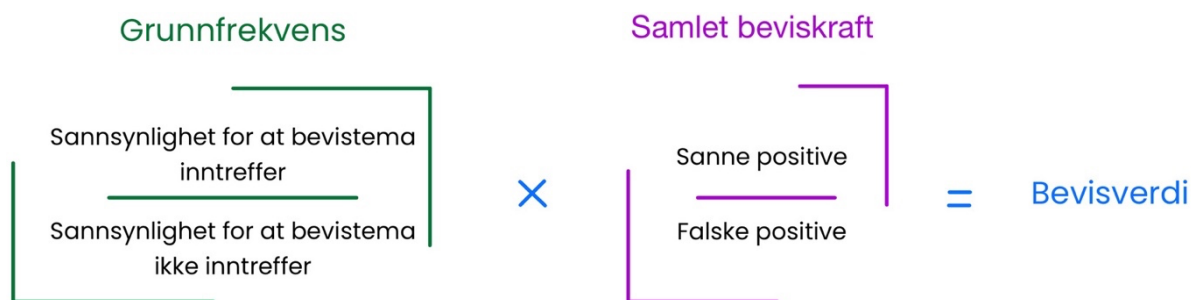
Det vil si at selv om hvert bevisfaktum isolert sett har for lav beviskraft, kan de oppfylle beviskravet samlet ved å multiplisere dem sammen.²³² Mange bevisfakta som peker i samme retning kan veie opp for fravær av sterke enkelte bevis. Når det er sagt krever multiplikasjon av beviskraftsleddene at bevisfaktaene er *uavhengige*. Hva som skjer når man multipliserer avhengige sannsynligheter kommer jeg tilbake til.

5.3.5 Bevisverdien

Bevisverdien er da lik grunnfrekvensen multiplisert med den samlede beviskraften, se Figur 3. Grunnfrekvensen er som sagt forholdstallet mellom sannsynligheten for at bevistema forekommer delt på sannsynligheten for at det ikke forekommer. Beviskraften er et forholdstall mellom sannsynligheten for sanne og falske positive.

²³¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.70

²³² (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.176



Figur 3

Det vil si at bevisverdien kan være lav selv om sannsynligheten for sanne positive er høy. Dette kan komme av at grunnfrekvensen er lav eller sannsynligheten for falske positive er høy.

Bevisverdien for at bilen er rød blir dermed $0.1764 \times 4 = 0.7059$. Beviskravet i Norsk rett er formulert som en sannsynlighet og ikke en odds. For å vurdere om beviskravet er oppfylt må man da konvertere oddsen til sannsynlighet. Dette gjøres ved hjelp av formelen

$Sannsynlighet = \frac{odds}{odds+1}$. Det vil si at sannsynligheten for at bilen er rød *gitt* at vitnet sier at den er rød er $\frac{0.7059}{0.7059+1} = 41\%$.

Til nå har jeg gjennomgått hva sannsynlighetsparadigmet mener med *sannsynlighet* og hvordan bevisbedømmer rent rasjonelt skal gå frem for å oppdatere overbevisningsgraden sin. I det følgende vil jeg drøfte hvorvidt denne metoden er rent praktisk gjennomførbar med hensyn til om den åpner for eller avhjelper kognitive feller. Dette vil jeg gjøre ved å diskutere hvert ledd i teoremet hver for seg.

5.4 Grunnfrekvensen

5.4.1 Hvorfor er det viktig å ta hensyn til grunnfrekvensen?

Regresjon mot middelerdien

Bevisvurderingen handler i sin kjerne om å skille på hvilke bevisfata som skyldes *tilfeldigheter*, det vil si fenomener som ikke har en meningsfylt *årsakskjede* til beviset, fra bevisfakta som faktisk forklarer beviset. For å greie å gjøre dette må man tenke statistisk,

ikke kausalt. For å bruke Kahnemans ord igjen «Kausalt finnes ikke støy noe steder, statistisk sett finnes den overalt.».²³³

Rasjonelt sett er grunnfrekvensen viktig fordi den tar hensyn til fenomenet som kalles *regresjon mot middelverdien*. Regresjon mot middelverdien er det rent matematiske fenomenet at tilfeldigheter gjør at hendelser svinger rundt gjennomsnittshyppigheten for en hendelsene.

Regresjon er en annen måte å forklare korrelasjon på. Korrelasjon er hvor mye to fenomener samvarierer. Det vil si beveger seg i takt. Dersom korrelasjonen er 1 beveger fenomen x og y seg perfekt i takt. Er den 0 er det ingen sammenheng i bevegelsene. Dersom det er ufullkommen korrelasjon, det vil si korrelasjon mindre enn 1, vil det alltid forekomme regresjon mot middelverdien. Dersom korrelasjonen er 0,3, betyr det at vi kan forklare 30% av bevegelsene mellom x og y, men 70% er opp til tilfeldigheter. 70% av variasjonen i datasettet vil dermed dras mot gjennomsnittshyppigheten. Denne variasjonen har imidlertid *ingen* forklaringskraft, slik de øvrige 30% har, det at 70% varierer rundt gjennomsnittsverdien er et rent matematisk fenomen.

Eksempelvis si at man statistisk skal bevise i en rettsak at et selskap systematisk har diskriminert på lønn mellom menn og kvinner. Man ser på forklaringsvariablene for hvorfor mennene i selskapet tjener mer enn kvinnene. X er da lønn, og Y er forklaringsvariabelen kjønn. Dersom forskjellen i ansattes lønn og kjønn samvarierer med 30%, kan det tyde på at 30% av lønnsforskjellen mellom menn og kvinners lønn skyldes at hun er kvinne. Men her må man også ha i mente at man undersøker en samvariasjon, ikke en kausalitet. 70% av lønnsforskjellen kan på sin side ikke forklares ut fra kjønn, og skyldes da *tilfeldigheter*.

Eller se tilbake til eksempelet med Tom W. Si at personbeskrivelsen stemmer. Det vil si at den har en viss prediktiv kraft for hvilken studieretning Tom W. har valgt å gå på. Men tiden har gått siden han var hos psykologen og han har antakelig endret seg en del i løpet av bachelorgraden. Si han fikk seg er kjæreste som fikk han til å åpne seg mer, og fant sin store lidenskap for kunst. Personskissen vil dermed ha noe forklaringskraft for hvilket studium Tom valgte, men langt i fra alt. Det vil si at det ikke er en 100% korrelasjon mellom personbeskrivelsen og valg av studiet, men noe, si da 30%. Den resterende forklaringskraften

²³³ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.220

på 70% kan ikke forklares av personskissen, men av andre ting som påvirkning fra kjæresten, livsendrende hendelser og andre *tilfeldigheter*.

Det vil si at dersom man ikke tar hensyn til grunnfrekvensen i bevisvurderingen, risikerer man å ikke ta tilstrekkelig høyde for at hendelser hovedsakelig inntreffer som følge av tilfeldigheter, som i Landås-saken. Dersom man ikke tar hensyn til grunnfrekvensen, vil man intuitivt *kun* se på beviskraften, hvilket vil tilegne beviset 100% forklaringskraft for bevistema, i stede for sine respektive 30%.²³⁴

5.4.2 Neglisjering av grunnfrekvens

Som jeg diskuterte i kapittel 4.3.2 leder representativitetsheuristikken til at man ignorerer grunnfrekvensen. Dette kommer av at en grunnfrekvens er et generelt faktum som ikke kan hektes på en stereotypi.²³⁵ Det er informasjon som kun *graderer* sannsynligheten for at et fenomen opptrer, men som ikke kan forklare *årsaken*. Si i rød bil-eksempelet at vitnet var en gammel litt tutlete mann med dårlig syn. Intuitivt danner man seg en stereotypi om et upålitelig vitne, som senker beviskraften til vitneuttalelsen. Det er dermed lett å ta hensyn til denne informasjonen når man skal gradere hvor *sannsynlig* det er at bilen er rød.

På den andre siden er et rent statistisk faktum om at 15 av 100 biler er røde, intuitivt vanskelig å bruke til å videre gradere sannsynligheten. Som nevnt kommer dette av at statistikken ikke sier noe om årsaken til utfallet. Den gamle tutlete mannen sier noe om årsaken til hvorfor man skal stole mindre på beviset, mens statistikken bare er en observasjon. Hjernen sliter dermed med å ta like mye hensyn til statistikken som det konkrete beviset i sannsynlighetsvurderingen. For å kunne ta hensyn til denne type informasjon må man ha anlegg for å tenke statistisk.

Det må også påpekes at det er i de færreste bevissituasjoner at grunnfrekvensen til bevistema eksplisitt blir nevnt. Som regel vil kun konkrete bevis som støtter påstanden bli presentert for retten. Det vil da være opp til dommeren selv å komme på at hen må ta hensyn til grunnfrekvensen. Men som jeg har diskutert i 4.2.5 tenker meningsskapingsmaskinen kausalt. Man utleder en generell årsakssammenheng fra en enkelt observasjon. Dersom man blir presentert for et vitne som sier at bilen er rød, vil man utlede fra denne observasjonen at bilen

²³⁴ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.387 flg.

²³⁵ (Kahneman, Tenke, fort og langsomt, 2013) s.182

var rød. Det krever statistisk tenkning å anslå grunnfrekvensen, og vekte den mot vitneuttalelsen. Man risikerer dermed at grunnfrekvensen systematisk blir neglisjert i bevisvurderingen, hvilket Landåssaken illustrerer.

5.4.3 Referanseklasseproblemet

Bayes teorem sier at det rasjonelle utgangspunktet for å starte et sannsynlighetsanslag for en hendelse er å regne ut grunnfrekvensen. For å regne ut grunnfrekvensen må man ha en statistikk på hvor hyppig fenomenet oppstår i *snitt*. Grunnfrekvenser utgjør dermed et objektivt ankerpunkt for et ellers subjektive sannsynlighetsanslag i bevisvurderingen. Men problemet oppstår når man ikke empirisk har data for et bevistema, noe som utgjør hovedregelen snarere enn unntaket. Da må bevisbedømmer foreta et subjektivt anslag av frekvensen til bevistemaet, som fører til det man i bevisteorien kaller for *referanseklasseproblemet*.

Et anslag av frekvens krever at man går motsatt vei av vitenskapelig metode. I stede for å foreta en *statistisk* vurdering, fra det generelle datasettet, til den spesielle variabelen, må man foreta en *kausal* vurdering. Man må abstrahere fra det unike tilfellet, til det generelle utvalget. Dette er særlig vanskelig fordi en person, gjenstand eller hendelse har tusenvis av karakteristikk å abstrahere fra.²³⁶ Man får dermed et problem med hvilke karakteristikk man skal selektere på, og hvor stort utvalget er når man først har valgt karakteristikk. Det blir da opp til den enkelte bevisbedømmer å vurdere.

Frekvensen kan variere dramatisk ut fra hvilke karakteristikk man velger. Eksempelvis i en sak mellom to studenter på Norges Handelshøyskole krevde den ene studenten en erstatning på 36 millioner kroner etter å ha blitt bitt i øye på et fuktig nachspiel.²³⁷ 15,9 millioner av denne erstatningssummen representerte hva studenten potensielt kunne ha tjent dersom han *ikke* hadde blitt bitt i øyet. Spørsmålet er da hvilken referanseklasse det er rimelig å ta utgangspunkt i når man skal estimere fremtidig tapt inntekt. Dersom man tar utgangspunkt i gjennomsnittlig inntekt i Norges befolkning er det klart at 15,9 millioner blir for høyt. Tilsvarende voldssaker har oppnådd en erstatning på 13 000kr. Men studenten argumenterte på sin side for at han som student ved Norges Handelshøyskole var en aspirerende næringslivstopp. Dette var dermed den naturlige referanseklassen å ta utgangspunkt i. Saken

²³⁶ (Cheng, 2009) s.2985

²³⁷ (Andersen, 2009)

illustrerer at frekvensen på hva som ansees som forventet fremtidig inntekt blir dramatisk forskjellig alt etter som man selekterer på gjennomsnittsinntekten hos Norges befolkning, eller snevrer det ned til flinke studenter ved Norges Handelshøyskole. Så hvordan skal man løse referanseklasseproblemet i retten?

For å gå nærmere inn på det spørsmålet må det først skilles mellom statistisk og juridisk grunnfrekvens. Dette kommer av at retten må ta hensyn til forskjellige prinsipper som ikke gjør seg gjeldene i vanlig forskning.²³⁸ Et slik prinsipp er uskyldspresumsjonen. I rettslig sammenheng defineres derfor grunnfrekvensen, ifølge Ekelöf m.fl., som «sannsynligheten for at gjerningsbeskrivelsen er sann på tidspunktet der påtalemyndigheten innledningsvis av hovedforhandlingen har lest gjerningsbeskrivelsen, men enda ikke har presentert noen bevis til støtte for tiltalen.»²³⁹ På dette tidspunktet vet bevisbedømmer naturlig nok at det foreligger en sannsynlighet som taler imot tiltalte, siden det først er tatt ut tiltale mot hen. Aktor tar ut tiltale mot de mistenkte der bevismaterialet etter aktors mening er sterkt nok til å få vedkommende domfelt.²⁴⁰ Det foreligger dermed en seleksjonsskjevhet i hvem som blir tiltalt. Dersom bevisbedømmer har erfart at 90% av alle tiltalte er skyldig, og tar inn denne statistikken i grunnfrekvensen, blir beviskraften i bevismaterialet regnet dobbelt. Man bryter dermed uskyldspresumsjonen som sier at tiltalte skal ansees som uskyldig til det motsatte er *bevist*. Uskyldspresumsjonen sier da at man må betrakte tiltalte som en *tilfeldig valgt person på gaten*.²⁴¹

I rettslig sammenheng er det to situasjoner *usikkerheten* i sannsynlighetsvurderingen er knyttet til.²⁴² Man må skille på om usikkerheten er knyttet til hvorvidt en *handling* har skjedd, eller *identifikasjon* av hvem som er gjerningsmann. Et eksempel på det første scenario er spørsmål om Peder Ås har unndratt skatt. Påtalemyndigheten må bevise at det har forekommet skatteunndragelse. Dersom *handlingene* kan bevises, er det Peder Ås som har foretatt dem. I det andre tilfellet er det gitt at hendelsen har skjedd, man skal bevise *hvem* som har utført handlingen. Eksempelvis om det var Peder Ås som drepte Marte Kirkerud. Disse to situasjonene krever to forskjellige måter å bestemme grunnfrekvensen på. Det må imidlertid påpekes at dette er en forenkling, det finnes også tilfeller der det både er usikkerhet knyttet til

²³⁸ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.115

²³⁹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.113

²⁴⁰ (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s.289

²⁴¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.116

²⁴² (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s.290

handling og identifikasjon.²⁴³ Dersom en pasient er død, er det usikkerhet knyttet til om den er forårsaket av en uaktsomhandling og hvem som eventuelt har handlet uaktsomt. Men for diskusjonens skyld holder jeg det enkelt med denne todelingen. I den første situasjonen anvender man prinsippet om bevistemaet sin alminnelighet, mens det andre kalles 1/N-prinsippet.

Prinsippet om bevistemaets alminnelighet

Prinsippet om bevistemaets alminnelighet sier at grunnfrekvensen for at tiltalte har handlet slik aktor påstår i gjerningsbeskrivelsen tilsvarer hvor ofte fenomenet som følger av gjerningsbeskrivelsen forekommer i populasjonen for øvrig.²⁴⁴ Si at det følger av gjerningsbeskrivelsen at Marte Kirkerud ble tatt for å urinere på offentlig sted 17. mai, kl 03:00, på Tryvann, i forbindelse med et russetreff. Man skal da sammenlikne denne konkrete situasjonen med et utvalg «tilsvarende» situasjoner. Men hvordan avgrense et utvalg av «tilsvarende» situasjoner på en ikke vilkårlig måte? Teoretikere som Ronald Allen og Michael Pardo mener at det ikke er mulig. Christian Dahlman og Eivind Kolflaath mener at det er mulig i prinsippet, men ikke i praksis.

James Franklin mener at referanseklassen skal avgrenses ut fra alle relevante karakteristikk som påvirker handlingen beskrevet i gjerningsbeskrivelsen.²⁴⁵ Utvalget må da konsentrere seg om handlingen «urinering på offentlig sted». Likevel bør den spesifiseres ytterligere? Det var tross alt natt til 17. mai i forbindelse med et russetreff. Er dette relevante karakteristikk som bør regnes inn? I første rekke støter man på to problemer:

For det første vil karakteristikkene øke grunnfrekvensen for at en tilfeldig valgt person urinerer på offentlig sted. Ved å spesifisere grunnfrekvensen ytterligere etter gjerningsbeskrivelsen kommer man i konflikt med uskyldspresumsjonen. Det er klart at sannsynligheten for at en beruset russ på et treff natt til 17.mai, vil urinere på offentlig sted er høyere, enn for at en tilfeldig person på gaten urinerer på offentlig sted. Grunnfrekvensen vil i seg selv indikere skyld, uavhengig av bevisføringen, og dermed bryte presumsjonen om at man er uskyldig til det motsatte er *bevist*. Dette problemet kalles *Naked statistical evidence*.²⁴⁶ Uskyldspresumsjonen legger dermed ytterligere en begrensning på den juridiske

²⁴³ (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s.290

²⁴⁴ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.117

²⁴⁵ (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s. 291

²⁴⁶ (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s.292

grunnfrekvensen. I tillegg til å sammenlikne med en tilfeldig valgt *person*, må det også være i en tilfeldig valgt *sammenheng*.²⁴⁷

For det andre vil det i praksis bli vilkårlig hva bevisbedømmer anser som nødvendige betingelser for handlingen beskrevet i gjerningsbeskrivelsen, og ikke minst vite hva som var relevante faktorer i den konkrete saken.²⁴⁸

Men si at man levde i en ideell verden der bevisbedømmer greier å avskjære utvalget til å være i tråd med uskyldspresumsjonen. Da kunne bevisbedømmer gått ut i verden og ført statistikk over hvor vanlig det er å urinere på offentlig sted ved å sende ut et spørreskjema til et representativt utvalg av befolkningen. Bevisdømmer kunne også ha gjort et estimat på hvor mange man antar urinerer på offentlig sted ut fra statistikken til politiet over hvor mange som faktisk får bøter for urinerer på offentlig sted.²⁴⁹ Dette er imidlertid ressurskrevende, som resulterer i at de fleste saker ikke har empiriske data til å regne ut grunnfrekvensen. Dette gjør at i praksis blir grunnfrekvensene anslått subjektivt av der resultatet vil variere ut fra hvem som foretar avgrensningen. Dette gjør vurderingen sårbar for kognitive skjevheter. Man risikerer da at like tilfeller ikke blir behandlet likt.

Ettersom denne avgrensningen er subjektiv, vil karakteristikken og størrelsen på utvalget man anslår være et resultat av representativitetsheuristikkens *stereotypier*. Hvilket kan føre til systematisk diskriminering av enkelte grupper som narkomane og minoritetsgrupper, som storsamfunnet har fordommer mot.

1/N-prinsippet

I situasjoner der det er klart at en gjerning har funnet sted, men man er usikker på *hvem* som har gjennomført handlingen estimeres sannsynligheten ved hjelp av 1/N-prinsippet. I det lukkede roms mysterier er avgrensning av utvalget ikke noe problem. I spørsmålet om hvem som har drept Dr. Mørk på Tudor Gods, er det bare seks mulige drapsmenn, siden det ikke var flere som oppholdt seg på godset under drapstidspunktet.²⁵⁰ Grunnfrekvensen er da $1/6 = 0,167$ sannsynlig for at tiltalte Miss Scarlett drepte Dr. Mørk.

²⁴⁷ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s. 119

²⁴⁸ (Dahlman & Kolflaath , The Problem of the Prior in Criminal Trials, 2021) s. 292

²⁴⁹ (NTB, 2019)

²⁵⁰ Eksempel hentet fra brettspillet Cluedo

I de fleste saker foreligger ikke de samme betingelsene som i det lukkede rom. Da oppstår problemet med hvordan man skal avgrense kretsen av mulige gjerningspersoner. Hvordan man avgrenser kretsen vil påvirke hvor sterk den konkrete bevisføringen må være for å tilfredsstillende beviskravet.²⁵¹ Dersom man inkluderer alt for få blir grunnfrekvensen for lav, og bevisførselen må overkompensere for å oppfylle beviskravet, som gjør at skyldige går fri. Blir for få mulige gjerningsmenn inkludert, blir grunnfrekvensen blir for høy, uskyldspresumsjonen blir da brutt og uskyldige blir dømt, og vice versa. Det har blitt gjort flere forsøk på å oppstille retningslinjer for grunnfrekvensen som tilfredsstiller uskyldspresumsjonen, Norman Fenton sitt forslag er det mest diskuterte.

Fenton foreslår at man skal begrense utvalget til personer som hadde lik mulighet til å begå handlingen som tiltalte, satt opp mot deres nærhet til åstedet.²⁵² Problemet med denne tolkningen er at det strengt tatt ikke er snakk om en grunnfrekvens lengre, da man må *bevise* tiltaltes nærhet til åstedet på gjerningsøyeblikket. I tillegg fungerer kun denne løsningen dersom tiltalte sist var observert i moderat avstand og tidsrom til åstedet. Eksempelvis i Baneheia-saken svarte Viggo Kristiansen på en tekstmelding som ble registrert av en telefonmast 19:37.²⁵³ Masten hadde svak dekning 150m og god dekning 400 meter unna åstedet. Retten la videre til grunn at drapene skjedde en gang mellom kl. 19:00 og 20:00. For det første må man gå langt inn i den faktiske bevisvurderingen for å fastslå hvor tiltalte sist ble observert. For det andre er det stor forskjell på hvor mange som befinner seg innenfor en 150m og 400 m radius av et åsted. Legger man til grunn 400m er det langt flere mulige gjerningsmenn, enn 150m. Hvilket vil få et stort utslag på hvor lav eller høy grunnfrekvensen blir. Igjen blir anslaget vilkårlig, fordi man ikke har tall på hvor mange som befant seg innen 400m eller 150m radius, og det blir opp til en subjektiv vurdering av dommeren.

Etablert litteraturs løsninger på referanseklasseproblemet

Dahlman og Eide foreslår å standardisere grunnfrekvensen for å unngå vilkårlighet i det subjektive anslaget.²⁵⁴ Eide foreslår at man skal nøytralisere grunnfrekvensen ved å sette den til 50%. Dette påpeker Dahlman vil bryte uskyldspresumsjonen, siden den sier at man er uskyldig til den motsatte er bevist. Er man én av to potensielle gjerningsmenn er man verken

²⁵¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s. 125

²⁵² (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.124

²⁵³ (Sommerset, Isungset, & Skramstad, 2020)

²⁵⁴ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.127-128 (Eide , 2016) s.107-108

skyldig eller *uskyldig*. I tillegg bryter man også med prinsippet om at man skal betraktes som en tilfeldig person på gaten. Siden det neppe er én av to som ville begått handlingen i gjerningsbeskrivelsen. Han foreslår dermed at tiltalte alltid skal betraktes som å ha blitt plukket ut fra et utvalg på enten 100, 1000 eller 10 000 personer. Hvor stort utvalget skal være er et rettspolitisk spørsmål, og må avgjøres ut fra lovgivers preferanser. Poenget er at grunnfrekvensen nøytraliseres for vilkårlighet, ved at den er lik i alle saker.

I en nyere artikkel av Dahlman og Kolflaath foreslår de å løse hele referanseklasseproblemet ved å se helt vekk fra grunnfrekvensen i bevisbedømmelsen.²⁵⁵ De begrunner dette med at dersom man har en slik regel som Dahlman foreslår om en standardisert grunnfrekvens, og et beviskrav som setter en terskel for hvor høy den endelige oddsen må være for at påstanden skal legges til grunn som rettsfaktum, så utgjør disse to reglene et normativt rammeverk for bevisvurderingen. Standarden for bevis kan da blir uttrykket som kombinasjonen av samlet beviskraft for rettsakens bevisfakta, som må til for å oppfylle beviskravet. Hvor sterk beviskraften må være er da avhengig av hvor høyt eller lavt man har satt den standardiserte grunnfrekvensen.

Hva vil så være motivasjonen til lovgiver være for å sette en høy eller lav grunnfrekvens? Dersom formålet med dette normative rammeverket er maksimere forventet nytte i form av å minimere antall feilaktige dom- og frifinnelser, mener Dahlman og Kolflaath at man like gjerne kan se helt vekk fra å prøve å sette en optimal kombinasjon av størrelse på grunnfrekvens og terskel på beviskravet. Dette kommer av at man tilpasser rammeverket for å oppnå den optimale beviskraften. I realiteten blir da det normative rammeverket en omvei, man kan heller regulere den optimale beviskraften direkte. Beviskravet blir dermed omdefinert til hva som ansees som den tilstrekkelige beviskraften.

Problemet med forlaget om å se helt vekk fra grunnfrekvensen slik jeg ser det er at man løser det filosofiske problemet med referanseklasser, ved å fjerne den praktiske løsningen på neglisjering av grunnfrekvens. Man rykker dermed tilbake til start, ved at man kun fokuserer på forklaringsvariabelen med 30% forklaringskraft og glemmer at 70% av variasjonen skyldes tilfeldigheter (regresserer rundt middelveidien). I delkapittel 5.6 vil jeg diskutere mer

²⁵⁵ (Dahlman & Kolflaath, The Problem of the Prior in Criminal Trials, 2021) s. 298

inngående hva jeg mener om hensiktsmessigheten av sannsynlighetsparadigmets detaljorienterte tilnærming til å konstruere en perfekt rasjonell fremgangsmåte.

5.5 Beviskraften

5.5.1 Hva er formålet med et beviskraftsledd

Det andre leddet i Bayes Teorem er likelihoodraten, nærmere kjent som beviskraftsleddet i bevisvurderingsteori.²⁵⁶ Beviskraft defineres som sannsynligheten for at et bevisfaktum har forårsaket et bevistema. Der grunnfrekvensen sier noe om hvor ofte bevistemaet vil oppstå ut fra tilfeldigheter, sier beviskraftsleddet noe om det er grunn til å tro at det foreligger en høyere eller lavere sannsynlighet for bevistemaet i det konkrete tilfellet.

5.5.2 Beviskraftsleddet flytter fokus til bevisets mengde og kvalitet

I delkapittel 4.3.3 diskuterte jeg at mengden og kvaliteten på bevisene *burde* svekke det indre signalet for hvorvidt en fortelling er plausibel eller ikke. Representativitetsheuristikken gjør imidlertid at man *kun* vurderer er histories troverdighet ut fra om den er *sammenhengende*, men ignorerer *mengden* og *kvaliteten* på informasjonen som har blitt presentert.²⁵⁷ Beviskraftsleddet tvinger imidlertid bevisbedømmer å aktivt ta stilling til *kvaliteten* på beviset ved å vurdere hvor pålitelig kilden er ved å anslå den sanne og falske positive. Videre blir bevisbedømmer tvunget til å vekte disse bevisfaktaene mot hverandre og gradere overbevisningen sin. På samme måte som grunnfrekvensleddet er beviskraftsleddet med på å rette opp kognitive skjevheter.

Leddets stiller imidlertid ingen minimumskrav til *mengden* bevisfakta som taler for eller mot et bevistema. WYSIATI-effekten vil føre til at man ikke stiller noen kritiske spørsmål til om bevismengden gjør bevistemaet godt nok opplyst. Denne robusthetsproblematikken står dermed fremdeles uløs ved bruk av Bayes teorem.

²⁵⁶ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.67

²⁵⁷ (Kahneman, Tenke, fort og langsomt, 2013) s.96

5.5.3 Feilslutninger som følge av utregningen av beviskraftsleddet

Intuitiv kausal tenkning

I denne oppgaven har jeg fastslått at bevisbedømmer intuitiv tenker kausalt, og det faller hen svært lite naturlig å tenke statistisk. Så selv om Bayes teorem er et verktøy for å rette opp denne skjevheten, er det så vanskelig for hjernen å omstille seg til den motsatte tankegangen at det er en del fallgruver på veien til å anvende teoremet riktig. I det følgende vil jeg gjøre rede for de vanligste feilslutningene bevisbedømmer begår når hen skal prøve å bruke Bayes teorem som er knyttet til beviskraften.

Forveksling av avhengige og uavhengige sannsynligheter

I delkapittel 5.3.4. poengterte jeg at dersom man skal multiplisere de enkelte beviskraftsleddene for å få den samlede beviskraften må hvert enkelt ledd være uavhengig av hverandre. Det at bevisbedømmer blander avhengige med uavhengige hendelser er imidlertid et noe utbredt problem i tilfeller der man har prøvd å anvende likelihoodrater i retten. Dette kan illustreres med den amerikanske saken *People vs. Collins*.²⁵⁸ Saken dreide seg om en kvinne som hadde blitt overfalt av to personer. Vitner hadde beskrevet gjerningspersonene som en hvit kvinne i mørke klær og en afrikansk utseende mann med skjegg og bart. I etterforskningen ble det funnet to personer som passet til denne gjerningsbeskrivelsen. Aktor regnet så ut sjansen for at et par tilfeldigvis skulle passet med denne beskrivelsen.

Regnestykket hen la til grunn var:

- Én av ti kvinner har hestehale.
- Én av tre kvinner er blonde.
- Én av ti menn med afrikansk etnisitet har skjegg.
- Én av fire menn med afrikansk etnisitet har bart.
- Én av 1000 par er mellom hvit kvinne og svart mann.
- Én av ti biler er gule.

Aktor multipliserte så sammen disse sannsynlighetene, havnet på $1/12\ 000\ 000$, og konkluderte med at det var en for liten sannsynlighet for at paret kunne være noen andre enn de tiltalte. Man ser tydelig i dette eksempelet at aktor ikke har tatt høyde for at flere av disse sannsynlighetene er avhengige. Eksempelvis har de med skjegg større sannsynlighet for å ha

²⁵⁸ (Eide, 2016) s.143

bart. Det var på den tiden også mer populært med gule biler i den svarte populasjonen enn for den øvrige. I ankesaken ble dermed denne utregningen justert ned til 1/1000, bevisverdien var da ikke like sterk lengre og paret ble frikjent. Saken viser at dersom man ikke tar høyde for avhengighet mellom bevisfakta i vurderingen av den samlede beviskraften, blir sannsynligheten grovt overvurdert.²⁵⁹

Men selv om man er klar over forskjellen på sannsynlighetsreglene for avhengige og uavhengige bevisfakta er det i praksis vanskelig å avgjøre om bevisfakta faktisk er uavhengig. I rettslig sammenheng er det flere omstendigheter som avgjør om bevisfakta er avhengige. Det mest åpenbare tilfellet er når vitner har snakket sammen.²⁶⁰ Minnene om hendelsen blir påvirket av hva den andre sier. Forklaringene vil da bli avhengige. Blir forklaringene helt like, mister den ytterligere forklaringen beviskraft, og de to vitnenes forklaringen skal ansees som ett bevisfaktum. Dersom forklaringene delvis blir påvirket av hverandre skal utregningen av beviskraften til den andre forklaringen vekte inn påvirkningen fra den første. Her oppstår to rent praktiske problemer. Det første er at man må inn i en bevisvurdering for om vitnene har påvirket hverandre. Det andre er at selv om man avdekker at de har snakket sammen, får man videre et praktisk problem med å tallfeste hvor stor påvirkningen har vært. Dette vil i stor grad ende opp som en subjektiv vurdering, som er sårbar for kognitive skjevheter.

Bevisfakta kan også være avhengig selv om det ikke foreligger noen sammensvergelse. Det kan også foreligge en *hendelse* rundt bevissituasjonen, som gjør bevisfaktaene avhengig. Det første bevisfaktumet øker sannsynligheten for hendelsen, og hendelsen øker sannsynligheter for et falskt positivt i det andre bevisfaktumet.²⁶¹ Eksempelvis er det funnet to fingeravtrykk fra tiltalte på en revolver, disse to fingeravtrykkene er ikke uavhengige bevisfakta. Funnet av det første fingeravtrykket økte sannsynligheten for hendelsen at tiltalte har holdt i revolveren, som igjen øker sannsynligheten for at man finner flere fingeravtrykk på revolveren.

På samme måte er bevisfaktaene i Collins-saken avhengig. Det at vitnet har observert at gjerningspersonen har skjegg, øker sannsynligheten for hendelsen at gjerningsperson har ansiktshår, som igjen øker sannsynligheten for at gjerningspersonen har bart. Det samme kan

²⁵⁹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018)

²⁶⁰ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.187

²⁶¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.193

gjelde for vitneforklaringer. Det første vitnet sier at bilen var blå, det var imidlertid dårlig lysforhold slik at det var vanskelig å skille blå, grønn og svart. Dette øker sannsynligheten for at det andre vitnet også sier at bilen er blå, selv om bilen faktisk var grønn.

Avhengigheten for denne typen bevisfakta er mye svakere enn for sammensvergelses. Man får da igjen et problem med å vurdere hvor stor avhengighet det er snakk om. Er bevisene *sterkt* avhengig, eller *svakt* avhengig. Dahlman foreslår at man skal behandle bevisfakta som er svakt avhengig som uavhengig, for å unngå en for kompleks bevisvurdering. Men siden System 1 sitt problem nettopp er å håndtere grader av sannsynlighet, vil ikke dette forslaget rette opp skjevheten. Det å gjøre subjektive anslag av hva som er sterkt og svak avhengighet vil i seg selv være en anstrengende øvelse, som lett havner i en enten eller vurdering.

Inverse feilslutninger: Aktor og forsvarers illusjoner

I beviskraftsleddet skal man anslå de sanne og falske positive til bevisdatumet, det vil si påliteligheten til beviset. De inverse feilslutningene går ut på at man tror at «sannsynligheten for et positivt testresultat, gitt uskyld» er lik som «sannsynligheten for at tiltalte er uskyldig gitt positivt testresultat».²⁶² Man blander altså sammen hva man anslår av sanne og falske positive i beviskraftsleddet, med hva man bruker Bayes teorem til å regne ut.

Sannsynligheten for et positivt testresultat gir kun uttrykk for presisjonen til testen. I Schweizer sin studie fra delkapittel 4.3.2 var presisjonen til en alkoholtest 95%. Det innebærer at den tar feil 5% av tiden. Det vil si at dersom politiet stopper 100 *påvirkede* personer på veien, vil 95 av dem teste positivt og 5 av dem teste negativt, selv om *alle* er påvirket. Men denne statistikken er kun gyldig når man *vet* at de man stopper er påvirket. Statistikken på 95% gir dermed uttrykk for beviset sin sanne positive.

Men problemet til bevisbedømmer er jo nettopp det, at hen ikke vet om tiltalte var påvirket eller upåvirket. For å regne ut sannsynligheten for at bilisten faktisk var påvirket, når det eneste beviset man har er en positiv test, er å bruke Bayes teorem. Som vi vet fra dette kapittelet, er det ikke nok med kun sannsynligheten for den sanne positive. Man må også ha sannsynligheten for den falske positive og grunnfrekvensen. I Schewizer sin studie var grunnfrekvensen for at en tilfeldig valgt bilist var påvirket 1 av 100. Sannsynligheten for den falske positive, det vil si statistikken over hvor presis testen er når man kun måler upåvirkede

²⁶² (Spiegelhalter, 2019) s.216

bilister, var 5%. Dette kan virke noe forvirrende da det er fort gjort å tro at sannsynligheten for den sanne og falske positive er avhengig av hverandre, $100-95=5$, men det er de ikke, den kunne like gjerne vært 3%.

Sannsynligheten for å være påvirket gitt en positiv test var da: $1/100 * 0,95 / 0,05 = 0,19 \rightarrow$ så må man konvertere odds til sannsynlighet $0,19 / (0,19 + 1) = 16\%$. I eksperimentet var det kun 9% av dommerne som svarte dette, majoriteten svarte 76 – 100%. Majoriteten svarte altså i sjiktet til den sanne positive som ble oppgitt på 95%. I tillegg til å ignorere grunnfrekvensen tyder det også på at de begikk en invers feilslutning.

Inverse feilslutninger er en veldig sterk illusjon som partene i en sak aktivt kan benytte seg av for å manipulere bevisbedømmer under prosedyren. Aktor kan bruke denne illusjonen til å argumentere for tiltaltes skyld. Logikken er da at siden sannsynligheten for at beviset skal matche noen andre enn tiltalte er så liten, er det usannsynlig at det er noen andre en tiltalte som er skyldig. Dette kalles aktors feilslutning.²⁶³ Si man har et blodspor som viser en blodtype kun 1 av 1000 personer har. Tiltalte har også denne blodtypen. Aktor prosederer da på at det er kun 0,001% sannsynlighet for at tiltalte er uskyldig. Dette er en så lav sannsynlighet, at det indikerer skyld. Hvis bevisbedømmer biter på dette argumentet har man begått aktors feilslutning.

Det er lett å begå feilslutningen fordi aktor unnlater å si noe om utvalget man trekker fra. Si den kriminelle handlingen har skjedd i en storby med 1 000 000 innbyggere. Dersom man ser blodsporet isolert sett vil det i denne storbyen være 1000 mulige gjerningsmenn. Sannsynligheten for at tiltalte er skyldig er da 1 av 1000, mens uskyldig er 999 av 1000. Ved å ikke se på utvalget man plukker den «tilfeldige» personen fra, ignorerer man det faktum at selv om en sannsynlighet er lav kan den gjelde mange personer. Et DNA-bevis med 1 til 1 milliarders sjanse for at en tilfeldig person matcher, vil ha 7 matcher på verdensbasis. Man blander altså sannsynligheten for positivt testresultat gitt skyld, med sannsynligheten for skyld gitt blodtype.

Det kalles forsvarers feilslutning når forsvarer snur på flisa og sier at sett blodsporet isolert vil 999 av 1000 med en slik blodtype være uskyldig i denne saken.²⁶⁴ Men da avleder forsvarer bevisbedømmerens oppmerksomhet vekk fra det faktum at mistanken rettet mot tiltalte neppe

²⁶³ (Corklin, 2019) s.74

²⁶⁴ (Corklin, 2019) s.74

kun er basert på blodsporet.²⁶⁵ Blodsporet er bare ett av mange bevis som indikerer skyld i saken. Hvilke øker sannsynligheten for skyld.

Utbredelsen av disse feilslutningene er relativt stor og etablert i saker som anvender statistiske bevis.²⁶⁶ Et velkjent eksempel på at forsvarer prøver å komme unna med forsvarers feilslutning var et ekspertvitne kalt inn av forsvarerne til O. J. Simpson.²⁶⁷ Simpson var tiltalt for å ha drept sin eks-kone Nicole Brown. Flere år tidligere hadde Simpson blitt dømt for vold mot Brown. Ekspertvitnet prøvde å bagatellisere dette faktumet med å si at det kun er 1 av 1000 kvinner som utsettes for vold av sine ektemenn som ender opp drept av dem, altså sannsynlighet for drap av ektefelle gitt vold fra ektemann. Problemet var at Brown allerede var drept. Man måtte da se på utvalget av drepte kvinner ikke voldsutsatte kvinner. Sannsynligheten for at dødsårsaken til en drept kvinne var vold fra sin ektemann var på sin side nærmere 80 av 100.

5.5.4 Skjevheter knyttet til anslaget av beviskraften

Hvem anslår beviskraften?

Til nå har jeg drøftet hvor lett det er å trå feil når man skal regne ut den enkelte og samlede beviskraften. Dette fordi hjernen er rigget til å tenke kausalt fremfor statistisk. Som igjen skaper illusjoner det er vanskelig å ignorere. Hittil har jeg imidlertid operert med eksempler der anslagene man skal regne ut fra er gitt på forhånd. Men hva skjer i selve anslaget av tallene man skal basere utregningen på. Er denne prosessen også sårbar for kognitive skjevheter? For å diskutere dette mer inngående må man skille på hvem som anslår disse tallene. På den ene siden har man eksperter og rettsmedisinere. Dommeren har i liten grad kompetanse til å overprøve deres uttalelser, og må i stor grad legge til grunn deres vurderinger direkte. På den andre siden har man dommeren selv som må anslå beviskraften til bevis som blir prosedert direkte for retten, slik som vitneuttalelser, sms'er, eposter og andre dokumentbevis.

Hvor sårbar bevisbedømmers vurdering er for kognitive skjevheter vil gå langs to akser. Den ene aksen går langs hvor stor grad bevisbedømmer må basere sine antakelser på *subjektive*

²⁶⁵ (Eide, 2016) s.82

²⁶⁶ (Eide, 2016) s.61 (Corklin, 2019) s.74-75

²⁶⁷ (Eide, 2016) s.141

meninger, kontra *objektive* datasett. Mens den andre aksen er går langs aksene i hvor stor grad man er i stand til å vurdere hvert enkelt bevisdatum isolert sett, eller om vurderingen blir påvirket av kunnskap om øvrige bevisdata.

En rettsmedisinere vil i stor grad kunne isolere seg til å vurdere hvert enkelt bevisdata, og basere anslaget på objektive datasett. Mens en dommer i stor grad er overlatt til å basere anslagene på subjektive meninger og har ikke anledning til å vurdere bevisdataene isolert. De fleste bevistema som dommeren skal vurdere har ikke noen datasett å lene seg på, og tiltalen blir lest opp før noen bevis blir presentert. Rettsmedisineren er dermed i mindre grad sårbar for kognitive skjevheter enn dommeren, men som jeg vil drøfte i det følgende er dette en fare for begge parter.

Rettsvitenskapelig bekreftelsesskjevheter

Rettsmedisinere er som sagt en klar kilde til å levere tall for å vurdere sannsynligheten for bevismiddelets pålitelighet i rettssaker. De uttaler seg blant annet om hvor konkluderende analyser av fingeravtrykk, våpen og DNA-bevis er. Bevisbedømmer har i liten grad kompetanse til å overprøve ekspertenes uttalelser, og må i stor grad legge til grunn deres vurderinger direkte. I skinnet virker tallene rettsmedisinere kommer med objektive og vitenskapelige, men spørsmålet er om også disse anslagene er sårbare for kognitive skjevheter?

Nyere forskning innen beslutningsteori har avdekket det som kalles den *rettsvitenskapelige bekreftelsestendensen*.²⁶⁸ Bekreftelsestendensen går ut på at dersom rettsmedisineren får det som kalles skjevhetsfremkallende informasjon, vil det påvirke konklusjonen til rettsmedisineren i favør av informasjonen. Det kan eksempelvis være en rettsmedisinere som skal tolke om et fingeravtrykk som er funnet på åstedet matcher et referanseavtrykk.²⁶⁹ Dersom rettsmedisineren har fått informasjon om at «vitner har observert vedkommende på stedet» eller «etterforskerne tror det er vedkommende», tolkes avtrykket i favør av informasjonen. Kontekstinformasjonen endrer hva granskeren ser etter i analysen av bevismiddelet, hen ser etter tegn som kan bekrefte kontekstinformasjonen og overser

²⁶⁸ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.252

²⁶⁹ (Kahneman, Sibony, & Sunstein, Støy, 2021) s. 251-252

motstridene tegn. Denne tendensen har blitt dokumentert i blodsprutanalyser, skjelettrest, rettspatologi og analyser av komplekse DNA-blandinger.

Bekreftelsestendensen kan også utvikle seg til en kaskade.²⁷⁰ Dersom det er et eller flere kontrolledd for analysen, og konklusjonen fra den første analysen ikke blir holdt skjult for det neste, vil kontrolleddet bli påvirket i sin tolkning. Hen vil da ha en tendens til å slutte seg til den første konklusjonen. Det samme fenomenet kan oppstå på tvers av fagdisipliner.²⁷¹ Eksempelvis dersom en kriminalteknisk odontolog vet at DNA-analysen har identifisert mistenkte, tenderer tolkningen av bitemerket til å matche med mistenkte.

Dette understreker viktigheten av god beslutningshygiene under etterforskningen. Itiel Dror som står bak de mest sentrale funnene i denne forskningen fremhever at *mer informasjon, ikke alltid er bedre*.²⁷² Spesielt ikke dersom den informasjonen kan lede til en prematur intuisjon. Rettsmedisinere og granskere bør dermed få tilgang til informasjon i det Dror kaller *lineær sekvensiell avsløring*. Et annet råd han kommer med er at granskere bør dokumentere vurderingene sine i hvert trinn. Det vil si at en fingeravtrykkgransker bør skrive ned analysen av fingeravtrykket fra åstedet, før hen starter å matche det mot referanseavtrykk og får vite kontekstinformasjon. Dette gjør at dersom hen endrer mening etter at kontekstinformasjon blir kjent, må dette begrunnes og dokumenteres. Man blir da sin egen kontrollør.

Dommerens anslag av beviskraften

Dahlmans metode

En rettsmedisinere som har god beslutninghygiene og baserer seg på pålitelige objektive datasett, har privilegiet med å foreta analysen via statistisk metode. Hen slutter fra det generelle til det spesielle om bevisdatumet sin beviskraft. Men dersom man ikke kan gå ut å observere verden slik den er, som å telle antall ganger en alkoholtest måler riktig og feil promille, må man forestille seg hvor hyppig testen vil treffe og bomme på prediksjonen. Christian Dahlman bruker drapet på Olof Palme for å eksemplifisere hvordan dette skal gjøres.²⁷³ Olof Palme, som var daværende statsminister i Sverige, ble skutt og drept på åpen gate da han var på vei hjem fra kino med sin kone Lisbeth Palme 28.februar 1986. Lisbeth

²⁷⁰ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.253

²⁷¹ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.254

²⁷² (Kahneman, Sibony, & Sunstein, Støy, 2021) s.258

²⁷³ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.72 flg.

Palmes vitnemål ble brukt som et viktig bevis i tingretten for å få dømt Christer Petterson. Han ble imidlertid frikjent i hovrätten på bakgrunn av manglende bevis. For å vurdere bevismakraften til Lisbeth Palmes vitnemål skal man ifølge Dahlman først anslå den sanne positive ved å anta at man har 100 tilfeller der en person blir drept på akkurat samme måte som Olof Palme, og at et vitne observerer morderen på akkurat samme måte som Lisbeth Palme. Det er så en videokonferanse tre år senere, slik det var i Plame-saken, der gjerningsmannen er med i oppstillingen. Den sanne positive er da hvor mange ganger Lisbeth greier å peke ut den skyldige i oppstillingen. Greier vitnet det 80 av 100 ganger, er den sanne positive 80%.

For å anslå den falske positive må man så vurdere hva som har påvirket vitnet i utpekingen av den angivelige gjerningsmannen. I denne saken snudde hun seg mot gjerningsmannen rett etter at skuddene ble avfyrt, hun forsto imidlertid ikke da at det var gjerningsmannen, og hun kunne ikke erindre ansiktstrekkene til vedkommende. Det at personen hun så var gjerningsmannen skjønnte hun først i senere tid. Ansiktet husket hun først da hun fikk se fantombilder av den angivelige gjerningsmannen. Det viste seg i ettertid at fantombildet var basert på en person som var irrelevant for saken, og altså ikke Christer Petterson. Minnet til Lisbeth Palme var med andre ord blitt kontaminert av fantombildet. Spørsmålet er da hvor mye dette har påvirket presisjonen til Lisbeth Palme for å peke ut riktig gjerningsmann.

Videre må man da forestille seg 100 mord der vitnemålet avgis på akkurat samme måte som i Palmesaken. I tillegg til at i videokonferansen tre år senere er det en mistenkt personen som likner på fantombildene, men der denne personen er uskyldig. Den falske positive er da hvor mange ganger vitnet peker ut den mistenkte som likner på fantombildet som gjerningsmann. Siden minnet hennes ifølge Dahlman er kontaminert av fantomtegningen vil hun risikere å peke ut den uskyldige mistenkte i 5 tilfeller. I de resterende 95 tilfellene vil hun peke ut den personen hun faktisk så.

Subjektivt anslag av bevismakraften

Man ser ut fra Dahlman sitt tankeeksperiment at det legges opp til en vurdering som ikke baserer seg på objektive data, men subjektive antakelser. Så vidt jeg har greid å lese meg frem til, er det sparsomt med forskning på konkrete kognitive skjevheter som påvirker det subjektive anslaget av sanne og falske positive. Jeg må da lene meg på en kvalifisert

spekulasjon ut fra drøftelsen i kapittel 4 om hva bevisbedømmer vil tendere å gjøre i slike subjektive vurderinger av sannsynligheter.

Spørsmålet «hva er sannsynligheten for at Lisbeth Palme peker ut Christer Petterson som gjerningsmann gitt at han faktisk var gjerningsmannen» og «gitt at han ikke var gjerningsmannen» er svært vanskelige spørsmål. Det er nærliggende å anta at man intuitivt *erstatte* dette spørsmålet med «hva er sannsynligheten for at Lisbeth Palme snakker sant, og ikke sant».

For å vurdere hva som påvirker oppfattelsen av troverdighet, er man tilbake i drøftelsen fra kapittel 4 om representativitetsheuristikker og forståelsesillusjonen (etterpåkløkskapsskjevheden og utfallsskjevheden). Hvis man setter seg inn i bevisbedømmers sko vil denne vurderingen skje under en rettsak, der Christer Petterson er tiltalt. Forståelsesillusjonen vil dermed gjøre bevisbedømmer forutinntatt for å lete etter forklaringer som tilsier at Petterson er skyldig. Hen vil dermed være tilbøyelig til å tro på Lisbeth Palme, og dermed overestimere den sanne positive.

Videre støter man på et nytt referanseklasseproblem, for hvilke karakteristika skal bevisbedømmer gå ut fra har påvirket minnet til Lisbeth Palme. Dersom Lisbeth Palme passer med stereotypien eller den underliggende fortellingen om en som snakker sant, vil bekræftelsestendensen gjøre at man husker de delene av bevisførselen som støtter opp under hennes troverdighet, og vil legge mindre merke til informasjon som tilsier det motsatte. WYSIATI-effekten vil da kun ta hensyn til den informasjonen som blir lagt merke til. Hvorvidt man blir overbevist av denne, kommer an på om den svekker eller styrker det *indre signalet*. Signalet vil i hovedsak bli styrket eller svekket ut fra om informasjonen gjør den underliggende fortellingen mer sammenhengende eller ikke.

Måten Dahlman forklarer hvordan Lisbeths vitnemål har blitt påvirket virker overbevisende fordi den har en sterk indre logikk, men det leder oppmerksomheten vekk fra mengden og kvaliteten på bevisene som taler for denne fortellingen. Det er nok svært mange andre *tilfeldige* faktorer som påvirker Lisbeth Palmes evne til å huske hvem hun så den dagen, men som ikke passer inn i en koherent fortelling. Intuisjonen vil dermed levere et anslag som er for ekstremt fordi de ikke tar hensyn til tilfeldigheter, og man vil ha en tilbøyelighet til å feste for stor lit til dem.

Numerisk anslag av overbevisning og prospektteorien

Dahlman går ikke så langt som å hevde at man skal konkludere med et spesifikt tall for sannsynligheten for sanne og falske positive, i tilfeller der man ikke har objektive kilder. Men mener heller at man skal anslå forholdstallet mellom dem som en rettesnor for å finne hvilken *størrelsesorden* beviskraften ligger innenfor.²⁷⁴ Det vil si at dersom man ikke har tilgang på informasjon som gjør det forsvarlig å gjøre et anslag av de konkrete falske og sanne positive sannsynlighetene, kan man gå rett på å vurdere forholdet mellom dem. Man skal da stille seg spørsmålet om det er ti, hundre, tusen eller hundretusen ganger mer sannsynlig for sant positivt fremfor falskt positivt. Man skal altså først anslå det numeriske tallet for sannsynligheten, for så regne ut sin overbevisningsgrad. Ut fra det jeg har drøftet over mener jeg at man i realiteten ender med å gjøre det motsatte. Man oversetter sin subjektive overbevisningsgrad til et numeriske tall. Men hvis jeg tar Dahlman på ordet, er bevisbedømmer i stand til å anslå sannsynligheter på denne måten, uten å falle i en kognitiv felle?

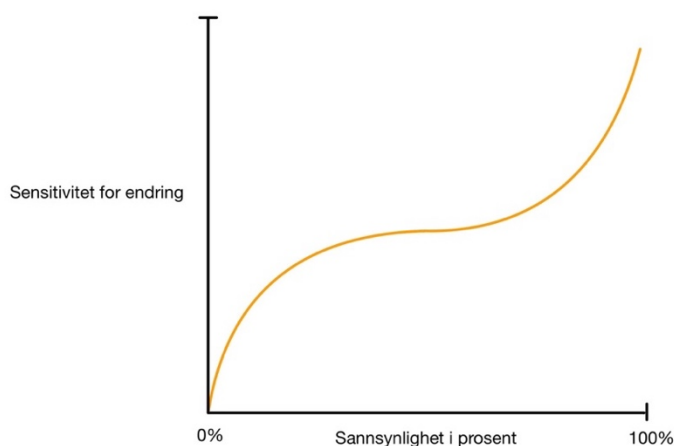
Den mest anerkjente beslutningsteorien kalles *prospektteori*.²⁷⁵ Den undersøker hvordan man vektet forskjellige sannsynligheter. Kort oppsummert sier den at individer overvurderer hendelser med lav sannsynlighet, mens man undervurderer hendelser med høy sannsynlighet. Man har med andre ord ikke en lineær forståelse av vektning av sannsynligheter. Teorien sier at når man skal vurdere hvor sannsynlig en hendelse er, har man to referansepunkter 0 og 100%, og man har avtakende sensitivitet for endring i sannsynlighet fra disse referansepunktene. Det vil si fra 0 til 1%, og 100% til 99% sannsynlighet har man høy sensitivitet, mens fra 1% til 100% har man lav sensitivitet.

Eksempelvis la oss si at man skal vurdere sannsynligheten av at en bilkollisjon inntreffer. Når man vurderer 0% sjans for kollisjon, har man et klart mentalt bilde av en uskadet bil. Men når man skal vurdere 100% sannsynlighet for en kollisjon, har man et mentalt bilde av en totalvraket bil. Disse to mentale bildene har et sterkt inntrykk knyttet til seg, og det er dette sterke inntrykket som gjør at man har høy sensitivitet for endring i sannsynligheten. På den andre siden har man et vagere bilde av hvordan en 10% kontra 20% krasjet bil ser ut, og man

²⁷⁴ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s. 76

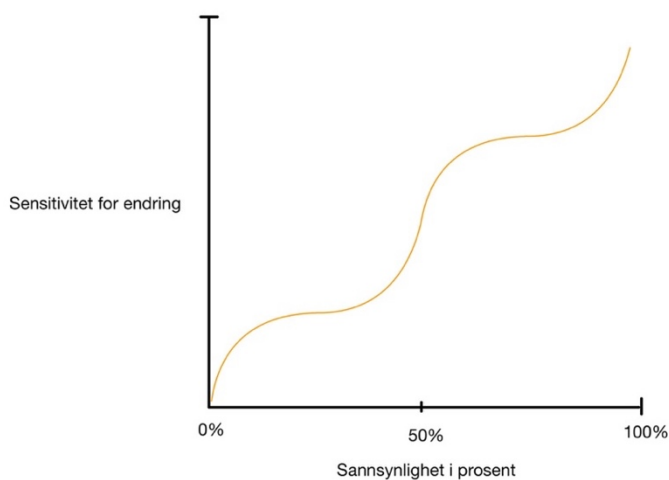
²⁷⁵ (Shley, Chan, Gunadi, & Ferecatu, 2021) s. 2

er dermed dårligere på å estimere forskjellen på disse sannsynlighetene, og har derav lavere sensitivitet. Man får dermed en sensitivitets kurve som er S-formet, illustrert i figur 4.



Figur 4

I tilfellet som er beskrevet får man intuitivt et grep om å diskriminere mellom tre kategorier av sannsynlighet; hvorvidt hendelsen «ikke har skjedd», «en mulighet for at hendelsen har skjedd» og «hendelsen har skjedd med sikkerhet». Nyere forskning har på sin side vist at sensitiviteten kan økes ved å legge inn flere referansepunkter.²⁷⁶ Dette kan illustreres med prediksjon av hvor mye vann som er i et glass. I utgangspunktet har glasset to referansepunkter; full eller tom. Man anslår da svært presist når sylindere er tom og når den er full, men relativt upresist anslagene imellom. Setter man imidlertid et merke på halvfullt, 50%, blir estimatene mellom 25%-75% mer presise, se figur 5.

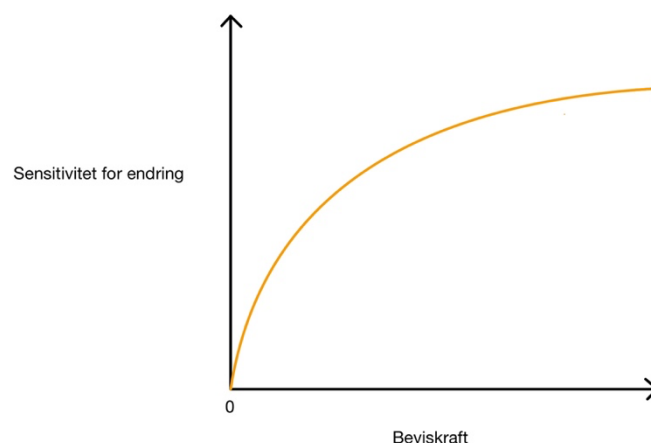


Figur 5

²⁷⁶ (Shley, Chan, Gunadi, & Ferecatu, 2021) s. 4

Når dette er sagt, er det enkelt å sette et merke på et fysisk objekt slik at det er enklere å estimere eksakt hvor halvfullt glasset er. Alle som skal estimere mengden vann i glasset ser det samme glasset, og har dermed det samme referansepunktet. Det blir en helt annen variasjon i estimatet når disse referansepunktene settes subjektivt.²⁷⁷ Hver enkelt bevisbedømmer vil over tid utvikle sine egne referansepunkter over hva som føles 10, 20 og 30% sannsynlig. Problemet med bevisvurderingen er at den ikke har noen fasit man kan trene sensitiviteten sin mot. Det vil vi at selv om beviskraftsanslaget over tid kanskje blir konsekvent for den enkelte bevisbedømmer, vil det være stor variasjon mellom bevisbedømmere, og dermed tilfeldig om noen treffer det «rette» anslaget.

Verre blir det når Dahlman foreslår at man ikke skal anslå presise sannsynligheter, for den sanne og falske positive, men heller se på forholdet mellom dem.²⁷⁸ Det vil si hvor mye mer sannsynlig det er at beviset stemmer, sammenliknet med at det ikke stemmer. Dersom man skal anslå beviskraften direkte, vil man slite med at man har kun ett referansepunkt; 0%. Sannsynlighetssensitiviteten blir dermed ikke S-formet, men avtakende, se Figur 6. Hvis man så skal stille seg spørsmålet om det er ti, hundre, tusen eller hundre tusen ganger mer sannsynlig for sant positivt fremfor falskt positivt, vil man med en avtakende sensitivitetsfunksjon oppleve intervallet ti til hundre tilsvarende som hundre til tusen. Dette til tross for at det andre intervallet er ti ganger så stort. Mangel på intuitiv sensitivitet i anslag av sannsynligheter vil dermed gjøre prosedyren Dahlman foreslår, om å anslå den numeriske beviskraften, lite treffsikker.



Figur 6

²⁷⁷ (Shley, Chan, Gunadi, & Ferecatu, 2021) s. 30

²⁷⁸ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.76

I Dahlmans eksempel blir beviskraften $0,8/0,05=16$. I motsetning til en sannsynlighet, der man måler anslaget opp mot 100%, er dette et absolutt tall. Det er lite informasjon i et absolutt tall. Skal 16 tolkes som sterk eller svak beviskraft? 16 er sterkt i forhold til 1, men svakt i forhold til 1000, og hvert fall svakt i forhold til 1 000 000. I Sverige, som anvender beviskraftsrater i retten for rettsmedisinske bevis, har de utviklet verbale skalaer for å sette beviskraften i perspektiv.²⁷⁹ Som man kan lese ut av tabellen tilsier en beviskraft på 16 en svak beviskraft sammenliknet med et DNA-bevis som vil ligge på +4, ifølge skalaen i tabell 2.²⁸⁰ Skalaen gir felles referansepunkter for beviskraften som vil gi økt sensitivitet i sannsynlighetsanslagene. Som man ser i tabellen under er den ikke lineær, slik at man likevel vil slite med avtakende sensitivitet i de høyeste intervallene.

| | | Beviskraft |
|----|--|-------------------|
| +4 | «resultaten talar ekstremt sterkt för att ...» | >1 000 000 |
| +3 | «resultaten talar sterkt för att ...» | 6 000 – 1 000 000 |
| +2 | «resultaten talar för att ...» | 100 - 6 000 |
| +1 | «resultaten talar i någon mån för att ...» | 6 - 100 |
| 0 | «resultaten talar varken för eller emot att ...» | < 6 |

Tabell 2

For å svare på spørsmålet som ble stilt innledningsvis; Er bevisbedømmer i stand til å anslå sannsynligheter numerisk, uten å falle i en kognitiv felle? Så er svaret mitt at dersom prospektteorien også rammer bevisbedømmere, vil de numeriske anslagene ikke være særlig presise på grunn av avtakende sensitivitet for endring i sannsynlighet. Man vil da i realiteten tolke skalaen motsatt vei. I stede for å anslå en beviskraft på 500 000, vil man kjenne etter på sitt indre signal om «resultaten talar sterkt för att...» bevistemaet stemmer, og oversette denne overbevisningsgraden via skalaen til 500 000. Som studier Dahlman selv referer til, viser,

²⁷⁹ (Morrison, 2012) s.6 (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.79

²⁸⁰ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.79

varierer hvordan bevisbedømmere tolker «taler for», «taler sterkt for» og «taler ekstremt sterkt for» påtakelig.²⁸¹

Dersom man bruker dette tallet til å videre regne ut bevisverdien ved hjelp av Bayes teorem, mener jeg at den endelige bevisverdien får en falsk vitenskapelig legitimering som gjør vurderingen mindre transparent. Dette fordi den ikke er åpen om hvor subjektiv vurderingen faktisk er.

5.6 Bayesiansk analyse kontra Bayes teorem

5.6.1 Superanalytikere

Når det er sagt er bayesiansk analyse den rasjonelle fremgangsmåten for å oppdatere sin subjektive mening, uavhengig om man har et objektivt eller subjektivt datasett å trekke konklusjoner fra. Man må imidlertid skille mellom når det er hensiktsmessig å bruke den matematiske formelen *Bayes teorem*, og tenkemåten *bayesiansk analyse*. Jeg mener at det er hensiktsmessig å bruke Bayes Teorem når man har tilnærmet objektive data å basere analysen på, og man har en konsekvent bevisbedømmer. Det vil si at det gir mening å bruke Bayes Teorem i for eksempel maskinlæring. Vurderingen baserer seg da relativt objektive data, og bedømmeren (maskinen) vil vekte og oppdatere sannsynlighetene likt hver gang. Men når det er snakk om et så subjektivt vurderingsgrunnlag som det ofte er i bevisbedømmelse, med en menneskelig bevisbedømmer som tenker naturlig kausalt fremfor statistisk, blir vurderingen mer transparent ved å bruke bayesiansk analyse.

Jeg introduserte bayesiansk analyse i eksperimentet med Tom W. Analysen tilsvarete teoremet, men formulert som verbale kontrollspørsmål. Kahneman oppsummerer med følgende spørsmål, der jeg har modifisert dem noe til å passe med bevisvurderingen:²⁸² 1) Foreta en intuitiv gjetning: Hva ville sannsynligheten for bevistemaet vært om beviset stemte? 2) Se etter hvor ofte bevistema oppstår i gjennomsnitt. 3) Anslå den prediktive kraften av informasjonen du har. 4) Justerer snittet i retning av den intuitive gjetningen, i den grad man anslår at bevisdataen kan predikere bevistema.

²⁸¹ (Dahlman, Beviskraft : metod för bevisvärdering i brottmål, 2018) s.80

²⁸² (Kahneman, Sibony, & Sunstein, Støy, 2021) s.387 flg.

Ta igjen eksempelet med vitnet som sier at bilen involvert i kollisjonen var rød. Dersom det stemmer at vitnet så en rød bil, er det sikkert at den involverte bilen var rød. Men sett bort fra informasjonen vitnet kommer med om fargen på bilen, hvor mange røde biler er på norske veier? Ut fra min erfaring tipper jeg 1 av 100.

Så vurderer jeg *vitnet* nærmere. Hvilke forutsetninger hadde vitnet for å observere fargen på bilen? Hvor sto vedkommende og hvilke lysforhold var det? Har vitnet noen insentiver til å lyve? Kan vitnet ha noen forbindelser til gjerningspersonen? Her stiller man seg altså kontrollspørsmål knyttet til *kvaliteten* på beviset, og hvor mye vekt man kan tillegge det. For å gjøre det enkelt, si at vitnet sto like ved og det var gode lysforhold. Men jeg vet ikke om det kan foreligge andre årsaker til at vitnet husker feil eller lyver. Så jeg konkluderer med høy prediktiv kraft, men ikke perfekt korrelasjon.

Jeg skal så konkludere om sannsynligheten er lav, høy, veldig høy eller veldig lav for om bilen var rød. For å konkludere tar jeg da utgangspunktet i snittet. Her kom jeg frem til ca 1 av 100, og dette er en ganske lav sannsynlighet. Retningen på min intuitive gjetning var at bilen var rød. Jeg landet så på at den prediktive kraften til bevisdataen var høy, men ikke veldig høy. Siden utgangspunktet for vurderingen er så lavt, greier ikke selv en høy beviskraft å trekke sannsynligheten for at bilen var rød til høyre enn til 50%. For at beviskravet skal være oppfylt kreves det med andre ord flere bevis som støtter den intuitive gjetningen.

Grunnen til at jeg tror at bayesiansk analyse er bedre egnet som bevisvurderingsmetode enn teoremet er at den er tro mot den intuitive kausale tankegangen. Kontrollspørsmålene gjør det lettere å se for seg sannsynlighetene man vurderer. De gir en romslig forståelse av den statistiske tenkemåten. Det er derfor det er lettere å forstå én prosent når den blir oppgitt som 1 av 100 enn 100%. 1 av 100 er et forhold man kan forestille seg, for eksempel én person mot en gruppe på 100 personer. Teoremet blir for abstrakt, fordi formelen er vanskelig å se for seg. Det er da man begynner å se de kognitive illusjonene, som inverse feilslutninger. Når man ikke greier å se det for seg greier man ikke å holde tungen rett i munnen. Metoden blir også mer transparent, fordi den kommuniserer tydelig at det er en subjektiv resoneringsmetode. Jeg mener dermed at bayesiansk analyse kan være et viktig supplerende verktøy for

bevisbedømmere under bevisvurderingen. Det ville dermed vært hensiktsmessig å trene bevisbedømmerne opp til å bli *superanalytikere*.²⁸³

Superanalytikere er individer som anslår prognoser som er bedre enn tilfeldig gjetning. I Philip Tetlock sitt Good Judgement Project som besto av titusener av vanlige folk, utgjorde rundt 2% superanalytikere. De ble blant annet spurt om å anslå «Hvor sannsynlig det var at Nord-Korea kom til å sprengre en atombombe innen utgangen av året», «Hvorvidt Russland offisielt kom til å annektere mer ukrainsk territorium i løpet av de neste tre månedene» ol. Det som kjennetegnet superanalytikerne, var deres evne til å bryte ned problemstillingen til mindre delspørsmål. De stiller seg selv spørsmålet «hva skal til for at svaret blir ja», og «hva skal til for at svare blir nei», i stedet på for å lytte til intuisjonen. De var også dyktige til å ta utgangspunkt i *grunnfrekvensen*. Ved å systematisk ta hensyn til grunnfrekvensen, tar man også systematisk hensyn til *tilfeldigheter*. Hvilket den menneskelige kognisjonen intuitivt er systematisk er dårlig på.

Superanalytikere scorer høyt på et mål av kognitiv stil som kalles «aktiv fordomsfri tenkning».²⁸⁴ Folk som scorer høyt på dette målet har en tendens til å oppdatere meningene sine etterhvert de får ny informasjon. For å oppsummere denne kognitive stilen med Johan Maynard Keynes sine ord «Når faktaene endrer seg, ombestemmer jeg meg. Hva gjør du?». De har en tenkemåte som Tetlock kaller en *evigvarende beta*. Begrepet evigvarende beta er lånt fra datavare-verdenen om programmer som ikke har som formål å lanseres som sluttversjon, men som til enhver tid brukes, analyseres og forbedres. En superanalytiker anvender med andre ord en uformell versjon av Bayes teorem. I stedet for å foreta faktiske numeriske anslag, bruker de formelen heller som hjelpespørsmål for å gjøre et mer presist sannsynlighetsanslag.

²⁸³ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.264 flg.

²⁸⁴ (Kahneman, Sibony, & Sunstein, Støy, 2021) s.269

6 Konklusjon

6.1 En praktisk bevisvurderingsteori

Denne oppgaven har drøftet følgende problemstilling:

I hvilken grad åpner og avhjelper de etablerte bevisvurderingsteoriene for kognitive skjevheter.

Bakteppet for denne problemstillingen er at norsk rett tradisjonelt har bygget på en forutsetning om at bevisbedømmer kan fastslå hva som har skjedd i en sak, ved hjelp av sin *sunne fornuft*. De siste femti årene med beslutningsteori har på sin side fastslått at det å stole på sin sunne fornuft gjør bevisvurderingen støyete og sårbar for kognitive skjevheter. Man trenger dermed en metode for å sikre at vurderingen for hvorvidt beviskravet er oppfylt skjer på en forutberegnelig måte. Denne oppgaven har dermed analysert de mest diskuterte bevisvurderingsteoriene i norsk juridisk litteratur, i lys av om de åpner for eller avhjelper kognitive skjevheter.

Det jeg i oppgaven har kommet frem til som hovedproblemet for bevisbedømmere i, er at man intuitivt tenker kausalt i bevisvurderingen. Kausal tankegang gjør at man konstruerer og leter etter årsak-virkning-sammenhenger som ikke nødvendigvis er der. Med andre ord tar bevisbedømmer ikke hensyn til *tilfeldigheter*.

Fortellingsperspektivet beskriver i utgangspunktet det bevisbedømmer *gjør*. Det vil si at perspektivet er ettergivende for denne skjevheten og erkjenner at den skjer. Tilhengere av dette perspektivet prøver på sin side å dressere trangen til å konstruere sammenhenger der det ikke er noen, ved å tvinge bevisbedømmer til å lage alternative fortellinger. Man skaper da det beslutningsteori kaller en *indre masse*. Hvilket vil modifisere det ekstreme intuitive anslaget. Problemet med denne fremgangsmåten er at den ikke har noen sikkerhetsmekanisme mot å være kritisk til mengden og kvaliteten til hvert enkelt bevis. Bevisene blir bare vurdert etter hvor godt de passer med fortellingen.

Sannsynlighetsparadigmet er på sin side normativt og sier noe om hva bevisbedømmer *burde* gjøre. Det man bør gjøre når man skal anslå sannsynligheter for om en hendelse har inntruffet er å tenke statistisk. Paradigmet legger til rette for statistisk tenkning ved å vise til Bayes

teorem. Ved å anvende Bayes teorem som bevisvurderingsregel, blir bevisbedømmer for det første tvunget til å ta stilling til bevisets grunnfrekvens, herunder sjansen for at omstendigheten er en *tilfeldighet*. For det andre må hen ta stilling til bevisets beviskraft, altså kvaliteten på beviset.

Problemet med sannsynlighetsparadigmet er hvordan man praktisk skal kunne anvende teoremet. Dahleman og Eide argumenterer for at man numerisk skal anslå grunnfrekvensen og beviskraften, for så å regne ut bevisverdien. Jeg mener at det er denne insisteringen på *regning* som er hovedproblemet. For når man prøver å regne ut bevisverdien, åpnes det for en rekke kognitive illusjoner bevisbedømmer sliter med å se forbi, nettopp fordi man intuitivt er rigget for å tenke kausalt. Dette resulterer i at anvendelsen av teoremet gjør resoneringen mindre transparent, enn om bevisbedømmer hadde vært åpen om at det var en ren subjektiv vurdering.

Men ren bayesiansk analyse er fremdeles en mer rasjonell måte å vurdere troverdigheten til bevisdata på enn ren «sunn fornuft». Man får et mer presist anslag av bevisverdien ved å stille seg et kontrollspørsmål om grunnfrekvens og beviskraft, enn å følge den rene magesfølelsen. Jeg ser på min side ikke problemet med å kombinere bayesiansk analyse med forklaringsperspektivets metode for å dressere fortellingene.

Hovedinnvendingen mot sannsynlighetsparadigmet som alminnelig bevisvurderingsregel er at perspektivet er låst til at man slutter *fra* bevisdata *til* rettsfakta. Med andre ord resonnementet går bare én vei.²⁸⁵ Dette forutsetter at et bevisdatum taler for seg selv, og et bevisdata ikke påvirker tolkningen av hverandre. Men som jeg har diskutert i fortellingsperspektivet blir ikke bevisdatum presentert i et vakuum. Allerede før rettssaken starter vet bevisbedømmer hva tiltalen gjelder. Alle bevis har dermed en kontekst de blir tolket inn i og bevisene som blir presentert gis mening i lys av informasjonen som allerede foreligger i saken. Man resonerer da også fra rettsfakta til bevisfakta. Men også på tvers mellom bevisfakta. I realiteten blir resonnementene sirkulære ikke lineære.

Fortellingsperspektivet er dermed bedre egnet som alminnelig metode for bevisvurdering. Men bayesiansk analyse kan utgjøre et essensielt verktøy i vurderingsleddet der bevisbedømmer vurderer hva som er de *etablerte omstendighetene*. Med andre ord burde

²⁸⁵ (Kolflaath, Bevisbedømmelse i praksis, 2013) s.94

bevisbedømmere trenes i bayesiansk analyse for å dressere fortellingene på en mer systematisk måte.

6.2 Bedre beslutningshygiene i retten

6.2.1 Gjennomgang av prosessreglene

Denne oppgaven har drøftet hvorvidt etablerte bevisvurderingsteorier åpner for eller avhjelper kognitive skjevheter. På veien har jeg imidlertid kommet over flere forslag til hvordan man kan forbedre beslutningshygiene i bevisvurderingen, som ikke nødvendigvis direkte relaterer seg til bevisvurderingsteori. Avslutningsvis ønsker jeg dermed å trekke frem tre funn som jeg mener er relativt lett tilgjengelige tiltak som kan iverksettes for å forbedre bevisvurderingsprosessen utover å forbedre bevisvurderingsteoriene.

For det første mener jeg det bør foretas en systematisk gjennomgang av prosessreglene med hensyn til kognitive skjevheter. Som jeg påpekte i delkapittel 4.4.2. risikerer man i straffeprosesslovens §289 a) å prime dommeren til at tiltalte er skyldig til den motsatte er bevist. Prosessreglene er altså i seg selv en kilde for kognitive feller. Men de kan også brukes aktivt til å legge til rette for god beslutningshygiene hvis man har dette i mente når reglene utarbeides.

6.2.2 Opplæring i fordomsfri tenkning

For det andre bør alle bevisbedømmere trenes til å bli superanalytikere som diskutert i 5.6.1. Når man skal vurdere de etablerte omstendighetene bør man: 1) Foreta en intuitiv gjetning: Hva ville sannsynligheten for bevistemaet vært om beviset stemte? 2) Se etter hvor ofte bevistema oppstår i gjennomsnitt. 3) Anslå den prediktive kraften av informasjonen man har. 4) Justerer snittet i retning av den intuitive gjetningen, i den grad man anslo at bevisdataen kunne predikere bevistema.

Lage vise masser på rådslagningsrommet

For det tredje mener jeg at det er mye å hente på å utnytte rådslagningen som dommerne har etter hovedforhandlingene til å skape en vis masse. I denne oppgaven har jeg gjort rede for hvilke statistiske skjevheter som er knyttet til den psykologiske tolkningsprosessen av bevis. Dette er empirisk identifiserte psykologiske skjevheter, som man kan iverksette kirurgiske

tiltak for å motvirke. Slike tiltak er å gå gjennom prosessreglene og ha gode bevisvurderingsmetoder.

Men beslutninger har flere slagsider enn dem man til nå har fått avdekket empirisk. Det er i tillegg svært individuelt i hvor stor grad skjevhetene påvirker den enkelte bevisbedømmer. Dette vil skape uønsket variasjon i vurderingene. Denne variasjonen kan korrigeres for ved hjelp av massenes visdom.

Rådslagningen kan da brukes som verktøy for å redusere psykologiske skjevheter fra bevisvurderingen. Gitt at dommerne får anledning til å danne seg selvstendige oppfatninger av saken, og gitt rom for å ytre dem. De psykologiske skjevhetene vil da dra i forskjellige retninger og kansellere hverandre ut.

Rådslagningen kan også være en arena der fagdommer kan veilede meddommerne til å anvende teknikkene fra forklaringsperspektivet, og bayesiansk analyse for å vurdere bevissituasjonen. På denne måten vil rådslagningen også korrigere for de statistiske skjevhetene.

I denne forbindelse er det passende å helt avslutningsvis nevne en viktig begrensning i metodene jeg har foreslått. Siktemålet med denne oppgaven har vært å avdekke og foreslå hvordan man kan rette opp kognitive skjevheter. For å oppnå forutberegnelighet. Men dette er ikke noen mirakelkur mot diskriminering i retten. Skjevheter er definert som avvik fra den verdien en vis masse ville kommet frem til. Men samfunnets verdigrunnlag vil betinge hva den vise massen ville vurdert som sant. Med andre ord kan det eksisterer skjevheter på samfunnsnivå, som rasisme og kjønnsdiskriminering. Rasisme er forutberegnelig i et rasistisk samfunn. Systematisk holdningsarbeid med å bekjempe feilaktige stereotyper vil derfor være avgjørende for at den materielle sannhets prinsipp skal overholdes.

Litteraturliste

- American Psychological Association. (2015). *APA Dictionary of Psychology, Second Edition*.
- Andersen , T. Ø. (2009, 02 29.). *Bitt i øyet - krever 36 mill*. Hentet fra E24.no:
<https://e24.no/naeringsliv/i/Mg0w9R/bitt-i-oeyet-krever-36-mill>
- Backer , I. L. (2015). *Norsk sivilprosess*. Universitetsforlaget .
- Baron, J., & Hershey, J. C. (1988). Outcome Bias in Decision Evaluation. *Journal of Personality and Social Psychology*, ss. 569-579.
- Bøhn, A. (2009). Ikke grunnlag for å avskaffe juryen. *Rett på sak* .
- Cheng, E. K. (2009). A Practical Solution to the Reference Class Problem . *Columbia Law Review*.
- Corklin, M. (2019). The effectiveness of baysian jury instructions in midigating the defence attorneys fallacy. *HOUSTON LAW REVIEW*.
- Dahlman, C. (2018). *Beviskraft : metod för bevisvärdering i brottmål*. Norstedts Juridik.
- Dahlman, C., & Kolflaath , E. (2021). The Problem of the Prior in Criminal Trials. I C. Dahlman, A. Stein, & G. Tuzet, *Philosophical Foundations of Evidence Law*. Oxford: Oxford press.
- Dahlman, C., Zenker, F., & Sarwar, F. (2016, Desember). Miss rate neglect in legal evidence. *Law, Probability and Risk*.
- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *the Proceedings of the National Academy of Science*. .
- Domstol.no*. (2021, Mai 11.). Hentet fra Juryordningen: <https://www.domstol.no/om-domstolene/de-alminnelige-domstolene/domstolenes-uavhengighet/juryordningen/>
- Dybvig, E. (2022, 01 28). *Gaute Drevdal frikjennes for åtte av ni voldtekter*. Hentet fra Morgenbladet.no: <https://www.morgenbladet.no/aktuelt/2022/01/28/gaute-drevdal-frikjennes-for-atte-av-ni-voldtekter/>
- Eide , E. (2016). *Bevisvurdering - Usikkerhet og sannsynlighet* . Cappelen Damm.
- Guthrie, C., Jeffrey, R., & Wistrich, A. (2007). Blinking on the bench: how judges decide cases. *Cornell Law Reveiw*.
- Guthrie, C., Rachlinski, J., & Wistrich, A. (2001). Inside the Judicial Mind. *Cornell Law Library*.

- Hagstrøm, V., & Stenvik, A. (2015). *Erstatningsrett*. Oslo: Universitetsforlaget.
- Harley, E. M. (2007). Hindsight bias in legal decision making. *Social Cognition*.
- Hastie, R., Schkade, D. A., & Paye, J. W. (1999). Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages. *Law and Human Behavior*.
- Hastie, R., & Pennington, N. (1993). The story model for juror decision making. I R. Hastie, *Inside the Juror: the Psychology of Juror Decision Making* (ss. 192-221). Cambridge: Cambridge University Press.
- Hjort, L. N. (2021). *Troverdighet i rettsalen* . Oslo: Pax .
- Hov, J. (2019). *Rettergang i sivile saker* . Oslo : Papinian.
- Jury analyst. (2021, 07. 22). *Danny Kahneman on Decision Hygiene*. Hentet fra <https://juryanalyst.com/>: <https://juryanalyst.com/blog/daniel-kahneman-juror-bias-cognitive-bias/>
- Kahneman, D., Sibony, O., & Sunstein, C. (2021). *Noise*. William Collins .
- Kahneman, D. (2013). *Tenke, fort og langsomt*. Pax Forlag.
- Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist* .
- Kahneman, D., Sibony, O., & Sunstein, C. (2021). *Støy*. Oslo: Pax forlag .
- Kjelby, G. J. (2015). Bevisrettens grunnprinsipper og hovedregler i straffesaker. I R. Aarli, M.-A. Hedlund, & S. E. Jebsen, *Bevis i Straffesaker* (ss. 79-152).
- Kjus, A. (2008). *Sakens Fakta* . Oslo: Unipub.
- Kolfaath, E. (2011). Bevist utover enhver rimelig tvil. *Tidsskrift for Rettsvitenskap*, ss. s.135-196 .
- Kolflaath, E. (2015). En Metode for Bevisbedømmelsen i Straffesaker. I R. Aarli, & M.-A. Hedlund, *Straffesaker i Bevisretten* (ss. 507-534). Oslo: Gyldendal.
- Kolflaath, E. (2018). Domskonferanser i straffesaker . *Lov og Rett*, ss. 359-377.
- Kolflaath, E., & Strandberg, M. (2015). Tre Epoker i Norsk Bevsiteori - med særlig vekt på straffeprosessen. I R. Aarli, & M.-A. Hedlund, *Bevisretten i Straffesaker* (ss. 33-47). Oslo: Gyldendal.
- Kolflaath, E. (2004). Bevisbedømmelse - Sannsynligheter eller fortellinger. *Jussens Venner* , ss. 281-304 .
- Kolflaath, E. (2013). *Bevisbedømmelse i praksis*. Fagbokforlaget .

- Lie, N. (2012). *Parts- og vitneavhør - i straffesaker og sivile saker*. Oslo: Cappelen Damm.
- Magnussen, S., & Stridbeck, U. (2020). Vurdering av troverdighet: Hva sier forskningen anno 2020? *Tidsskrift for strafferett*, ss. 6-17.
- Morrison, G. S. (2012). The likelihood-ratio framework and forensic evidence in court: a response to R v T. *The international journal of evidence & proof*.
- NOU 2002:11 "Dømt av likemenn" . (u.d.).
- NOU 2011:13 Juryutvalget: Når sant skal skrives. (u.d.). *NOU 2011:13 Juryutvalget: Når sant skal skrives*.
- NTB (19.07.13). *Her gir politiet flest "tissebøter"*. Hentet fra abcnyheter.no: <https://www.abcnyheter.no/nyheter/norge/2019/07/13/195593939/her-gir-politiet-flest-tissebøter>
- Nrk. (2017, 02. 01.). *Slik forklarer Jensen og Cappelen «blomsterspråket»*. Hentet fra Nrk.no: https://www.nrk.no/norge/slik-forklarer-jensen-og-cappelen-_blomsterspraket_-1.13354968
- NUT 1969:3 (1969). Innstilling om rettergangsmåten i straffesaker fra Straffeprosesslovkomiteen. (u.d.).
- Oeberst, A., & Goeckenjan, I. (2016). When Being Wise After the Event Results in Injustice: Evidence for Hindsight Bias in Judges' Negligence Assessments. *Psychology, Public Policy, and Law*.
- Olsbu, E., & Kristensen, E. (2020, Januar 29.). *Baneheia-saken: DNA-beviset som raknet*. Hentet fra Aftenposten : <https://www.aftenposten.no/norge/i/GGg2l6/baneheia-saken-dna-beviset-som-raknet>
- Ot.prp.nr.11 (2007-2008) Om lov om endringer i straffeprosessloven mv. (styrket stilling for fornærmede og etterlatte)
- Politiet: Kripos . (2021). *Kripos - Nasjonal drapsoversikt 2021* . Politiet.
- Rachlinski, J., & Wistrich, A. (2017). Judging the Judiciary by the Numbers: Empirical Research on Judges. *Cornell Law Faculty Publications*.
- Rachlinski, J., Wistrich, A., & Chris, G. (2005). Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding. *University of Pennsylvania Law Review*.
- Riksadvokaten. (2018). *Rundskriv nr. 3/2018*. Riksadvokaten .
- Robberstad, A. (2014). Eivind Kolflaath: Bevisbedømmelse i praksis. *Tidsskrift for rettsvitenskap*.
- Schei, T., Berg, M., Borgerud, I. M., Hellejsen, A., Bjella, K., Bårdsen, A., & Norden, D. B. (2001). *NOU 2001:32 A Rett på sak* .

- Schweizer, M. (2005). Kognitive Täuschungen vor Gericht – Eine Empirische Studie. *Dissertation der Rechtswissenschaftlichen Fakultät der Universität Zürich.*
- Shley, D., Chan, H.-y., Gunadi, M., & Ferecatu, A. (2021, November 10.). *How Categorization Shapes the Propability Weighting Function.* Hentet fra SSRN: <http://dx.doi.org/10.2139/ssrn.3959751>
- Simon , H. (1955, Februar). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, ss. 99-118.
- Sommerset, T., Isungset, O., & Skramstad, M. (2020, 10. 4.). *Var Viggo i bua eller i Baneheia?* Hentet fra Nr.no: https://www.nrk.no/sorlandet/xl/var-viggo-i-bua-eller-i-baneheia_-1.15136771
- Spiegelhalter, D. (2019). *The art of Statistics - Learning from Data.* Pelican.
- Strandbakken, A. (2003). *Uskyldspresumsjonen "In dubio pro reo".* Fagbokforlaget .
- Strandberg, M. (2012). *Beviskrav i sivile saker .* Fagbokforlaget.
- Taleb, N. N. (2004). *Fooled by Randomness.* London: Penguin books .
- Teigen, E. (2020, 01 01). *Nytt bevis kan snu opp ned på Sveriges største Metoo-sak.* Hentet fra Nettavisen.no: <https://www.nettavisen.no/okonomi/nytt-bevis-kan-snu-opp-ned-pa-sveriges-storste-metoo-sak/s/12-95-3423902832>
- Øyen , Ø. (2019). *Straffeprosess.* Fagbokforlaget.