

Thomas Hobbes' filosofi om mennesket og staten:  
En tolkning og kritisk vurdering

Henrik Sætra

*Mai 2009 (H08/V09)*

Masteroppgave  
Institutt for statsvitenskap  
Universitet i Oslo



Antall ord: 75.621

Inkl. innholdsfortegnelse, bibliografi og register.

## Sammendrag

Denne oppgaven består av to deler, der del 1 omhandler Hobbes' filosofi om individet, mens del 2 er en vurdering av hans politiske filosofi. Hobbes' filosofi om mennesket undervurderes ofte, og jeg presenterer en analyse av Hobbes' tenkning om dette temaet som viser at Hobbes ikke anser mennesker som rene kyniske egoister. Jeg sammenstiller så Hobbes' tenkning om mennesket med Maslow og eksperimentell økonomi og nevroøkonomi, og viser at Hobbes vanskelig kan hevdes å være "motbevist" av moderne teorier om mennesket. Det er imidlertid viktige elementer som fremkommer i disse nyere bidragene, som ikke diskuteres av Hobbes. Jeg ser videre på hvordan en utvidelse av forståelsen av menneskets natur, spesielt på bakgrunn av Maslows bidrag, kan synes å påvirke Hobbes' politiske filosofi. Jeg konkluderer med at et moderne Hobbesiansk resonnement ikke nødvendigvis må være likt Hobbes' eget; mennesker kan ha visse behov som gjør Hobbes' absolutte enevelde uholdbart og lite egnet for å bevare fred, samtidig som moderne demokratiske styreformer og historiske bevis sterkt svekker Hobbes' forsvar for eneveldet. Med bakgrunn i Hobbes' systematiske resonnement reviderer jeg visse deler av hans politiske filosofi, slik at jeg står igjen med en "Hobbesiansk filosofi" som er universell; universell i den forstand at filosofien ikke er historisk eller kulturelt betinget.

# Forord

Først av alt vil jeg takke mine veiledere i arbeidet med denne oppgaven: Leif Helland og Knut Midgaard (H-08/V-09). Jeg var så heldig å få muligheten til å følge Midgaards kurs “Spillteori, kommunikasjonsteori og forhandlingsteori” (STV4103B, V-08), og enda heldigere ved at Midgaard sa seg villig til å være involvert i veiledningen av min masteroppgave. Midgaard ønsket selv å ha med Leif Helland i dette arbeidet, og takket være Hellands og Institutt for statsvitenskaps velvilje har jeg vært så heldig å ha to svært dyktige hjelpere i mitt arbeid. Sjelden har det vært så sant at samtlige mangler ved oppgaven skyldes meg selv.

Jeg ønsker også å takke Pål Foss for hans provokative introduksjon til politisk teori ved Høgskolen i Østfold. Uten hans tydelige entusiasme og faglige tyngde i emnene som ble presentert, er det vanskelig å si om jeg noen gang hadde funnet fagområdet som har opptatt så store deler av årene som har fulgt.

Av mine venner vil jeg nevne to som har vært til stor hjelp i mitt arbeid med oppgaven. Kjetil August Tysse-Hostad har i løpet av mine studier vært en konstant kilde til tekstuell hjelp, og mangfoldige diskusjoner om diverse statsvitenskapelige emner har hjulpet meg til å utkrystallisere mine standpunkter. I tillegg til en takk for det faktum at han har funnet tid til å lese min oppgave og gi meg tilbakemeldinger, vil jeg gi Shahryar Attar Kashani en unnskyldning for alle avbrytelser i fritid og arbeid i løpet av min skriveprosess; det til tider ensomme skrivearbeidet lettes når man har venner man kan dele fantastiske sitater og morsomme idéer med.

Ingen fortjener dog mer takk enn min familie: Rita Karina Bjørsvik og Brage Bjørsvik Sætra. Deres evne til å tolerere min fysiske og psykiske fjernhet i denne perioden med skriving (og i løpet av mine studier generelt) er den eneste grunnen til at oppgaven eksisterer, og jeg ser frem til muligheten til å være mer tilstede. Også mine foreldre, Henning Sætra og Synøve Skaug, må takkes for all støtten og oppmuntringen jeg alltid

har fått.

Ellers en noe forsinket takk til Thomas Hobbes, som produserte et verk som for meg har vært utrolig stimulerende. Samtidig, dersom Hobbes kan sies ha stilt diagnosen, har jeg funnet vel så mye verdifult hos Marcus Aurelius, som i så tilfelle er medisinen man kunne trenge mot Hobbes' virkelighet.

# Innhold

<b>1</b>	<b>Innledning</b>	<b>6</b>
<b>I</b>	<b>Individet</b>	<b>8</b>
<b>2</b>	<b>Hobbes' syn på menneskets natur</b>	<b>9</b>
2.1	Introduksjon . . . . .	9
2.2	<i>Homo economicus</i> ' nyttefunksjon . . . . .	9
2.3	Nytte . . . . .	10
2.3.1	Det gode . . . . .	10
2.3.2	Makt . . . . .	13
2.3.3	Motivasjon . . . . .	15
2.3.4	Negativ nytte - <i>summum malum</i> . . . . .	17
2.3.5	Oppsummering . . . . .	18
2.4	Andres og egen nytte . . . . .	19
2.4.1	Kunnskap om andres motivasjon og preferanser . . . . .	19
2.4.2	Variasjoner i nytteverdinger og nyttesammenligning . . . . .	20
2.4.3	Godhet og veldedighet . . . . .	22
2.4.4	Rettferdighet og moral . . . . .	25
2.4.5	Relativ gevinst . . . . .	27
2.4.6	Innstilling til og vurdering av andre . . . . .	28
2.4.7	Oppsummering . . . . .	32
2.5	Kortsiktighet og langsiktighet . . . . .	32
2.5.1	Kortsiktighet . . . . .	33
2.5.2	Langsiktighet - grunnleggende i Hobbes' filosofi . . . . .	34
2.5.3	Generelt <i>krav</i> om langsiktig adferd for overlevelse/suksess . . . . .	36
2.5.4	Risikoaversjon . . . . .	37
2.5.5	Oppsummering . . . . .	37
2.6	Konklusjon . . . . .	38
<b>3</b>	<b>Hobbes' filosofi om individet vurdert i lys av nyere bidrag</b>	<b>40</b>
3.1	Innledning . . . . .	40
3.2	Sekundærlitteratur om Hobbes . . . . .	41
3.3	Psykologi – Humanisten og optimisten Maslow . . . . .	43

3.3.1	Menneskelig motivasjon og grunnleggende behov . . . . .	45
	Fysiologiske behov . . . . .	45
	Sikkerhetsbehov . . . . .	47
	Behov for tilhørighet og kjærlighet . . . . .	49
	Behov for anerkjennelse og respekt . . . . .	50
	Selvrealisering . . . . .	51
3.3.2	Innstilling til menneskets natur . . . . .	52
3.3.3	Instinkter, rasjonalitet og det gode . . . . .	55
3.3.4	På hvilke måter skiller Hobbes seg fra Maslow? . . . . .	57
3.3.5	Oppsummering . . . . .	59
3.4	Nevroøkonomi og eksperimentell økonomi – empiriske tilnærminger til mennesket . . . . .	60
3.4.1	Bakgrunn . . . . .	60
3.4.2	Metode . . . . .	62
3.4.3	Hovedbegreper . . . . .	64
	Sosiale preferanser og heterogenitet . . . . .	64
	Sterk gjensidighet og urettferdighetsaversjon . . . . .	66
	Altruistisk straff . . . . .	68
	Følelser . . . . .	72
	Instinkter og evolusjon . . . . .	73
3.4.4	Metodologiske innvendinger . . . . .	74
	Anonymitet og oppdemming av spill . . . . .	75
	Kontrolleffekter og etiske begrensninger . . . . .	78
3.4.5	Alternative forklaringer . . . . .	80
	“Godfølelse-altruisme” . . . . .	80
	Indirekte gjensidighet . . . . .	81
	Belønningsteori . . . . .	82
	Rykte . . . . .	83
	Kostbar signalisering . . . . .	84
3.4.6	Oppsummering . . . . .	85
3.5	Konklusjon . . . . .	86

## **II Samfunnet 88**

### **4 Hobbes’ resonnement om statens form 89**

4.1	Introduksjon . . . . .	89
4.1.1	Hobbes og autoritarisme . . . . .	90
	Politisk konkurranse og bindinger på staten . . . . .	92
	Legale, religiøse eller moralske bindinger på staten . . . . .	94
	Hobbes’ stat . . . . .	96
4.2	Hobbes’ resonnement . . . . .	98
4.2.1	Forberedende betraktninger om naturtilstanden . . . . .	99
4.2.2	De grunnleggende premissene . . . . .	101

	Behovet for en stat . . . . .	101
	Implikasjonene av synet på døden som det verste ondet . . . . .	106
	Maktens delbarhet . . . . .	111
	Demokrati og enevelde . . . . .	113
	Risikovillighet og ønsket om sikkerhet . . . . .	118
	Hobbes' "liberalisme" . . . . .	120
4.2.3	Resonnementets konklusjon . . . . .	123
	Kort tekstuell vurdering av resonnementet . . . . .	123
4.3	Konklusjon . . . . .	124
<b>5</b>	<b>En kritisk gjennomgang av Hobbes' politiske resonnement</b>	<b>127</b>
5.1	Innledning . . . . .	127
5.2	Argumenter mot staten som en nødvendighet . . . . .	128
5.2.1	Fellesgoder og samarbeid . . . . .	129
5.2.2	Frivillig samarbeid, kommutarianske samfunn og normer . . . . .	137
	Forskjellige løsninger på samhandlingsproblemer . . . . .	137
	Ulemper ved kommutarianske samfunn . . . . .	141
5.2.3	Hobbes og muligheten for utenomstatlige løsninger på naturtilstandens tragedie . . . . .	146
	Fremstillingen av Hobbes' filosofi i enkel spillform . . . . .	146
	Er naturtilstanden et Chicken- eller et Fangenes dilemma-spill? . . . . .	149
	Grunnleggerne av staten – hvem er de? . . . . .	152
	Hobbes' naturtilstand: alle mot én, alle mot alle, eller de mot dem? . . . . .	154
	Hobbes om frivillig samarbeid: små og store samfunn . . . . .	161
5.2.4	Oppsummering . . . . .	166
5.3	Mennesker og behovet for mer enn sikkerhet . . . . .	167
5.3.1	Innledende betraktninger om Hobbes' stat . . . . .	168
5.3.2	Frihet og demokrati - <i>behov</i> eller <i>ønsker</i> ? . . . . .	169
	Frihet som et grunnleggende menneskelig behov . . . . .	170
	Frihetens fordeler . . . . .	176
	Krever frihet demokrati? . . . . .	177
	Frihet som en selvfølgelighet og jorden som flat . . . . .	178
	Arbeidshypotese . . . . .	180
5.3.3	Hobbes' resonnement revurdert . . . . .	181
	Vurdering av premiss 2: <i>er</i> døden det verste ondet? . . . . .	182
	Vurdering av premiss 3: Hobbes' absolutisme . . . . .	185
	Vurdering av premiss 4 og 5: demokrati som ustabil og mennesker som ekstremt sikkerhetsfokuserte . . . . .	189
	Vurdering av premiss 6: Hobbes' "liberalisme" . . . . .	197
5.4	Konklusjon . . . . .	197
<b>6</b>	<b>Konklusjon</b>	<b>200</b>

# Kapittel 1

## Innledning

Thomas Hobbes (1588-1679) er av mange ansett som forfaderen til *det moderne*. Hans hovedverk, *Leviathan*, blir ofte omtalt som det største verket innen politisk filosofi skrevet på engelsk. Hobbes var en av de første systematiske tenkerne, noe som gjør at hans verker selv i dag fremstår som relevante.

Hva *menneskets natur* er har blitt forsøkt besvart uttallige ganger, og Hobbes' eget svar på dette spørsmålet fremstår som et av de mer sentrale, både på grunn av svarets kvaliteter, men også på grunn av de mange motstanderne av Hobbes' syn. Videre er Hobbes' tanker om mennesket en viktig del av hans politiske filosofi, noe som gjør en grundig vurdering av Hobbes' syn på mennesket nyttig.

Det er hevet over enhver tvil at Hobbes' bidrag til den politiske filosofien har mange sterke sider. Det har dog i tiden etter Hobbes blitt like tydelig at mange oppfatter alvorlige svakheter ved hans filosofi. Spesielt Hobbes' syn på mennesket og hans forkjærlighet for det absolutte eneveldet har til stadighet kommet under angrep, og det er disse temaene jeg ønsker å se på i denne oppgaven. Jeg vil vurdere spesielt Hobbes' politiske filosofi kritisk, for å avdekke hvorvidt det er behov for å revidere visse elementer av denne filosofien i lys av nyere historie. I motsetning til mye annen kritikk vil jeg imidlertid forsøke å yte Hobbes' filosofi mest mulig rettferdighet, ved at jeg ser på *helheten* og *sammenhengen* i hans tenkning, ikke løsrevne deler. Hobbes opplevde selv mye kritikk han anså som uholdbar:

[...] he [Lord Bishop of Derry] hath put together divers sentences picked out of my *Leviathan*, which stand there plainly and firmly proved, and sets



them down without their proofs, and without the order of their dependance one upon another; and calls them atheism, blasphemy, impiety, subversion of religion, and by other names of that kind. My request unto you [leseren] is, that when he cites my words for erroneous, you will be pleased to turn to the place itself, and see whether they be well proved, and how to be understood (Hobbes 2005b:281).

Det er viktig for meg å holde meg til hva Hobbes selv har sagt og forklart i denne oppgaven, da målet ikke er å fremsette et “moderne” resonnement for enhver pris. Det jeg anser som sentralt hos Hobbes vil beholdes, selv om det kan være kontroversielt, mens jeg argumenterer for at andre kontroversielle sider ved Hobbes bør tillegges liten vekt.

Jeg vil besvare to spørsmål i denne oppgaven:

- Hva er Hobbes’ syn på menneskets natur, og hvordan stemmer dette synet overens med moderne forståelser av mennesket?
- Hvordan ser Hobbes’ begrunnelse for sin foretrukne stat ut i “ren” form, og hvor tar Hobbes feil, dersom hans type stat *ikke* er en nødvendighet?

Del 1 av oppgaven (kap 2 og 3) omhandler det første leddet i problemstillingen. I kapittel 2 foretar jeg en analyse av Hobbes’ tekster, for å tydeliggjøre det jeg mener er en ofte oversett dybde i Hobbes’ tenkning om mennesket. I det følgende kapittelet ser jeg på mer aktuelle bidrag der menneskets natur behandles, spesielt i forbindelse med psykologen Maslow og eksperimentell økonomi og nevroøkonomi.

Det andre leddet i problemstillingen behandles i kapittel 4, der jeg fremsetter min versjon av Hobbes’ begrunnelse for den type stat han foretrekker, i form av et resonnement som bygger på seks premisser. I kapittel 5 ser jeg så nærmere på disse premissene, og jeg vil, der jeg finner det tillatelig, revidere premissene, for avslutningsvis å stå igjen med et mer moderne Hobbesiansk resonnement.

Del I

Individet

# Kapittel 2

## Hobbes' syn på menneskets natur

### 2.1 Introduksjon

En gjennomgang av Hobbes' filosofi har et naturlig utgangspunkt i *Leviathan* (1946), så også denne analysen av Hobbes' syn på menneskets natur. Det er imidlertid flere interessante variasjoner mellom Hobbes' forskjellige tekster, og jeg vil også benytte meg av innhold fra hans andre bidrag der disse utfyller, eller motsier, funnene fra *Leviathan*. Forskjellig kommentarlitteratur og relaterte bidrag vil i all hovedsak gjennomgås i neste kapittel.

I dette kapitlet legger jeg grunnlaget for en videre diskusjon av Hobbes' filosofi, med en gjennomgang av en del grunnleggende begreper i forbindelse med individers nyttefunksjoner. Ved å ta utgangspunkt i en helt enkel nyttefunksjon, vil jeg se på hvilke mangler denne har i forhold til en realistisk modell av menneskelig motivasjon, basert på Hobbes' filosofi. Mitt hovedformål er her å avdekke hva Hobbes faktisk sier om mennesker, og avslutningsvis i kapitlet vil jeg se kort på hvordan jeg ser for meg at en nyttefunksjon basert på Hobbes' syn på mennesket kunne tenkes å arte seg.

### 2.2 *Homo economicus*' nyttefunksjon

Et nyttig utgangspunkt i denne sammenheng er en helt enkel nyttefunksjon som ofte anvendes innen økonomi:

$$U_i = f_i(x_i), \quad f' > 0, f'' < 0 \quad (2.1)$$

Person  $i$ 's nytte gis av en strengt økende ( $f' > 0$ ) og konkav ( $f'' < 0$ ) funksjon av den delen av gevinst  $x$  som tilfaller  $i$  (her representert ved  $x_i$ ).  $x$  er i denne sammenheng f.eks. en direkte representasjon av et monetært utfall. Denne nyttefunksjonen er naturlig nok noe begrenset når det gjelder å predikere eller forstå kompliserte reelle situasjoner, men den gir oss viktige idéer om hva som mangler for at en nyttefunksjon skal bli mer realistisk i forhold til Hobbes' syn på mennesket.

For det første er det nødvendig å se på hva *nytte* innebærer for et Hobbesiansk individ; det synes åpenbart at en direkte representasjon av f.eks. monetær gevinst ikke er tilfredsstillende. For det annet må personers relasjoner, og deres oppfatninger av andre individers person og utfall, vurderes.

## 2.3 Nytte

Jeg tar som utgangspunkt at individer er *nytte-maksimerende*; de søker å oppnå en maksimal grad av nytte ut fra sine egne oppfatninger og vurderinger av situasjonene de befinner seg i. Dette er ikke ensbetydende med maksimering av f.eks. profitt, konsum eller lønn, som er lignende forutsetninger som ofte gjøres i økonomisk litteratur. Nytte er et langt bredere begrep enn disse økonomiske størrelsene, og mens man i de fleste tilfeller kan tilegne seg nytte f.eks. via profitt, er det også mange andre kilder til økt nytte. Bruken av *nytte* fremfor andre objektive størrelser gjør samtidig en formalisering av individers handlingsgrunnlag mer komplisert.

### 2.3.1 Det gode

Ifølge Hobbes har alle våre appetitter og aversjoner sitt opphav i de minste interne bevegelser. Disse *bestrebelsene* (eller impulsene) består enten i å bli tiltrukket av (appetitt) eller frastøtt fra (aversjon) noe (Hobbes 1946:31). Hvilket som er tilfelle er etter Hobbes' syn det avgjørende for hva som er *godt* og *ondt*; det en person har appetitt for, eller begjærer, blir av han kalt godt, mens det onde blir kalt nettopp det fordi personen hater eller avskyr det (Hobbes 1946:32). Verdt å merke seg her er også koblingen

Hobbes gjør mellom “livskraft” (vital motion) og appetitt; det kan synes som om det er et instinktivt element som sørger for overenstemmelse mellom hva individer *anser* som godt, og hva som faktisk *er* godt for dem, fysiologisk sett (Hobbes 1946:33). Jeg antar heretter at tilegnelse eller oppnåelse av det gode fører til økt nytte for individer.

Av goder skiller Hobbes mellom tre varianter: *forventede* goder, *manifesterte* eller *direkte* goder (gode effekter) og til sist gode *midler*, som er nyttige i anskaffelse av goder (Hobbes 1946:33). Det er altså tydelig at Hobbes’ individer er fremsynte; de tillegger goder i fremtiden, og midler for å anskaffe seg ytterligere goder, verdi i sine vurderinger, noe jeg kommer tilbake til senere i kapitlet.

Det ypperste et individ kan oppnå er for Hobbes *lykke* (felicity), en dynamisk tilstand som oppnås ved kontinuerlig tilfredsstillelse av sine ønsker (Hobbes 1946:39).

For there is no such thing as perpetual tranquility of mind, while we live here; because life itself is but motion, and can never be without desire, nor without fear, no more than without sense (Hobbes 1946:39).

Felicity is a continual progress of the desire, from one object to another; the attaining of the former, being still but he way to the latter (Hobbes 1946:63).

Dynamikken i denne tilstanden er på mange måter et resultat av Hobbes’ grunnleggende vitenskapsteori, der alt forstås som bevegelse (Hobbes 1946:7,8,9,39). Dersom et objekt ligger i ro, er det fordi krefter annullerer hverandre slik at man *oppfatter* ro, mens det mer korrekt kan karakteriseres som en situasjon med intens spenning<sup>1</sup>.

Videre kvalifiserer Hobbes svært ofte “det gode” med *tilsynelatende*, noe man i praksis også gjør i moderne teori når man ser på *forventet* nytte. “For hva en person som forsøker å se fremover i tid angår, dersom de gode konsekvensene veier tyngre enn de negative, er hele kjeden det forfatteren kaller *tilsynelatende*, eller *oppfattet god*” (Hobbes 1946:39). Enhver vurdering av det gode som ikke er umiddelbar og nær, tillegges altså en viss usikkerhet, som ifølge Hobbes er minst hos de med mye erfaring og fornuft (Hobbes 1946:39). Svært mange steder i *Leviathan* (og ellers) hevder Hobbes at mennesker er tilbøyelige til å feile, både pga. mangel på vett, vaner, subjektivitet, dårlige råd og mangelfull kompetanse/dårlige forutsetninger for å utlede endelige utfall der det er

---

<sup>1</sup>At dette er tilfellet hos mennesker som oppnår lykke blir tydeligere senere i kapitlet, men kort fortalt dreier det seg om en konstant søken etter og innsats for å oppnå *mer*, for å oppnå en *tilsynelatende* statisk tilstand med lykke.

mange mellomledd mellom nåtid og endelig resultat i en kausalkjede (Hobbes 1946:26, 68, 102, 123, 124, 128, 148, 169, 180, 241, 289, 1991:65).

Sometimes a man desires to know the event of an action; and then he thinketh of some like action past, and the events thereof one after another; supposing like events will follow like actions. [...] though such conjecture, through the difficulty of observing all circumstances, be very fallacious (Hobbes 1946:15).

Denne feilbarligheten er viktig for forståelsen av Hobbes, for ved den oppfatning at aktører alltid valgte “rett” ut fra sin egeninteresse – hadde forventningsrette antagelser – ville man kunne forestille seg et perfekt samfunn der alle avveide fremtidige konsekvenser mot umiddelbar gevinst; ifølge Hobbes’ resonnement ville de alltid ha fulgt loven og innfridd sine forpliktelser. *Fornuft og rettferdighet* er nemlig fullstendig forenlig for Hobbes; den eneste måten å sikre seg på i naturtilstanden er ved ved slå seg sammen med andre, “og han som erklærer at han synes det er fornuftig å bedra de som hjelper han, kan ikke [ved fornuft] forvente mer sikkerhet enn det han tidligere hadde ved sin egen makt” (Hobbes 1946:95). Bedrag, og brudd på avtaler/kontrakter, bringer oss nærmere naturtilstanden, og siden fornuft forbyr enhver handling som er skadelig for livet, finnes det kun én vei til den sikre og evige lykke<sup>2</sup>: “og det er ikke å bryte, men å holde, avtaler” (Hobbes 1946:96).

At mennesker misforstår hva som er til egen fordel, og at drifter noen ganger får utløp i handling uten at de blir kontrollert, er viktige årsaker til den kaotiske tilstanden og usikkerheten Hobbes beskriver. Dersom man skal forutsi en handling, er det altså ikke nok å analysere mulige utfall; man må analysere mulige kilder til forvrengt oppfatning av utfallet hos aktørene som velger. I en post factum analyse av en situasjon vil det derimot kunne være naturlig å se kun på utfall, dersom man f.eks. ønsker å vurdere graden av suksess ut fra observert handlingsrekke. Det må likevel også der tas høyde for kilder til forvrengning av aktørers oppfatning av utfallet, dersom man skal forstå hvorfor utfallet ble som det ble. Dersom en aktør “vant” og fikk et godt resultat, er ikke dette nødvendigvis en følge av målrettet og vellykket handling fra aktørens side; han, eller hans motparter, kan ha forventet helt andre utfall som følge av sine valg.

---

<sup>2</sup>Hobbes sier man aldri får evig sinnsro, men hevder her at vi *kan* oppnå evig lykke.

### 2.3.2 Makt

Når det er snakk om fremtidige goder, og midler til å oppnå goder generelt sett, er det viktig å se på Hobbes' definisjon av *makt*:

THE POWER *of a man*, to take it universally, is his present means, to obtain some future apparent good, and is either *original* or *instrumental* (Hobbes 1946:56).

Det er altså åpenbart at makt ikke er et gode som søkes for seg selv isolert sett, men fordi det er et middel til å oppnå fremtidige (tilsynelatende) goder. Makt er med andre ord *nyttig*, det er et gode av den tredje sorten nevnt i 2.3.1, det er et gode fordi det er et *middel*.

[. . .] desire for power, of riches, of knowledge, and of honour. All which may be reduced to the first, that is, desire of power. For riches, knowledge, and honour, are but several sorts of power (Hobbes 1946:46).

Hobbes reduserer her ønsket om rikdom, kunnskap og ære til et ønske om makt. Makt kan igjen reduseres til et ønske om å oppnå *lykke*, altså et ønske om å oppnå det gode. Det kan synes irrelevant å skille mellom disse forståelsene av makt, men det er sentralt når det gjelder den videre forståelsen av Hobbes. Mange ser ut til å fokusere uforholdsmessig mye på Hobbes' individers umettelige jakt på makt som noe negativt, mens jeg forstår denne søken etter makt som en helt naturlig bestrebelse etter det gode. Implikasjonene av disse forskjellige forståelsene er muligens nokså like, men den aversjonen mange har mot fremstillingen av mennesket som et herskesykt dyr, kan dempes noe med min forståelse av Hobbes på dette punktet. Kommentatorers fokus på søken etter makt kan dog forståes, da Hobbes skriver mye om temaet, og med en ordlyd som kan gi et feil inntrykk, dersom man ikke har hans foregående definisjoner av makt i bakhodet.

[. . .] the object of man's desire, is not to enjoy once only, and for one instant of time; but to assure for ever, the way of his future desire (Hobbes 1946:63).

And therefore the voluntary actions, and inclinations of all men, tend, not only to the procuring, but also to the assuring of a contented life [. . .] (Hobbes 1946:63)

So that in the first place, I put for a general inclination of all mankind,

a perpetual and restless desire for power after power, that ceaseth only in death (Hobbes 1946:64).

And the cause of this, is not always that a man hopes for a more intensive delight, than he has already attained to; or that he cannot be content with a moderate power: but because he cannot assure the power and means to live well, which he hath present, without the acquisition of more (Hobbes 1946:64).

Sitatene viser dog også *hvorfor* makt får denne fremtredende posisjonen i *Leviathan*. Søken etter goder foregår ikke bare i nåtid; individer trenger midler a) for å tilegne seg ytterligere goder for å oppnå lykke, og b) for å beskytte godene de allerede har. Sitatene har altså ikke forandret det fundamentale; makt har *ikke* blitt et mål i seg selv, men er fremdeles et middel for å nå ens egentlige mål: goder/lykke. Dersom makt utelukkende søkes for å oppnå noe *annet*, er det bemerkelsesverdig at dette kalles et (endelig) mål i seg selv. Hobbes uttrykker dette meget eksplisitt når han forklarer hvorfor man ønsker, og går med på, en statsdannelse; mennesker ønsker minimal motgang og sanselige gleder, noe som gjør dem tilbøyelige til å adlyde en samlet makt. Det samme gjelder frykten for døden (Hobbes 1946:64). Individer er altså villige til å ofre makt for å oppnå sikkerhet og muligheten til å leve et fredelig liv, med muligheter for å nyte de goder de måtte ønske (som sanselige gleder, kunst, kunnskap osv).

The final cause, end, or design of men, who naturally love liberty, and dominion over others, in the introduction of that restraint upon themselves, in which we see them live in commowealths, is the foresight of their own preservation, and of a more contented life thereby (Hobbes 1946:109).

*Tilfredshet* og moderasjon nevnes også flere steder av Hobbes, men begrepene er noe problematiske. Tilfredshet og moderasjon er faktisk tegn på en misforståelse, eller direkte "mangel", hos aktører, spesielt i naturtilstanden. Tilfredshet kan sees som en statisk tilstand der man er fornøyd med hva man har, noe som ikke stemmer overens med Hobbes' syn på *lykke*. Dersom man er tilfreds med hva man har, er det enten på grunn av a) at man ikke forstår at det man har i dag ikke vil bestå i morgen dersom man ikke aktivt søker mer, eller b) at man har dårlig fantasi og lite dømmekraft, "for det å ikke ønske noe, er være død" (Hobbes 1946:46,81,103). De "moderate" individene, her forstått som de tilfreds innenfor "rimelige grenser", vil etter Hobbes' oppfatning relativt raskt gå til grunne ved å bli tatt av dage av de som aktivt søker mer og mer



makt for å sikre seg og sitt. At disse aktørene eksisterer er for Hobbes selvsagt, og også naturlig, da den eneste måten å sikre seg selv på er å herske over flest mulig mennesker, helt til “han ikke ser noen makt av styrke nok til å true ham” (Hobbes 1946:68,81). Det er naturligvis store forskjeller mellom naturtilstanden og en etablert stat, men prinsippene Hobbes gjennomgår i bok 1 av *Leviathan*<sup>3</sup>, er likevel relevante når det gjelder menneskets natur.

For he that should be modest, and tractable, and perform all he promises, in such time, and place, where no man else should do so, should make himself a prey to others, and procure his own certain ruin, contrary to the ground of all laws of nature, which tend to nature’s preservation (Hobbes 1946:103).

*Moderasjon*, derimot, kan ansees som en taktikk, som i mange tilfeller vil være nødvendig, noe jeg også kommer tilbake til senere i kapitlet, i forbindelse med relativ gevinst og knappe goder i 2.4.5.

### 2.3.3 Motivasjon

Bildet Hobbes tegner av individers motivasjon er på ingen måte ukontroversielt, men hvorvidt dette stammer fra misforståelser, normative standpunkter eller feil hos Hobbes er mindre klart. Det som *er* klart, er at Hobbes’ filosofi er svært enkel på dette punktet – faktisk så enkel at den lett misforståes. Hobbes sier f.eks. om frivillige handlinger, at “formålet er *fordel for han selv* [aktøren]”, og at alt et menneske foretar seg gjøres til egen fordel (Hobbes 1946:86,102,124). En slik tilsynelatende “enkel” egoisme forstås derfor ofte som en *snever* egoisme der individer antas å være griske, have syke og uten empati. Hobbes ville nok ikke blitt spesielt indignert over en slik tolkning, siden han virker svært klar på at slike individer faktisk finnes. Problemet er at mange tolker Hobbes dithen at *alle* er slik, og dette er direkte feil, noe jeg vil komme nærmere inn på i min behandling av andres og egen nytte.

Som “begrensning” på søken etter egen fordel fremsetter Hobbes en plikt til å overholde avtaler – som samfunnskontrakten – selv om f.eks. kortsiktig profitt skulle tilsi å bryte denne, og dette rettfærdiggjøres ut fra individuell egeninteresse (Hobbes 1946:95, j.fr. 2.3.1). Hobbes påpeker dog også at selv om individers fornuft er forenlig med rettfæ-

---

<sup>3</sup>Bok 1: “OF MAN”.

dighet og loven, er menneskers *drifter* som oftest sterkere enn fornuften, noe som fører til mangelfull overholdelse av lover og avtaler, til tross for at respekt for loven ville være til det beste for alle (Hobbes 1946:122).

Min innvending mot eventuelle påstander om en “primitiv” egoisme hos Hobbes, er at man må lese mer enn én setning av Hobbes for å forstå hva han egentlig mener, da hver setning bygger på de foregående. Hvert ord han anser som viktig er utførlig definert, og i noen tilfeller på en måte som ikke er helt konvensjonell i henhold til dagens språk; å lese utdrag av Hobbes uten å forstå terminologien gir helt feil referanser, og i mange tilfeller en svært forvrengt forestilling om hva Hobbes faktisk sier.

[...] a man that seeketh precise truth had need to remember what every name he uses stands for, and to place it accordingly, or else he will find himself entangled in words, as a bird in lime twigs, the more he struggles the more belimed (Hobbes 1946:21).

Selv om det ligger en grunnleggende egeninteresse i bunnen av Hobbes’ individers motivasjon, gjør han det helt klart at man allikevel kan utføre “gode gjerninger”:

For no man giveth, but with intention of good to himself; because gift is voluntary; and of all voluntary acts, the object is to every man his own good [...] (Hobbes 1946:99).

Utsagnet om at motivasjonen for enhver frivillig handling er en fordel for aktøren selv, representerer noe av det som gjør Hobbes såpass interessant å se på i forbindelse med spillteori; hans individer kan ansees som nytte-maksimerende, og selv om han legger et svært komplekst resonnement til grunn for menneskelig motivasjon, er det langt mer systematisk og “enkelt” enn mange alternativer, og det egner seg derfor i mine øyne godt til formell modellering<sup>4</sup>.

Forfølgelse av egeninteresse er heller ikke en klanderverdig defekt hos mennesket for Hobbes, men det naturlige, og dermed det korrekte (Hobbes 1946:124,166)<sup>5</sup>.

---

<sup>4</sup>Nytte-maksimering og noen alternativer til dette gjennomgås under premiss 2 i 4.2.2.

<sup>5</sup>For en mer utfyllende behandling av nettopp dette se kapittel 3 og Oakeshott (1946:liv).

### 2.3.4 Negativ nytte - *summum malum*

And lastly the motive, and end for which this renouncing, and transferring of right is introduced, is nothing else but the security of a man's person, in his life, and in the means of so preserving life, as not to be weary of it (Hobbes 1946:87).

Hobbes' "negative" innstilling til mennesker og vår tilstand gjør det interessant å se på hvorvidt synet på *nytte* og *det gode* som noe utelukkende positivt (eller aktivt) muligens bør revurderes. Det kan selvsagt ikke benektes at det gode er godt – noe annet ville være absurd – men det er allikevel grunn til å se på i hvor stor grad *negativ nytte* kan være et interessant og nødvendig begrep for å fange opp nyansene i Hobbes' filosofi<sup>6</sup>.

[...] it may be observed that death, the involuntary cessation of desire and the pursuit which is the end of desire, is the thing of all others the most hateful; it is the *summum malum* (Oakeshott 1946:xxxiv).

Det ultimate ondet er å *ikke* leve, og utgangspunktet er, som de fleste forhåpentligvis er klar over, å leve. Det kan derfor være logisk å anse en fortsettelse av det som *er*, som det gode, og lykke, mens en endring i tilstanden er det onde.

En interessant følge av dette er at samtlige, uten unntak, vil oppleve det ultimate ondet: vi kommer alle til å dø. Dette gjør Hobbes' filosofi ganske interessant, siden også denne siden av hans tenkning er svært "mørk", eller ultrarealistisk om man vil. Mens andre kan synes å være opptatt av jakten på en opplyst tilstand, som for mange ser ut til å kunne føre til himmelen på jord for dem som oppnår den, er Hobbes opptatt av en kontinuerlig kamp for oppnåelse av de grunnleggende godene i livet, der døden ikke er en naturlig og nøytral hendelse, men *summum malum* – det verste som kan skje. Igjen, det er forståelig at en slik filosofi vekker avsky, eller fornektelse, hos mange,

---

<sup>6</sup>Kahneman og Tversky (1990) viste med sin *prospektteori* at det er godt belegg for å hevde at mennesker forholder seg forskjellig til utfall som gir tap og utfall som gir gevinst. Jeg vil ikke gå grundig inn på prospektteorien i denne oppgaven, men anser det som viktig å bemerke at det finnes visse begrensinger ved teorien om *forventet nytte*. Morrow (1994:48) påpeker at nyttefunksjoner kan gi mer risikovillighet for tap enn gevinster, men der man i prospektteorien har et rammeverk som lar nivåene mennesker anvender for å skille gevinst og tap variere (ut fra *rammeeffekter*), finnes ikke dette i standard nytte teori. Prospektteorien har imidlertid også problemer, spesielt forbundet med det faktum at den er langt mer komplisert enn standard nytte teori. Alt i alt fremstår det som tillatelig å hovedsakelig anvende standard nytte teori, men det er altså visse problemer man bør være oppmerksom på (Morrow 1994:48).

men ser man bak negativiteten i presentasjonen av Hobbes' resonnement, ser man at kjernen i hans filosofi hverken er mørk eller umenneskelig – den er realistisk. Hobbes' aksept av menneskets tilstand fører dog muligens til at argumentasjonen oppfattes noe annerledes for de som har et annet syn på mennesket og dets muligheter.

For Hobbes er *krig*, og spesielt borgerkrig, det verste mennesker kan oppleve, men det er nærliggende å anta at det er effektene, nemlig “kontinuerlig frykt, og fare for en voldelig død”, og livet som “ensomt, fattig, rått og kort”, som er det ytterste ondet (Hobbes 1946:82,120,219).

Of things held in propriety, those that are dearest to a man are his own life, and limbs; and in the next degree, in most men, those that concern conjugal affection; and after them, riches and means of living (Hobbes 1946:223).

Hobbes' syn på det verste ondet behandles videre i kapittel 4 og 5.

### 2.3.5 Oppsummering

Det er bragt på det rene at Hobbes forutsetter at individer er grunnleggende egeninteresserte, men at dette ikke resulterer i en forståelse av mennesker som kortsiktig egoistiske og ute av stand til å dra nytte av andres suksess (mer om dette i neste seksjon).

Nytte, som definert her, er dog forholdsvis vagt, og en “svakhet” er at det i tråd med Hobbes' filosofi utelukkende er basert på en subjektiv vurdering av et utfall. Nytte er fremdeles et vidt begrep, og en representasjon av nytte ved  $U_i$  i en nyttefunksjon må forståes som individets totale vurdering av et resultat, og ikke som en enkel funksjon av monetær gevinst e.l.

Videre er det som nevnt ovenfor store forskjeller på formålet med en nyttefunksjon; dersom noe skal predikeres, må det foretas en analyse av hva som er det *tilsynelatende gode* for aktøren, mens man ellers ofte kan se på det oppnådde godet dersom man analyserer post factum.

En tentativ definisjon av nytte for Hobbesianske individer kan formuleres som følger: *Nytte for Hobbes er oppnåelse av et gode nå, eller muligheten for å skaffe seg goder i fremtiden.*

## 2.4 Andres og egen nytte

Etter å ha sett på hva som er nytte for en aktør, er det viktig å se på hvordan *andre aktørers nytte* vurderes i den opprinnelige aktørens totale nyttekalkulasjon. Mange unnlater å behandle dette problemet overhodet, noe som etter min oppfatning gir svært mangelfulle vurderinger. Andre, derimot, er svært opptatt av dette, og har laget modeller der andres nytte faktisk i visse tilfeller tillegges *mer* vekt enn egen nytte. Jeg mener bestemt at andres nytte bør tas inn i en kompleks Hobbesiansk nyttefunksjon, da det har stor forklaringskraft, og Hobbes diskuterer selv dette spørsmålet i flere tilfeller. Det interessante med Hobbes er at han er meget bestemt på at formålet for enhver frivillig handling er egen fordel, samtidig som han har en helt klar formening om at andres nytte svært ofte, om ikke alltid, blir vektlagt.

### 2.4.1 Kunnskap om andres motivasjon og preferanser

Fremfor alt er det interessant å se hva vi kan *vite* om andres nytte. Introspeksjon vil ifølge Hobbes gi oss et godt grunnlag for å avdekke andres *motivasjon*. Dette fordi våre drifter er forholdsvis like, som “begjær, frykt, håp osv.” (Hobbes 1946:6). Det er imidlertid stor forskjell på *hva* mennesker begjærer, frykter og håper på, sier Hobbes, så det er derfor liten grunn til å anta at vi kan avdekke andres *preferanser* ved å se på våre egne (Hobbes 1946:6).

Om målene for våre drifter:

[. . .] for these [våre drifter] do so vary, and they are so easy to be kept from our knowledge, that the character of man’s heart, blotted and confounded as they are with dissembling, lying, conterfeiting, and erroneous doctrines, are legible only to him that searcheth hearts (Hobbes 1946:6).

Det er altså mulig å skjule våre preferanser bevisst ved bedrag, eller uforskyldt ved å generelt ha lite kunnskap om seg selv, verden eller hvordan ting fungerer og henger sammen. Selv om det er vanskelig, om ikke umulig, å lære andre fullstendig å kjenne, er ønsket om dette, og nysgjerrigheten som fører til at vi prøver, uunngåelig. Hobbes sier relativt lite direkte om behovet for kunnskap om andres motivasjon og preferanser, men den situasjonen han beskriver, med konkurranse om knappe goder, gjør det klart at kunnskap om andre gir makt. Å anta at Hobbes ser verdien i å ha kunnskap om andres

ønsker og intensjoner, synes for meg å være svært lite dristig, og kan også forsvares f.eks. ved å vise til at Hobbes anser spionasje som en viktig del av krigføring, og dermed en plikt for statsoverhodet – suverenen (Hobbes 1946:83, Hobbes 1991b:261-262). Kort sagt er det opplagt at informasjon om andres intensjoner gjør en bedre i stand til å forsvare seg, det gjør en i stand til å vanære andre dersom en kjenner en persons hemmeligheter, man vil lettere kunne bedra noen man har god kunnskap om, og en vil være i stand til å skaffe seg støtte ved å kunne tilby andre det de ønsker.

Det er imidlertid også et svært positivt element relatert til fullstendig informasjon<sup>7</sup> og heterogenitet: dersom vi ønsker forskjellige ting, kan vi ved å forhandle finne løsninger som er bra for flere, selv med knappe goder. Hobbes diskuterer dog ikke denne muligheten.

## 2.4.2 Variasjoner i nyttevurderinger og nyttesammenligning

Det er grunnleggende forskjeller i hvordan man vurderer det gode. Dette gjør det nødvendig å ta stilling til hvordan man skal vurdere andres tildelinger i aktuelle situasjoner. Skal jeg vurdere din nytte etter hva *du* synes om det du har fått, eller etter *min* vurdering? Eller skal kanskje din posisjon i utfallet vurderes ut fra min vurdering av din vurdering av ditt resultat? Alle disse variantene kan forsvares med plausible begrunnelser, og det gjenstår så å se hvilken fremgangsmåte som er mest i samsvar med Hobbes' filosofi.

From the same it proceedeth, that men give different names, to one and the same thing, from the difference of their own passions: as they that approve a private opinion, call it opinion; but they that mislike it, heresy: and yet heresy signifies no more than private opinion; but has only a greater tincture of choler (Hobbes 1946:67).

Selv om våre drifter er *sammenlignbare*, er Hobbes meget tydelig på at også de varierer (Hobbes 1946:6,46,67). Årsakene han oppgir til disse forskjellene er fysisk kroppslig forfatning, forskjeller i dannelselse, karakter og kultur. Årsakene er altså av to sorter: naturlige og tillærte (Hobbes 1946:46). Selv om jeg påvirkes av kulturen jeg tilhører, oppveksten mine foreldre gir meg og utdannelsen jeg får, vil det alltid være en del

---

<sup>7</sup>Her brukt som den spillteoretiske betegnelsen på en situasjon der spillere kjenner hverandres preferanser og strategier (Hovi 2008:31).

utenfor min eller andres kontroll, som bestemmes av mine gener og min fysiske sammensetning. Driftene varierer altså, men dette kan tenkes å gi sterkest utslag i *hvor sterkt* vi ønsker eller søker oss bort fra ting, ikke *hva* vi trakter etter.

Våre mål, appetitter og aversjoner er mer relevante enn drifter i forbindelse med vurdering av nytte, da det er disse som direkte avgjør målene for våre ønsker. I likhet med våre drifter varierer disse på bakgrunn av “individuell forfatning og dannelselse”, men de varierer ifølge Hobbes langt mer enn våre drifter, og som vi har vært inne på er disse også vanskeligere å avdekke i andre (Hobbes 1946:6,32). Et individs *forfatning* er forøvrig ikke konstant, men i “kontinuerlig mutasjon” slik at appetitter og det gode ikke engang er stabilt i ett enkelt individ (1946:32,104). At det er forskjeller individer imellom hva angår nyttevurderinger er altså helt tydelig, og når selv ett enkelt individ kan ha problemer med å være enig med seg selv over tid, finnes det “nær sagt intet objekt alle kan bli enige om som godt og ønskelig” (Hobbes 1946:32).

“En enkel bonde er klokere når det gjelder egne anliggender, enn hva kongelige rådgivere er når det gjelder andres”, sier Hobbes, og jeg er da tilbake til spørsmålet om hvordan andres nytte skal vurderes (Hobbes 1946:45). Først av alt er det her snakk om generelle nyttefunksjoner og nyttevurderinger. Hobbes gjør det klart at meglere, dommere og andre som skal løse andres konflikter, er forpliktet til å vise upartiskhet (Hobbes 1946:101,102). Det samme gjelder nødvendigvis ikke for individer generelt, da det er etablert at *egen fordel* er individers målsetning, og at oppnåelsen av goder nå eller i fremtiden er den direkte målestokken for nytte.

At egen nytte tillegges mest vekt i en situasjon synes opplagt, men dersom man forsøker å skille ut effekten av andres situasjon på egen nytte, som et eget ledd i en nyttefunksjon, er det interessant å se hvordan dette praktisk sett kan tenkes vurdert. I forbindelse med individers nyttefunksjoner synes det klart at en nyttefunksjon basert på Hobbes’ filosofi må inkludere en vurdering av andres situasjon – ikke utelukkende egen gevinst, som er tilfellet for *Homo economicus*. Dette innebærer som vi vil se i 3.4.3 at aktørene har *sosiale preferanser*.

Når det gjelder spørsmålet om hvorvidt Hobbesianske individer misliker ulikhet, er dette på mange måter et svært enkelt spørsmål; *relativ makt* og *eminens* er det viktige, så jo mer jeg har, jo bedre, og vice versa med din gevinst. (Jeg kommer snart tilbake til en mer utfyllende vurdering av relativ gevinst hos Hobbes.) Det er dog, spesielt i naturtilstanden, også en fare forbundet med å ha langt mer enn andre:

Also riches joined with liberality, is power; because it procureth friends, and servants: without liberality, not so; because in this case they defend not; but expose men to envy, as a prey (Hobbes 1946:56).

Midler gir en mulighet til å erverve tjenere, venner og allierte, så ulempen ved at man blir et mål for andre kan oppveies ved at man kan *kjøre* seg mer sikkerhet. Et interessant spørsmål er her hva det koster å kjøpe andres støtte. Det synes usannsynlig at full og pålitelig støtte koster særlig mye mindre enn det aktøren anser at den er verdt, så det *kan* tenkes at noe mindre midler i utgangspunktet kunne være å foretrekke. Det er imidlertid svært lite støtte i Hobbes' tekster for at det er et grunnleggende ønske for mennesker å være like andre, i den forstand at de skulle mislike å ha *mer* enn andre. Mennesker er like, men de anser hverandre ikke som like, og nettopp derfor omtaler naturlovene (nr. 9 og 10) dette spesielt (Hobbes 1946:80,100,101,120)<sup>8</sup>.

### 2.4.3 Godhet og veldedighet

One time, I remember, goeing in the Strand, a poore and infirme old man craved his [Hobbes'] Almes. He, beholding him with eies of pittie and compassion, putt his hand in his pocket and gave him 6d. Sayd a Divine that stood by, Would you have donne this, if it had not been Christ's command? Yea, sayd he. Why? quoth the other. Because, sayd he, I was in paine to consider the miserable condition of the old man; and now my almes, giving him some relief, doth also ease me (Aubrey 1999:157).

Et annet punkt som ofte bringes opp er at Hobbes beskriver *godhet* og *veldedighet* som "ønsket om godt for en annen", og at man har *god natur* dersom dette ønsket gjelder generelt for alle andre (Hobbes 1946:34). Hans definisjoner blir av mange tolket som at det finnes en slags "altruisme" hos Hobbes' individer, eller i alle fall en viss tilbøyelighet til *ikke* utelukkende å handle etter egen fordel. Dersom man leser hans andre definisjoner i samme kapittel, er det imidlertid lett å se at sorg, medfølelse, medmenneskelighet og andre størrelser utelukkende beskrives som et resultat av effekt på egen fordel. Selv om dette ikke nevnes direkte når det gjelder godhet, synes det plausibelt å anta at det samme er tilfellet også her, dersom man ser på Hobbes' andre tekster (Hobbes 1946:37).

---

<sup>8</sup>En mer inngående behandling av ulikhetsaversjon, som benyttet i eksperimentell økonomi, vil foretas i kapittel 3.



Det finnes dog ubetinget kjærlighet hos Hobbes; personer som elsker mennesker generelt betegnes som *snille*, og i *The Elements of Law* (1999)<sup>9</sup> omtaler Hobbes kjærlighet mellom bl.a. foreldre og barn som noe som nærmer seg betingelsesløs og *ekte* godhet/veldedighet (Hobbes 1946:35, 1999:56). Mennesker har *affeksjoner*, og “det antas alltid at jo nærere i slekt, jo mer affeksjon”, noe som bl.a. fører til at Hobbes mener suvereren vil velge nære slektninger som arvtagere fremfor andre der det er mulig (Hobbes 1946:128, 1999:56). Også dette er beheftet med noe usikkerhet; det er usikkert hvor genuint fri for egeninteresse denne affeksjonen er, da “det er opplagt at mennesker alltid mottar mest reflektert ære fra sine næreste slektninger” (Hobbes 1946:128). Den samme konklusjonen – at selv genuin familiær affeksjon er svært vanskelig å skille fra egeninteresse – finner vi når Hobbes snakker om ettermæle; man vil ikke oppleve sitt eget ettermæle på noe vis, men det er likevel ikke bortkastet å søke ære etter døden. Dette forklares ved at vi fryder oss ved å tenke på det, bl.a. på grunn av den fordel det kan gi våre etterkommere (Hobbes 1946:64). I teksten synes det klart at det er fryden ved tanken på denne æren som er avgjørende. Det er imidlertid et komplisert skille, siden fryden skyldes forventet fordel for våre etterkommere. Mer om dette følger under punktet om langsiktighet.

Det som er klart, er at når det selv i forbindelse med den næreste slekt i stor grad er egeninteresse som er avgjørende for våre handlinger, er det helt klart tilfellet når det gjelder andre. Videre, om jeg f.eks. gleder meg ved å gjøre noe godt for deg, kan dette neppe brukes som et argument *mot* egeninteresse som fundamental drivkraft.

There is yet another passion sometimes called love, but more properly good will or charity. There can be no greater argument to a man of his own power, than to find himself able, not only to accomplish his own desires, but also to assist other men in theirs: and this is that conception wherein consisteth charity (Hobbes 1999:56).

For no man giveth, but with intention of good to himself; because gift is voluntary; and of all voluntary acts, the object is to every man his own good [...] (Hobbes 1946:99).

Her blir altså veldedighet, og faktisk også “kjærlighet” generelt sett, forklart utelukkende ut fra egeninteresse; det å være “god”, og foreta veldedige handlinger, er et tegn på styrke og en måte å signalisere dette til andre på, noe som gir ære (Hobbes 1946:35,59).

---

<sup>9</sup> *The Elements of Law* består av bøkene “Human Nature” og “De Corpore Politico”.

Foruten å signalisere ære gir det en aktør *glede* å kontemplere sin egen makt, noe som er et umiddelbart gode, siden mennesker “kontinuerlig konkurrerer om ære og verdighet” (Hobbes 1946:111). Dette er nært knyttet til diskusjonen om *oppfattet makt*, som jeg kommer tilbake til i 2.4.6. Kort fortalt kan det tenkes at vellykket *signalisering* av makt faktisk *gir* makt, og at godhet og veldedighet på denne måten kan anses som rent egeninteresserte handlinger for å skaffe seg makt.

Det samme gjelder en av effektene av *godhet, vennlighet og veldedighet*: det gir venner og skaper forpliktelser, noe som utvilsomt er instrumentell makt for Hobbes, og er dermed nok en faktor som gjør godhet til noe som kan være motivert av ren egeninteresse (Hobbes 1946:56).

En som mottar veldedighet, eller generelt sett befinner seg på den svakeste siden i et usymmetrisk bytteforhold, er kanskje den beste dommeren hva angår forekomsten av “altruistisk veldedighet”:

To have received from one, to whom we think ourselves equal, greater benefits than there is hope to requite, disposeth to counterfeit love; but really secret hatred; and puts a man in the estate of a desperate debtor, that in declining the sight of his creditor, tacitly wishes him there, where he might never see him more. For benefits oblige, and obligation is thralldom; and unrequitable obligation perpetual thralldom; which is to one's equal, hateful (Hobbes 1946:65).

To give little gifts, is to dishonour; because it is but alms, and signifies an opinion of the need of small help (Hobbes 1946:58).

At “veldedighet” *kan* forklares med egeninteresse behøver ikke å innebære at aktørene involvert i situasjonen oppfatter det slik; én ting blir imidlertid klart: *transaksjoner er ikke nødvendigvis usymmetriske, selv om de muligens kan oppfattes slik*. Dette tolker jeg til å underbygge Hobbes' påstand om at veldedighet, som alle andre handlinger, for giver i realiteten er en balansert transaksjon, ved at han gleder seg selv ved at han hadde muligheten til å gi av sin overflod, han anser det muligens som gunstig for seg selv at mottager havner i en noe forbedret situasjon, og ikke minst står mottager – om han vil det eller ei – til en viss grad i gjeld til giver. Det er altså mer symmetri enn hva som oppfattes på det monetære, eller varebaserte, nivået; fysiske goder kan byttes mot glede, ære og makt.

But the affection wherewith men many times bestow their benefits on strangers, is not to be called charity, but either contract, whereby they seek to purchase friendship; or fear, which maketh them to purchase peace (Hobbes 1999:56).

#### 2.4.4 Rettferdighet og moral

Justice, and injustice, are none of the faculties neither of the body, nor mind. If they were, they might be in a man that were alone in the world, as well as his senses, and passions. They are qualities, that relate to men in society, not in solitude (Hobbes 1946:83).

Da egeninteressen så langt må sies å stå støtt som menneskers hovedmotivasjon i Hobbes' filosofi, gjenstår det å vurdere hvorvidt moralske retningslinjer kan sies å øve særlig innflytelse på våre aktører. Først og fremst er det nødvendig å se på hva moral er for Hobbes:

For moral philosophy is nothing else but the science of what is *good*, and *evil*, in the conversation, and society of mankind (Hobbes 1946:104).

Det mennesker i naturtilstanden kan enes om som godt – *fred* – er det avgjørende for hva som er samfunnets gode, og dermed også det som avgjør hva som kan betegnes som moralske dyder, nemlig “*rettferdighet, takknemlighet, beskjedenhet, rimelighet, nåde*” (Hobbes 1946:104). “Vitenskapen om dyder og laster er moralfilosofi. Derfor er doktrinen om naturlovene den sanne moralfilosofi” (Hobbes 1946:104). Mens det er heftige debatter om Hobbes' teori om moral og forpliktelse, vil jeg i denne oppgaven nøye meg med en forståelse av forpliktelse hos Hobbes som i all hovedsak består i å holde avtaler. Med samfunnskontrakten i kraft innebærer dette også en forpliktelse overfor naturlovene, og dermed også *fornuft*, så implikasjonene av denne forståelsen kontra andre tolkninger er forholdsvis små.

Justice therefore, that is to say, keeping of covenant, is a rule of reason, by which we are forbidden to do any thing destructive to our life; and consequently a law of nature (Hobbes 1946:96).

Naturlovene gir personer retningslinjer for deres adferd; disse må følges dersom en vil unngå naturtilstanden, noe man trygt kan anta at mennesker ønsker (Hobbes

1946:87,109). Først og fremst pålegges vi å søke fred der vi ser muligheter for å oppnå dette. For det andre må vi si oss villige til å legge ned vår rett til alt, for så å si oss fornøyde med den graden av frihet vi er villige til å gi andre (Hobbess 1946:85).

Whensoever a man transferreth his right, or renouceth it; it is either in consideration of some right reciprocally transferred to himself; or for some other good he hopeth for thereby. For it is a voluntary act: and of the voluntary acts of every man, the object is some *good to himself* (Hobbess 1946:86).

Videre krever naturlovene at vi holder avtaler (naturlov nr. 3), at vi er takknemlige (4), at vi er imøtekomende og tilpasser oss (5), at vi viser nåde (6), at fremtidig gode, og ikke hevngjerrighet, er vårt motiv i gjengjeldelse (7), at vi ikke krenker hverandre (8), at vi anser hverandre som like av natur (9), at vi er moderate og ikke krever flere rettigheter enn andre (10), at de som dømmer, ikke er partiske (11), at fellesgoder deles på en god måte (12, 13 og 14), at fredsmeglere vernes og får ferdes fritt (15) og at de med uenighet legger saken frem for en habil og nøytral megler, som dømmer via en god prosess (16-19) (Hobbess 1946:93,99-102).

At man bør følge naturlovene begrunnes utelukkende med fornuft og egen fordel, og det er uunngåelig å vende tilbake til Hobbess' poeng om at respekten for disse lovene begrunnes med egeninteresse (Hobbess 1946:95,96). Om noen tror at egen fordel taler for å bryte naturlovene, beror dette ifølge Hobbess på en misforståelse, siden naturlovene er det ytterste forsvar for enhver persons egen nytte, "for det er utenkelig at krig skal bevare liv, og fred ødelegge det" (Hobbess 1946:104).

Naturlovene kan definitivt sies å være en ledetråd for menneskelig adferd – en effektiv sådan – men å fremme dette som et bevis på at egeninteresse ikke er den eneste motivasjon er å misforstå Hobbess' generelle filosofi og hans naturlover.

Selv om "naturlig" moral ifølge Hobbess er svært begrenset, er det forsåvidt åpent for at individer tilskriver seg en annen og mer omfattende "moral". Kristendommen er her et godt eksempel. De som tror på og/eller følger dens forskrifter og anbefalinger, kan sies å følge en noe annen moral enn de som ikke bryr seg nevneverdig om disse.

Moral: concerned with the principles of right and wrong behavior and the goodness or badness of human character [...] *moral* refers to generally accepted standards of goodness and rightness in character and conduct

(New Oxford American Dictionary).

Det blir et spørsmål om hvordan man bruker ordet moral, og dersom standardene må være universelt godtatt, er det utelukket å diskutere “flere moraler”. Hobbes synes å være av den oppfatning at det kun finnes én moral: den som springer fra naturlovene. Når Hobbes så sier at Gud gjør mennesker i stand til å være dydige, både moralsk og intellektuelt, finnes det ingen tegn på at disse dydene er forskjellige fra de dydene som foreskrives av naturlovene (Hobbes 1946:280).

### 2.4.5 Relativ gevinst

VIRTUE generally, in all sorts of subjects, is somewhat that is valued for eminence; and consisteth in comparison. For if all things were equal in all men, nothing would be prized (Hobbes 1946:42).

Avgjørende for frekvensen av “gode handlinger” hos aktører som primært antas å være ute etter å sikre egen fordel, er forekomsten av *knappe goder* og *relativ gevinst/makt*. Jeg kommer nærmere tilbake til dette senere i oppgaven, men jeg ønsker allikevel å starte på temaet her, da det er sentralt i forbindelse med spørsmålet om hvordan mennesker vurderer andres nytte. Som sitatet ovenfor tydeliggjør, er det *eminens* som blir verdsatt når det gjelder dyder. Dette er et eksempel på *relativ makt* – makt som er verdifull kun fordi jeg har mer makt enn deg.

*Natural power*, is the eminence of the faculties of body, or mind: as extraordinary strength, form, prudence, arts, eloquence, liberality, nobility (Hobbes 1946:56).

*Absolutt nivå* er i mange tilfeller irrelevant, for uavhengig av hvor mye makt jeg har, vil du være mektigere dersom du har mer. Det samme gjelder *samlet makt*, som i allianser; “antallet som er tilstrekkelig for å være sikre, er ikke et bestemt antall, men avgjøres i sammenligning med motstanderen vi frykter” (Hobbes 1946:110). Dette begrepet er viktigst i konkurransesituasjoner; med et uendelig gode som mål ville f.eks. et absoluttnivå av makt som er tilstrekkelig for å utvinne en ressurs, være nok. Dersom det derimot er tale om et knapt gode, vil jeg i tillegg til å skulle utvinne ressursen, måtte kjempe mot andre om den.

Competition of riches, honour, command, or other power, inclineth to con-

tention, enmity, and war: because the way of one competitor, to the attaining of his desire, is to kill, subdue, supplant, or repel the other (Hobbes 1946:64).

Dette sitatet gjør det klart at rikdom, ære, autoritet og “annen makt” er knappe goder. For at jeg skal kunne skaffe meg disse, må andre nødvendigvis få mindre av dem. Dette er et svært sentralt poeng for Hobbes; mennesker er forholdsvis like i evnen til å utgjøre en trussel for hverandre, og siden vi ofte ønsker de samme tingene, og har relativt like forhåpninger om å oppnå dem, blir vi fiender (Hobbes 1946:64,80,81). “[P]å veien til målet, som i hovedsak er egen overlevelse, og noen ganger kun nytelse, søker vi å ødelegge og dominere hverandre” (Hobbes 1946:81).

Mennesker skiller seg her fra andre “sosiale dyr”:

[...] amongst these creatures, the common good differeth not from the private; and being by nature inclined to their private, they procure thereby common benefit. But man, whose joy consisteth in comparing himself with other men, can relish nothing but what is eminent (Hobbes 1946:111).

Som vist i avsnittet om forskjellige nytteverdier, kan betraktningene om relativ gevinst i visse situasjoner dempes av affeksjon for andre. Også spørsmål om hvorvidt andres styrke gir en selv makt vil her være sentralt, noe som kan være tilfellet i forbindelse med ens herskere, partnere, venner osv.

## 2.4.6 Innstilling til og vurdering av andre

Thomas Hobbs said that if it were not for the gallows, some men are of so cruell a nature as to take a delight in killing men more than I should to kill a bird (Aubrey 1999:157).

*Desire* of riches, COVETOUSNESS; a name used always in signification of blame; because men contending for them, are displeas'd with one another attaining them; though the desire in itself, be to be blam'd, or allow'd, according to the means by which these riches are sought (Hobbes 1946:34).

For å kunne godta denne “kyniske” fremstillingen av menneskelige relasjoner, er det viktig å se på hvordan andre individer – og *et menneskeliv* – verdsettes av Hobbes' individer. Dersom man leser Hobbes med en forestilling om at individet er ukrenkelig,

at et liv er uvurderlig, e.l., vil man enten ikke forstå, eller rett og slett forkaste, Hobbes' slutninger. Det er av fundamental betydning hvilken verdi hans individer legger i andres liv og velferd, og mens vi har sett at andres velferd nødvendigvis ofte må ofres for å oppnå egen fordel, gjenstår det å se om det finnes noen fundamental forskjell mellom synet på andres *liv* og *eiendeler*. Samtidig er det interessant å se hva Hobbes antar at vi føler for hverandre, hva slags *generell innstilling* mennesker har til hverandre.

And to consider the contrariety of men's opinions, and manners, in general, it is, they say, impossible to entertain a constant civil amity with all those, with whom the business of the world constrains us to converse: which business consisteth almost in nothing else but a perpetual contention for honour, riches, and authority (Hobbes 1946:460).

Hobbes hevder at der en angriper kun har én persons makt å frykte, vil man antageligvis måtte anta at angriperen, i fellesskap med andre, vil tilegne seg denne personens eiendeler. Ikke nok med det, Hobbes hevder videre at overtrederne også vil kunne ta livet av personen, eller frarøve ham friheten (Hobbes 1946:81). Ved å frarøve en person hans frihet må det antas at angriperen skaffer seg herredømme over personen, noe som igjen gir angriperen mer makt. Det interessante i dette utsagnet er imidlertid at det ikke synes å være noen hindring, annet enn makten hos den som forsvarer seg, for å ta et liv. Hvorvidt denne oppfatningen – at liv kan tas uten særlige skrupler – er basert på en reell tilbøyelighet hos mennesker, eller en irrasjonell frykt for slike mennesker, er noe mindre klart. Det er i alle fall klart at mennesker *antar* at det finnes andre som er villige til å drepe dem, og at “en person ikke kan avdekke hvorvidt en person som nærmer seg med voldelige intensjoner har til hensikt å drepe ham eller ikke” (Hobbes 1946:86). De som faktisk handler på en slik måte, og dreper andre mer eller mindre vilkårlig, betegnes dog av Hobbes som *ondskapsfulle*; det å skade noen uten grunn leder ifølge Hobbes til krig, som igjen er imot naturlovene, så det er helt klart klanderverdig og *ufornuftig* adferd (Hobbes 1946:96,100). Det er her viktig å påpeke at *ondskapsfulle* mennesker ikke regnes som ondskapsfulle av Hobbes fordi de får *glede* av andres ulykke; “Forakt, eller liten interesse, for andres ulykke, er hva menn kaller GRUSOMHET” (Hobbes 1946:37). Det er altså mangel på sympati som gjør mennesker ondskapsfulle, det er en passiv attributt. Det er dette som gjør noen i stand til å frarøve andre deres liv, frihet og eiendeler, ikke et aktivt *ønske* om å gjøre onde ting.

Let him therefore consider with himself, when taking a journey, he arms

himself, and seeks to go well accompanied; when going to sleep, he locks his doors; when even in his house he locks his chests; and this when he knows there be laws, and public officers, armed, to revenge all injuries shall be done him; what opinion he has of his fellow-subjects, when he rides armed; of his fellow citizens, when he locks his doors; and of his children, and servants, when he locks his chests. Does he not accuse mankind by his actions, as I do by my words? (Hobbes 1946:82)

Mennesket er ikke et sosialt dyr i den forstand at vi gledes av hverandres nærvær, ifølge Hobbes; tvert imot, det er en stor utfordring, spesielt i naturtilstanden (Hobbes 1946:81). Her dreier det seg bl.a. om knappe goder og *konkurransen*, som igjen fører til *usikkerhet*, og man har også spørsmålet om *ære* (Hobbes 1946:81).

“Enhver person søker å bli verdsatt av andre slik han verdsetter seg selv”, sier Hobbes, og videre, “når man ser hvilken verdi mennesker tilskriver seg selv, hvor mye respekt de ønsker fra andre, og hvor lite de verdsetter andre mennesker”, er det ikke overraskende at det oppstår konflikt (Hobbes 1646:81,118). Søken etter ære er utstrakt, og fører i siste instans til krig; ære fører til at individer bruker vold “over bagateller, som et ord, et smil, forskjellige meninger, og enhver annen form for nedverdiggelse, enten direkte på person eller indirekte på slekt, venner, nasjonen, deres yrke eller deres navn” (Hobbes 1946:81). Dette gjør det nødvendig å vurdere hvorvidt oppfattet “verdi” utgjør reell makt. Dersom så ikke er tilfellet må de som kriger og slåss over ære sies å være svært lite rasjonelle i henhold til Hobbes’ syn på det verste ondet og menneskers søken etter sitt eget beste – de risikerer i så fall livet uten grunn. Det som derimot synes mest plausibelt, er at en persons makt i stor grad finnes i andres oppfatninger om en selv, og at det å bli vanæret og nedverdiggelst, om enn kun verbalt, i så henseende helt klart kan oppfattes som et reelt tap av makt, i alle fall der dette skjer offentlig. Ære får mennesker til å angripe for å skape seg et *rykte*, eller *omdømme*, og når man ser Hobbes’ definisjon av *instrumentell makt*, blir det helt klart at rykte faktisk kan være makt: “Å ha ord på seg for å være mektig, er makt, fordi det hjelper til med å skaffe en person støtte fra dem som trenger beskyttelse” (Hobbes 1946:56,81).

Likheten i evnen til å drepe hverandre kan også understøtte behovet for et *rykte* om makt. (Hobbes 1946:80). “[F]or i naturtilstanden, kan forskjeller i styrke kun avgjøres ved kamp”, og “det er ikke alltid tydelige nok forskjeller i styrke eller klokskap mellom mannen og kvinnen til at retten [her til avkom] kan avgjøres uten krig” (Hobbes



1946:92,131). Det sistnevnte sitatet er noe malplassert, siden det helt tydelig omtaler forskjellen mellom mann og kvinne i forbindelse med rett til sitt avkom. Det underbygger allikevel det samme som det foregående sitatet: ved små fysiske forskjeller kan det være umulig å vite hvem som vil vinne uten kamp. Siden kamp er svært risikabelt, vil et *rykte* om makt kunne tenkes å være tilstrekkelig til å sikre seier, der motstanderen anser sannsynligheten for at han vil tape som for stor til at han vil risikere livet i kamp.

And because all signs of hatred, or contept, provoke to fight; insomuch as most men choose rather to hazard their life, than not to be revenged; we may in the eight place, for a law of nature, set down this precept, *that no man by deed, word, countenance, or gesture, declare hatred, or contempt of another* (Hobbes 1946:100).

Ingen av de tre opprinnelige grunnene til konflikt er *makt*, men makt er avgjørende i forbindelse med alle tre. “Den første [konkurranse] får mennesker til å angripe for gevinst, den andre [usikkerhet], for sikkerhet, og den tredje [ære], for omdømme” (Hobbes 1946:81).

Hobbes anser personers *verdi* å utelukkende være en “markedspris” for deres makt, som bestemmes av potensielle kjøpere. Han skiller seg dermed radikalt fra mange andre filosofer, som fremsetter en iboende selvstendig verdi i mennesker (Hobbes 1946:57). Makt er her “valutaen” for vår verdi, og verdien er heller ikke absolutt, siden den “avhenger av andres behov og vurderinger” (Hobbes 1946:57). Det er meningsløst å snakke om min verdi uten at det er snakk om min verdi for *andre*; “det er med menn som med andre ting; det er ikke selgeren, men kjøperen som bestemmer prisen. For la gjerne en mann verdsette seg selv så høyt han bare kan, noe de fleste gjør, deres verdi er uansett ikke noe annet enn hvor høyt andre verdsetter dem” (Hobbes 1946:57). Denne verdsettelsen av andre medfører enten å ære eller å nedverdige dem. Hva som blir tilfellet, avhenger av om vår verdsettelse er høyere eller lavere enn personens verdsettelse av seg selv (Hobbes 1946:57). Å *fortjene* noe er utelukkende bestemt av å ha rett til noe; man har ikke en iboende *rett* til noe annet enn hva lover eller kontrakt har lovet oss, noe som igjen går på at mennesket i en stat ikke har rett til annet enn hva staten gir det (Hobbes 1946:63). “Han som innfrir sin forpliktelse i en kontrakt først, *fortjener* det han er ment å motta ved at den andre innfrir sin forpliktelse, og har dette som *utestående*” (Hobbes 1946:89). *Verdighet* er dog noe annet, og går på hvorvidt man er *egnet*

for eksempel til visse posisjoner på bakgrunn av evner man besitter (1946:62).

### 2.4.7 Oppsummering

Mennesker interagerer, men siden de konkurrerer, og relativ makt og relativ gevinst er sentrale begreper for dem, har de ikke råd til å unne hverandre alt godt. Det er vanskelig å avdekke hverandres preferanser, men ved introspeksjon kan man lære *noe*. Det ligger imidlertid mye makt i å vite hva andre tenker og ønsker, og det er derfor sannsynlig at mennesker ikke er fullstendig åpne i signaliseringen av sine preferanser og intensjoner. “Spionasje” er en viss mulighet, men som i krig mellom stater, er det muligens også i mellommenneskelige relasjoner en fiendtlig handling å søke å avdekke andres motiver, planer og mål mot deres vilje.

Vi *ønsker* ikke å drepe hverandre, men det finnes mennesker som ikke har noe imot å ta andres liv. Siden det kan være mye å vinne på å ta et liv og tilegne seg en persons eiendom, er det derfor ikke nødvendig med aktiv ondskap for at menneskers situasjon blir vanskelig, usikker og voldelig. Moral i moderne forstand, og en høytsvevende forståelse av abstrakte prinsipper som rettesnor for menneskelig handling, er svært fremmed for en Hobbesiansk tenker. Men, naturlovene er fornuftslover som krever at man søker fred, og sørger derfor for at det er en viss mulighet for relativt fredelige samfunn.

## 2.5 Kortsiktighet og langsiktighet

When in the mind of man, appetites, and aversions, hopes, and fears, concerning one and the same thing, arise alternately; and divers good and evil consequences of the doing, or omitting the thing propounded, come successively into our thoughts; so that sometimes we have an appetite to it; sometimes an aversion from it; sometimes hope to be able to do it; sometimes despair, or fear to attempt it; the whole sum of desires, aversions, hopes and fears continued till the thing be either done, or thought impossible, is that we call DELIBERATION (Hobbes 1946:37).

Det å diskutere *hvorvidt* Hobbesianske individer er langsiktige virker muligens noe overflødig; de *er* langsiktige, og det uten noe snev av tvil. Allikevel er det interessant

å se litt nærmere på når, hvordan og når de eventuelt *ikke* er tilstrekkelig langsiktige. Taylor (1987) kritiserer Hobbes' manglende konkretisering av f.eks. hvordan mennesker diskonterer fremtidig nytte<sup>10</sup>. Tross mangelen på en diskusjon av diskontering i en form som enkelt låner seg til spillteoretisk behandling, mener jeg Hobbes er klar på at mennesker er langsiktige, og at de også til en viss grad diskonterer fremtidig nytte.

### 2.5.1 Kortsiktighet

Før jeg diskuterer langsiktighet i Hobbes' tekster, vil jeg se litt på hva Hobbes sier om det å *ikke* være langsiktig, rettere sagt *korttenkt/kortsiktig*. At det er langt vanskeligere å finne omtale av kortsiktighet enn langsiktighet hos Hobbes, forteller mye om hvor selvsagt han anser menneskers langsiktighet å være. Dersom et individ *overveier* noe, foretar han avveininger av mulige konsekvenser, og må derfor sies å være langsiktig (Hobbes 1946:37). Videre definerer Hobbes det å *ville* som den siste appetitt eller aversjon som velges i en vurdering, og siden en frivillig handling er en handling som springer fra *viljen*, må enhver frivillig handling utelukkes fra å kunne være kortsiktig (Hobbes 1946:37,38).

Vi sitter da igjen med *uoverveide* handlinger, og dersom disse har tilfredstillelse av et umiddelbart gode til hensikt, uten vekt på fremtidige konsekvenser, må de kunne kalles kortsiktige.

[P]assions unguided, are for the most part mere madness (Hobbes 1946:47).

As for the passions, of hate, lust, ambition, and covetousness, what crimes they are apt to produce, is so obvious to every man's experience and understanding, as there needeth nothing to be said of them, saving that they are infirmities, so annexed to the nature, both of man, and all other living creatures, as that their effects cannot be hindered, but by extraordinary use of reason, or a constant severity in punishing them (Hobbes 1946:194).

Våre *drifter* kan altså i visse tilfeller få utløp i handlinger, som ifølge Hobbes karakteriseres som "ren galskap" der de ikke kontrolleres, men som her betegnes som *kortsiktige* eller *uoverveide* handlinger. Noen av våre drifter (som hat, lyst, ambisjon og havepsyke) betegnes altså som svakheter hos mennesket (Hobbes 1946:194). Med *fornuft* er det

---

<sup>10</sup>Diskontering av fremtidig nytte innebærer at man vurderer midler i fremtiden som mindre verdt enn tilsvarende midler i nåtid. En mer teknisk anvendelse av begrepet følger i kapittel 4.

imidlertid mulig å få bukt med disse svakhetene, og fornuft er ifølge Hobbes' definisjon "beregning, det vil si addisjon eller subtraksjon, av konsekvensene som følger av generelle begreper det er enighet om betydningen av, for *merking* eller *signalisering* av våre tanker" (Hobbes 1946:25). Denne beregningen gjør altså enhver handling som følger, langsiktig, siden alle overskuelige konsekvenser er vurdert. Det å overveie konsekvenser og se fremover i tid er for Hobbes naturlig nok noe positivt:

Which kind of thoughts, is called *foresight*, and *prudence*, or *providence*; and sometimes *wisdom*; though such conjecture, through the difficulty of observing all circumstances, be very fallacious (Hobbes 1946:15).

Et viktig spørsmål når det gjelder langsiktighet, er hvorvidt det er *kompetanse* eller *innsats* som kvalifiserer til det. Er en person med mangelfull erfaring og lite vett langsiktig dersom han forsøker så godt han kan å overveie konsekvensene av en situasjon, eller kreves det en viss kompetanse? Ovenfor henviser Hobbes til fremsynthet som *klokskap* og *visdom*, men han kaller allikevel slike forsøk på å avdekke fremtidige konsekvenser som "meget feilbarlige"; det later altså til å være *innsatsen*, eller *prosessen* som kvalifiserer til langsiktighet (Hobbes 1946:15). Dette støttes videre når det gjelder rasjonell tenkning generelt, idet "utrente menn må feile i sine slutninger, og også professorer feiler ofte. Dette er også tilfellet i enhver annen form for argumentasjon; de dyktigste, mest oppmerksomme og best trente forleder seg selv, og utleder feilaktige konklusjoner" (Hobbes 1946:26). Uansett hvor inkompetent man er, vil man altså regnes som langsiktig dersom man foretar vurderinger og handler ut fra *vilje*. Men, de som har lite erfaring vil oftere enn andre forfølge handlinger rettet mot et tilsynelatende gode som kanskje viser seg å være noe helt annet enn forventet (Hobbes 1946:39).

### 2.5.2 Langsiktighet - grunnleggende i Hobbes' filosofi

The final cause, end, or design of men, who naturally love liberty, and dominion over others, in the introduction of that restraint upon themselves, in which we see them live in commowealths, is the foresight of their own preservation, and of a more contented life thereby (Hobbes 1946:109).

Hobbes illustrerer både styrken i individers langsiktighet og frykten for døden i sitatet ovenfor. Mennesker elsker frihet og herredømme over andre, men er allikevel villige til å gi avkall på begge deler for å få inngå i en stat. At dette er en langsiktig vurdering

synes opplagt, og store deler av resonnementet om den autoritære staten i *Leviathan* er basert på denne type langsiktighet.

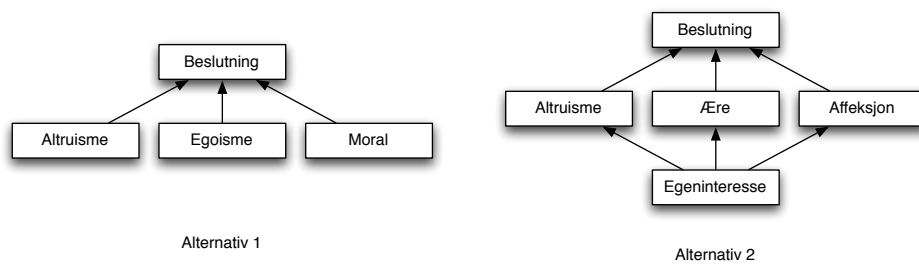
*Fornuft* er i seg selv langsiktig, ved at det dreier seg om avveininger (Hobbes 1946:25). Hobbes viser dog at vi også har drifter som til en viss grad underbygger vår langsiktighet. Ambisjon er nødvendigvis fremtidsrettet – en drift som får oss til å søke mer i fremtiden – og frykt er ifølge Hobbes det *eneste* som hindrer oss i å bryte loven for profitt eller nytelse, av respekt for fremtidige onder (Hobbes 1946:195). Hobbes legger også vekt på *pulchrum*, forventningene om noe godt, som et reelt (om enn muligens usikkert) gode (Hobbes 1946:33). Det må også påpekes at handlinger som intuitivt forstås som langsiktige, ikke nødvendigvis er dette for Hobbes:

Desire of fame after death does the same [som ønsket om å bli hyllet, det disponerer mennesker til prisverdige handlinger]. And though after death, there be no sense of the praise given us on earth, as being joys, that are either swallowed up in the unspeakable joys of Heaven, or extinguished in the extreme torments of hell: yet is not such fame vain; because men have a present delight therein, from the foresight of it, and of the benefit that may redound thereby to their posterity; which though they now see not, yet they imagine; and any thing that is pleasure to the sense, the same also is pleasure in the imagination (Hobbes 1946:64).

Dersom jeg foretar handlinger som kun vil ha innvirkning på mitt ettermæle, og dermed indirekte gir mine etterkommere nytte (dersom handlingene mine er positive), kan dette opplagt synes som et offer av umiddelbar nytte til fordel for fremtidig nytte. Hobbes anser imidlertid også dette som en balansert transaksjon, der gleden ved å forestille seg hva man har oppnådd (om enn i fremtiden), er det avgjørende for aktøren. Det er dog mindre sikkert hvorvidt slike betraktninger eksplisitt gjør seg gjeldende i våre aktørers vurderinger; det kan tenkes at de selv, i sine vurderinger, anser effektene etter deres egen død som viktige, men *grunnen* til at de velger dette er altså ingen uselviskhet, men et (muligens ubevisst) ønske om å tilfredsstille sin *appetitt* for å oppnå et godt ettermæle og sikre sine etterkommere. Denne forståelsen av motivasjon er sentral i min forståelse av Hobbes. “Altruisme”, eller uselviskhet, er ingen selvstendig kraft i menneskers motivasjon, men man kan ha preferanser som gjør at det å følge slike handlingsmønstre gir en selv økt nytte. Se figur 2.1<sup>11</sup>.

---

<sup>11</sup>Det er med dette tydelig at *egeninteresse* slik jeg behandler det her ikke er ment å være et empirisk relevant begrep som skilles fra f.eks. *altruisme*, da det omfatter enhver handling. Mitt formål i denne



Figur 2.1: Altruisme på primær- vs. sekundær-nivå

Langsiktighet kan også medføre uheldige konsekvenser, og det kan dermed synes som om det finnes et “korrekt” nivå; langsiktighet er ikke utelukkende positivt i seg selv. Hobbes advarer for eksempel mot å se altfor langt fremover i tid, da man risikerer at “ens hjerte dagen lang blir plaget av frykt for døden, fattigdom eller annen katastrofe, uten muligheter for fred eller pause i fortvilelsen, annet enn i søvn” (Hobbes 1946:70). En tendens til å se for langt fremover, eller legge for stor vekt på fremtiden, kan muligens også være en forklaring på at en del mennesker er det Hobbes kaller *ubesluttssomme*: de som “søker etter ting som bidrar lite til våre mål, og frykter ting som kun er små hindringer [...]” (Hobbes 1946:34). Slutninger trukket på bakgrunn av beregning av fremtidige konsekvenser er som Hobbes nevner iboende utsatt for feil. Det er da naturlig å anta at jo lenger frem man ønsker å beregne – jo flere ledd man må undersøke – desto større sjanse er det for å begå feil, og desto større konsekvenser vil feil i utgangspunktet for beregningen og opprinnelige antagelser få.

### 2.5.3 Generelt *krav* om langsiktig adferd for overlevelse/suksess

And from this diffidence of one another, there is no way for any man to secure himself, so reasonable, as anticipation; that is, by force, or wiles, to master the persons of all men he can, so long, till he see no other power great enough to endanger him: and this is no more than his own conservation requireth, and is generally allowed (Hobbes 1946:81).

Det er heller ikke tilstrekkelig å anse langsiktighet som noe positivt, det er ifølge Hobbes *essensielt* for personers overlevelse (Hobbes 1946:81). At det ikke finnes noen

---

oppgaven er ikke å fastsette en slik distinksjon mellom adferd motivert av forskjellige formål; jeg ønsker å påpeke hvordan en plausibel tolkning av Hobbes kan fremstå, og det faktum at man ikke behøver å forstå Hobbes som forkjemper for *egoisme* selv om han klart mener at *egeninteresse* er sentralt.

like fornuftig måte å sikre seg mot andre på som å komme andre i forkjøpet er et klart tegn på dette. Videre, som såvidt nevnt i starten av kapitlet, det å være moderat og medgjørlig – her forstått som ikke aktivt å søke makt for å sikre seg for fremtiden – er å gjøre seg selv til et offer. Individene som handlet slik, “ville frembragt sin egen undergang, i direkte motsetning til grunnlaget for naturlovene, som sikrer naturens bevarelse” (Hobbes 1946:103). Videre er det ikke engang nok med *noe* sikring av fremtiden, for det vil alltid finnes noen som søker enda mer makt enn hva de strengt tatt behøver, noe som vil medføre moderate individers undergang; utvidelse av eget herredømme er *nødvendig* for menneskers overlevelse (Hobbes 1946:81). Det er altså et reelt krav om en viss type oppførsel, som utelukkende er begrunnet i en uunngåelig fremtidig utvikling på grunnlag av menneskets natur og situasjon. Det må imidlertid bemerkes at denne adferden ikke utelukkende dreier seg om å gå i krigen for å underlegge seg eller drepe alle andre. Det må også bemerkes at dette er den *mest* fornuftige, men ikke *eneste* fornuftige, strategien. I kapittel 4 og 5 vil jeg se videre på hvorfor det er gode grunner til å hevde at moderasjon som strategi vil være det sikreste for mange.

#### 2.5.4 Risikoaversjon

I forbindelse med langsiktighet er det også relevant å diskutere risikovillighet. Umiddelbart vil det være naturlig å se på forbindelsen mellom risikoaversjon og høy diskontering av fremtids (nødvendigvis usikre) goder. Diskonteringsfaktoren er imidlertid *ikke* ensbetydende med risikoaversjon. Dette temaet behandles inngående under premiss 5 i 4.2.2 og i 5.3.3.

#### 2.5.5 Oppsummering

Hobbes' individer er langsiktige, eller i alle fall i stand til å være det, noe en gjennomgang av hans tekster viser uten særlig grad av tvil. Når det gjelder diskontering av fremtidig nytte er imidlertid Taylors (1987) innvending gyldig – Hobbes er lite spesifikk på dette punktet, men dette ser ikke ut til å skade hans teori nevneverdig; Hobbes er bl.a. klar på at mennesker har en egen evne til å oppfatte små og nødvendige betalinger som fryktelige, da de ikke er i stand til å se de overhengende farene slike betalinger f.eks. er ment å forhindre (Hobbes 1946:120). Det kan tenkes at Hobbes her ikke ser direkte på diskontering av farer i fremtiden, men heller en manglende evne til å oppfat-

te korrekt det som er fjernt fra en selv, men effekten blir i mange tilfeller den samme. Fraværet av en konkret diskusjon av diskontering hos Hobbes kan åpne for å anvende nyere kunnskap om menneskers faktiske diskontering av fremtidige goder. Hobbes ønsker å beskrive mennesker som de faktisk er, og dersom vi i dag vet ting om mennesket han ikke har beskrevet, og som ikke bryter med hva han *har* sagt, bør det til en viss grad være rom for å kombinere Hobbes' og vår kunnskap, for å få en mer komplett modell av mennesket.

Selve modelleringen av langsiktighet i spillteorien foregår ofte i det vi kaller spill med *uendelig tidshorisont*. Dersom et spill har *endelig tidshorisont*, er f.eks. baklengs induksjon en mulighet for å finne "rett løsning". Der spillet ikke har et bestemt antall runder – eller mangler en vel-definert stopp-regel, slik von Neumann og Morgenstern (2004) diskuterer – vet man ikke når spillet ender, og den definitive løsningen blir derfor langt vanskeligere å avdekke. Spillernes diskontering av fremtidig nytte (inkludert deres forventninger omkring hvorvidt spillet fortsettes nok en runde) har avgjørende betydning for hvordan de velger å spille. Denne typen formalisering av langsiktighet blir gjennomgått kort i forbindelse med premiss 1 i 4.2.2, og mer inngående i forbindelse med Taylor (1987) i 5.2.1.

## 2.6 Konklusjon

Jeg har i dette kapittelet foretatt en tekstuell analyse av Hobbes' tekster, med fokus på å avdekke Hobbes' syn på individet og menneskets natur. Gjennomgangen ble i stor grad foretatt med bakgrunn i spillteoretiske begreper, da jeg har hatt et ønske om å forberede marken for senere spillteoretiske betraktninger om Hobbes' filosofi.

Jeg har vist at Hobbes' individer ønsker å oppnå maksimal nytte i form av oppnåelse av goder nå eller i fremtiden. Til dette trenger de makt, men jeg hevder at makt ikke er et mål i seg selv; det er kun et middel for å oppnå goder. Jeg har sett på menneskets grunnleggende motivasjon, og mener bestemt at Hobbes' individer er *egeninteresserte*; formålet med deres handlinger er å oppnå egen fordel. Jeg har dog vist at dette ikke dreier seg om en kortsiktig egoisme. Mens egeninteresse er alene på det grunnleggende motivasjonsnivået, finnes det avgjørende nivået for beslutningstaking over dette, der f.eks. et ønske om godt for andre eller et ønske om å bedre andres situasjon for å styrke sine venner, kan føre til at individer utfører "gode" handlinger. Jeg har også



vist at langsiktighet er sentralt for Hobbes, ved at enhver frivillig handling blir tatt på bakgrunn av overveininger der fremtidige konsekvenser forsøkes kartlagt. Det er også en grunnleggende langsiktighet som muliggjør den statsdannelsen *Leviathan* i bunn og grunn omhandler; individer gir fra seg sine rettigheter og sin makt i bytte mot fred og økte muligheter for et langt og tilfreds liv.

Jeg mener videre en nyttefunksjon som skal kunne hevdes å være i tråd med Hobbes' filosofi bør inkludere det Taylor (1987:118) kaller "søfistikert altruisme". For Taylor er en aktør "altruistisk" dersom hans nyttefunksjon inkluderer andres gevinst:  $u_i = u_i(x_i, x_j)$ . Aktør  $i$ 's nytte ( $u_i$ ) blir utledet fra  $i$ 's gevinst ( $x_i$ ) og  $j$ 's gevinst ( $x_j$ ). (Vektingen av elementene er her uviktig.) Dette er det samme som å ha *sosiale preferanser*, og denne begrepsbruken blir etablert i neste kapittel. Dersom en aktør er *søfistikert* altruistisk blir nyttefunksjonen mer komplisert, og kan f.eks. se ut som ligning (2.2), der aktør  $j$ 's subjektive nytte ( $u_j$ ) også vurderes som positiv eller negativ for aktør  $i$ .

$$u_i = u_i(x_i, x_j, u_j), \quad j \neq i \quad (2.2)$$

Som nevnt i 2.4.2 hevder jeg Hobbes' individer helt klart har sosiale preferanser. Utover det mener jeg også at de må ansees som søfistikerte altruister. Hobbes legger sterk vekt på affeksjon og følelser, og det synes derfor relevant å inkludere dette i en nyttefunksjon, dersom man søker realisme i forhold til Hobbes' filosofi. Et godt eksempel på dette fikk jeg av Knut Midgaard: jeg, som far, vurderer om jeg skal gi min sønn tran. Jeg anser først og fremst det at sønnen min får tran som noe positivt ( $x_j$  er positiv og gir meg nytte ved at jeg anser min sønn som en "alliert/venn"). Det er allikevel et poeng at min sønn *hater* tran, og opplever det som fryktelig å måtte innta dette ( $u_j$  – min sønns subjektive oppfatning av nytte ved å innta tran – er svært lav/negativ). Det er opplagt at min nytte ved å gi min sønn tran endres dersom en vektlegging av hans egne følelser om situasjonen inkluderes i min nyttefunksjon. Dette gjør i mine øyne at man bør se nøye på mulighetene for å anvende slike nyttefunksjoner, men jeg vil i oppgaven forholde meg til konvensjonen, og vil bevege meg lite utover sosiale preferanser. Dette i stor grad fordi temaene jeg diskuteres ikke er avhengig av en komplisert nyttefunksjon, men også fordi bruken av slike nyttefunksjoner er svært lite utbredt, så jeg anser det prosjektet å skulle skape en realistisk Hobbesiansk nyttefunksjon som bedre egnet for senere arbeid.

# Kapittel 3

## Hobbes' filosofi om individet vurdert i lys av nyere bidrag

### 3.1 Innledning

I dette kapittelet vil jeg se på hvordan vi i dag forholder oss til de temaene jeg har undersøkt i det foregående kapittelet. Jeg finner det nyttig å sammenligne Hobbes' filosofi med to forskjellige tilnærminger til mennesket, og vil i hovedsak basere meg på psykologi og økonomi. Først vil jeg dog kort vurdere hvorvidt min tolkning av Hobbes deles av andre som har skrevet om ham. Med en så bred og god sekundærlitteratur ville det vært bemerkelsesverdig dersom det jeg hevder i kapittel 2, er helt nytt. Det er imidlertid mange forskjellige tolkninger av Hobbes, og jeg ønsker å se kort på hvordan noen av disse arter seg.

Fra psykologien ønsker jeg å se på Maslows teori om menneskelig motivasjon. Fra økonomien finner jeg det spesielt interessant å se på eksperimentell økonomi og nevroøkonomi, da disse fagområdene fokuserer på reell menneskelig adferd. De søker å avdekke (i alle fall deler av) menneskets natur via vitenskapelige eksperimenter der adferd observeres, samt ved å observere fysiologiske/nevrologiske reaksjoner i mennesker.

## 3.2 Sekundærlitteratur om Hobbes

Det er ingen enkel oppgave å oppsummere sekundærlitteraturen om Hobbes, da den er svært omfattende og innholdsrik. Jeg har heller ikke en slik litteraturgjennomgang som hovedmål i denne oppgaven. Jeg ønsker dog å belyse raskt hvordan min tolkning av Hobbes fra forrige kapittel kan vurderes i lys av annen kommentarlitteratur.

Jeg omtaler her i hovedsak litteratur som berører Hobbes og menneskets natur, ikke litteratur om Hobbes generelt. Et naturlig utgangspunkt er den innflytelsesrike “The Ethical Doctrine of Hobbes” (gjengitt i *Modern Political Theory* (Lively og Rose 1989)), originalt fra 1938, der A. E. Taylor hevder at Hobbes’ etiske (og deontologiske) doktrine kan frikobles fra hans syn på menneskelig motivasjon (Taylor 1989). Denne frikoblingen gjør Taylor i stand til å forsvare sluttproduktet til Hobbes, og samtidig imøtekomme kritikerne av det mange oppfatter som “psykologisk egoisme” hos Hobbes.

Oakeshott – i sin berømte introduksjon til *Leviathan* (1946) – følger en lignende linje, der den *røde tråden*<sup>1</sup> i Hobbes’ filosofi fremheves som det viktige, og at hans teori dermed ikke kan sees som en arkitektonisk struktur som faller sammen dersom fundamentet (her Hobbes’ fremstilling av menneskets natur) skulle vise seg å være sårbart (Oakeshott 1946).

Stuart Brown Jr. kommer så i en artikkel fra 1959 med kraftig kritikk at Taylors og Oakeshotts fremgangsmåte, og hevder det er umulig å skille Hobbes’ teori om psykologi fra hans teori om moral, og betegner forsøk på dette som å være på linje med å “kastere” Hobbes. Hobbes’ pessimisme er altså noe man må leve med dersom resten av teorien skal stå sterkt, selv om forutsetningen om psykologisk egoisme er “ufattelig grov og åpenbart feil” (Brown 1959). At Hobbes’ moralfilosofi og politiske filosofi er avhengig av hans psykologi synes plausibelt, og dette støttes av bl.a. Gauthier:

That Hobbes’s psychology is essential to his moral and political theories is one of the main theses of our argument (Gauthier 1969:1).

Malnes (1993:37) velger imidlertid å utforme en “Hobbesiansk teori” om internasjonal konflikt *uten* Hobbes’ syn på menneskets natur:

For convenience, I shall refer to these as the assumption of egoism and

---

<sup>1</sup>“The coherence of his [Hobbes’] philosophy, the system of it, lies not in an architechtonic structure, but in a single ‘passionate thought’ that pervades its parts“ (Oakeshott 1946:xix)

combativeness. These assumptions are necessary premises of the thesis of inevitable escalation, but they are bound up with a questionable – indeed, a cynical and pessimistic – view of human nature. If the general thesis of war – the major conclusion of Hobbes’ argument – is anchored in such assumptions, it will easily be rejected as excessively pessimistic. No wonder such gloomy premises yield such a dark conclusion! (Malnes 1993:37)

Det er imidlertid forståelig at man søker å utelate psykologi når man diskuterer stater fremfor individer. Malnes går så videre ved å reformulere Hobbes’ resonnement via en “mindre problematisk rute” (Malnes 1993:37). Jeg er, som det går frem i det foregående kapittelet, uenig i at Hobbes’ teori er tvilsom, kynisk og pessimistisk, og vil i dette kapittelet forsøke å vurdere Hobbes’ syn på mennesket ved å sammenligne det med utvalgte alternative teorier.

Gert (1965; 1967) ser ut til å løse noe av kontroversen omkring Hobbes’ psykologi, ved å benekte påstanden om at Hobbes står for egoisme overhodet, og da spesielt psykologisk egoisme. Gleden er imidlertid kortvarig; McNeilly (1966) forkaster Gerts tolkning av Hobbes, og hevder i stedet at Hobbes var inkonsekvent, og i realiteten beveget seg fra en forholdsvis snever egoisme i *The Elements of Law*, til å forkaste denne oppfatningen av mennesket i sine senere verker, som for eksempel *Leviathan*. Gerts artikler om temaet er dog meget interessante i forhold til mitt syn på Hobbes. Han forkaster tesen om at Hobbes’ filosofi om mennesket er basert på “psykologisk egoisme”, som han definerer som den “teorien som fremsetter at det eneste motivet for menneskelig handling er egeninteresse” (Gert 1965:341). Nøkkelordet for Gert er *motiv*, og han mener at når Hobbes sier at mennesker i enhver handling søker egen fordel, “er ingenting sagt om våre motiver for å ønske tingene vi ønsker, kun at det er en kausal forbindelse mellom noe som er fordelaktig for vår vitalitet og det faktum at vi ønsker dette” (Gert 1965:346). Her fremholder Gert tydeligvis et annet syn på *motivasjon* enn f.eks. Maslow, som jeg kommer til i 3.3. Mens Gert her krever at man må være *bevisst* på årsakene til en handling for at den skal kunne kalles motivert av disse årsakene, mener Maslow at også handlinger drevet av ubevisste behov/impulser må anees som motivert av disse. Dersom vi så anvender Maslows syn på motivasjon som både bevisst og ubevisst, og hevder at mennesket utelukkende er motivert av egeninteresse, får man det Gert kaller *tautologisk egoisme*.

Kavka (1986) påpeker med rette at begrepet tautologisk egoisme kun er interessant

for å vise at psykologisk egoisme er feil; han foretrekker det han kaller *dominerende egoisme* (predominant egoism), som innebærer at egoistiske motiver i overveiende grad dominerer, men at også andre motiver kan gjøre seg gjeldende, f.eks. i situasjoner der det å ikke forfølge egoistiske motiver er ufarlig. Som det vil fremgå i løpet av dette kapittelet anser jeg Kavkas begrep som en god tilnærming dersom man søker en empirisk relevant definisjon på Hobbesianske individers motivasjon. Gauthier fremmer forøvrig en lignende tanke, om enn noe fjernere fra Hobbes, ved at overlevelse ikke nødvendigvis ansees som det grunnleggende menneskelige motiv:

We must say, not that man is a self-maintaining engine, but that he is an engine which tends to seek to maintain itself. In other words, we may say that self-preservation is a very important human motive, but not the necessary basic motive (Gauthier 1969:24).

Minogue (1973) viser i sin relativt korte innledning til *Leviathan* (Hobbes 1973) en tilnærming til Hobbes som jeg føler meg meget komfortabel med. Minogue (1973:ix) mener det inntil nylig har vært moteriktig å vise avstand til Hobbes' syn på mennesket, men at dette – med rette – har endret seg. Minogue siterer Gert og påpeker at “den vanligste måten å tillegge grovhet til Hobbes' psykologi er å slutte seg til den oppfattning at Hobbes mener at mennesker i bunn og grunn er egoistiske dyr – noe som er feil” (Minogue 1973:x). Hans tolkning av Hobbes' filosofi er i stor grad i tråd med min gjennomgang i kapittel 2, og han påpeker at “[Hobbes'] utlegning av menneskelig adferd er mer nyansert enn den synes ved første øyekast” (Minogue 1973:xi). I motsetning til Oakeshott hevder Minogue at Hobbes' filosofi *ikke* utelukkende må ansees som et resultat av hans tid og omstendigheter, men at den er generell (Minogue 1973:xxiii, Oakeshott 1946:xi). Jeg er langt på vei enig med Minogue i dette, men vil i gjennomgangen av Maslow i det følgende avsnittet, og i kapittel 5, hevde at *visse deler* av Hobbes' filosofi må revideres dersom den skal fremstå som universell

### 3.3 Psykologi – Humanisten og optimisten Maslow

Først vil jeg vurdere motivasjonsteorien til Abraham Maslow (1908-1970) opp mot Hobbes' syn på mennesket. Jeg ønsker å sammenligne disse teoriene av flere grunner; først og fremst vil jeg få et annet, og i dag mer konvensjonelt, perspektiv på mennesket enn hva jeg har utledet fra Hobbes. En annen grunn til at jeg ønsker denne sammen-

ligningen er at det vil kunne gi Hobbes' psykologi mer "tyngde", dersom det finnes støtte for hans synspunkter i moderne litteratur. Likeledes vil det hjelpe oss å stille spørsmålsteget ved andre deler av hans teori, dersom forskningen på feltet har vist seg å fremme teorier i strid med Hobbes' tenkning. Utover dette er det interessant å kunne omtale og vurdere Hobbes i en mer moderne terminologi, noe som muligens kan gjøre hans teori enklere å forstå. Jeg ser også en egenverdi i å gå gjennom en fullstendig motivasjonsteori à la Maslows, for å få en komplett "moderne" struktur som Hobbes kan vurderes etter. Den siste grunnen er at jeg i kapittel 5 ønsker å vurdere alternativer til Hobbes' syn på mennesket i forbindelse med hans *resonnement* om staten; jeg håper å kunne utlede et slikt alternativ gjennom vurderingen av Maslow og eksperimentell økonomi og nevroøkonomi.

Hvorfor har jeg så valgt akkurat Maslow? Først og fremst anser jeg det som ønskelig å vurdere én bidragsyter på feltet, for å få en helhetlig alternativ teori å vurdere Hobbes' teori mot. Jeg kan ikke hevde å være en ekspert på psykologi, men jeg har valgt en klassiker innen feltet jeg *har* kjennskap til, siden han ofte anvendes i økonomi- og ledelseslitteratur. Maslow er også anerkjent, og ble i *Esquires* jubileumsutgave i forbindelse med dets 50-års jubileum valgt ut som den mest innflytelsesrike psykologen på midten av det tyvende århundre, og en av de viktigste for vår moderne forståelse av menneskets natur (Frager 1987:34). Den viktigste grunnen er imidlertid at Maslow fremstår som en svært interessant motvekt til Hobbes:

He was one of the foremost spokespersons of the humanistic or "Third Force" psychologies [...] The book [*Motivation and Personality*] has had a tremendous influence in creating a positive and whole view of human nature (Frager 1987:33).

Jeg har allerede påpekt at jeg har motforestillinger mot en fremstilling av Hobbes som en negativ tenker som fremmer et "ondt" syn på mennesket, og vil da vurdere dette videre i gjennomgangen av en tenker med en positiv teori. Maslows teori er positiv i den forstand at han fokuserer svært mye på menneskets muligheter fremfor begrensninger. Teorien er også positiv i den betydning at den er *optimistisk* hva gjelder menneskets natur og situasjon.

Først vil jeg gjennomgå Maslows diskusjon av grunnlegende behov og det hierarkiet disse danner. Jeg vil fortløpende sammenligne dette med hva jeg har funnet hos Hobbes. De elementære delene i Maslows og Hobbes' teorier er forholdsvis sammenlignbare, da

begge diskuterer menneskers behov og natur. Etter denne sammenligningen vil jeg se mer spesifikt på de delene hos Maslow man *ikke* finner igjen hos Hobbes, og hvilke implikasjoner disse utelatelsene kan tenkes å ha for Hobbes' fullstendige teori.

### 3.3.1 Menneskelig motivasjon og grunnleggende behov

De fleste har hørt om Maslows behovshierarki – spesielt de med kjennskap til psykologi eller teorier om ledelse. Maslow er en teoretisk orientert psykolog, men baserer seg i stor grad på empiri i utformingen av sine hypoteser.

Maslow deler inn menneskers *grunnleggende behov* i et hierarki, og jeg vil her følge hans inndeling, der fysiologiske behov kommer først. Etter dette plasseres sikkerhetsbehov, behov for tilhørighet og kjærlighet, behov for anerkjennelse og respekt og til sist selvrealiseringsbehov (Maslow 1987:59).

#### Fysiologiske behov

Først av alt har man fysiologiske behov, som for Maslow er primære. Dersom en person mangler *alt*, er det naturlig å tenke seg at han vil søke å tilfredstille disse behovene først:

The urge to write poetry, the desire to acquire an automobile, the interest in American history, the desire for a new pair of shoes are, in the extreme case, forgotten or become of secondary importance. For the human who is extremely and dangerously hungry, no other interests exists but food (Maslow 1987:58).

En svært interessant teori hos Maslow er at han anser behov som i stand til å *dominere* mennesker<sup>2</sup>. Ifølge Maslow oppleves ikke alle typer behov like sterkt *samtidig*; nye behov overtar når eksisterende behov blir tilfredsstilt. Det behovet som er mest fremtredende i en person kan – dersom frustrert – på mange måter “overta” personen. “For vår kronisk og ekstremt sultne person, kan utopia helt enkelt defineres som et sted med overflod av mat” (Maslow 1987:59). Alle målsetninger og fremtidsutsikter dreier seg om mat, og Maslow mener en slik person forestiller seg evig lykke bare han får noe å spise. “Det å leve blir definert i relasjon til å spise. Alt annen blir definert som uviktig. Frihet,

---

<sup>2</sup>Dette gjelder behov generelt, ikke utelukkende fysiologiske behov.

kjærlighet, fellesskapsfølelse, respekt, filosofi, kan feies til siden som verdiløst juggel, siden de ikke er i stand til å mette magen” (Maslow 1987:59).

At sultende mennesker søker mat må regnes som selvsagt, og er en representasjon av den enkleste formen for motivasjon man kan fremstille: et darwiniansk ønske om, og instikt for, overlevelse<sup>3</sup>.

Allerede her ser man svært interessante betraktninger som er relevante for Hobbes’ teori. Hobbes er helt opplagt klar over at mat er viktig, og han omtaler det i flere sammenhenger. Først og fremst anser han fysiske behov og “appetitter” som naturlige for mennesket, som f.eks. det å spise og puste. Hobbes kobler også dette til somatiske baser, ved at det er appetitter, men mest aversjoner, “fra[/for] noe de føler i kroppen” (Hobbes 1946:31,32).

Sult er ifølge Hobbes et behov vi deler med alle levende vesener. Dyr har imidlertid *kun* slike drifter, som “sult, tørste, lyst og sinne” (Hobbes 1946:15). Det interessante er at Hobbes også kommer svært nære å hevde at behov kan *dominere* individer, som f.eks. “i andre *dyr*, hvis overvekt av appetitter for mat og andre sensuelle gleder fjerner behovet for å vite årsaker [...]” (Hobbes 1946:35). Dette ligner mye på Maslows teori, men blir ikke utlagt som en fullstendig hypotese hos Hobbes, slik det blir hos Maslow. Det kan også synes som om Hobbes anser appetittene som substansielt like, kun utkonkurrert av hverandres styrke/mengde, mens Maslow med sitt hierarki skiller mellom behovene, og anser dem som substansielt forskjellige<sup>4</sup>.

I sine krav til en god motivasjonsteori forkaster Maslow fokuset på *sult* som universell basis for motivasjon, samt dyreforskning. Maslow påpeker nemlig en viktig sammenheng mellom fokuset på sult og dyreforskning: mye av det fokuset som har vært på sult kommer nettopp fra forskning på f.eks. rotter, som han mener hovedsakelig motiveres av fysiske behov (Maslow 1987:59). I dagens samfunn opplever mennesker svært sjelden slike behov som presserende, noe som gjør en tilnærming til mennesket utelukkende basert på sult lite fruktbar (Maslow 1987:59).

Det finnes en sentral likhet mellom Hobbes’ syn på *lykke* og Maslows syn på *behovstilfredstillelse*:

---

<sup>3</sup>Darwin nevner også reproduksjon, som kan forstås som overlevelse i bred forstand – overlevelse for våre gener.

<sup>4</sup>Hobbes vektlegger også det faktum at tilfredstillelse av grunnleggende behov – som det å skaffe mat – kan være såpass tidkrevende at man ikke får *mulighet* til å tilfredsstille andre behov. I så tilfelle vil altså behovet fremdeles være tilstede, men andre behov får prioritet (Hobbes 1946:103).



But what happens to their desires when there *is* plenty of bread and when their bellies are chronically filled? [...] *at once other (and higher) needs emerge* and these, rather than physiological hungers, dominate the organism. And when these in turn are satisfied, again new (and still higher) needs emerge, and so on (Maslow 1987:59).

Felicity is a continual progress of the desire, from one object to another; the attaining of the former, being still but the way to the latter (Hobbes 1946:63).

En sentral *ulikhet* er at man, ifølge Maslow, stadig endrer hva man søker ved slik behovstilfredstillelse, mens det for Hobbes ikke ser ut til å være predikert noe mønster i hvilke behov som overtar for de som er tilfredstilt<sup>5</sup>. Dette gir et svært viktig poeng som må diskuteres noe mer inngående mot slutten av denne behandlingen, ved at *Maslows individer kan sies å endres fundamentalt med hensyn til ønsker og behov*.

Maslow har et bredere arsenal av tidligere innsikt enn Hobbes når han utformer sin teori, noe vi bl.a. ser ved at *homeostasis*<sup>6</sup> anvendes som forklaring på kroppens signalisering av hva den trenger. Appetitter forstås til en viss grad som en direkte manifestasjon av kroppens reelle fysiske behov (Maslow 1987:59, Myers 2003:457). Det er selvsagt ikke å forvente at Hobbes anvender denne typen moderne begreper, og i mine øyne er følgende sitat et uttrykk for samme idé:

This motion, which is called appetite, and for the appearance of it *delight*, and *pleasure*, seemeth to be a corroboration of vital motion, and a help thereunto; and therefore such things as caused delight, were not improperly called *jucunda*, à *juvando*, from helping or fortifying (Hobbes 1946:33).

## Sikkerhetsbehov

Idéen om et grunnleggende behov for sikkerhet er *meget* sentral hos Hobbes. Dette vil komme tydelig til syne i kapittel 4 og 5, og jeg vil her hovedsakelig se på hva Hobbes og Maslow sier om temaet, mens vurderingen av implikasjonene kommer i de neste kapitlene.

---

<sup>5</sup>Hobbes skiller dog mellom forskjellige typer behov, men det er ikke noe systematisert hierarki som hos Maslow. Se mer under "Behov for selvrealisering" nedenfor.

<sup>6</sup>Kroppen forsøker automatisk å opprettholde en konstant normal blodomstrømning, og trenger da visse mengder vann, salt, sukker, proteiner osv. for å oppnå dette.

Maslow karakteriserer sikkerhetsbehovene som bestående av følgende elementer:

[...] (security; stability; dependency; protection; freedom from fear, anxiety, and chaos; need for structure, order, law, and limits; strength in the protector; and so on) (Maslow 1987:60).

Dette er ting også Hobbes fokuserer sterkt på, og Maslow påpeker at også slike behov kan *dominere* mennesker:

Again, as in the hungry human, we find that the dominating goal is a strong determinant not only of their current world outlook and philosophy but also of their philosophy of the future and of values. Practically everything looks less important than safety and protection (even sometimes the physiological needs, which, being satisfied, are now underestimated). A person in this state, if it is extreme enough and chronic enough, may be characterized as living almost for safety alone (Maslow 1987:60).

I naturtilstanden er det ifølge Hobbes svært liten grad av sikkerhet, da livet er preget av “kontinuerlig frykt, og fare for en voldelig død” (Hobbes 1946:82). Hvorvidt det verste ondet kan konstateres å være døden – noe Hobbes mener det er – vil jeg la stå åpent, men det synes meget sannsynlig at dette i det minste er tilfellet for mennesker dominert av sikkerhetsbehov.

Videre forklarer Maslow hvordan denne usikkerheten manifesterer seg i menneskers ønsker og mål. Det er meget interessant at Maslow mener sikkerhetsbehov ofte frembringer et sterkt ønske om “en beskytter, eller sterkere person eller system, som de kan stole på”, samt “en lettere aksept for diktatur eller militærstyrke” (Maslow 1987:61). Hobbes’ idéer vil da sannsynligvis fremstå som meget attraktive for mennesker med slike behov. En følt sikkerhetstrussel kommer ikke nødvendigvis bare i full krig, men ved “trussel mot lov, orden og samfunnets autoritet” (Maslow 1987:61). Dette omfatter Hobbes’ naturtilstand, som ikke består av åpen krigføring, men en kjent tilbøyelighet til dette (Hobbes 1946:82).

Det er påfallende hvor stor overenstemmelse det her er mellom Maslow og Hobbes, og man behøver ikke strekke seg spesielt langt for å anta at de to ville vært enige om at disse behovene er sentrale.

Som med de fysiologiske behovene er det ifølge Maslow svært sjelden man ser disse behovene utilfredstilt i særlig stor grad i moderne og fredelige samfunn. Dette medfører at

mennesker nedvurderer og “glemmer” at sikkerhet er viktig, noe som gjør at den typen tankegang man finner hos sikkerhetsdominerte individer virker fremmed, og kanskje til og med forkastelig, for individer opptatt av høyere behov (Maslow 1987:75,82,179). For Hobbes er naturtilstanden som vi vet svært forskjellig fra et moderne og stabilt samfunn, så forskjellen i fokus i de to tilstandene er helt naturlig.

I have also become convinced that getting used to our blessings is one of the most important nonevil generators of human evil, tragedy, and suffering. What we take for granted we undervalue, and we are therefore apt to sell a valuable birthright for a mess of pottage, leaving behind regret, remorse, and a lowering of self-esteem (Maslow 1987:179).

### **Behov for tilhørighet og kjærlighet**

Når det gjelder behov for tilhørighet og kjærlighet, ser vi i følgende sitat fra Fromm at slike “høyere” behov ikke på noen måte bør ansees som bagatellmessige:

The physiologically conditioned needs are not the only imperative part of man's nature. There is another part just as compelling, one which is not rooted in bodily processes but in the very essence of the human mode and practice of life: the need to be related to the world outside oneself, the need to avoid aloneness. To feel completely alone and isolated leads to mental disintegration just as physical starvation leads to death (Fromm 1994:17).

Hobbes skriver lite om behovet for tilhørighet, spesielt dersom det må være et selvstendig behov og ønske om dette, slik det er hos Maslow. Maslow anser disse behovene som svært viktige, og mener vi altfor ofte undervurderer eller overser disse; “[v]i har i stor grad glemt våre dyptliggende tendenser til å gjete, danne flokk, å bli medlem og å tilhøre” (Maslow 1987:62). Det er verdt å merke seg at Hobbes diskuterer grupper og allianser, men da i all hovedsak i form av instrumentell makt og sikkerhet (Hobbes 1946:56).

I forbindelse med kjærlighet (som Maslow påpeker ikke er det samme som sex), er det imidlertid mer samsvar mellom Maslow og Hobbes; selv om Hobbes ikke diskuterer det i særlig stor grad, nevner han det som svært viktig for mennesker:

Of things held in propriety, those that are dearest to a man are his own life,

and limbs; and in the next degree, in most men, those that concern conjugal affection; and after them, riches and means of living (Hobbes 1946:223).

At ekteskapelig kjærlighet her nevnes før rikdom og midler for overlevelse, bør være en klar indikasjon på at behovet er anerkjent av Hobbes. Hvorfor dette så i stor grad er utelatt i *Leviathan* kan deretter diskuteres, men jeg anser det som en følge av at Hobbes i hovedsak var opptatt av å forklare mennesket i den grad det påvirket hans teori om staten. Disse behovene kan da ha fremstått som mindre sentrale for Hobbes.

### **Behov for anerkjennelse og respekt**

Behovene for anerkjennelse og respekt er sterkt representert i Hobbes' tenkning:

And from hence it is, that kings, whose power is greatest, turn their endeavours to the assuring it at home by laws, or abroad by wars: and when that is done, there succeedeth a new desire; in some, of fame from new conquest; in others, of ease and sensual pleasure; in others, of admiration, or being flattered for excellence in some art, or other ability in their mind (Hobbes 1946:64).

I sitatet ovenfor ser man igjen at Hobbes foregriper den dynamikken Maslow systematiserer i form av ønsker som overtar når andre er tilfredsstilt. Videre er det også en åpenbar likhet i hvordan de ordlegger seg i forbindelse med slike behov:

All people in our society (with a few pathological exceptions) have a need or desire for a stable, firmly based, usually high evaluation of themselves, for self-respect or self-esteem, and for the esteem of others (Maslow 1987:63)

Lastly, considering what value men are naturally apt to set upon themselves; what respect they look for from others; and how little they value other men; from whence continually arise amongst them, emulation, quarrels, factions, and at last war, to the destroying of one another, and diminution of their strength against a common enemy [...] (Hobbes 1946:118)

Maslow peker her spesielt på "rykte eller prestisje (definert som respekt eller aktelse fra andre), status, berømmelse, ære, dominans, anerkjennelse, oppmerksomhet, viktighet, verdighet og verdsettelse" (Maslow 1987:63). Disse behovene har ifølge Maslow blitt undervurdert tidligere, f.eks. av Freud. Andre, som Alfred Adler, vektla dette sterkt,

og Maslow ser en sterk tendens til at dette nå anerkjennes som svært sentralt for mennesker (Maslow 1987:63).

Hobbes anser behovet for anerkjennelse og respekt (*ære*) som så viktig at det nevnes som en av de tre primære grunnene til konflikt (Hobbes 1946:81). I 2.4.6 har jeg behandlet dette mer inngående, og forklarer hvordan et behov for ære faktisk lett kan føre til vold ved mangel på anerkjennelse og respekt. Maslow går dog ikke spesielt inn på dette i forbindelse med konflikt og vold, men generelt sett kan man hevde at frustrasjon av behov kan tenkes å ha slike konsekvenser.

Siden disse behovene er såpass viktige, pålegger Hobbes suverenen å regulere ære og anerkjennelse (Hobbes 1946:118). Naturlov nr. 4 påpeker også at det er nødvendig at man viser hverandre *takknemlighet*, som også kan forståes som å anerkjenne andres godhet, noe som er sentralt for at de skal oppleve nytten de søker ved å utføre gode handlinger (Hobbes 1946:99).

## Selvrealisering

We should never have the desire to compose music or create mathematical systems, or to adorn our homes, or to be well dressed if our stomachs were empty most of the time, or if we were continually dying of thirst, or if we were continually theathened by an always-impending catastrophe, or if everyone hated us (Maslow 1987:49).

Selvrealisering kan nesten sies å være definisjonen av Maslows psykologi. Han anser psykisk “friske” og selvrealiserende individer som idealet, og diskuterer i store deler av sine verker hvordan man kan oppnå dette.

Maslows selvrealiserende individer er ikke genier à la Mozart, men mennesker som utvikler og beskjeftiger seg med det de er bra på (Maslow 1987:200). Selvrealiserende individer er kreative, åpne og “sunne” mennesker – “bra dyr” som han beskriver deres naturlighet – mens mange genier, “Wagner, f.eks., eller Van Gogh, Degas eller Byron”, opplagt ikke var psykisk friske (Maslow 1987:200).

Fokuset på selvrealisering gjenspeiles absolutt ikke av Hobbes, *men*, en del av Hobbes’ anerkjennelse av f.eks. kunst og vitenskap som grunner til å entre staten kan til en viss grad sammenlignes:

Desire of knowledge, and arts of peace, inclineth men to obey a common power: for such desire, containeth a desire for leisure; and consequently protection from some other power than their own (Hobbes 1946:64).

Hobbes ser altså ut til å forstå at mennesker har et behov for å utfolde seg utover det å sikre sitt daglige brød. Årsakssammenhengene er imidlertid noe forskjellige hos Hobbes og Maslow; mens Maslow mener behovet for selvrealisering kommer først når andre, mer grunnleggende, behov er tilfredsstilt, fremsetter Hobbes at dette behovet alltid finnes, ved at det er en grunn til å danne staten. Han ser imidlertid sammenhengen mellom ro og orden og slike “høyere” gjøremål, og benekter at slikt vil forekomme i naturtilstanden.

The merely surviving person will not worry much over the higher things of life, the study of geometry, the right to vote, civic pride, respect; he or she is primarily concerned with more basic goods (Maslow 1987:83).

*Desire*, to know why, and how, CURIOSITY; such as is in no living creature but *man*: so that man is distinguished, not only by his reason, but also by this singular passion from other *animals*; in whom the appetite of food, and other pleasures of sense, by predominance, take away the care of knowing causes [...] (Hobbes 1946:35)

Synet på krig og borgerkrig som ytterst uheldig kan ifølge Maslow med dette begrunnes i mer enn muligheten for døden, noe som også fremstår som en naturlig konklusjon for Hobbes:

No longer need we rest on sheer viability and survival as our only ultimate proof that poverty or war or domination or cruelty are bad, rather than good. We can consider them bad because they also degrade the quality of life, of personality, of consciousness, of wisdom (Maslow 1987:103)

### 3.3.2 Innstilling til menneskets natur

Hvordan arter så synet på menneskets natur seg hos en positiv tenker som Maslow? Først av alt forkaster han “egoisme-altruisme”-dikotomien, og behandler dermed et spørsmål jeg diskuterte i forrige kapittel. Jeg hevdet der at mennesker ifølge Hobbes er utelukkende egeninteresserte, men at dette ikke utelukker altruisme;

slik sett går Maslow svært godt overens med hvordan jeg forstår Hobbes (Maslow 1987:81,101,191).

If our instinctlike impulses, for instance, to love, arrange it so that we get more personal “selfish” pleasure from watching our children eat a special treat than from eating it ourself, then how shall we define “selfish” and how differentiate it from “unselfish”? (Maslow 1987:101)

Fragger (1987) siterer Maslow på at menneskets historie er en lang fortelling der menneskets natur har blitt undervurdert; “historien har praktisk talt alltid undervurdert menneskets natur” (Maslow 1987:158). Når det gjelder menneskets *positive potensiale* mener Maslow dette nær sagt alltid har blitt oversett, og at det er viktig å korrigere denne “forglemmelsen” (Fragger 1987:7).

For inscrutable reasons that only the intellectual historian may be able to unravel, Western civilization has generally believed that the animal in us is a bad animal, and that our most primitive impulses are evil, greedy, selfish, and hostile (Maslow 1987:92)

Maslow er sterkt uenig i en slik forståelse av “dyret” mennesket. “Arvesynden” er ikke et begrep Maslow slutter seg til, og han anser f.eks. Darwin som altfor opptatt av “dyriske” aspekter ved menneskets natur, slik at han utelukkende ser konkurranse, og ikke det samarbeidet som også finnes i de fleste menneskelige interaksjoner (Maslow 1987:92). Dette stammer ifølge Maslow fra vår tilbøyelighet til å forske på dyr, for deretter å slutte til mennesker. Ved at vi forsker på f.eks. ulver og rotter er det naturlig at vi anser oss selv som mer brutale og “onde” enn vi kanskje egentlig er, dersom vi antar at vår natur er lik.

Det er også interessant at Hobbes hevder mennesker ikke får noen glede av f.eks. å drepe andre. Dette gjelder *som regel*, men unntak finnes, og det er naturlig å forestille seg at Hobbes til en viss grad omtaler psykopater i sine mest negative vendinger om mennesket. Som vi ser nedenfor sier Hobbes “noen mennesker” – det er ingenting som tyder på at Hobbes antar at *alle* mennesker er psykopater, noe man kan få inntrykk av i noe Hobbes-litteratur.

It is helpful in understanding psychopaths to assume that they have no love identifications with other human beings and can therefore hurt them or even kill them casually, without hate, and without pleasure, precisely as

they kill animals who have come to be pests (Maslow 1987:129)

Thomas Hobbs said that if it were not for the gallows, some men are of so cruell a nature as to take a delight in killing men more than I should to kill a bird (Aubrey 1999:157).

Det er også viktig å skille mellom det å kalle menneskets natur ond, og det å erkjenne at menneskets natur kan få onde konsekvenser. Hobbes viser i sitt eksempel om mannen som reiser bevæpnet, låser døren sin og selv i sitt eget hus låser sine kister, at vi alle, akkurat som Hobbes selv, implisitt sier at andre ikke er til å stole på; men han er like klar på at “ingen av oss bebreider menneskets natur ved dette. Menneskers ønsker og andre drifter er ikke synder” (Hobbes 1946:83). Gauthier (1969:17) understeker det samme, når han i gjennomgangen av “alles krig mot alle” viser at denne krigen er et resultat av menneskers usikkerhet i naturtilstanden, og deres naturlige ønske om å sikre sin overlevelse; “Hobbes’ metafor misforståes grovt dersom man innbillen seg at den viser menneskers naturlige ondskap”.

It can certainly be granted by now that our knowledge is sufficient to reject any claim that human nature is, in its essence, *primarily, biologically, fundamentally* evil, sinful, malicious, ferocious, cruel, or murderous. But we do not dare to say there are *no* instinctlike tendencies at all to bad behaviour (Maslow 1987:124).

Nor dare we forget that human nature itself generates much evil, if not from inborn malice, then from ignorance, stupidity, fear, miscommunications, clumsiness, and so on (Maslow 1987:148).

Som avslutning på diskusjonen om menneskets natur vil jeg sitere Maslows forklaring på at et visst syn på livet kan føre til vold og en konfliktfylt tilstand. Dette er som oftest oversett, hevder Maslow, men det er i høyeste grad inkorporert hos Hobbes:

One reason for aggressive behaviour that is customarily overlooked, or if not then is phrased inaccurately, is the authoritarian view of life. If people were *actually* to live in a jungle in which all other animals were divided into two classes, those who could eat them and those they could eat, aggression would become a sensible and logical thing. People described as authoritarian must frequently tend unconsciously to envisage the world as just such a jungle. On the principle that the best defense is a good attack, these people are



apt to lash out, to strike, to destroy with no apparent reason whatsoever, and the whole reaction remains meaningless until it is realized that this was simply in anticipation of an attack by the other person. There are also many other well-known forms of defensive hostility (Maslow 1987:132)

### 3.3.3 Instinkter, rasjonalitet og det gode

Instinct, which we can define rigidly as a motivational unit in which the drive, motivated behaviour, and the goal object or the goal effect are all appreciably determined by heredity (Maslow 1987:51).

Instinkter er for Maslow en sentral del av mennesket, og siden instinkter må knyttes til menneskets natur, er det naturlig at Maslow heller ikke anser disse som noe negativt. Vi har sett på kroppens instinktive tendens til å sikre egen helse under bl.a. “Fysiologiske behov” ovenfor, og jeg sluttet at både Hobbes og Maslow ser slike instinkter som noe naturlig, og positivt, for organismen. Maslow påpeker også at nær sagt samtlige teoretiske retninger innen “psykiatri, psykoanalyse, sosialarbeid eller barneterapi har vært *nødt* til å fremsette en doktrine om instinktaktige behov, selv om de har vært uenige om så og si alt annet” (Maslow 1987:89).

As Eric Fromm has said, “reason by becoming a guard set to watch its prisoner, human nature, has become a prisoner itself and thus both sides of human nature, reason and emotion, were crippled” (Fromm 1947). We must all agree with Fromm that the realization of the self occurs not only by acts of thinking but rather by the realization of the total human personality, which includes the active expression not only of the intellectual but also the emotional and instinctlike capacities (Maslow 1987:158)

*Hele mennesket* er i fokus hos Maslow, og siden instinkter er en naturlig del av mennesket, tillegges det samme verdi som f.eks. fornuft. Her skiller Maslow seg sterkt fra Aristoteles, som ordnet menneskets evner i et hierarki med fornuften øverst (Maslow 1987:158). Med bakgrunn i dette dannes det en naturlig dikotomi, der *fornuften* og *følelser/instinkter* fremstilles som motstridende. Dette er ifølge Maslow helt feil, da disse aspektene ved mennesket i studier viser seg å være mer synergiske og forenlige enn motstridende (Maslow 1987:158).

Healthy reason, as definable today, and healthy instinctlike impulses point

in the same direction and are *not* in opposition to each other in the healthy person (although they *may* be antagonistic in the unhealthy). As a single example, all the scientific data now available indicate that it is psychiatrically desirable for children to be protected, accepted, loved, and respected. But this is precisely what children (instinctively) desire (Maslow 1987:93).

At Maslow her sier seg uenig med Aristoteles er noe man lett kan tenke seg at Hobbes ville frydet seg over, siden han hadde sterke motforestillinger mot “grekernes, og spesielt Aristoteles’, fåfengte og feilaktige filosofi” (Hobbes 1946:398):

And I believe that scarce anything can be more absurdly said in natural philosophy, than that which now is called *Aristotle’s Metaphysics*; nor more repugnant to government, than much of that he hath said in his *Politics*; nor more ignorantly, than a great part of his *Ethics* (Hobbes 1946:439).

Det er imidlertid hovedsakelig Aristoteles’ politiske doktrine som vies plass i *Leviathan*, ikke synet på instinkt og fornuft. Hobbes benytter ikke ordet “instinkt” i *Leviathan*, men snakker allikevel om *menneskets naturlige tilbøyeligheter* – noe jeg velger å tolke som et lignende begrep (Hobbes 1946:465). I *De Cive* forklarer han også hvordan mennesker ved en “naturlig impuls” søker seg vekk fra døden:

For every man is desirous of what is good for him, and shuns what is evil, but chiefly the chiefest of natural evils, which is death; and this he doth by a certain impulsion of nature, no less than that whereby a stone moves downward. It is therefore neither absurd nor reprehensible, neither against the dictates of true reason, for a man to use all his endeavours to preserve and defend his body and the members thereof from death and sorrows (Hobbes 2005a:8).

Mennesket søker naturlig å overleve, og naturlovene – som er *fornuftslover* – er veien for å oppnå dette der mennesker lever sammen; det er ingen grunnleggende strid mellom instinkt og fornuft hos Hobbes<sup>7</sup>.

At dyr har sterke instinkter er det liten tvil om, og dette er en av grunnene til at man anser instinkt og fornuft som motstridende, mener Maslow. Ved å forske på dyr, og anta at de samme instinktene gjelder hos mennesker, begår vi en stor feil, spesielt siden vi ofte identifiserer oss med dyr som “ulver, tigere, gribber og slanger, fremfor

---

<sup>7</sup>Det er heller ikke det *samme*, men som hos Maslow kan det være en viss synergi mellom disse.

bedre, eller i det minste mildere, dyr, som rådyr, elefanter eller sjimpanser” (Maslow 1987:92). Våre impulser er også langt svakere enn dyrs, hevder Maslow; mennesker styres ikke utelukkende av impulser, slik det kan hevdes at visse dyr gjør (Maslow 1987:91). Videre er menneskets instinkter forskjellige fra instinktene til en gribb, og Maslow mener de er langt mer positive.

Når det gjelder “det gode”, finner jeg mange likheter mellom Maslow og Hobbes, og da spesielt det faktum at organismens egne ønsker godtas som det gode (Maslow 1987:102)<sup>8</sup>. Dette er for Hobbes selve definisjonen på det gode der individer lever i naturtilstanden, og Maslow er langt på vei enig. At individers instinkter forstås som noe positivt, og noe som fremmer det gode, gjør at en stat som legger til rette for at instinkter får best mulig utløp fremstår som attraktiv. “Dersom våre iboende impulser forstås som noe beundringsverdig fremfor avskyelig, må vi søke å frigjøre disse slik at de får utfolde seg fritt fremfor å fange dem i tvangstrøyer” (Maslow 1987:103).

Hobbes og Aristoteles er faktisk enige om at menneskers egne vurderinger er det avgjørende for hva som er det gode i naturtilstanden. I en stat er det derimot loven som avgjør hva som er godt, ifølge Hobbes. Dette fordi loven eksisterer for å sikre statens mål, som igjen er menneskenes mål (Hobbes 1946:446). Mer om dette følger i kapittel 4.

### 3.3.4 På hvilke måter skiller Hobbes seg fra Maslow?

En av de største forskjellene mellom Maslow og Hobbes kan finnes i Maslows omtale av *frihet*. Maslow hevder det eksisterer visse *forutsetninger* for de grunnleggende behovene:

There are certain conditions that are immediate prerequisites for the basic need satisfactions. Such conditions as freedom to speak, freedom to do what one wishes so long as no harm is done to others, freedom to express oneself, freedom to investigate and seek information, freedom to defend oneself,

---

<sup>8</sup>“It may be objected by the technical philosopher “How can you prove that it is better to be happy than unhappy?” Even this question can be answered empirically, for if we observe human beings under sufficiently wide conditions, we discover that they, *they* themselves, *not* the observer, choose spontaneously to be happy rather than unhappy, comfortable rather than pained, serene rather than anxious. In a word, human beings choose health rather than illness, all other things being equal” (Maslow 1987:158).

justice, fairness, honesty, and orderliness in the group are examples of such preconditions for basic need satisfaction (Maslow 1987:64)

Videre hevder han at “hemmelighold, sensur, uærlighet og blokkering av kommunikasjon truer *alle* de grunnleggende behovene” (Maslow 1987:65). For Hobbes er det dog enkelt: mennesker velger overlevelse fremfor frihet dersom de må velge. Siden f.eks. sensur er *nødvendig* for å sikre menneskers liv og helse, er det fra Hobbes’ ståsted absurd å hevde at dette f.eks. truer menneskets sikkerhet og mulighet til å fø seg selv<sup>9</sup>. Dette er imidlertid muligens *for* enkelt, da mennesker som ikke er dominert av sikkerhetsbehov ifølge Maslow *ikke* har det samme synet på usikkerhet. Hvorvidt de *burde* ha det er én ting, men Maslow ser ut til å mene at det faktisk at de faktisk “glemmer” at slike behov er viktige når de er tilfredsstilt er den fremste grunnen til uro og konflikt (Maslow 1987:179).

Muligheten for at visse behov dominerer andre kan få store konsekvenser for Hobbes’ generelle teori. Det er på bakgrunn av dette at jeg ønsker å formulere en noe annen forståelse av mennesket som jeg vil bruke for å revidere Hobbes’ resonnement i 5.3. Dersom vi for øyeblikket antar at Hobbes har rett i at mennesker utsatt for sterk usikkerhet ville ønske og godta staten slik Hobbes foreskriver den, må vi se på muligheten for at mennesker etter statsdannelsen *endres*. Det de ønsker i naturtilstanden – grunnen til at staten dannes og får den formen Hobbes ønsker – blir glemt, og erstattet av andre behov. Dette fører til at individene i denne staten vil kunne tenkes å være svært lite fornøyd med Hobbes’ stat etter at de har glemt hvor viktig sikkerhet engang var.

Maslow diskuterer dette i forbindelse med *frihet* på følgende måte:

Whether or not this particular desire is universal we do not know. The crucial question, especially important today, is: Will men who are enslaved and dominated inevitably feel dissatisfied and rebellious? We may assume on the basis of commonly known clinical data that people who have known true freedom (not paid for by giving up safety and security but rather built on the basis of adequate safety and security) will not willingly or easily allow their freedom to be taken away from them. But we do not know for sure that this is true for people born into slavery (Maslow 1987:63n).

---

<sup>9</sup>Sensur og andre elementer ved Hobbes’ stat diskuteres i 4.1.1

Dette vil diskuteres nærmere i 5.3, der jeg vurderer Hobbes' resonnement ut fra en alternativ modell av mennesket, samt innstilling til usikkerhet og risiko generelt. Mennesker synes i visse situasjoner å være i stand til å verdsette frihet høyere enn sikkerhet, og det synes klart at dette vil kunne ha store konsekvenser for Hobbes' generelle teori om staten.

### 3.3.5 Oppsummering

Avslutningsvis vil jeg konkludere med at Hobbes og Maslow har mange likheter, men også noen viktige forskjeller. Jeg oppfatter denne gjennomgangen som bekreftende for mange av Hobbes' teorier om mennesket, og anser med det hele hans teori som noe styrket. Det er viktig å huske at *Leviathan* ble skrevet på midten av 1600-tallet, og det faktum at Hobbes var såpass nær det vi i dag anser som "sannheten" om menneskets natur, fremstår som imponerende.

Det er dog som sagt også forskjeller mellom Maslow og Hobbes, og disse forskjellene er viktige. Spesielt teorien om at mennesker endres fundamentalt når de domineres av forskjellige typer behov, synes svært interessant. Til slutt er det også fristende å påpeke en teori som åpenbares ved å lese Hobbes og Maslow side om side, selv om jeg på ingen måte her skal forsøke å psykoanalysere Hobbes; måten Maslow beskriver en person som er dominert av sikkerhetsbehov på, kan vanskelig benektes å passe svært godt på hvordan Hobbes fremstår. Det faktum at Hobbes levde i en tid med sterke konflikter og trusler mot statens autoritet gjør det nødvendig å vurdere hvorvidt Hobbes' teori som en helhet kan være noe farget av dette fokuset på sikkerhet<sup>10</sup>.

Anyone who attempts to make an emergency picture into a typical one and who will measure all of humanity's goals and desires by behaviour during extreme physiological deprivation is certainly being blind to many things (Maslow 1987:59).

Det er en ofte fremsatt kritikk av Hobbes at hans teori er situasjonsbestemt, og kun et resultat av hans historiske og kulturelle omgivelser, noe som ikke nødvendigvis er tilfellet. Jeg benekter dog som sagt ikke at det kan finnes *elementer* av Hobbes' teori som er preget av dette, noe jeg undersøker nærmere i kapittel 5.

---

<sup>10</sup>Det kan i samme stil hevdes at mange av dagens frihetsforkjempere er dominert av høyere behov, og overser mange viktige aspekter ved de grunnleggende behovene, da disse ikke er presserende for oss i dag.

## 3.4 Nevroøkonomi og eksperimentell økonomi – empiriske tilnærminger til mennesket

Jeg vil avslutte denne behandlingen av Hobbes' syn på individet og menneskets natur ved å se på funn fra eksperimentell økonomi og nevroøkonomi. Hovedgrunnene bak dette valget er: a) at disse retningene skiller seg betraktelig fra Maslow hva gjelder metode, og b) at dette er svært spennende forskningsfelt som er meget aktuelle.

*Eksperimentell økonomi* er et bredt forskningsfelt som “kombinerer innsikt fra økonomi, antropologi, evolusjonær og menneskelig biologi, sosialpsykologi og sosiologi” (Gintis et al. 2005a:xi). Formålet er å skape et realistisk bilde av menneskelig motivasjon, og som navnet tilsier baseres forskningen i stor grad på eksperimentell metode.

*Nevroøkonomi* er også preget av samarbeid på tvers av disipliner, og eksperimentell økonomi er sterkt representert også her. I tillegg til fagfeltene nevnt ovenfor, er *nevrologi* det definerende aspektet ved nevroøkonomi. Kritikken av reduksjonistisk neoklassisk økonomi er svært sterk blant disse forskerne, men også adferdspsykologi ansees som noe begrenset. Dersom man kan fremsette hypoteser om nevrologiske prosesser, som kan forklare eller understøtte den rene forskningen på faktiske valg, er håpet at dette vil gjøre det lettere både å utvikle bedre teorier og forkaste feilaktige teorier (Glimcher et al. 2009a:4).

Jeg vil ikke begi meg inn på de delene av nevroøkonomien som omhandler hjernens oppbygning og virkemåte, da dette hverken er mitt fagområde eller synes nødvendig for oppgaven; fokuset vil være på de (foreløpige) resultatene man har kommet frem til innen disse to feltene. Dette medfører at nevroøkonomi kan gi langt bedre innsikt i de temaene som her diskuteres enn hva jeg her formidler, men jeg mener allikevel at resultatene i seg selv er interessante i forbindelse med spørsmålet om hvor plausibelt Hobbes' syn på menneskets natur fremstår i lys av nyere bidrag.

### 3.4.1 Bakgrunn

Sometimes, economists wish to predict behaviour under completely novel conditions (for example, a new and untried public policy). There is no assurance that reduced form behavioral models will perform well in such contex-

ts, especially if the novel conditions are qualitatively distinct from any that have preceded them. In contrast, a good structural model, based on a deeper understanding of behaviour, may permit reasonable projections even when fundamental environmental changes occur. Many neuroeconomists believe that their field will provide such models (Bernheim 2009:123).

En fellesnevner for mye av forskningen innen fagfeltene jeg her diskuterer er et fokus på *psykologisk realisme*. Nært beslektet med tradisjonell adferdspsykologi, er det observerbar adferd og empirisk etterprøvbarehet som står i sentrum (Sanders 2002:45). Kahneman (2009:524) påpeker også at neuroøkonomien generelt sett har bekreftet forventningene til adferdspsykologer. Det er imidlertid like viktig å vite *hvorfor* mennesker handler som de gjør som å se på *hva* de gjør, og dette gjør at neuroøkonomi og eksperimentell metode skiller seg fra tradisjonell (spesielt neoklassisk) økonomi. Det er videre et mål at forklaringene som benyttes i økonomiske modeller skal være overførbare til andre aspekter ved menneskelig adferd, noe man oppnår dersom man benytter psykologisk realistiske modeller av mennesket (Glimcher et al. 2009a:4).

The ideas presented in this book are part of a continuous line of intellectual inheritance from Adam Smith and his friend and mentor David Hume, through Thomas Malthus, Charles Darwin, and Emile Durkheim, and more recently the biologists William Hamilton and Robert Trivers. But Smith's legacy also led in another direction, through David Ricardo, Francis Edgeworth, and Leon Walras, to contemporary neoclassical economics, that recognizes only self-interested behavior (Gintis et al. 2005b:3).

Det er forøvrig påfallende hvor mange av forfatterne innen feltet som bruker tid på å angripe "tradisjonell" økonomi, og kritikken tar flere former; Gintis et al. (2005b:6) peker på en antagelse innen økonomi og menneskelig biologi om at individer er "egen-anskuende" ("self-regarding")<sup>11</sup>. Dette innebærer at individer er *isolert egeninteresserte*, noe som ikke må forveksles med hva jeg omtaler som egeninteresse hos Hobbes. Mens Hobbes anser det som selvsagt at individer f.eks. kan hjelpe mennesker de er glade i, da dette gjør dem glade e.l., er dette *ikke* tilfellet ved egen-anskuende aktører (Gintis et al. 2005b:6,33). To implikasjoner følger av denne antagelsen om egen-anskuende aktører: de ser kun på utfall (ikke prosess), og de bryr seg kun om hva de *selv* får, *uavhengig av andres gevinst* (Gintis et al. 2005b:6). Denne forståelsen av mennesket –

---

<sup>11</sup>Se også 3.4.3.

*Homo economicus* – forkastes som empirisk uholdbar, og siden realisme etterstrebes anses en slik forståelse som uanvendbar (Gintis et al. 2005b:7).

It is hard to imagine a less inspiring account of our motives and our prospects. But if the ugly picture the conventional theory paints is *right*, then we'd be fools to avert our eyes from it. [...] It turns out, however, that the conventional theory *isn't* right. Individuals in collective action settings might not behave like saints, but they don't behave like fiends either (Kahan 2005:366).

Gjennomgående omtales bidrag som baserer seg på *Homo economicus*-lignende individer som “standard”, “konvensjonelle”, “mainstream” og “tradisjonelle”, som oftest i en negativ betydning av ordene (Falk og Fischbacher 2005:193, Sanfey og Dorris 2009:64, Gintis et al. 2005b:6, Kahan 2005:366, Bernheim 2009:119, Fehr 2009:215). Svært mange – også økonomer – er imidlertid klar over at mennesker ikke er fullt så enfoldige som *Homo economicus*, og dette gjelder i høyeste grad også Hobbes.

### 3.4.2 Metode

The *experimental economists* began with the viewpoint that economic principles should apply everywhere (as principles in natural and physical sciences are presumed to); their view was that when theories fail in simple environments, those failures raise doubt about whether they are likely to work in more complex environments (Glimcher et al. 2009a:4).

Kontrollerte eksperimenter er dagligdags innen mye forskning, men det har vært noe mindre utbredt i samfunnsvitenskap enn f.eks. innen naturvitenskap, mye grunnet vårt problematiske forskningssubjekt: *mennesket*. Utover problemene forbundet med å forske på organismer som er fullstendig klar over at de blir forsket på, er det åpenbart svært strenge etiske begrensninger på hva man kan utsette mennesker for i et eksperiment. Jeg vil se nærmere på de vanligste innvendingene mot eksperimentell metode i 3.4.4. Jeg er også av den oppfatning at det finnes innvendinger man sjelden ser, som angår temaer som er sentrale for Hobbes. Dette gjelder spesielt spørsmål om død og alvorlig fysisk skade, noe det er problematisk å undersøke i forbindelse med mennesker.

Forskjellige fagfelt har ofte noe forskjellig metodologi. Dette kan medføre visse vanskeligheter, og innen nevroøkonomi og eksperimentell økonomi løses dette i stor grad



ved at de fleste bidragene anvender “et felles spillteoretisk vokabular og en konsistent adferdspsykologisk metode” (Gintis et al. 2005a:xii).

Store deler av bidragene jeg har basert denne gjennomgangen på, som gjengis i samleverkene *Neuroeconomics* (Glimcher et al. 2009b) og *Moral Sentiments and Material Interest* (Gintis et al. 2005a), er basert på spill i laboratorier<sup>12</sup>.

Når det gjelder nevrologiske data er disse for det meste innhentet ved å avfotografere hjerneaktiviteten til subjekter som utfører handlinger i en fMRI-skanner (functional magnetic resonance imaging). Andre metoder kan også anvendes, som f.eks. ved å gi subjektene kjemikalier som påvirker hjerneaktivitet. fMRI fremstår som oftest som det beste alternativet, da det i forhold til alternativene er enkelt å implementere og ikke krever fysiske inngrep i hjernen til subjektene (Houser og McCabe 2009:57). Med fMRI- og PET(Positronic emission tomography)-skanning må subjektene ligge fullstendig rolig, så det er store begrensninger både i forhold til hvilke impulser de kan utsettes for (hovedsakelig tekstuell/audiovisuell) og hvordan de kan respondere. Det er altså ikke realistiske sosiale interaksjoner som testes på denne måten (Sanfey og Dorris 2009:72).

---

<sup>12</sup>Nevroøkonomi og eksperimentell økonomi baserer seg i stor grad på de samme spillene:

- Ultimatum-spillet (UG): Dette spillet har to roller: en tilbyder og en mottager. Tilbyderen har en viss delbar mengde av et gode, og må foreslå en fordeling av dette overfor mottageren. Forslaget er et ultimatum, da mottageren ikke er i en forhandlingsposisjon – han må enten godta eller avslå den foreslåtte fordelingen. Dersom forslaget godtas gjennomføres en fordeling i henhold til dette, men dersom mottager avslår faller godet bort for begge aktører.
- Diktator-spillet (DG): Dette er svært likt UG, men mottager har her ikke veto-makt. Det tilbyderen – diktatoren – foreslår gjennomføres.
- Tillit-spillet (TG): En *investor* starter med et visst beløp, og kan overføre penger til en *forvalter*. De vil i så tilfelle forrente seg. På slutten av spillet kan forvalteren velge hvor mye av det som er investert han ønsker å føre tilbake til investoren. Den første overføringen er dermed en indikasjon på tillit, mens tilbakeføringen er et signal om pålitelighet.
- Fangenes dilemma-spill (PD): Fangenes dilemma er et strategisk spill der hver aktør har to valg – samarbeide eller hoppe av. Gevinstene er symmetriske, og ordnes slik at summen av gevinstene er størst dersom begge samarbeider og minst dersom begge hopper av.
- Fellesgode-spillet (PG): Strukturen er forholdsvis lik som ved Fangenes dilemma, men dette spillet spilles som oftest med flere spillere. Hver spiller kan sette penger på sin private “konto”, eller han kan bidra med alt eller deler av sin formue til fellesskapet. Hver enhet som beholdes privat overføres umodifisert til neste runde, mens beløpet i fellespotten for det meste forrenter seg noe, før det så fordeles likt på spillerne. Dersom alle samarbeider får man i så tilfelle maksimal total gevinst. Spillet illustrerer gratispassasjer-problemet svært tydelig, ved at aktørene kan unnlate å bidra og fremdeles motta fordeler fra andres samarbeid og bidrag.

Ytterligere beskrivelse av de mest brukte spillene finnes f.eks. i Houser og McCabe (2009:55-57), Sanfey og Dorris (2009:64-66), Fehr (2009:218-9), Brosnan (2009:287).

Med aper gjelder tydeligvis andre retningslinjer, da disse ofte får boret et hull i kraniet, hvor det så føres en elektrode inn i den delen av hjernen man ønsker å måle. Denne teknikken gir muligheter for å forske på subjekter i realistiske situasjoner, og gir også langt mer presise målinger (Sanfey og Dorris 2009:67).

Ved å foreta målinger av aktivitet i de forskjellige delene av hjernen, mens subjektene utfører handlinger og utsettes for impulser, håper man å kartlegge hjernen, for bedre å forstå hvordan den fungerer.

En av hovedkonklusjonene fra disse forskningsfeltene er at “standard” økonomi tar feil med hensyn til menneskelig motivasjon. Dette er kanskje ikke spesielt revolusjonerende, ved at *Homo economicus* anses som *nyttig* – ikke *realistisk* – av de fleste økonomer. Det som derimot er interessant er at mange prediksjoner basert på neoklassisk økonomisk teori blir “motbevist” empirisk. Dermed setter man også spørsmålsteget ved *nyttien* av *Homo economicus*, og forskerne har kommet godt i gang med å utvikle alternativer som bedre representerer menneskets natur. Det er disse alternativene jeg ønsker å se nærmere på her.

### 3.4.3 Hovedbegreper

I forhold til formålet med denne oppgaven er det mest sentralt å se på hvilke teorier om mennesket man anser som mest plausible innen disse fagfeltene. Ren empiri og kvantitative analyser er selvsagt ikke uten interesse, men jeg vil her se mest på hvilke konklusjoner ekspertene innen disse feltene selv trekker om forhold som er interessante i forbindelse med Hobbes’ og Maslows syn på mennesket.

#### Sosiale preferanser og heterogenitet

Many influential economists have pointed out that people not only care about their own welfare but also for the well-being of others, and that this may have important economic consequences; these include one of the founding fathers of economics, Adam Smith (Smith 1976), and Nobel Prize winners such as Gary Becker (Becker 1974), Kenneth Arrow (Arrow 1981), Paul Samuelson (Samuelson 1993), Amartya Sen (Sen 1995) and Reinhard Selten (Selten 1998). Throughout most of its history, however, mainstream econo-

mics has relied on the simplifying assumption that material self-interest is the *sole* motivation of *all* people, and terms such as “other-regarding preferences” were simply not part of economists’ vocabulary (Fehr 2009:215).

*Sosiale preferanser* er et begrep som anvendes dersom en aktør inkluderer andres gevinst eller nytte i sin egen nyttevurdering. Andres (*j*’s) gevinst/nytte kan påvirke aktør *i*’s nytte positivt eller negativt, og sosiale preferanser er et samlebegrep som indikerer hvorvidt aktører er ”sosialt anskuende” (“other-regarding”, motsatsen til egenanskuende) (Fehr 2009:216).

At mange individer faktisk har sosiale preferanser er godt belagt, men det hevdes *ikke* at aktører som er egenanskuende ikke eksisterer (Fehr 2009:217). Det synes forøvrig like oppsiktsvekkende at det åpnes for at *Homo economicus* faktisk eksisterer, som at det finnes individer med sosiale preferanser.

Et annet begrep som viser seg å være viktig i denne sammenheng er *heterogenitet* når det gjelder individers preferanser:

However, the evidence also shows that there is considerable individual heterogeneity in social preferences: some people display little or no concern for their interaction partners, while others show strong social preferences (Fehr 2009:217).

Individer med sosiale preferanser fremmes altså ikke som en erstatning for tidligere mer “egoistiske” mennesker – de introduseres som et signifikant og viktig tillegg. Dette medfører at mye eksisterende teori som baserer seg på homogene aktører vil måtte gjennomgås, for å kontrollere hvorvidt deres antagelser og prediksjoner holder dersom man introduserer heterogenitet.

Fehr (2009:217) er klar på at “sosiale preferanser bør anses som en viktig komponent i individers nyttefunksjoner”, og med slike nyttefunksjoner er det mulig å representere svært forskjellige individer når det gjelder *hvor sterkt* andres nytte vektlegges. Der  $U_i = x_i$  illustrerer egenanskuende preferanser, er altså  $U_i = \alpha x_i + (1 - \alpha)x_j$  et eksempel på en nyttefunksjon som kan vise sosiale preferanser.  $x_i$  er her *i*’s gevinst, og  $x_j$  er gevinsten som tilfaller spiller *j*. For at sistnevnte funksjon skal illustrere sosiale preferanser må  $\alpha < 1$ <sup>13</sup>.

---

<sup>13</sup>Jeg hevder i 2.6 at sosiale preferanser bør inngå i en “Hobbessiansk nyttefunksjon”.

## Sterk gjensidighet og urettferdighetsaversjon

[...] strong reciprocity, is an historically evolved form of enlightened self- and kin-interest that falsely appears altruistic when deployed in social situations for which it was not an adaption (Gintis et al. 2005b:25).

Den fremste formen for sosiale preferanser er *sterk gjensidighet*. Fehr og Fischbacher (2005:151) forklarer sterk gjensidighet ved at “individer handler som om deres positive eller negative verdsetting av referanseagentens gevinst avhenger av denne agents handlinger”. Der mye økonomisk teori forutsetter at aktørene utelukkende vurderer *utfallet*, vektlegges her *prosessen*. Spesielt *intensjoner* trekkes frem som viktige. Det er bl.a. vist at aktører reagerer svært forskjellig på identiske forslag fremsatt av et menneske kontra en datamaskin, som oppfattes som intensjonsfri (Brosnan 2009:288).

Falk og Fischbacher (2005:197) fremholder at håndtering av intensjoner er nødvendig i en realistisk modell av mennesket, da motivene og intensjonene som tillegges andres handlinger er viktige for hvordan disse handlingene blir oppfattet. De er samtidig klare på at også utfall er viktig, men Fehr og Fischbacher (2005:153) understreker at sterk gjensidighet *ikke* er drevet av “forventninger om fremtidige økonomiske fordeler”. Sterk gjensidighet formaliseres på følgende måte:

$$U_i = \pi_i + \rho_i \varphi \sigma \quad (3.1)$$

$\pi_i$  er spiller  $i$ 's materielle gevinst. *Gjensidighetsparameteren*  $\rho_i$  representerer  $i$ 's preferanser for gjensidighet; “dersom  $\rho_i = 0$  har spilleren *Homo Economicus*-preferanser, noe som antas i standard spillteori” (Falk og Fischbacher 2005:193). Hvorvidt “standard” spillteori er begrenset til egen-anskuende aktører, og hva som her antas å være “standard” spillteori, vil jeg ikke berøre utover å sette et spørsmålstegn ved utsagnet. *Godhet/snillhet* gis ved  $\varphi$ , og kan være positiv eller negativ. Positiv  $\varphi$  innebærer at  $i$  anser  $j$ 's handlinger å være basert på gode intensjoner (Falk og Fischbacher 2005:194). *Gjengjeldelsesparameteren*  $\sigma$  bestemmer i hvilken grad  $i$  gjengjelder  $j$ 's handlinger, og “som en første tilnærming gis  $\sigma$  rett og slett ved spiller  $j$ 's gevinst” (Falk og Fischbacher 2005:194).

The product of the *kindness term* ( $\varphi$ ) and the *reciprocation term* ( $\sigma$ ) measures the reciprocity utility. If the kindness term is greater than zero, player  $i$

can *ceteris paribus* increase his utility if he chooses an action that increases player  $j$ 's payoff (Falk og Fischbacher 2005:194).

Sterk gjensidighet fremstår som en nyttig modell, da den inkluderer sosiale preferanser og det åpenbare faktum at vurderingen av andres nytte påvirkes av ens *innstilling* til personen. Hobbes er også klar på at (oppfattede) intensjoner, ikke utelukkende utfall, er viktige. Det å gi "liten hjelp" er f.eks. å vanære en person, da det formidler et negativt syn på mottager (Hobbes 1946:58). Hobbes sier også at mennesker er i stand til å skille mellom urett og tilfeldig skade, noe som opplagt innebærer at mennesker ikke utelukkende vurderer *utfall* (Hobbes 1946:111).

En annen modell for sosiale preferanser er *urettferdighetsaversjon*. Selv om denne modellen viser seg å være kvantitativt mindre viktig enn sterk gjensidighet, er den enklere å håndtere, og benyttes ofte som en nyttig tilnærming til *sterk gjensidighet* (Fehr og Fischbacher 2005:154):

Since models of inequity aversion are much simpler and more tractable than models of strong reciprocity, it is often convenient to "mimic" or to "black box" reciprocal behavior by inequity aversion (Fehr og Fischbacher 2005:154).

Ligning (3.1) ser enkel ut, men det blir raskt komplisert når man forsøker å sette seg inn i de forskjellige modifikasjonsvariablene. Brosnan (2009:286) og Falk og Fischbacher (2005:210) modellerer urettferdighetsaversjon som vist i (3.2), her som en  $n = 2$ -modell for spiller  $i$  og  $j$ :

$$U_i(X) = x_i - \alpha_i \max(x_j - x_i, 0) - \beta_j \max(x_i - x_j, 0) \quad (3.2)$$

der  $\alpha_i \geq \beta_i \geq 0$  og  $\beta_i < 1$ .

Ovenfor kan  $\alpha$  forstås som *misunnelse*, siden det med positiv  $\alpha$  vil være negativt for  $i$ 's nytte dersom  $j$  har en større gevinst enn  $i$ .  $\beta$  refereres ofte til som *skyldfølelse*, ved at positiv  $\beta$  fører til lavere nytte for  $i$  dersom han har mer enn  $j$ <sup>14</sup>. Det er imidlertid mulig å argumentere for at mer strategiske hensyn også kan føre til positiv  $\beta$ , noe jeg så kort på i 2.4.2.

---

<sup>14</sup>Brosnan (2009:286) kaller  $\beta$  "guilt", men dette er ikke det samme som "guilt-aversion", en annen teori, fremsatt av bl.a. Charness og Dufwenberg (2006). Det kan muligens være nyttig å anse  $\beta$  som en representasjon av uselvisk rettferdighetssans.

Som vi kan se i (3.2) er det teknisk sett *ulikhetsaversjon* som her demonstreres – ikke *urettferdighetsaversjon* – med mindre man først definerer urettferdighet som ulikhet, noe som ikke stemmer overens med Hobbes’ definisjoner<sup>15</sup>. Jeg vil heretter referere til denne modellen som *ulikhetsaversjon*. Dersom man skal vurdere urettferdighet i en slik modell – noe som absolutt ville vært interessant – måtte man enten vurdere personers *fortjeneste* eller deres formelle *rettmessige krav*. Begge deler hadde vært interessant, men det ville nødvendigvis blitt langt mer komplisert og uhåndterbart. Brosnan (2009) viser til andre mulige ligninger, der f.eks. en Rawlsiansk forståelse av rettferdighet legges til grunn. Se også Fehr (2009:216-7) og Falk og Fischbacher (2005:210-3) for grundige presentasjoner av andre modeller for sosiale preferanser.

Da jeg anser Hobbes’ individer å ha sosiale preferanser, er arbeidet med disse begrepene innen eksperimentell økonomi interessante m.t.p. videre spillteoretisk arbeid med en “Hobbesiansk nyttefunksjon”.

### **Altruistisk straff**

*Altruistisk straff* (og *betinget samarbeid*) inngår under begrepet *sterk gjensidighet* som diskuteres ovenfor. *Altruisme* er et begrep som anvendes svært mye i litteraturen, og jeg ser nærmere på hvordan dette brukes i 3.4.5. I forbindelse med *straff* brukes det her for å representere sanksjonsvirksomhet som utføres selv når det strider mot egenanskuende preferanser; strafferen pådrar seg en kostnad uten noen forhåpninger om direkte gevinst som følge at han straffer (Dayan og Seymour 2009:184).

Begrepet forklares best ved å vise hvordan dette kan måles med et fellesgode-spill, der hver runde består av to steg; først velger man hvor mye man skal bidra, og deretter får spillerne vite hvor mye de andre faktisk bidro. Det spesielle her er at man introduserer sanksjonsmuligheter, slik at hver spiller kan straffe gratispassasjerer (eller andre om de så måtte ønske<sup>16</sup>). Dette er imidlertid ikke kostnadsfritt for “strafferen”, da spiller *i* f.eks. må gi fra seg én enhet for at *j* skal miste tre enheter (Fehr og Fischbacher 2005:169).

Note that since punishment is costly for the punisher, the self-interest hy-

---

<sup>15</sup> “[...] when a covenant is made, to break it is *unjust*: and the definition of *injustice*, is no other than *the not performance of covenant*. And whatsoever is not unjust, is *just*” (Hobbes 1946:93).

<sup>16</sup> “Asosial straff” kan føre til lite heldige effekter av denne typen sanksjonsmuligheter, noe jeg vil se nærmere på i 5.2.2.

pothesis predicts zero punishment. Moreover, since rational players will anticipate this, the self-interest hypothesis predicts no difference in the contribution behavior between a public goods game without punishment and the game with a punishment opportunity. In both conditions, zero contributions are predicted (Fehr og Fischbacher 2005:169).

Det som gjør denne straffen “altruistisk” er altså at den som straffer pådrar seg en kostnad. I et perfekt scenario, der man klarer å isolere et spill fra alle bånd til verden rundt, vil slik straff kunne ansees som “altruistisk”, dersom det kun spilles én periode. Det er imidlertid et spørsmål om slik straff i så fall heller burde betegnes som irrasjonell og hevngjerrig, ved at man i et isolert engangsspill ikke kan anse læring hos den straffede som relevant – man oppnår altså utelukkende å minske begges gevinst<sup>17</sup>.

Det som er interessant med disse sanksjonene er at det er bevist at de kan føre til svært høy grad av samarbeid i fellesgode-spill (Boyd et al. 2005:215, Fehr og Fischbacher 2005:170, Sethi og Somanathan 2005:238). Der man i løpet av spillets gang i vanlige gjentatte fellesgode-spill, opplever minkende bidrag til fellespotten, endres dette drastisk når sanksjonsmuligheter innføres; graden av samarbeid øker raskt og nærmer seg det maksimale. Man opplever også etterhvert at *trusselen* om straff er tilstrekkelig for å opprettholde et høyt samarbeidsnivå (Fehr og Fischbacher 2005:170)<sup>18</sup>. I denne sammenheng kan reaksjoner mot f.eks. urettferdighet ansees som bindinger/trusler – ikke bare faktisk straff (Sethi og Somanathan 2005:238, Brosnan 2009:288).

Hobbess diskuterer sanksjoner, men bruker ordet *straff* kun om “onder påført av offentlig myndighet” med det formål å gjøre undersåtter mer tilbøyelige til å følge loven (Hobbess 1946:202):

From the definition of punishment, I infer, first, that neither private revenges, nor injuries of private men, can properly be styled punishment; because they proceed not from public authority (Hobbess 1946:203).

Hobbess er forøvrig klar over at sanksjoner forekommer også i naturtilstanden, men forventer helt tydelig uheldig sanksjonsvirksomhet her, noe som fører til at han fremsetter

---

<sup>17</sup> Dette bryter åpenbart med Hobbess’ forståelse av hvordan sanksjoner bør utøves, noe som behandles nedenfor.

<sup>18</sup> Dette i seg selv er en god grunn til å spørre seg hvorvidt “altruistisk” er en passende betegnelse for denne type oppførsel. Dersom aktørene individuelt oppnår høyere nytte ved at de håndhever en norm om å samarbeide, kan dette i høyeste grad like gjerne forståes som strategisk og egeninteressert adferd, dersom spillerne (korrekt) antar at også andre håndhever normer.

to naturlover som regulerer dette:

A sixth law of nature, is this, *that upon caution of future time, a man ought to pardon the offences past of them that repenting, desire it* (Hobbes 1946:99).

A seventh is, *that in revenges, that is, retribution of evil for evil, men look not at the greatness of the evil past, but the greatness of the good to follow* (Hobbes 1946:100).

Mennesker nektes altså å straffe vilkårlig, og forbys “å påføre straff med annet formål enn å irettesette overtrederen eller rettledede andre” (Hobbes 1946:100, 203). Begge lovene begrunnes med at de er nødvendige for “sikkerhet i fremtiden” (Hobbes 1946:100). Altruistisk straff utført i et engangsspill vil av Hobbes kunne kalles “ondsinnnet” (cruel) straff, et navn med noe andre assosiasjoner:

Besides, revenge without respect to the example, and profit to come, is a triumph, or glorying in the hurt of another, tending to no end; for the end is always somewhat to come; and glorying to no end, is vain-glory, and contrary to reason, and to hurt without reason, tendeth to the introduction of war; which is against the law of nature; and is commonly styled by the name of *cruelty* (Hobbes 1946:100).

Videre synes det som om Hobbes antar at det som oftest er mulig å identifisere overtreidere, noe som kreves for at altruistisk straff skal være mulig:

And those that deceive upon hope of not being observed, do commonly deceive themselves, the darkness in which they believe they lie hidden, being nothing else but their own blindness; and are no wiser than children, that think all hid, by hiding their own eyes (Hobbes 1946:194).

Hobbes mener straff er nødvendig, siden det ved siden av “ekstraordinær anvendelse av fornuften” er den eneste måten å forhindre kriminalitet som skyldes menneskers “hat, lyst, ambisjon og havepsyke” (Hobbes 1946:194). Hobbes anser heller ikke *utført straff* som det avgjørende elementet når det gjelder bekjempelse av kriminalitet, da *frykt* er den av våre drifter som er mest effektiv i forhold til forebyggingen av lovbrudd (Hobbes 1946:195):

For seeing the aim of punishment is not revenge, but terror [...] (Hobbes



1946:204).

Hvordan sanksjoner arter seg i spillteorien avhenger i stor grad av hvordan man modellerer sosiale situasjoner. Jeg vil i 5.2.2 komme nærmere inn på hvilke konsekvenser muligheten for sanksjoner kan ha, og hvordan dette stemmer overens med Hobbes' filosofi. Det er imidlertid tydelig at Hobbes anser muligheten for at mennesker sanksjonerer på svært lite konstruktive måter som stor, en mulighet som også dekkes innen nyere litteratur under betegnelsen *asosial straff*, som jeg ser på i 5.2.2.

Again, in all deliberations, and in all pleadings, the faculty of solid reasoning is necessary: for without it, the resolutions of men are rash, and their sentences unjust: and yet if there be not powerful eloquence, which procureth attention and consent, the effect of reason will be little (Hobbes 1946:460).

Et sentralt poeng i litteraturen er at altruistisk straff forekommer der individer bryter *sosiale normer* (Silk 2009:273, Sethi og Somanathan 2005:230, Bowles og Gintis 2005:386):

We propose that multi-individual negotiations result in the emergence of social norms that are collectively enforced. [...] cooperation can be stable in large groups, if noncooperators are punished and if those who do not punish noncooperators are also punished (Kaplan og Gurven 2005:95).

Individer straffer altså ikke for å håndheve *lover*, og siden det synes rimelig å anta at lover og normer ikke er fullstendig overlappende, kan man regne med at individer opprettholder sine egne "lover" ved bruk av slike mekanismer. Dette gjør et fullstendig legalistisk syn på det sosiale utilfredstillende, og en kritikk i tråd med dette kan tenkes å ramme Hobbes, på grunn av hans fokus på lover. For de som utformer politikk og lover er det heller ikke enkelt å opprette nye *normer*, og dette problematiske skillet mellom hva styresmaktene ønsker å oppnå og hva mennesker har av iboende preferanser og hva de følger av normer, problematiserer mye av Rational choice-skolens antagelser og forslag.

En inngående diskusjon av normer, sosial kapital og politikktutforming faller noe utenfor mitt formål med denne oppgaven, så jeg henviser her videre til Sethi og Somanathan (2005), Ostrom (2005) og Bowles og Gintis (2005).

## Følelser

I gjennomgangen av Maslow ble det klart at mennesker ikke er – eller bør strebe etter å være – rene rasjonalitetsmaskiner. Jeg hevder at menneskers underbevissthet i mange tilfeller er vel så viktig som hva de bevisst ønsker, noe som gjør det interessant å se på et begrep som *følelser*:

The primary research finding was that the group of participants that viewed the “sad” video (an excerpt from the movie *The Champ*) had an overall significantly higher rejection rate of unfair offers than those who watched either the neutral or the happy clip, indicating a demonstrable effect of negative mood on “emotional” decisions in the UG [ultimatum-spillet]. This is important, as it shows that subtle and transient emotional states, unrelated to the task at hand, can noticeably affect decisions to accept or reject monetary offers (Sanfey og Dorris 2009:77).

En persons humør varierer, og humør viser seg altså å ha en effekt på menneskers beslutninger. Utover den generelle følelsesmessige tilstanden “humør”, oppstår det kontinuerlig forskjellige følelser som reaksjoner på det man opplever. Disse følelsene kan være bevisste (f.eks. magefølelse) eller ubevisste, noe som gjør nevroteknologi svært interessant, ved at man kan observere hjerneaktivitet som ikke nødvendigvis er observerbar for agentene selv.

Forskning på hjerneskade og beslutningstagning har vist at skader som fører til følelsesmessige forstyrrelser har effekter på menneskers beslutningsevne; det er “eksperimentelt demonstrert at følelser spiller en avgjørende rolle i å ta beslutninger” (Sanfey og Dorris 2009:75, Mayr et al. 2009:313). Damasio (2009:210) påpeker videre at følelser ikke nødvendigvis påvirker beslutninger på en *negativ* måte, men at det kan være *positivt*. Hans teori om følelser er basert på “mekanismene som består av drifter og motivasjon, og også teorien om belønning og straff, som er de fundamentale bestanddelene i følelser-maskineriet” (Damasio 2009:210). I “Neuroscience and the Emergence of Neuroeconomics” legger (Damasio 2009) frem sin “somatisk markør”-teori, som også er underbygget av forskning på hjerneskadde. Man har også et begrep som “depressiv realisme”, som går på at deprimerte individer viser seg å være bedre til å vurdere usikre alternativer enn friske individer<sup>19</sup> (Dayan og Seymour 2009:185). Det bør være ukon-

---

<sup>19</sup>De friske individene er systematisk overoptimistiske.

troversielt å anse depresjoner som følelsmessige forstyrrelser, og det faktum at dette har innvirkning på beslutninger er altså svært godt dokumentert.

In the disciplines of psychology and neuroscience, another term that has been used to describe the appraisal or evaluation of events in the service of motivating action is *emotion* (Phelps 2009:233).

Jeg er lite teknisk i denne gjennomgangen av følelser, og ønsker primært å inkludere denne kilden til “irrasjonalitet” fordi Hobbes selv er meget opptatt av følelser; hans verker omtaler appetitter og aversjoner, drifter, begjær, kjærlighet, affeksjon osv., begreper jeg mener ligger nærmere forklaringer som legger vekt på følelser og ubevisst motivasjon enn hyperrasjonelle modeller der kun det valg som gir mest profitt, avkastning e.l. fremmes som gyldige valg.

### **Instinkter og evolusjon**

Jeg har i 3.3.3 sett på Hobbes’ og Maslows syn på instinkter, og vil her kun se kort på hvordan dette omtales innen eksperimentell økonomi og nevroøkonomi. *Evolusjon* vektlegges sterkt, og instinkter spiller en viktig rolle. Instinkter er, som følelser, en viktig kilde til ubevisst motivasjon.

Evolution appears to have endowed everything from the simplest organisms to us with powerful pre-specified, but inflexible alternatives (Konorski 1967, Dickinson 1980, Mackintosh 1983). These responses are termed *Pavlovian*, after the famous Russian physiologist and psychologist Pavlov. Immediately available rewards, such as food or water, and immediate threats, such as pain or predators (collectively called *unconditioned stimuli*), elicit a range of apparently unlearned, typically-appropriate, so-called consummatory, responses (Dayan og Seymour 2009:181).

Det er her sentralt å påpeke at instinkter/Pavlovianske reaksjoner kan stå i et spenningsforhold med strategiske og instrumentelle vurderinger. Dette påvirker prediksjoner fremsatt på bakgrunn av for individet eksterne objektive forhold, som f.eks. ved maksimering av en viss verdi-/nyttefunksjon. Dayan og Seymour (2009:182) har gjennomført meget interessante eksperimenter med duer, og de er i stand til å fremtvinge konflikter mellom instrumentelle vurderinger og Pavlovianske reaksjoner.

The consummatory responses associated with aversive outcomes appears to be more sophisticated than those for rewards, including increased heart rate and sweating during acute pain, fighting in the midst of a contest, and leg flexion in the face of foot-shock (Dayan og Seymour 2009:181).

Et essensielt poeng er at man ved økonomiske eksperimenter og etisk forsvarlig forskning ikke vil være i stand til å vurdere de sterkeste menneskelige instinkter – de som søker overlevelse. Såvidt meg bekjent er det umulig å gjennomføre eksperimenter med mennesker som fullt ut kan forsvare sitater som det følgende:

That is, an option which is presented as involving sure deaths will automatically engage a Pavlovian aversive withdrawal response decreasing its propensity to be chosen, that is absent for the option involving sure survival (Dayan og Seymour 2009:185).

Å bli presentert for fiktiv sikker død kan umulig hevdes å være det samme som å bli stilt overfor en troverdig situasjon der man faktisk er overbevist om at sikker død vil følge. Hvordan man har kommet frem til konklusjonen i sitatet ovenfor er noe uklart. Jeg mener allikevel at det er svært viktig å fremsette hypoteser også om disse aspektene ved mennesket, selv om man ikke får testet dem på linje med f.eks. menneskers økonomiske adferd. At mennesker har et svært sterkt forsvarssystem for å sikre overlevelse gjennomgås i detalj i boken *On Combat: The Psychology and Physiology of Deadly Conflict in War and in Peace* (Grossman 2008).

Alt i alt virker den tidligere gjennomgangen av instinkter i 3.3.3 forenlig med mye av litteraturen om emnet innen eksperimentell økonomi og nevroøkonomi. Instinkter ansees i hovedsak som gunstige mekanismer som har oppstått ved naturlig utvalg og evolusjon, og som nå arbeider *for* menneskets beste<sup>20</sup>.

#### 3.4.4 Metodologiske innvendinger

Det er ikke mitt formål å kritisere nevroøkonomi og eksperimentell økonomi her. Jeg er imidlertid av den oppfatning at ikke alle forklaringene som fremsettes innen disse fagfeltene er uproblematisk, og vil ikke basere meg fullstendig på begrepene som blir anvendt i disse forskningsgrenene senere i oppgaven. Det finnes alternativer til de for-

---

<sup>20</sup>Alternativt *gruppens* beste, men poenget blir det samme.

klaringene som vanligvis anvendes i disse fagfeltene, og disse alternativene kan gjøre det mulig å forklarere de empiriske resultatene på måter som stemmer bedre overens med Hobbes' filosofi. Maslow stilte seg også tvilende til hvorvidt man ved laboratorieforsøk faktisk er i stand til å måle det man faktisk ønsker:

It was the beautiful program of Watson that brought me into psychology. But its fatal flaw is that it's good for the lab and in the lab, but you put it on and take it off like a lab coat. [...] behaviour in the human being is sometimes a defense, a way of concealing motives and thoughts, as language can be a way of hiding your thoughts and preventing communication. [Maslow in Lowry, 1979, vol II. pp 1059-1060] (Frager 1987:37)

### **Anonymitet og oppdemming av spill**

Som vist i kapittel 2 har Hobbes et syn på mennesker som utelukkende egeninteresserte, i den forstand at de bevisst eller ubevisst søker egen fordel. Med bakgrunn i dette vil jeg vurdere hvorvidt en del av forklaringene som baserer seg på altruistiske modeller av mennesket er sårbare for kritikk, i første omgang metodologiske innvendinger.

Therefore [pga. instrumentell søken etter rykte], the one-shot character and the anonymity in simple social preference experiments are crucial for the clean documentation of social preferences. Repeated interactions and a lack of anonymity are confounds that need to be eliminated if one is seeking a clean measure of social preferences (Fehr 2009:219).

Jeg vil nedenfor behandle opparbeidelse av *rykte* som en sentral alternativ forklaring på en del av fenomenene som demonstreres innen eksperimentell økonomi. For at dette *ikke* skal være en gyldig innvending er man i laboratorieforsøk avhengig av å skape *anonymitet*<sup>21</sup>.

Dersom jeg, som subjekt i et eksperiment, *ser muligheten for* at andre på et eller annet tidspunkt vil kunne observere min adferd i spillet, vil dette åpne for at jeg handler på tilsynelatende "altruistiske" måter av strategiske grunner. Når jeg sier "ser muligheten for", er det fordi selv små sannsynligheter vil være av interesse her. Det er

---

<sup>21</sup>Det hevdes ofte at *engangsspill* er tilstrekkelig, men dette kan problematiseres, ved at mennesker lever videre etter spillets slutt, og de kan også ha problemer med å forholde seg til det begrensede formålet i slike spill, noe jeg kommer tilbake til nedenfor.

også utelukkende snakk om *subjektiv* oppfattelse av anonymitet. Hvorvidt subjekter *faktisk* er anonyme er fullstendig irrelevant, da det kun er deres egen oppfatning av dette som styrer deres adferd.

Double blind protocols affect behaviour in single-play dictator and trust games, establishing that people are more cooperative when third parties can know their decision [...] (Smith 2009:17).

[...] the prospect of signaling “good character” through one’s charitable contributions is an important motivator of seemingly altruistic behavior, as in Harbaugh (1998a, 1998b) and Ariely et al. (2007). Such motives should be particularly aroused when the charitable behavior is public. Interestingly, recent psychological work provides at least indirect evidence that it takes only very little to create a “public” situation. Bateson et al. (2006) showed that a poster depicting a pair of eyes led to an almost three-fold increase in prosocial behavior (payments into a shared coffee fund) compared with a control situation involving a picture of flowers. While the cause is rather trivial, this is a dramatic effect (Mayr et al. 2009:316).

Det skal altså svært lite til for å skape en “offentlig” situasjon, og dersom oppfattelsen av dette er tilstede – og det synes plausibelt å anta at dette ikke behøver å bli oppfattet på et bevisst nivå for å ha en effekt – vil strategiske og egeninteresserte motiver forbundet med elementer utenfor spillet kunne utgjøre en substansiell del av subjektens adferd.

As a theoretical complication, it is also worth considering the so-called “objective self-awareness” theory (Duval and Wicklund 1972). Supported by a host of empirical evidence (Silvia and Duval 2001), this theory states that cues that focus people’s attention on themselves (e.g., mirrors, cameras, or onlookers) also bring their own standards into the foreground (Mayr et al. 2009:317).

Det kan også tenkes at mennesker har en viss tendens til å undervurdere egen anonymitet. Dersom anerkjennelse er svært viktig for dem, og det å bli lagt merke til er avgjørende for deres selvfølelse, synes det plausibelt å anta at man innbiller seg at man er mindre anonym enn man egentlig er. Videre er fullstendig anonymitet like nødvendig overfor undersøker/eksterne personer som overfor mot-/medspiller, da muligheten for å “kommunisere” utover spillets rammer er tilstede i begge tilfeller.

Når det gjelder engangspill er det et noe annet – men relatert – fenomen som spiller inn: *oppdemming* av spillet. Dette gjelder som nevnt også anonymitet, da forskeren utilsiktet kan påvirke eksperimentet, selv om eksterne forhold ideelt sett ikke skal ha noen innvirkning på aktørene i et lukket eksperiment. Kanskje enda viktigere er det imidlertid at man klarer å skape et *reelt* og *korrekt oppfattet* engangspill der dette forsøkes.

In dictator games, altruistic behaviour substantially increases when people are informed that a second undefined task will follow the first. This discovery shows how strongly people can be oriented unconsciously to modify their behaviour if there is any inadvertent futurity in the context (Smith 2009:17).

Enhver mangel på fullstendig anonymitet vil kunne føre til at subjekter ser konsekvenser utenfor spillsituasjonen; *de spiller ikke lenger et engangspill*. Selv om subjekter instrueres til å la være, fremstår det som svært sannsynlig at aktører allikevel vil vurdere oppfatningen av seg selv; dersom forskere eller andre kan identifisere dem, vil muligheten for å imponere, eller “kommunisere”, være tilstede.

Det er ytterligere problemer forbundet med oppdemming, ved at personer kan velge å “spille” langt utover spillets grenser. “Lekkasjen” i spillet behøver altså ikke å stoppe ved effektene av å bli observert av forskeren. Man er også klar over at forskningen man er en del av vil analyseres, publiseres og sannsynligvis vil få konsekvenser langt utenfor laboratoriet. Man kan tenkes å ha en interesse i å handle på en viss måte, ved at man ser for seg at dette kan føre til endringer i samfunnet generelt. Kostnaden man da må bære for å handle suboptimalt i en spillsituasjon er for det meste meget begrenset, og ofte ikke engang reell.

Videre er det et poeng at ytterst få – om noen – reelle situasjoner mennesker opplever kan betegnes som engangspill. Dette gjør det nødvendig å vurdere *hvorvidt mennesker overhodet er i stand til å utelukke alle fremtidige betraktninger*, som etter all sannsynlighet helt automatisk gjennomføres i våre daglige valg og gjøremål<sup>22</sup>:

Sometimes it is argued that reciprocal behavior in anonymous one-shot experiments is due to subjects' inability to adjust properly to one-shot

---

<sup>22</sup>Dette resonnementet kan sammenlignes med Hayeks påstand om at mange menneskelige følelser og instinkter langt på vei er etterlevninger tilpasset den tiden vi levde i små og tette samfunn, noe som fører til at vi er ukomfortable og noe ineffektive i dagens anonymiserte samfunn (Hayek 1988).

interactions. One idea (see Binmore 1998) is that outside the laboratory, subjects are typically involved in a network of repeated interactions. It is well-known from repeated game theory that in repeated interactions, rewarding and punishing may be in the long-term self-interest of an individual. According to this argument, subjects who routinely interact in the repeated game of life import routines and habits that are appropriate for repeated interactions into the laboratory's one-shot situation because they do not understand the strategic differences between one-shot and repeated interactions (Fehr og Fischbacher 2005:156).

Hovedpoenget mitt er at det må betegnes som svært vanskelig å skape fullstendig oppdemming av et spill. Dette fører til at strategisk egeninteresserte betraktninger må vurderes der man ikke er fullstendig overbevist om at man har a) klart å skape et perfekt eksperiment med fullstendig opplevd anonymitet, b) fullstendig fravær av fremtidige konsekvenser der man søker å skape engangsspill og c) subjekter som på ingen måte velger å spille for å oppnå noe annet enn det forskeren ønsker at subjektet skal forfølge. At slike lekkasjer i spillet kanskje helst gjør seg gjeldende i subjektene underbevissthet gjør det ekstra vanskelig å avfeie denne kritikken.

Følgene av dette er at en del forklaringer basert på at oppdemming faktisk *er* oppnådd, som ofte ser på begreper som altruisme fordi langsiktig strategisk egeninteresse antas å være utelukket, kan betviles. Jeg mener slik tvil er på sin plass, og at "Hobbesianske" forklaringer basert på egeninteresse ikke kan ansees som motbevist, selv om *Homo economicus* på mange måter *er* motbevist.

### **Kontrolleffekter og etiske begrensninger**

The most obvious shortcoming is that subjects may behave differently in laboratory and in "real word" settings (Loewenstein 1999). [...] When subjects *can* think, so-called "experimenter effects" are common (Gintis et al. 2005b:5).

Kontrolleffekter er et opplagt problem ved denne typen eksperimenter, og er et samlebegrep for "forsøkskanineffekt, intervjuereffekt, instrumenteffekt, konteksteffekt etc." (Hellevik 2002:362). "Hawthorne"-eksperimentene har vist at vi er særdeles naive dersom vi antar at forskerens tilstedeværelse og oppfattede intensjoner er uviktige. Selv



om man er klar over at kontrolleffekter er et problem, synes det problematisk å forsikre seg om at man kan eliminere dem.

Videre er det spesielle etiske problemer forbundet med forsøk med mennesker:

Siden eksperimenter krever aktiv inngripen i sosialt liv på en helt annen måte enn ikke-eksperimentelle metoder vanligvis gjør, oppstår det spesielle forskningsetiske problemstillinger i tilknytning til eksperimentelle design. Det er opplagte grenser for hva forskere kan tillate seg å utsette andre mennesker for i forskningens navn (Skog 2004:80).

I tillegg er “problemet” med etiske begrensninger langt større i forbindelse med min oppgave enn ved forskning på økonomisk aktivitet, som ofte er hovedfokus innen eksperimentell økonomi. Et gjennomgående tema i oppgaven er det faktum at død, og stor fare for død, for Hobbes er svært sentralt, noe som også støttes bl.a. av Maslow (1987) og Grossman (2008).

Nevertheless, partly for ethical reasons having to do with the undesirability of actually relieving human subjects of part of their own wealth in an experiment, it has been very hard to study truly aversive learning and processing in a human economic context (Dayan og Seymour 2009:175).

Jeg vil hevde at mulighetene for fysisk skade og død umulig kan formidles på en troverdig måte i et forskningsetisk forsvarlig eksperiment. Dette innebærer at det er svært vanskelig å lære hvordan mennesker faktisk reagerer på slike fremtidsutsikter og valgmuligheter. I lys av dette har man i realiteten to valg: a) man kan ignorere disse aspektene ved menneskets liv og fremsette teorier og hypoteser utledet fra eksperimentelle data som universelle, eller b) man kan erklære sine konklusjoner som gyldige kun innenfor visse avgrensede situasjoner i menneskers liv, f.eks. økonomiske eller trivielle samhandlinger. Valg b) er langt mindre spennende, men jeg har store problemer med å akseptere valg a) som legitimt.

Det er allikevel svært viktig at forskerne fremsetter hypoteser også om disse aspektene av menneskets psykologi. Selv om eksperimentell metode muligens er lite egnet til å forske på menneskers innstilling til død og skade, synes det bedre å arbeide med antagelser om temaet enn å ignorere det.

### 3.4.5 Alternative forklaringer

Avslutningsvis vil jeg se på noen alternativer til de teoriene og forklaringene som synes mest utbredt innen eksperimentell økonomi. Disse alternativene er også synlige innen nevroøkonomi og eksperimentell økonomi, men de er noe mindre fremtredende enn hovedteoriene jeg har gjennomgått. Det interessante med de alternative teoriene jeg her vurderer, er at de åpner for at Hobbes' syn på mennesket fortsatt kan ansees som plausibelt.

#### “Godfølelse-altruisme”

**altruisme**, en, (fr. *altruisme* dannet av Comte til *autrui* ‘næste’ under innflydelse av lat. *alter* ‘den annen’), mest filos. næstekjærlighet; det etiske prinsipp å ofre sig for andre (mots. *egoisme*) (Norsk riksmålsordbok)

Again, he is only giving because he likes the feeling of giving (Mayr et al. 2009:309)

“Altruisme” brukes mye i forskningsfeltene jeg her behandler, og begrepsbruken varierer *svært* mye; *altruisme* brukes ofte i forbindelse med handlinger som helt opplagt kan forståes som motivert av egeninteresse, som f.eks. “gjengjeldende altruisme” (“reciprocal altruism”) og “ren altruisme” (“pure altruism”). “Biologer definerer altruisme som enhver adferd som er kostbar for aktøren og fordelaktig for mottageren” (Silk 2009:270). Når det så brukes når det gjelder “gjengjeldende altruisme”, som kort fortalt dreier seg om et udefinert bytte av tjenester (som stell blant aper), er det tydelig at begrepet ikke utelukker hva jeg anser som egeninteresse, selv om det i første instans kan synes å være individuelt ufordelaktige handlinger. Dersom man er av den oppfatningen at andre vil gjengjelde tjenesten man utfører senere, med mindre man selv er med på å bryte ned normen som tilsier gjengjeldelse, er det forenlig med egeninteresse å yte tjenesten. Individuelt, og på kort sikt, kan slike handlinger synes altruistiske. Men ser man på lengre sikt, og *gruppemessige* fordeler ved slike ordninger, kommer egeninteressen sterkt tilbake<sup>23</sup>.

---

<sup>23</sup>Den evolusjonære teorien om “gruppeutvalg” hevder at *naturlig utvalg* ikke bare skjer på individuell basis, men at vellykkede og veltilpassede *grupper* kan overleve til tross for at individene hver for seg muligens ikke kan anses som optimale. Dette diskuteres svært mye i litteraturen, og virker som en interessant forklaring på at adferd som ikke kan anses som (utelukkende) “egoistisk”, faktisk er svært utbredt; “So long as they can identify one another, trustworthy fair reciprocators actually achieve

Det synes forøvrig nyttig å anvende begrepet “kortsiktig altruisme” dersom man snakker om handlinger som pr. definisjon er altruistiske, men som allikevel fremstår som motivert av langsiktig egeninteresse. Taylor bruker dette begrepet, og hevder at “hver individuelle handling i et system preget av gjensidighet er *som oftest* karakterisert av en kombinasjon av hva man kan kalle kortsiktig altruisme og langsiktig egeninteresse” (Taylor 1982:28).

Nyttefunksjonen  $u_i = u_i(x_i, G)$ , der  $x_i$  er eget private gode mens  $G$  er et fellegode, sies å representere “ren altruisme”, siden  $G$  gir  $i$  økt nytte (Mayr et al. 2009:305). Det er imidlertid viktig å merke seg at man fremdeles snakker om individer som maksimerer sin egen nyttefunksjon.

Perhaps people not only care about the social problem, but also feel particularly good when they personally take voluntary action to help with the solution. As we will show, this can be modeled with a change in the utility-function, as is done in “warm-glow” models of altruistic behavior (Mayr et al. 2009:307).

Mens ren altruisme ikke forutsetter noen vektlegging av *prosess*, har man med “godfølelse-altruisme” (“warm-glow altruism”) – *uren altruisme* – inkludert prosess i nyttefunksjonen:  $u_i = u_i(x_i, g_i, G)$  (Mayr et al. 2009:308). Her representerer  $g_i$  eget bidrag til fellesgodet, og denne modellen er interessant fordi man kan oppnå nytte ved å  $g_i$  – man kan få en “godfølelse”. Årsaken til denne godfølelsen kan tenkes å være koblet til signalisering av makt og etableringen av rykte ved “finansielt imponerende og moralsk beundringsverdige handlinger” (Mayr et al. 2009:317). Dette gjør en slik modell godt egnet til å fremstille Hobbes’ filosofi som jeg har lagt frem i kapittel 2. Det kan imidlertid også tenkes at det ikke er bevisste og strategiske grunner bak slik adferd, noe jeg vil se på under “Belønningsteori” nedenfor.

## Indirekte gjensidighet

While the lack of specific balance contradicts reciprocal altruism, the presence of general balance is consistent with indirect reciprocity (Alexander 1987, Boyd and Richerson 1989) or costly signaling, where the return be-

---

higher material rewards over time than rational egoists! In other words, they can flourish” (Ostrom 2005:265).

nefit to the donor is food. If the return benefit is in another currency, such as increased mating opportunities, then a lack of general balance is not inconsistent with costly signaling (Kaplan og Gurven 2005:82).

Gjengjeldende altruisme betviles ofte med bakgrunn i manglende oversikt over ytelse og gjenytelse. Dersom jeg gir deg mat nå, krever teorien at jeg på et senere tidspunkt får igjen mat av *deg*. Dette ville i så tilfelle vært et helt opplagt tilfelle av en implisitt kontrakt, der ytelse og gjenytelse skjer på forskjellige tidspunkt, og burde muligens ikke kalles altruisme. Det er imidlertid meget interessant å se på muligheten for at det eksisterer sosiale strukturer/normer som fører til at slik adferd, med tidsforskyvninger av ytelse og gjenytelse, kan fungere langt bredere, og mindre spesifikt, enn hva man forventer med gjengjeldende altruisme, og dette dekkes av *indirekte altruisme*. Denne teorien om sosial samhandling synes i mine øyne svært plausibel, og den er også forenlig med en teori om egeninteresserte mennesker.

Indirect reciprocity, in Alexander's scheme, results when "the return explicitly comes from someone other than the recipient of the original beneficence" (Smith og Bird 2005:128)

## Belønningsteori

Therefore, a common reward mechanism is a crucial component of this system, and a large focus of the broader neuroeconomic endeavor in recent years has been to illuminate the neural processes involved in the encoding and representation of reward, and how these mechanisms may in turn underlie standard models of economic choice such as utility theory and its variants (Sanfey og Dorris 2009:73).

Hjernens "belønningssystem" kan på mange måter forstås som vårt interne insentivsystem. Disse insentivene regnes som "essensielle for individets og geners overlevelse, og understøtter elementære prosesser som å drikke, spise og reprodusere" (Schultz 2009:323). Belønning forekommer imidlertid også i forbindelse med andre menneskelige gjøremål, og kan forklare alt fra "leting etter mat til handel med aksjer på børsen" (Schultz 2009:323).

We can define rewards as objects or events that generate approach and consummatory behavior, produce learning of such behavior, represent positive

outcomes of economic decisions, and may engage positive emotion (Schultz 2009:323).

Uavhengig av årsakene til at man “selv-belønner” en viss type adferd, mener jeg handlinger må forståes som motivert av egeninteresse dersom den belønningen som (bevisst eller ubevisst) forventes, er årsaken til adferden. Denne typen forklaring er svært lik forklaringer basert på godfølelses-altruisme, og stemmer også godt overens med Hobbes’ syn på mennesket og dets appetitter og aversjoner.

## Rykte

Moreover, a skeptic might argue that because face-to-face interactions inevitably involve an individual’s reputation, the observed behaviors represent a combination of social preferences and *instrumental reputation-seeking*. The desire to acquire a reputation that is profitable in future interactions is a purely self-regarding motive that has nothing to do with social preferences, i.e., it represents a confound (Fehr 2009:219).

Jeg har i 2.4.6 etablert at *rykte* og *omdømme* er svært sentrale begreper for Hobbes. Jeg har argumentert for at oppnåelse av rykte og omdømme kan gi makt, og forfølgelsen av slike mål må dermed betegnes som i alle fall delvis – kanskje helt og holdent – egeninteressert.

Også her er det forbindelser til godfølelse-altruisme, ved at adferd som gir en slik godfølelse ofte “innebærer overveielser av et publikum og gjentatte interaksjoner”, som dermed gjør prestisje og signaleffekter viktige (Mayr et al. 2009:317).

Et fokus på rykte henger naturlig sammen med en sosial situasjon, da fullstendig anonymitet forhindrer muligheten for å skape et rykte. Jeg har i vurderingen av innvendinger mot eksperimentell metode sett på hvor vanskelig det kan være å etablere fullstendig anonymitet. Videre, selv med fullstendig *individuell* anonymitet kan personer vurdere sin *gruppes* rykte, dersom dette kan tenkes å fremkomme i resultatene fra eventuelle eksperimenter/forsøk.

Rykte forbindes også med spillteori, der det i stor grad er snakk om rykte som andres oppfatninger av hvilken *type* spiller en aktør er (Camerer 2009:204). Dersom *i* har et rykte for å være pålitelig og samarbeidsvillig, øker mulighetene for at også *j* vil tørre

å samarbeide. Har *i* rykte på seg for å være en “mad dog” i et Chicken-spill, øker sannsynligheten for at spiller *j* vil vike. Viktigst i forbindelse med samarbeid er det faktum at man opparbeider seg et rykte om pålitelighet og velvilje hos andre, noe som i mange tilfeller vil kunne føre til økte muligheter for gevinster.

Det bør generelt sett være lite problematisk å hevde at rykte kan ha en avgjørende effekt på menneskelig interaksjon, og denne typen forklaringer er sentrale i en teori om mennesker som utelukkende egeninteresserte (men ikke utelukkende egen-*anskuende*).

### Kostbar signalisering

[...] we argue that many of the phenomena classed as strong reciprocity (as defined in chapter 1) might be individually optimal (i.e., produce a net fitness benefit) and thus not require cultural or genetic group selection, at least not for their evolutionary origin (Smith og Bird 2005:115).

Adferd som ofte fremstilles som altruistisk eller irrasjonell ut fra et snevert og kortsiktig syn på egeninteresse, kan, som jeg har påpekt i forbindelse med teoriene ovenfor, gjerne være motivert av en noe bredere egeninteresse. *Kostbar signalisering* er en teori om *spredning av informasjon om evner og egenskaper*. Denne informasjonen må være nyttig både for den som signaliserer og for de som fanger opp signalene, og kostnaden ved å signalisere må være knyttet til den egenskapen som formidles (Smith og Bird 2005:116)<sup>24</sup>.

Det er en forskjell på den mer direkte strategiske søken etter makt via rykte som ble diskutert ovenfor og kostbar signalisering:

For example, when we say that signal observers may use the information they have received to choose someone as a (future) ally, we are *not* proposing that this is a favor reciprocated to the signaler, any more than a peahen that chooses the peacock with the showiest tail is “paying back” the cock for having expended high signaling costs (Smith og Bird 2005:117).

Det som kanskje er mest interessant med denne typen signalisering er at den hos mennesker ofte manifesterer seg i produksjonen av fellesgoder, noe som er gunstig for *grup-*

---

<sup>24</sup>Dette kan skje enten ved at signalisering er dyrere for individer med dårlig kvalitet når det gjelder egenskapen som signaliseres, eller eventuelt ved at signalene deres er mindre effektive (Smith og Bird 2005:116).

*pen* man tilhører, ikke kun individet. Kaplan og Gurven (2005:85) diskuterer f.eks. hvordan man i forskjellige kulturer har individer som skaffer gruppen over gjennomsnittlig mengde mat, uten å få betalt for dette på andre måter enn at de blir oppfattet som meget nyttige og av “høy kvalitet”, noe som igjen skaffer økte muligheter for reproduksjon osv. Det er i slike tilfeller noe vanskelig å trekke et skarpt skille mellom kostbar signalisering og instrumentell søken etter rykte, da de masseproduserende individene i det foregående eksempelet kunne forklares ut fra begge teorier. En viktig forskjell er som nevnt at kostbar signalisering krever at også *mottager* må nyte godt av signalet som sendes ut – i dette eksempelet kan det tenkes at det øker muligheten for at mottagerne av signalet er i stand til å skille ut genetisk “gode” partnere – mens dette ikke kreves der man ser på opparbeidelse av rykte som motivasjon.

Det er også hevdet at signalisering via produksjon av fellesgoder er gunstig på grunn av at mediet er godt egnet til å nå et bredt publikum (Smith og Bird 2005:122). Som i mange andre konkurransesituasjoner er det også en fare for *eskalering*: “signaliseringsnivåer har en tendens til å spinne ut av kontroll, som i et våpenkappløp der man forsøker å overgå sine motstandere” (Smith og Bird 2005:124).

### 3.4.6 Oppsummering

Eksperimentell økonomi og nevroøkonomi har vist seg å være svært fruktbare forskningsfelt, der mange tidligere sannheter har blitt forkastet og mange nye perspektiver har blitt fremstilt. Jeg har bevisst unngått å gå inn i det detaljerte nevrologiske materialet, og har med dette gått glipp av en del mulig informasjon. Jeg anser imidlertid de delene av faget jeg *har* sett på som nyttige og tilstrekkelige i denne sammenheng.

Det viktigste er, etter min mening, de empiriske funnene som er gjort; disse må kunne forklares på en eller annen måte dersom man skal kunne hevde å ha en realistisk modell av menneskelig adferd. De konvensjonelle forklaringene i disse fagområdene motsier i visse tilfeller Hobbes’ syn på mennesket, mens de alternative forklaringene jeg har vurdert er fullstendig forenlige med Hobbes’ filosofi. Det blir derfor svært interessant å følge utviklingen i disse fagområdene videre, spesielt i forbindelse med videreføringen av det gode grunnlaget for spillteoretiske formaliseringer der sosiale preferanser inngår.

## 3.5 Konklusjon

I dette kapitlet har jeg først vist at min tolkning av Hobbes ikke er uten presedens i sekundærlitteraturen, men det finnes selvsagt uoverenstemmelser når det gjelder hva Hobbes mente om mennesket. Jeg mener dog at min tolkning kan forsvares a) på grunn av at den er basert på en grundig gjennomgang av Hobbes' tekster og belagt med hans egne utsagn, og b) ved at også andre fremtredende forskere har fremmet lignende tolkninger.

Hva angår tolkningen av Hobbes' teori om mennesket isolert sett er dette kun én av to ting jeg ønsket å oppnå i dette kapitlet. "Kontrollen" av Hobbes mot Maslow var meget fruktbar, både på grunn av likheten jeg har argumentert for at eksisterer, og ikke minst på grunn av forskjellene. Spesielt Maslows syn på behovs evne til å *dominere* individer fremstår som sentralt, og vil som nevnt forfølges i kapittel 5. På bakgrunn av likheten mellom disse tenkerne synes det tillatelig å anvende den innsikten Maslow tilbyr, der Hobbes selv ikke har uttalt seg, og der Maslows idéer ikke bryter med Hobbes' syn på mennesket. Det fremstår som nyttig å kunne utvide Hobbes' teorier på denne måten, da det muliggjør konstruksjonen av en "modernisert" Hobbesiansk filosofi som er bedre rustet til å fortelle oss noe om dagens verden.

Nevroøkonomi og eksperimentell økonomi hadde en annen funksjon i denne sammenheng enn Maslow; det er her i hovedsak de empiriske fenomenene som er interessante, ved at de til en viss grad må anses som realiteter og må kunne forklares dersom en teori om mennesket skal være troverdig. Det er innen denne litteraturen et sterkt fokus på å angripe økonomiske modeller av mennesket, og jeg mener disse angrepene ikke er spesielt gyldige mot Hobbes, da det å fremstille ham som en simplistisk Rational choice-økonom – noe som ofte har blitt gjort – i mine øyne innebærer å overse sentrale deler av hans filosofi.

Det har imidlertid blitt påvist betydelige tilbøyeligheter til sosialt gunstig – og fredelig – adferd hos mennesker, og dette oppfattes av noen som problematisk i forhold til Hobbes' tilsynelatende konfliktorienterte syn på mennesket. Jeg har imidlertid påpekt visse svakheter ved forklaringene som ofte anvendes i litteraturen, og mener Hobbes' syn på mennesket *ikke* er uforenlig med langt mer sosial adferd enn hva som forventes med *Homo economicus*-modeller. Hvor konfliktfylt Hobbes' naturtilstand er undersøkes nærmere i 5.2.3.



Alt i alt finner jeg at Hobbes' teori om mennesket kan ansees som forholdsvis plausibel selv i dag. Dette til tross for at den naturlig nok er lite spesifikk på visse områder; den samsvarer ikke med dagens terminologi, og visse elementer kan klart hevdes å være feil. Jeg velger dog ikke å fordømme Hobbes på bakgrunn av dagens kunnskap om psykologi, men velger å se på de grovere linjene, og *der* mener jeg Hobbes kan ha hatt mye rett.

Del II

Samfunnet

# Kapittel 4

## Hobbes' resonnement om statens form

### 4.1 Introduksjon

I know not how the world will receive it [Hobbes' fullstendige doktrine i *Leviathan*], nor how it may reflect on those that shall seem to favor it. For in a way beset with those that contend, on one side for too great liberty, and on the other side for too much authority, 'tis hard to pass between the points of both unwounded (Hobbes 1946:2).

Som med så mange andre tidlige tenkere, er det vanskelig å plassere Hobbes i en av våre moderne kategorier. Er han en utilitarist, protoliberaler eller kanskje en totalitarist? Med litt kreativitet kunne han med bakgrunn i egne utsagn faktisk blitt fremstilt som hvilken som helst av disse, men det ville i så tilfelle ikke være en redelig måte å behandle Hobbes på. Det beste vi kan gjøre når det gjelder tenkere som tenkte før våre typologier, er å se på helheten i deres verker, for så å forsøke så godt vi kan å forstå hva de ønsker å formidle. Deretter må vi vurdere hvorvidt dette er forenlig med noen av våre "båser".

Føllesdal et al. (1996) påpeker et sentralt poeng i tolkningen av tekster: det er ofte store sprik mellom en forfatters originale *intensjoner* med sitt verk og det som formidles gjennom teksten som et selvstendig bidrag. Det er dermed viktig å avgjøre hva ens hensikt med tolkningen er: "er det å forstå og bedømme *verket*, eller er det å forstå,

eventuelt bedømme, forfatterens intensjoner?” (Føllesdal et al. 1996:103).

Det er min hensikt å vurdere hva Hobbes opprinnelig forsøkte å formidle med sine tekster. Utbredte tolkninger, som muligens har hatt stor innflytelse i perioden etter hans liv, vil derfor anses som fullstendig irrelevante i denne sammenheng dersom de tydelig strider mot hva Hobbes sier. Her blir det nødvendigvis en subjektiv oppfatning om hva man forestiller seg at Hobbes mente med sine tekster, og hvilke uttalelser som skal tillegges mest vekt, der han synes å ha kommet med motsigelser.

Dette kapitlet omhandler Hobbes' syn på staten; jeg vil se på hva slags stat Hobbes ønsker, hvordan den kan forstås med moderne terminologi, samt hvordan Hobbes argumenterer for nettopp sin foretrukne løsning. Jeg hevder at Hobbes' idealstat, som presentert i hans verker, er en *autoritær* stat, noe jeg vil forsvare i den første delen av dette kapitlet.

Hovedformålet med kapitlet er å etablere et sett med premisser som gir den konklusjonen at Hobbes' stat er nødvendig. Dette resonnementet er min versjon av Hobbes' rettferdiggjøring av hans type stat, men den er utelukkende basert på Hobbes' tekster. Sitatet i begynnelsen av kapitlet viser at Hobbes muligens anså seg selv som en mellomting mellom de mest liberale og de mest autoritære; med tanke på den tid som har passert siden Hobbes skrev, kan det imidlertid tenkes at det har skjedd en forskyvning i ordenes betydning, og jeg vil nå se hvordan vi i henhold til moderne terminologi kan forstå *autoritarisme*.

#### 4.1.1 Hobbes og autoritarisme

Authoritarian systems are political regimes which are characterized by a concentration of political power in the hands of a small group of elites who are not in any institutional sense accountable to the public. Thus, they lack those traits which distinguish liberal democratic orders – in particular, extensive civil liberties, rule of law, inter-party competition and representative government (Kuper og Kuper 1996:41).

Definisjonen ovenfor passer meget godt til min forståelse av autoritarisme, som igjen stemmer godt med Hobbes' filosofi. Dette er da også grunnen til at jeg velger å karakterisere Hobbes som en autoritær tenker. Vektleggingen av maktkonsentrasjon passer spesielt godt med Hobbes' tanker, selv om det for han ikke er et *krav* at denne mak-

ten skal samles hos få eller én. Enda viktigere er etter min mening det poenget at staten/suverenen ikke kan stilles til ansvar overfor borgerne eller institusjonelle konstruksjoner, noe jeg ser nærmere på nedenfor. Også manglene på “utstrakte sivile rettigheter, rettsstat, partipolitisk konkurranse og representativt styre” kan sies å stemme godt overens med Hobbes’ idéer om staten. Selv om enevelde ikke *kreves*, og konkurranse ikke fullstendig utelukkes, finnes det godt belegg for å hevde at definisjonen ovenfor er en god første tilnærming til Hobbes’ ideal, heretter referert til som “Hobbes’ stat”.

En annen utbredt forståelse av autoritarisme svarer til følgende definisjon fra Østerud:

Slike [autoritære] regimer er mer antipluralistiske enn anti-konstitusjonelle: de tillater ingen reell konkurranse om maktposisjonen, men aksepterer legale bindinger på statsstyret. Regimet kan være underlagt lovene og tradisjonelle institusjoner som religionen og moralen, samtidig som de foretrekker et avpolitisert samfunn (Østerud 2002:109).

I sin videre beskrivelse av autoritære regimer, inkluderer også Østerud trekk som er sentrale hos Hobbes:

Det autoritære regime aksepterer et skille mellom stat og samfunn, hvor det ikke er noe mål å omskape alle sosiale institusjoner etter statlig direktiv. Det forutsetter ingen stram ideologisk ensretting, men vil beskytte den politiske kontrollen mot uro og opposisjon (Østerud 2002:109).

Dette “tillegget” fra Østerud er nyttig, ved at det differensierer den *autoritære* staten fra den *totalitære*, som Østerud beskriver på følgende måte:

En mer vidtgående styringsmodell er det *totalitært* orienterte regime av fascistisk type. Fascismen tillater ikke uavhengige sosiale institusjoner, men krever innordning, tilpasning og full oppslutning. Fascismen er massemobiliserende, og krever politisering av tidligere autonome og ikke-politiserte områder. (Østerud 2002:109).

En fordel ved Østeruds definisjon er altså at autoritarisme og totalitarisme er to disjunkte begreper, mens totalitarisme for Kuper og Kuper (1996) kun er en spesiell type autoritarisme. Det finnes mange likheter mellom autoritarisme og totalitarisme, men en viktig differensierende faktor kan sies å være forholdet til *ideologi*. Den autoritære

staten krever ingen samfunnsmessig gjennomgripende ensretting, noe som må kunne sies å være en svært viktig forskjell. Synet på makt kan nok på overflaten synes ganske likt, men der absolutt og vilkårlig makt i totalitære systemer kan tenkes å forsvares med en ideologi, vil det autoritære regimet kun begrunne dette med at det er nødvendig for å opprettholde ro. Det er altså langt større rom for enighet om statens “formål” i et autoritært regime enn i et totalitært regime, og samtidig som det totalitære system har en mer altomfattende ideologi, har det også generelt sett et visst *krav* om å tilskrive seg denne.

For å se hvorfor Østeruds definisjon på det autoritære ikke passer til Hobbes, er det spesielt interessant å se på Hobbes’ syn på konkurranse og bindinger på staten.

### **Politisk konkurranse og bindinger på staten**

Det er ingen eksplisitt avvisning av politisk konkurranse hos Hobbes. Store deler av hans filosofi kan dog leses som en sterk advarsel mot slik konkurranse (Hobbes 1946:124,170 Se også premiss 4 i 4.2.2). Dette primært fordi menneskers ambisjoner, dersom de f.eks. ser muligheten for å oppnå suverenens stilling, vil være en kilde til konflikt. Konkurranse om posisjoner kan medføre en økt trang i samfunnet generelt til å markere seg, posisjonere seg og strategisk sett handle på en slik måte at en øker egne sjanser for å oppnå denne ypperste stilling i samfunnet, mens en gjør hva en makter for å svekke andres sjanser (Hobbes 1946:111,124,170). Hobbes foretrekker en eneherker som selv velger sin etterkommer, bl.a. for å minimere dette problemet<sup>1</sup>.

Det er imidlertid ikke et *krav* for Hobbes at suverenen skal være en eneherker/monark; demokrati og aristokrati er reelle alternativer, da “suverenen” er en *institusjon* i samfunnet, ikke en person (Hobbes 1946:124). Hobbes selv lener sterkt mot enevelde som den beste styreform, men dette begrunnes utelukkende med at han anser en slik styreform som best egnet til å oppnå målet med staten, som er å skape, og opprettholde, ro og orden (Hobbes 1946:64,87,112,116,219)<sup>2</sup>.

Et annet element som til en viss grad faller under temaet politisk konkurranse, er Hobbes’ syn på sensur, noe han anser som et nødvendig og helt naturlig verktøy for suverenen. Dette elementet er relevant på bakgrunn av det syn at konkurranse er best

---

<sup>1</sup>Mer om styreform under premiss 4 i del 4.2.2.

<sup>2</sup>En diskusjon om statens *mål* følger nedenfor.

dersom den er *reell* og *fri*, noe som langt på vei forutsetter fri tilgang på idéer.

[. . .] it is annexed to the sovereignty, to be judge of what opinions and doctrines are averse, and what conducing to peace; and consequently, on what occasions, how far, and what men are to be trusted withal, in speaking to multitudes of people: and who shall examine the doctrines of all books before they be published (Hobbes 1946:116).

Her ser vi altså at suverenen bestemmer hvilke doktriner som skal tillates. Suverenens midler ligger tilsynelatende i å innskrenke individers tale- og organisasjonsfrihet, samt i å kontrollere utgivelser av trykt materiale.

Spørsmålet er ikke nødvendigvis hvorvidt Hobbes tar feil i å åpne for denne type sensur og innskrenkninger av ytringsfriheten; i så tilfelle ville også dagens liberale samfunn kunne synes noe mangelfulle. Dersom en person kommer med rasistiske ytringer er dette forbudt i dagens Norge, og det samme gjelder f.eks. personen som ønsker å drive en nazistisk organisasjon. "Vel, dette er annerledes", vil man kanskje kunne hevde; det kan imidlertid være at det er nettopp slike tilfeller Hobbes diskuterer.

Formålet med sensur og kontroll over ytringer er å sikre statens overlevelse, det samme målet som gjelder alt annet staten foretar seg (Hobbes 1946:214). At individer som taler eller organiserer seg med det eksplisitte formål å spre hat mot andre grupper, kan neppe sies å fremme fred (Hobbes 1946:116)<sup>3</sup>.

Siden Hobbes ikke setter eksplisitte begrensninger på hvor utstrakt sensuren kan være, antar mange at hans stat faktisk ville benytte seg av dette og at Hobbes' stat derfor er preget av utstrakt sensurvirkosomhet. Det er imidlertid vanskelig å forestille seg at omfattende sensurvirkosomhet ikke ville bli oppdaget, og i så tilfelle ville suverenen selv skape kaos, ved at ingen lenger kunne stole på hva de leste eller hørte. En slik handling ville dermed kunne være i strid med naturlovene og selve formålet til staten.

Selv om Hobbes ikke krever fravær av politisk konkurranse, finnes det altså ikke i hans idealstat med enevelde. Videre, selv der man har en suveren forsamling og konkurranse om deltagelse i denne forsamlingen, vil Hobbes' syn på bl.a. sensur føre til at man kan sette spørsmålsteget ved hvor reell denne konkurransen kan sies å være. Det er altså ikke opplagt at Østeruds utelukkelse av politisk konkurranse i sin definisjon av autoritarisme

---

<sup>3</sup>Dessuten bryter slike aktiviteter med flere naturlover, og spesielt den 8. naturlov, som omhandler krenkelser: "Ingen mann skal ved handling, ord, overveielse eller gest erklære hat eller forakt for en annen"(Hobbes 1946:100).

skaper problemer i forbindelse med å kalle Hobbes autoritær; det er allikevel verdt å påpeke at Hobbes ikke krever et slikt fravær av konkurranse, noe som vil bli tydeligere i neste kapittel.

### **Legale, religiøse eller moralske bindinger på staten**

Mitt viktigste ankepunkt mot Østeruds definisjon av autoritarisme er det punktet at den autoritære staten kan akseptere bindinger. For Hobbes vil det være absurd å hevde at suverenen er bundet av loven, da suverenens vilje *er* loven (Hobbes 1946:116,173):

The sovereign of a commonwealth, be it an assembly, or one man, is not subject to the civil laws. For having the power to make, and repeal laws, he may when he pleaseth, free himself from that subjection, by repealing those laws that trouble him, and making of new; and consequently he was free before. For he is free, that can be free when he will: nor is it possible for any person to be bound to himself; because he that can bind, can release; and therefore he that is bound to himself only, is not bound (Hobbes 1946:173).

Siden makt ifølge Hobbes ikke kan deles (se premiss 3 i 4.2.2), er det ingen mulighet for at andre institusjonelle organer har makt over suverenen. Det maksimale man kan oppnå, er organer som opererer som agenter for suverenen, f.eks. et lovgivende organ som utfører sitt virke med autoritet “på lån” fra suverenen (Hobbes 1946:125). Suverenen kan ikke utskille visse elementer av den totale makten, hevder Hobbes, og suverenen er dermed til enhver tid i stand til å tilbakekalle den makten han har delegert, dersom dette er nødvendig. Dette temaet omtales mer inngående under premiss 3 i 4.2.2 og i 5.3.3.

Videre er heller ikke religiøse bindinger relevante for Hobbes. Hobbes anser religion for å være en av hovedkildene til uro og borgerkrig i hans tid, siden prioriteringen av Guds kontra suverenens ord ikke er tilstrekkelig avklart (Hobbes 1946:384). Store deler av *Leviathan* går med til å “angripe” religioner slik vi kjenner dem, og da spesielt katolisismen. Hans angrep er i stor grad fokusert på det faktum at presteskap og andre med makt over religioner er tilbøyelige til å forme religionen for å fremme sin egeninteresse; at mennesker med makt over religioner også har makt over mennesker er opplagt for Hobbes:

For he that pretends to teach men the way of so great felicity, pretends to



govern them; that is to say, to rule and reign over them; which is a thing, that all men naturally desire, and is therefore worthy to be suspected of ambition and imposture [...] (Hobbes 1946:283)

Siden religion i mange samfunn og tidsaldre ikke har vært underlagt staten, har dette vært en eksklusiv utenomstatlig arena for å tilegne seg menneskers lojalitet og makt, noe Hobbes anser som svært uheldig (Hobbes 1946:283)<sup>4</sup>.

Hva er så Hobbes' løsning på dette? Suverenen tildeles retten til å bestemme i religiøse spørsmål, via "vetomakt" over tolkninger. Mens det å godkjenne en tolkning ikke er det samme som å tolke selv, synes forskjellen relativt triviell her, da han eventuelt kan forkaste enhver tolkning han ikke anser som formålstjenlig (Hobbes 1946:188,246,255). At Guds oppfordringer først kan forståes som "fra Gud" etter at det har passert suverenens kritiske vurdering, bør være nok til å forkaste enhver pretensjon om at religion *binder* Hobbes' suveren.

Religion fremstår for Hobbes på mange måter som et pragmatisk "verktøy" underlagt staten. Med tanke på at menneskets natur er som den er, ser Hobbes tydelige fordeler ved at mennesker er religiøse. Som nevnt i 2.4.4 kan Gud gjøre mennesker dydige, "via doktrine, eksempel og noen ganger på naturlig vis" (Hobbes 1946:280). I tillegg er *frykt* det sentrale for å sikre gunstig adferd hos mennesker, og det synes åpenbart at prospektet om evig lidelse etter døden kan ha en sterkt oppdragende effekt på de som tror (Hobbes 1946:86,92).

Naturlovene, som for Hobbes *er* moral, kunne tenkes å være en annen kilde til binding:

The laws of nature oblige *in foro interno*; that is to say, they bind to a desire they should take place: but *in foro externo*; that is, to the putting them in act, not always (Hobbes 1946:103).

THE OFFICE of the sovereign, be it a monarch or an assembly, consisteth in the end, for which he was trusted with the sovereign power, namely the procuration of *the safety of the people*; to which he is obliged by the law of nature, and to render an account thereof to God, the author of that law, and to none but him (Hobbes 1946:219).

---

<sup>4</sup>Jeg mener her i hovedsak tilegnelse av lojalitet og makt via å ha kontroll over Guds befalinger, som nødvendigvis vil stilles over sivile lover av troende. At f.eks. kristendommens Gud befaler å adlyde den verdslige suverenen endrer ikke dette.

At Hobbes' suveren er bundet av moral er forsåvidt korrekt, men ikke slik vi er vant til å se dette. Naturlovene er fornuftsregler som sikrer fred, som er ethvert menneskes – og dermed statens – mål. Suveren er bundet av naturlovenenes moral, men det er viktig å påpeke at dette ikke er en *ekstern* binding, og faktisk heller ikke en ekstern moral<sup>5</sup>:

The interpretation of the law of nature, is the sentence of the judge constituted by the sovereign authority [...] (Hobbes 1946:181).

Man ser her at Hobbes' suveren – likesom han har makt til å tolke religion – har makt til å tolke naturlovene. Riktignok presenteres det her som en indirekte kontroll via dommere, men med Hobbes' syn på maktdeling er det opplagt at den makten dommeren her benytter seg av, i realiteten er suverenens. Dommeren innsettes ved lov, og loven er som vi har sett intet annet enn suverenens vilje. På bakgrunn av dette hevder jeg at naturlovenene ikke utgjør en *ekstern* moral; det er ikke et sett med regler som suveren blir tvunget til å overholde. Han kan selv tolke dem, og suverenens hovedmål forblir *folkets sikkerhet*. Naturlovene er imidlertid forenlige med dette målet, så det skal ikke forårsake problemer. Det er en viktig distinksjon hvorvidt naturlovene er rene fornuftsregler eller regler nedsatt av Gud. I Hobbes' opprinnelige fremstilling av naturlovene, i kapittel 14 og 15 av *Leviathan*, antas disse å være rene fornuftsregler, og jeg vil her holde meg til denne fremstillingen.

Det fremgår altså tydelig at Hobbes' suveren ikke i noen tradisjonell forstand er *bundet* av legale, religiøse eller moralske hensyn, slik man ut fra Østeruds definisjon ville kunne kreve for å karakterisere Hobbes' stat som autoritær.

## Hobbes' stat

At Hobbes' stat ikke passer spesielt godt med Østeruds definisjon av det autoritære, gjør det naturlig å vurdere hvorvidt Hobbes' stat kan ansees som totalitær.

Jeg nevnte i innledningen til kapittelet at Hobbes muligens også kunne forståes som en "protoliberaler", noe som diskuteres videre i premiss 6 i 4.2.2, i 5.2.3 og i hele 5.3. Det blir imidlertid raskt tydelig at en del av hans hovedidéer om makt og statens fullmakter er uforenlige med dagens forståelse av "liberale" samfunn<sup>6</sup>. Det som *er* klart, er at

---

<sup>5</sup>Hvorvidt suveren er bundet diskuteres også i 5.3.1.

<sup>6</sup>Se f.eks. sitat fra Kuper og Kuper (1996:41) ovenfor.

Hobbes var en av de første til å fremme liberale idéer som likhet for loven, rettsstatt og spesielt pluralisme. Problemet er imidlertid at det “liberale” aspektet ved hans teori ikke har noe forsvar mot det autoritære. Jeg diskuterer disse idéene under premiss 6 i 4.2.2.

Jeg nevnte statens “mål” – *telos* – og vil kort se på hvorvidt Hobbes’ samfunn kan hevdes å være ideologisk nøytralt, her grovt forstått som “uten mål”. Det er selvsagt problematisk å hevde at det er mulig å være uten mål, eller at en viss stat er ideologisk “nøytral”, men jeg vil i denne sammenheng anse *minimale mål* (som å sikre fred), for å være forenlig med en ideologisk nøytral stilling. En stat uten det mål å eksistere ville være vanskelig å forstå<sup>7</sup>.

And lastly the motive, and end for which this renouncing, and transferring of right is introduced, is nothing else but the security of a man’s person, in his life, and in the means of so preserving life, as not to be weary of it (Hobbes 1946:87).

THE OFFICE of the sovereign, be it a monarch or an assembly, consisteth in the end, for which he was trusted with the sovereign power, namely the procuration of *the safety of the people* [...] (Hobbes 1946:219)

Det er altså samsvar mellom individets motiv for å overføre sine rettigheter til suverenen og suverenens/statens mål. *Folkets sikkerhet* fremmes som statens *årsak*, og også dens mål. Jeg velger å anse dette som et minimalt mål av typen jeg nevnte ovenfor, et mål som ikke krenker noen, da vi som vist i 2.3.4 tar for gitt at det verste ondet er døden, eller krig. Ro, i form av fred, er altså kuren for det verste ondet, og mens det for Hobbes er det viktigste målet en stat kan ha, kan det ikke på noen meningsfull måte hevdes at Hobbes’ stat er totalitær ved å fremme dette.

Hobbes tillater som vi har sett f.eks. sensur, men det er liten støtte for å hevde at dette gjøres for å fremme én spesiell doktrine over andre. Hovedmålet med sensur er å hindre uroligheter og opprør, altså å sikre ro. Om noe, kan det hevdes at Hobbes ser på ideologisk ensretting og en reformasjon av mennesket som umulig, da naturen, og for han dermed også menneskets natur, er uforanderlig (Green 1993:128). Dersom Hobbes’

---

<sup>7</sup>En stat som anser seg selv kun som et instrument for å reformere mennesker slik at de en dag kan leve sammen i fred *uten* staten er et eksempel på en slik stat som her nevnt. Også denne må dog søke ro og stabilitet inntil målet er oppnådd, og det blir tydelig at denne staten ville hatt et svært omfattende mål i seg selv – nemlig å reformere mennesket.

resonnementet leder ham til å foreslå en “feil” type stat, som faktisk ville måtte forandre mennesker for at den skulle fungere, mener jeg dette er på grunn av feil i hans logiske resonnement, ikke ideologi, og det er nettopp denne argumentasjonen som vurderes i dette og det neste kapittelet.

Det er videre naturlig å konkludere med at Hobbes’ stat *ikke* er totalitær; den er autoritær, i alle fall dersom man anvender definisjonen til Kuper og Kuper (1996). Det faktum at Østeruds definisjoner ikke tillater en likefrem plassering av Hobbes’ stat er en svakhet som gjør det naturlig å ikke anvende disse definisjonene i denne sammenheng.

## 4.2 Hobbes’ resonnement

For å kunne vurdere “Hobbes’ resonnement”<sup>8</sup> på en god måte, ønsker jeg å utlede en mer stringent presentasjon av hans teori. Dette vil foregå ved at jeg fremstiller premisser og utleder en konklusjon med bakgrunn i Hobbes’ tekster, og hovedsakelig *Leviathan*. I dette kapittelet vil premissene *antas sanne*, og formålet er dermed å finne et sett med premisser som er *tilstrekkelige* og *nødvendige* for å nå den konklusjon at den Hobbesianske stat er den beste løsningen<sup>9</sup>. Ved at premissene må være tilstrekkelige kreves det at resonnementet må være fullstendig – at konklusjonen følger som en nødvendighet av premissene. At premissene må være nødvendige innebærer at premissene er en nødvendig del av resonnementet – at konklusjonen ikke blir den samme dersom enkelte fjernes.

Det er en del problemer forbundet med å gjøre et resonnement *tilstrekkelig*, da det i ethvert resonnement kan befinne seg *implisitte forutsetninger*:

Av og til kan disse forutsetningene være trivielle, så opplagte at det ikke er uenighet om dem. [...] Men av og til kan disse forutsetningene være meget problematiske. De deles ikke av vår tilhører, og det er nødvendig å begrunne dem nærmere. Av og til finner vi kanskje at heller ikke vi kan akseptere disse forutsetningene etter at vi først er blitt oppmerksomme på

---

<sup>8</sup>I det følgende refererer jeg ved “Hobbes’ resonnement” til min utlegning av det resonnementet som leder til konklusjonen om at Hobbes’ stat er nødvendig.

<sup>9</sup>Problematisering av resonnementets konklusjon, hvorvidt Hobbes’ stat er ønskelig, nødvendig, uunngåelig osv., blir behandlet nedenfor.

dem (Føllesdal et al. 1996:235).

Det er lett å bli forledet av en god forfatter med et godt grep på narrativer og retorikk; man kan få inntrykk av seriøsitet, man blir presentert med premisser man godtar og argumentasjonen virker upåklagelig. Allikevel kan det gjøres “sprang” i argumentasjonen der *implisitte forutsetninger* anvendes for å komme til ønsket konklusjon, uten at tilhører er oppmerksom på hva som har skjedd. Han vil da kunne høre konklusjonen, stusse litt siden han muligens forventet et annet utfall, men allikevel godta den, siden han ikke kan sette fingeren på hva som er feil. Hvorvidt Hobbes’ resonnement lider av en lignende svakhet gjenstår å se, og jeg mener min fremgangsmåte er godt egnet til eventuelt å avdekke dette. Min oppstilling av resonnementet innebærer som nevnt også en viss fare: min subjektive vektlegging av Hobbes’ poenger, samt mine antagelser om hvilke implisitte forutsetninger som eventuelt tas inn, farger nødvendigvis denne analysen. Jeg forsøker dog så langt som mulig å belegge mine valg av premisser, og hele resonnementet vil vurderes videre i neste kapittel.

Det vil nødvendigvis være umulig å innlemme all innsikt, og samtlige implisitte premisser, fra resonnementet som fremsettes i tekstuell form av Hobbes i *Leviathan*, men jeg anser allikevel forsøket på dette som fruktbart. Jeg foretrekker en fremgangsmåte som her stipulert, da det vil lette arbeidet med å vurdere resonnementet kritisk i kapittel 5. Jeg vil kunne vurdere hvert enkelt premiss for å avgjøre dets sannhetsverdi, og jeg vil, dersom nødvendig, kunne erstatte mine premisser utledet fra Hobbes med andre premisser, for så å vurdere hvilke implikasjoner disse alternative premissene kunne tenkes å innebære.

### 4.2.1 Forberedende betraktninger om naturtilstanden

Før jeg presenterer min versjon av Hobbes’ resonnement, vil jeg kort se på hvordan Hobbes selv forklarer at naturtilstanden blir konfliktfylt. Jeg har valgt å ikke analysere selve naturtilstanden i detalj her, og tar i resonnementet utgangspunkt i Hobbes’ forståelse av naturtilstanden som svært uønsket, samt uunngåelig, med mindre man oppretter en form for stat. Jeg ser imidlertid nærmere på hvordan naturtilstanden faktisk kan tenkes å arte seg i 5.2.3, da jeg mener Hobbes ofte blir misforstått når det gjelder hva han anser som naturtilstanden.

Kapittel 13 i *Leviathan*<sup>10</sup> er det sentrale kapittelet i denne forbindelse, og jeg vil her kort gjengi Hobbes' hovedpoenger.

Mennesker er fundamentalt sett like, i kraft av at selv den svakeste kan drepe den sterkeste (med list eller i samarbeid med andre). Denne likheten fører til like forventninger om å oppnå sine mål, noe som blir svært problematisk, siden godeknapphet gjør at ikke alle *kan* oppnå sine mål samtidig. Disse forholdene skaper *konkurransen*, og mennesker vil søke å ødelegge og underlegge seg hverandre.

Dersom man skaper eller eier noe, må man forvente at andre (alene eller i fellesskap) vil komme for å ta dette. Siden mennesker forutser denne utviklingen, og dermed blir redde for, og *usikre* på, hverandre, vil de søke å komme hverandre i forkjøpet – de vil tilegne seg herredømme over så mange de klarer, ved makt og list.

Det enkelte mennesket søker i tillegg å bli verdsatt av andre slik han verdsetter seg selv – han søker *ære* – og menneskets naturlige situasjon blir med dette ikke preget av lykke, men ulykke.

Det er dermed tre årsaker til at naturtilstanden betegnes av konflikt: konkurranse, usikkerhet og ære<sup>11</sup>.

Hereby it is manifest, that during the time men live without a common power to keep them all in awe, they are in that condition which is called war; and such a war, as is of every man, against every man. For WAR, consisteth not in battle only, or the act of fighting; but in a tract of time, wherein the will to contend by battle is sufficiently known: and therefore the notion of *time*, is to be considered in the nature of war; as it is in the nature of weather. For as the nature of foul weather, lieth not in a shower or two of rain; but in an inclination thereto of many days together: so the nature of war, consisteth not in actual fighting; but in the known disposition

---

<sup>10</sup>Kap 13: "OF THE NATURAL CONDITION OF MANKIND AS CONCERNING THEIR FELICITY, AND MISERY".

<sup>11</sup>I *De Cive* (Hobbes 2005a:1) kan man lese oversikten over kapittelet om naturtilstanden som en forklaring av naturtilstanden i stikkordsform: "[...] 2. That the beginning of civil society is from mutual fear. 3. That men by nature are all equal. 4. Whence the will of mischieving each other ariseth. 5. The discord arising from comparison of wits. 6. From the appetite many have to the same thing. 7. The definition of *right*. 8. A right to the end, gives a right to the means necessary to that end. 9. By the right of nature, every man is judge of the means which tend to his own preservation. 10. By nature all men have equal right to all things. 11. This right which all men have to all things, is unprofitable. 12. The state of men without civil society, is a mere state of war [...]. 13. War is an adversary to man's preservation. [...]"

thereto, during all the time there is no assurance to the contrary. All other time is PEACE (Hobbes 1946:82)

## 4.2.2 De grunnleggende premissene

### Behovet for en stat

**Premiss 1** *Fred* uten en overodnet samlende makt er ikke mulig

For if we could suppose a great multitude of men to consent in the observation of justice, and other laws of nature, without a common power to keep them all in awe; we might as well suppose all mankind to do the same; and then there neither would be, nor need to be any civil government, or commonwealth at all; because there would be peace without subjection (Hobbes 1946:110).

At staten er nødvendig for å oppnå fred er viktig å fastslå, da Hobbes' konklusjon er avhengig av et faktisk *behov* for en statsmakt for å oppnå menneskers felles mål – fred. I dette kapittelet vil premisset godtas, men det er et problematisk premiss, som vil bli vurdert videre i 5.2, der i forbindelse argumenter for anarki og kommunitarianske samfunn som alternativer til staten; det dreier seg bl.a. om teorier som hevder at “Fangenes dilemma”-situasjoner ved gjentatte spill *kan* få et pareto-optimalt utfall – selv uten mulighet for kommunikasjon og bindinger.

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^1$	3, 3	1, 4
	$\tau_i^2$	4, 1	2, 2

Fangenes dilemma

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^1$	$R, R$	$S, T$
	$\tau_i^2$	$T, S$	$P, P$

Fangenes dilemma med notasjon

Figur 4.1: Her vises et Fangenes dilemma-spill med ordinale preferanser, samt betegnelse på gevinstene, der  $T$ : Temptation to defect,  $P$ : Punishment for war,  $R$  Reward for cooperation og  $S$ : Suckers payoff.

Eksempelet i figur 4.1 og tilhørende diskusjon er hentet fra Hovi (2008)<sup>12</sup>. I tillegg til symbolene fra figuren, introduseres  $\delta$  ( $w$  hos Hovi) som *diskonteringsfaktor*<sup>13</sup>. Dette er et gjentatt Fangenes dilemma-spill, og diskontering av fremtidig forventet gevinst er dermed viktig. Videre antas her strategi  $\tau^1$  å være “tie/samarbeide”, mens  $\tau^2$  er “tyste/konflikt”. De ordinale preferansene fra figur 4.1 brukes her for eksempelets skyld som kardinale verdier man kan anvende matematisk, noe som kommenteres ytterligere nedenfor. Spillerne antas her å følge en “grim trigger”-strategi<sup>14</sup>. Hovi fremsetter følgende krav til rangeringen av gevinster, der den første begrensningen kreves for at spillet skal være et Fangenes dilemma-spill, noe som er tilfredsstilt i det venstre spillet i figur 4.1:

$$T > R > P > S, \quad 2R > T + S$$

$$U_i(\tau_i^1, \tau_j^1) = R + \delta R + \delta^2 R + \delta^3 R + \dots = \frac{R}{1 - \delta} \quad (4.1)$$

$$U_i(\tau_i^2, \tau_j^1) = T + \frac{\delta P}{1 - \delta} \quad (4.2)$$

Ligning (4.1) viser her forventet nytte for en aktør ved å samarbeide med en uendelig tidshorisont. Ligning (4.2) viser nytten for den som først bryter samarbeidet. Dette leder til følgende ulikhet som krav for at samarbeid skal kunne forventes:

$$\frac{R}{1 - \delta} \geq T + \frac{\delta P}{1 - \delta} \quad (4.3)$$

Ulikheten i (4.3) må tilfredsstilles for at en avtale skal være *selvhåndhevende*; forventet nytte ved å samarbeide må overstige (her eventuell være lik) nytten ved å bryte avtalen. Arrangert på en måte som gir oss kravet til spillernes diskonteringsfaktor:

$$\delta \geq \frac{T - R}{T - P} \quad (4.4)$$

Dersom vi forlater Hovi og går tilbake til Hobbes, er det interessant å se på hvilke

<sup>12</sup>Taylor (1987) bruker et lignende rammeverk, og jeg ser nærmere på dette i kapittel 5.

<sup>13</sup>Diskonteringsfaktoren utledes av en spillers forventning til at spillet skal fortsette ( $p$ ) og spillerens diskonteringsrate ( $r$ ):  $0 \leq \delta = \frac{p}{1+r} \leq 1$  (Hovi 2008:79).

<sup>14</sup>Strategien innebærer at partene begynner med samarbeid, for å så bytte til konflikt ved samtlige etterfølgende runder dersom den andre parten først bryter. Andre strategier, som i visse tilfeller må ansees å være mer realistiske, vil bli diskutert i 5.2.



implikasjoner disse funnene har. “Ulempen” ved  $\delta$  i denne forbindelse er at den består av to elementer: a) forventning om hvorvidt spillet vil fortsette flere runder og b) diskontering av fremtidig nytte. Begge elementene er svært viktige, og det kan derfor i visse tilfeller være noe uheldig å bruke ett symbol for å representere begge. Jeg vil her fokusere mest på sannsynligheten for at spillet fortsetter, da det er dette som synes mest relevant i forhold til Hobbes’ resonnement. Jeg har i 2.5 hevdet at Hobbes’ individer er langsiktige i sine betraktninger, så jeg ser ikke at ekstrem diskontering av fremtidig nytte skulle være årsaken til menneskers “problem” i naturtilstanden.

Det er visse problemer forbundet med dette eksempelet i forbindelse med borgere i konflikt i naturtilstanden. Jeg vil som nevnt utdype denne diskusjonen i neste kapittel, men vil påpeke et par hovedmomenter her. Ordinale verdier som gitt i figur 4.1 kan virke noe villedende, da de kan representere *meget* forskjellige kardinale verdier, noe som vil gi store utslag i kravet til  $\delta$  for å oppnå samarbeid. Videre er jeg usikker på om et spill med en slik preferansestruktur er best egnet for å illustrere Hobbes’ resonnement om naturtilstanden. Krig, og da spesielt døden, er det verste ondet for Hobbes, så det er ikke opplagt at gevinsten ved konflikt er større enn gevinsten ved å “tape”. Det er videre avgjørende for aktørene hvorvidt brudd og konflikt faktisk kan føre til en høy sannsynlighet for døden og brudd av spillet. I så tilfelle fremstår samarbeid på ubestemt tid svært attraktivt. En alternativ fremgangsmåte utforskes i premiss 2 og i 5.2.3.

Interessant her er det også å se på Hobbes’ gjennomgang av “sosiale dyr”, som faktisk demonstrerer denne spontane orden uten en eksplisitt formell statsmakt. Hobbes utdypet her de forskjellene mellom mennesker og dyr som gjør at mennesker *ikke* er i stand til å oppnå det samme.

It is true, that certain living creatures, as bees, and ants, live sociably one with another, which are therefore by Aristotle numbered amongst political creatures; and yet have no other direction, than their particular judgements and appetites; nor speech, whereby one of them can signify to another, what he thinks expedient for the common benefit: and therefore some man may perhaps desire to know, why mankind cannot do the same (Hobbes 1946:111).

Hobbes’ første punkt går på at mennesker konstant kjemper om “ære og verdighet”, en situasjon som ikke finnes hos dyrene det er snakk om. Denne kampen fører til

“misunnelse og hat, og til slutt krig” (Hobbes 1946:111).

Videre er det en sentral forskjell i forholdet mellom det individuelle og det kollektive gode. Det interessante her er at Hobbes ikke hevder at sosiale dyr forfølger det kollektive fremfor det individuelle godet, men at disse sammenfaller; siden de har en “naturlig tilbøyelighet til å fokusere på det private godet, sikrer de dermed felles goder” (Hobbes 1946:111). Mennesker, som vi har sett, forfølger også sitt eget gode; men mennesker, hvis “glede består i å sammenligne seg med andre mennesker, kan ikke sette pris på noe annet enn det som er eminent”, noe som indikerer at mennesker har langt større problemer med å tilegne seg felles goder naturlig (Hobbes 1946:111).

Den tredje forskjellen han påpeker er at dyrene mangler *fornuft*<sup>15</sup>. Dette fører til at dyrene ikke er kritiske i sin vurdering av for eksempel sine overordnede; “[de] ser ikke, og tror heller ei at de ser, noen feil i administrasjonen av sine felles gjøremål” (Hobbes 1946:111). Dette stemmer godt overens med Hobbes’ ønske om borgere som “blanke ark”, siden menneskers problemer skapes av deres ambisjoner og (ofte ubegrunnede) tro på at de selv er bedre egnet til å styre<sup>16</sup>. Dette fører ifølge Hobbes til problemer for samfunn, dersom det er svært mange som søker å endre tingenes tilstand på forskjellige måter samtidig (Hobbes 1946:111).

Et fjerde moment er *språk*; mens det unektelig er en fordel for mennesker å kunne kommunisere via språk, peker Hobbes på noen fordeler disse dyrene har ved *ikke* å ha denne muligheten: de kan ikke fremstille som sant det som er usant, og heller ikke som usant det som er sant (Hobbes 1946:111). De kan heller ikke feilaktig forsterke eller nedvurdere f.eks. betydningen av andres prestasjoner, noe som er en fordel da det ville kunne hindre mennesker i å nyte sine bragder til fulle. Andres anerkjennelse er som vi tidligere har sett viktig for mennesker, og dersom man ikke forventer å få anerkjennelse for sine prestasjoner, forsvinner noe av incentivet til å prestere.

En femte distinksjon er dyrenes manglende forutsetninger for å skille mellom urett og tilfeldig skade (Hobbes 1946:111)<sup>17</sup>.

Det siste, og med spillteoretiske øyne kanskje det mest interessante, punktet, er at dyrenes enighet er *naturlig*. “[M]enneskers [enighet] er kun ved kontrakt og dermed

---

<sup>15</sup>Begrepet ble diskutert i 2.5.

<sup>16</sup>“[...] the common people’s minds, unless they be tainted with dependence on the potent, or scribbled over with the opinions of their doctors, are like clean paper, fit to receive whatsoever by public authority shall be imprinted in them” (Hobbes 1946:221).

<sup>17</sup>Dette poenget er interessant i forhold til behandlingen av *intensjonalitet*, som ble berørt i 3.4.3.

kunstig, og det er ingen overraskelse dersom det kreves mer enn kontrakten for å gjøre enigheten konstant og varig” (Hobbes 1946:111).

Grunnen til at dette siste punktet er interessant for spillteoretikere er at mange faktisk hevder at det *er* mulig med en “naturlig” enighet/fred også mellom mennesker, som vist ovenfor, noe som vil behandles videre i neste kapittel.

Mens felles goder (for Hobbes noe som er godt for flere) og *fellesgoder* i økonomisk terminologi ikke er det samme, kan det allikevel i denne sammenheng være nyttig å se på fellesgoder og en del av litteraturen rundt dette. Pindyck og Rubinfeld (2005) definerer fellesgoder som ikke-ekskluderbare og ikke-rivaliserende; mennesker kan ikke ekskluderes fra å nyte godet, og marginalkostnaden ved å produsere nok av godet til én ekstra person er 0. Problemet i forbindelse med produksjonen av slike goder er at det “individuell rasjonelle” er å ikke bidra, siden den enkelte uansett vil kunne nyte godet – han kan velge å være *gratispassasjer*. Det oppstår dermed markedssvikt, ved at individer ikke vil produsere nok av slike goder uten koordinering. Ifølge Hobbes er altså slike problemer mindre fremtredende hos sosiale dyr, ved at det å forfølge det private godet ikke fører til slike problemer for dem – de spiller med andre ord ikke et Fangenes dilemma-spill. Det må imidlertid bemerkes at slik “individuell rasjonalitet”, ikke vil anses å være rasjonelt for Hobbes, da relasjoner til andre mennesker (med mindre man er en eremitt) er uunngåelig. Som vist i kapittel 2 er naturlovene fornuftslover som dikterer fornuftig adferd, og kortsiktig fokus kun på egen gevinst er ikke rasjonelt, med tanke på konsekvensene (Hobbes 1946:96,104, se 2.4.4)<sup>18</sup>.

Dette er imidlertid ikke et nytt fenomen, selv om *fellesgoder* på mange måter har blitt et moderne faguttrykk, spesielt innen Rational choice og økonomi. Thukydid skrev følgende lenge før mikroøkonomi og dagens politiske teori basert på økonomisk innsikt dukket opp:

[...] they devote only a fraction of their time to their general interests, spending most of it on arranging their own separate affairs. It never occurs to any of them that the apathy of one will damage the interests of all. Instead each state thinks that the responsibility for its future belongs to someone else, and so, while everyone has the same idea privately, no one notices that from a general point of view things are going downhill (Thukydid

---

<sup>18</sup>Hobbes’ resonnement i denne sammenheng refereres ofte til som hans svar til “tåpen”, da han i kapittel 15 av *Leviathan* klargjør sitt poeng ved å vise at “tåpen”, som mener det er best å bryte avtaler dersom man kan tjene noe på det, tar feil.

1972:120).

Hobbes peker også på et lignende fenomen:

For all men are by nature provided of notable multiplying glasses, that is their passions and self-love, through which, every little payment appeareth a great grievance; but are destitute of those prospective glasses, namely moral and civil science, to see afar off the miseries that hang over them, and cannot without such payments be avoided (Hobbes 1946:120).

Mennesker *ønsker* også staten, da den er nødvendig for å oppnå et “godt liv”<sup>19</sup>. Utover å være forferdelig ved at livet der er kort og brutalt, er det i naturtilstanden heller ikke mulig å oppnå visse goder mennesker ifølge Hobbes setter stor pris på. Det finnes ingen kunst, og heller ikke vitenskap, i naturtilstanden, siden dette krever en viss frihet til å bruke sine evner på andre gjøremål enn å forsvare seg (Hobbes 1946:64,161). Manglene i naturtilstanden diskuteres imidlertid nærmere i 5.2.3. Hovedpoenget i denne forbindelse er at mennesker ifølge Hobbes ikke kan oppnå *fred* uten staten.

### **Implikasjonene av synet på døden som det verste ondet**

**Premiss 2** Død – krig, og spesielt borgerkrig – er det verste ondet for mennesker

Et viktig element i forbindelse med vektleggingen av *stabilitet* i valgt statsform er hvor sterkt individer ønsker seg bort fra krigstilstanden. Jeg vil i denne forbindelse hevde at naturtilstanden, dersom den skal fremstilles som et enkelt tostrategi-spill, ikke nødvendigvis bør ansees som et Fangenes dilemma-spill. Jeg er ambivalent i forhold til hvor nyttig en såpass enkel fremstilling av en så komplisert tilstand er, men ser nytten i at *enkelte aspekter* ved situasjonen illustreres på en lettfattelig måte.

Krig er ifølge Hobbes nært knyttet til det verste ondet, og jeg mener dette må reflekteres i et naturtilstand-spill; det verste ondet er ikke nødvendigvis å bli underordnet en annen, slik det fremstilles i Fangenes dilemma. Ved et slikt utfall er det nemlig stor sannsynlighet for at man vil beholde livet, og dette blir i så tilfelle et langt bedre utfall enn å dø (eller bli utsatt for stor fare for å dø i kamp) i åpen konflikt. Leo Strauss ser ut til å ha samme oppfatning:

---

<sup>19</sup>Et poeng jeg også fant hos Maslow, se 3.3.1.

Fear for his life, which came upon the man in his struggle for triumph, moderates, even kills, the will to triumph, and makes him ready to submit, to leave triumph to the enemy, in order that he may save his own life (Strauss 1952:21).

		$k_j$				$k_j$					
		$\tau_j^1$	$\tau_j^2$			$\tau_j^1$	$\tau_j^2$				
$k_i$	$\tau_i^1$	3, 3	1, 4	$\tau_i^1$	4, 4	1, 3	$\tau_i^1$	3, 3	2, 4		
	$\tau_i^2$	4, 1	2, 2	$\tau_i^2$	3, 1	2, 2	$\tau_i^2$	4, 2	1, 1		
			Fangenes dilemma				Moderat selvhevdelse				Chicken

Figur 4.2: Preferansestruktur i Fangenes dilemma-, Moderat selvhevdelse- og Chicken-spill.

*Moderat selvhevdelse* er et mulig alternativ til Fangenes dilemma. Det synes dog problematisk å forene Hobbes' tenkning med spillets verdsetting av felles moderasjon (i naturtilstanden) som bedre enn herredømme. Dersom  $(\tau^1, \tau^1)$  forstås som et strategipar som innebærer etableringen av *staten*, vil jeg imidlertid anse Moderat selvhevdelse som et bedre alternativ enn Fangenes dilemma (og kanskje Utopia som bedre enn begge), men jeg ønsker her å modellere *naturtilstanden* og menneskers mulige strategier der – ikke hele Hobbes' filosofi. Måten Hobbes' filosofi formaliseres på behandles nærmere i forbindelse med Taylors behandling av Hobbes i 5.2.3.

En annet mulighet er å modellere naturtilstanden som et *Chicken*-spill, som vist i figur 4.2. Det er visse mangler også ved å fremstille naturtilstanden som et Chicken-spill, men det kan argumenteres for at en slik fremstilling er mer tro mot Hobbes' filosofi. For bedre å representere Hobbes' syn på naturtilstanden, mener jeg dette spillet også må inneholde usikre gevinster ved "konflikt"; mens man i et vanlig Chicken-spill antar det verste utfall for *begge* aktørene ved konflikt, synes det å være én "vinner" og én taper ved konflikt i et Hobbesiansk naturtilstand-scenario<sup>20</sup>. Jeg foreslår følgende gevinst for konflikt som mer i tråd med Hobbes' filosofi enn en fast og forutsigbar gevinst:

<sup>20</sup>Chicken er en "bilsport" der to aktører krasjer sine biler og der begge med stor sannsynlighet vil omkomme dersom begge velger "hard" strategi.

$$U_i(\tau_i^2, \tau_j^2) = K = p_i^{k_1} \bar{D} + p_i^{k_2} T + p_i^{k_3} H, \quad p_i^{k_1} + p_i^{k_2} + p_i^{k_3} = 1 \quad (4.5)$$

I ligning (4.5), som gir forventet gevinst ved *konflikt* ( $K$ ), representerer  $\bar{D}$  døden, og  $p_i^{k_1}$  gir  $i$ 's sannsynlighet for å få denne “gevinsten” ved krig/konflikt. Det er imidlertid også  $p_i^{k_2}$  sjanse for å tape men overleve (få gevinst  $T$  – trelldom) og  $p_i^{k_3}$  sjanse for å “vinne” (få gevinst  $H$  – herredømme). Vi har sett at det i naturtilstanden før kamp er usikkert hvem som er sterkest, så vi må uten spesielle omstendigheter anta at det mellom fremmede aktører, uten kjent rykte, vil antas 50% sjanse for seier ( $p_i^{k_3}$ ). Makt, avdekking av makt, verdi og verdsetting av selv og andre ble behandlet i 2.4.6.

Det at man ikke kan bedømme personers makt uten kamp betyr imidlertid ikke at individer selv alltid vil anta at mulighetene for å vinne ( $p^{k_3}$ ) og tape ( $p^{k_1} + p^{k_2}$ ) er like for seg selv og motstanderen. Hobbes er svært eksplisitt på at *mennesker generelt sett verdsetter seg selv for høyt*. Etersom *verdi* for Hobbes er definert som et mål på makt, som igjen må antas å påvirke sannsynlighetene for å seire ved konflikt, må vi slutte at mennesker oppfatter seg selv som mektigere enn de egentlig er. Dette medfører nødvendigvis at aktørenes subjektive nytteforventning for  $K$  blir uforholdsmessig stor, noe som kan tenkes å medføre mer konflikt enn hva man ville fått dersom aktørene var helt realistiske i sine vurderinger<sup>21</sup>.

Subjektive vurderinger av egen og andres makt er viktige i forbindelse med *rykte om makt*, som ble diskutert i 2.4.6 og 3.4.5. Når en person evaluerer sannsynligheten for å tape en konflikt, vil denne øke dersom han har grunn til å tro at motstanderen er langt sterkere enn han selv. Slik vil det, som jeg hevdet i kapittel 2 og 3, også kunne være svært relevant for en aktør å ha ord på seg for å være mektig. Andre vil i så tilfelle være mindre villige til å gå til kamp mot ham, da deres forventningsverdier for  $(\tau^2, \tau^2)$  blir lavere enn ved kamp mellom “likemenn”.

$$U_i(\tau_i^1, \tau_j^2) = T \quad (4.6)$$

$$U_i(\tau_i^2, \tau_j^1) = H \quad (4.7)$$

I ligning (4.6) og (4.7) ser vi at aktør  $i$  får gevinst “trelldom” (under  $j$ ) dersom han

---

<sup>21</sup>Deprimerte individer ville dermed muligens ha bedre muligheter for å leve fredelig, noe jeg så på i forbindelse med “depressiv realisme” i 3.4.3.

velger en myk strategi der motparten velger en hard strategi. Dersom rollene byttes får aktør  $i$  “herredømme” over  $j$ . Det er videre hevet over enhver tvil at herredømme er å foretrekke fremfor trelldom for Hobbesianske individer.

Det gjenstår så å bestemme utfallet for  $(\tau^1, \tau^1)$ . Det å velge en myk strategi kan som en første tilnærming forståes som å være “moderat” slik Hobbes bruker begrepet<sup>22</sup>. I denne sammenheng er det tale om strategisk moderasjon for å unngå visse ufordelaktige utfall, ikke *tilfredshet*. Forskjellen på disse begrepene ble vurdert i 2.5.

Also because there be some, that taking pleasure in contemplating their own power in the acts of conquest, which they pursue farther than their security requires; if others, that otherwise would be glad to be at ease within modest bounds, should not by invasion increase their power, they would not be able, long time, by standing only on their defence, to subsist. And by consequence, such augmentation of dominion over men being necessary to a man’s conservation, it ought to be allowed him (Hobbes 1946:81).

Ett alternativ er å forstå  $(\tau^1, \tau^1)$  som *modus vivendi*. Eventuelle konflikter *løses* ikke, men de eskalerer heller ikke til fysisk kamp. Slik å forstå vil utfallet være relativt godt for aktørene, siden dette kan forståes som en form for “fred”. En slik forståelse av fred er dog *ikke* forenlig med Hobbes’ begrepsbruk, da man per definisjon er i naturtilstanden (og dermed også i *krig*) dersom man ikke har opprettet en stat, noe jeg ser nærmere på i 5.2.3. Jeg anser derfor *modus vivendi* som en bedre beskrivelse av gjensidig moderasjon enn “fred”, og vil anvende dette som utfallet ved  $(\tau^1, \tau^1)$ .

Jeg ser imidlertid et forbedringspotensiale i en såpass begrenset fremstilling av Hobbes’ filosofi som jeg her fremsetter.  $(\tau^1, \tau^1)$  kan i min modell ansees som et relativt stabilt og greit utfall, dersom man har mulighet til å forplikte seg og spillet gjentas<sup>23</sup>. Dette gir åpenbart et noe misvisende bilde av Hobbes’ naturtilstand, da han er svært eksplisitt på at moderasjon i naturtilstanden *ikke* er nok til å sikre et trygt og godt liv, slik sitatet ovenfor viser. Ustabil våpenhvile er dermed et viktig utfall, og en slik fremstilling fremstår som plausibel. Denne problemstillingen vil vurderes videre i 5.2.3.

Med dette grunnlaget mener jeg at det er mulig å vurdere det jeg har kalt “Hobbes’ spill”, som vises i figur 4.3. Utfall  $(\tau^1, \tau^1)$  fører altså til en uendret situasjon for spillerne

---

<sup>22</sup>Taylor (1987) har et annet syn på hva det innebærer “å samarbeide”, og jeg ser nærmere på de forskjellige synene på dette i 5.2.2.

<sup>23</sup>Mer om mulighetene for forpliktelse i naturtilstanden i 5.2.3.

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
	$\tau_i^1$	$MV$	$T, H$
$k_i$	$\tau_i^2$	$H, T$	$K, K$
		Hobbes' spill	

Figur 4.3:  $MV$ : *Modus vivendi*,  $T$ : Trelldom,  $H$ : Herredømme,  $K$ : Konflikt (Ligning (4.5)).

– ingen gevinst og intet tap. Forutsetningen for at dette skal være et Chicken-spill, er at  $K < T < MV < H$ . Konflikt er verre enn trelldom, som er verre enn *modus vivendi*, som igjen er verre enn herredømme.

At død, og fare for død, for Hobbes er det verste ondet, ivaretas bedre i et Chicken-spill enn i Fangenes dilemma. Dette også fordi Hobbes er meget klar på at *krig* – og spesielt borgerkrig – er meget nært knyttet til fare for liv og lemmer. Videre er det naturlig å anta at åpen konflikt gir klart størst fare for død og skade. Verdien av  $K$  for den enkelte spiller (ligning (4.5)) kan dog tenkes å variere svært mye; konflikt kan være en kamp med sverd i en tid uten særlig kunnskap om medisin og legevitenenskap, eller det kan være en kamp på nevene i dagens Norge. Det skal lite til å forestille seg at dødeligheten er høyest i det førstnevnte tilfellet. Det må allikevel antas at kamp, slik jeg her vurderer det, må ansees som meget alvorlig, og individer som går til kamp må derfor forventes å bevæpne seg som best de kan. Jeg vil dog ikke forfølge en slik utvidelse av modellen, og vil kun anta at sannsynligheten for å dø eller skades hardt i kamp er stor nok til at utfall  $(\tau^2, \tau^2)$  av de fleste individer regnes som det verste ondet.

[...] opponents in the game will always try to exploit each other and they know it. Hobbesian man will not. He is not a utility maximizer, but a disaster-avoider. The proper response to my disarming myself is to disarm yourself, not to kill me: to seek peace, not to maximize advantage. I do not wish to be vulnerable not because I know you will exploit me if you can, but because I am not certain that you will not exploit me (Ryan 1996:224).

Dersom naturtilstanden er et Chicken-spill, er det gode grunner til å tro at man i det



store og hele faktisk klarer å realisere utfall  $(\tau^1, \tau^1)$ , noe jeg kommer nærmere tilbake til i 5.2.3. Dette kan synes å være i strid med standard spillteori, der Chicken (som engangsspill) har tre likevekter:  $(\tau^1, \tau^2)$ ,  $(\tau^2, \tau^1)$  og en blandet strategi der spillerne bestemmer strategi via lotteri. Det faktum at Hobbes' individer anser døden som det verste ondet – et onde som må unngås for enhver pris – gjør det imidlertid naturlig å behandle (majoriteten av) Hobbes' individer som “disaster-avoiders”, som Ryan (1996) hevder.

Hva innebærer så dette i et Chicken-spill? Siden det i den forenklede modellen av Hobbes' spill kun er ett utfall som har en påvirkning på sannsynligheten for å dø –  $(\tau^2, \tau^2)$  – er dette et utfall som må ansees som “katastrofe”. Ved å forfølge en maximin-strategi<sup>24</sup> vil spillerne med sikkerhet kunne unngå dette utfallet, da  $(\tau^2, \tau^2)$  ikke kan realiseres dersom (minst) en spiller velger strategi  $\tau^1$ . Kavka (1986:142) diskuterer hvorvidt Hobbes' individer er “disaster-avoiders”, noe som innebærer at man ved kjente ordinale sannsynligheter anbefales “å velge den handlingen som maksimerer mulighetene for å unngå alle katastrofale utfall”. Han påpeker videre at mange av Hobbes' resonnementer er noe uklare, og kan tolkes som å være basert på nyttemaksimering, maximin-strategier eller “disaster-avoidance”, og at det er vanskelig å si hvilken av disse man bør tillegge Hobbes (Kavka 1986:142-3).

I gjentatte spill er det forøvrig også innen standard spillteori enklere å se for seg at  $(\tau^1, \tau^1)$  kan realiseres, noe jeg kommer nærmere inn på kapittel 5.

## Maktens delbarhet

**Premiss 3** Makt *kan* ikke deles, kun midlertidig delegeres

Mens vi i dagens vestlige samfunn er vant til *maktdeling*, blir dette av Hobbes forkastet som fullstendig uholdbart, dersom mennesker skal kunne leve i fred:

For these [suverenens rettigheter] are incommunicable and inseparable (Hobbes 1946:118).

And this division is it, whereof it is said, *a kingdom divided in itself cannot stand*: for unless this division precede, division into opposite armies can never happen (Hobbes 1946:119).

---

<sup>24</sup>En beslutningsstrategi som innebærer å maksimere minimal gevinst.

For that [å dele en stats makt] were to erect two actors, that by opposing one another, must needs divide that power, which, if men will live in peace, is indivisible; and thereby reduce the multitude into the condition of war, contrary to the end for which all sovereignty is instituted. (Hobbes 1946:122).

I Hobbes' diskusjon om doktriner som fører til svekkelse eller sammenbrudd av en stat, er doktrinen om maktdeling et sentralt tema; denne doktrinen er "åpenbart og direkte i strid med essensen i et samfunn"; "[f]or hva annet er det å dele makten til et samfunn, enn å oppløse det; for delene av denne makten ødelegger hverandre" (Hobbes 1946:213). Hobbes snakker her *ikke* om at makten må ligge hos ett *individ*, men at den må samles i ett *organ*, som her kan være en kunstig person, som f.eks. der suverenen er en forsamling.

Det finnes åpenbart argumenter som taler for Hobbes' syn her, som at staten vil være mindre utsatt for forstyrrende intern maktkamp og at staten vil være handlekraftig. Det er imidlertid ikke tvil om at ulempene ved en slik maktkonsentrasjon også er store. Vi kjenner bl.a. Lord Actons berømte diktum om maktens korrupperende effekt. Også Locke har sterke motforestillinger mot den type makt Hobbes gir suverenen:

[F]or, if it be asked, what security, what fence is there, in such a state, against the violence and oppression of this absolute ruler? (Locke 2005:99)

Hobbes er, som vi vil se under premiss 4 og i 4.3, klar over disse ulempene, men anser dem som uunngåelige dersom man virkelig ønsker fred. Det er også et spørsmål hvorvidt makt som *oppfattes* som delt, faktisk *er* delt. For Hobbes vil en suveren kunne gi et organ lovgivende makt, men dette betyr ikke at makten deles. Selvbinding er, som vist ovenfor, ifølge Hobbes ikke en reell binding, og suverenen kan dermed når som helst tilbakekalle denne makten, noe som indikerer at den alltid har ligget hos suverenen.

For elective kings, are not sovereigns, but ministers of the sovereign; nor limited kings, sovereign, but ministers of them that have the sovereign power: nor are those provinces which are in subjection to a democracy, or aristocracy of another commonwealth, democratically or aristocratically governed, but monarchially (Hobbes 1946:125).

Jeg vil se mer kritisk på dette premisset i 5.3.3.

## Demokrati og enevelde

**Premiss 4** Et demokrati er mer utsatt for uroligheter enn et enevelde

And a representative of even number, especially when the number is not great, whereby the contradictory voices are oftentimes equal, is therefore oftentimes mute, and incapable of action (Hobbes 1946:108).

I dagens politiske klima er det forholdsvis vanskelig, og ikke minst uvant, å forsvare enevelde som styreform. Det er imidlertid ikke så lenge siden Europa ble styrt av mektige konger, og selv Norge har i dag en konge som formelt sett har mye makt. Hobbes er meget klar på at enevelde var hans foretrukne styreform, men som jeg også har vært inne på, er det nyttig å ha i bakhodet at Hobbes' syn begrunnes med *hva som er mest effektivt*, og ikke iboende egenskaper ved det å styres av én person.

For Hobbes var det en overvekt av fordeler for enevelde, og han begrunner sitt standpunkt ved å sammenligne enevelde og styre ved forsamling på seks punkter i kapittel 19 av *Leviathan* (Hobbes 1946:122)<sup>25</sup>:

Hobbes' første punkt går på forholdet mellom det kollektive og det individuelle gode. Hobbes foregriper en sentral tese i Rational choice-litteraturen idet han påpeker at enhver person som tjener som suveren (alene eller i komité) ikke slutter å bære sin "naturlige person". En person har sin personlige "hatt", og han har kanskje en hatt som representerer hans jobbstilling. Mens man i mange sammenhenger antar at individer handler som om de tar av sin personlige hatt når de opptrer i jobbsammenheng, påpekes det innen Rational choice-litteraturen at individer har en tendes til å beholde den personlige hatten på, selv når de ikler seg sin jobbhatt. Vi har altså flere hatter, og slutter aldri å være oss selv. Det er denne tanken Hobbes fremmer her, når han påpeker at "selv om han [suverenen] er nøye med å frembringe det felles gode når han opptrer i sin politiske stilling, er han mer, eller i det minste ikke mindre, opptatt av å frembringe individuelle goder for seg selv, sin familie, sine slektninger og venner" (Hobbes 1946:122). Som vist i diskusjonen om de sosiale dyr er det for mennesker ikke alltid samsvar mellom hva som er godt for felleskapet, og hva som er godt for individet. Hobbes mener mennesker velger det individuelle gode der det oppstår behov for å velge, og det er her hans argument for enevelde fremkommer: hos en eneherker konvergerer

---

<sup>25</sup>Kapittel 19: "OF THE SEVERAL KINDS OF COMMONWEALTH BY INSTITUTION, AND OF SUCCESSION TO THE SOVEREIGN POWER".

det felles og det individuelle gode, og der disse ligger nærmest hverandre, fremmes det felles gode best. “Rikdommen, makten og æren til en monark oppstår kun ved rikdommen, styrken og ryktet til hans borgere [...] mens i et demokrati, eller aristokrati, er det ikke alltid offentlig velstand som gir mest til den korrupte eller ambisiøse, men heller svikefulle råd, forræderske handlinger eller borgerkrig” (Hobbes 1946:123).

Det andre argumentet går på rådgivning, som er et svært viktig tema for Hobbes<sup>26</sup>. En monark har store fordeler på dette området, da han kan motta ”råd fra hvem, når og hvor han måtte ønske“ (Hobbes 1946:123). Dette gir monarken mulighet til å rådslå med eksperter på de feltene han måtte ønske, og ingen er utelukket. Han har også muligheten til å motta råd i det skjulte, noe Hobbes anser som viktig. Der suverenen er en forsamling blir det imidlertid mer problematisk; Hobbes hevder at kun de som i utgangspunktet har rett til å møte i forsamlingen, blir hørt, noe om er negativt, da disse “for det meste består av de som er mer kjent med tilegnelse av rikdom enn kunnskap” (Hobbes 1946:123). Videre, ved forsamlingsstyre gis råd i lange diskusjoner, der formålet er oppildning til, og overbevisning om, en viss fremgangsmåte, fremfor god gjennomføring av valgt strategi. Slike diskusjoner, der meninger fremlegges med “flammende lidenskap”, er ikke egnet til å skape forståelse, sier Hobbes, de skaper heller forvirring. Det er også et faktum at en forsamling vil ha problemer med å rådføre seg i hemmelighet (Hobbes 1946:123).

*Inkonsekvens* er Hobbes’ tredje argument mot forsamlingsstyre. Mens eneherskerens avgjørelser kun preges av den mangel på koherens som måtte stamme fra menneskets natur, har forsamlinger et langt større problem. De får et problem med sitt antall og sin varierende sammensetning; det som vedtas i dag kan ved neste korsvei reverseres, og mens dette av noen ansees som en fordel, blir det vanskeligere å arbeide med langsiktige planer (Hobbes 1946:123).

Et fjerde poeng er at det ikke finnes interessekonflikter (igjen, utover hva som eventuelt forårsakes av menneskets natur) hos en monark. I en forsamling kan derimot både misunnelse og egne interesser bringe uenighet til et konfliktnivå som forårsaker borgerkrig.

Hobbes’ femte poeng er meget sentralt, og kanskje spesielt interessant pga. innrømmelsen han her kommer med: “i et monarki har man den ulempen at enhver borger, ved

---

<sup>26</sup>Hele kapittel 25 i *Leviathan* (“OF COUNSEL”) tilegnes dette kapittelet, og det nevnes også hyppig ellers i hans verker.

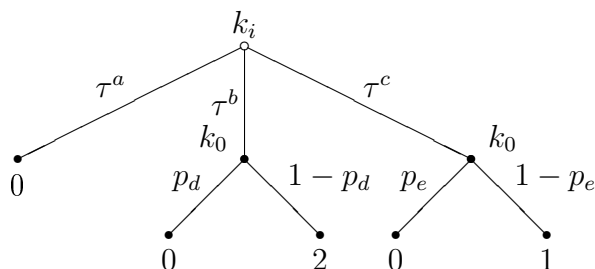
én manns makt kan fratas alt han eier, til fordel for en favoritt eller smigrer, og jeg innrømmer at dette er en stor – men uunngåelig – ulempe” (Hobbes 1946:123). Denne vilkårligheten er imidlertid ikke et argument for forsamlingsstyre, hevder Hobbes, da den like gjerne kan forekomme der, “for deres makt er den samme, de er utsatt for samme ondsinnede rådgivere, de kan forføres av dyktige talere, akkurat som en monark kan av sine smigrere” (Hobbes 1946:123). Videre vil det innad i en forsamling være en tendens hos medlemmene til å smigre hverandre når det gjelder én sak, for dermed å “kjøpe” seg støtte i andre saker. Også her ser vi hvordan Hobbes’ resonnement nærmer seg Rational choice, som gjennom begreper som “pork-barrel lovgivning” og “logrolling” på mange vis omtaler det samme som Hobbes. Videre er en monarks favoritter og slekt få i forhold til favorittene og slekten til samtlige i en forsamling (Hobbes 1946:123).

Hobbes’ sjettede punkt er også et tveegget argument. Han innrømmer her ulempene ved at makten i et monarki “kan nedarves til spedbarn, eller en som ikke kan skille godt fra ondt” (Hobbes 1946:124). Ved slike situasjoner blir makten i realiteten liggende hos andre, som kuratorer. Men, sier Hobbes, “å si at det er problematisk å legge suverenens makt i en mann eller forsamling, er å si at all regjering er et større besvær enn forvirring og borgerkrig” (Hobbes 1946:124). Det at makten kan tilfalle en “umyndig” person, innebærer ifølge Hobbes at det kan oppstå destruktiv kamp om posisjonen som kurator; han forsvarer allikevel eneveldet, siden monarken selv kan utnevne hvem som skal bekle en slik stilling. Ved problemer er dette dermed ikke på grunn av styreform, men “urettferdige borgere og deres ambisjoner”, som ikke avhenger av hvorvidt man styres av en forsamling eller én person (Hobbes 1946:124).

Hobbes avslutter så det sjettede punktet ved å hevde at *demokratier og aristokratier konstant er i samme posisjon som et enevelde der et barn er suverenen*:

For as a child wants the judgement to dissent from counsel given him, and is thereby necessitated to take the advice of them, or him, to whom he is committed; so an assembly wanteth the liberty to dissent from the counsel of the major part, be it good or bad. And as a child has need of a tutor, or protector, to preserve his person and authority; so also in great Commonwealths the sovereign assembly, in all great dangers and troubles, have need of *custodes libertatis*; that is, of dictators, or protectors of their authority; which are as much as temporary monarchs to whom for a time they may

commit the entire exercise of their power; and have, at the end of that time, been oftener deprived thereof than infant kings by their protectors, regents, or any other tutors (Hobbes 1946:122).



Figur 4.4: Et spill i ekstensiv form der forskjellene mellom demokrati og monarki illustreres. Aktør  $i$  forestilles å ha valget mellom ingen stat ( $\tau^a$ ), forsamlingsstyre ( $\tau^b$ ) og enevelde ( $\tau^c$ ). Etter første valg avgjøres utfall av sannsynlighet for krig,  $p_d$  ved demokrati og  $p_e$  ved enevelde.

I denne sammenheng kan det være illustrerende å se hva slags situasjon et hypotetisk individ før statsdannelsen står overfor. Dette vises i figur 4.4. Som vi har sett, forkaster Hobbes muligheten for et sivilt samfunn uten en stat, så  $\tau^a$  tillegges her verdien null, da dette representerer naturtilstanden. Forsamlingsstyre ( $\tau^b$ ) og en stat etter Hobbes' anbefalinger ( $\tau^c$ ) har to mulige utfall, der det er  $p_d$  sannsynlighet for krig under forsamlingsstyre og  $p_e$  sannsynlighet for krig under en Hobbesiansk stat. I begge tilfeller gir krig den laveste gevinst, her illustrert med den ordinale verdien 0.

$$p_d > p_e \tag{4.8}$$

Sannsynligheten for krig er, som vist i ligning (4.8), større ved forsamlingsstyre enn ved en Hobbesiansk stat, og dette er det sentrale for Hobbes. Han innrømmer at det finnes ulemper ved hans styreform, men mener at selv hans verste stat er bedre enn alternativene. For å illustrere at valget mellom demokrati og enevelde kan innebære en avveining, har jeg her gitt et *vellykket* forsamlingsstyre en høyere gevinst enn en vellykket Hobbesiansk stat. Dette for å illustrere at en slik stat på mange måter synes mer ønskelig enn en sterk autoritær stat, dersom begge er i stand til å frembringe fred. Denne rangeringen er altså basert på moderne (og antikke) idealer om staten, snarere enn Hobbes' tekst. Man er altså nok en gang i en situasjon der det er snakk om risikovurdering, og forholdet mellom sannsynlighetene og de mulige gevinstene må vurderes ut fra et individ med Hobbesiansk risikoaversjon.

$$E(\tau^b) = p_d 0 + (1 - p_d) 2 \quad (4.9)$$

$$E(\tau^c) = p_e 0 + (1 - p_e) 1 \quad (4.10)$$

Hobbes kan her forestille seg at  $p_e$  er såpass mye mindre enn  $p_d$  at hans stat ville vunnet ved en sammenligning av forventet gevinst:  $E(\tau^b) < E(\tau^c)$ , se ligning (4.9) og (4.10). Det er imidlertid også viktig å bemerke at krigsutfallet er meget sterkt uønsket, og at et Hobbesiansk individ dermed vil legge sterk vekt på å unngå dette utfallet. Det er innlysende i forbindelse med Hobbes' filosofi at dersom man erstatter de ordinale verdiene i figuren med kardinale verdier, vil krig som utfall få en sterkt negativ verdi; selv små forskjeller i sannsynligheten for krigsutbrudd vil dermed kunne få store konsekvenser for hvilket alternativ som fremstår som det beste.

At Hobbes *foretrekker* enevelde er klart, både ut fra argumentene som er presentert her og hans tallrike referanser til de feilaktige doktrinene om frihet og folkestyre i antikken. Hobbes anerkjenner videre kun tre styreformer: demokrati, aristokrati og monarki. Anarki, oligarki og tyranni er for Hobbes akkurat det samme, navngitt av de som er misfornøyde (Hobbes 1946:121,447). Hobbes vurderer styreformenes meritter utelukkende ut fra deres evne til å oppfylle statens mål:

The difference between these three kinds of commonwealth, consisteth not in the difference of power; but in the difference of convenience, or aptitude to produce the peace, and security of the people; for which end they were instituted (Hobbes 1946:122).

Dersom demokrati har vist seg å være det beste for å skape fred og sikre borgernes liv og eiendom, finnes det ingen indikasjon på at ikke denne styreformen ville foretrekkes. En slik forståelse, der man åpner for at Hobbes i en alternativ situasjon kunne godtatt og anbefalt demokrati, gir en mer tidløs Hobbesiansk filosofi. Det fører dog også til at man må se svært nøye på *Leviathan* for å skille de generelle prinsippene fra de historisk og kulturelt betingede. Det finnes helt opplagt mange fordeler ved demokratiet som ikke presenteres av Hobbes, og en videre vurdering av dette premisset foretas i 5.3.3.

## Risikovillighet og ønsket om sikkerhet

**Premiss 5** Mennesker er ikke villige til å ofre sikkerhet for andre goder

Jeg har allerede fastslått at død, og stor fare for død, er det verste ondet for Hobbes<sup>27</sup>. Det gjenstår imidlertid å klargjøre hvilke implikasjoner dette forholdet har på individers handlinger. En risikosøker ville ikke blitt nevneverdig påvirket av en ørliten mulighet for et negativt utfall, mens andre igjen ville forsake enorme gevinster dersom disse var forbundet med fare for livet. Ved å lese Hobbes fremkommer det tydelig at sikkerhet – minimalisering av risiko – er av stor betydning for mennesker. At selv den verste stat er bedre enn alternativet (borgerkrig/naturtilstanden), er et eksempel på dette (Hobbes 1946:120). Med et slikt syn kan man vanskelig hevde at det f.eks. kan være verdt å ofre sikkerhet for økt frihet, et tema som behandles inngående i 5.3.

Frank Knight får æren for å ha definert *usikkerhet* og *risiko* slik vi i dag bruker begrepene innen beslutningsteori (Weber og Johnson 2009:131). *Usikkerhet* i en beslutningssituasjon innebærer ifølge Harsanyi at “i alle fall visse objektive sannsynligheter er ukjente for [aktøren]”, mens man står ovenfor *risiko* i situasjoner der man kjenner de objektive sannsynlighetene (Harsanyi 1977:22). Weber og Johnson (2009:131) definerer usikkerhet som at “sannsynlighetene for forskjellige utfall ikke kan uttrykkes med matematisk presisjon”, mens Luce og Raiffa (1990:20) beskriver det som at “sannsynlighetene for [disse] utfallene er fullstendig ukjente, eller til og med meningsløse”. Det er dog viktig å være klar over at begrepene også kan bety noe annet enn hva konvensjonen innen økonomi skulle tilsi. von Neumann og Morgenstern (2004:163) omtaler *risiko* som “the worst that can happen under the given conditions”. *Norsk Riksmålsordbok* definerer *risiko* som “1) mulighet for tap, skade ell. lign.; vågnad [...] 2) eventuelt tap, eventuell skade”. For mer inngående behandling av den konvensjonelle bruken av begrepene se bl.a. Luce og Raiffa 1990:20, Harsanyi 1977:22, Render et al. 2006:71.

Fremfor alt må det gjøres klart at det er forskjell på å vurdere handlinger på bakgrunn av *sannsynlighet* og *mulighet*. Innen økonomi er det vanlig å anta at mennesker foretar beregninger på bakgrunn av subjektive sannsynligheter. For visse deler av realismeskolen i politisk teori er det derimot sentralt med *muligheter* for at visse utfall kan oppstå, nærmest uavhengig av sannsynligheter. Mens man ikke kan utelukke en mulighetsorientering hos Hobbes, blir det raskt klart at en *absolutt* mulighetsorientering,

---

<sup>27</sup>Under premiss 2 og i 2.3.4.



blandet med et syn på døden som det verste ondet, er empirisk feil, og jeg kan ikke tillegge Hobbes en teori ethvert oppegående menneske ville forstå var feil. Et eksempel på hvor absurd en slik teori ville være, finnes ved å se på at *alle* handlinger innebærer en viss sjanse for at døden inntreffer, enten som en direkte eller indirekte konsekvens. Mulighetsorienterte individer ville altså leve et liv der nær sagt samtlige handlinger de foretok seg gikk mot deres egen overbevisning; alternativt ville de kun anse sannsynligheter som var over en viss grense, som *mulige*. Det er uansett lite sannsynlig at slike mennesker f.eks. ville kjøre bil; i så tilfelle ville grensen være såpass høy at det uansett ville være mer nyttig å anse dem som sannsynlighetsorienterte. Det finnes også mye god kritikk av enkel *subjektiv forventet nytte/gevinst* som fremgangsmåte for å analysere beslutninger. Sentralt i denne kritikken er prospektteorien til Kahneman og Tversky (1990), og f.eks. “Allais og Ellesberg-paradokset” som bl.a. er beskrevet av Raiffa (2002:45).

Hobbes’ sikkerhetsfokus forutsetter at man ikke ønsker å gamble i utfall der man kan oppnå svært ufordelaktige utfall. Det er dette det dreier seg om når det i premisset forutsettes at individer ikke ofrer sikkerhet for andre goder. Hobbes mener man ikke vil ofre statens stabilitet for å oppnå f.eks. økt frihet. Det kan høres ut som om Hobbes her innfører en mulighetsorientert (kontra sannsynlighetsorientert) vurdering av staten, og dette kan faktisk ikke utelukkes. Det er imidlertid i så tilfelle forskjell på individers vurdering av individuelle gjøremål og spørsmål i forhold til statens form og forfatning. Det kan tenkes at Hobbes’ tanke er at mennesker ikke *kan/bør* gamble med sin sikkerhet i spørsmål om styreform dersom de skal oppnå trygghet, mens mange mennesker allikevel til en viss grad er villige til å ta sjanser i sitt private liv<sup>28</sup>.

Det er også interessant å vurdere hvorvidt kardinale verdier, som kan sammenlignes med andre goders verdi, kan innføres for døden o.l. Dersom dette kan gjennomføres på en god måte (noe som synes noe problematisk å få til, da man vanskelig får målt eller testet menneskers reaksjoner på f.eks. døden), synes det tilstrekkelig å forholde seg til vanlig kalkulerings av forventet verdi, selv om prospektteorien i visse tilfeller fremstår som en bedre *deskriptiv* teori. Det å dø vil nødvendigvis få en såpass negativ verdi at selv en liten mulighet for at det inntreffer vil ha store konsekvenser for aktørens vurderinger.

---

<sup>28</sup>Det samme problemet diskuteres av Slomp og LaManna (1996), og de mener Hobbes’ individers risikoaversjon ikke fører til et “vegetativt” liv, men at ingen mengde goder som f.eks. ære kan veie opp for å bli drept (Slomp og LaManna 1996:60).

Dette premisset er for meg det mest problematiske å forsvare, og det diskuteres ikke spesielt inngående eller konkret hos Hobbes. Den ekstreme risikoaversjonen som kreves for at Hobbes' resonnement skal være gyldig kan synes å være en *implisitt forutsetning* som ikke forsvares tilstrekkelig. At det *er* et sentralt premiss i resonnementet fremkommer tydelig fra hvordan jeg har formulert "Hobbes' spill" i premiss 2, da sannsynligheter er en uunngåelig del av begge strategier. Spiller *i*'s forventningsverdier for sine mulige strategier i Hobbes' naturtilstand-spill (figur 4.3) er som følger:

$$E_i(\tau_i^1) = p_{\tau_j^1}MV + p_{\tau_j^2}T \quad (4.11)$$

$$E_i(\tau_i^2) = p_{\tau_j^1}H + p_{\tau_j^2}[p_i^{k_1}\bar{\theta} + p_i^{k_2} + p_i^{k_3}H] \quad (4.12)$$

Spiller *i* har her et valg mellom to strategier, der strategi  $\tau^2$  innebærer en viss sannsynlighet for å dø. Ved å følge en maximin-strategi vil man da nødvendigvis velge strategi  $\tau^1$  dersom man bruker preferanserangeringen  $\bar{\theta} > K > T > MV > H$ . Ved vanlig nyttemaksimering, dersom vi her antar at  $K$  tillegges en sterkt negativ verdi (grunnet at den inneholder en ikke ubetydelig sannsynlighet for å dø), vil man med rene strategier forvente at aktørene velger  $\tau^1$ . Selv om spillerne skulle tillates å "gamble" med blandede strategier, ville man forvente en svært liten sannsynlighet for  $\tau^2$ . Samme resonnement kan antas for *j*. Man ville i så tilfelle kunne forvente et "fredelig" utfall, noe som også stemmer godt med den tolkningen av Hobbes' syn på naturtilstanden som presenteres i neste kapittel: naturtilstanden dreier seg ikke om konstant fysisk vold, men om en konstant *mulighet* for dette (Hobbes 1946:82).

## Hobbes' "liberalisme"

**Premiss 6** Staten skal ikke frata mennesker mer frihet enn nødvendig

But by safety here [i forbindelse med suverenens hensikt], is not meant a bare preservation, but also all other contentments of life, which every man by lawful industry, without danger, or hurt to the commonwealth, shall acquire to himself (Hobbes 1946:219).

But what is a good law? [...] It is in the laws of a commonwealth, as in the laws of gaming: whatsoever the gamesters all agree on, is injustice to none

of them. A good law is that, which is *needful*, for the *good of the people*, and withal *perspicuous* (Hobbes 1946:227).

Dette premisset hindrer den totalitære stat som konklusjon på resonnementet, noe som er et viktig poeng. Slik sitatet i åpningen av kapittelet viser, anser Hobbes selv sin doktrine som mellomliggende mellom de doktriner som ønsker for mye frihet og de som ønsker for mye autoritet. Det er altså ikke et resonnement fullstendig ubalansert i retning av mest mulig autoritet på alle områder<sup>29</sup>. Staten har et (minimalt) *mål*, og kun den autoritet som er strengt nødvendig for å sikre dette (i Hobbes' øyne), er ønskelig.

By LIBERTY, is understood, according to the proper signification of the word, the absence of external impediments: which impediments, may oft take away part of a man's power to do what he would; but cannot hinder him from using the power left him, according as his judgment, and reason shall dictate to him (Hobbes 1946:84).

Denne typen frihet ligger nærme beskrivelsen Berlin (2002c:169) gir på *negativ frihet*, som ofte blir brukt sammen med *positiv* frihet for å beskrive to meget forskjellige måter å forstå begrepet på. Mens negativ frihet ofte sies å være frihet *fra*, omhandler dens motsats frihet *til* å gjøre noe. Jeg vil ikke gå inn i denne debatten, men bemerker at Hobbes stiller seg i leiren for en minimalistisk forståelse av begrepet, samtidig som han forsvarer denne friheten sterkt. En sjømann som tvinges til å dumpe sin last i frykt for at skipet vil synke er fremdeles *fri*; han kan velge å la være, og frihet består altså av fravær av tvang eller hindringer. Frihet består *ikke* i å være i stand til å motså visse ting, eller i å være i stand til å utføre noe. Vann er utsatt for nødvendigheter i form av fysiske lover, men Hobbes mener det blir absurd å hevde at vann mangler frihet fordi det ikke kan renne oppover fremfor nedover (Hobbes 1946:137). Det samme får implikasjoner for Hobbes' individer, på grunn av hans mekanistiske materialisme; siden alt vi gjør er resultatet av materie som beveger seg på visse måter, er våre handlinger også et resultat av fysiske nødvendigheter. Dette, som med vannet, gjør oss allikevel ikke ufrie, da den eneste grunnen til at vi makter å gjøre det vi ønsker, er at ingen ting hindrer dette (Hobbes 1946:137). Denne determinismen hos Hobbes er kontroversiell for mange, da "fri vilje" blir en umulighet; per Hobbes' definisjoner følger våre handlinger av vår vilje, og dersom det ikke står noe i veien for dette, er vi fri. Dette er så langt

---

<sup>29</sup>Skinner (1996) påpeker imidlertid at det er et velkjent retorisk grep å innlede en tekst ved å påberope seg nøytralitet og saklighet, slik Hobbes med sitatet som her omtales gjør i *Leviathan*.

Hobbes strekker seg i denne saken, og et slikt syn er minst like vanskelig for mange å akseptere som den oppfatning at mennesker er utelukkende egeninteresserte. Gud sier at vi har fått fri vilje, men er fri vilje forenlig med en forståelse av at alt vi gjør, erfarer og tenker, kun er reaksjoner på forutgående hendelser, som for et allvitende vesen kunne forutsies uten problemer? Dette temaet diskuteres videre i 5.3.

Poenget med Hobbes' "liberalisme" er dog ikke viljens frihet, et begrep han selv avfeier som "absurd tale":

And therefore if a man should talk to me of a *round quadrangle*; or *accidents of bread in cheese*; or, *immaterial substances*; or of a *free subject*; a *free will*; or any *free*, but free from being hindered by opposition, I should not say he were in an error, but that his words were without meaning, that is to say, absurd (Hobbes 1946:27).

Poenget er at individer skal ha så mye frihet som mulig uten at det svekker statens muligheter til å oppfylle sitt mål. For Hobbes er frihet lik fravær av lov i en stat (Hobbes 1946:138,143). Menneskers frihet ligger kun i de ting "suverenen har bestemt: som f.eks. friheten til å handle, selge og ellers inngå kontrakter med hverandre; de kan velge eget bosted, sin egen diett, sitt yrke, de kan danne sine barn som de selv anser best, osv." (Hobbes 1946:139).

Det er fullt mulig å steile ved at jeg stiller dette i nærheten av "liberalisme", selv med gåsetegn. Jeg mener imidlertid det finnes gode grunner til dette: a) mennesker bør være fri til å forfølge sitt eget beste i størst mulig grad, b) individet selv er den beste dommer over egne anliggender, c) negativ frihet er det liberale idealet (ingen kan *tvinges til å være fri* f.eks.) og d) der loven ikke skaper hindringer er mennesket fritt, og Hobbes krever at det ikke skal lages lover som ikke er nødvendige (Hobbes 1946:45,227).

For the use of laws, which are but rules authorized, is not to bind people from all voluntary action; but to direct and keep them in such a motion, as not to hurt themselves by their own impetuous desires, rashness or indiscretion; as hedges are set, not to stop travellers, but to keep them in their way. And therefore a law that is not needful, having not the true end of a law, is not good. A law may be conceived to be good, when it is for the benefit of the sovereign; though it be not necessary for the people; but it is not so. For the good of the sovereign and people, cannot be separated. (Hobbes 1946:227).

At Hobbes mener suverenen kun bør implementere *gode* lover følger av at det gode i denne sammenheng er samfunnets gode, som er staten mål. Hobbes er tydelig på at frihet er viktig, så fremt det er kompatibelt med sikkerhet. Selv om sikringen av frihet synes svært mangelfull i forhold til dagens konvensjoner, synes Hobbes' intensjoner tydelige.

### 4.2.3 Resonnementets konklusjon

**Konklusjon** Hobbes' stat er *nødvendig* og *ønskelig*

Konklusjonen på Hobbes' resonnement er altså at Hobbes' stat er *nødvendig* for å unnsnippe naturtilstanden, og derfor også *ønsket* av mennesker. Jeg mener denne konklusjonen er gyldig, gitt premissene resonnementet bygger på, og jeg antar her som nevnt at disse premissene er korrekte, selv om jeg har stilt visse spørsmål som vil følges nærmere i neste kapittel.

#### Kort tekstuell vurdering av resonnementet

**Premiss 1** legger til grunn at menneskers naturlige situasjon er konfliktfylt, og at mennesker ikke er i stand til å unnsnippe denne situasjonen uten å danne en overordnet makt. Konklusjonen inneholder dermed et krav om at staten må dannes dersom mennesker ønsker å unnsnippe naturtilstanden.

**Premiss 2** fjerner tvilen som muligens finnes om hvorvidt mennesker kan være fornøyd i naturtilstanden. De ønsker *ikke* å forbli i denne situasjonen, da det ikke finnes ondt som er verre enn døden eller stor fare for døden. Med dette etableres det at staten er *ønsket*.

**Premiss 3** gjør det klart at makt ikke kan deles. Suverenen må ha absolutt makt, og dette premisset forhindrer derfor en moderne stat med konstitusjonell maktdeling og "checks and balances".

**Premiss 4** forklarer hvordan de forskjellige styreformene er forskjellige, og mer spesifikt hvorfor et enevelde for Hobbes er den formen som er best egnet til å skape og bevare ro og orden. Sammen med premiss 5 vil dermed demokrati/forsamlingsstyre måtte forkastes.

**Premiss 5** utelukker at mennesker kan tenkes å foretrekke f.eks. frihet fremfor sikkerhet. Sikkerhet for sine liv er det viktigste for individer, og de vil ikke bytte dette mot andre, mindre viktige, goder. Der man har valgmuligheter i forhold til statens form, må man ifølge Hobbes velge det som gir mest mulig sikkerhet

**Premiss 6** krever at det er en *grense* for hva staten kan foreta seg. Staten har et formål, som er å skape sikkerhet, og Hobbes gjør det klart at det ikke er nødvendig med full kontroll over alle aspekter av staten og menneskers liv for å oppnå dette.

**Konklusjon:** Staten må styres av suverenen, som helst er én person, og denne må ha all makt. Statens oppgave er å sikre sine borgeres liv, og der andre goder og sikkerhet er i konflikt, må sikkerhet prioriteres; staten er altså autoritær, ikke liberal. Den er dog heller ikke totalitær, da staten ikke kan gå utover sitt begrensede formål – mennesker skal ha så mye frihet som er forenlig med sikkerhet for alle.

I tabell 4.1 har jeg satt opp en vurdering av implikasjonene dersom enkelte premisser utelates fra resonnementet.

Premisser	Stat	Makt	Styreform	Regimetype
X23456	n/i	a	e	a
1X3456	n/i	a	e	a
12X456	n	a/i	e	a
123X56	n	a	d/a/e	a
1234X6	n	a	d/a/e	l/a
12345X	n	a	e	a/t
123456	n	a	e	a

Tabell 4.1: Premissimplikasjoner. Tabellforklaring: Premisser: X betyr at premisset er utelatt, Stat=n/i (Nødvendig/ikke nødvendig), Makt=a/i (Absolutt/innskrenket), Styreform=d/a/e (Demokrati/aristokrati/enevelde), Regimetype=l/a/t (Liberal/autoritært/totalitært)

### 4.3 Konklusjon

Mange har motforestillinger mot Hobbes' stat, og vi har sett at også Hobbes er klar over ulemper ved sitt ideal for staten. Han synes dog å være meget klar på at han har vurdert problemene, og han slutter at hans løsning ikke er til å unngå, dersom mennesker skal oppnå *fred*:

[...] not considering, that the state of man can never be without some incommmodity or other; and that the greatest, that in any form of government can possibly happen to the people in general, is scarce sensible in respect of the miseries, and horrible calamities, that accompany a civil war, or that dissolute condition of masterless men, without subjection to laws, and a coercive power to tie their hands from rapine and revenge [...] (Hobbes 1946:120)

*Hvorfor* det er så vanskelig for mange å si seg enig med Hobbes er noe jeg vil undersøke nærmere i neste kapittel. Spesielt viktige er spørsmålene om statens nødvendighet, Hobbes' absolutisme og valget av styreform.

Resonnementet, her gitt at premissene er sanne, fremstår for meg som gyldig. Det mest interessante i denne forbindelse gjenstår imidlertid, og det er å vurdere premissenes sannhetsverdi. Som jeg har vist i tabell 4.1 endres konklusjonen dersom enkelte av premissene fjernes; dersom historien og moderne bidrag kan sannsynliggjøre at premissene må fjernes eller modifiseres, må vi altså forvente at konklusjonen vil måtte endres.

Hobbes er forøvrig klar over at det finnes få eksempler på stater der hans foretrukne styreform, med den totale og samlede makt hos suverenen, har eksistert. Det er også et faktum at selv om jeg kaller Hobbes en autoritær tenker, anser jeg ikke de autoritære regimer som har eksistert som representative for Hobbes' ideal. Dette gjør det altså noe vanskelig å hevde med absolutt sikkerhet at historien har motbevist Hobbes, selv om det nok kan synes som at noen av hans meninger var *vel* ekstreme. Jeg mener derfor at man ikke automatisk kan forkaste Hobbes' filosofi uten å vise hvordan han tar feil, noe jeg vil fokusere på i resten av oppgaven.

For though in all places of the world, men should lay the foundation of their houses on the sand, it could not thince be inferred, that so it ought to be (Hobbes 1946:136).

Tabell 4.2: Hobbes' resonnement

P1	<i>Fred</i> uten en overodnet samlende makt er ikke mulig
P2	Død – krig, og spesielt borgerkrig er det verste ondet for mennesker
P3	Makt <i>kan</i> ikke deles, kun midlertidig delegeres
P4	Demokratier er mer utsatt for uroligheter enn ved et enevelde
P5	Mennesker er ikke villige til å ofre sikkerhet for andre goder
P6	Staten skal ikke frata mennesker mer frihet enn nødvendig
K	Hobbes' stat er <i>nødvendig</i> og <i>ønskelig</i>



# Kapittel 5

## En kritisk gjennomgang av Hobbes' politiske resonnement

### 5.1 Innledning

And as to the whole doctrine [som fremsatt i *Leviathan*], I see not yet, but the principles of it are true and proper; and the ratiocination solid. For I ground the civil right of sovereigns, and both the duty and liberty of subjects, upon the known natural inclinations of mankind, and upon the articles of the law of nature; of which no man, that pretends but reason enough to govern his private family, ought to be ignorant (Hobbes 1946:465).

Hobbes' resonnement om menneskets natur, problemene denne naturen forårsaker, og løsningen på disse problemene, er allment akseptert som skarpsindig. Det er imidlertid et stort problem: han kom frem til "feil" konklusjon! Feil i forhold til våre moderne idealer, og samtidig feil i forhold til hvordan vi ønsker å oppfatte oss selv og andre.

I kapittel 4 har jeg fremsatt det jeg anser for å være de sentrale premissene i Hobbes' resonnement, og det bør nå være mulig å vurdere gyldigheten av disse premissene. Oppgavens rammer begrenser mulighetene for en uttømmende gjennomgang av alle premissene, så jeg vil fokusere på to for meg avgjørende temaer i forhold til Hobbes' konklusjon: 1) menneskers behov for en stat overhodet og 2) individers forhold til frihet, demokrati og sikkerhet *etter* at de har kommet seg ut av naturtilstanden.

Det første punktet innebærer en vurdering av premiss 1 fra 4.2.2, som går på at staten er nødvendig for å unnsnippe naturtilstanden. Jeg ønsker å undersøke hvor sterke argumentene for anarki som et alternativ til staten kan sies å være, gitt Hobbes' syn på mennesket og dets situasjon. Det andre punktet favner langt videre, og vil kunne ha implikasjoner for de øvrige premissene, som omhandler døden som det verste ondet, maktdeling, eneveldets overlegenhet i forhold til andre styreformer, menneskers vilje til å utsette seg for usikkerhet og hvor mye frihet Hobbes mener undersåttende bør ha.

Underveis i kapittelet vil jeg undersøke hvorvidt premissene fremsatt i kapittel 4 er sårbare for kritikk. Premissene vil revideres der jeg finner grunnlag for å hevde at Hobbes tok feil. Avslutningsvis vil jeg ha et sett med premisser, der noen er forskjellige fra utlegningen i kapittel 4, og det vil være mulig å utlede en ny konklusjon på bakgrunn av dette nye settet med premisser.

## 5.2 Argumenter mot staten som en nødvendighet

**Premiss 1** *Fred* uten en overordnet samlende makt er ikke mulig<sup>1</sup>

Michael Taylor har skrevet flere bøker om anarki, (kommunitarianske) samfunn og frivillig samarbeid. Jeg har valgt å hovedsakelig basere meg på *The Possibility of Cooperation* (Taylor 1987), da han her har en relativt inngående diskusjon av Hobbes og samhandlingsproblemer. Taylor har i denne boken som hovedmål å utforme et sterkt angrep på det han kaller den "liberale" rettferdiggjørelsen av staten:

The most persuasive justification of the state is founded on the argument that, without it, people would not successfully cooperate in realizing their common interests and in particular would not provide themselves with certain public goods [...] (Taylor 1987:1).

Grunnen til at Taylors resonnement er av interesse, er at premiss 1 som her diskuteres er basert på den type argumentasjon Taylor angriper. Jeg har fremsatt den påstand at *fred* uten en overordnet samlende makt ifølge Hobbes ikke er mulig, og jeg har basert dette på min tolkning av *Leviathan*. Taylor ville med stor sannsynlighet godtatt

---

<sup>1</sup>Dette premisset ble opprinnelig etablert i 4.2.2. Jeg anvender her Hobbes' definisjon av fred – fravær av krig, se også 5.2.3 (Hobbes 1946:82).

denne tolkningen av Hobbes, og siden han selv anser Hobbes' bidrag som "det første fulle uttrykk for denne måten å rettferdiggjøre staten på" – og gjør utstrakt bruk av Hobbes i sitt resonnement – anser jeg en vurdering av premiss 1 på bakgrunn av Taylors resonnement som godt egnet i denne sammenheng (Taylor 1987:1).

Taylor argumenterer også sterkt for at Hobbes' naturtilstand må forstås som et Fangenes dilemma-spill (Taylor 1987:32,126,129,133). Jeg har under premiss 2 i 4.2.2 (premisset som legger til grunn at døden er det verste ondet) argumentert for at naturtilstanden kanskje heller bør forstås som et Chicken-spill, og jeg vil se på årsakene til våre forskjellige oppfatninger i 5.2.3.

### 5.2.1 Fellesgoder og samarbeid

De fem første kapitlene av *The Possibility of Cooperation* (Taylor 1987) gir oss Taylors spillteoretiske vurdering av fellesgode-problemer. Jeg vil først kort presentere Taylors resonnement om dette emnet. Deretter vil jeg se på hans tolkning av Hobbes i 5.2.3. Jeg finner Taylors drøfting av samhandlingsproblemer svært god, men jeg har visse problemer med hvordan han tolker Hobbes. Dette skaper imidlertid ikke alvorlige problemer, da Taylor selv har utført disse analysene adskilt; tolkningen av Hobbes kan, dersom det er ønskelig, forkastes, mens behandlingen av samhandlingsproblemer bevares.

"Problemet" med fellesgoder er at det som oftest fører til *samhandlingsproblemer*, som ifølge Taylor dreier seg om "at det er usannsynlig at rasjonelle egoister vil lykkes i å samarbeide om å fremme fellesinteresser" (Taylor 1987:3). Taylor peker på Garret Hardins "allmenningens tragedie" som et eksempel på et slikt problem (Taylor 1987:3). Samhandlingsproblemer oppstår der gevinster er *private* mens kostnader er *offentlige*; i allmenningens tragedie – eksemplifisert ved en felles beitemark – har hver enkelt bonde insentiver til å overbefolke beitemarken med kyr; fordelen for bonden ved å utplassere ekstra kyr er privat, mens kostnaden ved overbefolkning på beitemarken må deles av alle.

Det er også viktig å påpeke at når Taylor omtaler "rasjonelle egoister", er det snakk om det jeg her kaller *egen-anskuende* aktører; hans definisjon av "altruisme" svarer til min definisjon av å være *sosialt anskuende*. "Negativ altruisme" for Taylor er å fokusere på relativ gevinst og eminens, mens "positiv altruisme" går på vurderinger av andres nytte som *positiv* for en selv (Taylor 1987:33,109,111). "Sofistikert altruisme" er for Taylor

å ha nyttefunksjoner som inneholder andres *nyttevurderinger* – ikke bare gevinster; dette blir hos ham i all hovedsak forbigått som interessant, men svært komplisert og problematisk (Taylor 1987:111)<sup>2</sup>.

Før Taylors argumenter mot Hobbes fremsettes, er det nødvendig å gjøre rede for hvordan han formaliserer fellesgode-problematikken. Først og fremst forkaster han engangspill som svært lite relevante i analysen av reelle samhandlingsproblemer. Videre hevder han at dersom disse samhandlingssituasjonene skal modelleres som gjentatte spill, er det nødvendig med *ubestemt tidshorisont* (antall perioder,  $t$ , er ubestemt) (Taylor 1987:ix,xii).

Ifølge Taylor er det i denne sammenheng nødvendig å behandle populasjoner der  $n > 2$  dersom man skal få en realistisk fremstilling av sosiale problemer i større menneskegrupper. Taylor mener dessuten at det å basere seg på *parvise interaksjoner* ikke gir en tilfredstillende representasjon av individers samhandling (Taylor 1987:71). I denne forbindelse tar Taylor et forholdsvis kraftig oppgjør med Robert Axelrods bidrag til forståelsen av spontant samarbeid, som er basert på parvise interaksjoner der spillerne spiller hele spillet med samme motpart (Taylor 1987:71)<sup>3</sup>.

Axelrods bidrag er meget interessant, men jeg vil i denne oppgaven nøye meg med å se på Taylors bidrag, da dette eksplisitt knyttes til Hobbes, og etter min mening er tilstrekkelig for å vise at samarbeid *er* mulig uten tvang/ekstern koordinering – noe som også ville vært grunnen til å inkludere Axelrod (2006). Videre er Taylors innvendinger mot Axelrod gode, og jeg anser Taylors behandling av samhandlingsproblemer som den mest realistiske av de to.

*Betinget samarbeid* viser seg å være et svært viktig begrep i analysen av samhandlings-situasjoner. For at betinget samarbeid skal være mulig er det nødvendig med mer enn én tidsperiode:  $t > 1$ . Det er ifølge Taylor også nødvendig å forutsette utstrakt interaksjon i populasjonen, ikke bare interaksjon i isolerte par. En aktørs “avgjørelse av hvorvidt han skal bidra i produksjonen av fellesgoder vil generelt sett være avhengig av om *tilstrekkelig mange andre* samarbeider. Slike situasjoner kan ikke representeres som parvise interaksjoner” (Taylor 1987:71).

---

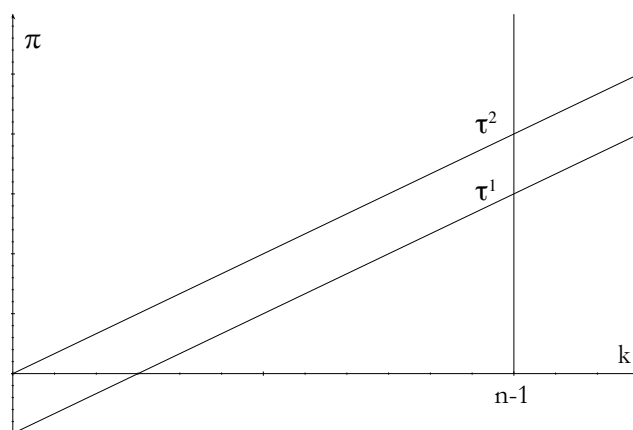
<sup>2</sup>Jeg argumenterer i 2.6 for at en nyttefunksjon med “søfistikert altruisme” kan være godt egnet for å representere Hobbes’ individer.

<sup>3</sup>Axelrod (2006) presenterer i sin bok *The Evolution of Cooperation* resultatene fra sine nå berømte data-turneringer. Der ble strategien tit-for-tat kåret til seierherre i den første turneringen. Også i en påfølgende turnering – der deltagerne hadde muligheten til å modifisere sine programmer i lys av resultatene av den første turneringen – viste tit-for-tat seg å være sterkest.

Hvordan man velger å formalisere interaksjoner i en populasjon har avgjørende betydning for ens resonnement. Det er dog problemer forbundet med alle måter å løse dette på, da realistiske formaliseringer vil anta en kompleksitet som blir vanskelig å håndtere. Selv om det kan synes mest realistisk å se på interaksjon mellom samtlige i en gruppe, slik Taylor gjør, støter man på problemer i forbindelse med a) hvordan man forholder seg til koalisjonsdannelse, og b) hvem eller hva individer sammenligner seg og sin situasjon med. Jeg holder meg i første omgang til Taylors fremgangsmåte, og vil diskutere hva jeg anser som mest hensiktsmessig i forbindelse med Hobbes' naturtilstand i 5.2.3.

Jeg vil i det følgende presentere Taylors behandling av  $n$ -person Fangenes dilemma-spill, og vil i forbindelse med vurderingen av hans tolkning av Hobbes i 5.2.3 se kort på hvordan Chicken-preferanser ville påvirket denne analysen. Denne gjennomgangen er en videreutvikling av formaliseringen av toperson-spill som ble gjennomgått under premiss 1 i 4.2.2.

Taylor anvender begrepet (*gjentatt*) spill ("supergame") for det endelige og altomfattende spillet der man har et gjentatt spill, og hver gjentakelse som *delspill* ("constituent game") (Taylor 1987:60). Taylor vurderer kun spill der det er snakk om gjentakelse av det *samme spillet*, noe jeg vil kommentere nærmere i 5.2.3. Jeg mener dette skaper problemer når det kommer til Taylors analyse av Hobbes, og ønsker å se på en alternativ fremgangsmåte til dette.



Figur 5.1:  $n$ -person Fangenes dilemma-spill

Strategiene i Taylors spill betegnes med  $C$  og  $D$  for henholdsvis *samarbeid* og *hoppe av*. Jeg vil som tidligere i oppgaven anvende  $\tau^1$  for  $C$  og  $\tau^2$  for  $D$  i denne gjennomgangen.

I figur 5.1 vises gevinsten for de to strategiene ved varierende samarbeidsnivåer i et engangsspill med Fangenes dilemma-struktur.  $\pi$  indikerer gevinst, mens  $k$  er antall som samarbeider (velger  $\tau^1$ ). Den vertikale streken  $n - 1$  illustrerer maksimalt antall andre samarbeidende, altså populasjonen  $n$  minus spilleren som betrakter figuren. Origo defineres til gevinsten knyttet til  $\tau^2$  dersom ingen samarbeider.

I denne figures vises det tydelig at  $\tau^2$  dominerer  $\tau^1$ , da den for alle verdier av  $k$  gir høyere gevinst. Uansett hvor mange andre som samarbeider, vil man (dersom man ignorerer alle fremtidige konsekvenser, eller disse konsekvensene er inkludert i gevinsten), tjene på å bytte fra  $\tau^1$  til  $\tau^2$ . Dette er grunnen til at det forventes samhandlingsproblemer i slike engangsspill.

I gjentatte spill finnes det imidlertid en strategi Taylor betegner med  $B_\sigma$ , som representerer “tit-for-tat” (“like for like”)<sup>4</sup>, der  $\sigma$  symboliserer hvor mange andre som må samarbeide for at aktøren med denne strategien skal velge  $\tau^1$ . Dersom  $\sigma = n - 1$  kreves det at *samtlig andre spillere samarbeider*, og i så tilfelle gir dette en likevekt, dersom aktørenes diskonteringsfaktor ikke er for lav (Taylor 1987:86). Denne strategien kan ikke illustreres i figur 5.1, da figuren kun illustrerer et delspill i det gjentatte spillet.

I Taylors modell er individens *diskonteringsfaktor* ( $\delta_i$ ) av sentral betydning. En for høy diskonteringsrate, med tilhørende lav verdsetting av fremtidig nytte, utelukker nærmest enhver mulighet for samarbeid i Fangenes dilemma-spill, selv der spillet gjentas. Jeg har i 4.2.2 sett på hvilken diskonteringsfaktor som er nødvendig for å skape samarbeid i et topersons Fangenes dilemma-spill med grim trigger-strategi.

Først vil jeg se hvordan Taylor legger frem delspillet det gjentatte spillet består av. Taylor utleder forutsetningene for selvhåndhevende samarbeid i et  $n$ -person spill ut fra funksjonene for gevinst ved valg av forskjellige strategier, og jeg vil her innføre noe av Taylors notasjon; strategi  $\tau^1$  gir  $f(k)$  og strategi  $\tau^2$  gir  $g(k)$ . Ligningene nedenfor gir gevinstene i hvert delspill (engangsspill):

$$f(k) = \frac{(k+1)r}{n} - 1 \quad (5.1)$$

$$g(k) = \frac{kr}{n} \quad (5.2)$$

---

<sup>4</sup>“Strategien tit-for-tat innebærer rett og slett å starte med samarbeid, for deretter å gjøre det samme som motspilleren gjorde i forrige periode” (Axelrod 2006:viii).

I disse ligningene representerer  $k$  antall andre som velger  $\tau^1$ , og  $r$  gir total gevinst for gruppen for hver enhet som ytes til fellesskapet.  $n$  representerer populasjonen, og gevinsten fra én enhet anvendt til fellesgodet gir altså en gevinst på  $\frac{r}{n}$  til hver aktør (Taylor 1987:96). Funksjonen  $g(k)$ , for strategi  $\tau^2$ , gir oss aktør  $i$ 's "gratis" nytte fra de andres samarbeid. I funksjon  $f(k)$ , for  $\tau^1$ , økes antall samarbeidende med én (siden aktøren slutter seg til samarbeidet) og  $i$ 's kostnad for å bidra ( $-1$ ) trekkes fra. Denne modellen er altså begrenset til en situasjon der man kun kan bidra med én enhet til fellesskapet.

Dersom  $r > 1$  vil ethvert bidrag være positivt for *fellesskapet*, mens man først ved  $r > n$  vil få en situasjon der det å bidra er *individuell fordelaktig*<sup>5</sup>. Dersom  $1 < r < n$  vil altså  $\tau^2$  være en dominant strategi i engangsspillet (et krav for at dette skal være et Fangenes dilemma-spill), siden aktørene ikke unilateralt kan øke sin egen nytte ved å bidra<sup>6</sup>.

Taylor vurderer så det gjentatte spillet, der alle spillerne velger strategi  $B_\sigma$ , der  $\sigma = n - 1$  (altså et krav om at alle andre samarbeider for at aktør  $i$  skal samarbeide) (Taylor 1987:96). Taylor viser at  $i$ 's forventede nytte av universelt samarbeid (strategi  $\tau^1$  for alle aktører i alle delspill) er  $\frac{f(n-1)\delta_i}{(1-\delta_i)}$ , og at han kan oppnå  $\frac{g(n-1)\delta_i + g(0)\delta_i^2}{(1-\delta_i)}$  ved å bytte til  $\tau^2$  i alle delspill (Taylor 1987:85)<sup>7</sup>. For at det ikke skal være lønnsomt å bytte til strategi  $\tau^2$  må altså følgende ulikhet tilfredsstilles:

$$\frac{f(n-1)\delta_i}{(1-\delta_i)} \geq \frac{g(n-1)\delta_i + g(0)\delta_i^2}{(1-\delta_i)} \quad (5.3)$$

Ved noe omskriving får vi følgende krav til diskonteringsfaktor, dersom samarbeid skal kunne forventes, i dette  $n$ -person spillet:

$$\delta_i \geq \frac{g(n-1) - f(n-1)}{g(n-1) - g(0)} \quad (5.4)$$

---

<sup>5</sup>Jeg ser her på gruppens samlede midler. Siden det koster 1 enhet å bidra, og avkastningen på bidraget er  $1r$ , må  $r > 1$  for at gruppen som helhet skal tjene på at noen bidrar.

<sup>6</sup>Det er her verdt å merke seg at for aktører med sterk vektlegging av relativ gevinst vil ikke engang  $r > n$  nødvendigvis være nok, da samtlige spillere vil nyte godt av aktørens bidrag, mens kun bidragsyteren pådrar seg en kostnad. Han vil dermed – tross positiv gevinst – komme dårligere ut sammenlignet med de andre.

<sup>7</sup>Merk forøvrig likheten med gjennomgangen av et toperson-spill (fra Hovi (2008)) under premiss 1 i 4.2.2, ligning (4.1) og (4.2).  $\frac{f(n-1)\delta_i}{(1-\delta_i)}$  er tilnærmet  $\frac{R}{1-\delta}$  og  $\frac{g(n-1)\delta_i + g(0)\delta_i^2}{(1-\delta_i)}$  er tilnærmet  $T + \frac{\delta P}{1-\delta}$ .

Taylor (1987:97) viser så at man ved å erstatte  $f$  og  $g$  med verdiene fra (5.1) og (5.2) får:

$$\delta_i \geq \frac{n-r}{(n-1)r} \quad (5.5)$$

From this, the required minimum discount factor can be calculated for any values of  $[n]$ <sup>8</sup> and  $r$ . Note that if there is virtually total discounting (i.e.  $[\delta]_i \rightarrow 0$ ), this equilibrium condition is satisfied for  $r > [n]$ , that is,  $[n]$  units of public good benefit must be produced out of each unit of contributions. If there is virtually *no* discounting (i.e.  $[\delta]_i \rightarrow 1$ ), then the equilibrium condition is satisfied for  $r > 1$ , that is, for *any* value of  $r$ . This is just as we would expect. If very great public good benefit can be produced out of each unit contributed, it will pay conditional cooperators to produce the public good even if their discount rates are very great, and conversely if discounting is negligible, conditional cooperation will pay even if a unit of contribution only just produces more than a unit of public good benefit (Taylor 1987:97).

Taylor går videre og behandler situasjoner der strategiene innad i en populasjon varierer, og viser hvordan man kan oppnå samarbeid der tit-for-tat spilleres  $\sigma$  (i strategien  $B_\sigma$  – tit-for-tat) er lavere enn  $(n-1)$ <sup>9</sup>. Slike situasjoner er interessante; mulighetene for å oppnå samarbeid øker, da det ikke lenger kreves *universelt* samarbeid. Jeg vil i denne sammenheng nøye meg med behandlingen jeg har presentert ovenfor, men vil poengtere at det finnes gode innvendinger mot Taylors videre resonnement.

Dersom man vurderer “samarbeidsgrupper” (koalisjoner av størrelse  $m$ ) mindre enn populasjonen ( $n$ ), og som i resten av Taylors resonnement anser alle individer som uavhengige av hverandre og uten kommunikasjonsmuligheter, kan det oppstå et *koordinasjonsproblem*. I et slikt rammeverk ansees koalisjoner som svært løse, og alle individer kan i prinsippet inngå i hver koalisjon. La oss så, for eksempelets skyld, anta at en gruppe på 10 personer kan oppnå samarbeid, og at populasjonen består av 30 individer. I denne beskjedne populasjonen vil det finnes over 30 millioner mulige koa-

---

<sup>8</sup>Taylor avender  $a_i$  for diskonteringsfaktoren  $\delta_i$  og  $N$  for  $n$ , og jeg har her erstattet for å få samme notasjon som i min tekst.

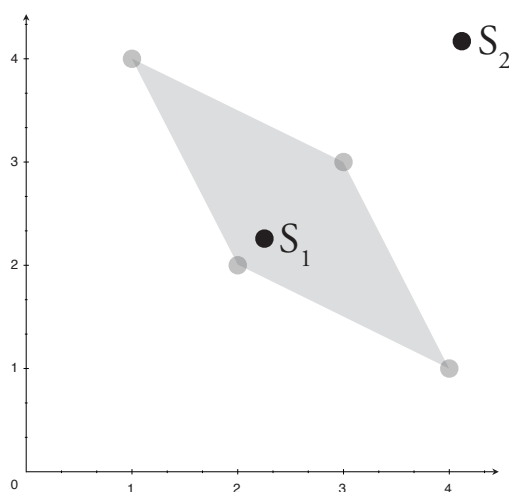
<sup>9</sup>Det er verdt å merke seg at at man med en populasjon med  $n > 2$  spillere som følger strategi  $B_{\sigma=n-1}$  like gjerne kunne sett på Grim Trigger. Brytes samarbeidet av én, vil det ikke kunne gjenopptas med mindre spillerne kommuniserer og forlater sin strategi minst én runde.



lisjoner à 10 personer<sup>10</sup>! Koordinasjonsproblemet består i at det synes lite sannsynlig at individene spontant ville være i stand til å enes om hvem som skal danne en slik koalisjon<sup>11</sup>.

Taylor hevder at selve *dilemmaet* i Fangenes dilemma ikke løses ved å se på gjentatte spill og betingede strategier, “men, det har blitt vist at *under visse forhold, kan det oppstå samarbeid, mellom noen eller alle spillere i det gjentatte spillet, uansett hvor mange spillere det er*” (Taylor 1987:104). Taylor har her vist en teoretisk mulighet. Gitt forutsetningene om 1) *n*-person spill, 2) egeninteresserte aktører (ifølge Taylor med *misunnelse* i Hobbes’ tilfelle) og 3) uendelig tidshorisont, er *alt mulig*. Samarbeid kan forekomme, og Taylor mener dette beviser at Hobbes tok feil: mennesker *kan* oppnå samarbeid uten staten, derfor er ikke staten *nødvendig*. Taylors løsning bygger på det såkalte *folketeoremet*<sup>12</sup>:

[...] When players are sufficiently patient, almost any feasible payoff allocation that gives each player at least his minimax security level can be realized in an equilibrium of the repeated game (Myerson 1991:331).



Figur 5.2: Utfallsrom ved Fangenes dilemma, og to hypotetiske statsløsninger,  $S_1$  og  $S_2$ .

<sup>10</sup>Ligningen  $\binom{n}{m} = \frac{n!}{m!(n-m)!}$  gir  $\frac{30!}{10!(30-10)!} = 30.045.015$ .

<sup>11</sup>Dette poenget skylder jeg Leif Helland.

<sup>12</sup>Myerson (1991:332) har følgende kommentar til navngivningen av dette teoremet: “We use here the phrase *general feasibility theorem* in place of “folk theorem”, because naming a theorem for the fact that it was once unpublished conveys no useful information to the uninitiated reader”.

At man sier seg enig i at den teoretiske muligheten for utenomstatlig samarbeid eksisterer, betyr allikevel ikke at man er nødt til å konkludere med at Hobbes tok feil. Hobbes er som jeg har påpekt svært opptatt av sikkerhet, og en viss *mulighet* alene for at samarbeid oppstår, kan ikke uten videre ansees som tilstrekkelig. Videre kan spontant samarbeid ansees som potensielt sårbart og ustabil.

I figur 5.2 vises utfallsrommet ved et Fangenes dilemma-spill (det skraverte området), og med folketeoremet kan vi slutte at alle løsninger her er mulige. Dersom man forstår Hobbes' stat som tilsvarende punkt  $S_1$ , viser Taylor at en slik løsning ikke er spesielt god, da man kan oppnå *mer* uten denne staten. Dersom man, som jeg mener Hobbes gjør, antar at hans stat er bedre enn enhver løsning uten en stat (f.eks.  $S_2$  i figuren), er det, tross Taylors resonnement, fullt mulig å hevde at staten er nødvendig. Dette innebærer at løsningen for Hobbes ligger *utenfor* det opprinnelige spillets rammer. Hvordan dette kan tenkes å foregå blir behandlet i 5.2.3.

Det må også spørres hvorvidt et resonnement basert på uendelig tidshorisont er egnet for å representere naturtilstanden. Livet er der kort og brutalt, og ved konflikt gir det lite mening å anta at denne konflikten kan vare evig. Det synes plausibelt å anta at det er en viss fare for å dø i naturtilstanden, og det synes også rimelig å forvente at én vinner og underlegger seg andre ved konflikt. Sett bort fra disse åpenbare problemene forbundet med Hobbes' filosofi, har mennesker også et tidsbegrenset liv, noe de er klar over. Det fremstår derfor som nyttig å introdusere muligheten for svært ufordelaktige brudd av spillet, men dette vil ikke forfølges videre i denne oppgaven. Det er dog nok en grunn til å ta visse forbehold når det gjelder relevansen av Taylors teoretiske mulighet i forbindelse med Hobbes' filosofi.

Det kan også være andre aspekter ved de samarbeidsformene som kan oppstå som gjør at denne teoretiske muligheten ikke nødvendigvis vil godtas av en Hobbesiansk tenker. Jeg vil i det følgende forlate Taylors formelle argument, da jeg finner det interessant å vurdere den typen utenomstatlig samarbeid Taylor selv fremmer som mest plausibel (kommunitarianske samfunn) opp mot hva Hobbes faktisk skriver om naturtilstanden og mulighetene for å unnslippe den.

[...] in the absence of the state social order can be maintained only if the relations between people are those characteristic of *community* (Taylor 1982:2).

## 5.2.2 Frivillig samarbeid, kommunitarianske samfunn og normer

Taylor ønsker å demonstrere hvordan mennesker *kan* finne grunnlag for å samarbeide uten tvang eller en statsmakt. Denne muligheten aksepterer jeg fullt ut, men jeg kan ikke nøye meg med dette i vurderingen av premiss 1 i min versjon av Hobbes' resonnement; det er nødvendig å se kort på hva denne muligheten for samarbeid kan innebære, og hvordan samarbeidsformen vil kunne tenkes å arte seg.

### Forskjellige løsninger på samhandlingsproblemer

Fremfor alt er det nødvendig å påpeke at det her er snakk om en *hypotetisk* mulighet for samarbeid. Gitt *visse forhold* i *visse modeller* er det grunn til å forvente samarbeid, men det er ingen garantier for at disse modellene representerer verden på en god nok måte. Intuitivt er det imidlertid svært enkelt for mennesker med erfaring fra det virkelige liv å godta at samarbeid faktisk forekommer i svært mange tilfeller, selv der mennesker tilsynelatende har svært gode muligheter for å utnytte hverandre.

Taylor definerer "samarbeid" ( $\tau^1$ ) i forbindelse med Hobbes som å "legge til side ens naturretigheter", mens han i den generelle gjennomgangen av samhandlingsproblemer definerer C som å handle med restriksjoner. Han mener selv at førstnevnte definisjon kun er en omformulering av den sistnevnte (Taylor 1987:131). Etter min mening innebærer denne fremstillingen en *for* voldsom forenkling, som ødelegger noe for plausibiliteten til Taylors modellering av Hobbes.

Det er stor forskjell på midlertidig å velge ikke å benytte seg av ens naturlige rettigheter i enhver situasjon, og det å permanent frasi seg disse rettigheten og overføre de til suverenen eller en annen. Jeg antar at Taylors modellering av Hobbes' naturtilstand og stat som et tostrategi-spill er årsaken til denne forenklingen. Det er min oppfatning at dette fører til en forvrengning av Hobbes' filosofi, og jeg vil i 5.2.3 kort utforske en alternativ fremgangsmåte.

Heretter vil jeg anvende begrepet "samarbeidsvilje" (det å følge strategi  $\tau^1$ ) som *menneskers vilje til å legge bånd på seg selv med sikte på at det skal være mulig å opprettholde en grad av ro og orden i en populasjon*. Dette innebærer at samtlige individer i utgangspunktet beholder sin rett til alt og alle, men at de altså velger å ikke benytte seg av

disse rettighetene der dette vil ødelegge den grad av ro og orden det hersker enighet om at er nødvendig. Dette er som nevnt i strid med Taylors begrepsbruk der han diskuterer Hobbes, men det gir i mine øyne en bedre mulighet for å utforske spontant samarbeid utenfor en stat. Et slikt samarbeid kan betegnes som *frivillig samarbeid*, ved at mennesker til enhver tid har rett til å bryte ut av samarbeidet. Taylors forståelse av samarbeid hos Hobbes er *ikke* det samme, da Taylor ser for seg at *Hobbes' individer former staten når alle velger  $\tau^1$*  – en tilstand der alle rettigheter overføres til suverenen. I det følgende vil strategi  $\tau^1$  innebære å handle med visse restriksjoner, mens  $\tau^2$  innebærer å se bort fra alle restriksjoner. Dette er dermed i tråd med Taylors strategier der han ikke diskuterer Hobbes spesielt.

Når det gjelder den videre diskusjonen av frivillig samarbeid velger jeg å se spesielt på én form for slikt samarbeid, og slutter meg i valget av denne formen til Taylors vurderinger i *Community, Anarchy and Liberty* (Taylor 1982):

Broadly speaking, three 'methods' have been proposed for ensuring the provision of the public good of social order. In practice the methods are found in various combinations. To characterise them briefly, they are: the state, the market and the community. [...] social order without the state can be maintained only in community. I shall not take very seriously the claim that the *market* is an alternative to the state as a sole means of ensuring the provision of social order. The less radical claim, made by many laissez-faire liberals, that *some* of the goods whose provision constitutes social order can be put on the market, has, I think, some validity. But the full claim, made especially in the United States by people calling themselves 'anarcho-capitalists' or 'libertarians', that *all* the components of social order could be satisfactorily put on the market is in my view indefensible (Taylor 1982:59).

[...] there are what we might call core characteristics of community, attributes possessed in *some* degree by *all* communities, and among these are the requirements that relations between members of a community be direct and many-sided and that they practise certain forms of reciprocity (Taylor 1982:3).

Det synes mest nærliggende å se på den utenomstatlige samarbeidsformen som fremstår som det mest realistiske alternativet til staten, og jeg velger derfor i det følgende

å se på *kommunitarianske samfunn*. Mens det finnes en teoretisk mulighet for frivillig samarbeid i uendelige populasjoner (det Taylor kaller en *markedsløsning* på samarbeidsproblemet) er jeg enig med Taylor i at små kommunitarianske samfunn fremstår som et langt mer plausibelt alternativ (Taylor 1982:2,59). Årsaken til at Taylor forkaster frivillig samarbeid i store populasjoner er hovedsakelig det faktum at de fleste mekanismer som muliggjør et slikt samarbeid krever små populasjoner:

Much voluntary cooperation in the provision of public goods is no doubt of this conditional kind. Unfortunately, it too is less likely to occur in large groups than in small ones, since a conditional cooperator must be able to monitor the behaviour of others in the group so as to reassure himself that they are doing their parts and not taking advantage of him. Clearly, as the size of the group increases, the mutual monitoring becomes increasingly difficult and the ‘tacit contract’ of conditional cooperation becomes increasingly fragile. In a relatively small group, on the other hand, especially one with an unchanging or only very slowly changing membership, people come into contact with and can observe the behaviour of most of their fellows, so that conditional cooperation is more likely to be workable. But even here, unless the group is *very* small, the arrangement is obviously a precarious one if it has no other supports (Taylor 1982:53).

If community is characterised by shared values and beliefs, direct and many-sided relations, and the practice of reciprocity, then it is clear that communities must be relatively small and stable (Taylor 1982:32).

Det er viktig å merke seg at jeg omtaler dette som *frivillig samarbeid*, da samarbeidet ikke forutsettes å være institusjonalisert, men spontant dannede (men stabile) grupper der medlemmene velger å samarbeide uten ekstern tvang. Slike “samfunn”, som jeg i det følgende vil omtale som *kommunitarianske* for å skille fra dagligdags bruk av ordet samfunn, er for Taylor en *løsning* på samhandlingsproblemet. Det er med hans terminologi en *ekstern* løsning, ved at selve spillet endres og nye “muligheter, holdninger og antagelser” introduseres for aktørene<sup>13</sup>. Videre er slike *samfunn* en *desentralisert* ekstern løsning, der en *sentral* ekstern løsning ville være en stat (Taylor 1987:21-3). Alternative løsninger som innebærer en statsdannelse vil bli behandlet videre i forbindelse med Hobbes’ absolutisme i 5.3.

---

<sup>13</sup>Taylor diskuterer også *interne* løsninger på slike samhandlingsproblemer, og disse utgjør det han kaller “markedsløsninger” på samhandlingsproblemet, som jeg ikke forfølger i denne oppgaven.

The members of such a group of people, or all of its active adult members, can wield with great effectiveness a range of positive and negative sanctions, including the sanctions of approval and disapproval – the latter especially via gossip, ridicule and shaming. Decentralized solutions can sometimes be effective where there is little community, but the size of the group would still have to be relatively small (as it must be in a community) (Taylor 1987:23).

Jeg har ikke her anledning til å foreta en full gjennomgang av alle former slike samfunn kunne tenkes å anta. Jeg vil nøye meg med det som blir sentralt for mitt prosjekt: å vurdere hvilke negative aspekter ved kommunitarianske samfunn som vil kunne ha innvirkning på Hobbes' syn på dette som et reelt alternativ til staten.

Denne diskusjonen vil være generell – altså ikke basert på Taylor (1982; 1987) alene – og jeg vil derfor også trekke på innsikt om temaet fra eksperimentell økonomi. Elementer av dette er diskutert kort i 3.4.3. På generelt grunnlag mener jeg brorparten av funnene fra eksperimentell økonomi både underbygger og forsterker Taylors generelle hypotese om at frivillig samarbeid er mulig. Sterk gjensidighet og/eller ulikhetsaversjon kan f.eks. *omforme* spill fra fangens dilemma-spill til andre slags spill der det er mindre problematisk å se for seg utstrakt samarbeid.

Sett at en aktør  $i$  har nyttefunksjonen (3.2) fra 3.4.3:  $U_i(X) = x_i - \alpha_i \max(x_j - x_i, 0) - \beta_j \max(x_i - x_j, 0)$  med en  $\alpha = 0.5$  og  $\beta = 0.5$ . Selv om denne modellen er rent hypotetisk, er det interessant å se hva slags effekter omformingen av spill kan ha. I figur 5.3 kan man se hvordan et Fangenes dilemma-spill med hensyn til objektive gevinster, kan omformes til et fundamentalt forskjellig spill når nyttefunksjonen for ulikhetsaversjon her anvendes. I spillmatrisen til høyre har vi nå et Moderat selvhevdelse-spill, der  $\tau^2$  ikke lenger er en dominant strategi, og man har to likevekter ( $(\tau^1, \tau^1)$  og  $(\tau^2, \tau^2)$ ). Da  $(\tau^1, \tau^1)$  er strengt foretrukket fremfor  $(\tau^2, \tau^2)$  er det plausibelt å anta at det Pareto-optimale utfallet velges (Taylor 1987:19).

Taylor er klar over at virkelige mennesker ikke er som *Homo economicus*, men velger å utføre brorparten av sin analyse av temaet *som om* de var det. Som han selv sier i innledningen til boken: “jeg er faktisk av den oppfatning at dette [oppfatningen av mennesket som “ren egoist”] er en god tilnærming til faktisk adferd i svært mange situasjoner der kollektiv handling er en mulighet” (Taylor 1987:109).

In the next three chapters, the payoffs are assumed *not* to reflect, *inter alia*,

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^1$	3, 3	1, 4
	$\tau_i^2$	4, 1	2, 2

Gevinster

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^1$	3, 3	-0.5, 2.5
	$\tau_i^2$	2.5, -0.5	2, 2

Nytteverdier

Figur 5.3: Ovenfor vises en Fangenes dilemma-struktur der verdiene er gitt i gevinster, og til høyre det samme spillet med nytteverdier utledet fra en nyttefunksjon (3.2) med ulikhetsaversjon og  $\beta = \alpha = 0.5$ .

- (i) incentives due to external coercion, including that applied or threatened by the state or any other external agency or by other members of the group (apart from the tacit threats and offers which may be thought to be embedded in conditional Cooperation – about which more later); (ii) altruistic motivation; and (iii) and ‘internal sanctions’ like guilt, loss of self-respect and so on, which may result from failure to conform to a norm, live up to one’s own ideals, perform one’s duties, or whatever (Taylor 1987:17).

Denne utelatelsen synes unødvendig i min diskusjon av muligheten for samarbeid. Ved å inkludere disse mulighetene for bredere menneskelig motivasjon og interaksjon – som det gjøres i eksperimentell økonomi – styrkes argumentene for samarbeid som en mulighet. I mine øyne har jeg dermed gitt Taylors innvending mot nødvendigheten av staten *størst mulig tyngde*, for å forsikre meg om at jeg ikke godtar Hobbes’ premiss via en stråmanns-argumentasjon i forhold til Taylor. En desto viktigere grunn til at de elementene Taylor utelater bør inkluderes, er at jeg mener å ha vist at Hobbes faktisk tillegger mennesker sosiale preferanser.

### Ulemper ved kommunitarianske samfunn

Mye kan sies om de attraktive sidene ved små, nære og trygge kommunitarianske samfunn. Jeg er her ikke interessert i å kartlegge de positive sidene ved slike samfunn, men finner det nødvendig å se på hvilke negative sider som finnes. Dette fordi jeg ønsker å se hva som taler *imot* å anse dette som en fullgod løsning på samhandlingsproblemet;

dersom det utelukkende er grunner *for* en slik løsning, vil behovet for staten være fraværende, men jeg anser altså ikke saken for å være så enkel.

Et spørsmål som melder seg i forbindelse med kommunitarianske normstyrte samfunn er hvilke *mål* som der kan forfølges. Jeg har i 4.1.1 sett kort på hva jeg regner som Hobbes' (og forsåvidt også andre liberaleres) idealer i denne sammenheng: *mest mulig nøytralitet*. Hvorfor samfunn dannes, og hvilke mål som deretter blir forfulgt, er et meget viktig spørsmål. *Pluralisme* er relatert til hvorvidt staten har et formål, og er et sentralt liberalt begrep. Jeg mener det er sterke argumenter for å hevde at kommunitarianske samfunn som oftest er mer ensrettede enn liberale samfunn, noe som strider med ønsket om legge til rette for pluralisme. Taylor er også oppmerksom på dette problemet, men peker på at det *finnes* visse kommunitarianske samfunn der man forsøker å unngå ensretting (Taylor 1982:164).

Det faktum at frivillig samarbeid ikke er institusjonalisert, fører til at slike samfunn må anees som *normstyrte*. Normer kan, under visse forutsetninger, være vel så sterke som lover, men det er en sentral forskjell: *det er ingen muligheter for å sikre at normer underbygger ønskelige trekk i mennesker eller samfunnet*. "Kommunitarianske samfunn virker fordi de er gode på å håndheve normer, og hvorvidt dette er en fordel avhenger av hva disse normene er" (Bowles og Gintis 2005:388).

Videre, uansett hvor vellykket et kommunitariansk samfunn kan synes fra innsiden, vil det alltid være et spørsmål om *hvem som utgjør samfunnet*. Majoriteten vil i de aller fleste tilfeller ha svært stor innflytelse på hvordan samfunnet fungerer, og det vil være dårlig vern mot majoritetstyranni og sosial eksklusjon. Mens likhet for loven og fastsatte regler for hvem som utgjør staten er prinsipper som danner grunnlaget for en liberal stat, har man altså betydelig risiko for undertrykkelse og eksklusjon av f.eks. etniske, politiske og holdningsmessige minoriteter i et slikt kommunitariansk samfunn:

The baggage of belonging, however, often includes poor treatment of those who do not. [...] When insider-outsider distinctions are made on divisive and morally repugnant bases such as race, religion, nationality, or sex, community governance may contribute more to fostering parochial narrow-mindedness and ethnic hostility than to addressing the failures of markets and states (Bowles og Gintis 2005:388).

*Altruistisk straff* ble diskutert relativt inngående i 3.4.3. Mens Taylor legger til grunn at tit-for-tat spillere kan "straffe" overtredere ved å bryte samarbeidet, er denne stra-



tegien – og “Penance”<sup>14</sup> – en noe “grov” strategi som innebærer at man nødvendigvis straffer *alle* dersom noen begår overtramp man ikke er villig til å tolerere. Fordelen med teorien om altruistisk straff er at slike sanksjoner er *selektive*. I små og nære samfunn er det god grunn til å anta at man har muligheten til å identifisere i alle fall noen av dem som utfører handlinger man mener må straffes. I slike situasjoner synes det mer effektivt å straffe kun overtredene, mens resten av gruppen kan fortsette å samarbeide. Denne teorien er altså en styrkning av Taylors resonnement i de tilfeller der det finnes tilstrekkelig informasjon, sanksjoner som kan individualiseres, en mulighet til å identifisere overtredere og muligheten til å kommunisere sin vilje til å fortsette samarbeidet til andre som er innstilt på det samme.

“Medaljen” har dog en bakside – *asosial straff* (antisocial punishment):

People might punish not only freeloaders, but cooperators too. For example, participants who had been punished in the past for contributing too little might retaliate against the cooperators because the cooperators are precisely those individuals most likely to punish the free-riding low contributors (Herrmann et al. 2008:1362).

At noen straffer individer som faktisk oppfører seg slik man generelt sett ønsker – og er avhengig av for å oppnå frivillig samarbeid – viser seg å være svært destruktivt. Det er påvist at individer kan oppnå svært høy grad av samarbeid dersom de har sanksjonsmuligheter, men det viser seg altså at asosial straff under visse forhold kan ødelegge den fordel man oppnår med disse sanksjonene (Herrmann et al. 2008:1362).

But what explains antisocial punishment? One plausible reason is that people might not accept punishment and therefore seek revenge. Revenge is a “human universal” and part of a culture of honor in many societies (Herrmann et al. 2008:1363)

Det faktum at staten ikke har monopol på utmåling og gjennomføring av *straff* kan altså by på visse problemer. Herrmann et al. (2008:1364) viser en sterk negativ sammenheng mellom forekomsten av asosial straff og oppnådd samarbeid; forskningssubjekter fra Boston og København utmerker seg som svært samarbeidsvillige og lite asosiale i sin

---

<sup>14</sup>*Penance* innebærer at “en spiller som avviker ved å velge Bryte, og som deretter vender tilbake til Penance, må gjøre opp for seg ved å være alene om å velge Holde i neste periode. Hvis den som bryter avtalen følger denne oppskriften, vil straffen vare i bare én periode, og partene vil deretter gjenopprette samarbeidet” (Hovi 2008:88). Strategien er dermed mer sofistiskert enn tit-for-tat, og gir bedre muligheter for å gjenopprette samarbeid ved avvik enn tit-for-tat.

sanksjonsvirksomhet, mens subjekter fra Aten og Muscat er de “verste” når det gjelder begge. Det er forøvrig svært enkelt å se for seg at samarbeid er vanskelig å opprettholde der de som faktisk forsøker å bidra til fellesskapet, opplever å bli *straffet* av dem som lar være. Frivillig samarbeid synes å være svært ømfintlig, spesielt siden alle til enhver tid er i besittelse av retten til å bryte samarbeidet. Det virker lite sannsynlig at man oppnår særlig høy grad av slikt samarbeid der man opplever denne typen adferd, med de uvennlige følelsene som med stor sannsynlighet følger slike interaksjoner.

Forfatterne selv peker på at fraværet av en rettsstat kan tenkes å påvirke forekomsten av asosial straff, noe som gjør dette fenomenet spesielt relevant i forbindelse med kommunitarianske samfunn<sup>15</sup>:

The strength of the rule of law in a society might also have an impact on antisocial punishment. If the rule of law is strong, people trust the law enforcement institutions, which are perceived as being effective, fair, impartial, and bound by the law (Herrmann et al. 2008:1363).

Resultatene av deres analyser bekrefter rettsstatens betydning, da personer fra samfunn der rettsstaten oppfattes som svak, er signifikant mer tilbøyelige til å utføre asosial sanksjonsvirksomhet. Det synes plausibelt å anta at man vil kunne oppleve mer asosial straff og hevn – privat rettfærdighet om man vil – i samfunn der en rettsstat ikke eksisterer overhodet. Dette er selvsagt noe avhengig av hvor effektivt samfunnet håndhever normer, samt hvor enkle disse normene er å forholde seg til og forstå. Det synes forøvrig lite problematisk å anta at man i et samfunn uten rettsvesen vil være noe mer avhengig av å håndheve sin egen rettfærdighet:

[. . .] if there be no power erected, or not great enough for our security; every man will, and may lawfully rely on his own strength and art, for caution against all other men (Hobbes 1946:109).

Til sist vil jeg gå inn på det som synes å være et universelt syn på kommunitarianske samfunn: *de er avhengige av å være små*. Dette kan være et alvorlig problem der man har en tett befolkning, da selv svært harmoniske grupper *innad*, kan tenkes å være langt mer konkurransepregede i forhold til *andre* samfunn. Man kan se for seg stammer – kommunitarianske samfunn – i Amazonas-jungelen; disse vil kunne fungere forholdsvis godt, sett bort fra en mangel på avansert vitenskap, lite frihet til å være annerledes

---

<sup>15</sup>Taylor (1982:81) understreker også at det han kaller “self-help justice” er en viktig kilde til “sosial kontroll” i kommunitarianske samfunn.

osv. Dersom man derimot forsøker å se for seg kommunitarianske “stammer” som et alternativ til staten f.eks. i befolkningstette deler av USA, synes det langt mindre sannsynlig at dette ville kunne fungere over tid; der mennesker lever tett på hverandre får man helt andre krav til håndtering av felles nyttige knappe ressurser, man vil ikke lenger ha muligheten til å utfolde seg mer eller mindre som man vil, og man vil kunne se hvordan andre samfunn styres, noe som kan skape misnøye og uro.

And in all places, where men have lived by small families, to rob and spoil one another, has been a trade, and so far from being reputed against the law of nature, that the greater spoils they gained, the greater was their honour; and men observed no other laws therein, but the laws of honour; that is, to abstain from cruelty, leaving to men their lives, and instruments of husbandry (Hobbes 1946:109).

Også dette innrømmer Taylor, og det leder ham til en meget pessimistisk konklusjon når det gjelder hvorvidt kommunitarianske samfunn (som ifølge Taylor er det eneste relle alternativet til en stat) kan ansees som et plausibelt alternativ til staten i dagens tett befolkede verden:

So the egalitarian anarchic community, though it can last for millennia if left alone, is terribly vulnerable to other states (Taylor 1982:130).

That this reaction should have been so common is little to be wondered at; but insofar as any lessons for the future can be drawn from this account of how anarchic communities first gave way to (primary or secondary) states, we can say that it inspires little optimism about the viability of anarchy [som for Taylor innebærer kommunitarianske samfunn] in a crowded world and even less optimism about the prospects for the emergence and durability of an anarchy set in a sea of states (Taylor 1982:139).

Kommunitarianske samfunn fremstår som en plausibel teori om hvordan mennesker kan tenkes å oppnå en viss grad av samarbeid uten en stat. Jeg har imidlertid påpekt noen mulige problemer med slike samfunn, som bl.a. ensretting, normer fremfor lover, manglende kontroll over samfunnets retning, majoritetstyranni og sosial eksklusjon, og til sist – men kanskje viktigst – forventede problemer der mange små samfunn eksisterer i umiddelbar nærhet av hverandre.

Én tradisjonell tolkning av Hobbes’ rettferdiggjøring av staten er at han antas å ha

trodd at mennesker uten en stat ikke er i stand til å samarbeide *overhodet*. Jeg har imidlertid et noe annet syn på annet syn på Hobbes' tenkning om dette, og mener han ser for seg at en viss grad av samarbeid *vil* forekomme i naturtilstanden, men at dette er av en slik art – og er såpass ustabil – at staten allikevel er en nødvendighet. Jeg må derfor se hvorvidt de problemene jeg har lagt frem i forbindelse med kommunitarianske samfunn omtales av Hobbes, og hvorvidt det er plausibelt å anta at han ville forkastet frivillig samarbeid i kommunitarianske samfunn som utilstrekkelig.

### 5.2.3 Hobbes og muligheten for utenomstatlige løsninger på naturtilstandens tragedie

Før jeg tar standpunkt til hvorvidt premiss 1 (påstanden om at staten er nødvendig for å oppnå fred) kan sies å svekkes av muligheten for utenomstatlig samarbeid, ønsker jeg å kommentere visse aspekter ved Taylors tolkning av Hobbes.

#### Fremstillingen av Hobbes' filosofi i enkel spillform

Først og fremst vil jeg her forholde meg til Taylors formalisering av Hobbes' filosofi, som jeg har noe vanskeligheter med å godta. Jeg hevder som nevnt i kapittel 4, under premiss 2, at Hobbes' naturtilstand muligens bør fremstilles som et Chicken-spill, men uoverensstemmelsen med Taylor går dypere enn som så; jeg ønsker ikke å modellere *både staten og naturtilstanden* i ett spill med kun to rene strategier, slik Taylor gjør.

Taylor anvender to modeller for Hobbes, én statisk og én dynamisk – begge Fangenes dilemma-spill (Taylor 1987:129). Jeg vil ikke gå inn på forskjellene mellom disse to modellene, men bemerker at Taylor selv ikke var fornøyd med noen av dem (Taylor 1987:138). Taylor hevder faktisk også at “Hobbes i realiteten kun behandler et engangsspill med Fangenes dilemma-struktur [...]” (Taylor 1987:33). Jeg sier meg enig med Taylor i at ingen av hans modeller er spesielt tro mot Hobbes' teori, for det første pga. hvordan spillet er formalisert:

One of the outcomes is called ‘the state of war’, or simply ‘War’ (*Lev* 96). This is the state of affairs which obtains when every individual seeks, in the absence of restraint, to maximize his utility (as Hobbes assumes he does). [...] Let us suppose that this strategy is *D* (Taylor 1987:130).

When men are not in the state of War, then there is ‘Peace’, says Hobbes. There are of course *three* outcomes other than War (assuming that there are only two strategies available to each player), but it is clear that Hobbes means that Peace corresponds to only one of these outcomes: it obtains when *no* individual chooses strategy *D*. [...] Thus, Peace obtains when every individual chooses strategy *C* (Taylor 1987:130).

*Fred* er, som vi har sett, definert som fraværet av *krig*, og krig innebærer for Hobbes en tilstrekkelig og kjent tilbøyelighet til kamp – ikke utelukkende faktisk fysisk kamp (Hobbes 1946:82).

Det problematiske ved Taylors fremgangsmåte er at staten *og* naturtilstanden her modelleres i samme spill, der de to tilgjengelige rene strategiene innebærer å beholde sine naturlige rettigheter (*D*) eller frasi seg dem (*C*); “vi kan nå si at strategi *C* innebærer å ‘legge til side ens naturrettigheter” (Taylor 1987:131). Det er tydelig at Taylor her snakker om å *frasi seg* sine rettigheter – slik Hobbes ser for seg ved dannelsen av staten – og ikke det frivillige valg midlertidig å begrense seg.

		$k_j$	
		$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^1$	<i>Fred</i>	1, 4
	$\tau_i^2$	4, 1	<i>Krig</i>

Taylor

		$k_j$		
		$\tau_j^0$	$\tau_j^1$	$\tau_j^2$
$k_i$	$\tau_i^0$	<i>Fred</i>	–	–
	$\tau_i^1$	–	<i>Modus Vivendi</i>	<i>T, H</i>
	$\tau_i^2$	–	<i>H, T</i>	<i>Krig</i>

Alternativ

Figur 5.4: Taylors fremstilling av Hobbes’ filosofi, og et noe mer nyansert alternativ.

I figur 5.4 ser vi Taylors modell til venstre. Et spørsmål tvinger seg frem: *Hvordan kan man si noe om behovet for å forlate naturtilstanden med denne modellen?* Da modellen ikke tillater *frivillig samarbeid* (siden Taylor forventer at staten dannes ved universelt samarbeid), kan man da si noe om mulighetene for dette med denne modellen? Frivillig samarbeid innebærer at man beholder sine naturrettigheter, men velger å begrense seg, og dette er det faktisk ingen mulighet i Taylors formalisering av Hobbes.

Til høyre i figur 5.4 illustrerer jeg hvordan jeg ser for meg at en slik modell bør være. Strategiparet  $(\tau^1, \tau^1)$  (begge samarbeider) er her ikke en overgang fra naturtilstanden til staten, men det jeg oppfatter som Hobbes' syn på naturtilstanden: ikke kamp, men en skjør "fred" der man hele tiden lever i fare for at kamp bryter ut (Hobbes 1946:82). Dersom man vil modellere naturtilstanden mener jeg altså at man kun behøver å utelukke strategi  $\tau^0$ , som er dannelsen av staten. Strategiene i spillet kan her forstås som å handle med eller uten restriksjoner.

Hovedproblemet er at Taylors spill har et utfall som "fanger", og dermed endrer spillet:  $(\tau^1, \tau^1)$ . Dersom alle faktisk velger  $\tau^1$  i én periode hos Taylor innebærer dette at staten formes, og da blir det meningsløst å snakke om at dette ikke er en likevekt fordi alle kan tjene på å bytte til strategi  $\tau^2$ . Poenget hos Hobbes er jo nettopp det at når staten faktisk etableres, vil lover og sanksjoner inntre, og det vil ikke kunne lønne seg å "hoppe av" etter dette.

Jeg foreslår å modellere interaksjoner hos Hobbes som *to* spill (se figur 5.5), der  $\Gamma^1$  (spill 1) representerer naturtilstanden og  $\Gamma^2$  er staten. Det er overgangen fra  $\Gamma^1$  til  $\Gamma^2$  som er interessant, og jeg mener  $\Gamma^1$  må forstås som et spill med en "utgangsstrategi"  $\tau^0$ , som endrer spillet til  $\Gamma^2$ .

		$k_j$				$k_j$	
		$\tau_j^1$	$\tau_j^2$			$\tau_j^1$	$\tau_j^2$
	$\tau_i^1$	3, 3	2, 4		$\tau_i^1$	2, 2	2, 1
$k_i$	$\tau_i^2$	4, 2	1, 1		$\tau_i^2$	1, 2	1, 1
		$\Gamma^1$				$\Gamma^2$	

Figur 5.5: En todelt fremstilling av Hobbes' filosofi, der  $\Gamma^1$  representerer naturtilstanden og  $\Gamma^2$  statstilstanden. Preferansene er her ordinale, og tallene i  $\Gamma^1$  og  $\Gamma^2$  kan ikke sammenlignes.

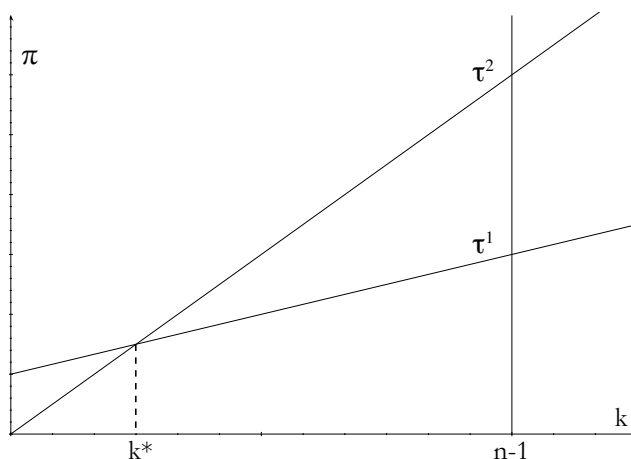
$\Gamma^1$  og  $\Gamma^2$  er statiske spill, der "dynamikken" ligger i at man kan flytte seg mellom spillene. Det er interessant å se hvordan f.eks. muligheten for statens sammenbrudd kan inkluderes i en slik fremgangsmåte, men en videre analyse av denne fremgangsmåten vil ikke forfølges videre i denne oppgaven.

Hvordan overgangen fra  $\Gamma^1$  til  $\Gamma^2$  skal foregå er åpent for vurdering, men for Hobbes er statsdannelsen alltid en mulighet (den eneste) der mennesker som søker seg vekk fra naturtilstanden velger å følge fornuften.

### Er naturtilstanden et Chicken- eller et Fangenes dilemma-spill?

Som vist til venstre i figur 5.5, og diskutert under premiss 2 i 4.2.2, fremstiller jeg Hobbes' naturtilstand som et Chicken-spill. Denne tolkningen er, som Taylor har påpekt, mer kontroversiell enn hans Fangenes dilemma-fremstilling. Det kan synes som om de forskjellige oppfatningene om hvordan Hobbes' politiske filosofi bør modelleres også er en del av årsaken til de forskjellige oppfatningene om hvilken preferansestruktur som best representerer Hobbes' individer i naturtilstanden.

Først og fremst har valget av spill forholdsvis lite å si for testingen av behovet for en stat (premiss 1), da jeg allerede har godtatt at det selv i et Fangenes dilemma-spill er muligheter for samarbeid.



Figur 5.6:  $n$ -person Chicken-spill

I figur 5.6 ser man et  $n$ -person Chicken-spill, og forskjellen fra 5.1 er tydelig: det er ingen dominant strategi i dette spillet. Taylor diskuterer også muligheten for å anvende Chicken:

The important feature of this game [Chicken] is that there are two equilibria and in each of these one player cooperates while the other has a 'free ride' on the public good provided out of his contribution, so that it will pay

each player to attempt to be the first to bind himself irrevocably to non-cooperation, if this is feasible. *This pre-commitment strategy ‘forces’ the other player into cooperation.* However, where each player is able to bind himself in this way both may realize the dangers of simultaneous binding and may forgo this for the fully cooperative outcome. This is, unfortunately, unstable (at least in the single play game) so that one can expect any cooperation to be fragile (Taylor 1987:39).

This analysis of pre-commitment as a risky decision gives us a further reason for believing that some cooperation will be forthcoming in chicken games, and hence that in those public goods interactions for which chicken is the most appropriate model *some* amount of the public good will be provided (Taylor 1987:47).

Forskjellen ligger altså i hovedsak i at man kan forvente *mer* samarbeid i et Chicken-spill enn i Fangenes dilemma, så om noe kan det synes som om jeg *svekker* Hobbes’ argument for staten ved denne tolkningen av naturtilstanden. Det er imidlertid interessant å undersøke nærmere hva som taler for denne tolkningen.

Først og fremst dreier det seg for meg om å legge vekt på usikkerhetsaversjon og døden/fare for døden som det verste ondet. For Taylor er derimot det å *blottlegge seg* det sentrale (Taylor 1987:131). Hobbes’ definisjon av *en naturlov* er her viktig:

A LAW OF NATURE, *lex naturalis*, is a precept or general rule, found out by reason, by which a man is forbidden to do that, which is destructive of his life, or taketh away the means of preserving the same; and to omit that, by which he thinketh it may be best preserved (Hobbes 1946:84).

Taylor vektlegger sterkt dette forbudet mot å gjøre det som “fjerner hans muligheter” for å sikre sitt eget liv; det verste ondet blir å frasi seg sine naturrettigheter dersom andre ikke gjør det, og dette blir resultatet for  $i$  ved  $(\tau_i^1, \tau_j^2)$  i Taylor sitt spill. Det er lett å se poenget, men det er for meg fremdeles uklart om det på noen måte kan være bedre for  $i$  å velge strategi  $\tau^2$ , f.eks. der man vet at ens motspiller er ondskapsfull, kjent for å myrde sine motstandere, og helt klart er sterkere enn en selv. For spiller  $i$  å velge  $\tau^2$  i en slik situasjon kan tolkes som å være “destruktivt for [ $i$ ’s] liv”, det ville fjerne hans muligheter for å bevare sitt liv og det ville være *å la være å gjøre det han tror gir best mulighet for å overleve*. Dette er selvsagt i et hypotetisk møte med én spesiell motstander, men jeg mener fremdeles det kan forsvares å anse  $(\tau^2, \tau^2)$



som svært ufordelaktig, selv i Taylors spill. Kavka (1986:111) hevder at det kun dreier seg om “små tap (f.eks. skade eller ressursforbruk)” der begge partene velger  $\tau^2$ , noe som for meg virker noe optimistisk. Det må imidlertid være slike vurderinger av utfall ( $\tau^2, \tau^2$ ) som fører til fremstillingen av naturtilstanden som et Fangenes dilemma-spill, med mindre man anser underkastelse som et *ekstremt* dårlig utfall, noe Hobbes ikke ser ut til å gjøre. Hobbes forklarer selv hvordan en aversjon mot risikabel kamp kan lede mennesker til underkastelse:

The second is, if a man taken prisoner in the wars, or overcome, or else distrusting his own forces, to avoid death, promises the conqueror or the stronger party his *service*, that is, to do all whatsoever he shall command him. [...] Now he that is thus tied, is called a *servant*; he to whom he is tied, a *lord* [...] There is therefore a confidence and trust which accompanies the benefit of pardoned life, whereby the *lord* affords him his corporal liberty (Hobbes 2005a:109).

Videre fremmer han et syn på slik underkastelse som vanskelig kan tolkes til at han anser det å være tjener som noe forferdelig:

For my part therefore I cannot find what reason a mere *servant* hath to make complaints, if they relate only to want of *liberty*; unless he count it a misery to be restrained from hurting himself, and to receive that life, which by war, or misfortune, or through his own idleness was forfeited, together with all manner of sustenance, and all things necessary to the conservation of health, on this condition only, that he will be ruled (Hobbes 2005a:121).

Det er forøvrig verdt å merke seg at det å underkaste seg suverenen innebærer samme “svake stilling” som det å underkaste seg en annen herre. Hobbes er imidlertid klar på at suverenen tilbyr langt bedre sikkerhet enn en herre i naturtilstanden, men det er et viktig poeng at Hobbes ikke anser denne typen underkastelse som forferdelig. I “mitt” spill er  $(\tau_i^1, \tau_j^2)$  bedre for *i* enn det var hos Taylor, da *i* her ikke gir fra seg sine naturrettigheter uten å vite hva som vil skje. Han underordner seg ved et slikt utfall sin overmann; det å velge strategi  $\tau^1$  her er altså ikke å gi avkall på alt til alle andre – ved  $(\tau_i^1, \tau_j^1)$  gir *ingen* fra seg sine naturrettigheter. Han gir avkall på det hans nye herre krever – og dette kravet kan ikke innebære *livet*<sup>16</sup>. Det synes plausibelt å anta at

---

<sup>16</sup>“But then the enemy, who has reached his goal, the safeguarding of the recognition of his superiority, of his honour, cannot, for honour’s sake, take his life; for ‘nothing but fear can justify the

$(\tau_i^1, \tau_j^2)$  fører til et “herre og tjener”-forhold, der  $i$  tjener  $j$  og med dette får muligheten til å leve et liv som kanskje er sikrere enn det var før, ved at  $j$  ifølge Hobbes må forsvare  $i$  dersom  $i$  skal ansees som hans eiendom:

[...] for as all other things, so servants also are acquired by war, whom in equity the lord must protect, if he will have them to be his (Hobbes 2005a:113).

Leo Strauss (1952) ser ut til å støtte en slik tolkning av underkastelse hos Hobbes, og har som vi tidligere har sett også hevdet at underkastelse er det beste valget der en aktør frykter for å miste sitt liv<sup>17</sup>. Hobbes gir to grunner for å underkaste seg:

But in they who submit themselves to another for fear, either submit to him whom they fear, or some other whom they confide in for protection. They act according to the first manner, who are vanquished in war, that they may not be slain; they according to the second, who are not yet overcome, that they may not be overcome (Hobbes 2005a:70).

Det er kontroversielt å si at Hobbes' individer anser underkastelse som bedre enn åpen konflikt, men jeg mener det finnes god støtte for et slikt syn. Det bør i alle fall vurderes seriøst, og med en slik forståelse av Hobbes får man også et mer realistisk bilde av hvordan naturtilstanden kunne tenkes å arte seg.

### Grunnleggerne av staten – hvem er de?

Jeg nevnte såvidt at tilgangen på strategi  $\tau^0$  (“utgangsstrategien” som tar spillerne fra naturtilstanden til statstilstanden) kan diskuteres, og at overgangen fra naturtilstanden til staten også er noe uklar. Taylor synes å ha et for restriktivt syn på muligheten for å danne staten<sup>18</sup>:

[...] it [fred] obtains when *no* individual chooses strategy  $D$ . (For he later argues that *everybody* must behave differently if society is to move out of the state of War into that of Peace.) (Taylor 1987:130).

---

taking away of another's life' ” (Strauss 1952:21).

<sup>17</sup>“Fear for his life, which came upon the man in his struggle for triumph, moderates, even kills, the will to triumph, and makes him ready to submit, to leave triumph to the enemy, in order that he may save his own life” (Strauss 1952:21).

<sup>18</sup>Det er imidlertid et spørsmål om hvor i boken man leser; visse steder virker det som Taylor anser innsettelsen av suverenen som mer tilgjengelig.

Taylor baserer denne forståelsen på Hobbes' omtale av *sosialkontrakten*:

*A commonwealth* is said to be *instituted*, when a *multitude* of men do agree, and *covenant, every one, with every one*, that to whatsoever *man*, or *assembly of men*, shall be given by the major part, the *right to present* the person of them all, that is to say, to be their *representative*; every one, as well he that *voted for it*, as he that *voted against it*, shall *authorize* all the actions and judgments, of that man, or assembly of men, in the same manner, as if they were his own, to the end, to live peaceably amongst themselves, and be protected against other men (Hobbes 1946:113).

Ut fra dette må det spørres hvordan en stat kan dannes; Taylor diskuterer en gitt befolkning, og hevder forholdsvis sterkt at *alle* må være enige (velge  $\tau^1$ ), og deretter *endre adferd* (Taylor 1987:130). At alle må endre adferd er korrekt, men at alle må være *enige* i at staten skal dannes er mer tvilsomt. I en gitt befolkning – la oss si på det geografiske området Grønland – kan det tenkes at en *folkemasse* går inn for å danne en stat, slik Hobbes sier det. Ifølge sitatet ovenfor er det altså *majoriteten* i denne folke-massen som bestemmer, og resten må innordne seg etter hva majoriteten har bestemt. Videre: *det er heller ikke nødvendig at denne folkemassen inkluderer hele populasjonen*; folkemassen kan igjen ansees som en majoritetskoalisjon i populasjonen.

Thirdly, because the major part hath by consenting voices declared a sovereign; he that dissented must now consent with the rest; that is, be contented to avow all the action he shall do, or else justly be destroyed by the rest. For if he voluntarily entered into the congregation of them that were assembled, he sufficiently declared thereby his will, and therefore tacitly covenanted, to stand to what the major part should ordain: and therefore if he refuse to stand thereto, or make protestation against any of their decrees, he does contrary to his covenant, and therefore unjustly. And whether he be of the congregation, or not; and whether his consent be asked, or not, he must either submit to their decrees, or be left in the condition of war he was in before; wherein he might without injustice be destroyed by any man whatsoever (Hobbes 1946:115).

På Grønland kan det altså dannes en stat av en majoritetskoalisjon i populasjonen, og dette innebærer at kravet om universell strategienighet i et gitt Taylor-spill ikke kan kreves på bakgrunn av Hobbes' filosofi. At alle må endre adferd trenger dermed ikke foregå

frivillig eller spontant. Så snart flertallet (eller en stor *nok* gruppering) har innsatt en suveren, kan minoriteten oppfordres til å endre adferd; dersom en slik oppfordring ikke har ønsket effekt, kan de faktisk rettmessig *tilintetgjøres*, ifølge Hobbes.

But new and thorny problems emerge. In particular, as we shall see, in those cases where *subsets* of the players find it collectively worthwhile to provide the public good, there arises a quite different strategic problem, which results from some players having an incentive to ensure that the subset which provides the public good does not include themselves. In this way, a Chicken-like game is generated by the Prisoners' Dilemma (Taylor 1987:82).

Slike argumenter som Taylor her fremmer, blir meningsløse i naturtilstanden dersom samarbeid likestilles med staten slik Taylor gjør i sin diskusjon av Hobbes, og blir kun et spørsmål om fellesgode-politikk der staten allerede eksisterer. Dersom en gruppe i Taylors spill da finner det verdt det å tilby et gode selv – og danner staten – vil de resterende enten måtte innordne seg dette eller holde seg utenfor staten. Der staten er dannet er det grunn til å tro at den, dersom sterk nok, vil fremtvinge a) innordning i staten eller b) fordrivelse.

### **Hobbes' naturtilstand: alle mot én, alle mot alle, eller de mot dem?**

Jeg har i 5.2.1 sett på to muligheter for å formalisere sentrale elementer i Hobbes' filosofi: parvise interaksjoner eller simultan og upersonlig massesamhandling. Taylor velger å se på hele populasjonen som samhandlende, og følger Nash (1997) i å behandle aktørene som fullstendig uavhengige av hverandre. Greit nok, naturtilstanden er jo en krig der alle kriger med alle?

[...] and such a war, as is of every man, against every man (Hobbes 1946:82).

Jeg finner at det å stoppe her og godta en slik fremstilling *ikke* er "greit nok". Det er et av paradoksene hos Hobbes at han omtaler krigen på denne måten, men det kan tenkes å bero på at vi som oftest tenker på "krig", og Hobbes' naturtilstand, som noe helt annet enn hva Hobbes har i tankene. Han hevder blant annet at indianerne i USA på hans tid faktisk levde i det han forstod som naturtilstanden (Hobbes 1946:83). Det er neppe noen som vil hevde at hver indianer faktisk kriget med hver indianer, at intet

samarbeid fant sted, eller at ingen koalisjoner – eller stammer om man vil – ble dannet der.

Robert Nozicks *Anarchy, State and Utopia* (1974) gir et svært interessant perspektiv på hvordan Nozick ser for seg at mennesker i naturtilstanden nødvendigvis vil danne grupper, som deretter konkurrerer med hverandre. Dette vil uunngåelig føre til at det dannes større og større grupper, noe som til slutt gjør (den minimale) staten til en uunngåelighet. Hobbes sier også eksplisitt at koalisjoner forekommer i naturtilstanden; den svakeste kan drepe den sterkeste, enten ved list eller “ved en sammensvergelse med andre”; en person som skaper verdier må passe seg, siden “andre må forventes å komme med forente krefter” for å ta det han har skapt; mennesker som er opptatt av konkurranse “bruker vold, for å gjøre seg til herre over andre menns person, deres koner, barn, og kyr”; det beste forsvar er å skaffe seg tilstrekkelig makt ved å være herre over “så mange personer han klarer” – alt dette i naturtilstanden (Hobbes 1946:80,81).

Dette poenget er heller ikke spesielt kontroversielt, og bl.a. Kavka (1986) mener at det opplagt tillates en viss grad av samarbeid i naturtilstanden:

[...] cooperation in the state of nature is possible, but it is precarious because of each party's lack of assurance of the other's reliability (Kavka 1986:154).

Også Gauthier (1969:113) peker på at Hobbes helt klart innser at (politiske) grupperinger kan oppstå i naturtilstanden, men at dette er av sekundær betydning, da Hobbes fokuserer på dannelsen av staten ved felles enighet og sosialkontrakten. Videre sier også Hoekstra (2007a) at det finnes grupperinger i naturtilstanden, og berører i sitatet nedenfor en annen kilde til oppfatningen om at *intet* samarbeid eksisterer i naturtilstanden:

In the most extreme sentence of his characterization of the natural condition he does say that there are ‘no arts, no letters, no society’, but there are many indications that society may be found there (Hoekstra 2007a:119).

Sitatet Hoekstra (2007a) her sikter til, der Hobbes på kraftfullt vis legger ut om elendigheten i naturtilstanden, er svært berømt, og kommer fra kapittel 13 i *Leviathan*:

In such condition, there is no place for industry; because the fruit thereof is uncertain: and consequently no culture of the earth; no navigation, nor use of the commodities that may be imported by sea; no commodious building;

no instruments of moving, and removing, such things as require much force; no knowledge of the face of the earth; no account of time; no arts; no letters; no society; and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short (Hobbes 1946:82).

Det fremstår som mest naturlig å tolke dette sitatet som en beskrivelse av et ekstremtilfelle, om man ikke skal tolke det som et retorisk grep for å få frem et poeng, som Hoekstra (2007a:112) gjør: “Hobbes’ hovedmål ved å tilby sitt portrett av lidelsen i naturtilstanden, er å skremme sine lesere til å holde fast ved den orden som allerede er etablert over dem”:

Most strikingly, his famous litany of what that condition lacks (‘there is no place for industry, . . . no culture of the earth, no navigation . . . , no account of time, no arts, no letters, no society’) is an adaption of a hyperbolic trope, characterizing uncivilized peoples by a negative list, which became conventional in the century after Columbus landed (Hoekstra 2007a:113).

Hoekstra (2007a:113) gir videre mange eksempler på andre som har laget slike lister “for å bevise betydningen av en bestemt kunst ved å beskrive menneskeheten uten denne kunsten”. Aristophanes, Suetonius og Horace ville fremme poesi; uten dette hevdet de at “mennesker var i en desperat kamp for ussel overlevelse, og uten fordelene ved sivilisasjon” (Hoekstra 2007a:113). Retorikere, sofister og filosofer har også hevdet at deres fag har reddet verden fra en tilstand lignende den beskrevet i Hobbes’ “litani”, og mens dette ikke betyr at Hobbes tar feil, betyr det i det minste at det er et svært velkjent – og effektivt – retorisk grep han her gjør.

I *De Corpore* har Hobbes et sitat som ligner mye på det det gjengitt fra *Leviathan* ovenfor:

Now, the greatest commodities of mankind are the arts; namely, of measuring matter and motion; of moving ponderous bodies; of architecture; of navigation; of making instruments for all uses; of calculating the celestial motions, the aspects of the stars, and the parts of time; of geography &c (Hobbes 1999:189).

Denne listen kommer her i forbindelse med Hobbes’ forsøk på å beskrive nytten av *filosofi*. Metoden han bruker er å se på de beste tingene mennesker har i hans samfunn,

for så å så se hva de mangler i de mest primitive samfunn. Det er dermed åpenbart at samme argumentasjon kunne vært ført til inntekt for en annen kunst, eller innsikt, man har ett sted, men ikke et annet. At Hobbes tillegger filosofi “menneskehetens viktigste goder” er interessant, men det endrer ikke det faktum at korrelasjon ikke nødvendigvis betyr kausalitet.

Det synes mest plausibelt å anta at Hobbes’ liste over mangler ikke *nødvendigvis* gjelder like sterkt i alle varianter av naturtilstanden. Hobbes hevder tross alt at man er i naturtilstanden i alle situasjoner der man ikke har én suveren som kan binde oss alle med ærefrykt (Hobbes 1946:82).

Dersom Hobbes’ sitat forståes bokstavelig, er det forståelig at man ser for seg den mest ekstreme elendighet i naturtilstanden. Like etter dette sitatet sier Hobbes at han ikke tror mennesker *generelt sett* har levd i en slik tilstand, men at det på hans tid var mange som levde slik. I hans beskrivelse av indianerne sier han at disse – *bortsett fra styret i små familier* – er uten en regjering. At Hobbes her, i selve eksempelet på hva som faktisk *er* naturtilstanden, sier at mennesker lever i små familier, er bemerkelsesverdig. Disse familiene er viktigere enn man lett kan få inntrykk av, spesielt siden Hobbes i *De Cive* er klar på at “en familie er en liten by”, og en *by* i denne boken er noe lignende hans “samfunn” (commonwealth) i *Leviathan* (Hobbes 2005a:84n). Dette er heller ikke en glipp av Hobbes, da han også senere i boken poengterer at “en stor familie er et kongedømme, og et lite kongedømme en familie” (Hobbes 2005a:108). Disse sitatene gjør det plausibelt å anta at Hobbes ikke bare snakker om mann, kone og barn når han sier familier eksisterer i naturtilstanden – større ansamlinger mennesker som lever under en herre er for Hobbes en *familie*.

Det faktum at Hobbes illustrerer naturtilstanden med en borgerkrig i (for Hobbes) moderne stater, bør være nok til ikke å ta hans liste over mangler bokstavelig og som universell; med mindre man da hevder at alle kart, klokker og kunnskap generelt sett forsvinner ved et krigsutbrudd, et standpunkt som ikke fremstår som særlig plausibelt. At det han nevner kanskje ikke vil *utvikles* i naturtilstanden er noe mer sannsynlig, men selv dette virker noe overdrevet. Dersom Hobbes faktisk begrenset naturtilstanden til en slik situasjon, er det *mye* i hans tekster som fremstår som fullstendig meningsløst, noe som taler imot å ta hans sitat bokstavelig og som en universell beskrivelse av naturtilstanden. Dersom man derimot anser sitatet som en beskrivelse av et verste tilfelle, eventuelt en overdrivelse for å overbevise om et poeng, er det uproblematisk.

Hobbes' bruk av retorikk i *Leviathan* er en annen grunn til å mistenke at sitatet om elendigheten i naturtilstanden som gjengis i *Leviathan* kan være en overdrivelse<sup>19</sup>. Skinner (1996:426) peker på at Hobbes i *The Elements of Law* og *De Cive* er av den oppfatning at vitenskapelige resonnementer er nok til å overbevise leserne, men at han i *Leviathan* har endret oppfatning, og der i langt større grad anvender retoriske virkemidler for å understreke sine poenger (Skinner 1996:176,426)<sup>20</sup>. Det kan derfor være nyttig å se hvordan Hobbes beskriver naturtilstanden i to tidligere verker; i *The Elements of Law* sier han følgende:

[. . .] as we know also that it is, both by the experience of savage nations that live at this day, and by the histories of our ancestors, the old inhabitants of Germany and other now civil countries, where we find the people few and short lived, and without the ornaments and comforts of life, which by peace and society are usually invented and procured [. . .] (Hobbes 1999:80)

Beskrivelsen av denne tilstanden i *De Cive* er svært lik:

They of America are examples hereof, even in this present age: other nations have been in former ages; which now indeed are become civil and flourishing, but were then few, fierce, short-lived, poor, nasty and deprived of all that pleasure and beauty of life, which peace and society are wont to bring with them (Hobbes 2005a:12).

Det er opplagt at manglene i en tilstand med vedvarende krig også her beskrives som svært alvorlige. Det er allikevel en markant forskjell mellom å mangle "ornamentene og bekvemmelighetene", eller "nyttelsen og skjønnheten", ved samfunn eller sivilisasjon, og det å mangle alt Hobbes lister opp i *Leviathan*. At livet er kort og brutalt fremstår som det sentrale, og dette er ut fra Hobbes' øvrige teori tilstrekkelig for å forklare hvorfor

---

<sup>19</sup>Jeg vil ikke analysere dette temaet inngående her, men føler det er nyttig å påpeke i forbindelse med Hobbes' beskrivelse av naturtilstanden, da den ofte blir lest for bokstavelig. Hobbes' retorikk er viktig, og man bør være klar over hvordan han anvender retoriske virkemidler; "[. . .] a considerable and quite diverse number of contemporary scholars of Hobbes, ranging from Victoria Kahn to David Johnston, who claim that Hobbes disguises but depends upon rhetoric [. . .] in order to manipulate his readers and develop a workable theory of state power and political authority. [. . .] These authors claim that Hobbes's famous and outspoken dislike for rhetoric is a ruse, allowing him to practice it in secret while denouncing the use of rhetoric by others (potentially dangerous rivals to his own rhetorical production of sovereign authority)" (Martel 2007:13).

<sup>20</sup>Dette kommer tydelig frem bl.a. i følgende sitat fra *Leviathan*: "Again, in all deliberations, and in all pleadings, the faculty of solid reasoning is necessary: for without it, the resolutions of men are rash, and their sentences unjust: and yet if there be not powerful eloquence [retorikk], which procureth attention and consent, the effect of reason will be little" (Hobbes 1946:460).



mennesker ønsker seg bort fra en slik tilstand.

Den atomistiske forståelsen av naturtilstanden som *alle mot alle*, bokstavelig talt, må nyanseres til å inkludere noe større molekyler. Dette behøver ikke endre det faktum at (nesten) alle mennesker har motstridende interesser, og dermed *kan forstås som fiender*. Det betyr at ikke alle aktivt kjemper mot alle hele tiden. Det utelukkes dermed heller ikke at mennesker i naturtilstanden kan finne interesser de deler med visse andre individer. Hobbes er også klar på at mødre f.eks. er *herrer* over sine barn i naturtilstanden: “for siden mennesker ved naturens nødvendighet søker det som fremstår som godt for oss, kan det ikke forstås slik at noen skal gi liv til en annen, slik at han med årene kan bli sterk, og med det en fiende. [...] Og derfor blir enhver kvinne som bærer barn både *mor* og *herre* i naturtilstanden” (Hobbes 2005a:116). At mor og barn, eller nær familie generelt sett, skulle være i krig, synes ikke å stemme overens med Hobbes’ fremstilling av menneskers natur som presentert i kapittel 2. Gauthier (1969:113) påpeker også denne tvetydigheten hos Hobbes:

There is an apparent incompatibility between Hobbes’s account of the family and his account of the state of nature. If men are born and raised in families, and if each family is a small-scale society, then how can men ever exist in the condition of war of all against all? (Gauthier 1969:118)

Gauthier (1969:118) “løser” dette problemet ved å anse naturtilstanden som en “logisk abstraksjon”, og hevder at menneskene Hobbes her omtaler ikke er “vesener født og oppdratt i en familie”. Å se naturtilstanden på denne måte løser riktignok de problemene jeg her har diskutert (hvorvidt det er en reell krig mellom *alle* og hvorvidt nær sagt *alt* mangler), men Hobbes’ øvrige argumentasjon ser tydelig ut til å være basert på reelle mennesker, med en menneskenatur som er kjent for enhver som kjenner seg selv. Av denne grunn forfølger jeg i denne oppgaven et syn på naturtilstanden som en reell tilstand med reelle individer, men bemerker at det finnes alternativer til denne forståelsen.

At individer til en viss grad kan samarbeide i naturtilstanden synes opplagt, og det kan tenkes motforestillingene mot dette delvis skyldes en annen feiloppfatning av Hobbes: at ingen type kontrakter eller avtaler kan inngås i naturtilstanden. Russel Hardin fremmer et slikt syn, og hevder at “[e]n kontrakt i naturtilstanden er ugyldig” (Hardin 1991:160). At Hobbes regnet med at slike kontrakter *kunne* forekomme, synes imidlertid tydelig å fremgå ved en lesning av hans tekster:

Covenants entered into by fear, in the condition of mere nature, are obligatory. For example, if I covenant to pay a ransom, or service for my life, to an enemy; I am bound by it: for it is a contract, wherein one receiveth the benefit of life; the other is to receive money, or service for it; and consequently, where no other law, as in the condition of mere nature, forbiddeth the performance, the covenant is valid (Hobbes 1946:91).

Frykten for en hvilken som helst forestilt åndelig makt er også ofte nok til å sikre at kontrakter opprettholdes, mener Hobbes. Det samme gjelder vektlegging av ære, og stolthet ved å vise at man ikke behøver å bryte avtaler. Disse forholdene kan i naturtilstanden gjøre at ord og avtaler blir styrket (Hobbes 1946:92). I naturtilstanden er det også en kontrakt mellom foreldre når det gjelder herredømme over avkom (Hobbes 1946:131). Ryan (1996:226) fremmer et lignende syn på kontrakter i naturtilstanden, og mener Hobbes “ikke mente at alle kontrakter i naturtilstanden var ugyldige uten en håndhevende makt”; dersom man forventer at motparten innfrir sin plikt, og det er trygt å overholde avtalen, er kontrakten gyldig også utenfor staten (Ryan 1996:226). At mennesker faktisk innfrir slike kontrakter skyldes som nevnt i stor grad frykt, noe Machiavelli (2003) også har bemerket:

Love endures by a bond which men, being scoundrels, may break whenever it serves their advantage to do so; but fear is supported by the dread of pain, which is ever present (Machiavelli 2003:66).

Det er altså liten tvil om at kommunikasjon og en viss grad av samarbeid ifølge Hobbes er mulig, og at avtaler til en viss grad kan inngås i naturtilstanden. “Til en viss grad”, da det er opplagt at Hobbes ikke anser stabiliteten i hverken de kontraktene jeg her snakker om, eller koalisjonene, som særlig god. At slike koalisjoner er ustabile hindrer dem allikevel ikke i å være sentrale i Hobbes’ naturtilstand. Det synes dermed rimelig å kreve at en god formalisering av naturtilstanden bør ta hensyn til dette. Hobbes er, som vi har sett, klar på at grunnen til at *usikkerhet* er et karakteristisk trekk ved naturtilstanden er at “andre må forventes å komme med forente krefter” for å ta det en skaper (Hobbes 1946:81).

von Neumann og Morgenstern (2004:15) poengterer at koalisjoner vil ha avgjørende betydning i spillsituasjoner der aktører kan tenkes å handle sammen:

E.g., it is clear that if certain great groups of participants will – for any reason whatsoever – act together, then the great number of participants

may not become effective; the decisive exchanges may take place directly between large “coalitions,” few in number, and not between individuals, many in number, acting independently. Our subsequent discussion of “games of strategy” will show that the role and size of “coalitions” is decisive throughout the entire subject (von Neumann og Morgenstern 2004:15).

Denne beskrivelse passer forøvrig godt med et annet bilde Hobbes gir på naturtilstanden: tilstanden suverener og stater befinner seg i, i forhold til hverandre (Hobbes 1946:83). Mens von Neumann og Morgenstern (2004) analyserer forhold mellom koalisjoner i spill der samarbeid er mulig, er Nash sin teori om ikke-kooperative spill, slik den fremgår i hans artikkel “Non-Cooperative Games” (Nash 1997), basert på “*fraværet* av koalisjoner, da det forutsettes at hver deltager handler uavhengig, uten samarbeid eller kommunikasjon med noen av de andre” (Nash 1997:14). Ut fra min lesning av Hobbes ser jeg ingen god grunn til å velge en slik fremstilling av hans naturtilstand, da menneskene der åpenbart *kan* kommunisere, og faktisk også til en viss grad inngå avtaler. Det Nash kaller kooperative spill synes som et interessant alternativ, der man har en “antagelse om at spillerne kan, og vil, samarbeide som de gjør i von Neumann og Morgensterns teori” (Nash 1997:25).

Måten “kooperative spill” defineres på utelukker imidlertid at dette kan anvendes i formaliseringen av naturtilstanden, da det innebærer at “spillerne kan kommunisere og danne koalisjoner som vil håndheves av en dommer” (Nash 1997:25). Det kan ikke forventes å eksistere en uavhengig dommer i Hobbes’ naturtilstand, men det er til tross for dette åpenbart at Hobbes ser for seg en samhandlingssituasjon som ligner mer på von Neuman og Morgensterns spill med koalisjoner enn Nashs spill med fullstendig fravær av kommunikasjon og avtaler. Det er imidlertid mulig å analysere mulig samarbeid også i ikke-kooperative modeller, for eksempel er “mange nyere, spillteoretisk baserte studier av samarbeidsprosesser basert på *ikke-kooperative* modeller” (Hovi 2008:33). Dette temaet fremstår som sentralt i videre spillteoretisk arbeid med Hobbes, men følges ikke videre i denne oppgaven.

### **Hobbes om frivillig samarbeid: små og store samfunn**

Etter at en har gjennomgått – og akseptert – muligheten for samarbeid i små (komunitarianske) samfunn gjenstår det å undersøke hvorvidt en slik mulighet påvirker Hobbes’ resonnement. Taylor mener Hobbes’ resonnement svekkes; han vil vise at blant

annet Hobbes' forsvar for staten er ugyldig, da samarbeid er mulig uten en stat. Må man så akseptere at han har utført det han ønsket, å "motbevise" Hobbes, fordi han har bevist at det finnes en viss mulighet for frivillig samarbeid? For meg er svaret på dette spørsmålet *nei*.

Først og fremst har jeg i diskusjonen om koalisjoner like ovenfor vist at Hobbes selv så muligheten for en viss type samarbeid i naturtilstanden. Indianerne – mennesker det ikke skal så mye mot til for å karakterisere som organisert i "kommunitarianske samfunn" – levde altså i naturtilstanden<sup>21</sup>:

It may peradventure be thought, there was never such a time, nor condition of war as this; and I believe it was never generally so, over all the world; but there are many places, where they live so now. For the savage people in many places of America, except the government of small families, the concord whereof dependeth on natural lust, have no government at all; and live at this day in that brutish manner, as I said before (Hobbes 1946:83).

Uansett hva man synes om hans beskrivelse av indianernes tilstand er det klart at Hobbes krever mer enn kommunitarianske samfunn som er relativt fredelige for å unnsnippe naturtilstanden. Hans videre beskrivelse av naturtilstanden har blitt sitert ofte:

Howsoever, it may be perceived what manner of life there would be, where there were no common power to fear, by the manner of life, which men that have formerly lived under a peaceful government, use to degenerate into, in a civil war (Hobbes 1946:83).

Hobbes presenterer i sitatet om en fallen stat en *illustrasjon* av naturtilstanden for å tydeliggjøre aspekter ved den som kan være vanskelige å forstå for mennesker med all sin erfaring fra statstilstanden. Denne illustrasjonen blir imidlertid ofte forstått som hans egentlige beskrivelse av naturtilstanden, noe jeg mener er noe feil. Indianerne levde faktisk i naturtilstanden, sier Hobbes, mens "man kan forestille seg" at livet der en moderne stat faller sammen, ville ligne. Men ligner det virkelig? I mine øyne gjør det det, og dette synes å være relatert til mitt argument om at kommunitarianske samfunn er *eksternt ustabile*. Indianer-stammer i USA hadde god plass, og klarte seg antageligvis forholdsvis godt, da de heller ikke dyrket jorden slik at det ble nødvendig for dem å

---

<sup>21</sup>Hoekstra (2007a:112) problematiserer Hobbes' anvendelse av indianere som eksempel ved å peke på at han kan ha hatt begrenset kjennskap til hvordan de faktisk levde.

forsvare land. “Stammer” i en moderne stat, med fast eiendom, ofte med uklare grenser pga. en fredelig utvikling over lang tid, og svært lite plass pr. person sammenlignet med indianerne i USA, vil med stor sannsynlighet finne det langt vanskeligere å etablere frivillig og harmonisk samarbeid i særlig utstrekning. Den knappheten Hobbes regner med i naturtilstanden, og usikkerheten ved å ha sine fiender rett utenfor døren, vil gjøre en borgerkrig i en tett befolket stat langt mer ekstrem enn indianernes tilstand. Denne typen argumentasjon stemmer godt overens med Taylors konklusjoner i *Community, Anarchy and Liberty* (Taylor 1982). Dette tyder på at “menneskers naturlige tilstand” kan innebære mange – og svært forskjellige – situasjoner. Disse situasjonene har ifølge Hobbes en sentral mangel felles: den sterke staten de trenger for å sikre fred.

Taylor selv innrømmer at hans “samfunn” nødvendigvis er små, og det later dermed til at muligheten for denne typen samarbeid vanskelig kan brukes til å endre Hobbes’ resonnement:

Nor is it the joining together of a small number of men, that gives them this security; because in small numbers, small additions on the one side or the other, make the advantage of strength so great, as is sufficient to carry the victory; and therefore gives encouragement to an invasion. The multitude sufficient to confide in for security, is not determined by any certain number, but by comparison with the enemy we fear; and is then sufficient, when the odds of the enemy is not of so visible and conspicuous moment, to determine the event of war, as to move him to attempt (Hobbes 1946:110).

And as for very little commonwealths, be they popular, or monarchial, there is no human wisdom can uphold them, longer than the jealousy lasteth of their potent neighbours (Hobbes 1946:172).

Hobbes ser ut til å diskutere ganske store samfunn – *Leviathan* er nødvendigvis stor<sup>22</sup> – og dette er grunnen til at jeg anser Hobbes’ syn på statens nødvendighet som plausibel dersom jeg holder meg til hans egen begrepsbruk. Som Ryan (1996:219) sier:

[T]he theory may be designed around the problem of sustaining and policing

---

<sup>22</sup> *Leviathan* er et enormt bibelsk sjømonster, og en vanlig tolkning er at Hobbes valgte dette navnet for å symbolisere egenskaper ved det han anså som en god stat. “[...] the mythological image is this not merely a superfluous, accidental ornament but serves a theoretical purpose in Hobbes’s argument” (Tralau 2007:61).

a *large*, prosperous society, where most people are known well only to a few friends but want to transact business and hold intellectual converse with distant strangers (Ryan 1996:219).

Min tolkning av Hobbes' syn på naturtilstanden fremstår muligens som noe ukonvensjonell, men den er ikke uten presedens. Det er forståelig at det oppstår forskjellige oppfatninger om naturtilstanden pga. Hobbes' noe tvetydige omtale av den, selve navnet på tilstanden og tidligere tenkeres anvendelse av begrepet. Når man ser for seg en *naturlig* tilstand ser man gjerne for seg en heller dyrisk og primitiv tilstand, men dette er altså feil i forhold til Hobbes (Hoekstra 2007a:114). Taylors resonnement er dermed i mine øyne allerede fra utgangspunktet maktesløst mot Hobbes' forsvar for staten, da Hobbes ikke rettferdiggjør staten pga. fullstendig mangel på samarbeid i naturtilstanden, men pga. manglende sikkerhet der et slikt samarbeid ikke er av tilstrekkelig utstrekning og ikke understøttes av suverenen.

Til syvende og sist blir vurderingen av hvorvidt Taylors "angrep" på Hobbes er vellykket et spørsmål om hvordan man tolker Hobbes. Dersom man forstår naturtilstanden som *alle mot alle* i *aktiv* krig (som Kavka gjør<sup>23</sup>), der mennesker mangler nær sagt alle fordeler ved sivilisasjon, er det viktig at Taylor viser at mennesker *kan* oppnå mer enn dette; dersom man tolker Hobbes slik jeg her argumenterer for, kan det hevdes at Taylor kun viser noe Hobbes selv har vurdert og forkastet.

Selv om man skulle slutte at staten ikke var *nødvendig* siden man kan oppnå *noe* med kommunitarianske samfunn, synes det som om Hobbes ville avfeie slike samfunn på nok et grunnlag: normstyrte og ofte hierarkiske samfunn kan tenkes å stride sterkt med Hobbes' krav om likhet og fravær av en objektiv "sannhet" om det gode.

Jeg har allerede diskutert hva Hobbes anser som statens formål og hvorfor jeg velger å anse Hobbes som en protoliberaler i 4.1.1 og under premiss 6 i 4.2.2. Jaume (2007) diskuterer i artikkelen "Hobbes and the Philosophical Sources of Liberalism" at det kan synes rimelig å anse visse aspekter ved Hobbes' filosofi som representative for senere

---

<sup>23</sup>Kavka (1986:91) velger å lese Hobbes' sitat om "ingen industri" osv. bokstavelig, og "denne sørgelige tilstanden som Hobbes beskriver ville sannsynligvis ikke bli frembragt av en krig mellom alle (etter hans definisjon), med mindre det forekom betydelig faktisk kamp". Problemet er jo at Hobbes beskriver krigstilstanden som preget av en *tilbøyelighet* til kamp, og man må dermed enten se bort fra eller omtolke denne beskrivelsen, eller konsekvensene Hobbes tillegger naturtilstanden. Kavka selv er eksplisitt på dette punktet, og forklarer at hans krav om *aktiv* krig er en direkte følge at det er nødvendig for å stemme overens med beskrivelsen Hobbes gir i sitatet jeg har hevdet kan være et retorisk grep.

liberal tenkning:

The state must be strong within its legitimate domain, but it becomes detestable once it stoops to destroy 'liberty in the proper sense'. This particular aspect of Hobbes's thought is welcomed by liberals such as Constant or Tocqueville, but they could hardly have been expected to render homage to Hobbes for it! (Jaume 2007:213)

Hobbes er absolutist, men det er ikke selvsagt at mennesker som er opptatt av selvutfoldelse og frihet til å ta sine egne avgjørelser er verre stilt i en Hobbesiansk stat enn i et sterkt normstyrt samfunn, noe jeg kommer nærmere inn på i 5.3.2. Jeg vil avslutte behandlingen av vurderingen av kommunitarianske samfunn som et reelt alternativ til staten med noen betraktninger om hvorfor Hobbes – som "liberaler" – sannsynligvis ikke ville ønsket eller godtatt slike samfunn.

Flertallets tyranni er et velkjent begrep, og det er grunn til å anta at kommunitarianske samfunn er spesielt utsatt for dette fenomenet. Mill poengterer at "folkets vilje" i realiteten er viljen til de mest aktive eller tallrike, noe som også bemerkes av Fromm (1994:113). Da disse kan tenkes å ønske å undertrykke minoritetene, er det "like nødvendig å forsikre seg mot dette som å forsikre seg mot ethvert annet maktmisbruk" (Mill 1985:62).

Society can and does execute its own mandates; and if it issues wrong mandates instead of right, or any mandates at all in things with which it ought not to meddle, it practices a social tyranny more formidable than many kinds of political oppression, since, though not usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating much more deeply into the details of life, and enslaving the soul itself (Mill 1985:63).

Det er også et viktig poeng at fraværet av en statsmakt ikke nødvendigvis medfører fravær av autoritet, og dermed ikke nødvendigvis fører til "frihet". "Autoritet behøver ikke å være en person som sier: du må gjøre dette, eller du kan ikke gjøre det" – autoritet kan internaliseres gjennom f.eks. samvittighet og pliktfølelse (Fromm 1994:165). Samvittighet er i stor grad – men ikke utelukkende – formet av sosialisering og normer, og det samme gjelder pliktfølelse. Trangen til å fjerne seg fra eksterne og eksplisitte autoriteter vil ifølge Fromm kunne føre til en situasjon som ikke nødvendigvis er bedre enn det å leve under en suveren (Fromm 1994:105). Der den synlige autoriteten er borte har man fortsatt det Fromm kaller "anonym" autoritet (Fromm 1994:166). "Den

er forkledd som sunn fornuft, vitenskap, psykisk helse, normalitet og folkeopinion. Den krever ikke noe annet enn det selvsagte. Den later heller ikke ikke til å bruke makt – kun mild overtalelse” (Fromm 1994:166).

But whereas in internalized authority the command, though an internal one, remains visible, in anonymous authority both command and commander have become invisible. It is like being fired at by an invisible enemy. There is nobody and nothing to fire back against (Fromm 1994:166).

The rulership of conscience can be even harsher than that of external authorities, since the individual feel its orders to be his own; how can he rebel against himself? (Fromm 1994:165)

#### 5.2.4 Oppsummering

Med min tolkning av naturtilstanden vil jeg hevde at Taylors resonnement om den teoretiske muligheten for utenomstatlig samarbeid ikke gir grunnlag for å endre premiss 1. Dersom Hobbes vet at mennesker samarbeider, og sier at det allikevel kun finnes én “løsning” på problemet han kaller naturtilstanden, synes det rimelig å hevde at staten må være en nødvendighet i et Hobbesiansk resonnement. Staten *form* diskuteres videre i den følgende delen av kapittelet.

Hobbes’ absolutistiske elementer er problematiske, men det forhindrer ikke at de liberale tendensene hos Hobbes er klare, og viktige å anerkjenne; det synes forholdsvis klart at Hobbes også på grunnlag av sine liberale prinsipper ville forkastet Taylors kommunitarianske samfunn som undertrykkende og for lite sikre. At Hobbes ønsket en absolutt stat betød ikke at han ville fortelle mennesker hvordan de burde leve – han forteller mennesker hvordan de etter hans oppfatning *må* leve dersom de skal unngå naturtilstanden:

What he did achieve was a remarkable breakthrough in self-and-other acceptance, given that people are passionately addicted to telling other people how they should live. Hobbes was more determined to resist that impulse than anyone before him, and possibly after him as well (Green 1993:112).



## 5.3 Mennesker og behovet for mer enn sikkerhet

I gjennomgangen av Maslow i 3.3 fant jeg en svært interessant teori: mennesker kan gjerne ha en fast natur, slik Hobbes antar, men de kan allikevel endres fundamentalt i gitte situasjoner når det gjelder hvilke behov som gjør seg gjeldende. Hobbes' syn på mennesket synes meget plausibelt i forbindelse med menneskers "lavere" behov, som de fysiologiske behov og sikkerhetsbehovene; Maslow påpeker dog at disse behovene ikke lenger er av særlig stor betydning for individer som opplever både fysiologisk og sosial sikkerhet – *andre* behov vil da dominere mennesket.

Denne teorien vil kunne vise seg å utgjøre en avgjørende innvending mot konklusjonen på Hobbes' resonnement i kapittel 4 – at Hobbes' stat er nødvendig og ønskelig. Der- som man skal si noe om hvorvidt Hobbes' filosofi er *universell*, synes det interessant å foreta en vurdering av Hobbes' resonnement m.t.p. mennesker som lever i dagens moderne samfunn. En av de viktigste forskjellene mellom individene Hobbes omtaler, og oss, kan sies å være vårt forhold til *frihet* og *demokrati*. I denne delen av oppgaven ønsker jeg derfor å utarbeide en arbeidshypotese som sier noe om menneskers forhold til disse begrepene, for deretter å vurdere Hobbes' resonnement ut fra en antagelse om at mennesker kan tenkes å ha et forhold til frihet og demokrati Hobbes selv ikke diskuterer.

Da jeg utvider Hobbes' forståelse av mennesket, ved å inkludere noe moderne psykolo- gisk innsikt, som f.eks. Maslows behovshierarki, er det ikke utelukkende Hobbes' egne betraktninger som vil vurderes i denne delen av oppgaven. Jeg vil fortløpende under- streke hvor det er snakk om hva Hobbes selv sier, og hva jeg mener kan ansees som en følge av Hobbes' *resonnement*. Mens Kavka (1986), Gauthier (1969) og andre lager "Hobbesianske teorier" der man forsøker å korrigere samtlige av Hobbes' feil, er jeg i denne delen mer konservativ, og endrer kun de delene av resonnementet der Hobbes uttaler seg tvetydig, der han ikke kommer med spesifikke synspunkter eller der utvik- lingen har ført til at Hobbes faktisk diskuterte noe annet enn hva vi i dag forstår med begrepene han anvender.

### 5.3.1 Innledende betraktninger om Hobbes' stat

Vurderingen av premissene 2 til 6 bygger på diskusjonen av Hobbes' stat i 4.1.1. Man kan dermed raskt konkludere med at Hobbes' stat vil ha en enehersker med absolutt makt; lovene bestemmes av suverenen, ingen kan begrense hans lovgivning og lovene kan ikke ramme ham<sup>24</sup>. Hobbes' stat er en rettsstat i den forstand at alle er like for loven, og lover kreves å være tydelige, nødvendige og kan ikke ha tilbakevirkende kraft. Den *vilkårligheten* som ofte tillegges Hobbes' stat, stammer fra det forhold at loven likestilles med suverenens vilje, og at den dermed teoretisk sett kan endres fort, kraftig og uforutsigbart. Suverenen er imidlertid "bundet" av visse fornuftsregler som Hobbes legger frem, bl.a. når det gjelder lover, og jeg antar med det at f.eks. prinsippene om at overtredere må ha tilstrekkelige forutsetninger for å gjøre seg kjent med lovene han straffes for å bryte – noe som selvsagt utelukker tilbakevirkende lovgivning – vil følges (Hobbes 1946:176-7).

Jeg har i vurderingen av hvorvidt Hobbes' stat er autoritær, i 4.1.1, poengtert at suverenen *ikke* er bundet i tradisjonell forstand, men finner det her allikevel nyttig å anse Hobbes' krav til suverenen som mer enn tomme besvergelses. Hobbes sier at "[d]ersom *monarken* utsteder dekret mot *naturlovene*, synder han mot seg selv", og at "enhver monark [...] er bundet av *naturlovene*" (Hobbes 2005a:102,103).

Jeg antar i det følgende at *suverenen faktisk regjerer slik Hobbes forutsetter*. Dette innebærer at suverenen kun gjør det som er nødvendig for å sikre oppnåelse av Hobbes' formål med staten:

Now all the duties of rulers are contained in this one sentence, *the safety of the people is the supreme law*. For although they who among men obtain the chiefest dominion, cannot be subject to laws properly so called, that is to say, to the will of men, because to be chief and subject, are contradictories; yet is it their *duty* in all things, as much as possibly they can, to yield obedience unto right reason, which is the natural, moral, and divine law (Hobbes 2005a:166).

Da jeg vil vurdere gyldigheten av Hobbes' resonnement, ønsker jeg å se hvorvidt det har svakheter selv i sin sterkeste form, med en "god" enehersker, som anerkjenner

---

<sup>24</sup>Lovene gjelder alle, men siden suverenen kan endre loven anser Hobbes ham for ikke å være bundet, da selv-binding for Hobbes er umulig dersom man selv kan fjerne bindingen (Hobbes 1946:173).

naturlovene og sin plikt til å fremme folkets sikkerhet<sup>25</sup>.

Hvilke rettigheter har så individer i Hobbes' stat? Hobbes avviker sterkt fra liberal statsteori, da Hobbes' individer forutsettes å ha overført alle sine rettigheter til suverenen (bortsett fra retten til å sikre sitt eget liv). Suverenen er bundet av statens formål, men det er ingen dekning for å hevde at Hobbes antar *ukrenkelige individuelle rettigheter* for suverenens undersåtter. Sikringen av ro og orden har absolutt prioritet over slike hensyn, og den likheten undersåttene har under Hobbes' suveren kan dermed forstås som en beskjeden likhet sett fra vårt ståsted – undersåttene er like i den forstand at alle er uten særlig sterke og omfattende rettigheter.

Det sentrale for vurderingen av Hobbes' stat som et alternativ til den moderne stat, er den moderne stats syn på rettsstaten som langt mer absolutt, den konstitusjonelle bindingen av makten (via lover *og* maktdeling) og det kjære forholdet til demokrati, både pga. legitimitet og som et gode i seg selv; demokratiet er en opplæringsinstitusjon, det gir mennesker en arena for deltagelse i håndteringen av viktige spørsmål og det legger til rette for debatt og meningskonkurranse.

### 5.3.2 Frihet og demokrati - *behov* eller *ønsker*?

Et *behov* kan forstås på flere måter, og jeg ønsker først å vurdere hvorvidt *frihet* bør ansees som et *universelt menneskelig behov* eller et noe svakere *ønske*. Jeg diskuterer her individuelle behov slik Maslow gjør i sin diskusjon av menneskelige behov. Det er videre mulig å diskutere behov i den forstand at samfunnet *behøver* visse institusjoner for å være stabile e.l., men det er altså ikke slike behov jeg her ser på.

Før jeg vurderer de øvrige premissene i resonnementet fra 4.2.2, må jeg utforme en konkret hypotese om hvordan menneskers behov kan tenkes å endres ved overgangen fra naturtilstanden til staten. Jeg vil derfor først se på hva Maslow og andre sentrale tenkere sier om *frihet*, for deretter å se kort på et begrep som ofte forbindes med frihet – *demokrati*.

Jeg vil ikke begi meg ut på debatten om positiv/negativ frihet. Jeg diskuterer *negativ* frihet slik Berlin (2002c) omtaler det<sup>26</sup>, ikke positiv frihet, som ifølge Fromm (1994:257)

<sup>25</sup>Innvendingene mot mangelen på skranker og tilitten til at én mann skal handle "rett" i en slik situasjon må ansees som svært sterke, men jeg vil ikke behandle disse problemstillingene spesielt her.

<sup>26</sup>"[...] liberty in this sense is simply the area within which a man can act unobstructed by others"

“består av den spontane aktiviteten til den totale og integrerte personligheten”. Hobbes definerer “liberty” som frihet *fra* – ikke *til* – som jeg har vist i premiss 6 i 4.2.2, noe som følger av at “det gode”, statens formål, er fullstendig subjektivt. Der argumenterte jeg også for at det ikke er selvmotsigende å snakke om frihet, selv om man har en mekanistisk filosofi som tilsier *nødvendighet*. Som Mill (2006:837) sier det: “Vi kan være fri, men andre kan allikevel ha grunner til å være helt sikre på hvordan vi vil benytte oss av denne friheten”; “[d]ersom vi kjente en person helt og holdent, og kunne avdekke all påvirkning han er utsatt for, kunne vi forutsagt hans adferd med like stor sikkerhet som med enhver fysisk hendelse” (Mill 2006:837).

### **Frihet som et grunnleggende menneskelig behov**

Under behandlingen av behovene for anerkjennelse og respekt i 3.3.1 har jeg sett på noen av Maslows tanker om frihet. Som jeg påpeker i 3.3.4 vektlegger Maslow og Hobbes dette begrepet svært forskjellig. En undergruppe av behov for *anerkjennelse og respekt*, består ifølge Maslow av et ønske om “styrke, oppnåelse, tilstrekkelighet, mestring og kompetanse, selvtillit i møtet med verden, samt uavhengighet og frihet” (Maslow 1987:63). Som vi så i kapittel 3 problematiserer han imidlertid selv at frihet inkluderes i denne listen:

Whether or not this particular desire is universal we do not know. The crucial question, especially important today, is: Will men who are enslaved and dominated inevitably feel dissatisfied and rebellious? We may assume on the basis of commonly known clinical data that people who have known true freedom (not paid for by giving up safety and security but rather built on the basis of adequate safety and security) will not willingly or easily allow their freedom to be taken away from them. But we do not know for sure that this is true for people born into slavery (Maslow 1987:63n).

Menneskers forhold til frihet må være sentralt i vurderingen av hvordan en god stat skal struktureres. Jeg finner det spesielt interessant å undersøke hvorvidt det er en grunnleggende konflikt mellom en absolutt suveren, slik Hobbes foreskriver, og menneskers grunnleggende behov. Grunnen til at muligheten for en slik konflikt er interessant, bør være åpenbar ut fra det resonnementet jeg har presentert i kapittel 4; premiss 2, 4 og 5, kan hevdes å være direkte feil dersom mennesker har et grunnleggende *behov* for en

---

(Berlin 2002e:169).

viss type frihet som ikke er tilgjengelig i Hobbes' stat, noe som helt klart vil kunne påvirke vurderingen av resonnementets konklusjon<sup>27</sup>.

Først og fremst vil jeg vektlegge Maslows plassering av frihet under *behov for anerkjennelse og respekt*. Som nevnt under 3.3.4 fremsetter Maslow også frihet som en *forutsetning* for grunnleggende behovstilfresstillelse (Maslow 1987:64). Det vil dog være merkelig å hevde at frihet er *nødvendig* for å stille sult eller sikre sitt eget liv, siden Maslow selv – i forbindelse med den sultne mann – hevder at “frihet, kjærlighet, fellesskapsfølelse, respekt, filosofi, kan feies til siden som verdiløst juggel, siden de ikke er i stand til å mette magen” (Maslow 1987:59)<sup>28</sup>.

Hva mener så andre tenkere om menneskers *behov* for frihet? Mill er en tenker det er naturlig å gå til når det er spørsmål om frihet. Jeg er imidlertid av den oppfatning at han i *On Liberty* er noe mer opptatt av å angripe totalitære og paternalistiske samfunnsformer enn å diskutere hvorvidt frihet er et grunnleggende og opplevd menneskelig behov i den forstand jeg diskuterer det her. Det er stor forskjell på om mennesker *ikke tåler å leve i ufrihet* og om de tåler det fint, men at *konsekvensene blir et “fattigere samfunn”*.

Mill snakker f.eks. om en “kjærlighet til frihet”, men det synes klart at han ikke forutsetter at dette gjelder mennesker generelt (Mill 1987b:280):

It is, perhaps, hardly necessary to say that this doctrine is meant to apply only to human beings in the maturity of their faculties. [...] Those who are still in a state to require being taken care of by others must be protected against their own actions as well as against external injury. For the same reason we may leave out of consideration those backward states of society in which the race itself may be considered in its nonage [...] Despotism is a legitimate mode of government in dealing with barbarians, provided the end be their improvement and the means justified by actually effecting that end. Liberty, as a principle, has no application to any state of things

---

<sup>27</sup>Premissene det her er snakk om dreier seg om det verste ondet, demokrati som mer utsatt for uroligheter enn enevelde og menneskers ekstreme risikoaversjon i beslutninger der alternativer som medfører økt fare for å dø inngår.

<sup>28</sup>Fra et effektivitetsbasert forsvar for frihet, som f.eks. Mill (1985) og von Mises (2007) kan sies å fremlegge, er det mulig å anse frihet som nødvendig for å skaffe best mulige og mest mulig effektive løsninger, som hjelper mennesker med å frembringe mat og sikkerhet. Denne funksjonen av frihet og demokrati vil jeg ikke forfølge i denne analysen, da Hobbes' idealstat har rom for fullstendig frie markeder osv., dersom man antar at dette er kompatibelt med en stats overlevelse (noe det bør være dersom det skal anees som positivt).

anterior to the time when mankind have become capable of being improved by free and equal discussion (Mill 1985:69).

Hvordan man definerer en “barbar” er uklart, men det kan synes som om Mill sikter til en manglende evne til å forbedre seg selv via verbale oppfordringer og råd; etter at denne evnen er oppnådd er det ikke lenger tillatelig å “forbedre” andre mot deres vilje (Mill 1985:69). Det er uansett snakk om frihet som aktuelt kun for mennesker (som individer og som rase) med evner av en viss modenhet. Dette varierer noe fra Maslows fremstilling, men det kan også tenkes at mennesker i “tilbakestående samfunnstilstander” rett og slett domineres av fysiologiske behov og sikkerhetsbehov. Alternativet er å tolke Mill slik at man antar at han snakker om “mennesker” som ikke har kommet til vårt stadium i evolusjonen – altså i praksis en annen rase – men jeg anser det som mest nærliggende å gå ut fra at han mener at mennesker som lever under primitive forhold, ikke har et behov for frihet slik han diskuterer det.

Berlin (2002c) behandler i større grad hvilken egenverdi frihet kan tenkes å ha, og hvorvidt frihet er et menneskelig *behov*:

Finally one may ask what value there is in liberty as such? Is it a response to a basic need of men, or only something presupposed by other fundamental demands? And further, is this an empirical question, to which psychological, anthropological, sociological, historical facts are relevant? Or is it a purely philosophical question, the solution of which lies in the correct analysis of our basic concepts, and for the answer to which the production of examples, whether real or imaginary, and not the factual evidence demanded by empirical enquiries, is sufficient and appropriate? (Berlin 2002a:50).

Det er her tre spørsmål som blir tatt opp: a) har frihet en iboende verdi, b) er frihet et grunnleggende menneskelig behov og c) hvordan finner man svar på disse spørsmålene? Jeg benekter ikke viktigheten av å behandle a), men vil forholde meg nøytral til dette spørsmålet her, idet jeg anser b) som det mest relevante for oppgaven. Menneskers fysiologiske behov kan til en viss grad avdekkes empirisk, men dette er vanskeligere med Maslows “høyere” behov; svaret på c) må sannsynligvis besvares både med empiri og filosofi; det vil f.eks. være mulig å observere menneskers reaksjoner på mangel på eller overflod av frihet, men disse observasjonene må med stor sannsynlighet suppleres

med filosofiske betraktninger om menneskers psykologiske behov<sup>29</sup>. Berlin spør selv hvorvidt empiri kan “motbevise” eventuelle oppfattede behov:

Could it be the case that if the evidence of the facts should go against us, we should have to revise our ideas, or withdraw them altogether, or at best concede that they – these propositions, if they are propositions – hold only for particular societies, or particular times and places, as some relativists claim? (Berlin 2002a:51)

For mitt formål i denne analysen er svaret på Berlins spørsmål *ja*, men dette innebærer ikke at jeg ikke ser verdien i filosofiske betraktninger som muligens ikke lar seg prøve empirisk. Når det gjelder den *universelle gyldigheten* av frihet som ideal ser man samme betraktninger hos Berlin som hos Mill:

It is true that to offer political rights, or safeguards against intervention by the State, to men who are half-naked, illiterate, underfed and diseased is to mock their condition; they need medical help or education before they can understand, or make use of, an increase in their freedom (Berlin 2002e:171).

Det er her igjen en antagelse om at andre behov har *prioritet* over frihet: “Den egyptiske bonden trenger klær og medisin før – og mer enn – personlig frihet [...]” (Berlin 2002e:172). Når det er snakk om en egyptisk bonde finner jeg det vanskelig å argumentere for at Berlin her sikter til pre-historiske mennesker som er fundamentalt forskjellige fra hva vi i dag anser som mennesker – *Homo sapiens*.

Berlin og Maslow (se sitat av Maslow (1987:63) ovenfor) ser også ut til å enes om en hypotese om at et *behov* for frihet til en viss grad kan være nært knyttet til en tidligere *erfaring* med frihet. Dette vil i så tilfelle innebære at ønsket om frihet (og evt. *demokrati*, mer om det senere) kan anees som en tillært appetitt. Det er dog ikke mulig å utelukke alternativet (at det er et grunnleggende og universelt behov med en sterk egenverdi), en mulighet Berlin (2002a) forfølger:

To this it is sufficient, perhaps, to say that those who have ever valued liberty for its own sake believed that to be free to choose, and not to be

---

<sup>29</sup>Maslow ser ut til å mene at empiri er overlegent andre kilder til vurderingen forskjellige *verdier*; “[i]t [at menneskers natur er god og iboende] means, for one thing, that it is no longer either necessary or desirable to deduce values by logic or to try to read them off from authorities or revelations. All we need do, apparently, is to observe and research. [...] The organism tells us what it needs (and therefore what it values) by sickening when deprived of these values and by growing when gratified” (Maslow 1987:102).

chosen for, is an inalienable ingredient in what makes human beings human; and that this underlies both the positive demand to have a voice in the laws and practices of the society in which one lives, and to be accorded an area [...] in which one is one's own master, a 'negative' area in which a man is not obliged to account for his activities to any man so far as this is compatible with the existence of organised society (Berlin 2002a:52).

Verdsettingen av frihet ser ut til å være forbundet med et spesielt syn på hva det innebærer å være menneske. Berlin hevder Locke, Mill og Tocqueville krever et minimum av frihet, da individer ved mangel på dette kan befinne seg "i en tilstand for snever selv for den minste påkrevde utviklingen av sine naturlige evner som alene gjør det mulig å forfølge – og faktisk begripe – de mange mål mennesker anser som gode eller hellige" (Berlin 2002e:171). Han ser videre på hva et eventuelt minimumskrav til frihet måtte innebære, og setter grensen der man ikke kan tåle mindre frihet uten å "bryte med essensen i menneskets natur"; hva denne essensen er, vil, som Berlin selv vedgår, "kanskje alltid være et spørsmål for uendelige debatter" (Berlin 2002e:173)<sup>30</sup>.

The dilemma is logically insoluble: we cannot sacrifice either freedom or the organisation needed for its defence, or a minimum standard of welfare. The way out must therefore lie in some logically untidy, flexible and even ambiguous compromise. Every situation calls for its own specific policy, since 'out of the crooked timber of humanity', as Kant once remarked, 'no straight thing was ever made' (Berlin 2002d:92).

Det er to dilemmaer forbundet med å anse frihet som et behov. Det første er det Berlin her peker på, nemlig det faktum at frihet og sikkerhet ikke i alle tilfeller er forenlig. Maslows behovshierarki er interessant her, siden de tre elementene som utgjør dilemmaet i Berlins sitat, befinner seg på ulike nivåer av hierarkiet. Et minimumsnivå av velferd er nødvendig for å tilfredstille fysiologiske behov, en viss grad av organisering er påkrevd for å tilfredstille sikkerhetsbehov, og til sist kreves frihet for å tilfredstille behov for anerkjennelse og respekt. Dersom vi – som Maslow – antar at mennesker "beveger seg" mellom disse behovene, og fokuserer svært sterkt på de behovene som er mest presserende for dem på et bestemt tidspunkt, er det lett å se at logikken blir "rotete" og løsningene "fleksible" og "tvetydige". Maslows perspektiv gjør det imidlertid enklere å se *hvorfor* dilemmaet er så vanskelig å løse, da *mennesker vektlegger elemen-*

---

<sup>30</sup>Det er verdt å merke seg at Berlin, på linje med Hobbes, setter den øvre grensen for frihet ved det som er forenlig med et "organisert samfunn".



*tene forskjellig avhengig av deres tilstand.* Vektleggingen vil også kunne påvirkes sterkt av situasjonen et samfunn befinner seg i, og vil måtte forventes å endres over tid. Krig og uro vil føre til borgere som krever mer sikkerhet, mens lange perioder med overflod vil føre til mennesker som har tilfredstilt de primære behovene, og som dermed vil fokusere på f.eks. frihet.

Det andre, og noe mer problematiske, dilemmaet oppstår som en følge av at Maslow plasserer behovet for frihet under behovet for “anerkjennelse og respekt”:

So much can I desire this [anerkjennelse som et verdifullt menneske], that I may, in my bitter longing for status, prefer to be bullied and misgoverned by some member of my own race or social class, by whom I am, nevertheless, recognised as a man and a rival – that is, as an equal – to being well and tolerantly treated by someone from some higher and remoter group, someone who does not recognise me for what I wish to feel myself to be (Berlin 2002e:203).

Det kan altså være en intern motsetning mellom anerkjennelse, respekt og frihet, noe som langt på vei er Fromms (1994) teori. Det moderne samfunnet gjør mennesker mer uavhengige og “fri”, men samtidig blir vi ”isolerte, ensomme og redde”; dette er frihetens “problem”, og Fromm mener det er avgjørende å være klar over begge disse virkningene av frihet (Fromm 1994:104). Berlin ser ut til å ha samme tanke:

The mass neurosis of our age is agoraphobia; men are terrified of disintegration and of too little direction: they ask, like Hobbes’s masterless men in a state of nature, for walls to keep out the raging ocean, for order, security, organisation, clear and recognisable authority, and are alarmed by the prospect of too much freedom, which leaves them lost in a vast, friendless vacuum, a desert without paths or landmarks or goals (Berlin 2002b:243).

Fromm hevder at mennesker har et iboende ønske om frihet, men at vi også har et “instinktivt ønske om underkastelse”, og spør seg: “finnes det en skjult tilfredstillelse i å underkaste seg, og hva er dens essens?” (Fromm 1994:5) Det er i dag bred enighet om at frihet er et positivt ladet begrep, men det er tilsvarende få bevis for at det er et universelt menneskelig behov på linje med et behov for mat og drikke. Uten empiriske bevis sitter jeg igjen med forskjellige teorier fra fremragende tenkere, og vil i slutten av denne gjennomgangen av frihet som et behov fremlegge en arbeidshypotese; denne vil fungere som en nyttig og plausibel forenkling som gjør det mulig å vurdere premissene

i Hobbes' resonnement.

Fromm, som Maslow, ser ut til å komme nærmest å kalle frihet et menneskelig behov, mens Berlin – en av vår tids største tenkere om emnet – ser ut til å støtte Hobbes' syn på prioriteringen av våre behov<sup>31</sup>:

All these facts show that the drive for freedom inherent in human nature, while it can be corrupted and suppressed, tends to assert itself again and again (Fromm 1994:xv).

Whatever his crudities and errors, on the central issue Hobbes, not Locke, turned out to be right: men sought neither happiness nor liberty nor justice, but, above and before all, security (Berlin 2002d:73).

### Frihetens fordeler

Jeg tar som nevnt ikke standpunkt til hvorvidt *frihet* har en egenverdi; der det som Mill (1985) påpeker kan påvises heldige konsekvenser av frihet, som kreativitet, vitenskap, og avdekking av verdifulle sannheter, er det her *disse* begrepene som ansees som gode. Dette åpner også for det faktum at f.eks. vitenskap ikke nødvendigvis er uløselig og eksklusivt knyttet til frihet. Det er dog åpenbart at frihet har vært medvirkende til overfloden vi i den vestlige verden i dag opplever.

Mills resonnement beror på at avdekkingen av de viktige idéene, og utviklingen av det geniale, er avhengig av frihet, noe han argumenterer sterkt for. Berlin påpeker imidlertid at forbindelsen mellom disse størrelsene kan være noe mer problematisk enn den synes hos Mill:

[. . .] the evidence of history tends to show (as, indeed, was argued by James Stephen in this formidable attack on Mill in his *Liberty, Equality, Fraternity*) that integrity, love of truth and fiery individualism grow at least as often in severely disciplined communities, among, for example, the puritan Calvinists of Scotland or New England, or under military discipline, as in more tolerant or indifferent societies; and if this is so, Mill's argument for

---

<sup>31</sup>Fromm hevder forøvrig at Hobbes vektla *makt* som det sentrale motiv for menneskelig adferd, men at motivet i realiteten kan forstås som drevet av et behov for sikkerhet: "To be sure, power over people is an expression of superior strength in a purely material sense. If I have the power over another person to kill him, I am "stronger" than he is. But in a psychological sense, *the lust for power is not rooted in strength but in weakness*" (Fromm 1994:160).

liberty as a necessary condition for the growth of human genius falls to the ground (Berlin 2002e:175).

Jeg vil ikke på noen måte avvise at frihet i sine mange former har svært mange beviselige fordeler, men Mill er i *On Liberty* i nærheten av å gjøre det jeg kritiserte Hobbes for i sin karakteristik av naturtilstanden i 5.2.3; at visse fordeler *kan* knyttes til frihet rettferdiggjør ikke alene at det knyttes *eksklusivt* til frihet.

### **Krever frihet demokrati?**

Det er viktig å påpeke at det ikke er en nødvendig forbindelse mellom den type frihet jeg her diskuterer og demokrati, noe Hobbes og Berlin er enige om:

Freedom in this sense [negativ frihet] is not, at any rate logically, connected with democracy or self-government. Self-government may, on the whole, provide a better guarantee of the preservation of civil liberties than other regimes, and has been defended as such by libertarians. But there is no necessary connection between individual liberty and democratic rule (Berlin 2002e:177).

If they suppose liberty to consist in this, that there be few laws, few prohibitions, and those too such, that except they were forbidden, there could be no peace; then I deny that there is more liberty in *democracy* than *monarchy*; for the one as truly consisteth with such a liberty, as the other (Hobbes 2005a:134).

Også Mill (1987a:194) benekter i sin behandling av forutsetningene for et permanent politisk samfunn at frihet er avhengig av demokrati; en av forutsetningene han fremsetter er at det i etableringen av staten fastsettes *noe* som er permanent og unntatt kritikk; dette kan f.eks. være – og Mill regner med at det i fremtiden (for Mill) *vil* være – prinsippene om individuell frihet og politisk og sosial likhet. Dette gjelder for monarki så vel som for demokrati (Mill 1987a:194). Vi har også sett at Mill fremholder enevelde som legitim statsform der mennesker ikke er i stand til å oppnå et tilfredsstillende samfunn gjennom demokrati (Mill 1985:69). I god Hobbesiansk ånd er Mill faktisk også enig i at demokrati har svakheter, noe jeg har gjennomgått under premiss 4 i 4.2.2:

No government by a democracy or a numerous aristocracy, either in its political acts or in the opinions, qualities, and tone of mind which it fosters, ever did or could rise above mediocrity except in so far as the sovereign many have let themselves be guided (which in their best times they always have done) by the counsels and influence of a more highly gifted and instructed *one* or *few*. The initiation of all wise or noble things comes and must come from individuals; generally at first from some one individual. The honour and glory of the average man is that he is capable of following that initiative; that he can respond internally to wise and noble things, and be led to them with his eyes open (Mill 1985:131).

Som jeg har vist i 5.2.2, i gjennomgangen av kommunitarianske samfunn, er det ikke bare autokratisk styre som er forbundet med alvorlige utfordringer når det gjelder frihet; Berlin (2002e) peker også på at mens en “liberalt anlagt despot” fint kan tenkes å gi sine undersåtter utstrakt frihet, kan et demokrati frarøve sine borgere det samme (Berlin 2002e:176).

I det følgende er det altså nødvendig å behandle frihet og demokrati som logisk separate begreper, selv om de ofte sammenstilles. Dette er vel å merke langt på vei en følge av at vi taler om negativ frihet. Berlin (2002e) påpeker at forskjellen mellom positiv og negativ frihet kommer tydeligst til syne når man spør *hva man er fri å gjøre* og *hvem som bestemmer hva en kan/ikke kan gjøre* (“hva som kontrolleres, ikke kilden til kontrollen”); “forbindelsen mellom demokrati og individuell frihet er langt svakere enn den har syntes for mange forkjempere for begge” (Berlin 2002e:176-7).

### **Frihet som en selvfølgelighet og jorden som flat**

At jorden er flat var en gang i historien like selvsagt som at frihet og demokrati er universelt ønsket og “slutten på historien”, som Fukuyama (2006) nylig hevdet med bemerkelsesverdig overbevisning. Få personer med kjennskap til menneskers tenkning – fra Confucius, via de berømte grekerne og hele veien opp til vår tid – kan med urokkelig sikkerhet hevde at dagens sannheter er endelige. Det synes klart at det har levd mange uhyre intelligente mennesker i løpet av de siste flere tusen årene, og så lenge mange av disse har akseptert teorier og oppfatninger vi i dag anser som uholdbare, er det nødvendig å innta en kritisk holdning til dagens oppleste og vedtatte sannheter.

[...] ages are no more infallible than individuals – every age having held many opinions which subsequent ages have deemed not only false but absurd; and it is as certain that many opinions, now general, will be rejected by future ages, as it is that many, once general, are rejected by the present (Mill 1985:78).

*Frihet og demokrati* kan sies å være to slike “sannheter” for oss i dag, og Hobbes fyller en viktig rolle i denne sammenheng: *hans filosofi fremmer enevelde som overlegent demokratiet og frihet som underordnet behovet for sikkerhet*. For at vi skal kunne utfordre hans resonnement er det nødvendig å gjøre det helt tydelig *hvorfor* frihet og demokrati er ønskelig. Det synes rimelig å anta at Mill ville oppfordre oss til å takke Hobbes for at han utfordrer våre oppfatninger:

If there are any persons who contest a received opinion, or who will do so if law or opinion will let them, let us thank them for it, open our minds to listen to them, and rejoice that there is someone to do for us what we otherwise ought, if we have any regard for either the certainty of the vitality of our convictions, to do with much greater labour for ourselves (Mill 1985:108).

Hobbes’ resonnement kan ikke avfeies med argumenter som at “Hobbes vektlegger ikke frihet”; man må vise *hvorfor* frihet er viktig, og hvorfor mangelen på frihet eventuelt svekker Hobbes’ resonnement og standpunkt. En av grunnene til å stille seg noe kritisk til frihet som et absolutt krav, er at denne oppfatningen er forholdsvis ny og særegen for vår tid og kultur:

In the second place, the doctrine is comparatively modern. There seems to be scarcely any discussion of individual liberty as a conscious political ideal (as opposed to its actual existence) in the ancient world. Condorcet had already remarked that the notion of individual rights was absent from the legal conceptions of the Romans and Greeks; this seems to hold equally of the Jewish, Chinese and all other ancient civilizations that have since come to light (Berlin 2002e:176).

Videre har historien vist at mennesker svært ofte har valgt bort sin frihet, selv i moderne tid:

We have been compelled to recognize that millions in Germany were as

eager to surrender their freedom as their fathers were to fight for it; that instead of wanting freedom, they sought for ways of escape from it; that other millions were indifferent and did not believe the defense of freedom to be worth fighting and dying for (Fromm 1994:3).

Dersom mennesker så skal ansees som suverene når det gjelder egne anliggender, må det problematiseres at de ofte velger “feil”<sup>32</sup>. Det hadde vært enklere å slå seg til ro med frihet som universelt ønsket dersom Mill ikke hadde funnet det nødvendig å kreve at “prinsippet om frihet kan ikke innebære et krav om at man skal være fri til å være ufri” (Mill 1985:173). Det at Mill ser det nødvendig å legge inn en slik begrensning av friheten kan tyde på at også han var klar over menneskers tendens til å søke seg vekk fra den type frihet som gjør dem usikre, isolerte og ulykkelige.

### Arbeidshypotese

Etter denne gjennomgangen av oppfatninger og vurderinger av *frihet* er det nå mulig å konkretisere en arbeidshypotese om menneskers forhold til dette begrepet. Jeg kan ikke hevde å ha avdekket nye og fundamentale sannheter om temaet utover Maslows betraktninger, men anser allikevel gjennomgangen som nyttig. For det første har jeg spesifisert hva slags frihet det her er snakk om, noe som er viktig, da definisjonene av frihet varierer svært mye. For det andre har jeg sett på en del sentrale bidrag for å få et bedre grunnlag for utformingen av arbeidshypotesen.

Jeg vil ta utgangspunkt i den mulighet at mennesker *ikke* har et fundamentalt og grunnleggende behov for frihet på linje med fysiologiske behov, sikkerhetsbehov og behov for kjærlighet og tilhørighet.

Når det gjelder behovene for anerkjennelse og respekt, synes det problematisk å betrakte *frihet* som en (universell) nødvendig betingelse for å tilfredsstille disse. Jeg har vist at det kan være svært sterke motsetninger mellom frihet og andre elementer under behovet for anerkjennelse og respekt; spesielt Fromm (1994:17) er sentral i argumentet for at menneskers instinkt for underkastelse må ansees som vel så sterkt som menneskers ønske om frihet. Det er imidlertid mulig å hevde at “behovet” for underkastelse

---

<sup>32</sup>“Speaking generally, there is no one so fit to conduct any business, or to determine how or by whom it shall be conducted, as those who are personally interested in it” (Mill 1985:180). “A plain husbandman is more prudent in affairs of his own house, than a privy-councillor in the affairs of another man” (Hobbes 1946:45).

er på behovsnivåene for sikkerhet eller tilhørighet, men det er fremdeles et faktum at svært mange mennesker i forholdsvis trygge og moderne samfunn har vist en oppsiktsvekkende evne til å underkaste seg, og faktisk til en viss grad ser ut til å trives i et slikt underkastelsesforhold til en suveren<sup>33</sup>.

[...] modern man still is anxious and tempted to surrender his freedom to dictators of all kinds, or to lose it by transforming himself into a small cog in the machine, well fed, and well clothed, yet not a free man but an automaton (Fromm 1994:xiii).

Selv om jeg ikke behandler frihet som et *behov* for hvert menneske, vil jeg anse muligheten for at frihet kan innta en *behovsaktig karakter*, der det har utviklet seg en *appetitt* for frihet, som meget plausibel.

Jeg vil med dette fremsette min arbeidshypotese, som er nettopp det: en tentativ fremstilling av et mulig syn på hvordan menneskers forhold til frihet kan forståes, som jeg anser som nyttig i den videre behandlingen av Hobbes' premisser:

#### **Arbeidshypotese:**

*Mennesker har intet universelt behov for frihet som er slik at de ødelegges dersom behovet ikke tilfredsstilles. Men de som har opplevd frihet som et sterkt samfunnsideal under fordelaktige forhold, som i dagens samfunn, kan utvikle et behovsaktig ønske om frihet.*

### **5.3.3 Hobbes' resonnement revurdert**

Det gjenstår nå å vurdere de premissene jeg har utledet fra Hobbes' filosofi i lys av arbeidshypotesen. Det tvinger seg med det frem en diskusjon om hva grunnlaget for denne vurderingen skal være: Hobbes' presentasjon av naturtilstanden og den påfølgende staten, eller dagens samfunn med individer som oss selv? Dette valget vil ha en avgjørende betydning for konklusjonen en kommer frem til, da *individer i Hobbes' naturtilstand ikke forutsettes å ha levd i et samfunn lik dagens, og heller ikke nødvendigvis vil komme i en situasjon der de etablerer en kjærlighet til frihet lik det man kan oppleve i dag*. Jeg vil her bygge videre på diskusjonen av naturtilstanden i 5.2.3.

---

<sup>33</sup>En slik vilje til å underkaste seg er forøvrig interessant i forbindelse med hvorvidt naturtilstanden er et Chicken-spill eller et Fangenes dilemma-spill.

Det forutsettes her – og jeg er klar over at dette er kontroversielt – at Hobbes' resonnement kan tenkes å være holdbart dersom man tar utgangspunkt i mennesker i en primitiv tilstand. Med "blanke ark" kan det synes som om Hobbes' politiske resonnement fremstår som gyldig – *gitt* at forutsetningen om at jeg omtaler en Hobbesiansk stat i sin beste utgave fremdeles står. Når jeg hevder at en slik utgave av hans filosofi kan synes å være gyldig, foretar jeg *ingen moralske vurderinger av en slik tilstand*; jeg benekter ikke at livet i en slik stat for oss kan ansees som fryktelig og umenneskelig, men jeg hevder en Hobbesiansk stat vil kunne være levedyktig, og at menneskene i en slik stat ikke nødvendigvis vil lide under mangelen på f.eks. institusjonalisert frihet og demokrati.

Jeg vil derfor se på premissene i resonnementet fra kapittel 4 i lys av *moderne mennesker*, og vil anta at disse menneskene faktisk svarer til de menneskene i arbeidshypotesen som har opplevd frihet som et ideal under fordelaktige forhold.

Premiss 1 er allerede behandlet i 5.2, og i naturtilstanden vil jeg fremholde at behovet for frihet uansett ikke vil gjøre seg gjeldende. Jeg vil derfor vurdere de resterende premissene, og spesielt de jeg antar vil påvirkes av mennesker med et sterkt ønske om frihet.

## **Vurdering av premiss 2: er døden det verste ondet?**

**Premiss 2** Død – krig, og spesielt borgerkrig – er det verste ondet for mennesker

Jeg har tolket Hobbes slik at døden og fare for døden er det verste Hobbes' individer kan forestille seg (jf. 2.3.4 og premiss 2 i 4.2.2). "Hobbes navnga trusselen om døden som basisen for sin psykologi", og "han var så overbevist om at hans overbevisning var universelt erfart at han gjorde dødsaversjon til grunnlaget for sin politiske filosofi" (Green 1993:113). Det er svært god dekning i Hobbes' tekster for å hevde at "frykten for døden spiller en avgjørende rolle i Hobbes' politiske filosofi" (Kavka 1986:82). Det kan dermed synes som om mange av de bidragene om Hobbes – kanskje spesielt fra økonomer – der f.eks. mangel på investering er hovedfokuset, går for langt i å fjerne seg fra hva Hobbes faktisk sier.

Slike bidrag er på ingen måte uten nytte, men jeg anser dem som mer interessante i den forstand at de viser *enkelte aspekter* ved f.eks. naturtilstanden, enn som sentrale



og uttømmende bidrag om Hobbes' tanker og intensjoner<sup>34</sup>.

Det er fullt mulig å angripe Hobbes når det gjelder hans fokus på døden (men det er vanskeligere å ignorere det og fremdeles hevde at man snakker om Hobbes' filosofi), og det har også blitt gjort svært ofte; Green (1993:113) anerkjenner den sentrale plassen frykten for døden har i Hobbes' filosofi, men mener at den "ikke frembringer reaksjoner som er like ensartede som han fastslår". Kavka (1986) er en av de viktigere bidragsyterne innen moderne forskning på Hobbes, og hans oppfatning er at Hobbes var for uklar når det gjelder menneskers dødsaversjon:

[...] Hobbes is simply unclear and inconsistent on the matter [dødsaversjon]. We are therefore free, in reconstructing his arguments, to adopt the more realistic view of peoples' attitudes toward death: that death is nearly universally shunned as a very great, but not necessarily the greatest possible, worldly evil (Kavka 1986:81).

Kavka mener altså at vi er fri til å være mer realistiske siden Hobbes er uklar, og han hevder videre at Hobbes' vektlegging av overlevelse "ikke binder ham til å ignorere vurderinger av livskvalitet for kun å vektlegge et langt liv" (Kavka 1986:82). Jeg tar meg som nevnt innledningsvis mindre friheter enn Kavka, da jeg er mer opptatt av å "oppdatere" Hobbes' opprinnelige resonnement enn å konstruere et lignende resonnement, der alle "feil" rettes, fra bunnen.

On the other hand, though death is the greatest of all evils (especially when accompanied by torture), the pains of life can be so great that, unless their quick end is foreseen, they may lead men to number death among the goods (Hobbes 1991a:48).

Sitatet ovenfor tolkes ofte til å bety at døden for Hobbes er det verste *enkeltondet*, men at mange mindre onder altså kan fremstå som så store at døden ikke lenger er verst; Hobbes (1997:116) hevder dog i *A Dialogue Between a Philosopher and a Student of the Common Laws of England* at enhver som tar sitt eget liv må være *gal* – ikke *compos mentis*<sup>35</sup>.

---

<sup>34</sup>"That Hobbes offers three different motives that would make the natural condition one of war may be seen as a strength of the theory, rather than a lack of economy. For it allows for greater human diversity than would reliance on a single motive, and it overdetermines the result" (Hoekstra 2007a:116).

<sup>35</sup>"[...] for naturally, and necessarily the Intention of every Man aimeth at somewhat, which is good to himself, and tendeth to his preservation: And therefore, methinks, if he kill himself, it is to be presumed that he is not *compos mentis*, but by some inward Torment or Apprehension of somewhat

Om døden *er* det verste ondet, eller *oppfattes* som det verste ondet er to forskjellige ting. Jeg opererer her med antagelsen om at døden for en organisme biologisk sett *er* det verste som kan inntreffe; Hobbes ser ut til å ha en lignende tanke, og han ser for seg at dette gjør seg gjeldende i menneskers adferd og ønsker via “naturlige impulser” (Hobbes 2005a:8). Det er også forskjell på å anse døden som det objektivt verste som kan skje, og det at individer alltid anerkjenner eller innser dette. Når Hobbes så sier at fryktelige omstendigheter kan lede mennesker til å anse døden som et gode, må det – ifølge hans egen logikk om godt og ondt – ansees som godt for individet, med mindre personen da er forvirret, eller rett og slett gal.

Hoekstra (2007a) har i mine øyne en god tilnærming til Hobbes, da han anser Hobbes’ premisser som bygget på “en rekke karakteristikk som er tilstrekkelig universelle til å være relevante for konklusjonen”; hans resonnement bygger på premisser som er “sanne generelt sett men ikke universelt” (Hoekstra 2007a:116). Spørsmålet blir med en slik tilærming hvorvidt *nok* personer f.eks. anser *frihet* som viktigere enn *livet*; det kreves ikke at *alle* mennesker har identiske oppfatninger og preferanser. Denne tilnærmingen fremstår som svært plausibel, og stemmer godt overens med min gjennomgang av Hobbes’ syn på mennesket i kapittel 2.

Man kan gjerne hevde at et liv i ufrihet er verre enn døden, men ville man faktisk tatt sitt eget liv dersom man ble fratatt sin frihet? Ord og litterære forsvar for frihet som det suverene godet synes billige sammenlignet med handling; handling fremstår her som sentralt, da ubevisst motivasjon ansees som vel så viktig som det et individ bevisst søker og ønsker. Det beste vi har er erfaringer som f.eks. Fromms omtale av tyskeres vilje til å oppgi sin frihet, og ut fra hva jeg vet er det svært få tilfeller av masse-selv-mord eller annen selv-destruktiv adferd, selv der de mest totalitære og undertrykkende regimer regjerer (Fromm 1994:3).

Fromm (1994) er klar på at frykten for å dø spiller en viktig rolle også for mennesker i dag:

Christianity has made death unreal and tried to comfort the unhappy individual by promises of a life after death. Our own era simply denies death and with it one fundamental aspect of life. [...] the fear of death lives an illegitimate existence among us. It remains alive in spite of the attempt to deny it [...] (Fromm 1994:245).

---

worse than Death, Distracted” (Hobbes 1997:116).

Utover kristendommen har mange filosofer også forsøkt å nedtone de tragiske aspektene ved døden, og da kanskje spesielt stoikerne:

We should apprehend, too, the nature of death; and that if only it be steadily contemplated, and the fancies we associate with it be mentally dissected, it will soon come to be thought of as no more than a process of nature (and only children are scared by a natural process) – or rather, something more than a mere process, a positive contribution to nature’s well-being (Aurelius 1964:49).

Slike doktriner, og alle alternativer som forsøker å dempe menneskets frykt for døden ved å skape håp om noe *mer*, kan muligens ha en sterkere effekt enn hva Hobbes faktisk tillegger dem. Det er et interessant spørsmål, som kan utforskes videre. I forbindelse med Hobbes i denne oppgaven anser jeg imidlertid en biologisk forståelse av døden som det verste ondet, og en forutsetning om at dette gjør seg gjeldende gjennom individers bevissthet og underbevissthet, som tilstrekkelig.

På bakgrunn av denne gjennomgangen synes det rimelig å beholde premisset slik det står, men jeg vil se videre på *hvordan mennesker faktisk forholder seg til dette ondet* under vurderingen av premiss 5 nedenfor.

### Vurdering av premiss 3: Hobbes’ absolutisme

**Premiss 3** Makt *kan* ikke deles, kun midlertidig delegeres

[H]e treats any division of sovereignty as tantamount to this conflict between sovereigns or would-be sovereigns; anything short of unified and absolute sovereignty is, or at least threatens quickly to become, a state of war (Hoekstra 2007a:111).

Dette premisset er en av svakhetene jeg mener man må leve med dersom man skal diskutere Hobbes’ filosofi. Han formulerer ikke dette som en svak anbefaling som kan anees å være relativ i forhold til hans tid, men som et *absolutt krav*. Jeg har fjernet meg fra tradisjonelle tolkninger av Hobbes der jeg mener å ha dekning for det, men det finner jeg altså ikke her. Den absolutte og uinnskrenkede makt er en av hovedidéene til Hobbes – å se bort fra dette vil være som å anse Mill som en nyttefilosof uten et *meget* skarpt forsvar for individers frihet.

Det må imidlertid påpekes at en del av kritikken mot Hobbes på dette punktet kan være basert på en misforståelse; Hobbes anser makt som *logisk udelelig*, men benekter ikke at mennesker allikevel vil prøve å dele eller midlertidig frasi seg denne makten (Oakeshott 1946:lvii). Morgenthau (1985:341) er like klar som Hobbes på at suverenitet er logisk udelelig: ”Dersom suverenitet betyr den øverste autoritet [noe det gjør hos Hobbes] følger det logisk at to eller flere enheter [...] ikke kan være suverene innen samme tid og sted. Han som er suveren er med logisk nødvendighet over alle andre”.

Ta Norge i dag: man vil kunne hevde at vi ikke har noen “suveren”, men at det er grunnloven som ligger til grunn og har prioritet over alt annet. Den sier også at makten er delt mellom lovgivende, utøvende og dømmende makt. Det er enkelt å anse dette som hele sannheten, og dermed anta at man i Norge har tre separate maktsentra, som kontrollerer hverandre. *Men*, grunnloven er kilden til all denne makten, og spesifiserer hvem som får “låne” den. Siden da grunnloven faktisk kan endres, hva skiller de med denne makten til å endre fra Hobbes’ suveren? Hobbes ville muligens sagt “vett”, da han anser det å ha en slik tilsynelatende åpenlys konkurranse om makten som ødeleggende, men ellers kan det slik sett være lite som skiller dagens Norge fra Hobbes’ idealer.

Også i dag, i nasjonalstatens og *det representative demokratiets* tidsalder, er folkets suverenitet – at folket er øverste myndighet – et kjennetegn på demokrati. I et representativt demokrati fremstår nok den lovgivende myndighet som den øverste. Forholdet mellom folket og den lovgivende myndighet kommer imidlertid straks til forgrunnen når det blir spørsmål om hvor tilfredsstillende en forfatning og et regime er, sett fra et demokratisk synspunkt (Midgaard 2004:24).

Kavka (1986:166) har, som Oakeshott, en “mild” forståelse av Hobbes’ absolutisme, som innebærer at absolutt makt ikke nødvendigvis betyr at én person, eller én forsamling, har all makt *direkte*<sup>36</sup>.

This interpretation of the Hobbesian notion of sovereignty renders the idea of absolute sovereignty less otiose than it appears at first glance, for now all unlimited and undivided sovereign power amounts to is a single person or assembly being in a position to appoint all those who exercise the various

---

<sup>36</sup>Kavka (1986) mener dog i motsetning til Hobbes og Morgenthau at en suveren med absolutt makt ikke er nødvendig for å hindre et tilbakefall til naturtilstanden. “Konstitusjonelle begrensninger på en regjerings makt er forenlig med varig ro, og det samme gjelder delingen av makten mellom adskilte styringsenheter” (Kavka 1986:167).

necessary powers of government (Kavka 1986:167).

Tuck (2007:187-8) viser på lignende måte hvordan Hobbes' suveren ikke er avhengig av å administrere den daglige "driften" av staten:

So Hobbes assumed that under normal circumstances government and its administration would be separated, and that this was the most 'natural' state of affairs. [...] Hobbes was also quite prepared to envisage a situation where the democratic sovereign (unlike at Rome or Venice) had virtually no power to intervene in the daily running of the state (Tuck 2007:187-8).

Det bør videre vurderes hvorvidt en *konstitusjon*, som spesifiserer og tydeliggjør lover og regler, og med det skaper større forutsigbarhet og fører til fred, nødvendigvis må godtas på bakgrunn av Hobbes' øvrige resonnement, dersom det viser seg å være best egnet til å sikre fred. Hobbes' eneste kriterier for å anbefale den stat han gjør er a) nødvendighet og b) hva som best gagnar folkets sikkerhet. Dersom en form for konstitusjonell maktdeling er *mulig*, og har vist seg å være mest stabil – *best for folkets sikkerhet* – synes det nødvendig at dette inngår i konklusjonen på et Hobbesiansk resonnement à la det jeg har fremsatt i kapittel 4.

Poenget om hvorvidt suverenen er over loven er en annen sak det er problematisk å støtte Hobbes i. *Men*, også i Norge i dag har "suverenen" muligheten til å endre lover, og dermed kan de ifølge Hobbes sies å være over loven, samtidig som de allikevel kan straffes av den. Oakeshott (1946:lvii) hevder at Hobbes' stat, med en absolutt suveren, ikke er et despoti, men en rettsstat. Dette synet deles bl.a. av Goldsmith (1996:283), som videre oppsummerer en relevant diskusjon fra *A Dialogue Between a Philosopher and a Student of the Common Laws of England* (Hobbes 1997):

Hobbes also has them agree on a suitably qualified version of his absolutism. When the Lawyer points out that statutes apparently restrict the powers of the sovereign, for example, to raise money without parliamentary consent and to send troops outside the kingdom, the Philosopher contends that all statutes restricting the powers of the monarch implicitly contain a reservation qualifying the restriction, viz., unless that action is necessary for the preservation of the kingdom. Although the Philosopher agrees that kings ought to respect such restrictive clauses, he nevertheless insists, and the Lawyer agrees, that they are not enforceable (Goldsmith 1996:290).

Det er imidlertid gode grunner for å hevde at Hobbes faktisk tok feil når han hevdet, uten særlige forbehold, at den eneste staten som ikke ville føre oss tilbake til naturtilstanden var den absolutte:

After all, we *know*, I think, that Hobbes's substantive theory can't be, in *general*, correct; since *constitutional democratic* institutions that violate his conditions for the Sovereign *have actually* existed and have not been noticeably less stable and orderly regimes than the kind of absolutism that Hobbes favored (Rawls 2007:85).

Allikevel, denne svakheten (eller særegenheten om man vil) mener jeg som sagt man må leve med i et Hobbesiansk resonnement. Selv om det *kan* hevdes at Hobbes' absolutisme forsvinner dersom man bruker hans sirkelargument<sup>37</sup>, er dette en lite nyttig, og tvilsom, fremgangsmåte. Hobbes er meget klar på at suverenen må være samlet og handlekraftig, *men*, det er allikevel viktig å være klar over at Hobbes' absolutisme ikke nødvendigvis innebærer at suverenen selv til enhver tid aktivt forvalter den makten han besitter:

Now the *power* in all kinds of government is equal [...] the conveniences or inconveniences of any government depend not on him in whom the authority resides, but on his officers [...] (Hobbes 2005a:141).

Hobbes forøvrig klar over det kontroversielle ved å kreve absolutt makt til suverenen:

This same supreme command and absolute power, seems so harsh to the greatest part of men, as they hate the very naming of them; which happens chiefly through want of knowledge, what human nature and the civil laws are; and partly also through their default, who, when they are invested with so great authority, abuse their power to their own lust (Hobbes 2005a:87).

At herskere har blitt korrumpert av makt var tydeligvis også klart for Hobbes, men han synes ikke å se det som en *nødvendighet* at makt korrupperer:

But it is the fault of the ruler, not of the government. For all the acts of Nero are not essential to monarchy [...] (Hobbes 2005a:133)

---

<sup>37</sup>“Sometimes, however, Hobbes uses a familiar *regress argument* for the Sovereign: such an unlimited power must exist, for if the supposed Sovereign is limited, it must be limited by some superior agency and then *that* agency is unlimited” (Rawls 2007).

Det er i denne forbindelse meget interessant å lese Hobbes' utlegning om hva som leder et menneske til *kriminalitet*, der han fokuserer spesielt på forventningen om å unnsnippe straff som et sentralt poeng (Hobbes 1946:193-4). Hobbes anser det som skjerpene dersom en lovbrøyer er mektig og man antar at han forventet å unnsnippe straff (Hobbes 1946:198). Vi har sett at suverenen ifølge Hobbes ikke kan anklages for å bryte loven, da den er hans vilje; det er uansett verdt å merke seg at Hobbes i andre sammenhenger anser det å ikke måtte stå til ansvar for sine handlinger som en sentral kilde til "urett". Hobbes' egen diskusjon om dette er imidlertid mildere enn den han gjengir i sin oversettelse av Aristoteles' *The Art of Rhetoric*; der hevdes det at de er mest tilbøyelige til åpenlyst å gjøre skade som bl.a. kan unnsnippe eller utsette en rettsak, de som kan bestikke dommerne, de som kan unnsnippe å betale bøter, de som ofte har sluppet unna tidligere og de som kan få benådning (Hobbes 1966:443). Maktens korrumperte effekt var altså ikke ukjent for Hobbes, selv om han ikke diskuterer problemene forbundet med dette i forbindelse med suverenen.

Man kan trygt konkludere med at Hobbes ikke tilbyr spesielt overbevisende argumenter for at hans absolutisme er ønskelig eller ufarlig. Man har allikevel med en teori som Hobbes' en motsats til dagens konvensjon, som har *visse* fordeler det kan være nyttig å vurdere; med absolutt makt er suverenen meget handlekraftig, og mulighetene for ødeleggende maktkamp kan sies å reduseres.

### **Vurdering av premiss 4 og 5: demokrati som ustabil og mennesker som ekstremt sikkerhetsfokuserede**

**Premiss 4** Et demokrati er mer utsatt for uroligheter enn et enevelde

**Premiss 5** Mennesker er ikke villige til å ofre sikkerhet for andre goder

Jeg har i gjennomgangen av premiss 4 og 5 i 4.2.2 beskrevet hva Hobbes faktisk sier om demokrati og menneskers manglende vilje til å tåle økt usikkerhet i bytte mot f.eks. demokrati. Jeg er dog, som tidligere nevnt, av den oppfatning at en isolert lesning av Hobbes' forskjellige diskusjoner kan føre til problemer, og slike problemer synes å gjøre seg gjeldende i forbindelse med Hobbes og demokrati.

Hobbes' forsvar for enevelde synes overveldende, men jeg vil her hevde at det er tre forhold som taler for å revidere premiss 4. For det første fremstår Hobbes' forsvar for enevelde som lite oppriktig. For det andre er det grunn til å spørre seg om det Hobbes

kaller demokrati er det samme som hva vi i dag mener med begrepet. For det tredje er det uunngåelig å bemerke at moderne historie har vist at demokratier (i alle fall i visse situasjoner) er vel så stabile som enevelder, og ut fra Hobbes' eget resonnement synes det opplagt at styreform bør velges ut fra hva som er best egnet til å skape fred. Jeg vil i det følgende, der ikke annet er nevnt, omtale demokrati og aristokrati under ett.

Hobbes' argumenter for enevelde ble gjennomgått under premiss 4 i 4.2.2. Det består i korte trekk i at: 1) det individuelt gode for suverenen og det kollektivt gode ligger nærmere hverandre ved enevelde enn demokrati, 2) en eneherker har bedre forutsetninger for å få god rådgivning enn en forsamling, 3) forsamlinger er mindre konsekvente enn en eneherker, 4) eneveldet er mindre utsatt for interessekonflikter, 5) ulempene ved favorisering av suverenums nærmeste er mindre med en eneherker enn ved forsamlingsstyre, siden han er én, mens man i en forsamling er mange og 6) mens monarkiet kan oppleve at makten tilfaller en umyndig, er demokratier og aristokratier ifølge Hobbes konstant i en slik situasjon, og eneherkeren har den fordel at han selv kan utnevne hvem som skal være kurator, noe som forhindrer uheldig konkurranse om denne posisjonen.

Hobbes gir inntrykk av at han skal *sammenligne* monarki og forsamlingsstyre i kapittel 19 i *Leviathan*. Det som fremstår som et åpenbart *forsvar* for enevelde utgis altså for å være nøytralt, og foretas ifølge Hobbes for å avdekke hva som er best egnet til å sikre fred og folkets sikkerhet (Hobbes 1946:122). Det er vanskelig å fri seg fra den følelsen at Hobbes i dette forsvaret for enevelde er sterkt interessert i å nå den konklusjonen at enevelde er den beste styreform, noe som svekker hans resonnement.

Utover det faktum at konklusjonen fra sammenligningen ikke kan sies å være like entydig som Hobbes fremstiller den, virker det som om Hobbes i sammenligningen utelater elementer som ville tale for forsamlingsstyre. F.eks. utelates problemet med arveretten i eneveldet fra selve sammenligningen, og behandles som et eget spørsmål, *etter* at Hobbes har konkludert med at enevelde er det beste alternativet. I en reell sammenligning synes det rimelig å forvente en viss balanse, og Hobbes er, som det følgende sitatet fra *De Cive* viser, opplagt klar over at problemet med arverett kun gjelder ved enevelde:

Now the right by which they may be continued, is that which is called the right of *succession*. Now because in a *democracy* the supreme authority is with the *people*, as long as there be any subjects in being, so long it rests



with the same person; for the people hath no successor. In like manner in an *aristocracy*, one of the nobles dying, some other by the rest is substituted in his place; and therefore except they all die together, which I suppose will never happen, there is no succession. The query therefore of the right of succession takes place only in an *absolute monarchy*. For they who exercise the supreme power for a time only, are themselves no *monarchs*, but *ministers* of state (Hobbes 2005a:122).

Når det gjelder Hobbes' omtale av styreform i *Leviathan*, synes det som om han her har rendyrket forsvaret for eneveldet, og utelatt åpenbare innrømmelser som synes naturlige i en sammenligning av styreformer. I *De Cive* kommer faktisk Hobbes med en slik innrømmelse:

But if the *people* in a democracy would bestow the power of deliberating in matters of war and peace, either on one, or some very few, being content with the nomination of magistrates and public ministers, that is to say, with the authority without the ministration; then it must be confessed, that in this particular *democracy* and *monarchy* would be equal (Hobbes 2005a:140).

Dette leder meg til det andre problemet med Hobbes' forsvar for eneveldet: hva slags demokrati har Hobbes i tankene? I *De Cive* definerer han *demokrati* som en regjering der "makten tilhører et råd der hver innbygger har en rett til å stemme" (Hobbes 1991b:192). Hobbes *kan* faktisk ikke diskutere vår tids demokrati som en reell mulighet, da hans absolutisme forhindrer blandede styreformer (Hobbes 2005a:95,96):

[...] I have already sufficiently proved (chapter XVIII) that all governments which men are bound to obey, are simple and absolute. [...] and in aristocracy and democracy, but one supreme assembly, with the same power that in monarchy belongeth to the monarch, which is not a mixed, but an absolute sovereignty (Hobbes 1946:361).

Hobbes diskuterer i sitt forsvar for eneveldet i stor grad demokratier der den suverene forsamlingen faktisk møtes for å drøfte og avgjøre spørsmål, noe som fremgår fra hans sammenligning; denne direkte formen for demokrati kan ikke sies å være spesielt representativ for vårt moderne demokrati. Tuck (2007:188) peker på at demokrati for Hobbes *kan* godtas, dersom den suverene forsamlingen "sover" og overlater all administrasjon til administratorer, men det er ikke et slikt demokrati Hobbes ser ut til å

sammenligne eneveldet med. Jeg nøyer meg med dette i forhold til det andre punktet, da det fremstår som opplagt vi i dag ikke mener det samme som Hobbes når vi snakker om *demokrati*.

Midgaard (2004:27) peker på at Platon, i direkte motsetning til Hobbes, anså rene styreformer som ustabile, og derfor foretrakk blandede styreformer. Platons syn på styreformenens stabilitet leder oss så til det tredje punktet: hva har historien vist når det gjelder hvilken styreform som er best egnet til å skape fred og sikkerhet? Jeg har allerede pekt på at Rawls (2007:85) mener historien har bevist at Hobbes tok feil på dette punktet, noe det synes å være tilnærmet universell enighet om:

Later history, we are tempted to say, has disconfirmed the assertions. Systems of government with limited and divided powers have proven to be long-lasting and stable and have provided security to many generations of their citizens (Kavka 1986:166).

Det synes, tross alt, klart at Hobbes ikke *utelukker* muligheten for demokrati (selv ikke slik han kjenner det), men at han fraråder det sterkt på bakgrunn av det han oppfatter som enveldets overlegenhet.

[...] no great popular commonwealth was ever kept up, but either by a foreign enemy that united them; or by the reputation of some eminent man amongst them; or by the secret counsel of a few; or by the mutual fear of equal factions; and not by the open consultations of the assembly (Hobbes 1946:172).

Hobbes er som man ser tydelig klar over at demokratier *har* fungert, men det kreves da spesielle forutsetninger. Min tolkning av Hobbes når det gjelder hvorvidt demokrati frarådes eller utelukkes som mulig styreform er heller ikke uten presedens:

[T]he civil society must be either a Monarchy or an Aristocracy or a Democracy. Which it is to be, is solely a question of which is most likely to produce the peace for which civil society is instituted [...] no constitution is without its defects. Which is best can be decided only by prudence [...] (Oakeshott 1946:xxxix).

To give Hobbes the greatest possible credit in his view of popular rule, note that it is clearly consistent with his general view of the rightness of a particular form of government. A form of government is right for a given

society and time if it is currently working (Hardin 1991:174).

Hobbes sier faktisk rett ut, at der et demokrati eksisterer, er dette den beste styreform:

And of the three sorts, which is the best, is not to be disputed, where any one of them is already established; but the present ought always to be preferred, maintained, and accounted best; because it is against both the law of nature, and the divine positive law, to do any thing tending to the subversion thereof (Hobbes 1946:361).

Det synes med dette utsagnet tydelig at Hobbes' forsvar for eneveldet *må* anees som et resultat av at han selv levde i et kongedømme, og i tillegg levde i en periode da han faktisk erfarte mange problemer med demokratiet:

Hobbes singles out democracy for criticism, but his 'presentism' suggests the possibility that on this subject he may not think what he says or say what he thinks (Hoekstra 2007b:216).

Hva er så Hobbes' grunnlag for å anbefale enevelde? *Utelukkende dets – slik Hobbes ser det i sin situasjon – overlegne evne til å skape sikkerhet!* Det er altså *pragmatiske* årsaker til at enevelde foretrekkes, og grunnen til at Hobbes ser enevelde som det beste alternativet er igjen basert på hans syn på menneskers natur og adferd, samt hans erfaringer, som ikke kan anees som representative for de erfaringer vi har i dag.

Disse tre momentene fører igjen til at jeg ser det som tvingende nødvendig å revidere Hobbes' anbefaling av en gitt styreform. Jeg har i denne delen av oppgaven også "utvidet" den forståelsen av mennesket jeg utledet fra Hobbes' *Leviathan*. Jeg mener jeg ikke har fjernet meg fra Hobbes' tanker, men at jeg har strukturert menneskers behov på en måte Hobbes selv var i nærheten av, men som man finner et langt renere uttrykk for hos Maslow, som vist i 3.3. Menneskers ønsker endres stadig, og nye behov kommer i forgrunnen når andre blir tilfredsstilt. Men, der Hobbes ikke forutsetter et mønster i hvilke behov som vil ta over – og der sikkerhet tilsynelatende alltid har en sentral plass – får vi med Maslow et hierarki som tilsier at mennesker etterhvert som de grunnleggende behovene tilfredsstilles, ikke lenger vil være like opptatt av sikkerhet.

Dette i seg selv har ikke *nødvendigvis* så store konsekvenser, da jeg mener det ikke kan vises at mennesker har et universelt behov for demokrati eller en spesiell form for frihet. *Men*, jeg vurderer her mennesker som beskrevet i arbeidshypotesen – mennesker med

kjennskap til, og en form for kjærlighet til, frihet – og dette spesialtilfellet kan være veldig interessant. Machiavelli (2003) har klare oppfatninger om at mennesker som er vant med visse idealer kan skape problemer for en suveren:

Anyone who becomes master of a city accustomed to freedom and does not destroy it may expect to be destroyed by it; for such a city may always justify rebellion in the name of liberty and its ancient institutions (Machiavelli 2003:28).

Siden *frihet* i arbeidshypotesen er implementert som et *behovsaktig ønske*, med bakgrunn i et forhold til begrepet som noe positivt, anser jeg det her som nyttig å inkludere det samme synet på *demokrati*. Jeg ser dog som sagt ikke den nødvendige koblingen mellom negativ frihet og demokrati, spesielt ikke et absolutt demokrati som her diskuteres. Menneskers forhold til begrepene forstås som deres subjektive oppfatning, og siden den type individer jeg ønsker å vurdere er ment å representere mennesker med erfaring med liberale demokratier, som det vi i dag opplever f.eks. i Norge, synes det også tillatelig å introdusere et kjært forhold til demokrati i arbeidshypotesen.

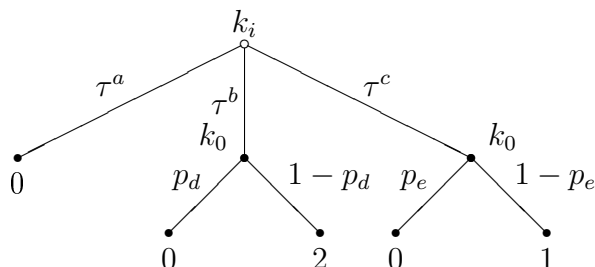
#### **Reformulert arbeidshypotese:**

*Mennesker har intet universelt behov for frihet eller selv-styre som er slik at de ødelegges dersom behovene ikke tilfredsstilles. Men de som har opplevd frihet og selv-styre som sterke samfunnsidealene under fordelaktige forhold, som i dagens samfunn, kan utvikle et behovsaktig ønske om frihet og selv-styre.*

Hobbes og demokrati er relativt lite diskutert, noe jeg finner noe merkelig. Selv om Hobbes beskrev demokrati som uholdbart, synes det plausibelt at hans *systematiske resonnement* vil føre til en konklusjon som tilsier at demokrati er beste styreform i visse tilfeller, spesielt dersom det har vist seg å være mest stabilt. Alternativet til å godta at at Hobbes' resonnement fører til demokrati som styreform for moderne mennesker er å hevde at demokrati *ikke* er bevist å være best for oss, eller at Hobbes hadde en irrasjonell forkjærlighet for enevelde som ville være umulig å endre selv med relativt klare bevis. Ingen av disse alternativene fremstår som spesielt plausible.

Min påstand er a) at mennesker med et kjært forhold til demokrati og frihet vil tillegge disse begrepene en høyere verdi enn de som ikke har det samme forholdet til begrepene, og b) at mennesker som lever et godt og forholdsvis tilfredsstillt liv *ikke* vil fremvise den ekstreme risikoaversjonen i forhold til utrygghet som jeg har antatt i gjennomgangen av Hobbes' resonnement. Dessuten er det rimelig å anta at mennesker som har opplevd

et stabilt demokrati ikke oppfatter risikoen ved demokrati eller aristokrati som like presserende som Hobbes ser ut til å anta at den er.



Figur 5.7: Gjengivelse av figur 4.4. Aktør  $i$  forestilles å ha valget mellom ingen stat ( $\tau^a$ ), forsamlingsstyre ( $\tau^b$ ) og enevelde ( $\tau^c$ ). Etter første valg avgjøres utfall av sannsynlighet for krig,  $p_d$  ved demokrati og  $p_e$  ved enevelde.

Under premiss 4 og 5 i 4.2.2 har jeg fra før diskutert figur 5.7. Jeg hevdet der at Hobbesianske individer må forventes å velge  $\tau^c$ . Dette gir best sikkerhet der Hobbes' forutsetning om at  $p_e < p_d$  er gyldig, da en eventuell fordel ved demokrati for Hobbes ikke kan sies å veie opp for den ekstra usikkerheten dette medfører.

Begge faktorene i den foregående vurderingen endres dersom vi antar at mennesker med en kjærlighet til frihet og demokrati, i en stabil og tilfredsstilt tilstand, står overfor samme valg. Dersom vi nå antar at  $E(\tau^b) > E(\tau^c)$ , enten ved at de kardinale verdiene i figuren viser en langt større fordel for demokrati som vellykket utfall, eller at forskjellen mellom  $p_e$  og  $p_d$  er liten nok, vil individene faktisk ønske å leve i et demokrati.

Det kan her gjerne forutsettes at Hobbes' resonnement i kapittel 4 er gyldig, og at de samme individene jeg her diskuterer valgte  $\tau^c$  – enevelde – da de befant seg i naturtilstanden og var dominert av et behov for sikkerhet<sup>38</sup>. Situasjonen jeg nå diskuterer inntreffer når disse menneskene har dannet staten og fått tilfredsstilt behovet for sikkerhet; deres ønske om frihet og demokrati vil kunne gjøre seg gjeldende igjen, og de vil ønske å velge på nytt – *de ønsker nå et demokrati*. Machiavelli (2003) ser for seg at man kan få problemer med rastløse og misfornøyde borgere dersom disse fratas verdier de verdsetter og ønsker:

But in republics there is greater vigor, greater hatred, greater desire for revenge, and the memory of earlier freedom cannot and will not let them

<sup>38</sup>Menneskene forutsettes da også å ha levd i et "godt" samfunn med frihet og demokrati som sterke idealer før de havnet i naturtilstanden.

rest. Thus, the surest procedure is either to destroy them or to live in them (Machiavelli 2003:29).

Dersom dette kravet om frihet og selv-styre er svært sterkt, vil Hobbes' suveren ha to valg: a) staten omformes til et demokrati eller b) suverenen må med sin makt sensurere og fortie – eller tilintetgjøre – alle som er med på å fremme disse doktrinene.

Det synes rimelig å anta at alternativ b) for Hobbes *ikke* er akseptabelt *dersom* demokrati kunne tenkes å være stabilt, og faktisk kunne sikret fred for en gitt befolkning<sup>39</sup>! Hobbes fortviler over sine samtidiges kjærlighet til Aristoteles og antikkens demokrati-idealene, og ønsker helt klart å sensurere denne typen doktriner; men, vi kan ikke i dag si oss enige i Hobbes' grunnlag for dette standpunktet: at demokrati er ustabil og uunngåelig vil lede mennesker tilbake til naturtilstanden.

Hoekstra (2007b:215) påpeker også at Hobbes selv innser at mennesker kan mistrives uten muligheten til å være en del av suverenen. Hobbes nevner i *The Elements of Law* dette som en årsak til opprør:

Thirdly, the other sort of discontent which troubleth the mind of them who otherwise live at ease, without fear of want, or danger of violence, ariseth only from a sense of their want of that power, and that honour and testimony thereof, which they think is due unto them (Hobbes 1999:163).

Det er forøvrig verdt å merke seg at Hobbes igjen nærmer seg Maslows tanker, ved at han tillegger denne misnøyen mennesker som "forøvrig lever et bekymringsfritt liv". Hobbes ønsker at staten så langt det er mulig lar mennesker leve som de selv ønsker, så lenge den aldri viker fra sitt formål, som er å opprettholde freden. Dersom man så kan vise, noe jeg mener man kan, at et demokrati faktisk *er* et levedyktig alternativ til enevelde i visse tilfeller, eller at et enevelde med sterkt frihetssøkende og demokratisk orienterte undersåtter ville stå maktesløs mot folkets krav – med mindre folket ble undertrykket i en slik grad at *ingen* ville ønsket en slik stat – synes det åpenbart at et Hobbesiansk resonnement fører til en demokratisk stat. Det synes i alle tilfeller klart at en universell Hobbesiansk teori ikke tar et fast standpunkt til hvilken styreform som foretrekkes.

---

<sup>39</sup>Jeg må her forutsette at en eventuell overgang til demokrati er mulig uten at man havner i naturtilstanden. Hobbes hevder, som vi har sett, at det eksisterende alltid er best, men dette synes mer som et argument ment for å roe undersåtter enn noe som kan anvendes til å si noe om foretrukket styreform.

Jeg ønsker med det å revidere begge premissene som her ble omtalt. Premiss 4 endres til å kreve at den mest stabile styreform *ut fra forholdene* velges. Premiss 5 kvalifiseres noe, ved at menneskers risikoaversjon ikke anses som absolutt, slik den var i kapittel 4. Dette på grunnlag av den mulighet at individer i staten, med et forhold til frihet og demokrati, og uten ekstrem mangel på sikkerhet, *vil* kunne tenkes å ofre *noe* sikkerhet for å få frihet og demokrati.

**Revidert premiss 4** Den mest stabile styreform varierer med individers idealer

**Revidert premiss 5** Mennesker er generelt sett ikke villige til å ofre sikkerhet for andre goder

### Vurdering av premiss 6: Hobbes' "liberalisme"

**Premiss 6** Staten skal ikke frata mennesker mer frihet enn nødvendig

Dette premisset endres ikke, da jeg fremdeles anser den tidligere etableringen av premisset i 4.2.2 som gyldig. Mens det uten tvil er problemer forbundet med Hobbes' manglende forsvar for individuelle rettigheter, synes det klart at elementer av Hobbes' filosofi foregrep mange av liberalismens idealer, som likhet for loven, en vektleggelse av at hver person selv skal kunne velge og forfølge det han selv anser som godt og fokuset på at staten skal gjøre så lite som mulig utover det å sikre fred og folkets sikkerhet.

## 5.4 Konklusjon

I tabell 5.1 har jeg kvalifisert to premisser (4 og 5) og konklusjonen (revidert premiss/konklusjon er markert med \*) fra resonnementet i kapittel 4. Årsaken til disse endringene er et krav om at resonnementet skal være universelt – det skal kunne gjelde alle mennesker i alle kjente tidsperioder. Det synes klart at enevelde ikke kan være et *krav*, da demokrati og frihet for mange mennesker ansees som viktige verdier de faktisk vil være villige til å ofre noe sikkerhet for å oppnå, dersom de ikke er dominert av sikkerhetsbehov. Videre kan det hevdes at nyere historie har vist at demokratier *kan* være stabile – vel så stabile som våre moderne enevelder – og slik sett synes dilemmaet forbundet med å velge demokrati *eller* sikkerhet kunstig.

P1	<i>Fred</i> uten en overodnet samlende makt er ikke mulig
P2	Død – krig, og spesielt borgerkrig er det verste ondet for mennesker
P3	Makt <i>kan</i> ikke deles, kun midlertidig delegeres
P4*	Den mest stabile styreform varierer med individers idealer
P5*	Mennesker er generelt sett ikke villige til å ofre sikkerhet for andre goder
P6	Staten skal ikke frata mennesker mer frihet enn nødvendig
K*	Den absolutte staten – med styreform etter forholdene – er nødvendig og ønsket

Tabell 5.1: Oppdatert versjon av Hobbes' resonnement. Premissene som er endret er markert med \*.

Jeg har i den første delen av kapittelet konkludert med at jeg mener staten, for Hobbes, er nødvendig, og at frivillig samarbeid og organisering i mindre kommunitarianske samfunn vurderes og forkastes av Hobbes. Jeg har likeså hevdet at Hobbes anser makten som udelelig, slik at enhver suveren har all makt.

Alt i alt har jeg tatt et noe kontroversielt grep: jeg hevder at deler av Hobbes' resonnement må ansees som kulturelt og historisk betinget, og at hans filosofi derfor ikke kan hevdes å være universelt gyldig. Dette gjelder spesielt Hobbes' syn på styreformmer.

Jeg har så revidert Hobbes' resonnement der det synes å være dekning for dette. *Folkets sikkerhet* må ansees å ha absolutt prioritet; dersom en viss type individer trenger en annen styreform enn Hobbes' foretrukne enevelde for å være sikre, vil ikke et Hobbesiansk resonnement anbefale enevelde.

At mennesker ikke alltid er like risikoaverse som Hobbes forutsetter, bygger etter min mening på Hobbes' noe primitive forståelse av menneskelig psykologi. Maslows behovshierarki kom først midt i det tyvende århundret, og jeg mener dette perspektivet på menneskers behov er svært nyttig. Det gjør det i alle fall mulig å forklare hvorfor mennesker muligens ikke er like risikoaverse som Hobbes fremstiller dem, og det uten at Hobbes nødvendigvis tok feil: han forklarte *ett aspekt* ved mennesker svært godt, men overså muligens andre.

En svakhet ved Hobbes' resonnement er det absolutistiske elementet. Jeg har gjennom oppgaven skilt mellom Hobbes' *krav* og *anbefalinger*, da jeg mener det er en viktig forskjell på disse. At Hobbes i sin tid *anbefaler* enevelde kan muligens forklares som kulturelt og historisk betinget, men han *krever* en absolutt suveren, på et grunnlag



som ikke synes å være betinget på samme vis, så dette mener jeg man ikke kan fjerne dersom man skal kunne hevde å snakke om Hobbes' filosofi. Jeg har videre hevdet at dette spørsmålet til en viss grad kan reduseres til en semantisk diskusjon, der moderne maktdeling ikke nødvendigvis bryter med Hobbes' absolutisme, dersom det faktisk i siste instans er mulig å spore all makt til ett organ.

Jeg ønsker dog ikke å unnskyldes Hobbes her, da det ikke synes som om Hobbes' *intensjon* var å tillate en såpass sterk faktisk maktdeling som vi har i dag. Hobbes anså suverenens behov for absolutt makt som sentralt, og mens han innså at denne makt-konsentrasjonen kunne være problematisk, mente han alternativet uansett ville være verre.

Den oppdaterte konklusjonen, at den absolutte staten, med styreform etter forholdene, er nødvendig og ønsket, fremstår som en meget "Hobbessiansk" konklusjon. Det er imidlertid *ikke* Hobbes' konklusjon, men den følger av hans resonnement, dersom man utvider (men ikke bryter med) Hobbes' modell av menneskelig psykologi, og tar til etterretning at demokratier har vist seg å være mer stabile enn enevelder i moderne tid. Videre, da Hobbes opplagt ikke hadde kjennskap til vår tids måte å utøve demokrati på, synes det urimelig å kreve at Hobbes forkastelse av *en annen type demokrati* må beholdes.

# Kapittel 6

## Konklusjon

Jeg har i denne oppgaven gått bredt ut i min analyse av Hobbes' syn på mennesket og staten. Først etablerte jeg at Hobbes har et langt mer nyansert, og mindre kynisk, syn på mennesket enn mange ser ut til å anta. Jeg vurderte så Maslow, opprinnelig i et forsøk på å se hvordan Hobbes' filosofi kunne forståes i moderne terminologi, men fant samtidig viktige likheter mellom disse tenkerne. Denne likheten er interessant, da Hobbes' og Maslows rykte vanskelig kunne vært mer forskjellig i forhold til deres syn på mennesket.

Videre så jeg på aktuelle perspektiver på mennesket fra eksperimentell økonomi og nevroøkonomi. I disse fagfeltene presenteres mennesket noe annerledes enn hva som er tilfellet hos Hobbes, men jeg har pekt på at det er grunn til å vurdere hvorvidt forklaringene som brukes av de fleste forskerne i fagfeltene muligens ikke er de eneste forklaringene som er tilgjengelige. De empiriske resultatene, som fremmes som kontroversielle i forhold til konvensjonell neoklassisk økonomi, er *ikke* nødvendigvis uforenlige med Hobbes' syn på mennesket, og jeg har kort presentert noen alternative forklaringer som kan underbygge et slikt standpunkt.

I det fjerde kapittelet har jeg så sett på hva Hobbes sier om statens form og funksjon. Jeg har der presentert et logisk resonnement i tråd med Hobbes' filosofi, og dette resonnementet, med den konklusjon at Hobbes' absolutistiske stat er nødvendig og ønskelig, fremstår for meg som gyldig, gitt premissene. Det er imidlertid først i kapittel 5 jeg vurderer hvorvidt selve premissene Hobbes ser ut til å forutsette, er sanne.

Jeg har så sett på argumentene for anarki (eller "frivillig samarbeid") som et reelt al-

ternativ til staten; at slikt samarbeid er mulig slår meg som overveiende sannsynlig. Mitt formål var å vurdere hvorvidt muligheten for frivillig samarbeid eller kommunitarianske samfunn påvirker Hobbes' resonnement, og jeg hevdet at slike argumenter *ikke* er egnet til å rokke ved Hobbes' resonnement. Dette fordi Hobbes klart utelukker denne typen samarbeid som en "løsning" på naturtilstanden; han anerkjenner muligheten for samarbeid i mindre grupper, men anser det som utilstrekkelig. Videre er det også i forhold til krav om likhet, og maksimal grad av frihet som er forenlig med *fred*, aspekter ved kommunitarianske samfunn som gjør det naturlig å anta at Hobbes ville mislike disse.

Min konklusjon i denne oppgaven er for det første at Hobbes ser ut til å ha forutsatt en for begrenset modell av mennesket til at hans filosofi kan anees som universell. Nærmere bestemt virker det som om Hobbes kan ha fokusert uforholdsmessig mye på mennesker dominert av sikkerhetsbehov. Jeg har hevdet at dette er uten store konsekvenser (og faktisk helt korrekt) i, og like etter, naturtilstanden, men at det etter statsdannelsen kan medføre problemer; mennesker som er mette og føler seg trygge, og som tilfredsstillende visse forutsetninger om tidligere erfaringer, kan oppleve å ønske demokrati og frihet.

Dette leder til det andre punktet i konklusjonen, nemlig at Hobbes' enevelde *ikke* kan anees som det best egnede for å opprettholde fred for enhver populasjon. I avslutningen av oppgaven har jeg hevdet at et Hobbesiansk resonnement derfor ikke krever enevelde. Dette fører til at jeg fjerner meg fra Hobbes' egen tekst, ved at jeg anser hans forsvar for eneveldet som lite holdbart. Det er imidlertid ikke med dette sagt at jeg fjerner meg fra Hobbes' *intensjoner* med sin filosofi, da jeg sterkt vektlegger hans kriterium for valg av styreform, som er evnen til å bevare fred og folkets sikkerhet.

At Hobbes' syn på styreform må revideres har blitt spesielt tydelig etterhvert som moderne historie har vist at konstitusjonelle demokratier fremstår som vel så stabile som enevelder. For at Hobbes' filosofi skal fremstå som universell, må Hobbes' forsvar for eneveldet forkastes som kulturelt og historisk betinget.

"Hovedproblemet" man står igjen med i en Hobbesiansk filosofi som her fremsatt er dermed Hobbes' absolutisme. Det er mulig å fjerne Hobbes' krav om dette i forsøket på å lage en moderne Hobbesiansk teori, noe som har blitt gjort av flere, men det synes også nyttig å bevare Hobbes' sentrale karakter. Hobbes er for mange selve ansiktet til statsabsolutismen, og dersom også dette elementet fjernes, kunne man like gjerne

utformet en ny teori uten referanser til Hobbes. Jeg har forsøkt å tolke, og oppdatere, Hobbes' filosofi på beste måte, *uten* å forsøke å skape et glansbilde der alle kontroversielle elementer er fjernet. Det er nyttig å vurdere en filosofi som Hobbes' kritisk, men det er etter min mening like nyttig å forsøke å se hvilke innsikter Hobbes faktisk har å tilby den moderne leser. Jeg støtter meg med dette til Rawls' oppfordring om hvordan Hobbes bør leses:

Finally, as an aside, let me say that in looking at a text of this sort [*Leviathan*], which is so large, and with so many elements in it, if you are to get as much out of it as you can, you must try to interpret it in the best and most interesting way. There is no point in trying to defeat it, or to show the author was wrong in some way. The thing is to make as much out of it as you can and to try to get a sense of how the overall view might go, if you put it in the best way. Otherwise, I think it is a waste of time to read it, or to read any of the important philosophers (Rawls 2007:52).

# Bibliografi

- Aubrey, J. (1999). *Aubrey's Brief Lives*. Jaffrey, New Hampshire: Godine.
- Aurelius, M. (1964). *Meditations*. London: Penguin Books.
- Axelrod, R. (2006). *The Evolution of Cooperation*. New York: Basic Books.
- Berlin, I. (2002a). Five essays on liberty: Introduction. I H. Hardy (Red.), *Liberty*. Oxford: Oxford University Press.
- Berlin, I. (2002b). John Stuart Mill and the ends of life. I H. Hardy (Red.), *Liberty*. Oxford: Oxford University Press.
- Berlin, I. (2002c). *Liberty*. Oxford: Oxford University Press.
- Berlin, I. (2002d). Political ideas in the 20th century. I H. Hardy (Red.), *Liberty*. Oxford: Oxford University Press.
- Berlin, I. (2002e). Two concepts of liberty. I H. Hardy (Red.), *Liberty*. Oxford: Oxford University Press.
- Bernheim, B. D. (2009). The psychology and neurobiology of judgement and decision making: What's in it for economists? Se Glimcher et al. (2009b), Kapittel 9.
- Bowles, S. og H. Gintis (2005). Social capital, moral sentiments, and community governance. Se Gintis et al. (2005a), Kapittel 13.
- Boyd, R., H. Gintis, S. Bowles, og P. J. Richerson (2005). The evolution of altruistic punishment. Se Gintis et al. (2005a), Kapittel 7.
- Brosnan, S. F. (2009). Responses to inequity in non-human primates. Se Glimcher et al. (2009b), Kapittel 19.
- Brown, S. M. J. (1959). Hobbes: The Taylor thesis. *The Philosophical Review* 68(3), 303–323.
- Camerer, C. F. (2009). Behavioral game theory and the neural basis of strategic choice. Se Glimcher et al. (2009b), Kapittel 13.
- Charness, G. og M. Dufwenberg (2006). Promises and partnership. *Econometrica* 74(6), 1579–1601.

- Damasio, A. (2009). Neuroscience and the emergence of neuroeconomics. Se Glimcher et al. (2009b), Kapittel 14.
- Dayan, P. og B. Seymour (2009). Values and actions in aversion. Se Glimcher et al. (2009b), Kapittel 12.
- Falk, A. og U. Fischbacher (2005). Modeling strong reciprocity. Se Gintis et al. (2005a), Kapittel 6.
- Fehr, E. (2009). Social preferences and the brain. Se Glimcher et al. (2009b), Kapittel 15.
- Fehr, E. og U. Fischbacher (2005). The economics of strong reciprocity. Se Gintis et al. (2005a), Kapittel 5.
- Fragar, R. (1987). *Motivation and Personality*, Kapittel Foreword: The Influence of Abraham Maslow. I Maslow (1987).
- Fromm, E. (1994). *Escape from Freedom*. New York: Henry Holt and Company.
- Fukuyama, F. (2006). *The End of History and the Last Man*. New York: Free Press.
- Føllesdal, D., L. Walløe, og J. Elster (1996). *Argumentasjonsteori, språk og vitenskapsfilosofi*. Oslo: Universitetsforlaget.
- Gauthier, D. P. (1969). *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford: Oxford University Press.
- Gert, B. (1965). Hobbes, mechanism, and egoism. *The Philosophical Quarterly* 15(61), 341–349.
- Gert, B. (1967). Hobbes and psychological egoism. *Journal of the History of Ideas* 28(4), 503–520.
- Gintis, H., S. Bowles, R. Boyd, og E. Fehr (Red.) (2005a). *Moral Sentiments and Material Interests: The Foundation of Cooperation in Economic Life*. Cambridge: MIT Press.
- Gintis, H., S. Bowles, R. Boyd, og E. Fehr (2005b). Origins, evidence and consequences. Se Gintis et al. (2005a), Kapittel 1.
- Glimcher, P. W., C. F. Camerer, E. Fehr, og R. A. Poldrack (2009a). A brief history of neuroeconomics. Se Glimcher et al. (2009b), Kapittel 1.
- Glimcher, P. W., C. F. Camerer, E. Fehr, og R. A. Poldrack (Red.) (2009b). *Neuroeconomics: Decision making and the brain*. London: Academic Press.
- Goldsmith, M. M. (1996). Hobbes on law. I T. Sorell (Red.), *The Cambridge Companion to Hobbes*, Kapittel 11. Cambridge: Cambridge University Press.
- Green, A. W. (1993). *Hobbes and Human Nature*. New Jersey: Transaction Publishers.

- Grossman, L. C. D. (2008). *On Combat: The Psychology and Physiology of Deadly Conflict in War and in Peace*. Milstadt, Illinois: Warrior Science Publications.
- Hardin, R. (1991). Hobbesian political order. *Political Theory* 19(2), 156–180.
- Harsanyi, J. C. (1977). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge, Massachusetts: Cambridge University Press.
- Hayek, F. A. (1988). *The Fatal Conceit: The Errors of Socialism*. Chicago: The University of Chicago Press.
- Hellevik, O. (2002). *Forskningsmetode i sosiologi og statsvitenskap*. Oslo: Universitetsforlaget.
- Herrmann, B., C. Thöni, og S. Gächter (2008). Antisocial punishment across societies. *Science* 319, 1362–1367.
- Hobbes, T. (1946). *Leviathan*. London: Basil Blackwell.
- Hobbes, T. (1966). *The English Works of Thomas Hobbes of Malmesbury - Volume VI*. Aalen: Scientia Verlag.
- Hobbes, T. (1973). *Leviathan*. London: Everyman's Library.
- Hobbes, T. (1991a). De homine. I B. Gert (Red.), *Man and Citizen (De Homine and De Cive)*. Indianapolis, Indiana: Hackett Publishing Company.
- Hobbes, T. (1991b). *Man and Citizen (De Homine and De Cive)*. Indianapolis, Indiana: Hackett Publishing Company.
- Hobbes, T. (1997). *A Dialogue Between a Philosopher and a Student of the Common Laws of England*. Chicago: The University of Chicago Press.
- Hobbes, T. (1999). *Human Nature and De Corpore Politico*. Oxford, New York: Oxford University Press.
- Hobbes, T. (2005a). *The English Works of Thomas Hobbes of Malmesbury - Volume II*. New York: Elibron.
- Hobbes, T. (2005b). *The English Works of Thomas Hobbes of Malmesbury - Volume IV*. New York: Elibron.
- Hoekstra, K. (2007a). Hobbes on the natural condition of mankind. I P. Sringborg (Red.), *The Cambridge Companion to Hobbes's Leviathan*, Kapittel 4. Cambridge: Cambridge University Press.
- Hoekstra, K. (2007b). A lion in the house: Hobbes and democracy. I A. Brett og J. Tully (Red.), *Rethinking the Foundation of Modern Political Thought*, Kapittel 11. Cambridge: Cambridge University Press.

- Houser, D. og K. McCabe (2009). Experimental neuroeconomics and non-cooperative games. Se Glimcher et al. (2009b), Kapittel 5.
- Hovi, J. (2008). *Spillteori: En innføring*. Oslo: Universitetsforlaget.
- Jaume, L. (2007). Hobbes and the philosophical sources of liberalism. I P. Sringborg (Red.), *The Cambridge Companion to Hobbes's Leviathan*, Kapittel 8. Cambridge: Cambridge University Press.
- Kahan, D. M. (2005). The logic of reciprocity: Trust, collective action, and law. Se Gintis et al. (2005a), Kapittel 12.
- Kahneman, D. (2009). Remarks on neuroeconomics. Se Glimcher et al. (2009b), Kapittel 33.
- Kahneman, D. og A. Tversky (1990). Prospect theory: an analysis of decision under risk. Se Moser (1990), Kapittel 6.
- Kaplan, H. og M. Gurven (2005). The natural history of human food sharing and cooperation: A review and a new multi-individual approach to the negotiation of norms. Se Gintis et al. (2005a), Kapittel 3.
- Kavka, G. S. (1986). *Hobbesian Moral and Political Theory*. Princeton, New Jersey: Princeton University Press.
- Kuhn, H. W. (Red.) (1997). *Classics in Game Theory*. Princeton, New Jersey: Princeton University Press.
- Kuper, A. og J. Kuper (Red.) (1996). *The Social Science Encyclopedia*. Abingdon, Oxon: Routledge.
- Lively, J. og A. Rose (Red.) (1989). *Modern Political Theory*. London: Routledge.
- Locke, J. (2005). *Two Treatises of Government and A Letter Concerning Toleration*. Lawrence, Kansas: Digireads.com Publishing.
- Luce, R. D. og H. Raiffa (1990). Utility theory. Se Moser (1990), Kapittel 1.
- Machiavelli, N. (2003). *The Prince*. New York: Bantam Books.
- Malnes, R. (1993). *The Hobbesian Theory of International Conflict*. Oslo: Scandinavian University Press.
- Martel, J. R. (2007). *Subverting the Leviathan: Reading Thomas Hobbes as a Radical Democrat*. New York: Columbia University Press.
- Maslow, A. H. (1987). *Motivation and Personality*. Delhi: Pearson Education.
- Mayr, U., W. T. Harbaugh, og D. Tankersley (2009). Neuroeconomics of charitable giving and philanthropy. Se Glimcher et al. (2009b), Kapittel 20.
- McNeilly, F. S. (1966). Egoism in Hobbes. *History of Philosophy* 16(64), 193–206.



- Midgaard, K. (2004). Gode styreformer og demokrati. I K. Midgaard og B. E. Rasch (Red.), *Demokrati – vilkår og virkninger*, Kapittel 2. Bergen: Fagbokforlaget.
- Mill, J. S. (1985). *On Liberty*. London: Penguin Books.
- Mill, J. S. (1987a). Coleridge. I A. Ryan (Red.), *Utilitarianism and Other Essays*. London: Penguin Books.
- Mill, J. S. (1987b). Utilitarianism. I A. Ryan (Red.), *Utilitarianism and Other Essays*. London: Penguin Books.
- Mill, J. S. (2006). *The collected works of John Stuart Mill: A System of Logic, Ratiocinative and Inductive*. Indianapolis: Liberty Fund.
- Minogue, K. (1973). *Leviathan*, Kapittel Introduction. I Hobbes (1973).
- Morgenthau, H. J. (1985). *Politics Among Nations: The Struggle for Power and Peace*. New York: McGraw-Hill.
- Morrow, J. D. (1994). *Game Theory for Political Scientists*. Princeton: Princeton University Press.
- Moser, P. K. (Red.) (1990). *Rationality in Action: Contemporary Approaches*. Cambridge: Cambridge University Press.
- Myers, D. G. (2003). *Psychology*. New York: Worth Publishers.
- Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Cambridge: Harvard University Press.
- Nash, J. (1997). Non-cooperative games. Se Kuhn (1997), Kapittel 3.
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Oakeshott, M. (1946). *Leviathan*, Kapittel Introduction. I Hobbes (1946).
- Ostrom, E. (2005). Policies that crowd out reciprocity and collective action. Se Gintis et al. (2005a), Kapittel 9.
- Phelps, E. A. (2009). The study of emotion in neuroeconomics. Se Glimcher et al. (2009b), Kapittel 16.
- Pindyck, R. S. og D. L. Rubinfeld (2005). *Microeconomics*. New Delhi: Prentice-Hall of India.
- Raiffa, H. (2002). *Negotiation Analysis: The Science and Art of Collaborative Decision Making*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Rawls, J. (2007). *Lectures on the History of Political Philosophy*. Cambridge: Harvard University Press.

- Render, B., R. M. Stair, og M. E. Hanna (2006). *Quantitative Analysis for Management*. New Jersey: Pearson Prentice Hall.
- Ryan, A. (1996). Hobbes's political philosophy. I T. Sorell (Red.), *The Cambridge Companion to Hobbes*, Kapittel 9. Cambridge: Cambridge University Press.
- Sanders, D. (2002). Behaviouralism. I D. Marsh og G. Stoker (Red.), *Theory and Methods in Political Science*, Kapittel 2. Hampshire: Palgrave Macmillan.
- Sanfey, A. og M. Dorris (2009). Games in humans and non-human primates: Scanners to single units. Se Glimcher et al. (2009b), Kapittel 6.
- Schultz, W. (2009). Midbrain dopamine neurons: A retina of the reward system? Se Glimcher et al. (2009b), Kapittel 21.
- Sethi, R. og E. Somanathan (2005). Norm compliance and strong reciprocity. Se Gintis et al. (2005a), Kapittel 8.
- Silk, J. B. (2009). Social preferences in primates. Se Glimcher et al. (2009b), Kapittel 18.
- Skinner, Q. (1996). *Reason and Rhetoric in the Philosophy of Hobbes*. Cambridge: Cambridge University Press.
- Skog, O.-J. (2004). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming*. Oslo: Gyldendal Akademisk.
- Slomp, G. og M. M. A. LaManna (1996). Hobbes, Harsanyi and the edge of the abyss. *Canadian Journal of Political Science* 29(1), 47–70.
- Smith, E. A. og R. B. Bird (2005). Costly signaling and cooperative behavior. Se Gintis et al. (2005a), Kapittel 4.
- Smith, V. L. (2009). Experimental economics and neuroeconomics. Se Glimcher et al. (2009b), Kapittel 2.
- Strauss, L. (1952). *The Political Philosophy of Hobbes*. Chicago: The University of Chicago Press.
- Taylor, A. E. (1989). The ethical doctrine of Hobbes. Se Lively og Rose (1989).
- Taylor, M. (1982). *Community, Anarchy and Liberty*. Cambridge: Cambridge University Press.
- Taylor, M. (1987). *The Possibility of Cooperation*. Cambridge: Cambridge University Press.
- Thukydid (1972). *History of the Peloponnesian Wars*. New York: Penguin Books.
- Tralau, J. (2007). Leviathan, the beast of myth: Medusa, Dionysos, and the riddle of

- Hobbes's sovereign monster. I P. Sringborg (Red.), *The Cambridge Companion to Hobbes's Leviathan*, Kapittel 2. Cambridge: Cambridge University Press.
- Tuck, R. (2007). Hobbes and democracy. I A. Brett og J. Tully (Red.), *Rethinking the Foundation of Modern Political Thought*, Kapittel 10. Cambridge: Cambridge University Press.
- von Mises, L. (2007). *Human Action: A Treatise on Economics*. Indianapolis: Liberty Fund.
- von Neumann, J. og O. Morgenstern (2004). *Theory of Games and Economic Behavior*. Princeton, New Jersey: Princeton University Press.
- Weber, E. U. og E. J. Johnson (2009). Decisions under uncertainty: Psychological, economic, and neuroeconomic explanation of risk preference. Se Glimcher et al. (2009b), Kapittel 10.
- Østerud, Ø. (2002). *Statsvitenskap: Innføring i politisk analyse*. Oslo: Universitetsforlaget.

# Register

- absolutisme, 111–112, 185–189, 201
  - kontroversielt, 188
- adferd, 25–26
  - og hjerneskode, 72
  - prediksjon av, 12, 73
- adferdspsykologi, 60–61
- affeksjon, *Se* kjærlighet
- allianser, *Se* koalisjoner
- altruisme, 22, 24, 36, 69, 129
  - definisjon av, 80
  - gjengjeldende, 80
  - godfølelse-, 81
  - kortsiktig, 80
  - og egeninteresse, 80
  - ren, 81
  - sofistikert, 38–39, 129
- altruistisk straff, 68–71, 142
- anerkjennelse, *Se* behov, ære
- anonymitet, 75–76
- asosial straff, 143–144
- autoritarisme, 90–98, 125
- autoriteter, 121, 165–166
- avtaler, *Se* kontrakter
- behov, 45–52, 169–176
  - anerkjennelse/respekt, 50–51, 104, 170–171, 175, 180
  - fysiologiske, 45–47
  - og dominans, 45–46, 48, 58, 59, 193–194, 197
  - selvrealisering, 51–52
  - sikkerhets-, 47–49, 118
  - tilfredsstillelse av, 11, 46–47
    - forutsetninger for, 57, 171
  - tilhørighet/kjærlighet, 49–50
  - undertrykkelse av, 51, 196
- behovshierarki, 45–52, 174
- belønning
  - og egeninteresse, 83
  - og følelser, 72
  - teorien om, 82
- borgerkrig, *Se* krig
- demokrati, *Se* styreformer
- det gode, *Se* goder
- dyder, 25
- dyreforskning, 46
- døden, 17–18, 29, 48, 118, 150, 182–185
  - frykten for, 29, 34, 152, 182–184
  - og forskningsetikk, 74, 79
  - og naturtilstanden, 106
  - og religion, 95
  - og spillteori, 119
- egeninteresse, 15, 18, 21–24, 28, 36, 38
  - begrensning av, 15, 25
  - isolert, 61
  - langsiktig, 80
  - og fornuft, 26
- egoisme, 16, 18, 36, 38
  - dominerende, 42
  - og altruisme, 52–53
  - psykologisk, 41–42
  - snever, 15
  - tautologisk, 42–43
- eksperimentell økonomi, 60–87, 140
  - metode, 62
    - etiske begrensninger, 79
    - innvendinger, 74–79
- eminens, *Se* relativ gevinst
- enevelde, *Se* styreformer
- ensretting, 91–92, 97
- evolusjon, *Se* instinkter
- familier, *Se* naturtilstanden
- fellesgoder, 84–85, 104–106, 113, 129–136
  - gratispassasjer, 105
- folketeoremet, 135–136
- fornuft, 12, 15, 26, 35, 55–56, 104
  - og drifter, 33
- forsamlingsstyre, *Se* styreform, demokrati
- fred, 100–101, 147
- fri vilje, 121–122
- frihet, 57, 58, 121, 178–181
  - egenverdi, 172–174
  - erfaring med, 173
  - fordeler ved, 176–177

- minumumskrav, 174
- negativ, 121–122, 169–170, 178
- og demokrati, 177–178
- og lov, 122
- og nødvendighet, 121–122, 169–170
- og sikkerhet, 58–59, 119, 174–175, 184, 197
- organisasjons-, 93
- positiv, 121–122, 169–170, 178
- som behov, 170–176, 181
- som problem, 175
- ytrings-, 93
- frivillig samarbeid, 103, 104, 135, 137, 161–166
- følelser, 55, 72–73
  
- goder, 10–12, 14, 17
  - enighet om, 21
  - felles, *Se* fellesgoder
  - knappe, *Se* knappe goder
  - kollektive/individuelle, 104, 113–114, 129, 133, 161
  - tilsynelatende, 11, 34
  - vurdering av, 11, 21, 57, 180
- godhet, 22–25
  
- herredømme, 29, 36, 107, 108
- Hobbes' stat, 90–98, 135, 168
- homeostasis*, 47
- Homo economicus*, 9, 21, 61, 62, 64, 65, 140
  
- implisitte forutsetninger, 98–99, 119
- indirekte gjensidighet, 82
- instinkt, 46
- instinkter, 10–11, 55–57, 73–74
  - og dyr, 56
  - og evolusjon, 73
  - som positivt, 55
- intensjoner, 19, 29, 66–67, 104
  
- kamp, 110, 151
  - usikkert utfall, 30, 108
- kjærlighet, 22–24, 49
  - som behov, *Se* behov
- knappe goder, 27–28, 100, 144
- koalisjoner, 27, 134–135, 153, 160–161
  - familier, 157
  - i naturtilstanden, 154–159
- kommunitarianske samfunn, 136–146, 161–166
  - negative sider ved, 141, 164–166
  - størrelse, 138–139, 144, 163
- konflikt, *Se* kamp, 54, 103, 107
  - årsaker til, 30, 31, 100
- konkurransen, 32, 100
  - politisk, 92–94
  
- konstitusjoner, 187
- kontrakter, 12, 15, 31
  - håndheving av, 161
  - i naturtilstanden, 159–160
  - selvhåndhevende, 102, 132
  - sosialkontrakten, 153
- kontrolleffekter, 78
- koordinasjonsproblemer, 134–135
- kortsiktighet, *Se* langsiktighet
- kostbar signalisering, 84
  - og fellesgoder, 84
- krig, 18, 103, 106–111, 154, 157, 182–185
  - og styreform, 116
  
- langsiktighet, 11, 32–38
  - diskontering, 32
  - forkjøp, 36, 54, 100
  - kompetanse/innsats, 34
  - og kortsiktighet, 33
  - som nødvendig, 36
  - ulemper ved, 36
- liberalisme, 121, 122, 197
  - Hobbes og, 96
  - og kommunitarianske samfunn, 164–166
- lover, *Se* straff, 57, 70, 71, 94, 122, 168
  - og suverenitet, 186, 187
- lykke, 11, 13, 14, 17
  - som behovstilfredsstillelse, 46–47
  
- majoritetstyranni, 142, 165
- makt, 13–15, 23–24, 27, 31, 185–189
  - delegasjon av, 94, 112, 186
  - instrumentell, 23–24, 30
  - korrupperende effekt av, 188–189
  - naturlig, 27
  - og vilkårlighet, 112, 114–115, 168
  - oppfattet, 23–24, 30
  - som logisk udelelig, 185
  - vurdering av, 108
- maktdeling, 111–112, 185–189
  - tilsynelatende, 112
- markedssvikt, 105
- mennesket, 174
  - ambisjoner, 92, 104
  - feilbarlighet, 11–12
  - forfatning, 21
  - fornuft og instinkt, 55–57
  - innstilling til andre, 28–32
  - likhet, 26, 28, 30, 100
  - modell av, 167, 201
  - og samarbeid, 86
  - og sammenligning, 20, 104

- og språk, 104
- som dyr, 53
- svakheter ved, 33–34
- verdi, 30, 31
- menneskets natur, 9–39, 49, 52–55
  - konsekvenser av, 32, 36–37, 54–55, 100
  - som god, 16, 54, 55
  - som undervurdert, 53
- midler, *Se* makt, 13
- modeller
  - nytte, 64
  - realisme, 61, 64
- moderasjon, 14–15, 26, 36–37, 107, 109
- Modus vivendi*, 109
- moral, *Se* naturlover, 25–27, 95
- motivasjon, 15–16, 19, 42
  - og sult, 45–46
  - ubevisst, 35, 42, 72–74, 82–83, 184
- naturlover, 25–27, 150–151, 168
  - og staten, 95–96
  - og straff, 69
  - som fornuftslover, 56, 105
  - som moral, 25
- naturrettigheter, 137, 147
- naturtilstanden, 147, 181
  - familier, 157, 159
  - formalisering av, 103, 106–111, 136–138, 146–161
  - koalisjoner, *Se* koalisjoner
  - konflikt i, 99–101, 107–108, 151
  - kontrakter i, 159–160
  - mangler ved, 106, 155–159
  - samarbeid i, 154–161, 164
  - sikkerhetsbehov og, 48
  - som Chicken-spill, 107–111, 146, 149–152
  - som Fangenes dilemma-spill, 103, 106, 110, 146
  - som logisk abstraksjon, 159
  - verste ondet, *Se* onder
- nevroøkonomi, 60–87
  - metode, 63
  - etiske begrensninger, 79
- normer, 142
  - og straff, 71
- nytte, 10, 17–22
  - funksjon, 9, 18, 21, 65
    - Hobbess, 38–39
  - maksimering, 10, 16, 110–111, 120
  - forventet, 11, 17, 118–119
  - negativ, 17
  - og prosess, 66
  - og utfall, 66
  - sammenligning av, 20
  - vurdering av andres, 19–22, 27–32, 65
- omdømme, *Se* rykte
- onder, 10
  - verste, 17–18, 106–111, 182–185
- ondskap, 29, 32, 54
- other-regarding, *Se* preferanser, sosialt anskuede
- pluralisme, 142
- preferanser, 19
  - avdekking av, 19, 32
  - egen-anskuende, 61, 66, 129
  - og heterogenitet, 20, 65
  - ordinale, 101
  - sosiale, 21, 38–39, 64–68
    - nyttefunksjon, 65
  - sosialt anskuede, 65, 129
- prospektteori, 17, 118–119
- psykopater, 53
- rasjonalitet
  - individuell, 105, 133
- relativ gevinst, 27–28, 32
- relativ makt, *Se* relativ gevinst
- religion, 26, 94–95, 185
- retorikk, 156–159
- rettferdighet, 12, 15, 25, 67
- rettigheter, 31, 169
- rettsstat, 143–144, 187
- risiko, 117–120
  - aversjon, 116, 194
  - definisjon av, 118
  - mulighetsorientering, 118, 119
- rykte, 30, 50, 83–85, 108
  - grupper og, 83
  - og anonymitet, 83
  - som makt, 30–31, 81, 83, 108
- rådgivning, 114
- samarbeid, 69, 129–146
  - betinget, 130
  - frivillig, *Se* frivillig samarbeid
  - i Chicken-spill, 150
  - i naturtilstanden, *Se* naturtilstanden
  - markedsløsning, 138–139
  - mulighetene for, 135–137
  - selvhåndhevende, 102, 132
  - ved kontrakt, 104
- samhandlingsproblemer, 129–130
- sanksjoner, *Se* straff, altruistisk straff

self-regarding, *Se* pref., egen-anskuende  
 selvmord, 183  
 sensur, 92–93  
     og frihet, 58  
 sikkerhet, *Se* behov, 14  
     og livskvalitet, 183  
 sosiale dyr, 103–105  
 sosiale preferanser, *Se* preferanser  
 spillteori, 16  
     Chicken-spill, 107–111, 149  
     delspill, 131  
     diskontering, 32, 101–102, 132–134  
     dominans, 132, 149  
     eksperimentell metode, 63  
     engangsspill, 132–133  
         og langsiktig egeninteresse, 78  
         og oppdemming, 76–78  
     Fangenes dilemma, 101–102, 107, 131–135, 140  
     fullstendig informasjon, 20  
     gjentatt spill, 101–102, 130–135  
     Hobbes' spill, 109  
     ikke-kooperative spill, 161  
     interaksjonsform, 130–131, 154  
     koalisjoner, *Se* koalisjoner  
     kommunikasjon, 134–135, 160–161  
     kooperative spill, 161  
     likevekt, 110  
     Moderat selvhvedelse, 107, 140  
     naturtilstanden, 106–111  
     omforming av spill, 140  
     rykte, 83  
     tidshorisont, 38, 130, 136  
 spontan orden, *Se* frivillig samarbeid  
 stater  
     behovet for, 101–106, 128–166  
     etableringen av, 152–154  
     fordeler ved, 51, 106  
     og bindinger, 94–96, 168, 187  
     og formål, 91–93, 96–97, 117, 121, 122, 142  
     og ideologi, 91  
     som ønskelige, 14, 106–111  
 sterk gjensidighet, 66  
 straff, 69, 142–144, 189  
     altruistisk, *Se* altruistisk straff  
     hevn, 26  
     i naturtilstanden, 69  
     selektiv, 142–143  
 strategier  
     Disaster-avoidance, 110–111  
     Grim trigger, 101–102  
     maximin, 110–111, 120  
     Penance, 142–143  
     tit-for-tat, 132, 142  
 styreformer, 113–117, 189–197  
     blandede, 192  
     demokrati, 178–180, 190–197  
     enevelde, 179  
         Hobbes' forsvar for, 113–116, 189–192  
     kriterium for valg av, 117, 193  
     og frihet, 177–178, 193–194, 196  
     og stabilitet, 106, 192, 197  
     valg av, 116, 195–197, 201  
*summum malum*, *Se* døden  
 suverenen, 92  
     og makt, 94–96  
 talefrihet, *Se* frihet  
 tilfredshet, 14  
 totalitarisme, 91, 121  
 ulikhet, 21–22  
 ulikhetsaversjon, 21–22, 67–68  
     nyttefunksjon, 67  
 underbevissthet, *Se* motivasjon  
 underkastelse, 106, 108, 150–152  
     som behov, 175, 180  
 urettferdighet, 67  
 urettferdighetsaversjon, *Se* ulikhetsaversjon  
 usikkerhet, *Se* behov, sikkerhet, 12, 32, 54, 100,  
     150, 160  
     definisjon av, 118  
     i naturtilstanden, 58  
 veldedighet, *Se* godhet  
 verdi, 30, 31  
     menneskets, *Se* mennesket  
     vurdering av, 31, 108  
 vilkårlighet, *Se* makt  
 vitenskapsteori, 11  
     mekanisme, 121–122  
 ære, *Se* rykte, 22–25, 30–32, 50–51, 100, 103  
     knappt gode, 28  
 økonomi, 105  
     eksperimentell, *Se* eksperimentell økonomi  
     nevro-, *Se* nevroøkonomi  
     tradisjonell, 61, 62, 64