

Ansvar og kontroll

En vurdering av John Martin Fischer og Mark Ravizzas
teori om moralsk ansvar

Even Flaatten



Mastergradsoppgave i filosofi ved Institutt for filosofi,
ide- og kunsthistorie og klassiske språk

UNIVERSITETET I OSLO

15.05.2012

År: 2012

Tittel: Ansvar og kontroll

Forfatter: Even Flaatten

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Avhandlingens problemstillinger er: Kan vi ha ansvar for våre handlinger dersom handlingene er kausalt determinert? Og hvordan skal i så fall dette ansvaret beskrives?

John Martin Fischer og Mark Ravizza behandler disse problemstillingene i sin bok "Responsibility and Control" (1998). De hevder at de elementene som betinger moralsk ansvar ikke er avhengig av stillingtaken til kausal determinisme, alternative muligheter eller fri vilje. Denne filosofiske posisjonen blir kalt *semikompatibilisme*. Avhandlingens oppgave består i å gi en vurdering av Fischer og Ravizzas synspunkter, samt å skissere en alternativ tilnærming.

Innledningsvis viser Fischer og Ravizza til at moralsk ansvar er personers særegne form for ansvar, og at det finnes andre ansvarsformer enn moralsk ansvar. Med henvisning bl.a. til Aristoteles hevder de at det finnes to betingelser for moralsk ansvar. De to betingelsene blir kalt for kunnskapsbetingelsen og den frihets- eller kontrollrelevante betingelsen. Fischer og Ravizza fokuserer i sitt arbeid på den sistnevnte betingelsen. De viser også til at moralsk ansvar er nært knyttet til det Peter Strawson kaller reaktive holdninger – som krenkelse, takknemlighet osv. – og til ros og klander. Det er imidlertid ikke slik at en person nødvendigvis har moralsk ansvar, dersom han blir møtt med reaktive holdninger eller med ros eller klander.

Den sentrale tesen til Fischer og Ravizza er at moralsk ansvar er betinget av veiledningskontroll (*guidance control*). En stor del av deres bok går derfor ut på å gjøre rede for hva veiledningskontroll er. Veiledningskontroll er blant annet betinget av moderat respons på grunner (*moderate reasons-responsiveness*). Forfatterne gjør også oppmerksom på at veiledningskontroll og respons på grunner ikke primært er egenskaper hos den enkelte aktør, men egenskaper ved den mekanismen som styrer aktørens atferd. En mekanisme har moderat respons på grunner hvis:

- den som regel er mottakelig for grunner,
- mottakeligheten er realitetsforankret, og
- den reagerer rasjonelt i minst ett av alle mulige, tenkelige scenarier

Fischer og Ravizza skiller mellom veiledningskontroll og regulativ kontroll. Regulativ kontroll innebærer tilgang til alternative muligheter. Med en deterministisk forutsetning har ikke mennesker regulativ kontroll. Det kreves imidlertid ikke tilgang til alternative muligheter for veiledningskontroll. En person kan respondere på grunner og ha veiledningskontroll, selv om han ikke har tilgang til alternative muligheter.

Veiledningskontroll er en relativt stabil egenskap som kan opprettholdes, selv om aktøren reagerer irrasjonelt i enkeltsituasjoner. En person kan også ha veiledningskontroll, selv om han i enkelttilfeller mangler en situasjonsbetinget form for kontroll (f.eks. mister balansen og faller).

Veiledningskontroll er ikke bare betinget av moderat respons på grunner, men også av eierskap til den operative mekanismen - dvs. at aktøren tar ansvar for sin atferd. I avhandlingen blir det stilt spørsmål til denne betingelsen. Det blir uttrykt tvil om å ta ansvar er en nødvendig betingelse for moralsk ansvar, og det blir spurt om en person kan *ha* ansvar selv om han ikke *tar* ansvar.

I forbindelse med innføringen av å ta ansvar som betingelse for veiledningskontroll, knytter Fischer og Ravizza sin teori til Galen Strawsons subjektivistiske tilnærming til moralsk ansvar. De åpner imidlertid for at moralsk ansvar også inneholder "objektive" elementer. Og de imøtegår Peter Strawsons utsagn om at det ikke tillates eksternt rasjonelt forsvar av reaktive holdninger. Dette er grep som blir kritisert i avhandlingen. Peter Strawsons syn på eksternt rasjonelt forsvar av reaktive holdninger blir forsvart mot Fischer og Ravizzas kritikk. Og det blir også rettet kritikk mot Fischer og Ravizzas subjektivistiske tilnærming til moralsk ansvar og å ta ansvar.

I stedet for en subjektivistisk tilnærming blir det skissert en sosial tilnærming til moralsk ansvar. Ut fra en slik tilnærming har en aktør moralsk ansvar innen et sosialt system når systemet på legitimt vis vurderer at han har moralsk ansvar.

Forord

Jeg takker min veileder Kjell Eyvind Johansen for kritiske og gode råd, og for å ha vist stor tålmodighet, under arbeidet med avhandlingen. Jeg takker også "seniorgruppa på filosofi" v/gruppas innpisker og inspirator Ellen Voldstad for god og varm støtte og stimulerende og spennende diskusjoner.

Innholdsfortegnelse

Sammendrag	3
Forord	5
Innholdsfortegnelse	7
INTRODUKSJON	9
1. Innledning	9
1.1 Formål, begreper, metode	9
1.2 Noen filosofiske hovedretninger	12
DEL I FISCHER OG RAVIZZAS TEORI OM MORALSK ANSVAR	14
2. Sentrale elementer i Fischer og Ravizzas teori	14
2.1 Oversikt	14
2.2 En grunnholdning til moralsk ansvar	16
2.3 Hva er moralsk ansvar?	17
2.4 Betingelser for moralsk ansvar	19
2.5 Et trekk ved Fischer og Ravizzas argumentasjonsteknikk	22
3. Drøfting av noen forutsetninger og grunnleggende begreper	23
3.1 Deterministisk agnostisisme	23
3.2 Determinisme og indeterminisme	25
3.3 Determinisme og mangel på alternative muligheter	26
3.4 Semikompatibilisme og frihet	28
3.5 Fri vilje og autonomi	29
4. Determinisme, moralsk ansvar og rasjonell overveielse	32
4.1 Om den deterministiske utfordringen	32
4.2 Rasjonell overveielse og alternative muligheter	34
5. Mekanisembasert perspektiv	35
6. Reaktive holdninger	37
6.1 Peter Strawsons teori om reaktive holdninger og moralsk ansvar	38
6.2 Fischer og Ravizzas svar til Strawson	39
6.3 Vurdering av Fischer og Ravizzas kritikk	42
6.4 Sosial praksis assosiert til reaktive holdninger	47
7. Determinisme og legitim klander	50
7.1 Perebooms kritikk av Fischer og Ravizza	51
7.2 Fischers og Perebooms diskusjon	52
7.3 Om legitim klander	53
8. Moralsk ansvar og kontroll	55
8.1 Regulativ kontroll og veiledningskontroll	55
8.2 Respons på grunner	57
8.3 Eierskap til den operative mekanismen og å ta ansvar	60
8.4 Joseph Raz' kritikk	61
8.5 Vurdering av kontrollbegrepet	63

9.	Eierskap til den operative mekanismen	67
9.1	Eleonore Stumps kritikk	67
9.2	Fischer og Ravizzas syn	70
9.3	Å ta ansvar som historisk fenomen	72
9.4	Noen spørsmål	74
10.	Drøfting av det teoretiske grunnlaget	80
10.1	Galen Strawsons subjektivistiske teori om fri vilje og ansvar	80
10.2	Vurdering	84
DEL II EN ALTERNATIV TILNÆRMING TIL ANSVAR		88
11.	En skisse	88
11.1	Grunnlaget	88
11.2	En sosial tilnærming til moralsk ansvar	90
11.3	Utfordringer	96
AVSLUTNING		97
Litteraturliste		99

INTRODUKSJON

1. Innledning

1.1 Formål, begreper, metode

Formål og oppgave

Avhandlingens formål er å svare på disse problemstillingene: Kan vi ha moralsk ansvar for våre handlinger dersom determinismen er sann, dvs. at handlingene er determinerte? Og hvordan skal i så fall moralsk ansvar beskrives?

John Martin Fischer og Mark Ravizza behandler disse problemstillingene i sin bok "Responsibility and Control" (Fischer and Ravizza 1998). De hevder at de elementene som betinger moralsk ansvar ikke er avhengig av stillingtaken til kausal determinisme. De tar ikke stilling til om våre handlinger er determinerte. De tar heller ikke stilling til spørsmålet om vi kan handle annerledes enn vi faktisk gjør (dvs. om vi har alternative muligheter), eller om vi har fri vilje, men hevder at moralsk ansvar finnes, selv om en ikke tar stilling til disse spørsmålene. Denne filosofiske posisjonen blir kalt *semikompatibilisme*. De viktigste elementene i Fischer og Ravizzas teori¹ om moralsk ansvar er:

- En moralsk ansvarlig person ser seg selv som passende kandidat for reaktive holdninger.
- Moralsk ansvar er betinget av en form for kontroll, som Fischer og Ravizzas kaller veiledningskontroll (*guidance control*).

Fischer og Ravizza gir dermed sitt svar på avhandlingens problemstillinger. De svarer imidlertid ikke fullt og helt på problemstillingene, for de behandler bare *en* betingelse, som de kaller for den frihets- eller kontrollbetingelsen for moralsk ansvar.

¹ Med ordet *teori* vises i denne avhandlingen til den underliggende strukturen som ligger til grunn for systematisk tenkning, som strukturen i Fischer og Ravizzas tanker om moralsk ansvar i boka *Responsibility and Control*. Det stilles ikke spesifiserte vitenskapsteoretiske krav for at tenkningen skal ha karakter av teori.

Avhandlingens oppgave er knyttet til forholdet mellom problemstillingene og det svaret som Fischer og Ravizza gir på disse problemstillingene i hovedsak i boka "Responsibility and Control", men i noen grad også i andre verker. Oppgaven faller i to deler. Del I består i en vurdering av Fischer og Ravizzas teori om moralsk ansvar, slik denne blir uttrykt i første rekke i boka "Responsibility and Control" og dernest i noen andre deler av deres forfatterskap og hos forfattere som de bygger på. I avhandlingens Del II skisserer jeg en alternativ tilnærming til moralsk ansvar.

Noen sentrale begreper

De følgende definisjonene hentes fra de forfatterne som er sentrale for avhandlingen: John Martin Fischer, Mark Ravizza, Robert Kane, Derk Pereboom og Peter van Inwagen. Disse forfatterne definerer imidlertid flere ord ulikt. Derfor er det behov for en klargjøring

Alternative muligheter. Synonym² i denne avhandlingen: Alternative handlingsmuligheter. En aktørs mulighet til å handle annerledes enn han faktisk gjør. Eller i fortidsform: En aktørs mulighet til å ha handlet annerledes enn han faktisk gjorde. Blir også omtalt som frihet, evne eller makt (*power*) til å gjøre noe annet.

Atferd. Fischer og Ravizza bruker atferd (*behavior*) som en generell betegnelse for handlinger, unnlater og produksjon av konsekvenser av handlinger og unnlater. I avhandlingen har ikke unnlater og konsekvenser noen sentral plass. *Handling* blir derfor brukt som generell betegnelse, dvs. i tilfeller hvor det også kan inkludere unnlater og konsekvenser.³

Determinisme. Synonym i denne avhandlingen: Kausal determinisme. Fischer og Ravizza bruker uttrykket *kausal determinisme*, og definerer det slik: Tesen om at en fullstendig påstand om faktaene på et gitt tidspunkt, sammen med en fullstendig påstand om naturlovene, medfører all sannhet om hva som skjer etter det gitte tidspunktet (Fischer og Ravizza 1998: 14).

² Synonymer er ord og uttrykk med lik betydning.

³ På tilsvarende måte brukes også ordet *aktør* på generelt vis, både om et menneske som handler og et menneske som unnlater å handle.

Fischer og Ravizza bruker trolig uttrykket *kausal determinisme* for å skille det fra teleologisk determinisme (fatalisme, Guds styrelse osv.). Imidlertid defineres ordet *kausal* også på mange ulike måter, og determinisme blir derfor ikke så mye mer presist med tillegget *kausal* enn uten dette tillegget. Bortsett fra i referanse til F&Rs tekst bruker jeg derfor uttrykket *determinisme*, ikke *kausal determinisme*.

Frihet. Frihet har ulik betydning hos de forfatterne som det refereres til. Det knyttes blant annet til muligheten for alternativ handling (Kane 2002: 18), og til moralsk ansvar (Fischer 2006: 6). Siden ordet er så mangetydig, prøver jeg å unngå å bruke det. Det blir imidlertid brukt med referanse til enkelte forfattere i den betydningen som de tillegger ordet.

Fri vilje. To definisjoner av fri vilje er akseptable, sett i lys av denne avhandlingens formål (Fischer et.al. 2007: 1 - se for øvrig drøfting av begrepet *fri vilje* i kapittel 3):

1. Fri vilje krever ”alternative muligheter” eller makt (*power*) til å handle annerledes.
2. Fri vilje krever at vi er ”ultimate kilder” for våre frie handlinger, eller ultimate kilder for vår vilje til å utføre frie handlinger.

Moralsk ansvar. I avhandlingen brukes også kortformen *ansvar*. I sin redegjørelse for hva moralsk ansvar er, tar Fischer og Ravizza det de kaller et Strawsoniansk perspektiv: å være moralsk ansvarlig er å være en passende kandidat for reaktive holdninger (F&R 1998: 7). Denne definisjonen blir videreutviklet og utdypet på flere måter. En kan si at en klargjøring av hva dette innebærer, er en sentral del av avhandlingens formål og innhold.

Metode

Avhandlingen har en problemdrøftende karakter. En viktig del av oppgaven består også i redegjørelse for og klargjøring av de sentrale begrepene. Det vil bli lagt noe vekt på tekstanalyse ved gjennomgang av sentrale verker.

Avhandlingen spenner over et stort felt. Det har noen konsekvenser for hvor dypt man kan gå inn i de enkelte deler – og hvor avgjørende argumentene kan sies å være. Det kan være av interesse at også Fischer og Ravizza viser en beskjedenhet når det gjelder å gi avgjørende argumenter for sine standpunkter:

... we aim to give...”philosophical explanations”, not to do ”coercive philosophy” ... (W)e are claiming that we have offered a strong plausibility argument...., but not an argument that any rational agent is compelled to accept. (Fischer and Ravizza 1998: 11)

1.2 Noen filosofiske hovedretninger

Det er to filosofiske hovedretninger i synet på moralsk ansvar – kompatibilisme og inkompatibilisme. Skillet er basert på de ulike oppfatninger om forholdet mellom fri vilje og determinisme, som også innebærer ulike syn på forholdet mellom moralsk ansvar og determinisme. Kompatibilismen hevder at determinisme og fri vilje (og moralsk ansvar) er kompatible. Inkompatibilismen hevder at determinisme og fri vilje (og moralsk ansvar) er inkompatible.

Debatten mellom kompatibilisme og inkompatibilisme består i en debatt mellom to tilsynelatende riktige ideer, nemlig at mennesket har fri vilje og at alt har en årsak. Spørsmålet er så om noen av disse ideene må vike, eller om uoverensstemmelsen mellom de stridende ideene kan løses.

Kompatibilistene hevder at uoverensstemmelsen kan løses. De hevder at vi har fri vilje selv om kausal determinisme er sant. Eller sagt på en annen måte: Spørsmålet om mennesker har fri vilje eller moralsk ansvar er ikke avhengig av stillingtaken til kausal determinisme. I artikkelen ”Free Agency” (1975) skiller Gary Watson mellom ulike kompatibilistiske retninger og setter et vannskille ved Harry Frankfurts arbeider “Alternate Possibilities and Moral Responsibility” (1969) og “Freedom of the Will and the Concept of a Person” (1971). Han hevder at tidligere kompatibilistiske filosofer som David Hume har brukt et frihetsbegrep som innebærer at en person er i stand til å handle annerledes enn han faktisk gjør. Frankfurt legger til grunn at en person har fri vilje hvis han har frihet til å ville det han ønsker å ville. Han hevder at vi kan ha fri vilje på denne måten uten å være i stand til å handle annerledes enn vi faktisk gjør.

Semikompatibilismen utgjør en egen kompatibilistisk retning. I følge semikompatibilismen er spørsmålet om mennesker har moralsk ansvar ikke avhengig av stillingtaken til om vi har fri vilje. Dette standpunktet tas av Fischer og Ravizza i boka ”Responsibility and Control” som gjennomgås i denne avhandlingen.

Inkompatibilistene hevder at determinisme og fri vilje utelukker hverandre. En kan trekke frem to grupper av inkompatibilisme: libertarianisme og hard inkompatibilisme. Libertarianere hevder at et totalt fravær av fri vilje er utenkelig (se f.eks. Peter van Inwagen "An Essay on Free Will" 1983). Libertarianere finner derfor at determinisme er usant. Hard inkompatibilisme innebærer et deterministisk og inkompatibilistisk ståsted som har til konsekvens at fri vilje ikke finnes.

Noen filosofer faller utenfor skjemaet over kompatibilisme og inkompatibilisme. En kan her trekke frem Peter Strawson og Galen Strawson. Verken far eller sønn Strawson kan uten nærmere drøfting klassifiseres innen en av de to filosofiske hovedretningene om moralsk ansvar.

DEL I FISCHER OG RAVIZZAS TEORI OM MORALSK ANSVAR

2. Sentrale elementer i Fischer og Ravizzas teori

2.1 Oversikt

Det er mange filosofer som hevder et kompatibilistisk standpunkt, altså at fri vilje (og moralsk ansvar) kan gå sammen med determinisme og mangel på alternative muligheter.⁴

Det er derimot få filosofer som har tatt på seg oppgaven å gjøre rede for moralsk ansvar uten å knytte det til fri vilje. En av de få filosofer som har gjort dette, er John Martin Fischer, til dels i samarbeid med Mark Ravizza.⁵

Denne filosofiske posisjonen blir kalt for semikompatibilisme. På dette grunnlaget har Fischer og Ravizza utviklet en teori om moralsk ansvar. De sentrale elementene i teorien er:

1. Det kan ikke utelukkes at kausal determinisme er sant
2. Det kan ikke utelukkes at kausal determinisme er usant
3. Moralsk ansvar finnes, uavhengig av stillingtaken til spørsmålet om kausal determinisme er sant eller usant
4. Personers særegne ansvarsform er moralsk ansvar. Andre skapninger kan ha andre former for ansvar
5. Å være moralsk ansvarlig er å se seg selv som passende kandidat for reaktive holdninger.
6. Ved studiet av moralsk ansvar må man se på mekanismen som ligger til grunn for utøvelsen av ansvar (den operative mekanismen), og ikke bare på aktøren som utøver ansvaret
7. Moralsk ansvar er betinget av veiledningskontroll
8. Veiledningskontroll er betinget av:

⁴ Mangel på alternative muligheter innebærer ikke nødvendigvis kausal determinisme. Selv om bevegelsesmønsteret i atomkjernene synes å ha en indeterministisk karakter, innebærer ikke det med nødvendighet at de kan bevege seg på en annen måte enn de faktisk gjør.

⁵ Synspunkter som har referanse til flere av Fischers verker blir relatert til Fischer alene. Men konkrete tekstreferanser blir naturligvis knyttet til det enkelte verk, enten til ett av Fischers verker eller til hans samarbeid med Mark Ravizza.

- moderat respons på grunner
- den operative mekanismen er aktørens egen, dvs at aktøren har eierskap til den operative mekanismen

I kortform kan en si at Fischer og Ravizzas teori er tosidig. For det første gir de en redegjørelse for hva moralsk ansvar etter deres syn *ikke* er avhengig av, nemlig kausal determinisme. For det andre beskriver de hva moralsk ansvar er konstituert og betinget av: reaktive holdninger og veiledningskontroll.

På dette grunnlaget har avhandlingens Del I fått denne strukturen:

I kapittel 2 presenteres hovedelementene i Fischer og Ravizzas teori om moralsk ansvar. I kapitlet beskrives også deres grunnholdning til moralsk ansvar og et trekk ved deres argumentasjonsteknikk. I kapittel 3-4 presenteres og drøftes noen premisser for deres teori, nemlig at de ikke tar stilling til spørsmålene om determinisme, alternative muligheter og fri vilje. Deres mekanismebaserte syn på moralsk ansvar, som også kan sies å utgjøre et premiss for teorien, behandles i kapittel 5. I kapittel 6-9 presenteres og drøftes hovedelementene i teorien, nemlig at moralsk ansvar er nært knyttet til reaktive holdninger, ros og klander, og at det er betinget av veiledningskontroll. Fischer og Ravizza har en subjektivistisk tilnærming til sin teori om moralsk ansvar. I kapittel 10 blir det gitt en redegjørelse og drøfting av denne tilnærmingen.

Jeg benytter meg av en rekke forfattere til hjelp for drøfting av Fischer og Ravizzas teori. Drøftingen av forholdet mellom kausal determinisme og moralsk ansvar (tema nr. 1-3 fra listen ovenfor) berører hovedretningene i diskursen om determinisme og fri vilje. Her blir det vist til Peter van Inwagen som representant for libertarianismen og Derk Pereboom som representant for hard inkompatibilisme. En utfordring formulert av Edmund Henden blir også belyst. Ideen om at moralsk ansvar bygger på reaktive holdninger (tema nr. 5) har Fischer og Ravizza hentet fra Peter Strawson. Standpunktet om at moralsk ansvar er betinget av en form for kontroll – nemlig veiledningskontroll (tema nr. 6-8) - blir utfordret av Joseph Raz. Når det gjelder ideen om at veiledningskontroll er betinget av eierskap til den operative mekanismen, knytter Fischer og Ravizza dette til synspunkter fra Eleonore Stump og til en subjektivistisk teori utviklet av Galen Strawson.

Blant de kritikerne av F&R som er trukket fram, er det bare Pereboom som gir konkret kritikk av Fischer og Ravizzas standpunkter og tekster. De andre forfatterne begrunner sitt eget filosofiske standpunkt uten referanse til Fischer og Ravizza, og også uten noen spesifikk referanse til deres semikompatibilistiske posisjon. Når det f.eks. gjelder van Inwagens libertarianske argumenter, kan en si at disse er rettet mot kompatibilismen i sin alminnelighet. Van Inwagens kritikk er ikke spesifikt rettet mot Fischer og Ravizzas semikompatibilisme. Etter mitt syn er imidlertid kritikken fra disse forfatterne relevant, selv om den ikke er direkte rettet mot Fischer og Ravizza og deres filosofi.

Med det opplegget som er skissert her, blir ikke alle deler av Fischer og Ravizzas bok "Responsibility and Control" like grundig behandlet. Det legges vekt på behandlingen av de to betingelsene for veiledningskontroll: respons på grunner og eierskap til den operative mekanismen. Dessuten vektlegges deres synspunkter på utfordringen fra determinismen. Det legges ikke vekt på ansvar for unnlatelser og konsekvenser. Fra deres tidligere bok "Perspectives on Moral Responsibility" (1993) trekkes særlig fram deres drøfting av Strawson syn på eksternt rasjonelt forsvar for reaktive holdninger.

I min drøfting har jeg særlig vektlagt de emnene hvor jeg tviler på Fischer og Ravizzas argument: om eierskap til den operative mekanismen og om eksistensen av eksternt rasjonelt forsvar for reaktive holdninger. Jeg har langt på vei utelatt noen egen drøfting av de emnene hvor jeg er enig med F&R; nemlig om utfordringen fra determinismen, og om respons på grunner. I noen tilfeller blir det bare stilt spørsmål, mens det anses å falle utenfor avhandlingens ramme å forsøke å besvare spørsmålet. Dette gjelder f.eks. spørsmålet om forholdet mellom veiledningskontroll og kompetanse, som blir stilt i forbindelse med behandlingen av Joseph Raz' artikkel om ansvar og kontroll. Det gjelder også flere enkelt-elementer, særlig knyttet til de to betingelsene for veiledningskontroll - moderat respons på grunner og eierskap til den operative mekanismen.

2.2 En grunnholdning til moralsk ansvar

Fischer gir et synspunkt på forholdet mellom determinisme og moralsk ansvar i boka "My Way":

.... our basic status as distinctively free and morally responsible agents should not depend on the arcane ruminations ... of the theoretical physicists and cosmologists.....Our fundamental nature as

free, morally responsible agents should not depend on whether the pertinent regularities identified by the physicists have associated with them (objective) probabilities of 100 percent (causal determinism) or say, 98 percent (causal indeterminism). Given that we think of ourselves as morally responsible agents in control of our behaviour (in the relevant way), how could the discovery that laws of nature have 100 percent probabilities associated with them, rather than 98 percent....., make us abandon our view of ourselves as persons, as morally responsible agents in control of our behaviour? This just seems highly implausible and unattractive to me. (Fischer 2006: 6)

Fischer skiller her mellom fysikk på den ene side og moral og selvforståelse på den annen. Han sier at vår selvforståelse ikke kan bygge på de skiftende sannheter i utviklingen innen naturvitenskapen. En kan si at han har en grunnleggende oppfatning om at moralsk ansvar finnes. Han vil holde på ideen om at vi er moralsk ansvarlige aktører uavhengig av vitenskapelige funn i nåtiden og framtiden.

2.3 Hva er moralsk ansvar?

Fischer og Ravizza åpner sin bok "Responsibility and Control" med å skrive om en vase som blir knust. De antar at du ville reagere på *en* måte hvis det var en klønete katt som uheldigvis hadde revet ned vasen, og at du hadde reagert svært forskjellig om det var en ondskapsfull gjest som hadde knust den. Du ville se på katten som kausalt ansvarlig, men ikke moralsk ansvarlig, mens gjesten ville være både kausalt og moralsk ansvarlig.⁶

Med dette eksemplet setter F&R moralsk ansvar i sammenheng med kausalitet. Moralsk ansvar er ikke en årsak, men synes å ha en nær sammenheng med visse årsakstyper. F&R finner kausalt ansvar tilknyttet mange ulike årsakstyper. Men *moralsk ansvar* er bare knyttet til enkelte årsakstyper, utført av visse aktører. Fischer og Ravizza kaller disse aktørene for personer. De søker å forklare hva moralsk ansvar er ved å skille mellom personer og andre skapninger (*creatures*). De skriver:

An important difference between persons and other creatures is that only persons can be morally responsible for what they do. When we accept that someone is a morally responsible agent, this typically involves more than holding a particular belief about him; it entails a willingness to adopt certain attitudes toward that person and to behave toward him in certain ways. (Fischer and Ravizza 1998: 1)

⁶ Fischer og Ravizza understreker at det i tillegg til moralsk ansvar finnes en rekke andre former for ansvar (innenfor "den menneskelige sfære" – mitt tillegg), for eksempel juridisk ansvar, organisasjonsansvar, rolleansvar osv. En kunne ha ønsket seg en drøfting av forholdet mellom moralsk ansvar og disse andre formene for menneskelig ansvar. Men Fischer og Ravizza presenterer ikke noe forsøk på å definere det moralske ansvars plass i et totalbilde av menneskelig ansvar.

I følge F&R kan vi reagere med krenkelse, indignasjon, takknemlighet eller liknende overfor moralsk ansvarlige personer,⁷ men ikke overfor andre skapninger eller ikke-levende fenomener. Således vil ikke en forsker bli krenket av en rotte som agerer på en måte som avkrefter hypotesen i et vitenskapelig forsøk. Forskeren kan bli frustrert og skuffet over rottas uventede handlingsmønster, ”... but it seems wholly out of place to feel *indignant* at the rat or to *resent* the rat for its behavior” (op.cit.: 6). Disse holdningene (krenkelse osv.) kaller Fischer og Ravizza for reaktive holdninger.

Det er et tilsvarende forhold mellom moralsk ansvar og ros og klander (eller kritikk – altså negativ kritikk – hvis en vil bruke litt mer moderne språkform). Man kan gi en person ros eller klander. Men F&R hevder at en katteeier ikke vil klandre sin katt som river ned en vase. Katten har dermed heller ikke moralsk ansvar for å rive ned vasen.

Det er ikke bare dyr og døde gjenstander som er uten moralsk ansvar. Heller ikke alle mennesker har moralsk ansvar. Siden moralsk ansvar er en betingelse for å være en person, er ikke ethvert menneske en person, i følge Fischer og Ravizza.⁸

En person kan ha moralsk ansvar for klanderverdige handlinger, for rosverdige handlinger, og også for moralsk nøytrale handlinger (op.cit: 8). (Hvis det aksepteres at man kan ha moralsk ansvar for moralsk nøytrale handlinger, kan uttrykket være uheldig og misvisende. Kanskje bør det vurderes å bruke et annet uttrykk, for eksempel *personlig ansvar*.)

For Fischer og Ravizza gir beskrivelsen av de reaktive holdningene en nøkkel til forståelsen av moralsk ansvar. De hevder at reaktive holdninger konstituerer moralsk ansvar. De gir imidlertid ikke en fast og konsekvent beskrivelse av moralsk ansvar relatert til reaktive

⁷ Å ha moralsk ansvar og å være moralsk ansvarlig (person) blir brukt som synonymer. Begge uttrykkene refererer til en persons ansvar for en handling (eller unnløstelse), uavhengig av om handlingen er gjenstand for ros eller klander. Å være moralsk ansvarlig må altså ikke forveksles med å være ansvarsfull, som normalt er rosverdig, og kun unntaksvis kan være klanderverdig.

⁸ I følge Fischer og Ravizza er en person et menneske som har moralsk ansvar. Et menneske uten moralsk ansvar er ikke en person. Jeg bruker ikke den terminologien. Bortsett fra i direkte referanse til Fischer og Ravizzas tekst setter jeg person identisk med menneske. Det kan også stilles spørsmål om et dyr kan ha moralsk ansvar, og derfor er en person. Jeg avviser ikke at dyr kan ha moralsk ansvar, men jeg ser bort fra denne mulige betydningen av ordet *person*. Her brukes dette ordet bare om mennesker.

holdninger. Deres bok kan sies å være en innføringsbok, hvor leseren får gradvis dypere innsikt i hva moralsk ansvar, og beskrivelsen av moralsk ansvar endrer seg med dette. Fischer og Ravizza kan sies å ende med denne definisjonen:⁹ Å ha moralsk ansvar er å se seg selv som passende kandidat for reaktive holdninger.¹⁰ Men det er ikke et entydig og uproblematisk forhold mellom reaktive holdninger og moralsk ansvar. Dette blir det gjort rede for i kapittel 6.

2.4 **Betingelser for moralsk ansvar**

Hva kreves for å være moralsk ansvarlig, og hva kjennetegner de mennesker som ikke er det? Som et grunnlag for drøftingen av dette spørsmålet legger Fischer og Ravizza vekt på en viktig distinksjon:

... it is useful to distinguish between someone *being* a morally responsible agent and such an agent being responsible *for* particular events such as actions, omissions, and consequences. Following this taxonomy, we may maintain our earlier distinction between the houseguest and the kitten – that the guest is a morally responsible agent, whereas the kitten is not – and at the same time consistently maintain that *in this particular instance* the guest is *not* morally responsible for breaking the vase, since the damage resulted from something – say, a seizure – over which she had no control. (Fischer and Ravizza 1993: 6)

En kan altså skille mellom to former for unntak fra moralsk ansvar:

- aktøren er ikke moralsk ansvarlig aktør (omtalt i avhandlingen som *aktørbasert ansvar*)
- aktøren er moralsk ansvarlig, men han var ikke moralsk ansvarlig for den aktuelle atferden (omtalt som *tidsbegrenset, situasjonsbetinget* eller *atferdsbetinget ansvar*)

Fischer og Ravizza skriver at de har utviklet et syn på moralsk ansvar “that takes its cue from Peter Strawson’s landmark essay, ‘Freedom and Resentment’” (F&R 1998: 5). På samme måte som F&R skiller også Strawson mellom to kategorier av unntak for moralsk ansvar. Han hevder at den ene kategorien kan karakteriseres ved påstander som “Han er bare et barn” eller “Han er håpløs schizofren”. Til den andre kategorien

⁹ Dette er ikke en vitenskapsfilosofisk avhandling, og vitenskapsteoretiske uttrykk som *teori, definisjon, begrep, årsak* o.l. blir brukt uten stor grad av presisjon.

¹⁰ Behandlingen av moralsk ansvar blir av praktiske grunner noe avgrenset i avhandlingen. Kollektivt moralsk ansvar blir ikke inkludert. Ansvar for unnlaterer blir i liten grad tatt med. Ansvar for konsekvenser av handlinger og unnlaterer blir også i liten grad berørt, iallfall når det gjelder de indirekte og/eller langsiktige konsekvensene. Ansvar for andre mennesker blir inkludert i den grad det dreier seg om *atferd* overfor andre mennesker.

... belong all those which might give occasion for the employment of such expressions as 'He didn't mean to', 'He hadn't realized', 'He didn't know'; and also all those which might give occasion for use of the phrase 'He couldn't help it', when this is supported by such phrases as 'He was pushed', 'He had to do it' (Strawson: 50)¹¹

Når det gjelder disse utsagnene av den andre kategorien skriver Strawson at de har noe viktig til felles (selv om de også har betydningsfulle ulikheter):

None of them (*dvs. utsagnene*) invites us to suspend towards the agent, either at the time of his action or in general, our ordinary reactive attitudes. They do not invite us to view the *agent* as one in respect of whom these attitudes are in any way inappropriate. They invite us to view the *injury* as one in respect of which a particular two ways one of these attitudes is inappropriate. They do not invite us to see the *agent* as other than a fully responsible agent. They invite us to see the *injury* as one for which he was not fully, or at all, responsible. (op.cit: 50-51)

Strawson skriver ikke mye om betingelsene for å unnta en person eller en handling (eller unnløstelse) for moralsk ansvar. Men Fischer og Ravizza bidrar til en fortolkning, bl.a ved å vise til Aristoteles. De skriver: "These conditions are "negative" in the sense that they indicate in which an agent can fail to be morally responsible: ignorance and force" (F&R 1998: 12). Betingelsene for å bli unntatt for ansvar er altså knyttet til uvitenhet eller tvang. Fischer og Ravizza viser at de positive betingelsene for ansvar korresponderer med de negative betingelsene. På dette grunnlaget omtaler de de to positive betingelser for moralsk ansvar: en betingelse om fornuftbasert kunnskap (*epistemic condition*) og en frihetsrelevant betingelse (*freedom-relevant condition*). Den sistnevnte betingelsen omtales også som kontrollbetingelse, som også er den betegnelsen som blir brukt i avhandlingen. Fischer og Ravizzas fokus er rettet mot kontrollbetingelsen, som de beskriver slik:

It specifies that the agent must not behave as he does as the result of undue force; that is, he must do what he does freely. Alternatively, one could say that the freedom-relevant condition specifies that the agent must *control* his behaviour in a suitable sense, in order to be morally responsible for it. (op.cit.: 13)

I boka "Responsibility and Control" legger F&R vekt på kontroll heller enn frihet. Det er derfor naturlig å vektlegge den alternative forståelsen av kontrollbetingelsen, nemlig at aktøren må kontrollere sin atferd på en passende måte for å være moralsk ansvarlig for den.

¹¹ Sidetallene i Strawsons artikkel viser til sidetall i Fischers antologi "Moral Responsibility" (1986)

Om kunnskapsbetingelsen skriver Fischer og Ravizza: "It captures the intuition that an agent is responsible only if he both knows the particular facts surrounding his action, and acts with proper sort of beliefs and intentions" (op.cit.: 13). De gjør det klart at det ikke blir gitt noen drøfting av denne betingelsen i deres bok.

Fischer og Ravizza gir noen eksempler på atferd som kan unntas fra moralsk ansvar. Som eksempel på uvitenhet omtaler de en person som kjører over en kattunge som han ikke ser. Sjøføren mangler relevant kunnskap knyttet til sin handling, og er derfor ikke moralsk ansvarlig ("it would be odd to judge that the driver is morally responsible for the kitten's untimely death", op.cit.: 12). Som eksempel på tvang nevner de personer som blir utsatt for manipulasjon, hjernevask, hypnose eller liknende. Gjennom den tvangen som de er utsatt for, mangler de den formen for frihet eller kontroll som kreves for moralsk ansvar.

Fischer og Ravizzas gjør rede for kontrollbetingelsen som de omtaler som veiledningskontroll, og hevder at denne i sin tur er betinget av moderat respons på grunner og eierskap til den operative mekanismen (evnen til å ta ansvar). I avhandlingen blir det stilt spørsmål om en alternativt kan se på veiledningskontroll som bare betinget av moderat respons på grunner, og ikke av eierskap til den operative mekanismen. Denne fortolkningen av moralsk ansvar kan karakteriseres som et minimumssyn, fordi det innebærer få betingelser for moralsk ansvar og få individer som blir unntatt fra moralsk ansvar. Det er dermed en fortolkning som innebærer at et stort antall personer *har* moralsk ansvar.

I tillegg til disse to synspunktene på moralsk ansvar kan det vises til Gary Watson, som hevder at deltakelse i det moralske fellesskapet er en betingelse for moralsk ansvar (Watson 1987: 271). Michael McKenna følger opp Watsons syn, og hevder at det er en forutsetning for moralsk ansvar at aktøren tilskrives eller tillegges moral (*is morally addressed*). McKenna trekker fram et eksempel med en person, Robert Harris, som brutalt myrdet to unge menn for å bruke bilen deres i forbindelse med et bankran. Han viderefører dette argumentet og skriver: "Harris cannot be morally addressed ... unless Harris understands our demands (but understanding here appears to imply a shared framework of values, and Harris lacks this)" (McKenna 1998: 128). I følge McKenna kreves det først at aktøren forstår samfunnets krav og kjenner dets moralske forventninger. Dermed kan det etableres et felles verdigrunnlag, og han kan bli medlem av det felles moralske samfunnet. På dette

grunnlaget kan han bli tillagt moral, som i sin tur er en betingelse for å være moralsk ansvarlig.

En kan si at oppfatningen til Watson og McKenna har karakter av et moralsyn på moralsk ansvar, og at oppfatningen har noe mer eksplisitt moralsk karakter enn F&Rs betingelser knyttet til respons til grunner og eierskap til den operative mekanismen. Men Fischer og Ravizzas argument for å inkludere *å ta ansvar* som betingelse for moralsk ansvar har likevel en viss likhet med Watsons argument for deltakelse i det moralske fellesskapet som betingelse for moralsk ansvar, i det begge argumentene forutsetter noen personlige trekk utover et rent minimum for å være moralsk ansvarlig.

I kapittel 9 blir det gitt en drøfting av Fischer og Ravizzas syn på moralsk ansvar betinget av å ta ansvar. Dette synet blir vurdert i forhold til det såkalte minimumssynet, hvor å ta ansvar ikke inkluderes som betingelse for moralsk ansvar. Watsons og McKennas syn på moralsk ansvar blir ikke vurdert.

2.5 Et trekk ved Fischer og Ravizzas argumentasjonsteknikk

Deler av Fischer og Ravizzas argumentasjonsteknikk blir knyttet til det som gjerne blir kalt Frankfurt-type historier (*Frankfurt-style cases*). I boka "Four Views on Free Will" kritiserer Robert Kane denne argumentasjonsteknikken:

... these "Frankfurt-type examples"... are pretty strange and esoteric examples on which to base a general belief that moral responsibility is compatible with determinism. Do they have anything to do with our ordinary world where there probably are no Frankfurt controllers (at least not yet)?... these unusual Frankfurt-type examples don't really prove what they are supposed to prove anyway. In particular, they do not show that moral responsibility is compatible with never having the ability to do otherwise and so they do not show that moral responsibility is compatible with determinism. (Fischer et.al. 2007: 168)

Kane hevder altså at Frankfurts eksempler er urealistiske, og at man ikke kan generalisere på grunnlag av dem. Fischer er langt på vei enig i dette. Riktignok forsvarer Fischer Frankfurt-eksemplene (Fischer et.al. 2007: 189), og skriver at selv om de ikke inntreffer i virkeligheten, illustrerer de hvordan friheten kan virke i en deterministisk virkelighet. Han mener altså at disse eksemplene gjør det lettere for oss å forstå en deterministisk virkelighet, som ellers er vanskelig for oss å fatte. Han understreker eksemplenes pedagogiske funksjon, ved at de gir et bilde – om enn lite realistisk – av en deterministisk virkelighet.

Jeg tar her et mellomstandpunkt. Jeg er enig med Fischer at Frankfurt-eksemplene kan ha en pedagogisk funksjon selv om de er lite realistiske. Men etter mitt syn kan de også i noen tilfeller virke forvirrende og kunne bidra til å bringe debatten inn i et lite fruktbart spor. I avhandlingen blir det vist et eksempel på dette.

3. Drøfting av noen forutsetninger og grunnleggende begreper

I dette kapitlet blir det gitt en drøfting av noen grunnleggende begreper: kausal determinisme (i denne avhandlingen også omtalt i kortformen *determinisme*), indeterminisme, alternative muligheter og fri vilje. Nærstående uttrykk, som frihet og autonomi, blir også berørt i drøftingen. Dessuten blir Fischer og Ravizzas agnostiske holdning til eksistensen av determinisme, alternative muligheter og fri vilje gjenstand for særlig drøfting.

Disse begrepene er svært komplekse og mangetydige. Drøftingen er derfor nødvendigvis begrenset, men jeg antar at den er akseptabel sett i forhold til avhandlingens formål.

3.1 Deterministisk agnostisisme

Fischer og Ravizza hevder at vi ikke vet om doktrinen om kausal determinisme er sann, men vi vet heller om den er usann (Fischer and Ravizza 1998: 14-15). Fischer kaller en slik posisjon hvor man ikke tar stilling til alternativene, for *agnostisisme*. Denne agnostiske posisjonen synes også å omfatte spørsmålet om den frie vilje. Med visse reserver som det blir gjort rede for i kapittel 3.5, tar han ikke stilling til om fri vilje er kompatibel med determinismen (Fischer et.al. 2004: 4).

Siden Fischer og Ravizza ikke tar stilling om determinismen er sann, kan en anta at deres teori om moralsk ansvar også bygger på denne forutsetningen om ikke-stillingtakeren, altså en forutsetning om deterministisk agnostisisme. Men er det tilfelle? Klarer de å opprettholde sitt agnostiske standpunkt, eller må de legge enten deterministiske eller indeterministiske forutsetninger til grunn for sin teori om moralsk ansvar?

I kapittel 8 blir det gjort rede for Fischer og Ravizzas to kontrollbegreper, veiledningskontroll og regulativ kontroll. Deres sentrale tese er at moralsk ansvar er betinget av veiledningskontroll. Moralsk ansvar kan ikke være betinget av regulativ kontroll, siden regulativ kontroll bl.a. forutsetter tilgang på alternative muligheter. På grunn av sin ikke-stilling-

taken til doktrinen om kausal determinisme kan F&R ikke slå fast at mennesker har tilgang på alternative muligheter, og de kan derfor ikke hevde at moralsk ansvar er betinget av regulativ kontroll, som nettopp forutsetter tilgang på alternative muligheter.

På tilsvarende vis kan de ikke hevde at moralsk ansvar er betinget av fri vilje, siden de forholder seg agnostisk både til determinismen og til den frie viljes kompatibilitet til determinismen.

Men ville Fischer og Ravizzas teori om moralsk ansvar blitt annerledes dersom de la en annen forutsetning til grunn og forutsatte at determinismen er usann? Selv skriver de at synet på determinisme ikke skal ha noen avgjørende innflytelse på vår oppfatning av oss selv som moralsk ansvarlige aktører: «...*the sort of discovery represented by causal determinism should not in itself issue in our changing our very basic ... view of ourselves as morally responsible agents*» (Fischer and Ravizza 1998: 15).

Jeg er enig i dette. Vi kan se på oss selv som ansvarlige, både med et deterministisk og et indeterministisk grunnsyn. Men blir *beskrivelsen* av ansvar lik i begge tilfeller?

Jeg tviler på det. Fischer og Ravizza gir god begrunnelse for at moralsk ansvar er betinget av veiledningskontroll, dersom vi ikke tar stilling til determinisme og fri vilje ("forstått på en spesiell måte" – jf. kap. 3.4). F&R's begrunnelse er solid og god (selv om man kan være uenig i noen elementer). Men jeg kjenner ingen filosofer som hevder at determinisme er usann og som begrunner moralsk ansvar på den samme måten.

Hvordan ville så Fischer og Ravizza bygd opp sin teori under forutsetning om at doktrinen om kausal determinisme er usann? Det kan jeg naturligvis ikke vite, men jeg kan slå fast at de nevnte hindrene for knytte moralsk ansvar til regulativ kontroll og fri vilje ville være borte. F&R kunne da ikke - med krav på seriøsitet - unnlate å ta opp spørsmålet om hvordan moralsk ansvar forholder seg til fri vilje og regulativ kontroll. Jeg mener at det ikke er usannsynlig at deres teori ville bli påvirket av en slik drøfting og blitt annerledes dersom de la til grunn at determinismen er usann.

Jeg ser det altså slik at F&R i *utgangspunktet* stiller seg agnostisk til spørsmålet om determinismen er sann og om fri vilje er kompatibel med determinismen. Men i sitt arbeid

unnlater de forutsetninger om indeterminisme, alternative handlingsmuligheter og fri vilje. I *realiteten* forutsetter de derfor at determinismen er sann og at fri vilje ikke er kompatibel med determinismen. En kan si at de erklærer seg som agnostikere, men at de i sin teoribygging opptrer som om de tror på determinismen og ikke tror på den frie vilje. Noe spissformulert kan en si at de har en agnostisk grunnholdning og fører et deterministisk resonnement.

Det er altså et dilemma knyttet til den agnostiske posisjonen. Mitt svar på dette dilemmaet representerer ikke noen endelig løsning, men det er likevel det minst dårlige svaret etter mitt syn:

1. Vi kan ikke vite om determinismen er sann eller usann.
2. Det er derfor viktig å undersøke begge alternativer – og konsekvensene av dem.
3. Fischer og Ravizza gir et bidrag til slik undersøkelse ved – i realiteten, selv om de har et annet utgangspunkt – å beskrive moralsk ansvar dersom determinismen er sann. Andre filosofer har gitt sitt bidrag ved å beskrive moralsk ansvar dersom determinismen er usann.
4. Jeg kan ikke si at slike undersøkelser vil gi svar på hvilket alternativ som er det rette. Men det er å håpe at slike undersøkelser vil kunne hjelpe den enkelte til å finne ut hva han tror (mest) på.

Jeg følger Fischer og Ravizza ved at jeg støtter deres agnostiske holdning og i hovedsak gjennomfører mine drøftinger under forutsetning om at determinismen er sann. Men i noen sammenhenger gir det også mening å trekke inn det agnostiske utgangspunktet.

3.2 Determinisme og indeterminisme

Fischer og Ravizza presenterer tre definisjoner med ulikt presisjonsnivå av kausal determinisme. Deres ”mellom-presise” definisjon er:

Causal determinism is the thesis that, for any given time, a complete statement of the facts about that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time. (op.cit.: 14)

Fischer og Ravizza erkjenner at mange av dagens fysikere vil si at doktrinen om kausal determinisme er falsk (bl. a. på grunn av studier av bevegelser i atomkjernene). Men de skriver også at nye undersøkelser kan vise at de indeterminerte bevegelsene som dagens teorier postulerer, kan være resultat av inadekvat informasjon og analyse, og at man i fremtiden kan utvikle en fullstendig deterministisk teori om universet.

Videre hevder de at indeterministiske bevegelser i atomkjernen og menneskets frie vilje har lite med hverandre å gjøre. Det er ikke lettere å forklare menneskets frie vilje og moralske ansvar, dersom en legger tilfeldige (*random*) indeterministiske bevegelser til grunn enn på grunnlag av et deterministisk syn. Tvert imot, i følge F&R blir utfordringen i å gjøre rede for menneskets moralske ansvar nesten den samme om man trekker inn indeterminisme på mikronivået som om man bygger på et rent deterministisk syn. De skriver:

Although contemporary physicists tend to believe that causal determinism is false, they believe that something very much like it is true: a doctrine we shall call "almost causal determinism."....We would claim that the threat from almost causal determinism to moral responsibility is very similar to the threat from causal determinism. (op.cit.: 15)

De finner derfor grunn til å ta utfordringen fra determinismen alvorlig selv om det er sterk oppslutning om indeterminismen blant dagens fysikere. Og de understreker at vi per i dag ikke vet om doktrinen om kausal determinisme er sann, men vi vet heller ikke om den er falsk.

3.3 *Determinisme og mangel på alternative muligheter*

I følge Fischer og Ravizza representerer doktrinen om kausal determinisme en utfordring for ideen om at vi har moralsk ansvar for våre handlinger. I blant trekker de også inn ideen om at mennesker mangler alternative muligheter til handling, og at dette kan representere en utfordring for moralsk ansvar. Blant annet er Frankfurt-type eksemplene orientert omkring ideen om det i gitte situasjoner ikke finnes alternative muligheter til handling.

Hva er så forholdet mellom determinisme og mangel på alternative muligheter? Følger den ene doktrinen med nødvendighet av den andre, eller finnes det tilfeller hvor det ene kan forekomme, men ikke det andre?

Etter mitt syn innebærer ikke mangel på alternative muligheter nødvendigvis kausal determinisme. Selv om bevegelsesmønsteret i atomkjernene synes å ha en indeterministisk karakter, innebærer ikke det at en kan påvise at elementene i atomkjernene på et gitt tidspunkt kunne ha beveget seg på en annen måte enn de faktisk gjorde. For meg gir det altså mening å tenke seg indeterminerte bevegelser uten alternative bevegelsesbaner.

Det kan synes lettere å forsvare ideen om mangel på alternative muligheter enn doktrinen om kausal determinisme. Doktrinen om kausal determinisme bygger på meget kompliserte vitenskapelige teorier og vitenskapsteoretiske overveielser. Ideen om mangel på alternative muligheter kan derimot bygge på følgende argument:

Vi er - som mennesker - ikke i stand til å gå tilbake i tid. Når person P gjorde handling a_1 på det fortidige tidspunktet t_1 , kan vi derfor ikke gå tilbake til tidspunktet t_1 og se om han kunne ha gjort handlingen a_2 istedenfor a_1 .

Dette er et forholdsvis enkelt, empirisk basert argument. På dette grunnlaget finner jeg at setningen ”*Det er ikke påvist empirisk at mennesket har alternative handlingsmuligheter*” står forholdsvis sterkt. Jeg antar at denne setningen er lettere å forsvare enn setningen ”*Det er ikke påvist at indeterminisme er sant*”.

Det er viktig at de forutsetninger som man bygger på, er enkle og lette å forsvare. Det er et argument for å basere avhandlingen på ideen om at det ikke er påvist at mennesket har alternative handlingsmuligheter.

Tilsynelatende argumenterer Derk Pereboom imot dette standpunktet. Med henvisning til Fischer og Ravizzas skille mellom aktuell og alternativ sekvens hevder han at man bør knytte prinsippene for forklaring av frihet og ansvar til den enkelte handlingens kausale historie, og ikke til alternative muligheter (Pereboom 2001: 37). Jeg tviler imidlertid på om en kan utlede dette fra F&Rs anbefaling om å studere mekanismen i den aktuelle sekvensen (se nærmere om dette i kap. 3.5).

Imidlertid bruker den refererte litteraturen oftere forutsetningen om determinisme enn forutsetningen om mangel på alternative muligheter. Det kan også nevnes at forskjellen mellom determinisme og mangel på alternative muligheter ofte er uten praktisk betydning. For eksempel legger Fischer og Ravizza liten vekt på denne forskjellen. Jeg følger denne praksis og legger språkbruken nær til de tekstene som refereres, uten å gå nærmere inn på de spørsmålene som er drøftet i dette underkapitlet.

3.4 Semikompatibilisme og frihet

Fischer og Ravizzas filosofiske posisjon blir kalt for semikompatibilisme. Semikompatibilismen kan bli definert på forskjellige måter. I debattboka "Four Views on Free Will" blir semikompatibilismen definert slik:

... the view that holds that responsibility is compatible with determinism, combined with agnosticism about whether free will understood in some particular way might not be compatible with determinism. (Fischer et.al. 2007: 4)

For å få klarhet i denne definisjonen er det viktig å vite hva Fischer mener med fri vilje. Fischer unngår i stor grad å bruke dette uttrykket, og han gir ingen presis definisjon. Jeg finner ikke noen annen knagg å henge hans *free will understood in some particular way* på enn *freedom to do otherwise*. Fri vilje (forstått på en spesiell måte) er altså frihet til å gjøre noe annet, eller frihet til alternative muligheter. Det følgende sitatet kan imidlertid gi inntrykk av en nyanseforskjell mellom *fri vilje* og *frihet til å gjøre noe annet*:

The account of guidance control ... helps us to reconcile causal determinism with moral responsibility for actions, even if causal responsibility is inconsistent with freedom to do otherwise. We shall contend that the case for incompatibility of causal determinism and freedom to do otherwise is different from (and stronger than) the case for the incompatibility of causal determinism and moral responsibility for actions. (Fischer and Ravizza 1998: 5)

Fischer uttrykker altså en agnostisisme om fri vilje (forstått på en spesiell måte) er kompatibel med determinisme. På tross av dette ser han et relativt sterkt grunnlag for inkompatibilitet mellom determinisme og frihet til å gjøre noe annet, sett i forhold til grunnlaget for inkompatibilitet mellom determinisme og moralsk ansvar. Jeg holder imidlertid fast ved at fri vilje forstått på en spesiell måte er frihet til å gjøre noe annet. Jeg forstår dette slik: Fischer finner at begrunnelsen for inkompatibilitet mellom determinisme og frihet til å gjøre noe annet (=fri vilje forstått på en spesiell måte) er sterkere enn begrunnelsen for inkompatibilitet mellom determinisme og moralsk ansvar. På dette grunnlaget har han utviklet sin semikompatibilisme som legger til grunn at moralsk ansvar er kompatibelt med determinisme, mens det kan være at frihet til å gjøre noe annet (=fri vilje forstått på en spesiell måte) ikke er kompatibel med determinisme.

Fischer og Ravizza bruker også ordet *frihet* i sammensetningen *frihetsrelevant betingelse*, som er en av betingelsene for moralsk ansvar (op. cit.: 13). Her brukes ordet *frihet* om mangel på tvang. Med den deterministiske forutsetningen dreier dette seg om en avgrenset

form for frihet, som ikke innebærer frihet til å gjøre noe annet. Den deterministiske forutsetningen innebærer også en avgrenset form for tvang. Den «deterministiske tvangen» som vi alle - i følge det deterministiske synet - er utsatt for, er ikke inkludert i denne formen for tvang. Den avgrensede formen for tvang omfatter fysisk tvang (for eksempel innesperring) og psykisk tvang (for eksempel hypnose). Fischer og Ravizza hevder også at man kan mangle veiledningskontroll og unntas fra moralsk ansvar hvis man ikke tar ansvar. Å ta ansvar blir da en komponent i frihetsbetingelsen for moralsk ansvar, og manglende tilfredsstillende av denne betingelsen er begrunnet i tvang. En kan derfor slutte at ikke å ta ansvar er en tvangsmessig defekt som medfører at aktøren mangler frihet, veiledningskontroll og moralsk ansvar. Når F&R bruker ordene *fri* og *frihet* uten nærmere spesifikasjoner, synes det å være i betydningen *mangel på tvang* slik det blir omtalt her.

Selv om Fischer og Ravizza omtaler frihet og fri vilje, har de ikke noen sentral plass i deres ansvarsbegrep – i motsetning til for eksempel hos Frankfurt. Frankfurt bruker begrepet *fri vilje* for å beskrive, begrunne og forklare menneskets moralske ansvar. F&R, derimot, uttaler seg nesten ikke om fri vilje, bortsett fra at Fischer hevder at han stiller seg agnostisk til spørsmålet om fri vilje – forstått på en spesiell måte - er kompatibel med kausal determinisme. I de verkene som jeg viser til, uttaler de seg heller ikke mye om frihet. Jeg vil, i likhet med Fischer og Ravizza, være forsiktig med bruken av ordene *frihet* og *fri vilje*. I det følgende vil jeg imidlertid se nærmere på en side ved *fri vilje*, og i tillegg se på forholdet mellom fri vilje og autonomi.

3.5 *Fri vilje og autonomi*

I boka "Four Views on Free Will" blir det gitt denne beskrivelsen av forsøkene på å definere *fri vilje*:

Much of the tradition has taken "free will" to be a kind of power or ability to make decisions of the sort for which one can be morally responsible, but philosophers have also sometimes thought that free will might be required for a range of other things, including moral value, originality and self-governance. Two other claims often made about free will are hotly disputed among philosophers; and authors of this volume will take different sides on these claims. One is the claim that free will requires "alternative possibilities" or the power to do otherwise, and the other is the claim that free will requires that we are "ultimate sources" of our free actions or the ultimate sources of our wills to perform free actions. (Fischer et.al. 2007: 1)

Jeg kan ikke se at Fischer foretar noe klart valg mellom disse definisjonene av fri vilje, men at han heller velger å være varsom med å bruke uttrykket. Dette sitatet ovenfor kan

imidlertid tjene som inngang til et innspill fra Derk Pereboom. I boka “Living without Free Will” drøfter Pereboom ulike intuisjoner om fri handling (*free action in the sense required for moral responsibility*). To av Perebooms intuisjoner har betydelig likhet med de to sistnevnte påstandene (*claims*) om fri vilje hos Fischer et.al. Det dreier seg om disse to intuisjonene:

- (1) An action is free in the sense required for moral responsibility only if the agent could have done otherwise than she actually did.
- (3) An action is free in the sense required for moral responsibility only if it is not produced by a deterministic process that traces back to causal factors beyond the agent’s control. (Pereboom 2001: 1-3)

Med henvisning til Fischer og Ravizzas skille mellom aktuell og alternativ sekvens (se kap. 5.1) foretrekker Pereboom intuisjon (3). Han viser at ved manipulering av hjernen til en person (Jones) gir den alternative sekvensen det samme resultatet som den aktuelle sekvensen. Men denne manipuleringen av hjernen trenger ikke være en del av den aktuelle sekvensen, og behøver derfor ikke være bestemmende for at Jones handler som han gjør i den aktuelle sekvensen. Pereboom følger Fischer i det synet at mekanismen i den aktuelle sekvensen er avgjørende for vurderingen om en aktør har ansvar for en handling i en gitt situasjon. Han uttrykker dette slik: “I maintain ... that the most plausible and fundamentally explanatory incompatibilist principles concern the causal history of an action, and not alternative possibilities.” (op. cit. 1-2)

Med henvisning til Fischer og Ravizzas skille mellom aktuell og alternativ sekvens, mener altså Pereboom at fri vilje skal knyttes til handlingens kausale historie, og ikke til alternative muligheter. Jeg tviler imidlertid på om en kan utlede dette fra F&Rs anbefaling om å studere mekanismen i den aktuelle sekvensen. Fischer og Ravizzas anbefaling inneholder et nesten selvinnsende råd, nemlig at selve analyseobjektet *moralsk ansvar* bør analyseres i forhold til den aktuelle sekvensen hvor ansvaret inngår. Deres anbefaling inneholder ikke noe råd om hvordan *fri vilje* skal defineres, og heller ikke noe råd om at en skal basere seg på en forutsetning om kausal determinisme framfor mangel på alternative muligheter.

I tillegg til at jeg tviler på Perebooms argument, kan det nevnes at hans standpunkt er tatt ut fra en inkompatibilistisk grunnholdning som jeg ikke deler.

Imidlertid er det ikke nødvendig for denne avhandlingens formål å avgjøre om fri vilje skal relateres til alternative muligheter eller ultimate kilder (*ultimate sources*). For mitt formål kan *fri vilje* forstås på begge disse to måtene.

En annen forståelse av fri vilje er utviklet av Harry Frankfurt. I følge Frankfurt er en handling fri når aktørens første ordens ønske fører til en handling som samsvarer med aktørens andre ordens ønske. Fri handling innebærer frihet til å handle slik aktøren ønsker å handle – både etter første og andre ordens ønske. Og fri vilje innebærer frihet til å ville det man ønsker å ville – også dette både etter første og andre ordens ønske. For at en aktør skal kunne ha fri vilje, må han derfor ha en viss kontroll over bevissthet og motivasjon, slik at han kan bringe første ordens ønske i samsvar med andre ordens ønske.

For Frankfurt er det en nær relasjon mellom fri vilje og autonomi.¹² I det følgende skal jeg gi et kort riss av noen ulike måter å forstå autonomi på, som har en parallell til ulike syn på moralsk ansvar. I den forbindelse kan en oversiktsartikkel i Stanford Encyclopedia of Philosophy, hvor det skilles mellom grunnleggende autonomi (*basic autonomy*) og ideell autonomi (*ideal autonomy*), være opplysende:

... we must keep separate the idea of basic autonomy, the minimal status of being responsible, independent and able to speak for oneself, from ideal autonomy, an achievement that serves as a goal to which we might aspire and according to which a person is maximally authentic and free of manipulative, self-distorting influences. (Christman, Stanford Encyclopedia: 3)

En viktig forskjell mellom ideell autonomi og grunnleggende autonomi er at ideell autonomi forutsetter en betydelig personlighetsutvikling, mens grunnleggende autonomi forutsetter svært liten grad av personlighetsutvikling. En annen forskjell er at ideell autonomi kan synes å være uoppnåelig, mens grunnleggende autonomi ikke er uoppnåelig, men tvert imot er en egenskap som de aller fleste mennesker har. Frankfurts fri vilje-begrep kan sies å stå i en mellomstilling mellom ideell autonomi og grunnleggende autonomi. Jeg antar at slik fri vilje ikke er uoppnåelig, men at den krever større grad av personlighetsutvikling enn det som kreves for grunnleggende autonomi.

¹² Se f.eks. Frankfurt, "Autonomy, Necessity, and Love", 129-141 og Velleman, "Identification and Identity", s. 91-123.

En kan sammenlikne betingelsene for de ulike fortolkningene av autonomi med de ulike fortolkningene av moralsk ansvar (se kap. 2.4). Betingelsene for moralsk ansvar etter minimumssynet svarer i noen grad til betingelsene for grunnleggende autonomi, selv om minimumssynet ikke eksplisitt omfatter betingelser om uavhengighet og evne til å snakke for seg selv. Fischer og Ravizzas betingelser for moralsk ansvar synes å være til dels snevrere og til dels mer omfattende enn betingelsene for grunnleggende autonomi. Watsons moralfortolkning av moralsk ansvar antas å være mer omfattende enn kravene til grunnleggende autonomi.

Det er mange filosofer som baserer ansvar på autonomi, dvs. at man må ha autonomi for å kunne ha ansvar. I denne avhandlingen brukes kontrollbegrepet istedenfor autonomi. Det er betydelige berøringspunkter mellom autonomi og kontroll, og mange vil hevde at autonomi er overordnet i forhold til kontroll. Men forholdet mellom autonomi og kontroll blir ikke vurdert i denne avhandlingen. Jeg nøyer meg med å vurdere forholdet mellom kontroll og ansvar.

4. Determinisme, moralsk ansvar og rasjonell overveielse

I denne avhandlingen blir det brukt to typer av argumenter mot determinismen. Fischer og Ravizza beskriver utfordringen for moralsk ansvar dersom alle våre handlinger er styrt av hendelser i fortiden. Et annet argument mot determinismen presenteres av Peter van Inwagen. Han hevder at rasjonell overveielse (*rational deliberation*) er inkompatibel med determinismen. Edmund Henden imøtegår dette, men påstår at aktøren må ha falsk tro dersom determinismen er sann.

Det blir gitt noen svar på utfordringene fra determinismen i dette kapitlet. Men en kan også si at hele avhandlingen representerer et forsøk på å svare på den deterministiske utfordringen til forestillingen om at vi har moralsk ansvar.

4.1 Om den deterministiske utfordringen

F&R gjennomgår tre argumenter for at vi *ikke* er moralsk ansvarlige dersom kausal determinisme er sann. De skiller mellom direkte og indirekte utfordringer til moralsk ansvar fra determinismen. Det sentrale prinsippet i den direkte utfordringen kalles Principle of the Transfer of Nonresponsibility. Dette prinsippet innebærer at hvis ingen er

moralsk ansvarlig for en ting p, og ingen er moralsk ansvarlig for at tingen p leder til tingen q, så er heller ingen moralsk ansvarlig for tingen q.

La oss si at p skjedde en gang langt tilbake i fortiden, og at q er en aktuell handling utført av personen T.

Hvis p skjedde en gang langt tilbake i fortiden da det ikke fantes mennesker, var ingen mennesker moralsk ansvarlig for p. Hvis det er naturlovene som styrer fra p gjennom et enormt antall mellomstadier til q, og determinismen er sann, er ingen mennesker moralsk ansvarlig for overgangen fra p til q. Det fremgår da av The Principle of the Transfer of Nonresponsibility at ingen er moralsk ansvarlig for den aktuelle handlingen q.

F&R presenterer to versjoner av den indirekte utfordringen fra determinismen. Den ene versjonen likner på eksemplet ovenfor. Den andre versjonen bygger på et prinsipp som kalles Principle of the Fixity of Past and Laws. F&R antar at personen S gjør handlingen A på tidspunktet T3. The Principle of the Fixity of Past and Laws innebærer at når situasjonen i fortiden på tidspunktene T1 og T2 er bestemt, og de aktuelle naturlovene som styrer personen S er bestemt, så må S gjøre A på tidspunktet T3. Det følger altså av dette argumentet at S ikke kan gjøre noe annet enn det han faktisk gjør – han mangler den type kontroll som innebærer alternative handlingsmuligheter.

Fischer og Ravizza avslutter denne drøftingen slik:

We have laid out very powerful and disturbing challenges to moral responsibility based on the truth of causal determinism. We human beings place great weight on the natural view that we are morally responsible agents in control of our behavior ... And yet a potent set of challenges can be mounted to these ideas. What makes these challenges particularly worrisome is that they themselves are based on cogent and seemingly unassailable aspects of common sense.... The challenges.... come from plausible and commonly accepted principles. If we are to be secure in our view of ourselves....., the challenges from causal determinism must be addressed. (Fischer and Ravizza 1998: 25)

I følge F&R består altså utfordringen i å løse uoverensstemmelsen mellom troverdige ideer. Deres svar er at det finnes en form for kontroll som ikke innebærer determinisme eller alternative muligheter, nemlig veiledningskontroll, og at moralsk ansvar er basert på denne formen for kontroll. I kapittel 8 blir det gjort rede for Fischer og Ravizzas begrep om veiledningskontroll.

4.2 Rasjonell overveielse og alternative muligheter

En type argument mot determinismen (eller rettere sagt: *for* alternative muligheter) blir reist av Peter van Inwagen. Han hevder at rasjonell overveielse (*rational deliberation*) forutsetter tro på – og et reelt valg mellom - alternative muligheter:

In my view, if someone deliberates about whether to do A or to do B, it follows that his behaviour manifests a belief that it is *possible* for him to do A – that he *can* do A, that he has within his power to do A – and a belief that it is possible to B. (van Inwagen 1983: 155)

For å klargjøre sitt standpunkt avgrensers van Inwagens sitt argument til visse typer av valg.¹³ For det første må det dreie seg om verdivalg hvor alternativene er inkommensurable. Man skal altså ikke kunne kalkulere hvilket alternativ som er det beste. Alternativene kan heller ikke være åpenbart ulikeverdige, slik at det ene alternativet er åpenbart bedre enn det andre. Ofte vil det dreie seg om valg mellom umiddelbar lyst og hensynet til andre, eller valget mellom umiddelbar lyst og fremtidige fordeler. Men andre typer av verdivalg kan også være aktuelle (Henden 2010: 3).

Van Inwagen utelukker ikke at den frie vilje griper inn i valg hvor man kan kalkulere hvilket alternativ som er det beste. Men han holder denne type av overveielser utenfor diskusjonen, fordi han anser det *mindre opplagt* at den frie vilje trer i kraft i slike kalkulerende overveielser enn i verdibaserte overveielser.

Van Inwagen mener at troen på at man har et reelt valg er en nødvendig betingelse for rasjonell overveielse. Han hevder at oppfatningen om at en aktør kan foreta rasjonell overveielse av ulike handlingsalternativer er inkompatibel med oppfatningen om at det finnes forutbestemte betingelser som umuliggjør alle alternativer unntatt ett.

Dette standpunktet blir kalt overveielsesinkompatibilisme (*deliberation incompatibilism*). I artikkelen ”Deliberation Incompatibilism” drøfter Edmund Henden dette standpunktet. Henden foretar en revisjon av van Inwagens syn, og hevder at rasjonell overveielse *ikke* er inkompatibel med determinisme. Men han framhever likevel at det er en forutsetning for

¹³ Gjennomgangen både av van Inwagens og Hendens synspunkter bygger på Hendens artikkel ”Deliberation Incompatibilism” (2010).

rasjonell overveielse at det finnes alternative muligheter. I følge Henden må en person som overveier, tro at det finnes alternative muligheter for å gjøre den ekstra anstrengelsen som overveielser krever. Hvis determinismen er sann, finnes ikke alternative muligheter. Hvis determinismen er sann, må derfor seriøse overveiere ha falsk tro om sin egen aktivitet mens de overveier. Hvis determinismen er sann, vil derfor den rasjonelle overveielser bygge på falsk tro, knyttet til den falske forestillingen at det finnes alternative muligheter. Dette, skriver han, vil være en trist nyhet for oss alle (Henden 2010: 20).

Fischer erkjenner at vi ofte tror at vi har reelle alternativer når vi overveier. Vi tror at vi kan velge hva vi skal gjøre, «... that it's up to us. All this is false, if determinism is true. We are, in that case, commonly deluded about the nature of our agency.» (Fischer 2006: 184). Fischer mener ikke som Henden at vi *må* ha falsk tro under overveielse, men han erkjenner at vi ofte har det. Han erkjenner dette som en realitet, men det later ikke til at han tar det like tungt som Henden gjør.

Andre filosofer står nærmere Henden i synet på at vi må ha falsk tro, men de synes likevel å kunne akseptere dette. Således skriver Galen Strawson at vi ikke er virkelig (*truly*) ansvarlige aktører i det hele tatt, selv om vi ikke kan la være å tro at vi er det (jf. kap. 10). Randolph Clarke interesserer seg også for Fischers syn, og kommenterer det slik:

Can we easily give up such [false] beliefs when we deliberate? I'm confident we can do it on occasion, if, say, we want to be sure not to presume something that, for all we know, might be false. But whether we can do it on a consistent basis is another matter. It strikes me that holding such beliefs is at least a very strong psychological habit. This aspect of our ordinary self-understanding might be resilient in the following sense: it might give way when pressed upon, but bounce right back once the pressure is off. (Clarke 2010: 247)

Jeg mener dette er en god beskrivelse av den deterministiske overveier. Den deterministiske overveier vil trolig gå inn i og ut av den falske tro på at han har reelle alternativer. Dette kan føles tungt, som Henden gir uttrykk for, men i følge Fischer og Clarke *må* det ikke føles så tungt - slik jeg forstår dem. Jeg deler denne sistnevnte oppfatningen.

5. Mekanisembasert perspektiv

Fischer og Ravizza anvender ulike tilnæringer og perspektiver på moralsk ansvar. Det mekanisembaserte perspektivet er ett av disse. Det gis en beskrivelse av det mekanisembaserte perspektivet.

baserte perspektivet, og det stilles spørsmål om dette perspektivets forhold til de andre perspektivene og tilnærmingene som de bruker i sin redegjørelse om moralsk ansvar.

Fischer og Ravizzas syn

Fischer og Ravizza understreker betydningen av at det legges et mekanismebasert - og ikke aktørbasert – perspektiv på analysen av moralsk ansvar og kontroll. For å klargjøre forskjellen mellom de to perspektivene presenterer F&R historien om verkstedeieren Leno. Leno har bestemt seg for å svindle kunden Mrs. Pratt. Men hvis Leno skulle forandre mening, vil sønnen Nick (som er nevrokirurg) manipulere Lenos hjerne, slik at han likevel vil komme til å svindle Mrs. Pratt. Historien har to alternative forløp, begge med det samme resultatet. I det ene forløpet – den aktuelle sekvensen – bestemmer Leno seg for å svindle Mrs. Pratt, og han fullfører sitt forsett. I den alternative sekvensen bestemmer han seg for *ikke* å svindle Mrs. Pratt; men på grunn av forhold som han ikke har kontroll over, svindler han henne likevel.

F&R hevder at i den aktuelle sekvensen er Leno ansvarlig for sin handling. Men i den alternative sekvensen forekommer en *ansvarsundergravende mekanisme* som medfører at Leno i dette tilfellet ikke er ansvarlig for handlingen. Dette viser, i følge F&R, at det er nødvendig ikke bare å studere *hva* Leno faktisk gjør, men også studere mekanismen (den operative mekanismen) som ligger til grunn for det som han faktisk gjør, og vurdere om denne er ansvarsundergravende eller ikke. Det er altså ikke tilstrekkelig med et aktør- og atferdsperspektiv, det må også anvendes et mekanismebasert perspektiv.

Fischer og Ravizza understreker også at vurderingen av Leno og mekanismen som ligger til grunn for hans atferd, skal bygge på beskrivelse av den aktuelle sekvensen - og ikke den alternative sekvensen. Den alternative sekvensen tjener først og fremst som et hjelpemiddel for å tydeliggjøre deres poeng.

Vurdering

Fischer og Ravizza forteller historien om Leno for å vise nødvendigheten av et mekanismebasert perspektiv. Dette innebærer ikke at aktørperspektivet helt og fullt kan erstattes med et mekanismeperspektiv. Det kan vises ved historien om Jennifer, som blir fortalt i kapittel 8. Jennifer foretar en del ufornuftige handlinger. F&R hevder at hvis Jennifer aldri

hadde handlet fornuftig, dvs. i samsvar med grunner, ville hun ha manglet veiledningskontroll og moralsk ansvar. Men hun handler av og til fornuftig, og de anser derfor at hun *har* veiledningskontroll og moralsk ansvar. Fischer og Ravizza gjør altså først en vurdering av Jennifer som person, og finner at hun ikke er utsatt for en ansvarsundergravende mekanisme. På dette grunnlaget finner de at hun har moralsk ansvar for sine handlinger.

Fischer og Ravizza anvender en rekke ulike perspektiver og tilnærminger til sin teori om moralsk ansvar. I tillegg til de perspektivene som er nevnt her, har de en subjektivistisk tilnærming til ansvar ved sin vektlegging av *å ta* ansvar. De har også et historisk perspektiv på ansvar ved at i det minste en av betingelsene for veiledningskontroll (eierskap til den operative mekanismen eller *å ta* ansvar) anses *å være* et historisk fenomen. De har lagt liten vekt på *å* vurdere forholdet mellom de ulike tilnærminger og perspektiver, og de har heller ikke drøftet om andre tilnærminger eller perspektiver burde ha vært trukket inn for *å* forklare hva moralsk ansvar er. Jeg kan i denne sammenhengen gi innspill til en tilnærming som Fischer og Ravizza i liten grad har trukket inn:

I den aktuelle sekvensen om Leno, hvor han ikke er manipulert, anses han ansvarlig for sin handling. Men i den alternative sekvensen hvor han blir manipulert av sin sønn, er han ikke ansvarlig. Leno skifter altså mellom *å være* og *ikke være* ansvarlig. Det er behov for en teori som forklarer hvordan slike skifter finner sted. I dette tilfellet inntreffer endringen ved innblanding fra en annen person, nemlig Lenos sønn Nick. På denne bakgrunn kan en spørre om en sosial tilnærming som bl.a. omfatter ulike former for sosial kommunikasjon og samhandling, vil kunne gi et fruktbart tilskudd til F&Rs teori om moralsk ansvar. Dette spørsmålet følges opp i avhandlingens kapitler 10-11.

6. Reaktive holdninger

Med utgangspunkt i Peter Strawsons essay "Freedom and Resentment" spør Fischer og Ravizza om moralsk ansvar består i *å være* mottaker (*recipient*) av reaktive holdninger. De finner at dette ikke er et tilstrekkelig krav for moralsk ansvar. Fischer og Ravizza imøtegår også Strawsons oppfatning om at det ikke kan gis eksternt rasjonelt forsvar for de reaktive holdningene. I dette kapitlet blir synspunktene til Strawson og F&R gjennomgått og vurdert.

6.1 Peter Strawsons teori om reaktive holdninger og moralsk ansvar

På grunnlag av Strawsons artikkel beskriver Fischer og Ravizza reaktive holdninger slik: ”Strawson uses the term ‘reactive attitude’ to refer to this range of attitudes that belong to [our] involvement or participation with others in interpersonal human relationships.” (Fischer og Ravizza 1998: 5). Strawson nevner mange eksempler på reaktive holdninger og trekker særlig fram krenkelse og takknemlighet (*resentment and gratitude*). I utgangspunktet dreier det seg om holdninger og følelser som aktiveres når andre menneskers handlinger reflekterer visse holdninger overfor oss («of goodwill, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other» (Strawson: 49)). Men en person også kan vise reaktive holdninger på vegne av andre – på grunn av handlinger utført mot andre.

Strawson omtaler også handlinger eller holdninger (*practices or attitudes*) assosiert med reaktive holdninger. Det dreier seg bl.a. om straff, klander, fordømmelse (*moral condemnation*) og anerkjennelse (*approval*). Strawson interesserer seg mer for de reaktive holdningene enn for sosial praksis basert på de reaktive holdningene:

These practices and attitudes (*straff, fordømmelse osv.*) permit, where they do not imply, a certain detachment from the actions or agents which are their objects. I want to speak, at least at first, of something else: of the non-detached attitudes and reactions of people directly involved in transactions with each other; of the attitudes and reactions of offended parties and beneficiaries: of such things as gratitude, resentment, forgiveness, love and hurt feelings. (Strawson: 48)

Strawson interesserer seg for reaktive holdninger fordi de er umiddelbare (*non-detached*). Normalt sett¹⁴ kan ikke et menneske være uten reaktive holdninger. Holdningene er dypt forankret i oss. De er noe som *er* der. De kan endres og justeres, men ikke fjernes:

Inside the general structure or web of human attitudes and feelings there is endless room for modification, redirection, criticism, and justification. But questions of justification are internal to the structure or relate to modifications internal to it. The existence of the general framework of attitudes

¹⁴ Det er uklart om en person kan være uten reaktive holdninger. Det må i så fall skje på denne måten: Reaktive holdninger har karakter av deltakende holdninger (*participant attitudes*). Motsatsen til deltakende holdninger er objektive holdninger. En person som bare har objektive holdninger, vil ikke ha noen reaktive holdninger. Strawson uttaler seg ikke om dette er mulig.

itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an external 'rational'¹⁵ justification. (op.cit. 64)

Det finnes altså en generell struktur eller et rammeverk av menneskelige holdninger og følelser. Sett i forhold til dette rammeverket skisserer Strawson to alternativer for justering, endring, kritikk og forsvar av menneskelige holdninger og følelser: enten med ståsted innenfor rammeverket (internt rasjonelt forsvar) eller utenfor rammeverket (eksternt rasjonelt forsvar). I følge Strawson kan det bare tillates forsvar for holdningene innenfor rammen av den generelle strukturen. Dette rammeverket av holdninger er noe som vi er gitt gjennom den realiteten som det menneskelige samfunn er. En kan ikke komme utenfor selve rammeverket og forsvare holdninger fra utsiden. Fischer og Ravizza omtaler Strawsons tanker på denne måten:

A central insight of Strawson's position is the suggestion that our questions about the property of the reactive attitudes should be taken as essentially practical questions, not theoretical ones Strawson holds that it is these reactive attitudes and practices themselves that are constitutive of responsibility. Thus, our ascriptions of responsibility do not find their ground in the truth of some metaphysical proposition (about, say, a creature possessing contra-causal freedom); rather, responsibility is grounded in nothing more than our adopting these attitudes toward one another. (Fischer and Ravizza 1993: 16)

F&R skriver altså at Strawson bygger ansvar på intet annet enn vår tilegnelse av reaktive holdninger. I kapittel 6.3 skal jeg se nærmere på om dette er en riktig tolkning av Strawson.

6.2 Fischer og Ravizzas svar til Strawson

Fischer og Ravizza har noen kritiske vurderinger av Strawsons teori. Deres kritikk er knyttet til disse temaene:

- Utfordringen knyttet til at noen samfunn kan holde mennesker ansvarlig som, intuitivt sett, ikke er det
- Spørsmålet om behandling av kritikk og endring av eksisterende praksis med grunnlag i Strawsons teori om moralsk ansvar
- Forholdet mellom reaktive holdninger og moralsk ansvar

¹⁵ Strawson setter 'rational' .i anførselstegn. I avhandlingen blir det imidlertid ikke gitt noen definisjon av rasjonell som kan bidra til å skille mellom Strawsons 'rational' justification og ordinær rasjonalitet. Jeg har derfor valgt å fjerne anførselstegnene.

Ansvarliggjøring av mennesker uten moralsk ansvar

Først retter de oppmerksomheten mot et samfunn – som vi kan kalle Andeby – hvor noen mennesker holdes ansvarlig selv om de intuitivt ikke er det:

.... one wonders what should be said about situations in which communities hold people responsible who intuitively are not. Does the mere fact that certain attitudes are taken toward an agent establish that he is an appropriate candidate for this treatment?

Imagine, for example, a society in which severely retarded or mentally disturbed individuals are resented, blamed and harshly punished for their failure to adhere to the norms of the community. (Perhaps the society attributes their failure to poor character or an evil nature.) Even though all the members of this community strongly feel moral outrage, resentment, and indignation, should the mere fact of their commitment to these feelings preclude further worry that the adoption of such attitudes toward the agents is not justified? (Fischer and Ravizza 1993: 18)

F&R synes å mene at det er intuitivt sikkert at de omtalte tilbakestående individene ikke er moralsk ansvarlige. De skriver ikke hvorfor dette er intuitivt sikkert, men det kan antas å være knyttet til en manglende evne hos de mentalt tilbakestående. De kan ikke tilpasse seg Andeby-samfunnets krav. De er ikke i stand til å følge fellesskapets moralske regler, og derfor har de heller ikke moralsk ansvar for sine handlinger.

Selv om de ikke er i stand til oppføre seg annerledes, blir de forfulgt og straffet for sine avvik. Etter våre moralnormer er dette en urimelig behandling av mentalt tilbakestående individer. Men med Strawsons teori vil det ikke være tilgjengelig noen ekstern instans utenfor Andeby som har noen autoritet til å kritisere Andeby-befolkningens reaktive holdninger. Det generelle rammeverket av holdninger og følelser som også omfatter de reaktive holdningene, tillater ikke eksternt rasjonelt forsvar verken av Andeby-befolkningens holdninger eller deres kritikers holdninger.

Kritikk av eksisterende praksis

Fischer og Ravizza hevder at Strawson relativiserer tilskrivning av ansvar til de enkelte samfunn eller grupper. På denne bakgrunn reiser de spørsmålet om muligheten for å kritisere og revidere eksisterende praksis: “The close connection that Strawson makes between being responsible and actually being the recipient of reactive attitudes raises questions about the ability of his theory to criticize and revise existing practices.” (1993: 18).

Forholdet mellom reaktive holdninger og moralsk ansvar

Fischer og Ravizza hevder: “Strawson’s theory holds that being morally responsible is nothing other than being a recipient of these [reactive] attitudes and a participant in the associated practices.” (op.cit.: 17) Etter deres oppfatning er den ubetingede koplingen av moralsk ansvar til reaktive holdninger og tilhørende praksis en avgjørende svakhet med Strawsons teori. De svarer på denne svakheten med å videreutvikle Strawsons teori til noe som de kaller for en Strawsoniansk teori:

This type of theory holds that morally responsible agent are not just those who, as a matter of practice, are recipients of the reactive attitudes; rather, agents are morally responsible if and only if they are *appropriate* recipients – that is, they are rationally accessible to the reactive attitudes. (op.cit.: 18)

Fischer og Ravizza hevder at å være moralsk ansvarlig er å være passende kandidat for reaktive holdninger.¹⁶ Å være mottaker av reaktive holdninger er ikke det samme som å være *passende* mottaker av – eller passende kandidat for – reaktive holdninger. En aktør kan bli møtt med reaktive holdninger uten å være passende kandidat for slike holdninger; og en aktør kan være passende kandidat for reaktive holdninger, men ikke bli møtt med reaktive holdninger.

Det kan nevnes eksempler på avvik mellom å være mottaker av reaktive holdninger og passende kandidat for reaktive holdninger. Fischer og Ravizza omtaler en person som ødelegger en vase, men som ikke har moralsk ansvar for dette, fordi hun fikk et (epileptisk) anfall som hun ikke hadde kontroll over. La oss anta at hun ble møtt med reaktive holdninger fra en tilskuer ikke forsto at hun var syk. Hun ble altså møtt med reaktive holdninger, selv om hun ikke var passende kandidat for dette under sitt anfall.

Fischer og Ravizza skriver om denne tilnærmingen som identifiserer moralsk ansvarlig med å være passende kandidat for reaktive holdninger, at den “views the reactive attitudes as having some grounding in theoretical considerations, and it looks for a rational justifi-

¹⁶ Dette er et syn som Fischer og Ravizza hevder i ”Perspectives of Moral Responsibility” (1993), og som de tar utgangspunkt i i ”Responsibility and Control”. I ”Responsibility and Control” ser de moralsk ansvar i et subjektivistisk perspektiv, og deres definisjon blir derfor justert: Moralsk ansvar er *à se seg selv* som passende kandidat for reaktive holdninger. Det er mulig at denne justeringen til et subjektivistisk perspektiv også innebærer at de ikke lenger insisterer på eksistensen av eksternt rasjonelt forsvar for reaktive holdninger.

cation for taking the attitudes toward some agents and not others” (op.cit.: 19). De hevder at denne tilnærmingen krever et teoretisk grunnlag for å avgjøre hvem som er moralsk ansvarlig og hvem som ikke er det. De reaktive holdningene til andre personer må bygge på teoretiske betraktninger, og det trengs et rasjonelt forsvar for å kunne ha reaktive holdninger til noen aktører og ikke til andre.

6.3 *Vurdering av Fischer og Ravizzas kritikk*

I min vurdering av Fischer og Ravizzas kritikk snur jeg på deres argumentasjon og trekker fram deres siste utfordring først, og deretter de andre spørsmålene som de reiser:

- Finnes det rasjonelt forsvar for reaktive holdninger?
- Aksepterer Strawson teoretiske betraktninger om moralsk ansvar?
- Kan det gis eksternt rasjonelt forsvar for det som er intuitivt riktig?
- Finnes det en eller flere definisjoner av veiledningskontroll og moralsk ansvar?
- Har en mottaker av reaktive holdninger moralsk ansvar – uansett?
- Kan Strawson beskrive og forklare kritikk og endring av eksisterende praksis?

Teoretiske betraktninger og rasjonelt forsvar for reaktive holdninger

Fischer og Ravizza hevder at deres tilnærming som identifiserer moralsk ansvar med å være passende kandidat for reaktive holdninger, må bygge på teoretiske betraktninger og rasjonelt forsvar av de reaktive holdningene. Bryter dette med Strawsons syn?

Strawson skriver at det finnes et uendelig rom for justering, kritikk og forsvar innenfor den generelle strukturen av holdninger og følelser. Han hevder at vi kan justere de reaktive holdningene, men at dette skjer innenfor – og ikke utenfor - rammeverket av våre holdninger og følelser. En må kunne anta at det ikke bare er ujusterte og umiddelbare (*non-detached*) reaktive holdninger som konstituerer moralsk ansvar, men at også justerte og middelbare reaktive holdninger kan konstituere moralsk ansvar. En må også kunne anta at slik justering, kritikk og forsvar inkluderer teoretiske betraktninger, slik at en persons justerte reaktive holdninger kan bygge på teoretiske betraktninger. Dette gjelder også for forsvar av holdninger. Forsvar av holdningene er interne i forhold til strukturen, og knytter seg til justeringer internt til den. Strawson avviser derfor ikke at vi kan gi reaktive holdninger et rasjonelt forsvar. Det han går i mot, er at vi kan gi et *eksternt* rasjonelt forsvar for de reaktive holdningene.

Men hvilken holdning har Strawson til filosofers teoretiske betraktninger om moralsk ansvar? Aksepterer han slike betraktninger, eller avviser han dem? Jeg legger til grunn at Strawsons egen artikkel om moralsk ansvar inneholder teoretiske betraktninger, og at han derfor ikke er imot *alle* teoretiske betraktninger om moralsk ansvar, men at han er kritisk til visse typer av teoretiske betraktninger. En kan vise til at han kritiserer bruken av metafysiske proposisjoner, f.eks. utsagn om at skapninger har kontra-kausalt frihet (Fischer og Ravizzas formulering). Han kritiserer også en overintellektualisering av fakta i form av en enøyd utilitarianisme. Han aksepterer heller ikke eksternt rasjonelt forsvar for reaktive holdninger. Hva slags teoretiske betraktninger er det da han aksepterer? Etter mitt syn representerer Galen Strawsons subjektivistiske teori om fri vilje og moralsk ansvar et mulig svar på Peter Strawsons utfordring. Selv om Galen Strawson anvender en annen terminologi enn sin far, kan en spørre om også Galen Strawsons tilnærming innebærer en avvisning av eksternt rasjonelt forsvar for holdninger som konstituerer moralsk ansvar. Dette spørsmålet tas opp i kapittel 10.

Intuisjon og eksternt rasjonelt forsvar

Fischer og Ravizza spør hvordan man skal forholde seg til et samfunn hvor noen mennesker som intuitivt sett ikke er ansvarlig, likevel holdes ansvarlig. Er det slik at Strawsons teori innebærer at tvilsomme holdninger og brutal praksis må aksepteres, fordi moralsk ansvar identifiseres med de til enhver tid rådende reaktive holdninger? F&R illustrerer sitt spørsmål med omtale av et lokalsamfunn (som jeg kaller Andeby) hvor mentalt tilbakestående individer blir fordømt og straffet for sine avvik. Fischer og Ravizza forutsetter som intuitivt sikkert at de mentalt tilbakestående individene i Andeby ikke er moralsk ansvarlig for sin manglende tilpasningsevne, og at de ikke skal fordømmes og straffes for dette.

På bakgrunn av Strawsons utfordring kan en så spørre om det kan gis eksternt rasjonell kritikk av – eller forsvar for - Andeby-befolkningens holdning til de mentalt tilbakestående individene. Eller for å snu på spørsmålet: Kan det gis eksternt rasjonelt forsvar for F&Rs holdning - at de mentalt tilbakestående individene i Andeby er uten moralsk ansvar?

Vi kan anta at to personer – f.eks. Fischer og Ravizza – er uenig i holdningene til innbyggerne i Andeby. F&R hevder at Andeby-befolkningens behandling av de mentalt tilbakestående ikke er god, og de går i dialog med innbyggerne for å overbevise dem at de

tar feil. Vi forutsetter at det ikke finnes noen ekstern instans som kan forsvare F&R's syn, og heller ikke noen slik instans som kan forsvare Andeby-innbyggernes syn. F&R er i mindretall i forhold til alle innbyggerne i Andeby, og er ikke medlemmer av dette samfunnet. De har ikke spesielle maktmidler og bruker lite tid på å diskutere med Andebys innbyggere. Jeg antar at Fischer og Ravizzas argumenter ikke slår igjennom under disse forutsetningene, og at Andebys innbyggere opprettholder sine opprinnelige holdninger.

Nå kan en si at F&R har intuitivt eller selvinnlysende rett, og at Andeby-samfunnet mangler kunnskap og innsikt og tar derfor feil. Med et slikt ståsted kan det gis forsvar for Fischer og Ravizzas reaktive holdninger. Men for å imøtegå Strawson er det ikke tilstrekkelig å framheve de mentalt tilbakeståendes manglende moralske ansvar som *intuitivt* sikkert, en må også kunne gi et eksternt rasjonelt forsvar for denne holdningen. Er det gitt et eksternt rasjonelt forsvar for Fischer og Ravizzas syn ved at mange (f.eks. i det opplyste miljøet i den vestlige kulturkrets) finner at deres syn er intuitivt eller selvinnlysende rett? Jeg kan ikke se det.

Intuisjon er ikke det samme som eksternt rasjonelt forsvar. Intuisjon er heller å forstå som et premiss som man ikke gir noe rasjonelt forsvar for, og som det kan oppfattes som unødvendig å gi et eksternt rasjonelt forsvar for. Jeg deler Fischer og Ravizzas reaktive holdninger i denne saken. Jeg støtter heller ikke Andeby-befolkningens behandling av mentalt tilbakestående individer, slik det er blitt beskrevet her. Men har jeg gitt et eksternt rasjonelt forsvar for disse holdningene ved å erklære meg enig med dem? Nei, det har jeg ikke.

En eller flere definisjoner

Hvis det bare finnes en mulig oppfatning av veiledningskontroll og moralsk ansvar, kan andre oppfatninger og synspunkter avvises. Er det da slik at det bare finnes en mulig oppfatning? Fischer og Ravizza legger vekt på veiledningskontroll som en betingelse for moralsk ansvar. Og en person har veiledningskontroll dersom han har moderat evne til respons på grunner og eierskap over den operative mekanismen, dvs. at han er i stand til å ta ansvar. Kan en hevde at disse oppfatningene om veiledningskontroll og moralsk ansvar er de eneste gjeldende? Hvis en også kan påvise at de mentalt tilbakestående ikke tilfredsstiller kravene til veiledningskontroll og moralsk ansvar, kan en konkludere at Andeby-befolkningen tar feil når de stiller de mentalt tilbakestående til ansvar for sin atferd.

Det blir vist i kapittel 9 at enkelte elementer i definisjonen av veiledningskontroll kan være gjenstand for diskusjon. Det dreier seg blant annet om spørsmålet om moralsk ansvar inkluderer betingelse om å *ta* ansvar og ha eierskap til den operative mekanismen som styrer aktørens handlinger. Jeg antar derfor at det kan finnes flere oppfatninger, eller flere definisjoner, av moralsk ansvar. Og jeg kan ikke se at det er gitt argument mot at alternative definisjoner kan innebære et annet syn på de tilbakestående ansvar enn Fischer og Ravizza har.

Sosiale og individuelle reaktive holdninger

Jeg har lagt til grunn at det ikke er påvist at det gis et eksternt rasjonelt forsvar for reaktive holdninger, og at det derfor kan synes som en moralsk ansvarlig person er lik en mottaker av reaktive holdninger. Det synes vanskelig å finne grunnlag for tilbakevisning av tilskrivning av ansvar. For en videre vurdering av spørsmålet vil jeg trekke fram enda et eksempel: Fischer og Ravizza har nevnt eksemplet med kvinnen (la oss kalle henne Anna) som knuste en vase under epileptisk anfall. Hun ble møtt med reaktive holdninger fra en tilskuer ikke forsto at hun var syk. La oss anta at tilskueren etterpå skjønnte at Anna var syk, trakk tilbake sine reaktive holdninger og skjønnte at Anna ikke var moralsk ansvarlig for å knuse vasen.

Hvis forutsetningen om manglende eksternt forsvar for reaktive holdninger innebærer at moralsk ansvar er lik de reaktive holdningene som gjelder til enhver tid, kan denne fortellingen bli fortolket slik: Anna har først ansvar for å knuse vasen, for tilskueren har reaktive holdninger og *tillegger* Anna ansvar. Hun har deretter ikke lenger ansvar for det samme forholdet, fordi tilskueren justerer sine reaktive holdninger og tillegger henne ikke lenger ansvar for å knuse vasen. Fortellingen blir ytterligere komplisert hvis det er flere tilskuere til stede: Tilskuer A skjønner at Anna har et epileptisk anfall og tillegger henne ikke ansvar for å knuse vasen, mens tilskuer B tillegger henne ansvar for å knuse vasen. Er da Anna både moralsk ansvarlig og ikke moralsk ansvarlig for å knuse vasen på samme tid?

Ved å skille mellom sosialt baserte og individuelt baserte reaktive holdninger blir det gitt et forsøk på fortolkning av denne fortellingen i avhandlingens Del II.

Om kritikk og endring av eksisterende praksis

Gir Strawsons teori anledning til kritikk og endring av eksisterende praksis i et lokalsamfunn? For å vurdere det kan vi se på noen ulike scenarier i det omtalte lokalsamfunnet Andeby.

La oss se på et scenario: Samfunnet utenfor Andeby reagerer på forholdene der. Påtalemyndigheten tar ut siktelse mot noen av dem som har utøvd trakassering og vold mot personer i minoritetsgruppen. Det blir også igangsatt en mediekampanje mot Andeby og de holdningene som er fremherskende der. Aktivistene i Andeby blir hengt ut og beskrevet som ondskapsfulle og tvilsomme individer i denne kampanjen. Disse tiltakene virker langt kraftigere enn Fischer og Ravizzas samtaler med innbyggerne i Andeby. På bakgrunn av stor oppmerksomhet omkring saken, sosialt press, og i tillegg maktutøvelse fra påtalemyndighetens side, endres Andeby-innbyggernes praksis. Med andre ord: Dersom mottiltakene mot majoritetsbefolkningen i Andeby er tilstrekkelig kraftige, er det grunn til å anta at de vil endre sin praksis - og en slik endring kan skje, selv om en ikke kan påvise noe eksternt rasjonelt forsvar for det omliggende samfunnets reaktive holdninger.

Så til et annet scenario: To personer, Fischer og Ravizza, reiser til Andeby og oppholder seg der i lange perioder. De argumenterer for et annet syn enn Andebyinnbyggernes syn på de mentalt tilbakestående. De bruker lang tid på å samtale med byens befolkning. De har ingen spesielle maktmidler til rådighet, og ingen andre instanser bruker makt for å få Andeby til å endre sin praksis. Etter et år har likevel et flertall av byens befolkning forandret oppfatning, fordi Fischer og Ravizza har overbevist dem om at det er galt å kreve atferd av mentalt tilbakestående som de er i stand til å tilfredsstille.

Kan dette andre scenariet forklares uten eksistensen av eksternt rasjonelt forsvar for reaktive holdninger? Dersom det finnes eksternt rasjonelt forsvar for Fischer og Ravizzas reaktive holdninger, virker det troverdig at de før eller senere vil kunne overbevise Andeby-befolkningen om sitt syn. Men vil F&R kunne overbevise Andebys befolkning hvis det *ikke* finnes noe eksternt rasjonelt forsvar for deres reaktive holdninger?

Strawson erkjenner at fornuften spiller en viktig rolle innenfor strukturen av menneskelige holdninger. McKenna og Russell uttrykker det slik: «Strawson is clear that reason has an

important role to play ‘inside the general structure or web of human attitudes and feelings’ that he is concerned with.» (McKenna and Russell: 11) En betydelig del av våre holdninger bygger på fornuftsargumenter, og vi kan endre holdninger på grunnlag av fornuftsargumenter. Fornuften har altså stor betydning for våre vurderinger og holdninger, selv om det skulle vise seg at verken fornuften, vurderingene eller holdningene lar seg forsvare eksternt. Det synes derfor ikke å bryte med Strawsons teori at Andeby-befolkningen endrer sine reaktive holdninger på grunn av fornuftsargumenter.

Jeg har søkt å vise at disse to scenariene kan passes inn i Strawsons teori. Jeg vil derfor ikke hevde på bakgrunn av denne historien at Strawsons teori gjør det umulig å forklare kritikk og endring av eksisterende praksis i et lokalsamfunn.

Oppsummering

Fischer og Ravizza kritiserer Strawson for:

- Hans teori medfører forsvar for intuitivt gale standpunkter, som ansvarliggjøring av individer som er uten ansvar
- Teorien forklarer ikke kritikk og endring av eksisterende praksis
- Teorien setter likhet mellom moralsk ansvar og å være mottaker av reaktive holdninger

De legger også vekt på at reaktive holdninger må kunne bygge på teoretiske betraktninger.

I min vurdering har jeg funnet lite grunnlag for kritikken av Strawson, bortsett fra at det i visse tilfeller virker urimelig å identifisere en moralsk ansvarlig person med mottaker av reaktive holdninger. Det blir gitt et forslag til løsning av dette dilemmaet i avhandlingens Del II.

6.4 Sosial praksis assosiert til reaktive holdninger

Strawson omtaler også sosial praksis eller holdninger (*practices or attitudes*) assosiert med reaktive holdninger¹⁷, som bl.a. fordømmelse og anerkjennelse. Et inntak til Strawsons

¹⁷ På norsk er det naturlig å bruke entallsformen sosial praksis, selv om Strawson bruker flertallsformen *social practices*. Kanskje mener Strawson at det også finnes annen praksis enn sosial praksis som er assosiert til reaktive holdninger. Uansett gjelder omtalen her bare *sosial praksis* assosiert til reaktive holdninger.

drøfting av praksis kan være hans beskrivelse av optimisten og pessimisten, eller – som andre kanskje ville ha sagt det - kompatibilisten og inkompatibilisten.¹⁸ I det følgende skal vi se nærmere på den optimistiske posisjonen, som Strawson karakteriserer slik:

- (1) the facts as we know them do not show determinism to be false
- (2) the facts as we know them supply an adequate basis for the concepts and practices which the pessimist feels to be imperilled by the possibility of determinism's truth (P. Strawson: 46)

Strawson er kritisk til denne posisjonen:

The optimist's style of over-intellectualizing the facts is that of a characteristically incomplete emiricism, a one-eyed utilitarianism. He seeks to find an adequate basis for certain social practices in calculated consequences, and loses sight ... of the human attitudes of which these practices are the expression. (op.cit.: 64)

Her beskriver Strawson optimistens posisjon svært snevert. Det kan knapt være tvil om at det finnes andre kompatibilistiske eller "optimistiske" posisjoner enn den enøyde utilitarismen. Det erkjenner da også Strawson i avslutningen av sitt essay:

If we sufficiently, that is radically, modify the view of the optimist, his view is the right one. It is far from wrong to emphasize the efficacy of all those practices which express or manifest our moral attitudes..... What *is* wrong is to forget that these practices, and their reception, really *are* expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes. (op.cit.: 93)

Her utvides perspektivet ved at moralske holdninger trekkes inn i tillegg til det kalkulerende utilitaristiske perspektivet.

Fischer og Ravizza har også drøftet optimistens posisjon. De omtaler en optimist-strategi som de kaller *sosial-reguleringsmodell*. De skriver at denne strategien

.... justifies our responsibility-reactions by appeal to the useful consequences that follow from these reactions. Such a consequentialist justification understands our responsibility-reactions primarily as an impersonal, objective means of regulating behavior. According to this view, the point of these reactions is to educate moral agents and to provide inducements for them to behave in certain ways. Threats of punishment and blame exert pressure on agents to conform their behavior to the laws and norms of society. Similarly, promises of praise and reward offer powerful incentives for agents to strive for the ideals that have currency in a particular moral community. (Fischer and Ravizza 1993: 11-12)

¹⁸ Jeg forutsetter at Strawsons ordvalg er bevisst, og bl.a. kan ha til formål å omgå fastlåste posisjoner i diskusjonen mellom kompatibilisme og inkompatibilisme. Men det er uansett knapt noen tvil om at optimisten representerer en kompatibilistisk posisjon, og pessimisten en inkompatibilistisk posisjon.

Også Fischer og Ravizza finner at denne modellen har et noe snevert nytteperspektiv. De viser til Susan Wolf og skriver:

.... we do not hold someone responsible first and foremost because we calculate that such an ascription of responsibility will have an instrumental value in producing some desired results; rather we "find ourselves reacting to the actions and characters of others, approving of some, disapproving of others. Unless there is some reason to restrain ourselves, we simply express what we feel. (Fischer and Ravizza 1993: 14)

Fischer og Ravizza framhever altså våre reaksjoner i første rekke som uttrykk for hva vi føler (Strawson omtaler disse uttrykkene som uttrykk for moralske holdninger), og dernest som kalkulerende atferd. En kan imidlertid hevde at det er noe ufullstendig eller polarisert i beskrivelsen, både til Fischer og Ravizza og til Strawson. Består den sosiale praksis assosiert med de reaktive holdningene¹⁹ bare av de to "motpolene" kalkulasjoner og uttrykk for umiddelbare følelser og holdninger? Eller består den sosiale praksis av flere viktige elementer? Et bidrag til svar på dette spørsmålet gis av Lawrence Stern. Han skriver om involverte²⁰ holdninger:

.... Not to be calculating is to feel, and express one's feelings. One mode of nonmanipulative action is what I call *dialogue*. Dialogue consists in relying on the other person's rational and moral capacities in one's dealings with him. Dialogue may aim at a result. One may initiate it with a view toward changing the other person's beliefs or conduct in some precise respect. However, it does not focus on the result with the intensity of manipulation. For it allows the other person to change us. (Stern 1974: 74-75)

Stern hevder altså at involverte holdninger innebærer følelser, uttrykk for følelser og dialog. Og dialog innebærer å stole på andres rasjonelle og moralske evner. Det innebærer også en villighet til å la den andre personen forandre oss. Dialog kan være målrettet, men ikke manipulativ. Jeg forstår Stern slik at dialog kan være eksempel på sosial praksis assosiert med reaktive holdninger, slik ros og klander er det. Hvis dette er rett forstått, innebærer sosial praksis langt mer den formen for kalkulasjon som Strawson tillegger sin optimistiske posisjon. Det innebærer også mer enn kalkulasjon og umiddelbare reaksjoner basert på moralske holdninger, som er Strawsons modifiserte posisjon, og som F&R synes å slutte seg til.

¹⁹ I følge Fischer og Ravizza er ros og klander, belønning og straff eksempler på sosial praksis. Strawson framhever særlig moralsk fordømmelse og anerkjennelse.

²⁰ "Involverthet" er et trekk ved de reaktive holdningene. Det framgår at Stern inkluderer både reaktive holdninger og sosial praksis i sitt begrep om involverte holdninger.

Jeg vil her trekke fram enda ett element som har en sentral plass i Strawsons filosofi, og som kanskje hører hjemme her, nemlig: internt rasjonelt forsvar, kritikk og endring av reaktive holdninger. Strawson har som kjent avvist eksistensen av eksternt rasjonelt forsvar av reaktive holdninger, men han erkjenner at det finnes internt rasjonelt forsvar (og kritikk) av slike holdninger. Hvilken karakter har slike internt rasjonelle prosesser? De må uten tvil ha subjektive komponenter ved at aktøren gjennomgår og vurderer – og eventuelt forsvarer – egne reaktive holdninger. Men det ville være underlig om ikke et slikt internt rasjonelt forsvar også har sosiale komponenter: Aktøren henter nødvendigvis impulser fra sosiale omgivelser når han vurderer (forsvarer eller kritiserer) sine reaktive holdninger. Internt rasjonelt forsvar av reaktive holdninger synes altså å ha komponenter av sosial praksis assosiert med reaktive holdninger. Det er likevel grunn til å understreke en tydelig forskjell mellom praksisformene: Rasjonelt forsvar og rasjonell kritikk har en rasjonell, kognitiv karakter; mens fordømmelse og anerkjennelse og ros og klander har en normativ og moralsk karakter.

Drøftingen i dette underkapitlet kan oppsummeres slik: For Strawson har det vært veldig viktig å understreke at moralsk ansvar har en umiddelbar og udiskutabel basis i de reaktive holdningene. Dette programmet har ført til en noe anti-teoretisk tilnærming (men neppe så anti-teoretisk som noen av hans kritikere hevder). På noen områder kan man med utgangspunkt i Strawsons grunnsyn gi betydelige teoretiske bidrag uten å komme i konflikt med hans syn om at det ikke kan gis noe eksternt rasjonelt forsvar av reaktive holdninger. Et tema som med fordel kan videreutvikles, er sosial praksis assosiert til reaktive holdninger. Som nevnt gir Lawrence Stern et godt tilskudd til en litt for polarisert framstilling av Strawson og Fischer og Ravizza. I tillegg spør jeg om elementer i det interne rasjonelle forsvaret av reaktive holdninger kan sees som deler av denne sosiale praksisen. Jeg gir mitt bidrag til denne drøftingen i avhandlingens Del II.

7. Determinisme og legitim klander

Fischer og Ravizza omtaler ros og klander som nært knyttet til reaktive holdninger. De bruker et felles uttrykk - ansvarsreaksjoner (*responsibility-reactions*) - for disse holdningene, følelsene og handlingene som omfatter både reaktive holdninger og ros og klander (Fischer and Ravizza 1993: 6). Også Strawson legger vekt på den nære relasjonen mellom reaktive holdninger og handlinger eller holdninger som straff, klander, fordømmelse og

anerkjennelse. Strawson viser størst interesse for de reaktive holdningene. Andre viser større interesse for det som Strawson kaller sosial praksis (*social practices*) basert på de reaktive holdningene, slik som ros og klander. Blant de sistnevnte er Derk Pereboom.

Pereboom hevder at vi ikke kan bli klandret legitimt for gale handlinger, dersom all atferd er styrt av en deterministisk kausal prosess. I motsetning til de fleste andre filosofene som blir omtalt i avhandlingen, gir Pereboom en eksplisitt drøfting av Fischer og Ravizzas standpunkt. I dette kapitlet blir det også gjort rede for Fischers svar til Pereboom. Fischer reiser tvil om det er nødvendig forbindelse mellom moralsk ansvar og klander. Til slutt i kapitlet gir jeg en vurdering ved å drøfte begrepet *legitimitet*.

7.1 Perebooms kritikk av Fischer og Ravizza

Pereboom stiller denne betingelsen for moralsk ansvar: "moral responsibility requires actions to have indeterministic actual causal histories, or more fundamentally, to have causal histories that make agents ultimate sources of their actions" (Pereboom 2001: xv). Mennesker kan altså ikke ha moralsk ansvar hvis alt er styrt av deterministiske naturlover, i følge Pereboom. Han utdyper dette standpunktet:

The intuition is that if all of our behavior was "in the cards" before we were born, in the sense that things happened before we came to exist that by way of a deterministic causal process, inevitably result in our behavior, then we cannot legitimately be blamed for our wrongdoing. (op.cit.: 89)

Sentralt i Perebooms begrunnelse er hans historie om professor Plums drap på fru White, i fire versjoner (*cases*). I Case 1 er Plum utviklet av spesialister innen nevrovitenskapene, og disse spesialistene kan manipulere ham direkte gjennom radio-teknologi. I Case 2 er han også utviklet av nevroforskere, men de kan ikke kontrollere ham direkte. I Case 3 er han et alminnelig menneske, bortsett fra at han er avgjørende preget (*determined*) av en rigorøs oppdragelse. I Case 4 er Plum et alminnelig menneske med normal oppdragelse.

I alle fire versjonene er Plum egoistisk, men han handler likevel *ikke alltid* egoistisk. Han har evnen til å respondere på fornuftsgrunner. Han blir ikke drevet av en tvangsmessig drift (*irresistible desire*). Han er altså beskrevet slik at han skal kunne tilfredsstille Fischer og Ravizzas krav til en person med veiledningskontroll og moralsk ansvar. I alle fire versjonene dreper likevel professor Plum fru White for å oppnå noen personlige fordeler,

noe som begrunnes i hans sterkt egoistiske karakter. I Case 4 beskriver Pereboom dette slik:

Plum's killing of Ms. White comes about as a result of his undertaking the moderately reason-responsive process of deliberation ..., and he does not act because of an irresistible desire. He has the general ability to grasp, apply and regulate his behavior by moral reasons, but in these circumstances the egoistic reasons are very powerful, and together with the background circumstances they deterministically result in the act of murder. (op.cit.: 115)

I følge Fischer og Ravizza er Plum ikke moralsk ansvarlig hvis han er utsatt for direkte manipulasjon av hjernen. Han er derimot moralsk ansvarlig hvis:

- det ikke forekommer noen ansvarsundergravende mekanisme
- han responderer på fornuftsgrunner (*has the ability to grasp, apply and regulate his behavior by moral reasons*)
- han handler ikke på grunnlag av tvangsmessig drift (*irresistible desire*)

Disse kravene tilfredsstilles i flere av tilfellene (versjonene). Det forutsettes derfor at F&R finner Plum moralsk ansvarlig i de tilfellene.

Pereboom er uenig i dette. Han skriver at Plum ikke er moralsk ansvarlig i Case 1, fordi han er determinert av nevroforskeres handlinger utenfor egen kontroll. Han hevder videre at det ikke er noen avgjørende forskjell på Plums atferd i de fire versjonene. Når man kan slå fast at Plum ikke er moralsk ansvarlig i Case 1, er han derfor heller ikke moralsk ansvarlig i de andre tilfellene (versjonene).

7.2 Fischers og Perebooms diskusjon

I den delen av debatten mellom Fischer og Pereboom som det blir vist til i avhandlingen, er uenigheten knyttet til ros og klander særlig viktig. Fischer og Ravizza hevder at en person med moralsk ansvar er kandidat for ros og klander. Det er dette som skiller personer fra andre skapninger, og som – med visse unntak – skiller menneskene fra dyrene. Vi reagerer ikke på dyr med ros og klander, og vi tillegger dem heller ikke moralsk ansvar. Når det gjelder (de fleste) mennesker derimot, reagerer vi med ros og klander *fordi* vi tillegger dem moralsk ansvar for deres handlinger. Fischer og Ravizza hevder at hvis en person ikke er utsatt for en ansvarsundergravende mekanisme og *tar* ansvar – og derfor har veiledningskontroll – så har han moralsk ansvar og er kandidat for ros og klander.

I boka *Four Views on Free Will* diskuterer Fischer og Pereboom eksemplene med professor Plum. Fischer hevder at man må skille mellom moralsk ansvar og ros og klander. Han bruker et eksempel om en kvinne som er blitt utsatt for vold i ekteskapet gjennom mange år. Til slutt dreper hun sin mann i fortvilelse. Fischer antar at hun er moralsk ansvarlig for sin handling – altså at hun er en person med veiledningskontroll - men at hun likevel ikke er moralsk klanderverdig tatt i betraktning de tragiske omstendighetene som handlingen ble foretatt under (Fischer et.al. 2007: 186).

Fischer mener derfor at professor Plum kan være ansvarlig for sin handling i Case 1 og 2, selv om han ikke kan klandres fordi hans handling var manipulert av nevroforskere. Fischer skiller mellom å være aktuell kandidat (*apt candidate*) for reaktive holdninger (og ros og klander) og rettferdiggjørelse av applikasjonen (*justification for the application*) av slike holdninger. Professor Plum kan altså være aktuell kandidat for ros og klander i Case 1 og 2. Men det kreves ytterligere betingelser (nemlig at han ikke er hjernemanipulert eller ansvarsundergravd på annen måte) for at de reaktive holdningene og ros og klander skal kunne rettferdiggjøres (op.cit.: 185). Det er nærliggende å anta (men ikke helt klart) at Fischer med dette mener at Plum har moralsk ansvar i alle fire versjonene, men at han bare kan klandres i Case 3 og 4.

Pereboom følger Fischer et stykke på vei, men ikke hele veien. Han er enig i at moralsk ansvar ikke nødvendigvis utløser ros eller klander: "once it settled that an agent is morally responsible, it then needs to be determined whether what he did was wrong, and if he did not, whether he could have or should have understood that it was" (op. cit.: 197). Først når det er klarlagt om aktøren kunne ha forstått - eller *burde* ha forstått - at han handlet galt, kan han klandres for sin handling. Men Pereboom mener at dette ikke svekker hans poeng. Han forutsetter at professor Plum forsto at det var galt å drepe White. Plum kunne derfor ikke ha moralsk ansvar for sin udåd uten samtidig å bli klandret for det, i følge Pereboom.

7.3 Om legitim klander

Jeg skal ikke her ta stilling til diskusjonen mellom Fischer og Pereboom, da dette innspillet er brakt inn for et avgrenset formål, nemlig å se nærmere på uttrykket *legitim klander*. Pereboom forutsetter at vi må kunne bli legitimt klandret for våre feil for å ha moralsk ansvar for dem. Han hevder videre at det ikke kan gis legitim klander dersom all vår atferd

”lå i kortene” fra før vi ble født. For å vurdere dette utsagnet er det viktig å klargjøre hva *legitim* klander er.

Pereboom legger dette resonnementet til grunn: Legitim klander rettes mot den opprinnelige årsaken til en feil. Med deterministisk forståelse er den opprinnelige årsaken av ukjent karakter, uendelig langt tilbake i tid og lå i kortene før aktørens fødsel. Klander kan derfor ikke rettes legitimt mot annet enn denne ukjente årsaken uendelig langt tilbake i tid. Verken aktøren eller hans atferd er opprinnelig årsak til feilen. Med denne forståelsen av *legitimitet* – og forutsatt at determinismen er sann - kan det ikke rettes legitim klander mot den personen som gjør en gal handling.

Denne forståelsen av legitimitet kan knyttes til drøftingen i forrige kapittel om rasjonelt forsvar for reaktive holdninger. Det kan antas at det er en relasjon mellom rasjonelt forsvar for reaktive holdninger og legitimering av klander: Hvis det kan gis et rasjonelt forsvar for en reaktiv holdning, vil det også ofte (ikke alltid, det kan være riktig å la være å klandre, selv om den reaktive holdningen var fullt forsvarlig) gis rasjonelt forsvar for den klandereren som har sitt utspring i den reaktive holdningen. Hvis det *ikke* gis et rasjonelt forsvar for en reaktiv holdning, antas det at det *ikke* kan gis noe rasjonelt forsvar for klandereren som har utspring i den reaktive holdningen. Det antas videre at et rasjonelt forsvar for klander er en form for legitimering av klander. Det antas også at det kan skilles mellom ekstern og intern legitimering av klander, på tilsvarende måte som det skilles mellom eksternt og internt rasjonelt forsvar for reaktive holdninger.

Pereboom spør ikke om klandereren kan forsvares innenfor rammeverket av de holdninger og den sosiale praksis hvor handlingen finner sted. Han etterspør en legitimering av klandereren som er knyttet til bakenforliggende årsaker til objektet for klandereren. Han etterspør dermed en legitimering av klandereren som er ekstern i forhold de reaktive holdningene som klandereren har sitt utspring fra. Jeg antar altså at Pereboom etterspør et eksternt rasjonelt forsvar – eller en *ekstern* legitimering - for klander av en aktør som gjør en gal handling.

I kapittel 6 ble muligheten for eksternt rasjonelt forsvar av reaktive holdninger drøftet, og det ble ikke funnet klar påvisning av at slikt forsvar kan gis. Hvis vi gjør de ovenfor nevnte antakelsene, vil mangel på eksternt rasjonelt forsvar av reaktive holdninger også innebære en mangel på ekstern legitimering av klander. Så langt gir derfor drøftelsen i forrige kapit-

tel støtte til Perebooms syn – at man ikke legitimt kan klandre en person for hans gale handlinger.

Men legitimitet kan også forstås på annen måte. Legitim betyr opprinnelig lovlig, avledet av *lex*, som betyr lov. Det er vanlig å skille mellom to former for lovlighet: legitimitet og legalitet. Det legale er i samsvar med landets lover. Legitimitet, derimot, er ikke bare relatert til lovverket. Legitimitet kan ha referanse til mange ulike tankesystemer og mange ulike sosiale grupperinger. Legitimitet forstått på denne måten befinner seg i en sosial kontekst eller en sosial setting. En kan derfor spørre om det kan gis legitim klander, dersom denne klanderen er tankemessig og/eller verdimessig forankret i et tankesystem som har legitimitet innen den sosiale settingen som man befinner seg i. Det gis en vurdering av dette spørsmålet i avhandlingens Del II.

8. Moralsk ansvar og kontroll

Fischer og Ravizza hevder at moralsk ansvar er betinget av en spesiell form for kontroll – veiledningskontroll. I dette kapitlet beskrives først forholdet mellom veiledningskontroll som er kompatibel med determinisme, og regulativ kontroll som *ikke* er kompatibel med determinisme. Dernest blir det gjort rede for den ene betingelsen for veiledningskontroll, respons på grunner. Den andre betingelsen - eierskap til den operative mekanismen – blir i hovedsak omtalt i kapittel 9.

Det blir også presentert en kritikk av kontrollbegrepet av Joseph Raz (Raz 2010). Han tar avstand fra sammenkoplingen av ansvar og kontroll, og gir uttrykk for at man kan ha ansvar uten å ha kontroll. Til slutt gir jeg min egen vurdering av spørsmålet.

8.1 Regulativ kontroll og veiledningskontroll

For Fischer og Ravizza er *kontroll* den sentrale betingelsen for moralsk ansvar. F&R skiller imidlertid mellom *regulative control* og *guidance control*. *Regulative control* innebærer tilgang til alternative muligheter (Fischer 2006: 39). *Guidance control* innebærer ikke tilgang til alternative muligheter. Den direkte oversettelsen av de to kontrollbegrepene er *regulativ kontroll* og *veiledningskontroll*.

Regulativ kontroll er en overordnet form for kontroll. En kan si at den regulative kontrollen ikke kontrollerer handlingene direkte, men er en kontroll av det som kontrollerer handlingene.

Veiledningskontroll er knyttet til rasjonalitet og grunner (*reasons*) og til evnen til å ta styring over sitt liv og sine handlinger. Det bør også nevnes at F&Rs rasjonalitetsbegrep både omfatter moralske vurderinger (verdirasjonalitet) og kognitiv tenkning (målrasjonalitet), trolig med hovedvekt på det første. I norsk oversettelse er ikke uttrykket *veiledningskontroll* helt tilfredsstillende. I stedetfor kunne en bruke uttrykket *personlig kontroll* for å vise at det er en kontroll som har sitt utspring i den enkelte person – i motsetning til regulativ kontroll, som i følge Fischer og Ravizza er utenfor menneskelig kontroll.

En av F&Rs oppgaver er å vise at en person kan ha veiledningskontroll uten å ha regulativ kontroll, dvs. kontroll uten alternative muligheter. De trekker fram et eksempel om Sally som kjører bil. Hun ønsker å svinge til høyre. F&R antar at hun som et resultat av sin intensjon dreier rattet og styrer bilen mot høyre. Videre antar de at dersom hun hadde intensjon om å svinge til venstre, ville hun ha dreid rattet den andre veien, og bilen hadde svingt til venstre. F&R skriver at Sally kontrollerer bilen, og at hun har en viss kontroll over bilens bevegelser.

Insofar as Sally actually guides the car in a certain way, we shall say that she has "guidance control". Further, insofar as Sally also has the power to guide the car in a different way, we shall say that she has "regulative control". (op.cit: 31)

F&R hevder at Sally har veiledningskontroll når hun er fri til å sørge for at bilen svinger den veien hun vil. Regulativ kontroll innebærer frihet *både* til gjøre handling A (svinge til høyre) *og* gjøre noe annet (for eksempel svinge til venstre). Fischer og Ravizza viderefører sitt eksempel, og skriver at Sallys bil har en "kjørerinstruksjon" som innebærer at instruktøren ville ha hindret henne i å svinge mot venstre istedenfor mot høyre, dersom hun hadde prøvd på det. I dette tilfellet har hun fremdeles veiledningskontroll over bilen, siden hun kan svinge bilen til høyre slik hun selv ønsker. Men hun har ikke regulativ kontroll. Hun kan nemlig ikke gjøre noe annet enn å svinge til høyre. Hvis hun forsøker på det, vil instruktøren hindre henne.

I generell form er historien om Sally slik: En person p utfører en handling a som fører til resultatet r. Hvis han ikke hadde gjort handlingen a, hadde imidlertid en annen person (eller kraft) virket inn og utført handlingen b, som også hadde ført til resultatet r. Dette er en type historie som ble brukt i en viktig artikkel av Harry Frankfurt, "Alternate Possibilities and Moral Responsibility" (Frankfurt 1969), og de blir derfor ofte kalt Frankfurt-type historier eller Frankfurt-type eksempler.

8.2 Respons på grunner

Fischer og Ravizza hevder at en aktør viser veiledningskontroll over en handling, når handlingen styres av den mekanismen hos ham som responderer på grunner ("an agent exhibits guidance control of an action to the extent that the action issues from his own reasons-responsive mechanism" – F&R 1998: 170). Respons på grunner består av tre ledd: mottakelighet for grunner (*receptivity to reasons*), valget om å handle i samsvar med grunnene, og selve handlingen. De to siste leddene blir behandlet under ett og omtalt som reaksjonsevne (*reactivity*). Uttrykket "*respons på grunner*" er et uvant uttrykk på norsk. Noe forenklet kan en si at det dreier seg om evne til rasjonell respons eller simpelthen rasjonalitet. Når F&R skriver om respons på grunner, omtaler de mekanismer og ikke aktører. I dette kapitlet knytter jeg imidlertid repons på grunner ikke bare til mekanismer, men også til aktører – uten at det synes å bryte med Fischer og Ravizzas intensjoner.

Forfatterne skiller mellom tre grader av respons på grunner: *sterk, moderat og svak respons*. Sterk respons på grunner inntreffer under disse betingelsene:

.... if K (K er en mekanisme) were to operate and there were sufficient reason to otherwise, the agent would recognize the sufficient reason to do otherwise and thus choose to do otherwise and do otherwise. (Fischer and Ravizza 1998: 63)

Sterk respons på grunner omfatter altså følgende mekanisme: en gjenkjennelse (*recognition*) av grunner, et valg om å handle i samsvar med grunnene, og selve handlingen. F&R mener at en person ikke trenger å ha sterk respons på grunner for å ha moralsk ansvar og veiledningskontroll. De begrunner dette ved hjelp av et eksempel:

Jennifer går på basketballkamp. Det fører til at hun ikke rekker å slutføre et arbeid i tide. Hun erkjenner at det ville være fornuftig å fullføre sitt arbeid, men hun har svak vilje og går i stedet til basketballhallen. Hun erkjenner hva som er fornuftig, men hun handler ikke

fornuftig, dvs. i samsvar med grunner. Hun har derfor ikke sterk respons på grunner. I et alternativt scenario blir Jennifer fortalt at det koster 1000 dollars å se kampen. I så fall vil hun ikke gå og se kampen. I et tredje scenario blir hennes arbeid forsinket fordi hun engasjerer seg i veldedighetsarbeid. På dette grunnlaget mener F&R at Jennifer har veiledningskontroll, selv om den mekanismen som styrer henne ikke er drevet av sterk respons på grunner. Kravet om sterk respons på grunner er derfor for streng betingelse for veiledningskontroll.

Svak respons på grunner defineres slik: "there exist *some* possible scenario in which there is sufficient reason to do otherwise, the agent recognizes this reason, and the agent does otherwise." (op.cit.: 63) Mens sterk respons på grunner innebærer en mekanisme hvor en aktør *alltid* vil handle i samsvar med grunner, innebærer svak respons på grunner at mekanismen *minst en gang* vil være i samsvar med grunnene.

For å beskrive svak respons på grunner (svak rasjonalitet) gir Fischer og Ravizza et eksempel med en person som dreper mennesker på en ferge. Han ville begått denne udåden i alle mulige scenarier bortsett fra ett: dersom en person på det laveste dekket hadde røykt på en Gambier-pipe. Dersom dette scenariet inntreffer, tilfredsstiller mannen på fergen kravet til svak rasjonalitet. Men han har ikke veiledningskontroll, for hans rasjonalitet er *for* svak.

F&R søker derfor å beskrive en moderat form for respons på grunner som befinner seg mellom de to ytterverdiene (sterk og svak respons på grunner), og som forutsettes å danne grunnlag for veiledningskontroll. I beskrivelsen av moderat respons på grunner skiller de mellom mottakelighet eller reseptivitet (*receptivity*) og reaksjonsevne (*reactivity*). Det er asymmetri mellom disse to elementene ved at det stilles store krav til aktørens mottakelighet og mindre krav til reaksjonsevnen. F&R skriver: "In order for an agent to be morally responsible for an action .. the actual mechanism that issues in his action must be at least "regularly" receptive to reasons." (op.cit: 71) Aktøren må altså ha en mekanisme som vanligvis – eller som regel - er mottakelig for grunner. Mottakeligheten må også være forståelig, basert på realiteter. De viser igjen til aktøren som dreper mennesker på en ferge i alle tilfelle unntatt ett; dersom en person røyker på en Gambier-pipe. Denne personen opptrer rasjonelt - eller fornuftig - i bare ett av en rekke scenarier, og viser derfor ikke

regelmessig mottakelighet for grunner. Hans handlingsmønster er heller ikke forståelig. Hans motivasjon for å handle moralsk forsvarlig i det ene tilfellet (at en person på nedre dekk røyker på en Gambier-pipe) er ikke basert på realiteter som har betydning for handlingen, og framstår derfor som uforståelig. Han har derfor ikke en mottakelighet som kreves for moderat respons på grunner. (I denne historien beskrives en persons reaksjoner og ikke hans mottakelighet (reseptivitet), men jeg forstår det slik at beskrivelsen av reaksjonene gir innsikt i personens mottakelighet. Beskrivelsen av reaksjonene viser at det dreier seg om en person med svekket mottakelighet.)

En kan også vise til Fischer og Ravizzas historie om Jennifer som i noen tilfelle utfører det arbeidet hun er forpliktet til, og i noen tilfelle lar det være. Om henne skriver F&R at hennes handlinger er forståelige og moralsk ansvarlige. Siden hun også responderer på grunner med en viss regularitet, hevder de at hun har en mottakelighet (reseptivitet) som tilfredsstillende kravet til moderat respons på grunner.

Når det gjelder reaksjonsevnen har F&R den intuisjon at denne er udelelig. Det innebærer at en mekanisme som reagerer rasjonelt i *en* situasjon, er i stand til å reagere rasjonelt i *alle* situasjoner. De skriver det slik: “Our point is that, holding fixed the actual kind of mechanism, reactivity is all of a piece: if the mechanism can react to any reason to do otherwise, it can react to all such reasons.” (op.cit.: 74) Det er derfor tilstrekkelig med svak reaksjonsevne, slik at man reagerer fornuftig i minst ett mulig scenario, for å kunne ha veiledningskontroll.

En mekanisme har dermed moderat respons på grunner hvis:

- den som regel er mottakelig for grunner,
- mottakeligheten er realitetsforankret, og
- den reagerer rasjonelt i minst ett av alle mulige, tenkelige scenarier

Hvis en vil knytte denne forståelsen til den enkelte aktør, kan en si at en aktør tilfredsstiller kravet om respons på grunner hvis mekanismen som styrer hans handlinger har moderat respons på grunner.²¹

Kommentar

Det er god grunn til å akseptere de sentrale elementer i denne framstillingen, etter min mening. Blant annet synes det nokså innlysende at det ikke kan kreves sterk respons på grunner for veiledningskontroll og moralsk ansvar, for det er vel få eller ingen mennesker som alltid er mottakelig for grunner og handler i samsvar med dette. Og de to eksemplene som er nevnt på moderat og svak respons på grunner (Jennifer og mannen på fergen), er tydelige eksempler på de to formene for respons. Men eksemplet om mannen som dreper alle på fergen unntatt han som røyker på en Gambierpipe, er svært ytterliggående og etterlater en usikkerhet om hvor grensen går mellom moderat og svak respons på grunner. Det hadde derfor vært ønskelig med en utdypet beskrivelse som gir klarere beskjed om hva det vil si å være som regel mottakelig for grunner, og derfor moralsk ansvarlig; og hvordan en aktør er som *ikke* som regel er mottakelig for grunner – og derfor *ikke* er moralsk ansvarlig.

8.3 Eierskap til den operative mekanismen og å ta ansvar

Fischer og Ravizza hevder at den andre betingelsen for veiledningskontroll – i tillegg til moderat respons på grunner – er eierskap til den operative mekanismen for veiledningskontroll og moralsk ansvar, eller sagt på en annen måte: den operative mekanismen må være aktørens egen. Fischer og Ravizza hevder at en persons eierskap til den operative mekanismen består i at han *tar ansvar*. De skriver at evnen til å ta ansvar er basert på tre typer av holdninger – eller tre betingelser (Fischer and Ravizza 1998: 210-213):

- Personen må se på seg selv som en aktør, dvs. at visse ting i verden er resultat av hans valg og handlinger.

²¹ Fischer og Ravizza forutsetter at det finnes eksternt rasjonelt forsvar for reaktive holdninger. Jeg antar at de også forutsetter at det også finnes et eksternt grunnlag for bedømmelse av grunner (dvs. et grunnlag for å skille mellom akseptable og uakseptable grunner). I denne avhandlingen blir derimot lagt til grunn at det finnes et *internt*, og ikke eksternt, forsvar for reaktive holdninger og et internt grunnlag for bedømmelse av grunner. Det innebærer en annen forståelse av definisjonen av *moderat respons på grunner* enn den forståelsen som F&R forventes å ha.

- Han må se på seg selv som et aktuelt mål for reaktive holdninger, dvs. at andre mennesker kan reagere med moralsk anerkjennelse eller avvisning (eller andre holdninger) på hans atferd.
- Disse to nevnte betingelsene må være basert på personens egen oppfatning av hva som er sant og riktig ("*the individual's evidence*")

I kapittel 9 blir det gitt en egen behandling av denne betingelsen for veiledningskontroll – å ha eierskap over den operative mekanismen og å ta ansvar.

8.4 Joseph Raz' kritikk

En forfatter som utfordrer Fischer og Ravizzas syn, er Joseph Raz (Raz 2010). Raz' innspill er riktignok ikke rettet mot Fischer og Ravizza, og jeg kjenner ikke til noe svar fra F&R på hans artikkel. Noen synspunkter i hans artikkel kan imidlertid være et godt utgangspunkt for drøfting av Fischer og Ravizzas syn på ansvar og kontroll.

Raz skiller mellom tre typer av ansvar. Vi skal her se på to av dem (Raz 2010: 433):

- 1 Mennesker er ansvarlig¹ hvis og bare hvis de har evnen til rasjonell handling (*the capacity for rational action*). (Raz presiserer her at vår *evne til rasjonell handling* er mer omfattende enn bare våre *rasjonelle evner (rational capacities)*. Rasjonelle evner er knyttet til evnen til å tenke og å ta beslutninger osv., mens evne til rasjonell handling også omfatter sansning, hukommelse og kroppskontroll for å kunne handle effektivt.)
- 2 Mennesker som er ansvarlig¹ er ikke nødvendigvis ansvarlig² for alt de gjør. Når de for eksempel handler under hypnose, er de ansvarlig¹, men ikke ansvarlig². For å være ansvarlig² må deres handlinger være knyttet til deres evne til rasjonell handlemåte (*rational agency*).

I resten av artikkelen forklarer Raz hva han legger i begrepet ansvarlig². I det følgende blir derfor Raz' begrep ansvarlig² omtalt som *ansvarlig* eller simpelthen som Raz' ansvarsbegrep.

Raz innleder sin artikkel med følgende påstand: "we can ... be responsible ... for actions we do not adequately control, for negligent action, and for non-intentional omissions" (Raz 2010: 433). Han illustrerer sitt syn med å beskrive en drukkenbolt som går ut av en bar. Drukkenbolten sjangler, dumper bort til et bord og velter et vinglass som knuser. I tillegg til kontrollelementet trekker Raz også inn spørsmålet om drukkenboltens handlinger er intendert. Han skriver:

Breaking the glass was unintentional and uncontrolled. But walking out was intentional, and yet not properly controlled. Control is a matter of degree. The drunk controlled his walking sufficiently to get out of the bar, but not well enough to avoid bumping into the furniture. (op.cit: 433)

Med bakgrunn i dette eksemplet hevder han at det er tre nødvendige og tilstrekkelige betingelser for at en person skal kontrollere sin handling:

- (1) either one performs it because one intends to do so, or one performs it, aware that one does so, by performing another action which one intends to perform.
- (2) The performance is guided by one's intentions and one's beliefs, so that to the extent that one's factual beliefs are true one does not, in performing the action, do anything else which one believes one should not (on balance) do;
- (3) In so far as realisation of the previous conditions depends on control of one's body they are securely realised. (op.cit.: 434)

Raz omtaler *The Control Principle* slik: "we are responsible ... for X if and only if X is under our rational control, or only because, and to the extent that X has aspects which are under our rational control". Hovedelementet i Kontrollprinsippet er altså at vi har ansvar for en handling hvis og bare hvis handlingen er under vår rasjonelle kontroll. Raz hevder at dette er en utilstrekkelig betingelse for ansvar.²² Han viser til eksemplet med drukkenbolten, og hevder at han er ansvarlig for å gå ut av baren. Drukkenbolten (la oss kalle ham Dick) er altså ansvarlig for en handling som riktignok er intendert, men ikke under adekvat kontroll.

Raz imøtegår også det synet at vi er ansvarlig for intenderte handlinger og bare dem. Han hevder at vi kan være ansvarlig for ikke-intenderte unnlater, og det kan være intenderte handlinger som vi ikke er ansvarlig for. Raz hevder derfor at et intensjonsprinsipp konstruert på tilsvarende måte som Kontrollprinsippet er utilstrekkelig betingelse for ansvar.

Raz' svar er knyttet til det han kaller *domain of secure competence* (område for sikker kompetanse). Han skriver:

We develop our sense of who we are in part by evolving and becoming aware of our secure competence. That recognition takes a normative form: if an act is a kind within our secure competence we hold ourselves responsible ... for it ..., or for doing whatever we happen to do in failing to perform it.... . (op.cit.: 435)

²² Raz hevder at Kontrollprinsippet ikke er en *tilstrekkelig* betingelse for ansvar. Men det er grunn til å spørre om *kontroll* derimot er en *ikke nødvendig* betingelse for ansvar - i følge Raz.

Så hvis vi oppfatter evnen til *ikke* å knuse glass som innenfor området for vår sikre kompetanse, har vi ansvar også når vi mangler adekvat kontroll og knuser glasset.

Nedenfor blir det stilt noen spørsmål på grunnlag av Raz' artikkel, blant annet om forholdet mellom sikker kompetanse og kontroll, og om forholdet mellom kontroll og ansvar.

8.5 Vurdering av kontrollbegrepet

Joseph Raz hevder at mennesker kan ha ansvar uten adekvat kontroll. Han skriver at drukkenbolten Dick har ansvar (riktignok et avledet ansvar, altså redusert ansvar), men han har ikke kontroll idet han sjangler og knuser glasset. Tisynelatende svekker Raz' argument Fischer og Ravizzas teori om moralsk ansvar basert på veiledningskontroll. Men det byr på noen problemer å vurdere Raz' synspunkter opp imot Fischer og Ravizzas, fordi forfatterne bruker de sentrale begrepene *ansvar* og *kontroll* på ulikt vis. Det er derfor behov for først å klargjøre hvordan forfatterne definerer de sentrale begrepene. Jeg vil se om dette kan gi svar på Raz' utfordring.

Fischer og Ravizzas og Raz' ansvarsbegrep

Fischer og Ravizzas ansvarsbegrep er basert på veiledningskontroll, moderat respons på grunner og eierskap til den operative mekanismen. I tillegg er deres ansvarsbegrep basert på reaktive holdninger, ros og klander - altså andres reaksjoner – og dessuten kommer det inn et subjektivt element med kravet om aktøren må se seg selv som ansvarlig og ta ansvar. Raz' definisjon er derimot avgrenset til aktørens evner og egenskaper med fokus på de rasjonelle evnene: rasjonell handling og rasjonell handlemåte. Raz trekker også inn begrepet *avledet ansvar (derivative responsibility)*. Han hevder at Dick har ansvar for å ta beslutningen om å gå ut av baren uten hjelp, og et avledet ansvar for å gå, å sjangle, å dumpe bort til bordet og å velte vinglasset.

Fischer og Ravizza hevder at noen mennesker er utsatt for ansvarsundergravende mekanismer, slik at de ikke er moralsk ansvarlig for sine handlinger. Dette gjelder for manipulerede individer eller individer som er mer eller mindre fullstendig uten kontroll. Kan det tilsvarende gjelde for Raz' drukkenbolt, slik at han ikke er ansvarlig verken for å gå ut av rommet, sjangle, støte bort til bordet eller knuse glasset? Det er mulig at F&R ville hevde det, selv om Raz tilskriver Dick et avledet ansvar for å knuse glasset.

Fischer og Ravizzas og Raz' kontrollbegrep

En kan imidlertid omskrive fortellingen om drukkenbolten Dick og beskrive den edru personen Didrik som mister balansen og knuser et vinglass. En edru og for øvrig tilregnelig person (etter F&R's språkbruk: kandidat for reaktive holdninger) som et øyeblikk mister balansen og knuser et glass, kan neppe sies å være utsatt for en ansvarsundergravende mekanisme. Jeg forutsetter at både Fischer og Ravizza og Raz vil si at han har moralsk ansvar. Men Raz vil hevde at han mangler kontroll. Hva vil så F&R svare?

I følge Fischer og Ravizza er veiledningskontroll blant annet betinget av moderat evne til å respondere på grunner. F&R skriver om Jennifer som går på baseballkamp istedenfor å gjøre hjemmearbeid når de skal forklare hva de mener med dette. De skriver om Jennifer at hun ikke må reagere på fornuftsgrunner i *alle* situasjoner for å ha veiledningskontroll. Det er tilstrekkelig at hun er en person som reagerer på fornuftsgrunner *noen ganger*. På dette grunnlaget antar F&R at hun hadde en viss reseptivitet overfor fornuftsgrunner *også* da hun brøt med disse grunnene og gikk på baseballkamp istedenfor å gjøre sitt hjemmearbeid, dvs. at hun hadde evne til å se at hun burde gjøre sitt arbeid, men at hun likevel foretrakk å gå på baseballkamp. Hvis hun i tillegg til dette også kan *ta* ansvar, har hun veiledningskontroll og moralsk ansvar.

Veiledningskontroll er altså et begrep som beskriver en persontype med visse egenskaper. Dette er egenskaper som personer ikke er født med, men som de fleste erverver seg. Men noen erverver seg aldri disse egenskapene. Når en person har fått veiledningskontroll, er dette en ganske stabil tilstand. Han må bli utsatt for alvorlige skader for å miste en veiledningskontroll som han har ervervet. Man mister ikke veiledningskontroll i en uforutsett situasjon, hvor man i følge dagligtalen kan "miste kontrollen". Veiledningskontroll er altså en relativt stabil personegenskap, og ikke en egenskap ved en handling, hendelse eller situasjon. Selv om veiledningskontroll er en relativt stabil egenskap ved en person og ikke en situasjonsbetinget egenskap som varierer med personens ulike handlinger i ulike situasjoner, understreker Fischer og Ravizza at aktørperspektivet på veiledningskontroll er utilstrekkelig. Istedenfor å studere aktørene må en se på de operative mekanismene som ligger til grunn for aktørenes handlinger.

Ved å gjøre det klart at veiledningskontroll er en relativt stabil mekanisme hos den enkelte person – og ikke et situasjons- eller atferdsbetinget fenomen – har Fischer og Ravizza besvart den mest åpenbare innvendingen som framkommer på grunnlag av Raz' artikkel. En kan altså ha veiledningskontroll selv om en har situasjonsbetinget mangel på kontroll (mister balansen og knuser et glass). Det innebærer at drukkenbolten Dick har veiledningskontroll hvis han:

- som regel er reseptiv overfor grunner
- reagerer på grunner noen ganger

Dick kan derfor ha veiledningskontroll – og dermed ansvar - selv om han i en akutt situasjon mangler kontroll. På tilsvarende måte kan den edru personen Didrik ha veiledningskontroll og ansvar, selv om han i en akutt situasjon mangler (en form for) kontroll.

Likevel er det usikkert hvor dekkende veiledningskontroll er som betingelse for moralsk ansvar. Nedenfor trekker jeg fram et par forhold som *kan* indikere at veiledningskontroll spiller en noe mer beskjeden rolle enn F&R hevder. Det vil imidlertid kreve grundigere undersøkelser enn det er plass til her, å avklare om disse argumentene virkelig innebærer at veiledningskontroll skal ha en mer beskjeden plass i en undersøkelse av moralsk ansvar.

Situasjonsbetinget mangel på ansvar

Hvordan kan vi forklare en tidsavgrenset og situasjonsbetinget mangel på ansvar hvis veiledningskontroll bare forklarer ansvar på aktørnivået, og ikke forklarer tidsavgrensede variasjoner i ansvaret knyttet til i ulike situasjoner og i forhold til ulike handlinger? La oss anta en versjon 2 av fortellingen om den edru personen Didrik, hvor han blir dyttet over bordet og knuser glasset. Har han moralsk ansvar for å knuse glasset ved versjon 2? Både i versjon 1 og 2 er Didrik en person med veiledningskontroll og moralsk ansvar. I begge versjoner mangler han likevel en form for situasjonsbetinget kontroll. I versjon 1 er han uvøren og uoppmerksom og knuser glasset. Ved å være oppmerksom kunne han ha unngått å knuse glasset. Som en person med veiledningskontroll antas det derfor at han har moralsk ansvar for å knuse glasset. Men hva med versjon 2? I versjon ble han dyttet, og hadde ikke i sin makt å unngå å falle over bordet og knuse glasset. Hadde han også i denne versjonen av fortellingen ansvar for å knuse glasset?

Moralsk ansvar er i utgangspunktet basert på en persons veiledningskontroll, og ikke kontroll i enkeltsituasjoner. Med disse premissene skal Didrik ha moralsk ansvar både i versjon 1 og 2 av fortellingen. Men sunn fornuft tilsier at han har ansvar for å knuse glasset i versjon 1 når han faller, men ikke i versjon 2 når han blir dyttet. Hvordan skal da forholdet mellom kontroll og ansvar forklares?

Det kan gis ulike forklaringer på den tilsynelatende uoverensstemmelsen i forholdet mellom kontroll og ansvar på aktørnivået og på atferdsnivået. En forklaring kan gis ved å hevde at Didrik kunne ha handlet annerledes i versjon 1. Han kunne ha vært oppmerksom og latt være å knuse glasset. Imidlertid synes F&Rs deterministiske forutsetning å skape problemer for denne forklaringen. Det er vanskelig å tilpasse en forklaring om at Didrik kunne ha handlet annerledes til den deterministiske forutsetningen i avhandlingen.

En kan også rette søkelyset på tiden før Didrik mister balansen. Som en person med veiledningskontroll har han normalt også kontroll over sin atferd. Han forutsettes altså å ha atferdskontroll en viss tid før han knuser glasset, både i versjon 1 og 2. Og i begge versjoner endrer han seg fra atferdsbetinget kontroll til ikke-kontroll. Ansvar for endringen er imidlertid ulik i de to versjonene. I versjon 1 har han ansvar for overgangen fra oppmerksomhet til uoppmerksomhet – og fra kontroll til ikke-kontroll. Han har ansvar for å ha brakt seg selv ut av atferdskontroll, og for konsekvensene av dette. Det har ikke inntruffet noe som skulle frata ham hans veiledningskontroll. Han har derfor ansvar for å knuse glasset i versjon 1. I versjon 2 blir han brakt midlertidig ut av atferdsbetinget kontroll av utenforliggende forhold (forhold utenfor ham selv). En kan si at disse forholdene virker ansvarsundergravende og fritar ham for ansvar for å knuse glasset. Men blir han brakt ut av veiledningskontroll i versjon 2? Slik F&R framstiller det, er veiledningskontroll en aktørbasert, og ikke en situasjonsbetinget egenskap. Betyr det at mangel på ansvar i dette tilfellet ikke er betinget av manglende veiledningskontroll, og derfor må begrunnes på annen måte? En kunne ønske at dette spørsmålet ble vurdert av Fischer og Ravizza, men det er ikke blitt gjort. Dette er derfor vanskelig å svare på spørsmålet på grunnlag av Fischer og Ravizzas tekst.

Kontroll og kompetanse

Både Raz og F&R opererer med noen grunnleggende egenskaper knyttet til moralsk ansvar. For Raz er *området for sikker kompetanse* et grunnleggende begrep. For den edru personen Didrik er trolig evnen til ikke å knuse glass innenfor området for sikker kompetanse, også når han mangler adekvat kontroll og knuser glasset. Og Fischer og Ravizza vil trolig si at Didrik har veiledningskontroll, også når han mangler adekvat kontroll og knuser glasset. Det er altså en parallellitet mellom Raz' kompetansebegrep og F&Rs kontrollbegrep. En kan derfor spørre om det er kontroll eller kompetanse som *egentlig* betinger moralsk ansvar. Eller er det både kontroll og kompetanse?

En kan også vise til F&Rs redegjørelse om to betingelser for moralsk ansvar, kunnskaps- og kontrollbetingelsen. Fischer og Ravizza har gjort oppmerksom på at deres fokus er rettet mot kontrollbetingelsen. Kan Raz' kompetansebegrep knyttes til den andre betingelsen – kunnskapsbetingelsen? Dette spørsmålet kan stilles, men det er vanskelig å gi et klart svar, bl.a. fordi den gjennomgåtte litteraturen i liten grad gir noen drøfting av kunnskapsbetingelsen.

9. Eierskap til den operative mekanismen

I tillegg til respons på grunner er veiledningskontroll også betinget av aktørens eierskap over den operative mekanismen. I dette kapitlet blir Fischer og Ravizzas syn på eierskap til den operative mekanismen gjennomgått og drøftet. For å kunne gjennomføre drøftingen blir også historiske aspekter ved ansvarsbegrepet vurdert. Til slutt i kapitlet gis en samlet drøfting.

9.1 Eleonore Stumps kritikk

I artikkelen "Responsiveness and Moral Responsibility" fra 1987 skriver Fischer at dersom mekanismen som styrer en handling responderer på grunner, er aktøren moralsk ansvarlig (og har veiledningskontroll): «an agent is morally responsible for performing an action insofar as the mechanism that actually issues in the action is reason-responsive» (Fischer 1987: 85-86). Denne artikkelen, og andre tidligere av Fischers tidlige arbeider, skiller seg fra Fischers senere arbeider på noen områder:

- Fischer skiller ikke mellom svak og moderat respons på grunner.

- Han har lagt til kravet om eierskap til den operative mekanismen som en betingelse for veiledningskontroll og moralsk ansvar.

Fischers tidlige arbeid er blitt kritisert av flere forfattere, blant annet Eleonore Stump.

Stumps fortelling om Raymond

Stump illustrerer sitt syn med en fortelling om Raymond i filmen *The Manchurian Candidate* som var programmert av kaprere til å drepe sin sjef. Stump legger fram flere alternativer av fortellingen. I ett av alternativene hindrer kaprerne Raymond i å drepe sjefen, fordi de er alliert med Raymonds sjef. Stump skriver:

In this case Raymond meets all of Fischer's conditions for weak reasons-responsiveness: there is some possible world in which there is sufficient reason for Raymond to do otherwise than kill his boss, and in that world Raymond uses the same mechanism he used in the actual world in killing his boss, but he doesn't kill him. And yet, even though Fischer's conditions for weak reasons-responsiveness are met, Raymond clearly is not morally responsible for killing his boss in the actual world, in which he is controlled and programmed by his captors. (Stump 1993: 259)

Stump hevder at Raymond tilfredsstiller kravene til svak respons på grunner. Men han er ikke moralsk ansvarlig for å drepe sin sjef, fordi han er kontrollert og programmert av kaprerne. Fischers betingelse om svak respons på grunner for å ha moralsk ansvar må derfor suppleres, slik at tilfeller som Raymond holdes utenfor – og ikke blir betraktet som moralsk ansvarlige. Stump foreslår:

To rule out this sort of case involving Raymond, what we need is an account of the sort Aquinas provides, which requires that the mechanism be the agent's own intellect and explains the crucial role of the intellect in volition, thereby denying moral responsibility to an agent whose intellect is programmed or whose will is coerced. (op.cit.: 260)

Stump viser altså til Thomas Aquinas og hevder at mekanismen som styrer handlingen, må være aktørens eget intellekt.

Fischer og Ravizzas justering av teorien om moralsk ansvar

På bakgrunn av kritikken fra Stump og andre filosofer innfører Fischer og Ravizza noen nye elementer i sin teori om moralsk ansvar i boka «Responsibility and Control»:

- De strammer inn kravet til respons på grunner. De krever ikke lenger bare svak respons, men moderat respons som innebærer sterk grad av reseptivitet for grunner.

- De legger til kravet om eierskap til den operative mekanismen som en betingelse for veiledningskontroll og moralsk ansvar.

Spørsmålet er så om F&Rs grep gir svar på Stumps utfordring. Vi skal se på hver av de to endringene som de foretar, og vurdere i hvilken grad de bidrar til en mer dekkende beskrivelse av veiledningskontroll og moralsk ansvar.

Vurdering

Stump har vist at Raymond har svak respons på grunner. Men har han også moderat respons på grunner? Fischer og Ravizza hevder at en mekanisme har moderat respons på grunner hvis den som regel er mottakelig for grunner, hvis mottakeligheten er realitetsforankret, og hvis mekanismen reagerer rasjonelt i minst ett av alle mulige, tenkelige scenarier. Hvordan passer Raymonds reaksjoner til denne definisjonen av moderat respons på grunner?

I et av alternativene blir Raymond programmert til ikke å drepe sin sjef. En kan si at mekanismen som styrer hans handling, reagerer rasjonelt i dette alternativet, og at Raymond derfor reagerer rasjonelt i minst ett av alle mulige, tenkelige scenarier. Men tilfredsstiller Raymond de andre betingelsene for moderat respons på grunner? Er den mekanismen som styrer Raymond som regel mottakelig for grunner? Og er mottakeligheten realitetsforankret? Jeg antar at mekanismen som styrer Raymond, så lenge han er programmert, ikke er mottakelig for grunner, men er mottakelig for kaprernes tvilsomme hensikter. Videre antar jeg at mottakeligheten ikke er realitetsforankret, men forankret i de samme tvilsomme hensiktene. Siden Raymond ikke tilfredsstiller alle tre betingelsene, antas han ikke å ha moderat respons på grunner, og han har derfor heller ikke veiledningskontroll eller moralsk ansvar.

Stump hevder at vi må trekke inn at (den operative) mekanismen må være aktørens eget intellekt og forklare intellektets avgjørende rolle for å svare på utfordringen som historien om Raymond representerer. Jeg har imidlertid ovenfor argumentert for at det er tilstrekkelig å stramme inn kravet til respons på grunner for å kunne konkludere om det vi alle er enige om, nemlig at Raymond - som programmert og kontrollert av kriminelle kaprere - ikke har moralsk ansvar for sin handling. Jeg kan ikke se at historien om Raymond *i tillegg* krever at den operative mekanismen må være aktørens eget intellekt. Jeg avviser ikke at

det kan være slik at den operative mekanismen (som ligger til grunn for moralsk ansvar) må være aktørens eget intellekt. Det kan synes til og med sannsynlig at det er slik. Men jeg kan ikke se at historien om Raymond viser at mekanismen må være aktørens eget intellekt. Og hvis det er slik at dette saksforholdet har kommet inn i Fischer og Ravizzas tenkning på uklare premisser, kan det få følger for måten de har behandlet saksforholdet på. Jeg skal gjøre nærmere rede for dette i det følgende.

9.2 Fischer og Ravizzas syn

I «Responsibility and Control» hevder Fischer og Ravizza at det kreves eierskap til den operative mekanismen for veiledningskontroll og moralsk ansvar - i tillegg til moderat respons på grunner. Dette er svært likt Eleonore Stumps syn, som er at mekanismen må være aktørens eget intellekt. F&R's syn kan imidlertid sies å være noe videre enn Stumps forslag, idet de skriver at mekanismen må være aktørens egen uten å begrense det til aktørens intellekt.

Fischer og Ravizza hevder at en aktør som har eierskap over den operative mekanismen, ser seg selv som passende kandidat for reaktive holdninger – og *tar ansvar*. De omtaler tre betingelser for å ta ansvar, knyttet til personens selvbilde og vurdering av evidens (se kap. 8.3). Å ta ansvar synes å være både en nødvendig og tilstrekkelig betingelse for eierskap til den operative mekanismen. Ved å ta dette grepet finner F&R det nødvendig å trekke inn bakgrunnen for – eller historien bak - mekanismen som styrer aktørens handlinger: «We contend that an agent's being morally responsible is *genuinely historical* in the sense that it requires that the agent have previously taken responsibility.» (F&R 1998: 214). Å ta ansvar er altså er historisk, og ikke et ahistorisk, fenomen. Det blir orientert nærmere om dette nedenfor.

For å forklare sitt syn bringer Fischer og Ravizza en historie om en person - Judith – som blir manipulert til å slå sin venninne. De presenterer historien om Judith i tre versjoner. I den første versjonen blir Judith manipulert (ved elektronisk stimulering av hjernen) slik at hun *må* slå venninnen Jane. Fischer og Ravizza skriver at mekanismen som leder til denne handlingen ikke er moderat responsiv på grunner, fordi stimuleringen av hjernen fører til at Judith har et uimotståelig ønske om å slå Jane. Judith har derfor ikke moralsk ansvar for å slå Jane. I Versjon 2 blir Judith manipulert til å få en sterk, men ikke uimotståelig, drift for

å slå Jane. Dersom slaget ville medføre at mange uskyldige mennesker ville dø, ville Judith la være å slå Jane. Fischer og Ravizza skriver om dette:

Our account of taking responsibility holds that agents typically have taken responsibility, in the past, for certain salient kinds of mechanisms of practical reasoning (uninfluenced by subliminal advertising, direct electronic stimulation of the brain, and so forth, and unreflective habit Given that Judith does not know about the manipulation of the scientist, and has not explicitly considered such manipulation, it is plausible to say that Judith has not taken responsibility for the mechanism that actually issues the action... (Fischer and Ravizza 1998: 233)

Mekanismen som leder til Judiths handling er i dette tilfellet moderat responsiv på grunner, fordi Judiths ønske om å slå Jane ikke er uimotståelig. Heller ikke i dette tilfellet har altså Judith moralsk ansvar for å slå Jane, i følge Fischer og Ravizza.

I forbindelse med Versjon 3 viser F&R til den tredje betingelsen for å ta ansvar, nemlig at aktørens syn på seg selv må være basert på hans bevis på en passende måte (*his evidence in an appropriate way*). I Versjon 3 er Judith blitt stimulert elektronisk til å ha et relevant syn på seg selv. Hun har ikke formet dette synet på en passende måte, og har derfor ikke moralsk ansvar, selv om hun har moderat respons på grunner.

Strukturen i denne historien likner på Stumps historie om Raymond. På tilsvarende måte som Stumps historie om Raymond, bruker F&R historien om Judith for å begrunne at eierskap til den operative mekanismen er betingelse for veiledningskontroll og moralsk ansvar. De skriver: "In certain cases involving direct manipulation of the brain (and similar influences), it is natural to say that the mechanism leading to the action is not ... the agent's own." (op.cit.: 230) F&R legger altså til grunn at en person som er manipulert, ikke er moralsk ansvarlig. Betingelsen om moderat respons på grunner er etter deres syn ikke tilstrekkelig for å utelukke at manipulerte personer betraktes som moralsk ansvarlige. Eierskap til den operative mekanismen må også inkluderes som betingelse for moralsk ansvar.

Noen innvendinger

Fischer og Ravizza tar opp to innvendinger mot å knytte moralsk ansvar til å ta ansvar (op.cit.: 217):

- Enkelte personer vil kunne bli fristet til ikke å ta ansvar - for å slippe å bli holdt ansvarlig.

- Deres definisjon på moralsk ansvar er ikke uttømmende, fordi det kan oppstå tilfeller hvor aktøren av en eller annen grunn ikke har tatt ansvar, men hvor han likevel er moralsk ansvarlig. Hva med dem som foretar brutale eller onde handlinger – skal ikke de bli omfattet av ansvarsbegrepet? Skal ikke de bli stilt til ansvar for sine handlinger?

Når det gjelder den første innvendingen, skriver de:

...the price of the refusing to take responsibility may be as high as that of accepting responsibility. If one fails to believe that one is an agent in this sense [that he sees himself as the source of his behavior – in a certain way], then one will presumably make no effort to affect the world by one's choices....., and whatever upshots happen to occur will be entirely accidental. Such an individual's behaviour would surely threaten others (as well as himself), and this alone would seem to give the community sufficient reason to intern the individual, or at least to significantly restrict his access to the rest of the community. This forced sequestration surely is not something to be sought. (op.cit.: 218-219)

De hevder altså at de som ikke tar ansvar, ikke er moralsk ansvarlige og ikke er kandidater for reaktive holdninger. De kan altså foreta brutale handlinger uten å være moralsk ansvarlig for disse handlingene. De kan eventuelt bli stilt rettslig til ansvar for sine handlinger, men moralsk sett kan de ikke bli stilt til ansvar for sine handlinger.

Når det gjelder den andre innvendingen, viser F&R til sin subjektivistiske posisjon. De gir et eksempel med en seiler. En person som ikke tar ansvar, er som en person som seiler for vær og vind uten å forsøke å styre båten. De skriver:

”The basic idea is that an individual who really does not see him-self as an agent and a fair target for the reactive attitudes cannot be deemed genuinely active and responsible. In *not seeing himself* in a certain way, he *fails to be* a morally responsible agent. Lacking the required view of himself, he *is* essentially passive, buffeted by forces that assail him.” (op.cit.: 221)

F&R's syn er at personer som ikke tar ansvar og ikke er moralsk ansvarlige, stiller seg utenfor samfunnet og den daglige samhandling med andre mennesker. Dette er ingen fristende posisjon. I praksis er det langt på vei umulig ikke å ta ansvar, og dermed å stille seg utenfor samfunnet. Fischer og Ravizza hevder derfor at man i praksis ikke innsnevrer begrepet moralsk ansvar ved å knytte begrepet til å ta ansvar.

9.3 Å ta ansvar som historisk fenomen

Som nevnt hevder Fischer og Ravizza at å ta ansvar – eierskap til den operative mekanismen - er et historisk fenomen. De skiller mellom genuint historiske fenomener og fenomener som ikke er historiske eller bare tilsynelatende er historiske. De skriver at de ikke-historiske fenomenene er tidskuttfenomener (*current time-slice phenomenons*) som kun er

avhengig av ”snap-shot”-egenskaper og ikke av historiske fakta, mens «certain phenomena are historical in the sense that they depend (in some interesting way) on features of their history» (Fischer and Ravizza 1998: 171).

De tilsynelatende historiske fenomenene kan også kalles for epistemiske historiske fenomener. Et eksempel på et epistemisk historisk fenomen hentes fra Freud. Freud hadde til behandling en pasient, Elisabeth, som var fysisk lammet av hysteri. Det viste seg at seksuell undertrykkelse på bakgrunn av opplevelser som barn var årsaken til hennes tilstand. Fischer og Ravizza skriver at hennes lammelse kan iakttas i den aktuelle situasjonen. Det var ikke nødvendig å kjenne hennes opplevelser som barn for å kunne fastslå hennes funksjonssvikt (og at hun er uten ansvar for sin reduserte mobilitet). Det dreier seg derfor ikke om et genuint historisk fenomen, fordi vår kunnskap om hennes fortid endrer ikke vår bedømmelse av hennes ansvar i nåtiden. Det er derimot et epistemisk historisk fenomen, fordi vi får utvidet forståelse av Elisabeths situasjon gjennom kunnskap om hennes fortid. Fischer og Ravizza uttrykker det slik:

What is distinctive about an epistemically historical phenomenon is that features of the past point us to snapshot features of the present; features of the past are *useful* in helping us characterize and understand the present, but all we “really” need to understand the present is contained in the snapshot features of the present. (op.cit.: 190)

Som kontrast til dette epistemisk historiske fenomenet beskriver F&R et fenomen som de hevder er genuint historisk. De gir to versjoner av en fortelling om en beruset sjåfør som kjører ned et barn. I det ene tilfellet hadde han drukket kraftig før han satte seg til å kjøre. I det andre tilfellet hadde han fått injisert alkohol uten sin viten og vilje. I det første tilfellet er han ansvarlig for sin beruselse, og dermed også ansvarlig for konsekvensene av dette. I det andre tilfellet er han ikke ansvarlig, han manglet både viten om og kontroll over at han fikk alkohol. Fischer og Ravizza skriver derfor at kunnskap om den historiske bakgrunnen er nødvendig for å bedømme om sjåføren hadde moralsk ansvar eller ikke. Vi trenger historisk kunnskap om en person for å klargjøre om han har moralsk ansvar. Moralsk ansvar er derfor et genuint historisk fenomen, og ikke noe epistemisk historisk fenomen eller et snap-shot- eller tidskuttsfenomen.

For Fischer og Ravizza utgjør fortellingen om sjåføren et bilde på noe annet som er deres egentlige budskap, nemlig at å *ta ansvar* er et historisk fenomen. Som vist ovenfor, hevder

de at å ta ansvar er del av en prosess. De erklærer seg derfor uenig med Frankfurt, som hevder at ansvar er strukturelt og ahistorisk:

... he infers from of continuous manipulation that the subject's "history is utterly episodic and without inherent connectedness." But surely this is a spurious transition, for continuous manipulation is compatible with continuity and intelligibility. ... the clear implication seems to be that a certain degree of historical continuity and connectedness is essential to moral responsibility, which would appear to be incompatible with a current slice model of moral responsibility. (op.cit.: 199)

9.4 Noen spørsmål

Jeg stiller dette spørsmålet til Fischer og Ravizzas framstilling:

Er moderat respons på grunner en tilstrekkelig betingelse for veiledningskontroll, eller er eierskap til den operative mekanismen, dvs. å ta ansvar, en nødvendig tilleggsbetingelse?

For drøftingen av dette spørsmålet skal noen forhold vurderes:

- 1 Er å ta ansvar – og dermed eierskap til den operative mekanismen og moralsk ansvar – historiske fenomener? Og utgjør å ta ansvar en egen betingelse for veiledningskontroll og moralsk ansvar, fordi det er et historisk fenomen?
- 2 Hvilken karakter har F&Rs tre betingelser for å ta ansvar?
- 3 Gir F&Rs historie om Judith tilfredsstillende begrunnelse for at eierskap til den operative mekanismen er en nødvendig betingelse for veiledningskontroll?
- 4 Finnes det tilfeller hvor en aktør har ansvar selv om han ikke tar ansvar?

Til slutt i kapitlet blir det vurdert om eierskap til den operative mekanismen bør opprettholdes som en betingelse for veiledningskontroll, men gis en mer beskjeden plass enn Fischer og Ravizza har gitt det.

Moralsk ansvar som historisk fenomen

For å forklare at moralsk ansvar er et historisk fenomen forteller Fischer og Ravizza om en sjåfør som kjører i beruset tilstand. Etter mitt syn illustrerer ikke denne fortellingen F&Rs poeng på en helt god måte. Jeg stiller spørsmål ved versjonen hvor sjåføren får injisert alkohol uten viten og vilje. Selv om sjåføren ikke er kjent med inntaket av alkohol, bør han vel kjenne virkningen. En kan si at han bør merke sin svekkelse, selv om han ikke kjenner årsaken til denne svekkelsen. Han bør også unnlate å kjøre bil i svekket tilstand, selv om han ikke kjenner årsaken til svekkelsen. En kan derfor motsi F&R, og hevde at sjåføren har ansvar for å sette seg i bilen i svekket tilstand og kjøre på barnet, selv om han har fått injisert alkohol uten viten og vilje.

En kan imidlertid presisere injeksjonsversjonen av fortellingen: Sjåføren fikk injisert små mengder alkohol, så vidt over promillegrensen, som det ikke var grunn å merke. Han kan derfor ikke anses moralsk ansvarlig for å sette seg i bilen og kjøre i beruset tilstand, for han har ikke noen erkjennelse om sin beruselse og kan ikke forventes å ha slik erkjennelse. For å kunne vurdere sjåførens moralske ansvar, må man derfor kjenne den historiske bakgrunnen i forbindelse med inntaket av alkohol, dvs. om han selv drakk eller om han fikk alkohol injisert uten å være kjent med det.

Vi kan også trekke inn Fischer og Ravizzas fortelling om Jennifer. Fischer og Ravizza gjør rede for ulike fortidige handlinger, og på den bakgrunnen finner de at Jennifer som regel er mottakelig for grunner (i tillegg til å tilfredsstille de andre kravene til moderat respons på grunner). De konkluderer derfor at Jennifer har veiledningskontroll og moralsk ansvar. Deres bedømmelse er altså ikke grunnet på en enkeltobservasjon, men på flere observasjoner over en historisk periode.

En kan merke seg at i fortellingen om Jennifer dreier spørsmålet seg om hun har moderat respons på grunner. Det kan derfor synes som å ta ansvar ikke er alene om å være et historisk fenomen, men at det ligger historiske trekk også til grunn for respons på grunner. Er det da slik at begge betingelsene for veiledningskontroll - både moderat respons på grunner og å ta ansvar - er historiske fenomener?

Før vi kan svare på dette, må det klarlegges om fortellingen om Jennifer er genuint historisk eller bare tilsynelatende (epistemisk) historisk. Freuds pasient Elisabeth som var lammet av hysteri, danner et eksempel på epistemisk historisk fenomen. Fortellingen om Elisabeth viser ikke til et genuint historisk fenomen, fordi vår kunnskap om hennes fortid ikke endrer vår bedømmelse av hennes manglende ansvar i nåtiden. Kunnskapen om hennes fortid er ikke egnet til å dømme om ansvar, men bare til å *forklare* hvorfor hun er en person uten moralsk ansvar. Når det gjelder Jennifer, derimot, forstår jeg F&R slik at kunnskapen om hennes fortid brukes til å bedømme om hun er en person med moralsk ansvar. Så langt jeg kan skjønne, er Jennifers moralske ansvar genuint historisk, fordi bedømmelsen av hennes moralske ansvar ikke er basert på en enkeltobservasjon, men på flere observasjoner over en historisk periode.

Nå kan det innvendes at fortellingen om Jennifer viser en annen type historisk prosess enn prosessen fram til å ta ansvar. Fischer og Ravizza synes å mene at prosessen fram til å ta ansvar er en fremadskridende personlighetsutviklingsprosess, hvor personen på et gitt tidspunkt når det nivået av modenhet at han tar ansvar. Fortellingen om Jennifer avslører derimot vekslende tilfeller hvor hun noen ganger er mottakelig for grunner, og noen ganger ikke er det, og hvor bedømmelsen av ansvar bygger på en samlet vurdering av disse ulike enkelthandlingene. Bedømmelsen av Jennifers respons på grunner kan derfor synes mer episodisk basert enn bedømmelsen av en utviklingsprosess fram til å ta ansvar.

På den annen side kan en forsvare at Jennifers moralske ansvar er et historisk fenomen. For det første gir ikke F&R beskrivelse av genuine og epistemiske historiske fenomener noe klart grunnlag for å hevde at Jennifers moralske ansvar *ikke* er et historisk fenomen. Dernest kan det hevdes at Jennifers utvikling fram til moderat respons på grunner kan være resultat av en prosess, selv om *observasjonene* gir et episodisk inntrykk.

På denne bakgrunn kan en spørre om det historiske ikke er noe særskilt ved å ta ansvar, men at respons på grunner også må anses å være et historisk fenomen. En kan også spørre om det faktum at å ta ansvar synes å være et historisk fenomen, er tilstrekkelig begrunnelse for å framstille det som en særskilt betingelse for veiledningskontroll – hvis det er tvilsomt om det historiske er noe særskilt ved å ta ansvar.

Tre betingelser for å ta ansvar

Fischer og Ravizza gjør rede for tre betingelser for å ta ansvar, og dermed for eierskap til den operative mekanismen (jf. kap. 8.3). De tre betingelsene er knyttet til personens selvbilde og vurdering av evidens: han må se seg selv som aktør, se seg selv som aktuelt mål for reaktive holdninger, og basere disse to atferdsformene på egen oppfatning av evidens. En kan spørre: Hvilket bidrag gir disse betingelsene til å forklare veiledningskontroll, som moderat respons på grunner ikke gir? Med andre ord: Hva er forholdet mellom betingelsene for å ta ansvar og betingelsen om moderat respons på grunner?

En kan si at de tre betingelsene for å ta ansvar handler om respons på grunner med en subjektiv tilnærming. Å se at visse ting i verden er resultat av valg og handlinger, er en evne til å se årsaker og konsekvenser. Det er en evne som en kan forutsette at en person

med moderat respons på grunner behersker. Det subjektive elementet kommer inn ved at personen ser seg selv i denne sammenhengen av årsaker og konsekvenser. Et tilsvarende argument kan føres om de to andre betingelsene. Moderat respons på grunner forutsettes å omfatte evnen til se på personer som aktuelle mål for reaktive holdninger. Det subjektive elementet kommer imidlertid inn når det er personen selv som er det aktuelle målet for reaktive holdninger. Og moderat respons på grunner må også omfatte en evne til å vurdere evidens. Det subjektive elementet kommer også her inn ved betingelsen om at evidens skal være basert på personens egen oppfatning.

Spørsmålet er så: Hvor viktige eller nødvendige er disse subjektive elementene ved å ta ansvar for veiledningskontroll og moralsk ansvar? Fischer og Ravizza søker å gjøre rede for dette blant annet gjennom en fortellingen om Judith.

Historien om Judith

Ovenfor er det beskrevet en historie i tre versjoner om en person - Judith – som blir manipulert til å slå sin venninne. I hvilken grad gir denne historien en begrunnelse for at moderat respons på grunner er utilstrekkelig for veiledningskontroll, og at veiledningskontroll i tillegg er betinget av eierskap til den operative mekanismen?

Mitt syn er dette: Versjon 1 er grei. Mennesker kan bli hypnotisert, hjernevasket eller grovt manipulert, slik at de ikke kan anses som moralsk ansvarlig for sine handlinger. Versjon 2 og 3 rammes etter mitt syn av kritikken til Robert Kane (se kapittel 2.5), nemlig at Judith ikke har noe å gjøre med vår verden hvor det sannsynligvis ikke finnes noen Frankfurt-kontrollører. Men jeg antar at Versjon 2 kan omarbeides til en troverdig form. Jeg foreslår denne omarbeidingen:

I Versjon Omarbeidet 2A (O2A) blir manipulasjonsofferet – la oss kalle henne Julie – utsatt for hypnose, hjernevask eller annen grov manipulasjon og hun manipuleres til å utføre en rekke handlinger. Men det går en grense for hvor langt manipulasjonen rekker. I Versjon Omarbeidet 2B (O2B) prøver manipulator å påvirke henne til å begå et mord. Da vegrer hun seg og vil ikke følge manipulators ordre. Er det da slik, som F&R hevder, at Julie mangler veiledningskontroll – og dermed også mangler moralsk ansvar - i alle situasjoner (både O2A og O2B) fordi hun er så sterkt utsatt for manipulasjon? Eller er det slik

at hun bryter ut av manipulasjonen når hun vegrer seg for å begå et mord (Versjon O2B), og at hun i det hun bryter ut av manipulasjonen også har veiledningskontroll? Hvis det siste er tilfelle, vil Versjon O2A være slik: Julie er manipulert, har ikke moderat respons på grunner og har ikke veiledningskontroll. I Versjon O2B har Julie brutt ut av manipulasjonen. Som ikke-manipulert har den mekanismen som ligger til grunn for hennes handlinger moderat respons på grunner, og Julie har også veiledningskontroll.

Forstått på denne måten viser ikke historien om Julie at betingelsen om moderat respons på grunner er utilstrekkelig for veiledningskontroll. Hvis vi antar at historien om Judith er konstruert og urealistisk, og at historien om Julie er en sannsynlig omarbeiding av historien om Judith til en realistisk form, kan heller ikke historien om Judith gi oss noen overbevisende begrunnelse for at betingelsen om moderat respons på grunner er utilstrekkelig for veiledningskontroll.

Jeg kan derfor ikke se at Fischer og Ravizza har vist gjennom sin historie om Judith at moderat respons på grunner er utilstrekkelig for veiledningskontroll, og at veiledningskontroll (og moralsk ansvar) i tillegg er betinget av eierskap til den operative mekanismen.

Å ta ansvar og ha ansvar

La oss se på en annen forståelse av å ta ansvar enn Fischer og Ravizzas. Det forutsettes at en av de følgende to betingelsene må gjelde for at personen P tar ansvar:

- P erkjenner at han har gjort handling a
- P erklærer at han ville ha tatt ansvar for handling a dersom han hadde gjort handling a

La oss også se på et eksempel: Karl er en ”gjenganger-kriminell” som har utført en rekke lovbrudd gjennom mange år. Etter å ha vært ute av fengsel i en periode, blir han brakt til avhør, siktet for nye lovbrudd. Han nekter å ha utført handlingen, til tross for sterke beviser. Han viser heller ikke noen sterk interesse for å forsvare seg og motbevise at han utførte lovbruddene. Han viser forakt og manglende interesse både for fellsskapets regler, slik disse kommer til uttrykk i straffeloven, og for de rettslige prosedyrene i form av avhør, rettssak og dom.

Karl tilfredsstillter ikke den første betingelsen for å ta ansvar, for han erkjenner ikke å ha utført noe lovbrudd. Han tilfredsstillter heller ikke den andre betingelsen, for han viser ingen respekt eller interesse for fellesskapets syn på at et menneske skal ta ansvar for sine handlinger. Han viser en konfrontasjonslinje overfor rettssamfunnet: Han og rettssamfunnet er to uforsonlige motstandere. Han ser det slik at han gjør hva han gjør, og at han ikke har noe å svare for (dvs. å ta ansvar for) overfor rettssamfunnet. Sett på denne måten tar ikke Karl ansvar (iallfall ikke overfor rettssamfunnet, men kanskje overfor noen andre). Skal man så trekke den slutning at Karl ikke er moralsk ansvarlig fordi han ikke tar ansvar?²³ Er det ikke tvert imot av interesse å slå fast at nettopp Karl, som unndrar seg ansvar og ikke er villig til å kommunisere i forhold til fellesskapets spilleregler i det hele tatt, *har* moralsk ansvar?

Det er viktig å merke seg at det her brukes en annen definisjon av å ta ansvar enn F&Rs definisjon. Det er også viktig å merke seg at Karl *kan* være en person som tar ansvar, selv om han ikke tar ansvar overfor rettssamfunnet, som han anser som sin fiende. Likevel kan en spørre hva Fischer og Ravizza vinner – og hva de taper – på å inkludere å ta ansvar som betingelse for moralsk ansvar. Gjennom drøftingen i avhandlingen er det ikke blitt klarlagt at å ta ansvar er en *nødvendig* betingelse for moralsk ansvar. Er det så en *tilrådelig* betingelse? Jeg tviler på dette dersom betingelsen forstås slik at Karl ikke er moralsk ansvarlig for sin atferd (inkludert lovbrudd). Dersom å ta ansvar er betingelse for å ha ansvar, skal dette etter mitt syn forstås slik at tilfeller som Karl tar ansvar og har ansvar.

Til oppsummering vil jeg understreke at jeg ikke har grunnlag for å avvise Fischer og Ravizzas syn på at eierskap til den operative mekanismen og å ta ansvar er betingelse for veiledningskontroll og moralsk ansvar, men jeg ser ikke at de har vist det gjennom den argumentasjonen som er gjennomgått her. Jeg viser også til at inkludering av å ta ansvar som betingelse for å ha ansvar *kan* ha implikasjoner som jeg stiller meg tvilende til.

²³ Karl kan naturligvis være rettslig ansvarlig for sine lovbrudd selv om han på grunn av antatt karakterbrist skulle bli vurdert til ikke å være moralsk ansvarlig.

Omdefinering av eierskap til den operative mekanismen

Selv om det skulle vise seg at eierskap til den operative mekanismen ikke utgjør en selvstendig betingelse for veiledningskontroll, kan denne faktoren likevel ha en plass i bildet – men en mer beskjeden plass enn Fischer og Ravizza har gitt den. F&R hevder at personer som ikke er utsatt for ansvarsundergravende mekanismer som hypnose, elektronisk stimulering av hjernen eller annen form for manipulasjon, har eierskap til den operative mekanismen. Forstått på denne måten kan eierskap til den operative mekanismen antas å være en av flere mulige betingelser for respons på grunner. Det er interessant oppgave å undersøke forholdet mellom eierskap til den operative mekanismen og respons på grunner på disse premissene, men det vil ikke bli gjort i denne avhandlingen.²⁴

10. Drøfting av det teoretiske grunnlaget

Fischer og Ravizza hevder at en person må se seg selv som passende kandidat for reaktive holdninger for å ha moralsk ansvar. De omtaler dette som en subjektivistisk tilnærming. De legger Galen Strawsons subjektivistiske teori om fri vilje og moralsk ansvar til grunn for sitt syn. Det blir her gjort rede for Strawsons teori, og Fischer og Ravizzas bruk av denne teorien blir vurdert.

10.1 Galen Strawsons subjektivistiske teori om fri vilje og ansvar

Fischer og Ravizza skriver:

... we agree with Galen Strawson in embracing a “subjectivist” approach to moral responsibility. Recall that this sort of approach requires that an agent has a certain kind of view of himself, in order to be morally responsible for his behavior. And this is precisely the case, on our account of taking responsibility (and moral responsibility). In order to be morally responsible, a person must see himself as agent who is an appropriate candidate for the reactive attitudes. (Fischer og Ravizza 1998: 228-229)

I kapittel 6 blir det gjort rede for Fischer og Ravizzas syn på at en person har moralsk ansvar når han er passende kandidat for reaktive holdninger. Her presiserer de dette synspunktet ved å klargjøre at han må *se seg selv* som passende kandidat for reaktive holdninger for å være passende kandidat for reaktive holdninger. De ser på dette som en

²⁴ I følge Fischer og Ravizza er å ta ansvar lik eierskap til den operative mekanismen. Det innebærer imidlertid ikke at å ta ansvar er lik eierskap til den operative mekanismen i det tilfellet som omtales her.

subjektivistisk tilnærming til moralsk ansvar, og finner et grunnlag for sin tilnærming i Galen Strawsons subjektivistiske teori om fri vilje og moralsk ansvar.

Galen Strawson hevder at nesten alle teorier om frihet er objektivistiske (Strawson, G. 1986, s. 25). Han nevner to typer av objektivistiske teorier: libertarianske og kompatibilistiske. Han analyserer disse teoriene, og finner at ingen av dem gir noen god forklaring på hva friheten egentlig er.

Strawson beskriver libertarianismens dilemma slik:

Consider any description D given by libertarians of the antecedents of an action A that they claim to be a free action, which allegedly makes it clear *why* A is a free action, and from which it does *not* follow that A's antecedents are partly indeterministic in nature. Given such a description one may say this: (1) let us suppose that these antecedents have determinedly come to be the way they are; for D does not rule this out. (2) But D is meant to show why A is free. (3) So to say that if these antecedents have determinedly come to be the way they are then the action is not after all a free one, as these libertarians must, is to admit that D has not after all made clear why A is a free action. (4) It follows that any adequate libertarian account of free action which purports to show how and why actions can be free must point to actual indeterministic elements among their antecedents, and must in addition explain how the indeterministic elements contribute to the actions' freedom. (op. cit.: 32)

Strawson hevder at libertarianismen ikke gir en positiv, detaljert beskrivelse av hva frihet er. Den positive beskrivelsen av fri handling må inkludere en beskrivelse av forutgående hendelser (*antecedents*). Disse hendelsene kan ikke være av deterministisk art, for determinisme kan ikke skape frihet. De forutgående hendelsene må altså ha indeterministisk karakter. Strawson hevder at libertarianerne ikke kan påvise at slike indeterministiske hendelser – eller indeterministiske elementer – peker henimot en fri handling (og langt mindre at indeterministiske hendelser *fører* til en fri handling, da en slik nødvendig sammenheng mellom en fri handling og en forutgående hendelse er i strid med selve begrepet *fri handling*). I tillegg til dette krever Strawson at libertariansk teori må kunne forklare *hvordan* indeterministiske elementer bidrar til at en handling blir fri. Heller ikke dette klarer libertariansk teori, i følge Strawson. Libertariansk teori kan derfor ikke levere noen tilfredsstillende teori om friheten, verken i form av fri handling, frie aktører, fritt valg eller fri vilje (s. 59).

Strawson viser også til at de libertarianske teoriene forutsetter at fri vilje finnes og at determinismen er falsk. Hvis en så hevder at determinismen er falsk, er det nærliggende å anta at indeterminismen er sann. Men fri vilje er ikke det samme som indeterminisme. Fri

vilje er selvdeterminering, i følge Strawson. Og libertarianske teorier klarer ikke å gjøre rede for forholdet mellom indeterminisme og selvdeterminering. På samme måte som disse teoriene ikke klarer å beskrive friheten, klarer de heller ikke å forklare ha selvdeterminisme er.

Når det gjelder kompatibilismen skriver Strawson at dersom determinismen er sann, er en aktørs nåtidige handlinger determinert av egne og andres tidligere handlinger – og av naturlovene som styrer handlingene. Dersom determinismen er sann, kan derfor ikke aktøren være fullt ut fri (*really free*), og hvis han ikke er fri kan han heller ikke være fullt ansvarlig (*truly responsible*) for sine handlinger. Strawsons argument likner på det som i kapittel 4 blir omtalt som The Principle of the Fixity of Past and Laws, og han finner dette argumentet tilstrekkelig tungtveiende til å avvise kompatibilismen.

Ved å gjennomgå libertarianismen og kompatibilismen hevder Strawson å ha utført en uttømmende analyse av det han kaller objektivistiske teorier om friheten, ettersom den første type av teorier avviser determinismen og den andre aksepterer muligheten av at determinismen er sann. (Den tredje hovedtypen av teorier om fri vilje og determinisme - hard inkompatibilisme - er ikke aktuell for Strawsons analyse, for hard inkompatibilisme representerer en teoriretning som hevder at fri vilje ikke finnes. En gruppe av teorier som blir omtalt som revisjonisme, blir heller ikke eksplisitt trukket inn i hans analyse, imidlertid kan en i en viss forstand hevde at han selv representerer denne fjerde hovedtypen.)

Strawson hevder derfor at fri vilje og moralsk ansvar objektivt sett ikke finnes, fordi ingen aktuelle objektivistiske teorier kan gi belegg for at det finnes. Likevel er det et uomtvistelig faktum at vi – eller i alle fall de fleste av oss - erfarer friheten, dvs. at vi erfarer at vi har fri vilje og kan handle fritt. Ikke bare erfarer vi friheten: vi *tror* på friheten, og vi kan ikke la være å tro på friheten. Hans grunnleggende ide er altså: “we are not really free and truly responsible agents at all, even if we cannot help believing that we are” (op. cit.: 311). I denne setningen finner Galen Strawson et konsentrert uttrykk for sin subjektivistiske teori om friheten. Han knytter teorien til Kants filosofi: “every being that cannot act except *under the idea of freedom ... is just for that reason from a practical point of view really free*” (op. cit.: 68).

I følge den subjektivistiske tilnærmingen må vi altså med nødvendighet tro at vi er fri, og det er nettopp dette som er betingelsen for vår frihet. Strawson gir et eksempel:

[A man] has a general conception of himself as a free agent. He faces a button. The good and bad consequences of pressing it or not pressing it, in the next twenty minutes, are finely balanced and momentarily different. He may be familiar with determinism, and believe it is true. But, he reflects, although he will, as soon as the twenty minutes are up, be able to say truly that what he did was determined, that in no way changes the fact that he now knows that he is able, now, to choose completely freely. (op.cit.: 70)

Strawson gir også et eksempel på en person som ikke har noen som helst ide om hva det vil si å være en fri aktør. Han skriver at en slik person som ikke har noen begreper om frihet, heller ikke kan være fri. Det framgår altså av dette eksempelet at det er mulig ikke å erfare – eller tro på – friheten. Det å ikke erfare friheten representerer imidlertid et unntak, og gjelder personer som utviklingsmessig er kommet til kort – de er av ulike grunner utviklet, slik at de ikke kjenner noen frihet.

Strawson konkluderer med at frihet *er* ideen om frihet, og at ideen om frihet er en nødvendig betingelse for friheten. Bare subjektivistiske teorier kan derfor være korrekte teorier om frihet.

Peter Strawson er blitt kritisert for å ta en for anti-teoretisk posisjon i sitt syn på moralsk ansvar. Det er mulig å se på Galen Strawsons subjektivistiske teori om fri vilje og ansvar som et svar på kritikken mot Peter Strawsons artikkel om ansvar. En kan si at Galen Strawson kommer disse kritikerne i møte når han utvikler en teori – en subjektivistisk teori – om ansvar. Men med sin subjektivistiske teori søker han samtidig å være i samsvar med hovedtrekk i Peter Strawsons syn: [Som mennesker] «får vi et generelt rammeverk om holdninger gjennom deltakelse i samfunnet. Det er derfor ikke behov for - og det kan heller ikke gis tillatelse til - noen ekstern rasjonell begrunnelse eller rettferdiggjørelse» [av disse holdningene] (Strawson, P.: 210). Galen Strawsons teori *kan* sees som en teori som ikke gir noen *ekstern* begrunnelse, men – fordi den er subjektivistisk – gir en *intern* begrunnelse av de reaktive holdningene.

Fischer og Ravizzas subjektivism er av en annen karakter enn Galen Strawsons. Fischer og Ravizza avviser ikke objektive elementer ved moralsk ansvar ("Of course, this subjectivity does not imply that there are not important 'objective' elements of our accounts of taking responsibility and moral responsibility" – F&R 1998: 223). Det er derfor tvilsomt

om Fischer og Ravizzas teori om moralsk ansvar bør omtales som en subjektivistisk teori. Den bør heller omtales som en teori med et subjektivistisk element.

10.2 Vurdering

Det blir ikke her foretatt noen drøfting av Galen Strawsons subjektivistiske teori, men Fischer og Ravizzas tilpasning til denne teorien blir vurdert. Det blir også stilt spørsmål om man kan velge andre tilnærminger til moralsk ansvar enn F&Rs subjektivistiske tilnærming, dersom man aksepterer Galen Strawsons avvisning av objektivistiske teorier.

Fischer og Ravizzas subjektivistiske tilnærming

Det er noen problemer med å forstå F&Rs subjektivistiske posisjon og forholdet mellom deres posisjon og Strawsons subjektivismen. Problemene er knyttet til:

- Hvordan er forholdet mellom subjektivismen og semikompatibilisme i F&Rs teori?
- Hvordan er forholdet mellom subjektivismen og F&Rs påstand om at det finnes et eksternt forsvar for reaktive holdninger?
- Hvordan er forholdet mellom Galen Strawsons subjektive syn på friheten og F&Rs ignorering av den frie vilje?
- Mener F&R at en aktør som ser seg selv som passende kandidat for reaktive holdninger, er moralsk ansvarlig – uansett?

Jeg skal utdype disse problemstillingene i det følgende: Strawson hevder at kompatibilismen representerer en objektivistisk teoriretning. F&R gir på sin side uttrykk for en semikompatibilistisk posisjon, som neppe avviker fra annen kompatibilisme med hensyn på subjektivismen / objektivisme. Betyr det at de har en objektivistisk, semikompatibilistisk grunnholdning med et visst innslag av subjektivismen? Eller dreier det seg om en subjektivismen med enkelte objektive elementer? De hevder selv at deres subjektivismen ikke innebærer at det ikke finnes viktige objektive elementer i deres beskrivelse av moralsk ansvar. Men de gir ikke noe svar på hvordan denne blandingen av subjektive og objektive elementer er.

Når det gjelder deres syn på eksternt rasjonelt forsvar for reaktive holdninger, argumenterer de i "Perspectives of Moral Responsibility" (1993) for at slikt forsvar finnes. Kan et slikt argument forenes med en subjektivistisk posisjon? Eller har de endret oppfatning, slik at de i 1993 mente at det finnes et eksternt rasjonelt forsvar for reaktive holdninger, mens

de i "Responsibility and Control" (1998) har forlatt dette synet og har inntatt en subjektivistisk posisjon? Det er vanskelig å svare på dette spørsmålet, for Fischer og Ravizza drøfter ikke den eventuelle eksistensen av eksternt rasjonelt forsvar i "Responsibility and Control".

Sentralt i Strawsons subjektivistiske filosofi står hans syn på friheten, inspirert av Kant: alle vesener som bare kan handle under frihetens ide, er virkelig frie (Strawson, G. 1986: 68). Det er friheten som er det sentrale, mens ansvar blir utledet av friheten. Men Fischer og Ravizza ignorerer den frie vilje, og heller ikke friheten har noen sentral plass hos dem. Fischer har en liknende holdning til moralsk ansvar som Galen Strawson har til frihet. Fischer gir uttrykk for en uforbeholden tro på personers moralske ansvar, uavhengig av alle vitenskapelige resultater og andre forhold som måtte kunne virke inn. Spørsmålet er så om moralsk ansvar har en noe annen karakter enn frihet, dvs. at moralsk ansvar har en mer relasjonell og sosial karakter enn frihet, og at tilnærmingen derfor blir ulik med moralsk ansvar som det umiddelbart girte enn med frihet som det umiddelbart girte.

Fischer og Ravizzas subjektivistiske tilnærming kommer til uttrykk ved å definere en aktør som ser seg selv som passende kandidat for reaktive holdninger, som moralsk ansvarlig. Men er en aktørs syn på seg selv som passende kandidat for reaktive holdninger nødvendig og tilstrekkelig betingelse for moralsk ansvar? Har aktøren moralsk ansvar hvis han mangler evnen til å vurdere seg selv? Har han moralsk ansvar, fordi han subjektivt sett oppfatter seg som passende kandidat for reaktive holdninger? Eller kan vi vise til et eksternt rasjonelt forsvar for reaktive holdninger som imøtegår hans selvbilde som moralsk ansvarlig og passende kandidat for reaktive holdninger? Hvis det kan gis et slikt eksternt rasjonelt forsvar som går imot hans selvbilde, kan vi avvise dette selvbildet og si at han *ikke* er passende kandidat for reaktive holdninger, selv om han oppfatter seg slik.

Fischer og Ravizza framhever et subjektivistisk element ved moralsk ansvar, nemlig å se seg selv som kandidat for reaktive holdninger. De utelukker ikke at deres teori også har objektive elementer, men de gjør ikke rede for forholdet mellom de subjektive og objektive elementene. De har heller ikke omtalt og drøftet om det finnes tilfeller hvor en aktør ikke er moralsk ansvarlig, selv om han ser seg selv som passende kandidat for reaktive holdninger. En kan anta at de gir plass for et eksternt syn som kan korrigere aktørens syn

på seg selv som passende kandidat for reaktive holdninger. Men Fischer og Ravizza gir ikke eksplisitt uttrykk for dette. Det er derfor ikke helt klart hvilket svar de vil gi til disse spørsmålene.

Alternativer til Fischer og Ravizzas subjektivism

I denne avhandlingen er det ikke foretatt noen drøfting av Galen Strawsons subjektivistiske tilnærming til fri vilje og moralsk ansvar og hans avvisning av objektivistiske teorier. Men det er framsatt en antakelse om at det er fellestrekk ved Galen Strawsons avvisning av objektivistiske teorier om fri vilje og ansvar og Peter Strawsons avvisning av eksternt rasjonelt forsvar for reaktive holdninger. Og det er i avhandlingen ikke funnet påvist eksternt rasjonelt forsvar for reaktive holdninger. På dette grunnlaget spør jeg hvilke alternativer man har dersom man – under opprettholdelse av F&Rs semikompatibilistiske grunnlag - aksepterer Galen Strawsons avvisning av de objektivistiske teoriene. Finnes det alternativer til Fischer og Ravizzas subjektivistiske tilnærming, dersom man holder objektivistiske teorier – med Galen Strawsons forståelse av hva objektivistisk teori er – utenfor?

Jeg antar at slike alternativer finnes. F&Rs tilnærming har tydelig karakter av intrasubjektivisme, altså en subjektivism knyttet til det enkelte subjekt. En *intersubjektiv teori* om moralsk ansvar – dvs en teori som omfatter mange subjekter og interaksjonen mellom disse - kan få en annen utforming en F&Rs teori.²⁵ Jeg antar at *sosiale teorier* om moralsk ansvar også utgjør et alternativ til objektivistiske teorier. Det kan diskuteres om sosiale og intersubjektive teorier er ett og det samme, om den ene type teori inngår som en undergruppe under den andre typen, eller om det dreier seg om ulike typer av teorier. Denne drøftingen blir ikke tatt her. Det blir ikke gjort noen vurdering av subjektivistiske teorier med intersubjektiv karakter. Men i det neste kapitlet skal jeg gi noen innspill til en sosial teori om moralsk ansvar.

²⁵ Det er mulig at Galens Strawsons teori kan beskrives som intersubjektivistisk. Strawsons teori blir imidlertid ikke ansett som et alternativ i denne sammenhengen, fordi den ikke anses å ha et semikompatibilistisk grunnlag.

Hva er forholdet mellom subjektivistisk og sosial teori om moralsk ansvar? ²⁶ Hvilke karakteristiske likheter og forskjeller er det mellom disse to teoritypene? Når det gjelder likheter, kan det trekkes fram at verken subjektivistiske teorier eller sosiale teorier forutsetter at moralsk ansvar - og de reaktive holdningene som ansvaret bygger på - kan gis et eksternt rasjonelt forsvar. I begge typer av teorier er ansvar avgrenset til den (individuelle eller sosiale) konteksten som teorien fastsetter. Og i begge tilfeller kan en *moralsk ansvarlig person* defineres som passende kandidat for reaktive holdninger. Men hvis vi antar at moralsk ansvar er å være passende kandidat for reaktive holdninger, vil den subjektivistiske teorien vektlegge at aktøren ser seg selv som passende kandidat for reaktive holdninger, mens den sosiale teorien vil vektlegge at det sosiale systemet ser aktøren som passende kandidat for reaktive holdninger.

Det vil ofte være samsvar mellom de to tilnærmingene: Når det sosiale systemet vurderer en aktør som passende kandidat for reaktive holdninger, vil ofte aktøren se seg selv på denne måten. Men i noen tilfeller ser trolig aktøren seg selv som passende kandidat for reaktive holdninger, mens andre ikke gjør det. Og det kan også forekomme at andre ser på ham som en passende kandidat for reaktive holdninger, mens han ikke ser seg selv slik.

²⁶ I omtalen av subjektivismen legges Fischer og Ravizzas intrasubjektivistiske forståelse til grunn. Det er derfor en intrasubjektivistisk versjon av subjektivismen som her blir sammenliknet med en sosial teori om moralsk ansvar.

DEL II EN ALTERNATIV TILNÆRMING TIL ANSVAR

11. En skisse

I dette kapitlet skisseres²⁷ en sosial tilnærming til moralsk ansvar. Det blir gitt en begrunnelse for valg av tilnærming, og trukket fram noen utfordringer som er knyttet til denne tilnærmingen. Det kan nevnes at en sosial tilnærming til ansvar med de forutsetningene som er lagt til grunn her, er lite utforsket etter det jeg kjenner til. Skissen kan derfor oppfattes som et forslag om å gjøre nærmere undersøkelser av denne tilnærmingen.

11.1 Grunnlaget

Noen forutsetninger

Det er noen spørsmål som det synes vanskelig eller umulig å besvare. Et slikt er spørsmålet om det finnes alternative handlingsmuligheter. Det forutsettes å være menneskelig umulig å gå tilbake i tid og undersøke om en person på et gitt tidspunkt kunne ha handlet annerledes enn han faktisk gjorde. Det er altså ikke vist at en person har alternative muligheter på et gitt tidspunkt. Men det er heller ikke vist at en person *ikke* har alternative muligheter på et gitt tidspunkt.

Et annet spørsmål som er vanskelig å besvare, omhandler eksternt rasjonelt forsvar for reaktive holdninger. Det har ikke lyktes gjennom de drøftinger som har vært gjort i denne avhandlingen, å vise at slikt eksternt rasjonelt forsvar finnes. Men det er heller ikke vist at eksternt rasjonelt forsvar for reaktive holdninger *ikke* finnes.

Når man ikke klarer å gi svar, er det viktig å undersøke de ulike svaralternativene. En bør ikke la et alternativ, eller en tilnærming, være uutforsket. Når det gjelder de to nevnte spørsmålene, finnes to prinsipielle alternativer (hovedalternativer) og en rekke tilnærminger som tilhører ett av de to hovedalternativene (samt tilnærminger som ikke så lett kan klassifiseres eller som representerer en kombinasjon av de to hovedalternativene). De to hovedalternativer er:

²⁷ Denne presentasjonen av en alternativ tilnærming til moralsk ansvar inneholder ikke et fullt gjennomført argument, og blir derfor omtalt som skisse.

- Alternativ 1. Det forutsettes at det man spør etter og ikke finner, likevel finnes.
Alternativ 2. Det forutsettes at det man spør etter og ikke finner, ikke finnes - siden man ikke har klart å påvise at det finnes.

I denne skissen bygger jeg på alternativ 2, altså at det ikke finnes alternative muligheter og heller ikke eksternt forsvar for reaktive holdninger. For meg er det et tillokkende alternativ å forutsette at det man ikke har funnet, heller ikke finnes. Jeg påstår ikke at dette alternativet er best, eller at alternativ 1 er utelukket. Men jeg hevder at alternativ 2 er minst like godt som alternativ 1.

Denne avhandlingen bygger på Fischer og Ravizzas semikompatibilistiske standpunkt, som innebærer at man i sin undersøkelse av moralsk ansvar ikke forutsetter fri vilje. Når en legger dette til grunn sammen med forutsetningene i Alternativ 2, begrenses omfanget av aktuelle tilnærminger til moralsk ansvar. I det forrige kapitlet er det trukket fram to aktuelle tilnærminger, den (intra)subjektivistiske og den sosiale tilnærmingen. Disse blir sammenliknet nedenfor.

Sosial og subjektivistisk tilnærming til moralsk ansvar

En sammenlikning av sosial og subjektivistisk tilnærming til moralsk ansvar kan ta utgangspunkt i en kommentar fra Fischer og Ravizza til Peter Strawson. De hevder at Strawson blander sammen å bli holdt ansvarlig og å være ansvarlig:

..... Strawson's theory may reasonably be said to give an account of what it is for agents to be held responsible, but there seems to be a difference between being held responsible and actually being responsible. Surely it seems possible that one can be held responsible even though one in fact is not responsible, and conversely that one can be responsible even though one is actually not treated as a responsible agent. By understanding responsibility primarily in terms of our actual practices of adopting or not adopting certain attitudes toward agents, Strawson's theory risks blurring the difference between these two issues. (Fischer and Ravizza 1993: 18)

Fischer og Ravizza bruker uttrykket *ansvarsreaksjon* som samlebegrep for reaktive holdninger og sosial praksis basert på reaktive holdninger, som bl.a. ros og klander. Å holde en aktør ansvarlig for en handling kan også sees som en form for ansvarsreaksjon. Når F&R hevder at at en aktør som blir holdt ansvarlig, ikke alltid *er* moralsk ansvarlig, innebærer det at ikke alle ansvarsreaksjoner medfører moralsk ansvar. Hva kreves så for at en ansvarsreaksjon skal medføre moralsk ansvar? I kapittel 7 blir det vist til at ansvarsreaksjonen (klanderen) må være legitim for å medføre moralsk ansvar. Det blir også hevdet at det ikke gis noe eksternt rasjonelt forsvar av holdninger og reaksjoner. Men et *internt* rasjonelt for-

svar for våre reaksjoner kan likevel ha legitimitet. Det må da forutsettes at det interne rasjonelle forsvaret av reaksjonen forankres i et tanke- eller verdisystem som befinner seg innen det samme rammeverket som forsvaret av reaksjonene befinner seg i. Tankesystemet som reaksjonen forankres i, kan være en del av aktørens tanker, men det kan ikke bare være et individuelt tankesystem. Tankesystemet må være relatert til aktørens sosiale omgivelser og være forankret i et rammeverk, som vi i det følgende vil kalle for en sosial setting. Hvis en ser på legitimitet på denne måten, kan ikke en reaksjon legitimeres ved en intrasubjektivistisk tilnærming.²⁸

11.2 En sosial tilnærming til moralsk ansvar

En sosial tilnærming til ansvar kan ta utgangspunkt i en rekke ulike forfattere. Her blir det tatt utgangspunkt i Gary Watsons forståelse av ansvar som tilregnelighet, og hvor det å holde hverandre ansvarlig involverer en sosial setting. En sosial setting inneholder sosiale systemer, strukturer, prosesser og aktiviteter. På slikt grunnlag skisseres mulige svar på noen dilemmaer, bl.a. hvordan legitimering av ansvar skjer.

To former for ansvar

I artikkelen ”Two Faces of Responsibility” (Watson 1996) beskriver Watson to måter å forstå ansvar på. I sin beskrivelse av den ene forståelsesmåten viser Watson til John Dewey: “Responsibility is ... one aspect of the identity of character and conduct. We are responsible for our conduct because that conduct is ourselves objectified in actions.” (Dewey: 160-61) Watson omtaler dette som ansvar forstått som attributability: Ansvar sees som et aspekt ved identiteten til aktøren og hans handling. Denne forståelsen av ansvar er blitt kritisert av flere forfattere. Watson uttrykker sin kritikk av ansvar som attributt slik:

Holding people responsible is not just a matter of the relation of an individual to her behavior; it also involves a social setting in which we demand (require) certain conduct from one another and respond adversely to one another’s failures to comply with these demands. (Watson 1996: 229)

²⁸ F&R vil kanskje hevde at de ikke bare vektlegger aktørens egne holdninger. De kan si at en aktør *reelt* må ta ansvar for å ha moralsk ansvar, og at det ikke er tilstrekkelig at han selv sier at han tar ansvar. Men hvis de etablerer en slik *reell* posisjon utenfor aktøren selv, står de i fare for å forlate sin subjektivistiske posisjon. Jeg finner derfor den subjektivistiske posisjonen problematisk når en skal gjøre rede for legitimiteten i bedømmelsen av moralsk ansvar.

Den andre formen for ansvar som innebærer krav og respons i forhold til hverandres atferd, omtaler Watson som ansvar som tilregnelighet (accountability). Han vektlegger denne ansvarsformen. Imidlertid understreker han også betydningen av ansvar som attributt. Begge forståelsesmåter har sin verdi. Det dreier seg derfor om å supplere – ikke å erstatte - attributtforståelsen med en forståelse av ansvar som tilregnelighet.

Sosial setting og sosialt system

Watson hevder at det å holde mennesker ansvarlig involverer en sosial setting hvor vi krever en viss type atferd fra hverandre, og responderer negativt på de andres manglende evne til å rette seg etter disse kravene. Hva er så en sosial setting? Watson har ikke gjort rede for dette. Det kan derfor være vanskelig å gi et godt svar. Jeg velger imidlertid å legge vekt på to forhold, nemlig at det kan eksistere sosiale settinger for sosiale systemer på mange nivåer – for store samfunn og for små organisasjoner,²⁹ og at sosial setting har karakter av et altomfattende rammeverk innenfor en sosial virkelighet.

En sosial setting forutsettes altså å være hele rammeverket av sosiale og sosialrelaterte strukturer, prosesser og aktiviteter, inkludert sosial praksis som er omtalt i kapittel 6.4. En sosial setting kan imidlertid omfatte mer enn de elementene av sosial praksis som er omtalt foran i avhandlingen. Den sosiale settingen omfatter ikke bare nytterelaterte handlinger, som en viss tolkning av utilitarianismen kan gi inntrykk av. Det er heller ikke tilstrekkelig å føye til uttrykk for holdninger og følelser, som en tolkning av Strawson kan gi inntrykk av. Kanskje er det heller ikke tilstrekkelig å legge til mellommenneskelig kommunikasjon gjennom dialog og kognitive prosesser gjennom internt rasjonelt forsvar for reaktive holdninger, for den sosiale settingen forutsettes å omfatte all sosial virkelighet innenfor det rammeverket som settingen setter en grense for.

Er Watsons sosiale setting det samme som Strawsons rammeverk av holdninger og følelser og sosial praksis assosiert med holdninger og følelser? Eller kan sosial setting forstås som en del (den ”sosiale komponenten”) av rammeverket av holdninger og følelser? Strawson hevder at rammeverket av holdninger og følelser danner en generell struktur eller vev, og

²⁹ Det finnes sosiale systemer på mange nivåer. Her omtales *sosial organisasjon* som samleuttrykk for gruppe, lokalsamfunn, organisasjon, nasjonal subkultur osv, og *samfunn* eller *storsamfunn* som brukes om stat, nasjon, internasjonalt samfunn osv.

at det er noe som vi er gitt gjennom den faktiske eksistensen av det menneskelige samfunn. Strawsons rammeverk har altså en sosial basis – som Watsons setting har det. Og Watsons negative respons svarer til Strawsons krenkelse, klander og moralsk fordømmelse (*resentment, blame, moral condemnation*). Og begge filosofer forutsetter at den negative responsen er svar på krav, forventninger eller behov som ikke blir tilfredsstillende møtt. Det er altså nær forbindelse mellom de to begrepene, selv om de ikke nødvendigvis er like.

I en sosial setting befinner det seg ulike sosiale systemer. Med den sosiale tilnærmingen som legges til grunn her, defineres moralsk ansvar av det sosiale systemet som en aktør befinner seg i – i tillegg til den sosiale settingen som han befinner seg i. En kan hevde at en aktør har moralsk ansvar når det sosiale systemet ser ham som passende kandidat for reaktive holdninger – under forutsetning av at systemets vurdering av og dom³⁰ om saken er legitim, dvs er relatert til anerkjente tanke-systemer i den sosiale settingen. Hvis en utelater referansen til reaktive holdninger, kan en si: En aktør har moralsk ansvar innen et sosialt system, når systemet på legitimt vis vurderer at han har moralsk ansvar.

Legitimering av ansvar

Hvordan får et sosialt systems vurdering og dom om moralsk ansvar legitim status? For å gi svar på dette, kan en stille de følgende spørsmålene:

- Hvilke kriterier er det sosiale systemets vurderinger og dommer basert på?
- Hvordan legitimerer det sosiale systemet en dom?
- Hva skjer når det er uenighet eller konflikt mellom ulike sosiale systemers dommer om moralsk ansvar?

Hvilke kriterier er det sosiale systemets vurderinger og dommer basert på? Det er grunn til å forvente at det sosiale systemets vurderinger og dommer er basert på de betingelser som er omtalt i avhandlingen, som kontroll, kunnskap, deltakelse i det moralske fellesskapet og/eller evne og vilje til å ta ansvar. Men vurderingene kan også være basert på andre kriterier som ikke er nevnt her.

³⁰ Tilskrivning av moralsk ansvar er av konkluderende art og har karakter av dom. Men fordi tilskrivning av ansvar oftest skjer uformelt, blir slik tilskrivning til dels omtalt som *vurdering* av ansvar.

Hvordan legitimerer det sosiale systemet en dom? Det forsettes at dommer om moralsk ansvar legitimeres - som det er gjort rede for med ros eller klander - ved at de er i samsvar med anerkjente tanke-systemer innen den aktuelle sosiale settingen. Det forutsettes også at slike tanke-systemer blir anerkjent blant annet gjennom oppslutning fra sosiale systemer og aktører med makt og innflytelse.

Men hva skjer hvis det sosiale systemet blir utfordret av andre systemer, enten utenfra eller innenfra? Dersom det ikke finnes eksternt rasjonelt forsvar for dom om ansvar, er dommen avhengig av det ståstedet som det dømmende sosiale systemet har, dvs. av den sosiale settingen som dette systemet befinner seg i. Hvis den sosiale settingen er det norske samfunnets opinion, slik det for eksempel kommer til uttrykk etter en offentlig debatt, framstår denne opinionens dom om moralsk ansvar som legitim. Men hvis den sosiale settingen er en subkultur eller en annen nasjon som er i opposisjon til den norske opinionen, vil ikke denne opinionens oppfatning om moralsk ansvar uten videre framstå som legitim. Her kan en igjen bruke eksemplet med Andeby: Sett fra Andebybefolkningens ståsted representerer deres oppfatning av at de tilbakestående individene er ansvarlige for sine avvik, en legitim dom om moralsk ansvar.³¹ Men hvis den norske opinionen følger Fischer og Ravizza og mener at de tilbakestående ikke er ansvarlige for sine atferdsavvik, vil Andebys dom over de tilbakeståendes moralske ansvar være illegitim sett fra den norske opinionens ståsted. Vurdering og dom om moralsk ansvar varierer altså med den sosiale settingen.

Det er gjort rede for at en aktør har moralsk ansvar innen et sosialt system, når systemet på legitimt vis vurderer (og dømmer) at han har moralsk ansvar. Og vurdering og dom legitimeres ved å ha relasjon til et anerkjent tanke-system innen den sosiale settingen hvor vurdering og dom foretas. En kan derfor si at en aktør har moralsk ansvar i en sosial setting, hvis vurderingen av (og dommen om) moralsk ansvar har legitimitet innen denne settingen.

³¹ Hvis en forutsetter at Andebys standpunkt er ekstremt, vil trolig ikke legitimiteten rekke lenger enn til Andebys grenser. Men hvis standpunktet er mindre ekstremt, kan legitimiteten rekke lenger. Hvis Andeby aksepterer at noen mentalt tilbakestående mennesker ikke er moralsk ansvarlige, men at de omtalte individene er ansvarlige fordi de er lite tilbakestående, kan deres standpunkt ha legitimitet også utenfor deres eget lokalsamfunn.

Hvilke krefter blir iverksatt for å finne løsninger hvis det ikke finnes et eksternt rasjonelt forsvar for dommene? Hvis det ikke finnes et eksternt rasjonelt forsvar, må andre krefter enn rene fornuftsargumenter virke for å avgjøre hvilken dom som skal framstå som legitim. Med en sosial tilnærming blir slike krefter omtalt som sosiale krefter. Dersom *ett* ord skulle bli valgt om de sosiale kreftene, ville *makt* bli valgt.

Hvordan blir uenighet om dommer om ansvar løst? Dersom de virkende kreftene er sosiale krefter omtalt som makt, vil makten avgjøre hvilken dom om ansvar som er legitim. Hva er så makt i denne sammenhengen? I hvilken grad er makten rasjonell, og i hvilken grad er den irrasjonell? I hvilken grad, under hvilke betingelser, og på hvilken måte gir makten ytelser til befolkningen? Og i hvilken grad, under hvilke betingelser, og på hvilken måte *tar* makten ytelser fra befolkningen? Det er umulig å foreta en sosialteoretisk maktanalyse innenfor rammen av denne avhandlingen. Jeg vil begrense meg til disse bemerkningene: Det blir lagt til grunn i avhandlingen at det finnes en rasjonalitet, selv om det skulle vise seg at det ikke finnes et eksternt rasjonelt forsvar for våre reaksjoner (og dommer). Vi kan altså forutsette at prosessen for å legitimere dommer om moralsk ansvar blir drevet av sosiale krefter som *kan* omtales som makt, og at det kan gis et internt rasjonelt forsvar for de dommene som blir avsagt. Sett fra maktens ståsted er dommene både rasjonelle og legitime. Sett fra en posisjon utenfor makten kan dommene framstå som både irrasjonelle og illegitime. Men sett fra en posisjon utenfor makten kan de også framstå som litt rasjonelle og litt irrasjonelle, og dermed med svekket legitimitet, men ikke helt uten legitimitet.

Makten er ikke alltid hos øverste formelle myndighet. Et eksempel på dette beskrives i kapittel 6 i det første scenariet om lokalsamfunnet Andeby. Her opprettholder Andeby sin ansvarsdefinisjon på tross av innblanding fra storsamfunnets representanter. Da viser Andebybefolkningen også at de har makt til å definere ansvaret i sitt samfunn.

To eksempler

Det kan nevnes et par eksempler på bruk av sosial tilnærming til ansvar. Med det første eksemplet søker jeg å vise at moralsk ansvar ikke baseres på enkeltindividers reaktive holdninger. Det andre eksemplet søker å vise at assymetri i reaksjonen på gode og onde handlinger kan gis en sosial begrunnelse.

Det første av de to eksemplene er omtalt i kapittel 6 og dreier seg om en kvinne, Anna, som knuser en vase under epileptisk anfall. Anna blir møtt med ulike reaksjoner: noen forstår at hun har anfall og fordømmer henne ikke for å knuse vassen, mens andre forstår ikke at hun er syk og har derfor reaktive holdninger fordi hun knuser vassen. Det ble stilt spørsmål om Annas moralske ansvar varierer med tilskuernes reaksjoner dersom det ikke finnes noe eksternt rasjonelt forsvar for (eller kritikk av) de reaktive holdningene som hun blir møtt med. Men med en tilnærming som vist ovenfor, har en enkeltpersons reaktive holdning bare legitimitet ved å forankres i et sosialt system med legitimitet i den aktuelle sosiale settingen. Og det er bare legitime reaktive holdninger som konstituerer moralsk ansvar. Reaktive holdninger uten legitimitet konstituerer *ikke* moralsk ansvar. Den reaktive holdningen til tilskueren som tillegger Anna moralsk ansvar for sitt epileptiske anfall, har ikke legitimitet, og konstituerer derfor ikke moralsk ansvar. En eventuell mangel på eksternt rasjonelt forsvar for reaktive holdninger endrer ikke noe ved dette. På dette grunnlaget kan en gå imot Fischer og Ravizzas oppfatning om at moralsk ansvar identifiseres med reaktive holdninger dersom det ikke kan gis noe eksternt rasjonelt forsvar for de reaktive holdningene.

Det andre eksemplet er presentert av Susan Wolf, som skiller mellom en person som kjenner seg tvunget til å gjøre en god handling (for eksempel redde et barn som holder på å drukne) og en person som kjenner seg tvunget til å gjøre en dårlig handling (Fischer and Ravizza 1998: 55-58). Blant de sistnevnte nevner hun tvangsmessige handlinger på grunn av narkotikaavhengighet, kleptomani, hypnose og deprivert barndom (*deprived childhood*). Livredderen er i stand til å handle i samsvar med det Sanne og det Gode, og er derfor moralsk ansvarlig for sin handling, i følge Wolf. En person som handler tvangsmessig på grunn av narkotikaavhengighet eller lignende er derimot ikke i stand til å handle i samsvar med det Sanne og det Gode, og er derfor ikke moralsk ansvarlig. På dette grunnlaget utvikler Wolf en assymetri-tese om at man gir livredderen ros, men man unnlater å klandre den narkotikaavhengige for sitt misbruk.

Fischer og Ravizza søker å unngå Wolfs assymetri-tese ved å overføre oppmerksomheten fra aktøren som utfører en handling til mekanismen som styrer denne handlingen. De hevder at dersom det inntreffer en ansvarsundergravende mekanisme, har aktøren ikke ansvar for sin handling (se kap. 5). I følge F&R skal man rette oppmerksomheten mot slike under-

gravende mekanismer, og ikke mot den enkelte aktør og dennes gode eller onde handlinger.

Ved en sosial tilnærming kan Wolfs og F&Rs fortolkninger bli supplert. De ulike reaksjonene overfor livredderen og den narkomane kan forklares med på atferdsregler innen det sosiale systemet som Wolf viser til. Atferdsreglene bygger på et assymetrisk syn på ros og klander, hvor ros anses å være en byggende reaksjon, mens klander kan være en nedbrytende reaksjon som i visse tilfeller bør unngås. Dette synet innebærer at det ikke er assymetri i moralsk ansvar mellom gode og onde handlinger, men det er en sosialt begrunnet assymetri i reaksjonene på disse handlingene.

11.3 *Utfordringer*

Det er betydelige utfordringer knyttet til tilnærmingen som er skissert i dette kapitlet. Blant utfordringene er å klargjøre ulike sosiale systemers ansvarsbegreper, og å klargjøre hva slags legitimitet de har til å dømme om ansvar. Det er også en utfordring å klargjøre hva som skjer når det er konflikt mellom ulike sosiale systemer eller tankesystemer om definisjonen av moralsk ansvar. Innebærer det at ”makten rår”, og at fornuften bare fungerer innenfor rammen av de premissene som makten setter?

Å svare på dette er en utfordring som ikke skal tas i denne avhandlingen.

AVSLUTNING

To sentrale elementer i den gjennomgåtte teorien til Fischer og Ravizza om moralsk ansvar er:

- det semikompatibilistiske grunnlaget
- oppfatningen om at moralsk ansvar er betinget av veiledningskontroll

I denne avhandlingen gis det støtte til disse to sentrale elementene i teorien. Men det reises likevel tvil om noen av Fischer og Ravizzas argumenter. Deres argumenter for at det finnes et eksternt rasjonelt forsvar for reaktive holdninger, er ikke helt overbevisende. Og det kan også rettes kritikk mot deres subjektivistiske tilnærming til moralsk ansvar.

Som et alternativ til deres tilnærming blir det skissert en sosial tilnærming til moralsk ansvar. Ut fra en slik tilnærming har en aktør moralsk ansvar innen et sosialt system når systemet på legitimt vis vurderer at han har moralsk ansvar. Og legitimitet gis gjennom relasjon til et anerkjent tankesystem innen den sosiale settingen som danner rammeverket for vurderingen. En aktør har derfor moralsk ansvar i en sosial setting, hvis vurderingen av (og dommen om) moralsk ansvar har legitimitet innen denne settingen.

Litteraturliste

- Alexander, J.C. 1988. "The New Theoretical Movement". I Neil J. Smelser 1988.
- Arpaly, Nomy. 2006. *Merit, Meaning, and Human Bondage. An Essay on free Will*. Princeton: Princeton University Press.
- Beaty, Michael D., red. 1990. *Christian Theism and the Problems of Philosophy*. Library of Religious Philosophy, vol 5. Notre Dame: University of Notre Dame Press
- Bok, Hilary. 1998. *Freedom and Responsibility*. Princeton: Princeton University Press.
- Buss, Sarah og Lee Overton, red. 2002. *Contours of Agency: Essays on Themes fram Harry Frankfurt*. Cambridge: MIT Press.
- Christman, John. *Autonomy in Moral and Political Philosophy*. Stanford Encyclopedia of Philosophy.
- Clarke, Randolph. 2010. "Determinism and Our Self-Conception." *My Way. Philosophy and Phenomenological Research*, 1: 242-250.
- Dewey, John. 1891. *Outlines of a Critical Theory of Ethics*. Optrykk: Hillary House, 1957.
- Fischer, John M., red. 1986. *Moral Responsibility*. London: Cornell University Press.
1987. "Responsiveness and Moral Responsibility." I Schoeman 1987: 81-106.
2006. *My Way. Essays on Moral Responsibility*. Oxford: Oxford University Press.
- Fischer, John M. og Mark Ravizza, red. 1993. *Perspectives on Moral Responsibility*. Ithaca: Cornell University Press.
1998. *Responsibility and Control. A Theory of Moral Responsibility*. New York: Cambridge University Press.
- Fischer, John M., Robert Kane, Derk Pereboom og Manuel Vargas. 2007. *Four Views on Free Will*. Malden MA: Blackwell Publishing.
- Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66/23: 829-39.
1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68: 5-20.
1994. "Autonomy, Necessity and Love." I Frankfurt. 1999: 129-141.
1999. *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.
- Habermas, Jürgen. 1992. *Faktizitat und Geltung. Beitrage zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt am Main: Surkamp. (Engelsk oversettelse: 1994. *Between Facts and Norms*. Cambridge MA: MIT Press.)
- Harman, Gilbert. 2000. *Explaining Value and other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- Henden, Edmund. 2010. "Deliberation Incompatibilism." *dialectica*. Oxford.
- Homans, George. 1961. *Social Behavior: Its Elementary Forms*. New York: Hartcourt, Brace and World.
- Kane, Robert, red. 2002. *The Oxford Handbook of Free Will*. Oxford: Oxford University Press.
- McKenna, Michael. 1998. "The Limits of Evil and the Role of Moral Address. A Defense of Strawsonian Compatibilism." *The Journal of Ethics*, 2: 123-142.
- McKenna, Michael og Paul Russell, red. 2008. *Perspectives on P.F.Strawson's "Freedom and Resentment"*. Burlington VT: Ashgate.
- Pereboom, Derk. 2001. *Living without Free Will*. Cambridge UK: Cambridge University Press.

- Raz, Joseph. 2010. "Being in the World." *Ratio*. Volume 23, Issue 4: 433-452.
- Schoeman, Ferdinand, red. 1987. *Responsibility, Character and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press.
- Slagstad, Rune, red. 1975. *Positivism, dialektikk, materialisme*. Oslo: Universitetsforlaget.
- Smelser, Neil J, red. 1988. *Handbook of Sociology*. Newbury Park: Sage Publications
- Stern, Lawrence. 2012. "Freedom, Blame and Moral Community." *The Journal of Philosophy*, 3: 72-84.
- Strawson, Galen. 1986. *Freedom and Belief*. Oxford: Clarendon Press.
- Strawson, Peter. 1962. *Freedom and Resentment*. Proceedings of the British Academy 48
Optrykk i Fischer 1986: 45-66.
- Stump, Eleonore. 1990. "Intellect, Will and the Principle of Alternate Possibilities." I
Beaty 1990: 254-285. Optrykk: Fischer and Ravizza 1993: 237-262.
- Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- Velleman, David J. 2002. "Identification and Identity." I Buss, Sarah og Lee Overton, red.
2002: 91-123.
- Watson, Gary. 1975. "Free Agency." *Journal of Philosophy*, 72: 205-220.
1987. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." I
Schoeman 1987: 256-286.
1996. "The Two Faces of Responsibility." *Philosophical Topics*. 24: 227-248.
red. 2003. *Free Will*. Oxford: Oxford University Press.
- Wolf, Susan. 1987. "Sanity and the Metaphysics of Responsibility." I Schoeman 1987: 46-
62.
1990. *Freedom within Reason*. New York: Oxford University Press.