

“Kunne jeg ha handlet annerledes?”

Betydninger av kontrafaktisk tenkning i en
pedagogisk kontekst

Kari M. Manum



Masteroppgave i pedagogikk, allmenn studieretning
Det utdanningsvitenskapelige fakultet
Pedagogisk forskningsinstitutt

UNIVERSITETET I OSLO

November 2016

“Kunne jeg ha handlet annerledes?”

Betydninger av kontrafaktisk tenkning i en pedagogisk kontekst

© Kari M. Manum

2016

“Kunne jeg ha handlet annerledes?”

Kari M. Manum

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

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

TITTEL:

“KUNNE JEG HA HANDLET ANNERLEDES?”

Betydninger av kontrafaktisk tenkning i en pedagogisk kontekst

AV:

Kari M. Manum

EKSAMEN:

Masteroppgave i pedagogikk,
allmenn studieretning

SEMESTER:

Høst 2016

STIKKORD:

Pedagogisk filosofi, kontrafaktisk tenkning, fri vilje, determinisme, indeterminisme, kausalitet, modalitet, “hindsight bias”, handlingsbegrepet, sammenligning, retrospeksjon, prospeksjon, aktørskap

Sammendrag

I det daglige antar vi gjerne en form for frihet i den betydningen at vi i en del tilfeller hadde handlingsalternativer tilgjengelige, og at vi besitter en kapasitet til å kunne ha valgt annerledes enn vi faktisk gjorde. Den mentale konstruksjonen av hypotetiske alternativer til faktiske hendelser sammenfattes innenfor en rekke fagområder som kontrafaktisk tenkning eller resonnering, dvs. noe som står i motsetning til det faktiske. Kontrafaktiske betingelser og utsagn gjennomsyrrer språket vårt, og kontrafaktisk tenkning regnes for å være en essensiell egenskap ved det mentale landskapet og intelligensen vår for å gi mening til verden rundt oss, og for å gjøre bedømmelser og ta beslutninger. Kontrafaktisk tenkning utgjør dermed et sentralt element i læringsprosesser. Ruth M. J. Byrne (2005) hevder i tilknytning til kontrafaktisk tenkning at rasjonaliteten og forestillingsevnen har felles utgangspunkt i det kognitive. Det å være rasjonell innebærer å kunne forestille seg, og det å forestille seg innebærer å være rasjonell. Med tanke på hvilket samspill det synes å være mellom tanker om det faktiske og det kontrafaktiske, synes det på sin plass å hevde at kontrafaktiske tanker og forestillinger om virkeligheten ikke bare er avgjørende for *hvordan* - men også for *at* vi i det hele oppfatter og forstår oss selv, andre og verden. Til og med framstår det som umulig å si noe fornuftig om betydningen av kontrafaktisk tenkning uten å benytte kontrafaktiske resonnementer og påstander for å gjøre nettopp dette. Hvis vi spør oss: "Hvordan ville vi ha opplevd virkeligheten uten å tenke kontrafaktisk om den?", er det i seg selv en kontrafaktisk tanke. Det er for eksempel dette Douglas Richard Hofstadter kaller for "the subjunctive instant replay" (Hofstadter, 1979, s. 632, 640).

Begrepet kontrafaktisk stammer fra filosofiske tekster, der den logiske statusen ved mulighet og resonnementer rundt sannsynlighet har blitt drøftet grundig. Ettersom kontrafaktiske betingelser ikke kan undersøkes eller bekreftes empirisk, har de skapt mye hodebry for filosofene. Dermed er det ikke sagt at de kontrafaktiske betingelsene ikke også skaper hodebry for oss utenfor det filosofiske rommet. Deltagere i sosiale praksiser behersker kontrafaktisk tenkning, uten at de nødvendigvis har et bevisst forhold til implikasjonene og konsekvensene av virksomheten. Men selv om vi med letthet tenker og argumenterer med kontrafaktiske betingelser, karakteriseres den sosiale diskursen samtidig av pågående konflikter og forhandlinger om hva som egentlig skjedde, hva som kunne ha skjedd og hva som burde ha skjedd.

Det handler om individers subjektive og intersubjektive virkelighetsforståelse av hva som *er*, *bør være* og *kan være* i den normative sosiale konteksten. De underliggende metafysiske oppfatningene om kausalitet, fri vilje og ansvarlighet er vanligvis ikke under debatt. Denne type antagelser spiller en sentral rolle for hvordan man tenker om menneskets muligheter, tilbøyeligheter og begrensninger, som igjen er sentrale utgangspunkt i oppdragelsen og opplæringen.

I de senere årene har den kontrafaktiske tenkningen og dens funksjoner fått betydelig oppmerksomhet innen pedagogikkens hjelpedisipliner som for eksempel filosofi, historie, sosiologi og psykologi. Innen pedagogisk teori er temaet derimot utforsket. Med et bredt anlagt perspektiv gjør jeg en teoretisk analyse av den kontrafaktisk tenkningens natur og betydninger situert i en pedagogisk kontekst. Fenomenet med kontrafaktisk griper dypt inn i virkelighetsforståelsen vår og er et uunnværlig analytisk verktøy ved legitimering av handlinger. Et mål med oppgaven er derfor å vise at temaet fortjener en større plass i den pedagogiske bevisstheten.

Innholdsfortegnelse

1	Innledning	1
1.1	Problemstilling og struktur	2
1.2	Avgrensninger	6
2	Kontrafaktisk tenkning	7
2.1	De temporale dimensjonene ved kontrafaktisk tenkning	9
2.2	Kontrafaktisk tenkning innen ulike fagfelt og samfunnsområder	11
2.2.1	Hverdagslivet	14
2.2.2	Forskning generelt	16
2.2.3	Psykologien	16
2.2.4	Filosofien	18
2.2.5	Historiefaget	23
2.2.6	Pedagogikkfaget og utdanningsfeltet	24
3	Kontrafaktiske påstander	27
3.1	Modalitet og modal logikk	27
3.2	Kontrafaktisk tenkning – modal tenkning	28
4	De metafysiske rammene	29
4.1	Kontrafaktisk tenkning og metafysiske antagelser	29
4.2	Tesen om determinismen	31
4.3	Indeterminismen	33
4.4	Problemet med determinismen	37
4.5	Ekstern kompatibilisme og pedagogisk filosofi	40
5	Frihet	46
5.1	Frihetens ontologiske dimensjoner	46
5.1.1	Et moderne frihetsideal – eksempel fra skolen	52
5.2	Autonomi	54
5.2.1	Frihet versus autonomi	55
5.2.2	Moralsk, politisk og personlig autonomi	56
5.2.3	Autonomi som abstrakt ideal	57
5.2.4	Fornuftens allmakt	59
5.2.5	Skillet mellom å være selvstyrt og styrt utenfra	60
5.2.6	Autonomihemmende versus autonomifremmende	61

5.2.7	Oppdragelse og utdanning som iscenesettelse	62
5.2.8	Frihet og autonomi i den generelle delen av læreplanen	65
5.2.9	Autonomi og grunner for handling	66
6	Handlingsbegrepet	74
6.1	Begrepsavklaring	75
6.2	Handlingsbegrepet innenfor ulike disipliner	82
6.3	Handling og aktør	83
6.3.1	Hvem kan handle?	87
6.4	Handlinger versus hendelser	88
6.4.1	Eksempler fra skolen	89
6.5	Handlingsforklaringer	92
7	Kontrafaktisk livsverden	95
7.1	Det ”utematisk fungerande”	95
7.2	Instrumentelle og utopiske kontrafaktisiteter	100
7.3	Kontrafaktiske følelser og sosial kontroll	102
8	Avslutning	105
	Litteraturliste	108

1 Innledning

I det daglige antar vi gjerne en form for frihet i den betydningen at vi i en del tilfeller hadde handlingsalternativer tilgjengelige, og at vi besitter en kapasitet til å kunne ha valgt annerledes enn vi faktisk gjorde. Den mentale konstruksjonen av hypotetiske alternativer til faktiske hendelser sammenfattes innenfor en rekke fagområder som kontrafaktisk tenkning eller resonnering, dvs. noe som står i motsetning til det faktiske. Kontrafaktiske betingelser og utsagn gjennomsyrrer språket vårt, og kontrafaktisk tenkning regnes for å være en essensiell egenskap ved det mentale landskapet og intelligensen vår for å gi mening til verden rundt oss, og for å gjøre bedømmelser og ta beslutninger. Kontrafaktisk tenkning utgjør dermed et sentralt element i læringsprosesser. Ruth M. J. Byrne (2005) hevder i tilknytning til kontrafaktisk tenkning at rasjonaliteten og forestillingsevnen har felles utgangspunkt i det kognitive. Det å være rasjonell innebærer å kunne forestille seg, og det å forestille seg innebærer å være rasjonell. Med tanke på hvilket samspill det synes å være mellom tanker om det faktiske og det kontrafaktiske, synes det på sin plass å hevde at kontrafaktiske tanker og forestillinger om virkeligheten ikke bare er avgjørende for *hvordan* - men også for *at* vi i det hele oppfatter og forstår oss selv, andre og verden. Det framstår til og med som umulig å si noe fornuftig om betydningen av kontrafaktisk tenkning uten å benytte kontrafaktiske resonnementer og påstander for å gjøre nettopp dette. Hvis vi spør oss: "Hvordan ville vi ha opplevd virkeligheten uten å tenke kontrafaktisk om den?", er det i seg selv en kontrafaktisk tanke. Det er for eksempel dette Douglas Richard Hofstadter kaller for "the subjunctive instant replay" (Hofstadter, 1979, s. 632, 640).

Begrepet kontrafaktisk stammer fra filosofiske tekster, der den logiske statusen ved mulighet og resonnementer rundt sannsynlighet har blitt drøftet grundig. Ettersom kontrafaktiske betingelser ikke kan undersøkes eller bekreftes empirisk, har de skapt mye hodebry for filosofene. Dermed er det ikke sagt at de kontrafaktiske betingelsene ikke også skaper hodebry for oss utenfor det filosofiske rommet. Deltagere i sosiale praksiser behersker kontrafaktisk tenkning, uten at de nødvendigvis har et bevisst forhold til implikasjonene og konsekvensene av virksomheten. Men selv om vi med letthet tenker og argumenterer med kontrafaktiske betingelser, karakteriseres den sosiale diskursen samtidig av pågående konflikter og forhandlinger om hva som egentlig skjedde, hva som kunne ha skjedd og hva som burde ha skjedd.

Det handler om individers subjektive og intersubjektive virkelighetsforståelse av hva som *er*, *bør være* og *kan være* i den normative sosiale konteksten. De underliggende metafysiske oppfatningene om kausalitet, fri vilje og ansvarlighet er vanligvis ikke under debatt. Denne type antagelser spiller en sentral rolle for hvordan man tenker om menneskets muligheter, tilbøyeligheter og begrensninger, som igjen er sentrale utgangspunkt i oppdragelsen og opplæringen.

I de senere årene har den kontrafaktiske tenkningen og dens funksjoner fått betydelig oppmerksomhet innen pedagogikkens hjelpedisipliner som for eksempel filosofi, historie, sosiologi og psykologi. Innen pedagogisk teori er temaet derimot utforsket. Med et bredt anlagt perspektiv gjør jeg en teoretisk analyse av den kontrafaktiske tenkningens natur og betydninger situert i en pedagogisk kontekst. Hovedvekten ligger på litteratur. Fenomenet med kontrafaktisk tenkning griper dypt inn i virkelighetsforståelsen vår, og er et uunnværlig analytisk verktøy ved legitimering av handlinger. Et mål med oppgaven er derfor å vise at temaet fortjener en større plass i den pedagogiske bevissthet.

1.1 Problemstilling og struktur

“Kunne jeg ha handlet annerledes?”

Tilsynelatende framstår dette som et enkelt spørsmål, men ved nærmere undersøkelser er det mer komplisert enn først antatt. Hva er domenet for en slik undersøkelse? Hva er involvert i et svar på spørsmålet, og hva er underliggende når vi stiller et slikt spørsmål?

Oppgavens skrives innenfor området pedagogisk filosofi, og spørsmålet tjener som inngang til en teoretisk analyse av kontrafaktisk tenkning i den pedagogiske konteksten. Hva må til for at vi skulle kunne ha handlet¹ annerledes? Den generelle, og mer eller mindre reflekterte virkelighetsforståelsen knyttet til handingslivet, impliserer sentrale ontologiske spørsmål om virkelighetens beskaffenhet. Det innbefatter de metafysiske posisjonene *determinisme* og *indeterminisme*, og herunder de moralfilosofiske begrepene *fri vilje* og *ansvarlighet*. I lys av

¹ Begrepet ”handling” benyttes snarere i formelle og teoretiske sammenhenger enn i dagliglivet. Donald Davidson bemerker dette: “The word ‘action’ does not very often occur in ordinary speech, and when it does it is usually reserved for fairly portentous occasions” (Davidson, 2001, s. 5). Men ettersom begrepet har tilstrekkelig resonans i *common sense* har jeg valgt å benytte det.

de metafysiske rammene, undersøker jeg betydninger av kontrafaktisk tenkning med hensyn til autonomi- og frihetsproblematikk.

Kontrafaktiske påstander kan ikke underlegges empiriske undersøkelser, og er allerede i seg selv en type undersøkelser av hypoteser vi har om virkeligheten, enten det skjer i dagliglivet eller i vitenskapelig forskning. Det å resonnerer kontrafaktisk kan derfor også kalles for *modal tenkning*. Ved å henvende oss innover i oss selv - inn i vårt eget sinn, forsøker vi å finne ut av universets ytterste grenser i form av hva som er mulig, umulig og nødvendig. Men disse tankeeksperimentene forutsetter i seg selv metafysiske antagelser om virkelighetens beskaffenhet, som fri vilje, kausalitet og determinisme/indeterminisme. Jon Elster hevder for eksempel: "Counterfactual propositions are law-users and not law-confirmers. If, on the basis of a certain generalization, a counterfactual proposition is asserted, this does not constitute further proof of the generalization" (Elster, 1978, s. 191).

Ettersom den kontrafaktiske tenkningen, når den inkluderer menneskelige aktiviteter, baserer seg på antagelsen om at en situasjon eller hendelse kunne ha resultert i flere mulige utfall, innebærer det samtidig et visst handlingsbegrep. Jeg hevder dermed at aktørskap manifesterer seg blant annet gjennom kontrafaktisk tenkning, ettersom kontrafaktiske refleksjonsprosesser står sentralt i hypotesedannelser om årsakssammenhenger og aktørskap.

Mot slutten av oppgaven tar jeg meg friheten til å utvide Richard Peters (1966) klassiske dannelsbegrep om *innvielse*, fra "utdanning som innvielse i en rasjonell livsverden" til å implisere "utdanning som innvielse i en kontrafaktisk livsverden".

I tillegg til at drøftingen og argumentasjonen skjer lys av de metafysiske rammene, lar jeg den tematiske analysen skje innenfor konteksten av en påstand jeg presenterer i oppgavens begynnelse. Den handler om at utdanningsfeltet og pedagogikkfaget stiller i særklasse sammenlignet med andre fagdisipliner og samfunnsområder, i det å være et *pre-kontrafaktisk* prosjekt som normativ sosial praksis. Med dette mener jeg at skolens hovedanliggende er å realisere mål og verdier, som kan kalles for ideelle tilstander og størrelser, og som dermed er *pre-kontrafaktiske* i sin natur. Med pre-kontrafaktisk (eller før-kontrafaktisk) menes at det skjer en sammenligning mellom det som forventes å skje og alternative scenarioer. I dette tilfellet handler det om ønskede alternative scenarioer. Det dreier seg om å stabilisere, justere eller forandre ut fra en ønsket fremtidstilstand. At skolen er et pre-kontrafaktisk prosjekt er ikke nødvendigvis oppsiktsvekkende i seg selv, i og med det samfunnsmandatet skolen har og

har hatt lenge. Det som derimot er oppsiktsvekkende, er at det ikke er mer oppmerksomhet rundt implikasjonene for de unges utvikling, læring og identitetskonstruksjon, å bli kontinuerlig vurdert og dermed vurdere seg selv opp mot en idealtilstand. En del av de utopiske målene er ikke ment å realiseres, men fungerer som en slags rettesnor for opplæringen, deriblant målene om frihet, autonomi, ansvar, og likhet. Disse målene sammen med mer trivielle, men like fullt nåtidens ideelle kompetansemål, utgjør en del av det normative spenningsfeltet skolens ulike aktører befinner seg innenfor: Spenningen mellom det som *er* og det som *bør* være. Jo mer disse aktørenes innsats og prestasjoner vurderes innenfor skolesystemet, desto mer engasjerer de seg i visse typer kontrafaktisk tenkning gjennom retrospeksjon. I denne *pre-kontrafaktiske* konteksten innvies elevene i en *kontrafaktisk livsverden*. Samtidig med skiftet i den temporale dimensjonen, forvandles skolens betingelser i indikativ² form før eller senere til elevens betingelser i konjunktiv³ form: Det som *kan* og *bør* skje, erstattes med det som *kunne ha* og *burde ha* skjedd. Det er denne siste formen for kontrafaktisk tenkning som er oppgavens hovedfokus, men den står i et gjensidighetsforhold til den første, og må forstås i lys av denne.

Den pedagogiske relevansen ligger i at oppgavens problemstilling og tematikk retter seg mot grunnleggende metafysiske antagelser knyttet til handlingslivet. Tematikken går dermed inn til kjernen av flere pedagogiske grunnlagsproblem, som autonomi, virkelighetsoppfatning, identitetskonstruksjon og handlingslegitimering for å nevne noen. I innledningen til artikkelsamlingen *Humaniorastudier i pedagogikk* uttrykker Kvernbekk følgende om hva pedagogisk filosofi kan bidra med: “Pedagogisk filosofi kan analysere underliggende antagelser, påpeke usagt og utenkte implikasjoner, snu tilvante forestillinger på hodet og problematisere vedtatte sannheter, og slik bidra til kritisk refleksjon og bevisstgjøring” (Kvernbekk, 2011, s. 11). En fundamental virksomhet i handlingslivet vårt er å trekke slutninger og legitimere handlinger på bakgrunn av oppfatninger, både de vi er bevisst og mer underliggende ontologiske oppfatninger om virkelighetens. Det å undersøke fenomenet med kontrafaktisk tenkning i den pedagogiske konteksten, kan bidra til å snu tilvante forestillinger på hodet og problematisere vedtatte sannheter. Temaet er til stor del ignorert i pedagogisk teori og praksis, og er dermed ikke typisk for den pedagogisk-filosofiske diskursen. Lars Løvlie sier: ”Dagens pedagogiske filosofi lar seg ikke avgrense, men består snarere av en

² Indikativ: ”...modus som uttrykker at en ytring skal tolkes som et konstaterende eller fortellende utsagn...” (Bokmålsordboka, 2016)

³ Konjunktiv: ”...verbalkategori, modus som uttrykker at et utsagn er tenkt, uvirkelig, mulig eller ønskelig” (SNL, 2016).

rekke temaer, problemer og stier som utforskes i kritisk dialog” (Løvlie, 2003, s. 82). I oppgaven undersøker jeg et fenomen og en problemstilling som er underliggende og grunnleggende i all menneskelig virksomhet, og dermed også for oppdragelse og undervisning. Samtidig er det en undersøkelse av om “stien” via det *kontrafaktiske landskapet* kan representere en fruktbar åpning inn det pedagogisk-filosofiske feltet, og om den til og med kan utgjøre et bidrag til den analytiske verktøykassen for den virksomme pedagogen; et bidrag til kritisk refleksjon og innsikt.

Oppgaven er strukturert på følgende måte: I kapittel 2 redegjør jeg for ulike aspekter ved kontrafaktisk tenkning, også innenfor pedagogikkens hjelpevitenskaper som filosofi, historie og psykologi. Jeg presenterer to tabeller som gir en oversikt over de temporale dimensjonene ved kontrafaktisk tenkning, og kontrafaktisk tenkning innen ulike vitenskapelige disipliner og samfunnsområder. Informasjonen i tabellene er hentet fra vitenskapelig litteratur knyttet til de forskjellige fagområdene. I dette kapittelet presenteres også oppfatningen om skolen som en pre-kontrafaktisk sosial normativ praksis.

I Kapittel 3 redegjør jeg for kontrafaktiske påstanders natur, og herunder om modalitet og modal/formal logikk.

I kapittel 4 tar jeg for meg de metafysiske rammene determinisme og indeterminisme, og ser også på hva som skiller kausalitetsbegrepet fra determinisme. Videre redegjør jeg for de filosofiske posisjonene kompatibilisme og inkompatibilisme. Jeg sammenligner Peter van Inwagens (2002) argument for tesen om determinismen med Baruch Fischhoff (1974) sin redegjørelse for tendensen til ”snikende determinisme”. Jeg avslutter med å vise til hvordan et område av metafysikkens fri vilje-problem overlapper med noen av de mest sentrale spørsmålene innen pedagogisk filosofi.

Kapittel 5 er en oppfølger til kapittel 4, der jeg tar for meg frihetens ontologiske dimensjoner ved hjelp av teoriene til Immanuel Kant, Jean-Paul Sartre og Maurice Merleau-Ponty. I forlengelsen av frihetsbegrepet redegjør jeg for autonomibegrepet og drøfter sentraliteten av kontrafaktisk tenkning for forståelsen av og idealet om autonomi.

I kapittel 6 redegjør jeg for handlingsbegrepet. Videre ser jeg nærmere på noen sentrale filosofiske tilnærminger til handling og aktørskap, for å se om det kan kaste lys over hva som skjer når vi tenker kontrafaktisk om oss selv og handlingene våre. Jeg argumenterer for

påstanden om at aktørskap manifesterer seg blant annet gjennom kontrafaktisk tenkning. Med bakgrunn i dette, argumenterer jeg for at forventningene og kravene til elevene i dagens målstyrte skole, framstår som en form for disiplinering som er i strid med idealene om autonomi. Til slutt ser jeg på hva handlingsfilosofien sier om handlingsforklaringer. Så vendes oppmerksomheten mer mot hva intensjonal handling kan sies å være.

I kapittel 7 utvider jeg den klassiske definisjonen av utdanning som ”innvielse i en rasjonell livsverden”, til å implisere utdanning som ”innvielse i en *kontrafaktisk livsverden*”.

Hovedargumentet for dette er at den kontrafaktisk tenkningen utgjør en grunnleggende bestanddel ved rasjonaliteten vår, med de implikasjonene det måtte ha. Jeg belyser videre maktperspektiver ved kontrafaktisk tenkning i skolen, gjennom begrepsparet *instrumentelle* og *utopiske* kontrafaktisiteter. Videre skisserer jeg en hypotese om at skolens kontrollordninger og vurderingssystemer innebærer en sterk sosial kontroll, i form av å fungere skamproduserende.

Kapittel 8 er en oppsummerende avslutning av det jeg har funnet og drøftet i undersøkelsen.

1.2 Avgrensninger

Kontrafaktisk tenkning er et komplekst fenomen med forgreininger til en rekke epistemologiske og ontologiske temaer. Innenfor rammen av en masteroppgave, er det ikke mulig å berøre alle spørsmålene som kan stilles i tilknytning til tematikken. Jeg tar derfor sikte på å redegjøre for noen av de mest sentrale trekkene eller fenomenene som enten forutsetter eller forutsettes av kontrafaktisk tenkning. Gjennom et bredt kildetilfang oppnår jeg større variasjon i å belyse problemstillingen, samtidig som dette kan gå på bekostning av begrepsmessig stringens. Jeg opererer med en pluralistisk forståelse av begrepet kontrafaktisk, og binder meg således ikke til en bestemt fortolkning av begrepet.

2 Kontrafaktisk tenkning

I dette kapittelet skal jeg sirkle inn begrepet *kontrafaktisitet*, og se på utvalgte dimensjoner ved den kontrafaktiske tenkningen som jeg vurderer er vesentlige for å belyse problemstillingen. Jeg presenterer først to tabeller som skal anvendes i argumentasjonen. Tabell 1.1 er en oversikt over de temporale dimensjonene ved kontrafaktisk tenkning, og tabell 1.2 er en oversikt over kontrafaktisk tenkning innen ulike vitenskapelige disipliner og samfunnsområder. Informasjonen i tabellene er hentet fra vitenskapelig litteratur knyttet til de forskjellige fagområdene.

Tabell 1.1 Oversikt over de temporale dimensjonene ved kontrafaktisk tenkning (KT)

Type KT	Post-kontrafaktisk (kontrafaktisk)	Pre-kontrafaktisk	Pre-faktisk
Data	Historiske "fakta"	Historiske data + antesiperte sosiale "fakta" + hypotetiske scenarioer	Historiske data + hypotetiske scenarioer
Metode	Retrospeksjon: Sammenligner faktiske utfall med alternative utfall	Prospeksjon: Sammenligner forventede scenarioer (antesiperte sosiale "fakta") med alternative scenarioer	Prospeksjon: Sammenligner alternative scenarioer
Bias	Hindsight bias: "snikende determinisme"	a) Hindsight bias: "snikende determinisme" b) Skråsikkerhet/overdreven pessimisme/overforenklet syn på framtiden "Nirvana-feilslutningen" (Demsetz 1969)	

Tabell 1. Oversikt over de temporale dimensjonene ved kontrafaktisk tenkning.

Tabell 1.2 Oversikt over kontrafaktisk tenkning innen ulike vitenskapelige disipliner og samfunnsområder.

Tabell 1.2 Oversikt over kontrafaktisk tenkning (KT) innen ulike vitenskapelige disipliner og samfunnsområder

Type kontrafaktisk tenkning	Common sense		Vitenskap og/eller offentlig institusjon, samfunnsområder			
	Hverdagslivet	Filosofi	Historie	Psykologi; kognitiv og sosialpsykologi	Rettssystemet	Pedagogikk/utdannings- vitenskap/utdanningspolitikk skole/utdanning
a) Ureflekterte, ubevisste	X			X	X	X
b) Metoderedskap; analytisk rammeverk		X	X	X	X	X
c) Metadebatt om kontrafaktiske påstander og betingelser		X	X	X	X	
d) Drøfting av naturen til kontrafaktisiteter		X	X			

	Hverdagslivet	Filosofi	Historie	Psykologi; kognitiv og sosialpsykologi	Rettssystemet	Pedagogikk/utdannings- vitenskap/utdanningspolitikk skole/utdanning
1) Fokus på bias: hindsight" og "snikende determinisme"			X	X		
2) (Post-)kontrafaktisk prosjekt som normativ, sosial praksis					X	X
3) (Pre-)kontrafaktisk prosjekt som normativ sosial praksis						X
4) Fokusområder og spørsmål	Hypotetisk tenkning, intuitive kausale analyser på formen "hvis, så", "hvis bare"	Kausalitet, fri vilje- determinisme, autonomi, etikk og moralfilosofi, vitenskapsteori, mulig verden- semantikken	Kausalanalyser: Kontrafaktisk historieskriving	Kausalanalyser. KT som kognitivt fenomen; kausalitet, emosjoner, persepsjoner, motivasjoner, mål, evalueringer	Kausale analyser; attribusjon av ansvarlighet, "but for"- analyser	Kausalanalyser. Intervensjoner, tiltak Adferdsmodifikasjon Evalueringer: Alternative historier (Taleb 2004, 2010)
5) Temporalitet ved kausalanalyser knyttet til aktørskap og ansvarlighet	Fortid, fremtid	Fortid, fremtid	Fortid. (Bedømmer handlinger over lengre perioder/helt liv)	Fortid, fremtid	Fortid. (Bedømmer skadelige utfall av enkelte, bestemte handlinger)	Fortid, fremtid

Forklaring tabell 1.2
X = større grad av
x = mindre grad av

2.1 De temporale dimensjonene ved kontrafaktisk tenkning

I tabell 1.1 om de temporale dimensjonene ved kontrafaktisk tenkning, har jeg tatt utgangspunkt i at vi sammenligner scenarioer eller hendelser, og at dette gjøres både retrospektivt og prospektivt. Innen de fleste fagområder er man mest opptatt av den kontrafaktiske tenkningen som handler om fortidige hendelser, og det er også denne typen jeg drøfter mest. Men jeg har allikevel valgt å inkludere det prospektive aspektet ved kontrafaktisk tenkning, fordi jeg mener det er en sentral kategori i sammenheng med utdanning og pedagogikk.

Det *prospektive* er for det første et viktig bakgrunnstappe for retrospeksjon og kontrafaktisk tenkning. Tolkningen av et visst utfall gjennom retrospeksjon skjer gjerne i lys av det prospektive og framtidens antatte mulighetsrom, der visjoner, idealer, forhåpninger, ønsker og krav til disse spiller en viktig rolle. Våre underliggende prospeksjoner synes vesentlige for de normative implikasjonene; hvordan den kommende virkeligheten optimalt bør bli og hva som bør skje for at dette skal komme i stand. Men våre prospeksjoner opptrer ikke bare som bakgrunnstappe for retrospeksjon og kontrafaktisk tenkning. I visse teoretiske sammenhenger drøftes det også hvordan vi resonnerer kontrafaktisk om framtidige hendelser (Hoch & Roediger, 1985; Svetlova, 2009). Noen teoretikere innenfor psykologisk forskning har gitt denne type ”kontrafaktisk” tenkning egne begreper, som det å tenke *pre-kontrafaktisk* (Kofta, Weary, & Sedek, 1998; Markman & Miller, 2006) eller *pre-faktisk* (Epstude, Scholl, & Roese, 2016). Noen opererer i denne sammenhengen også med *post-kontrafaktisk* (Markman & Miller, 2006) for å skille mellom de ulike kontrafaktisitetene. Noen vil kanskje hevde at det er å tøyse begrepet litt vel langt, da vi ikke kjenner framtidens faktisiteter eller aktualiteter. Vi kan bare ane og etter beste evne forsøke å predikere framtidige scenarioer. Ved retrospeksjon kjenner vi utfallet, og det er om dette vi tenker kontrafaktisk. Dette innebærer ikke at de fortidige faktaene er uten fortolkning eller at de ikke er teoriladet i Norwood Russell Hanson sin (1958) betydning. Men sammenlignet med framtidige hendelser, stiller de allikevel i en annen kategori angående det vi vanligvis kaller fakta.

Tenkningen vår omfavner et ubegrenset tidsspenn hva gjelder fortid, nåtid og framtid, og det er en innebygget asymmetri mellom de temporale dimensjonene. Litt forenklet så oppfatter vi framtiden som åpen, mens fortiden er fastsatt. Med retrospeksjon som metode reflekterer vi

over fortidige hendelser, mens overveielsene våre er forbeholdt framtiden. Å reflektere over fortiden involverer også å reflektere over om vi kunne ha overveid annerledes enn vi gjorde, eller om vi kunne ha handlet annerledes på bakgrunn av de overveielsene vi gjorde. Dette poengterte Aristoteles slik: "...for neither does one deliberate about what has happened in the past, one deliberates about what still lies in the future and may happen or not; what has happened cannot be made not to have happened..." (Aristotle, 1926, VI, ii, 6, 1139b). I dette perspektivet oppfatter vi det som at tiden beveger seg fra framtid til fortid, og fra det som er mulig til det som er nødvendig.

Verken *pre-kontrafaktisk* eller *pre-faktisk* fungerer derfor som presise beskrivelser eller begreper, da *pre-kontrafaktisk* feilaktig antar at framtiden inneholder faktisiteter, mens *pre-faktisk* ikke fanger opp poenget med at scenarioer sammenlignes. På tross av denne upresisheten velger jeg å benytte begrepet *pre-kontrafaktisk*, ettersom det allerede er i bruk i noe litteratur. Ved *pre-kontrafaktisk* tenkning sammenligner vi antaserte sosiale fakta, eller det vi forventer skal skje, med ulike hypotetiske scenarioer. Et eksempel er frafallet i videregående skole. Det har for eksempel vært et forventet scenario at ungdom fortsetter å droppe ut av videregående skole, hvis man ikke foretar seg noe. Ved *pre-kontrafaktisk* tenkning sammenlignes den forventede situasjonen og de hypotetiske årsakene til den, med alternative scenarioer. I tabell 1.1 har jeg for ordens skyld tatt med kategorien og begrepet *pre-faktisk*, fordi det har blitt anvendt i en ganske nylig studie (Epstude et al., 2016). Jeg har imidlertid gitt det en egen definisjon, som sier at man sammenligner noenlunde likeverdige framtidige scenarioer.

I oppgaven har jeg valgt å inkludere tenkning om framtidige scenarioer, fordi jeg mener at utdanningsfeltet og pedagogikkfaget stiller i særklasse sammenlignet med andre institusjonelle områder og fagdisipliner, i det å være et *pre-kontrafaktisk* prosjekt og en *normativ sosial praksis*. Det vil si at skolens hovedoppgave er å realisere mål og verdier, som i sin natur kan kalles for ideelle tilstander og størrelser. Janicke Heldal Stray påpeker hvordan skolens mål er formulert både i et nytte- og dannelsesperspektiv, der siktepunktet er "elevenes fremtidige funksjonalitet i arbeidslivet og som borgere" (Stray, 2010, s. 6). Hva som skjer i praksis er et annet spørsmål jeg ikke går videre inn på her. Men uansett hvem eller hva som bestemmer premissene og rammene for pedagogisk praksis; om det er enhetsskoleideologien eller målstyringsideologien som ligger bak den rådende utdanningspolitikken, så handler det om å stabilisere, justere eller forandre ut fra en ønsket fremtidstilstand. Samtidig er det klart at

styring og planlegging av skole handler om å analysere det fortidige, både det uforutsette og ulike igangsatte tiltak og intervensjoner, samt å vurdere utfallet av disse.

2.2 Kontrafaktisk tenkning innen ulike fagfelt og samfunnsområder

Kontrafaktisk tenkning er ikke et nytt tema i vitenskapelig sammenheng. Innen filosofien kan interessen spores tilbake til den greske antikken. Selve begrepet kontrafaktisk stammer nettopp fra filosofiske tekster, der den logiske statusen ved modale påstander og resonnementer rundt sannsynlighet har blitt nøye drøftet. De siste 30-40 årene har oppmerksomheten rundt kontrafaktisitet både som begrep og kognitivt fenomen fått et betydelig oppsving innenfor en rekke fagområder som historie, psykologi, sosiologi, filosofi og det juridiske området. Innen pedagogikkfaget er betydningen av kontrafaktisk tenkning nærmest utforsket, og noen av de teoretiske kildene for oppgaven er derfor hentet fra pedagogikkens såkalte hjelpevitenskaper som filosofi, psykologi og historie. Kontrafaktisk tenkning er et komplekst fenomen som inkluderer en rekke epistemologiske og ontologiske temaer. I og med at kontrafaktiske tankeprosesser er så fundamentale for både det mentale livet og følelseslivet, skjer disse prosessene på flere nivåer; helt fra den hverdagslige ubevisste implementeringen av kontrafaktiske påstander, til filosofiske debatter om naturen til kontrafaktisiteter. De ulike vitenskapelige disiplinene gjør seg nytte av kontrafaktiske resonnementer og drøfter betydningen av og sannhetsgehalten ved kontrafaktiske påstander på forskjellige måter, alt etter relevansen for faget deres. Det ligger i sakens natur at det stilles ulike spørsmål knyttet til kontrafaktisk tenkning innenfor de ulike disiplinene. Det er også en viss tverrfaglig utveksling, da ulike forskningsfunn og debatter innen for eksempel historiefaget er relevant for psykologien og vise versa. Likeså med betydningen av de filosofiske drøftingene for de andre fagene. Det er særlig to problemstillinger knyttet til kontrafaktisk tenkning som synes å være relevante på tvers av flere disipliner. Den ene handler om *bias*, dvs. skjevhet eller forutinntatthet ved vurderinger, som kontrafaktisk tenkning gjerne er beheftet med. Denne kan dermed være betydningsfull både for validiteten ved rene analytiske resonnementer og fortolkning av empiri. Den andre problemstillingen er av mer vitenskapsteoretisk karakter, og omhandler de underliggende metafysiske antagelsene ved kontrafaktisk tenkning og såkalt modal tenkning, som for eksempel kausalitet, fri vilje, determinisme og indeterminisme.

Tabell 1.2 er en oversikt over de ulike typene av kontrafaktisk tenkning som er karakteristisk for ulike vitenskapelige disipliner og samfunnsområder, samt en egen kategori for hverdagslivet (*common sense*). Formålet med denne oversikten er å belyse hvordan det forholder seg med benyttelsen av og bevisstheten om kontrafaktisk tenkning i den pedagogiske konteksten.

Det er flere grunner til at jeg har valgt å se ”betydninger av kontrafaktisk tenkning i den pedagogiske konteksten” i relasjon til andre fagfelt og samfunnsområder. En ting er at pedagogikkfaget generelt benytter seg av de såkalte hjelpedisiplinene filosofi, historie, psykologi og sosiologi. En annen ting er at det rett og slett er forsket mer på betydninger av kontrafaktisk tenkning i de andre disiplinene, dvs. innen pedagogikk er betydninger av kontrafaktisk tenkning nærmest utforsket. En annen grunn er at jeg oppfatter at pedagogikkfaget og utdanningsfeltet i større grad enn andre fag- og samfunnsområder kan kalles et *pre-kontrafaktisk prosjekt*. I begynnelsen av arbeidet var fokuset mitt rettet mot betydninger av kontrafaktisk tenkning i forbindelse med retrospeksjon. Men etter hvert som arbeidet skred frem, har jeg blitt mer oppmerksom på at innholdet i en del av de retrospektive refleksjonsprosessene hos både elever og lærere, skjer i konteksten av institusjonen der målene først og fremst er framtidsrettede. Styringen av skolen dreier seg i stor grad om å stabilisere, justere eller forandre ut fra en ønsket fremtidstilstand. Skolens ulike aktører befinner seg dermed i et spenningsfelt jeg betegner som *pre-kontrafaktisk* (se tabell 1.1). På bakgrunn av *historiske fakta* og *antesiperte sosiale ”fakta”* samt *hypotetiske scenarioer*, sammenlignes forventede scenarioer med *alternative scenarioer*. Dette kan gjelde resultater på internasjonale tester, fravær, frafall i videregående skole, mobbing, trivsel i skolen osv. De historiske ”faktaene” er fortolkninger av fortiden, samt utfallet av skolens spesifikke satsningsområder. Utfallet er da også utgangssituasjonen for videre planlegging. Jeg velger så å legge til at det utover å være *pre-kontrafaktisk*, også er et *normativt sosialt prosjekt* (tabell 1.2). Skolen er framtidsrettet med uttalte verdier og visjoner om en ønsket samfunnsutvikling. Inni dette pre-kontrafaktiske prosjektet befinner det seg en rekke ulike aktører, for eksempel skoleeiere, rektorer, lærere og elever. Jo mer disse aktørenes innsats vurderes innenfor skolesystemet, jo mer engasjerer de seg i kontrafaktisk tenkning gjennom retrospeksjon (eller det jeg i tabell 1.1 også har kategorisert som post-kontrafaktisk tenkning). Det innebærer at man ved retrospeksjon sammenligner faktiske utfall med alternative utfall. Det er denne formen for kontrafaktisk tenkning som er oppgavens hovedfokus. Blant skolens aktører er

elevene den gruppen som i størst grad vurderes og som vurderer seg selv under andres tilsyn, og som i minst grad vurderer andre.

Filosofifaget får betydelig mer oppmerksomhet enn de andre fagområdene, da det naturlig nok først og fremst er denne disiplinen som beskjeftiger seg med epistemologiske og ontologiske temaene knyttet til kontrafaktiske påstander og betingelser. Det er en kjensgjerning at oppgaven med å belyse og drøfte spørsmål om kontrafaktisitetens natur synes å havne på bordet til filosofen. I de følgende kapitlene vil jeg utdype tabell 1.2.

I tabell 1.2 har jeg kategorisert typer av kontrafaktisk tenkning etter følgende nivåer:

a) *Ureflekterte, ubevisste*. Jeg siterer Hofstadter: "The manufacture of 'subjunctive worlds' happens so casually, -so naturally, that we hardly notice what we are doing" (Hofstadter, 1979, s. 640). "Subjunctive" kan oversettes med hypotetisk konjunktiv; de tingene vi forestiller oss som ikke har skjedd.

b) *Metoderedskap; analytisk rammeverk*. Dette innebærer anvendelsen av kontrafaktiske kausalanalyser som metode. Anvendes innen alle fagområder.

c) *Metadebatt om kontrafaktiske påstander og betingelser*. Med dette menes at faget innbyrdes drøfter implikasjoner av kontrafaktisk resonnering. Historiefaget er et eksempel på en vitenskapelig disiplin der metoden *kontrafaktisk historieskriving* har vært heftig debattert. Et annet eksempel er det juridiske området, der man har drøftet betydningen av kontrafaktiske kausalanalyser ("but, for") knyttet til legale disputer:

In law, as in the rest of life, we indulge, indeed, require, many speculations on what might have been. Although such counterfactual thinking often remains disguised or implicit, we encounter it whenever we identify a cause, and quite often when we attempt to fashion a remedy. [...] Yet, as the examples above suggest, troublesome might-have-beens abound in legal disputes (Strassfeld, 1992, s. 342-343)

d) *Drøfting av naturen til kontrafaktisiteter*. Jeg henviser her først og fremst til debatter innen filosofi og vitenskapsteori om den epistemologiske statusen til kontrafaktiske betingelser og hypoteser.

1) *Fokus på bias: "hindsight" og "snikende determinisme"*. Dette innebærer hvorvidt det er et uttalt fokus på "skjevhet" eller forutinntatthet ved kontrafaktiske analyser, som for eksempel "snikende determinisme" (Fischhoff, 1974). Dette fenomenet har opptatt både historikere og psykologer.

2) *(Post-)kontrafaktisk prosjekt som normativ, sosial praksis*. Med dette menes hvorvidt et felt eller fagområde kan karakteriseres ved først og fremst å være rettet mot det historiske ved handlinger og hendelser. Etter min mening er det først og fremst innen rettssystemet at fortiden undersøkes i et normativt lys. Innen utdanningsfeltet gjøres det til en viss grad, men hovedfokuset er allikevel på det framtidige.

3) *(Pre-)kontrafaktisk prosjekt som normativ sosial praksis*. Med dette mens hvorvidt et felt eller fagområde kan karakteriseres ved først og fremst å være rettet mot framtiden. Her mener jeg at utdanningsfeltet utmerker seg, ved å ha eksplisitte målsetninger og visjoner for framtidens samfunn og mennesker.

4) *Fokusområder og spørsmål*. Dette er de spørsmålene som stilles innenfor de ulike fagområdene/feltene, og temaer man er spesielt opptatt av.

5) *Temporalitet ved kausalanalyser knyttet til aktørskap og ansvarlighet*: Viser hvilket tidsperspektiv man er rettet mot ved kausalanalyser knyttet til aktørskap og ansvarlighet; fortid eller fremtid.

2.2.1 Hverdagslivet

Karakteristisk for den hverdagslige kontrafaktisk tenkningen, er at den både er intuitiv og skjer relativt ureflektert. Selv om det kan være mye uenighet om innholdet i de hypotetiske scenarioene vi skaper, så stille vi sjelden spørsmålstegn ved det faktum at vi tenker kontrafaktisk. Den kontrafaktiske tenkningen og kommunikasjonen skjer ofte på formen: "hvis, så" eller "hvis bare". For eksempel kan en lærervikar tenke slik: "Hvis jeg hadde visst at eleven hadde en dysleksidiagnose, hadde jeg ikke rettet alle feilene og kommentert norskprøven slik jeg gjorde det". Et annet eksempel er eleven som tenker: "Hvis jeg bare hadde husket å ta med gymtøyet den siste gymtimen før ferien, så hadde jeg ikke fått den siste anmerkningen som medførte at jeg fikk nedsatt karakter i orden og oppførsel".

Selv om *kontrafaktisk tenkning* kan synes å referere til et rent kognitivt fenomen, er mange av våre daglige kontrafaktiske tankeprosesser nært knyttet til følelseslivet. De mest typiske ”kontrafaktiske følelsene” er anger og skyld (Byrne, 2005), og disse er igjen nært knyttet til sosiale vurderinger og tilskrivelser av handlingsfrihet og ansvarlighet. Det er mange av følelsestilstandene våre som viser seg å opptre samtidig med at vi sammenligner scenarioer. Et trekk ved en del følelser er at de er nært knyttet til forestillinger om en annen tilstand eller situasjon enn den nåværende. Eksempler på slike kontrafaktiske følelser kan være skam, lettelse, tilfredshet, sympati, lengsel, håp, sinne og sorg. I følgende sitat oppsummerer William James (1912) dynamikken mellom forestillinger om det som kunne ha vært og følelser:

[...] and eventually our remorse and regrets; and the error included in remorse and regrets, the error of supposing that the past could have been different, justifies itself by its use. Its use is to quicken our sense of what the irretrievably lost is. When we think of it as that which might have been ('the saddest words of tongue or pen'), the quality of its worth speaks to us with a wilder sweetness; and, conversely, the dissatisfaction wherewith we think of what seems to have driven it from its natural place gives us the severer pang (James, 1912)

Han påpeker at ”feilen” i å anta at fortiden kunne ha vært annerledes, rettfærdiggjør seg selv når vi tenker slik. Gjennom denne ”feilaktige antagelsen”, blir vi klar over hva som for alltid er tapt. Jeg tolker ikke James dithen at han argumenterer for at det er ”feil” å oppfatte virkeligheten slik, men snarere at det er en illustrasjon av dynamikken mellom forestillinger og følelser. Videre påpeker han hvordan tankene om det som *kunne ha vært* har en forlokkende effekt (”speaks to us with a wilder sweetness”), og omvendt om misnøyen vi opplever når noe synes å ha blitt revet vekk fra sitt ”naturlige sted”. Dette er samtidig en påpekning av at vi i en del tilfeller oppfatter det faktiske som kontrafaktisk, dvs. når ting som skjer ikke burde ha skjedd. Denne dobbeltheten ved de kontrafaktiske følelsene knyttet til dagliglivets virkelige hendelser er karakteristisk.

I tabell 1.2 har jeg begrenset den hverdagslige kontrafaktiske tenkningen til kategorien *ureflekterte og ubevisste*. Bevisstgjøring av og refleksjon over den hypotetiske tenkningen knyttet til hverdagslivet har først og fremst pågått i psykologiske studier.

2.2.2 Forskning generelt

Ved teoretisering innen forskning er kontrafaktisk tenkning et sentralt verktøy, både implisitt som en underliggende analytisk logikk og eksplisitt som metode. Kontrafaktiske kausalanalyser benyttes som analytisk rammeverk og metoderedskap innen alle fagfelt. Tankeeksperimenter er ansett for å være et uvurderlige analytisk verktøy, og særlig på de områdene der det ikke er mulig å bevise noe gjennom eksperimenter eller observasjonelle studier.

Det ekte eksperimentet er i sin natur kontrafaktisk, da man spør hva som ville ha skjedd dersom for eksempel et bestemt tiltak ikke hadde blitt innført. Men det kan også være en bevisst underliggende analytisk logikk i case-studier og naturlige eksperimenter, der man tester ut kausale hypoteser gjennom kontrafaktiske tankeeksperimenter. Den kontrafaktiske forståelsen av kausalitet er kanskje den vanligste innen årsaksforklaringer og analyser i vitenskap. Kausale forklaringer tar ofte form av beskrivelser av hvordan ting kunne vært annerledes (Anscombe, 1975; Sloman, 2005). En slik forståelse tar utgangspunkt i at kausalforklaringer innebærer påstander om en kontrafaktisk betingelse. Ved første øyekast kan det synes som at det kontrafaktiske rammeverket som metoderedskap er mest aktuelt for utvendige forklaringer av samspillet mellom hendelser, og dermed en type kausalforklaringer i den naturvitenskapelige tradisjonen. Men sett i lys av hvor essensiell den kontrafaktiske tenkningen er for virkelighetsforståelsen, kan det argumenteres for at det kontrafaktiske rammeverket er en underliggende analytisk logikk, også der målet er å forstå menneskers handlinger utfra den meningen de selv de. Dette kan begrunnes med at vi forstår handlinger gjennom deres potensielle forskjellighet. Aristoteles påpekte for eksempel i den nikomakiske etikk: “The class of things which admit of variation includes both things made and actions done” (Aristotle, 1926, VI, iii, 4-iv, 1140a). Således innbefatter det å ha vitenskapelig kunnskap kontrafaktisk kunnskap, og herunder kunnskap om virkelighetens modale struktur.

2.2.3 Psykologien

Siden 1970-tallet har kontrafaktisk tenkning som kognitivt fenomen fått betydelig oppmerksomhet innenfor empirisk psykologisk forskning, og da særlig innenfor kognitiv psykologi og sosialpsykologien. Spørsmålene som stilles er gjerne hva slags faktorer som generer kontrafaktiske tanker, hva er de underliggende prosessene og hva er konsekvensene av å resonnerer kontrafaktisk. Et annet spørsmål er innholdet i kontrafaktiske tanker; Hva er

det som ”binder” de kontrafaktiske tankene, eller hvilke er de typiske begrensningene for de fiktive konstruksjonene om virkeligheten? Dette er samme spørsmålet som Hofstadter stiller, selv om han undersøker det analytisk. Innen den *kognitive psykologien* har man vært opptatt av å undersøke kontrafaktisiteter, for å belyse grunnleggende prosesser ved for eksempel læring og hukommelse. Pionerarbeidet til Kahneman og Tversky (Kahneman, Tversky, & Slovic, 1982) viste at mennesker har en tendens til å ”omgjøre” mentalt uforutsette og eksepsjonelle hendelser mer enn de gjør om på vanlige hendelser. Innen *sosialpsykologien* har man vært opptatt av å undersøke individets kognitive funksjoner og emosjoner innen en bredere sosial kontekst.

Forskningen innen psykologien på 1980 og 1990-tallet hadde mye fokus på bias og på negative konsekvenser av kontrafaktisk tenkning. I løpet av det siste desenniet synes det å ha skjedd en dreining mot et funksjonelt perspektiv. Psykologen Neal J. Roese som har gjort en rekke studier på kontrafaktisk tenkning i senere tid, har et funksjonelt perspektiv på kontrafaktisiteter: ”In this view, counterfactual thinking may be seen primarily as a useful, beneficial, and utterly necessary component of behaviour regulation” (Epstude & Roese, 2008, s. 2). Vi skal se at nytteperspektivet på kontrafaktisk tenkning går igjen, ikke bare hos psykologiske forskere, men også hos filosofene, dypt inne i *fri vilje-determinisme* debatten.

Innen psykologien er det forsket mye på ulike konsekvenser av å resonnere kontrafaktisk. Jeg vil redegjøre nærmere noe som kan true validiteten ved undersøkelsene vi gjør i retrospekt, som gjøre ved tilbakeskuende. I den banebrytende artikkelen *Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty* omtalte Baruch Fischhoff (1974) en spesiell form for hindsight bias som han kalte ”snikende determinisme”. Det dreier seg om at vi ved retrospeksjon konstant overdriver hva som kunne vært antatt ved prospeksjon. I tillegg til tendensen ved retrospeksjon å oppfatte det som skjedde som uunngåelig, så ser folk på det som skjedde som uunngåelig *før* det skjedde. Hermed oppstår tendensen til ”snikende determinisme”. En konsekvens av dette er at folk forventer både av seg selv og andre at de skulle kunne ha vært mer forutseende enn hva som faktisk var mulig. De til og med minnes på en feilaktig måte sine egne prediksjoner, slik at de i retrospeksjon overdriver hva de visste på forhånd (Fischhoff, 1974). Fischhoff var imidlertid ikke den første som påpekte dette fenomenet. Historikeren Georges Florovsky (1969) viste til hvordan tendensen til determinisme er implisert i den retrospektive metoden:

In retrospect we seem to perceive the logic of events, which unfold themselves in a regular order, according to a recognizable pattern, with an alleged inner necessity, so that we get the impression that it really could not have happened otherwise.
(Florovsky, 1969)

Vi synes dermed å tillegge de fortidige hendelsene en indre nødvendighet og oppfatter at de ikke kunne ha skjedd på en annen måte enn de gjorde. Men samtidig står tendensen med ”snikende determinisme” i strid med en annen oppfatning som sier at vi kunne ha handlet annerledes, og at utfallet kunne ha blitt et annet.

I tabell 1.1 har jeg tatt med noen andre bias eller typiske skjevheter i vurderinger ved *pre-kontrafaktisk tenkning*. En tendens er å være skråsikker på hvordan noe vil utspille seg, og dette kan lede til både overdreven optimisme og pessimisme på vegne av framtidige hendelser. En annen tendens er den såkalte ”Nirvana-feilslutningen”, der sammenligningen skjer mellom en idealsituasjon (eller ideal norm) og en eksisterende situasjon (Demsetz, 1969, s. 1). Nirvanatilnærmingen handler om å finne diskrepanser mellom det ideelle og det reelle. Hvis det kan påvises slike, er det en tendens at det konkluderes med at det reelle er ineffektivt og ikke bra nok. Det er et kjent retorisk grep hos politikere; at sett i lys av hvor bra ting kunne ha vært, framstår realitetene som verre enn de egentlig er. Dermed er veien duket for økonomiske bevilgninger og annen form for støtte til intervensjoner og skolereformer.

2.2.4 Filosofien

I det daglige stiller vi sjeldent spørsmålstegn ved den generelle sannhetsgehalten til kontrafaktiske påstander. Men det er ikke dermed sagt at det er alltid enighet om hva som kunne eller burde ha skjedd, og hvordan virkeligheten skulle kunne ha sett ut hvis ting hadde fått utspille seg på en annen måte. Innen filosofien har derimot gyldigheten ved kontrafaktiske påstander vært et pågående kontroversielt tema. I og med at kontrafaktiske påstander ikke kan undersøkes empirisk, men er avhengig av konteksten for de konklusjonene man trekker, representerer de et problem i forhold til sannhetsverdien. Ikke sjelden munner diskusjoner om kontrafaktiske påstander ut i forslag eller påstander om virkelighetens natur, og hvor skillene går mellom hva som er faktisk og hva som er det kontrafaktisk.

I debatten om frihet og moralsk ansvarlighet er selve byggesteinen i premissene for fri vilje en kontrafaktisk betingelse. For at man skal kunne holde en person ansvarlig for en handling, er det en nødvendig betingelse at personen kunne ha handlet annerledes. “And almost all philosophers agree that a necessary condition for holding an agent responsible for an act is believing that that agent could have refrained from performing that act” (Inwagen, 1975, s. 50).

Kontrafaktisiteter har lenge spilt en betydelig rolle i filosofisk tankegang og metodologi. Allerede fra før-sokratisk tid har filosofer forsøkt å belyse virkelighetens beskaffenhet ved å sette den opp mot kontrafaktiske modus ved situasjonen. Xenophanes fra Colophon baserte seg for eksempel på tankeeksperimenter for å begrunne relativiteten til menneskelig kunnskap: “No man knows, or even will know, the truth about the gods and about everything I speak of: for even if one chanced to say the complete truth, yet one knows it not [...]” (Kirk, Raven, & Schofield, 1983, s. 179). Både Platon og Aristoteles drøftet den epistemologiske statusen til hypotetiske konjunktiver, og Gottfried Leibniz argumenterte på 1600-tallet i sin modale metafysikk for at en alternativ virkelighet er mulig, så lenge den ikke står i motsetning (eller kontradiksjon) til logikkens formelle lover. Leibniz er dessuten ansett som forløperen for utviklingen av modal logikk i det tjuende århundre. Det vi i dag kjenner som ”mulig verden”-semantikken oppstod med David Lewis sin modalrealisme på 1970-tallet. Før dette ble lignende temaer som ”hva hvis”-scenarioer og fiksjoner behandlet av en rekke filosofer. Fra det nittende århundre kan jeg særlig nevne filosofene Jeremy Bentham (1932) og Hans Vaihinger (1968) som på hver sin måte forsøkte å rehabilitere *fiksjonens* status. Bentham ville vise at fiksjon og fakta er to uatskillelige aspekter ved den samme kognisjonen (de Champs, 2011, s. 1). For ham er fiksjon eksistensmodusen til alt som er virkelig. Vaihinger som var kantianer, utviklet en hel filosofi om fiksjonalisme i storverket *Als ob* (eng. *What if*) (Vaihinger, 1968). Han hevdet at all erkjennelse er basert på fiksjoner, som han igjen så på som nyttige oppfinnelser. Dette inkluderte også begrepet om frihet: “We encounter at the very threshold of these fictions one of the most important concepts ever formed by man; the idea of freedom; human actions are regarded as free, and therefore as ‘responsible’ and contrasted with the necessary course of natural events” (Vaihinger, 1968, s. 43). Den ontologiske statusen Bentham og Vaihinger gir fiksjoner står dermed i kontrast til hverandre. En annet eksempel på en filosof som står i kontrast til Vaihingers anti-metafysiske posisjon, er Ernest Bloch. I hans håpsfilosofi *Das Prinzip Hoffnung* (eng. *The Principle of Hope*) innebærer

håpsbegrepet en ikke-empirisk realitet. Han tillegger dermed håpet og *det-ennå-ikke-værende* en ontologisk status (Bloch, 1986).

Sosiologen Roland Paulsen tar utgangspunkt i de to motstridende perspektivene på det kontrafaktiske hos Vaihinger og Bloch. Han kategoriserer de i *instrumentelle* (Vaihinger) og *utopiske* (Bloch) kontrafaktisiteter. Paulsen argumenterer for at kontrafaktiske forestillinger er nødvendige for å overskride makt. Instrumentelle kontrafaktisiteter er løsninger som gjør at maktstrukturer forblir uendrede, mens utopiske kontrafaktisiteter er løsninger som utfordrer dominerende maktstrukturer (Paulsen, 2014, s. 160). Denne kategoriseringen sammen med perspektivet på maktoverskridelse, vil jeg benytte meg av i en drøfting i kapittel 7.2.

Nelson Goodman skrev i det klassiske essayet *The problem of Counterfactual Conditionals*: “The analysis of counterfactual conditionals is no fussy little grammatical exercise. Indeed, if we lack the means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science” (Goodman, 1973, s. 3). Siden Goodman satte søkelyset på temaet, har kontrafaktiske betingelser blitt gitt betydelig oppmerksomhet.

Kausalitetsteorier med utgangspunkt i det kontrafaktiske har fått et stort oppsving på tvers av fagdisipliner, og er i dag en av de dominerende teoriene om kausalitet (Kvernbekk, 2016). I neste avsnitte vil jeg ta en nærmere kikk på Lewis teori om *mulige verdener*, for å la det føre inn i en debatt om skillet mellom det faktiske og det kontrafaktiske.

Lewis posisjon kalles *modal realisme*. Han tar bokstavelig den måten vi snakker om det som kunne ha skjedd, og om det som kunne ha vært. Styrken i argumentasjon hans ligger i forslaget om at mulige verdener ikke er så fremmede for oss som det metafysiske navnet kan synes å innebære. For å forstå Lewis prosjekt må vi først se på hvordan Hume definerte kausalitet; den såkalte regularitetsanalysen:

”We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. Or in other words where, if the first object had not been, the second never had existed” (Hume, 2004, s. 48-49)

Hume definerte dermed kausalitet i to vendinger. Lewis utviklet sin teori ved å ta fatt i den siste vendingen, og foreslo en kontrafaktisk analyse av kausalitet. En sentral idé i analysen er

det han kaller sammenlignbar likhet mellom ulike verdener. En verden er nærmere den aktuelle verden enn en annen verden, hvis den første ligner mer på den aktuelle verden enn den andre. Ved å drøfte ting som ikke har skjedd, har han funnet et verktøy for kausalitet som fungerer i et ikke-deterministisk system.

Lewis teori om kausalitet i termer av kontrafaktisk avhengighet gjenspeiler dermed den intuitive oppfatningen vi har om årsak og virkning i det daglige. Selv om vi vanligvis ikke tenker over det, så impliserer vi den kontrafaktiske avhengigheten idet vi hevder at A forårsaker B; for dersom A ikke hadde inntruffet, så hadde ikke heller B inntruffet. Ta eksempelet med en lærer som hevder at årsaken til at en elev har vært trist den siste uken, er at foreldrene ligger i skilsmisse. Da oppfatter vi implisitt at hvis foreldrene ikke hadde ligget i skilsmisse, så ville ikke eleven ha vært trist. Et annet eksempel kan være kritikken av dagens skolesystem, en målstyrt skole som de siste årene har hatt et økende fokus kompetansemål og testing. Det har samtidig vært en oppsiktsvekkende økning i alvorlige stressrelaterte helseproblemer hos barn og unge. Kritiske røster hevder at disse plagene har bakgrunn i at de unge blir utsatt for umulige krav, fordi de uansett aldri kan bli gode nok innenfor dette systemet. Den kontrafaktiske betingelsen implisitt i dette, er at god psykisk helse fremmes når barn og unge opplever at de strekker til og at det de presterer er godt nok.

For filosofer er det nærliggende å hevde at kausalitet involverer kontrafaktisiteter. Som Tim Crane (2003) sier: "Philosophers put this by saying that causation involves counterfactuals: truths about matters 'contrary to fact'" (Crane, 2003, s. 56). Videre påpeker Crane: "So we could say that, if we believe that A caused B, we commit ourselves to the truth of the counterfactual claim: 'If A had not occurred, B would not have occurred'" (Crane, 2003, s. 56). Selv om dette kan synes som en påpekning av noe trivielt, så er det interessant å merke seg at gjennom å hevde en kausal sammenheng mellom hendelser så har man samtidig forpliktet seg til en påstand om en kontrafaktisk betingelse.

Noe av problemet med den kontrafaktiske analysen, er at vi har beveget oss fra en empirisk verden til en mulig verden. Men så kan vi hevde at det er nærmest umulig å beskrive den empiriske verdenen uten å ta i bruk eller i det minste anta kontrafaktisiteter. Selv om Lewis teori om mulige verdener har fått stor innflytelse, er det mange som ikke slutter seg til den modale realismen; idéen om at mulige verdener virkelig eksisterer. Lewis sier selv at det ikke kan bevises at dette er sant, men han hevder at teorien hans er løfterik i og med at *hvis* det er

sant, *så* er mange av problemene innen språkfilosofien og logikken løst (Lewis, 1973)⁴. Her er det verdt å legge merke at Lewis forslag innebærer en kontrafaktisk betingelse. Det blir litt som å si at i en mulig verden er det mulig at mulige verdener eksisterer.

Det er kanskje ikke så rart at Lewis modale realisme er ansett for å være kontroversiell. Men hvis vi skal ta på alvor hvordan vi faktisk tenker og uttrykker oss, og hvordan vi i stor grad forstår virkeligheten gjennom å tenke kontrafaktisk om den, bør vi kanskje hevde at vår gjengse virkelighetsoppfatning er mer kontroversiell enn vi vanligvis tenker. Men selv om vi ser bort fra *mulig verden* – semantikken og ”bare” forholder oss til kontrafaktisk tenkning som sådan, så finnes det kritiske stemmer. Jon Elster (1978) inntar for eksempel et kritisk standpunkt til kontrafaktisk resonnering når han påpeker et skille mellom det å benytte lover og det å bekrefte de: ”Counterfactual propositions are law-users and not law-confirmers. If, on the basis of a certain generalization, a counterfactual proposition is asserted, this does not constitute further proof of the generalization”. Men samtidig mener han at kontrafaktiske spørsmål kan forsvares og legitimeres under visse begrensede omstendigheter (Elster, 1978, s. 191). Et klassisk forsvar for kontrafaktisiteter er at alle forklaringer til en viss grad impliserer og dermed støtter sine egne kontrafaktisiteter.

Ettersom det kontrafaktiske er så nært knyttet til oppfatninger om kausalitet, vil jeg drøfte dette mer utførlig ved å ta utgangspunkt i et spørsmål som Douglas Hofstadter stiller om kontrafaktisiteter. Hofstadter spør: ”Why do some counterfactuals strike us as ’less counterfactual’ than other counterfactuals?” For som han sier videre, så er det åpenbart at det som ikke skjedde ikke skjedde, og at det ikke finnes grader av ”didn’t-happen-ness” (Hofstadter, 1979, s. 638). Det at noe *nesten* skjedde ligger i sinnet, og ikke i de eksterne faktaene, påpeker han. Her vil nok de fleste si seg enig i at *nesten-hendelser* ligger i sinnet, men det er mer uenighet om hva slags forhold de såkalte eksterne faktaene har til sinnet.. Richard Ned Lebow (2000) hevder at ethvert skarpt skille mellom faktisiteter og kontrafaktisiteter hviler på tvilsomme ontologiske påstander. Dette begrunner han med at filosofer lenge har anerkjent at fakta er en sosial konstruksjon. I motsetning til fysiske gjenstander er sosiale fakta refleksjoner av begreper vi bruker for å beskrive den sosiale virkeligheten, og ikke en beskrivelse av virkeligheten i seg selv. Lebow konkluderer med at: “Every good counterfactual thus rests on multiple factuals, just as every factual rests on counterfactual assumptions – and these assumptions too often go unexamined” (Lebow,

⁴ Mine uthevinger.

2000). Lebow sin påstand griper inn i flere sentrale vitenskapsteoretiske problemstillinger. Den ene handler om divergerende ontologiske antagelser om virkelighetens beskaffenhet, og epistemologiske oppfatninger om hvilken tilgang vi har til denne ”virkeligheten”. Den andre problemstillingen handler om kausalbegrepet: De ontologiske oppfatningene om forholdet mellom det faktiske og det kontrafaktiske, synes å henge sammen med oppfatninger om kausalitet. De kontrafaktiske scenariene krever at vi gjør en slags ”mental gest” i motsetning til en fysisk gest. I stedet for å hevde noe om årsakssammenhenger i fenomener som har utspilt seg, går vi opp ”mentale kausale stier” der utfallet blir et annet enn det vi kan observere eller slutte oss til om noe utenforliggende. Men visse teoretikere hevder at kausalitet kun er et fortolkningsskjema og ikke en observerbar regularitet mellom fenomener. Da blir i stedet spørsmålet hva som skiller mellom ulike fortolkningsskjemaer. Ifølge en slik oppfatning har vi et fortolkningsskjema som skiller mellom det faktiske og det kontrafaktiske, uten at disse fortolkningene henviser til noe objektivt utenfor oss selv. Per Arne Tufte forklarer hvordan debatten mellom posisjonene kausal *realisme* og *konstruktivisme*: ”...vedrører spørsmålet om kausalitet eksisterer uavhengig av vår bevissthet, eller er knyttet til vår bevissthet og oppfattelse av verden rundt” (Tufte, 2013, s. 342).

Med tanke på hvor mange tilfeldigheter som tilsynelatende er avgjørende i for eksempel befruktningsprosessen, så er hvert eneste menneskes eksistens karakterisert ved den kontrafaktiske fortellingen om *nesten-hendelsen*: ”Jeg ble nesten ikke til”. Men at jeg for eksempel skriver om kontrafaktisk tenkning, oppleves i høyeste grad som et faktum som innebærer noe mer enn en sosial konstruksjon. På tross av dette mener jeg at Lebow sin påstand om at det er et gjensidighetsforhold mellom det vi omtaler som det faktiske og det kontrafaktiske, og at vi har mye å hente i å undersøke de underliggende antagelsene, virker svært rimelig.

2.2.5 Historiefaget

Innen historieforskningen er det en pågående kontrovers om kontrafaktiske analyser er egnet som metoderedskap. Slik jeg oppfatter det og har notert det i tabell 1.2, er historiefaget sammen med filosofifaget representert i alle nivåene av ulike typer kontrafaktisk tenkning. I den norske skolen er det nettopp i forbindelse med historiefaget at elevene får gjøre seg kjent med begrepet kontrafaktisk. På tross av at kontrafaktisk historieskriving som analytisk metoderedskap er regnet for å være kontroversielt innen historiefaget, så lærere elevene om

dette. Det ble introdusert i forbindelse med innføringen av Kunnskapsløftet (LK06). I læreplanen står det: ”Mål for opplæringa er at eleven (etter 10. årssteget) skal kunne: - finne døme på hendingar som har vore med på å forme dagens Noreg, og reflektere over korleis samfunnet kunne ha vorte dersom desse hendingane hadde utvikla seg annleis” (LK06, 2005). Eleven får øve seg i å tenke kontrafaktisk om historiske hendelser, men de blir ikke gjort oppmerksomme på at kontrafaktisk tenkning er noe vi holder på med hele tiden, og den fundamentale rollen det spiller hvordan vi forstår virkeligheten og legitimerer handlinger.

2.2.6 Pedagogikkfaget og utdanningsfeltet

Kontrafaktisk tenkning i den pedagogiske konteksten skjer på mange nivåer og tar seg ulike uttrykk. Men det finnes knapt noen litteratur som drøfter metateoretiske perspektiver knyttet til kontrafaktisk tenkning. Jeg henviser her til mangel på litteratur der begreper som *kontrafaktisk* eller *mulig-verden* benyttes eksplisitt. I tabell 1.2 ser vi at den ubevisste og ureflekterte typen av kontrafaktisk tenkning er høyt representert man innen pedagogikkfaget og utdanningsfeltet. Dette begrunner jeg med at skolen er et i særklasse *pre-kontrafaktisk* prosjekt, og at dette danner en viktig kontekst for alle aktørene i skolen. For å sette det på spissen, så er skoleeiere, rektorer og lærere kronisk opptatt av ting som kunne og burde ha vært annerledes. Samtidig er mye av skolens virksomhet preget av hverdagstenkning. Videre er kontrafaktisk tenkning som analytisk rammeverk og metoderedskap mye anvendt innen pedagogisk forskning. Det finnes for øvrig mange tilgrensende begreper og perspektiver - transcendentale idéer, utopisk tenkning, forestillingsevnen, moralsk dannelse og lignende. Men de dypere implikasjonene av kontrafaktisk tenkning som sådan, er et nærmest utforsket område innenfor det pedagogiske feltet.

Når som helst det konkluderes utfra kontrafaktiske betingelser, så er idéen om potensielle utfall i sentrum. Begrepet potensialitet er i seg selv ikke entydig og har en heller spekulativ natur, da det refererer til enhver mulighet som noe kan sies å ha. På tross av dette, spiller begrepet om potensialitet en stor rolle i pedagogisk tenkning, praksis og utdanningspolitikk. Uansett hvor vagt begrepet måtte være, så er et udiskutabelt mål med utdanning det å oppfylle potensial. Det innebærer å skape noe eller å få noe til å skje som er karakterisert ved sitt fravær; det er (ennå) ikke her. Sentralt i denne idéen i utdanningssammenheng, om dette som både er fraværende og potensielt samtidig, er at det ikke kommer til å frambringes av seg selv, hvis vi lar ting være som de er. For å bli det som det er ment å bli, trengs det bestemte

påvirkninger, bestemte strukturer og bestemte former for kontroll. Dette igjen krever refleksjon og planlegging. Et uunnværlig analytisk verktøy i slike prosesser, er kontrafaktisk tenkning om hvordan ting kunne ha blitt annerledes, dersom andre valg hadde blitt tatt, og man hadde handlet annerledes. Kontrafaktisk tenkning i den pedagogiske konteksten skjer på mange nivåer og tar seg ulike uttrykk. Jeg har allerede nevnt idealene om frihet og autonomi. Andre eksempler er en type slagord som betegnelsen på stortingsmelding nr. 16: "...og ingen sto igjen..."⁵, som er oversatt fra den amerikanske rapporten: "No child left behind" (NCLB). Den amerikanske skolereformen var basert på premisset om at målstyring ved bruk av nasjonale standarder, ville minske forskjellen mellom elevene. Denne kan kalles en pre-kontrafaktisk påstand. Blant skoleledere og politikere er det vanlig å benytte kontrafaktiske påstander for å attribuere årsaker og legitimere tiltak og handlinger. Lærere sammenligner kontinuerlig utfall eller resultater med forutbestemte mål og idealer. Sist, men ikke minst, så tenker elevene kontrafaktisk om hvem de kunne eller burde ha vært, og om hva de kunne ha gjort annerledes. Kontrafaktiske tanker om selv de mest verdslige ting bygger på ontologiske oppfatninger og antagelser.

I dette kapittelet har jeg sett på noen utvalgte dimensjoner av kontrafaktisk tenkning, som jeg vurderer som sentrale for å belyse problemstillingen. Jeg har argumentert for at selv om oppgaven først og fremst handler om kontrafaktisk tenkning knyttet til retrospeksjon, så spiller det prospektive en vesentlig rolle. I tabell 1.2 har jeg vist hva slags typer kontrafaktisk tenkning som er karakteristisk for ulike vitenskapelige disipliner og samfunnsområder. Oversikten viser samtidig at innen pedagogisk teori og praksis er temaet til stor del ignorert.

Den kontrafaktiske tenkningen i hverdagslivet pågår mer eller mindre automatisk. Kontrafaktiske tanker og "kontrafaktiske følelser" er nært sammenknyttet og spiller en sentral rolle for sosiale vurderinger og tilskrivninger av handlingsfrihet og ansvarlighet.

Innen psykologien har man vært opptatt av bias i forbindelse med kontrafaktisk tenkning. Den meste framtreddende tendensen er fenomenet "snikende determinisme" (Fischhoff, 1974). Tendensen til å oppleve at de fortidige hendelsene skjedde med nødvendighet står allikevel i motsetning til den grunnleggende oppfatningen om at vi har fri vilje, og kunne ha handlet annerledes. Vi har sett at det er sterk uenighet om hva slags skille det er mellom det faktiske og det kontrafaktiske. De motstridende posisjonene gjenspeiler divergerende ontologiske

⁵ Fra stortingsmelding nr. 16 (2006-2007)

antagelser om virkelighetens beskaffenhet, og epistemologiske oppfatninger om hvilken tilgang vi har til denne ”virkeligheten”. De ontologiske oppfatningene om forholdet mellom det faktiske og det kontrafaktiske, synes å henge sammen med oppfatninger om kausalitet.

Fra dette kapitlet vil jeg ta med meg begrepsparet instrumentelle og utopiske kontrafaktisiteter fra Paulsen (2014), for å drøfte maktperspektiver i kapittel 7.2.

3 Kontrafaktiske påstander

I dette kapitlet vil jeg se nærmere på hva som karakteriserer kontrafaktiske påstander. Først redegjør jeg for modalitet og modal/formal logikk, og deretter tar jeg for meg kontrafaktiske påstander og hva som skiller de fra formale påstander. Kontrafaktisk resonnering innebærer at vi stiller en type betingelser på formen: ”hvis, så” og ”hvis, bare”, og benytter modale hjelpeverb som *skulle*, *kunne*, *ville*, *burde* etc. Med disse *modale operatorene* skapes en form for mentale rom der vi kan undersøke kontrafaktiske situasjoner eller mulige verdener, og vurdere den epistemiske statusen av disse som *mulig* eller *nødvendig*. Derfor kan kontrafaktisk tenkning også kalles for modal tenkning.

3.1 Modalitet og modal logikk

Begrepet modalitet stammer fra det latinske begrepet *modus*, som betyr *måte*, og et utsagns modalitet uttrykker på hvilken måte et utsagn er sant. I motsetning til *kategoriske* påstander som sier hvorvidt noe er aktuelt og kan underkastes empiriske undersøkelser, så kan modale påstander ikke undersøkes eller bekreftes empirisk.

Modal logikk kalles den delen av logikken som tar for seg den logiske strukturen ved de nevnte begrepene som *nødvendig*, *mulig*, *kontingent* og *umulig*. Innen filosofisk logikk har dette vært en sentralt emne, helt fra oldtidsfilosofene og fram til i dag.

Innen formallogikken søker man ved resonnementer å abstrahere generelle prinsipper som er uavhengige av konteksten resonnementene opptrer i. Argumenters gyldighet bestemmes utfra argumentasjonens struktur eller form, for eksempel ved å undersøke forbindelsen mellom premisser og konklusjon. Hvis vi vet at setningen A er sann, så vet vi også at setningen “enten A eller B” er sann. Det var Aristoteles som grunnla den formale logikken, da han søkte etter eksakt kunnskap ved å undersøke betingelsene for bevisførsel. Emnet var rent logisk, og handlet ikke om hvorvidt man kunne finne sanne premisser (Næss, 1963).

To skjemaer inne den formale logikken har lenge vært kjent under navnene *modus ponens* og *modus tollens*:

<i>Modus ponens:</i> Hvis a så b a <hr/> b	<i>Modus tollens:</i> Ikke b , derfor ikke a ikke a <hr/> Ikke b
---	---

3.2 Kontrafaktisk tenkning – modal tenkning

Innen matematikken fungerer “hvis a så b ” - formen utmerket, fordi det som er matematisk usant forblir usant, og det som er sant, forblir sant. Men i det virkelige livet vil vi ofte vite hvordan ting hadde vært hvis de var annerledes. Et utsagn på formen “hvis...så”, kalles en betingende påstand. Et “hvis...så” -utsagn der vi snakker om hva som ville ha skjedd hvis noe var sant, kalles en kontrafaktisk (betingende) påstand. I motsetning til påstander innen formallogikken, er kontrafaktiske påstander meningsløse uten informasjon om den konteksten de opptrer i. For at påstanden hvis “ a så b ”, skal kunne kalles kontrafaktisk, må minst en av antagelsene være åpenbart usann både for den som uttrykker den og i forhold til delt kunnskap om saksforholdene. Mens den logiske “hvis a så b ”, ikke krever at vi vet hva a eller b faktisk sier noe om, så krever påstanden “Hvis A var sant, så ville B være sant” at vi også får utfyllende informasjon om de kontekstuelle betingelsene før vi kan vurdere den.

Vi har sett hvordan at modale påstander og dermed kontrafaktiske påstander ikke kan testes ut empirisk og er dermed avhengig av den konteksten de opptrer i. Omgangen med modaliteter i tanken og språket har dermed viktige normative implikasjoner. Robert Stalnaker påpeker dette aspektet ved modale setninger: ”Sentences involving modals are usually to be construed relative to all possible worlds consistent with the speaker’s knowledge, or with some set of presuppositions, or with what is morally right, or legally right, or normal, or what is within someone’s power” (Stalnaker, 1999, s. 35-36).

4 De metafysiske rammene

Denne delen innledes med en begrunnelse for å sette oppgavens problemstilling og tematikk inn i et metateoretisk rammeverk. Deretter tar jeg for meg de de metafysiske posisjonene *determinisme/indeterminisme*. Når *fri vilje* sammenholdes med *determinisme*, får vi posisjonene *kompatibilisme* og *inkompatibilisme*. Så tar jeg for meg forskjellen mellom determinisme og kausalitet. Videre sammenligner jeg og drøfter den filosofiske posisjonen om determinisme med tendensen til ”snikende determinisme”. Til slutt viser jeg til hvordan et område av metafysikkens fri vilje-problem overlapper med noen av de mest sentrale spørsmålene innen pedagogisk filosofi. Ettersom hverken determinisme/indeterminisme og fri vilje er klart avgrensede størrelser eller begreper, og tilfanget av litteratur og perspektiver er overveldende, vil jeg i dette kapitlet redegjøre for begrepene utfra et hensyn om relevans for den videre drøftingen av oppgavens problemstilling.

4.1 Kontrafaktisk tenkning og metafysiske antagelser

Den akademiske striden om determinisme og fri vilje har pågått i flere tusen år og har engasjert utallige filosofer. En del av litteraturen kjennetegnes av høyt abstraksjonsnivå med sofistikerte utlegninger, og kan i en del tilfeller synes utilgjengelig for andre enn fagfolk og spesielt interesserte. Den norske filosofen Lars. Fr. H. Svendsen henviser til blant annet dette i boken *Frihetens filosofi*:

Det beste ville være å sette en stor parentes rundt hele debatten om hvorvidt fri vilje er forenlig med determinisme – eller indeterminisme, for den saks skyld og heller gå mer pragmatisk til verks for å se hvordan utbredte forestillinger om frivillighet, tilregnelighet og ansvar lar seg belyse ut fra hva vi faktisk *vet* om oss selv og verden i stedet for å bli sittende fast i en metafysisk spekulasjon (Svendsen, 2013, s. 23)⁶.

Han påpeker videre at de metafysiske problemstillingene ikke har funnet noen tilfredsstillende løsning. Selv om uttrykket “metafysisk spekulasjon” framstår som noe ledende og som et ledd i Svendsens argumentasjon, er det samtidig lett å være enig med han, hvis det å engasjere seg i den metafysiske debatten om fri vilje og determinisme kan føre til at man mister kontakten

⁶ Svendsen sin utheving.

med de fenomenologiske erfaringene som danner utgangspunktet for kontroversen. Men hva vet vi faktisk om oss selv? Når jeg velger å plassere oppgavens tematikk inn i dette klassiske metateoretiske rammeverket, er det fordi den kontrafaktiske tenkningen som utgjør en grunnleggende bestanddel ved vår rasjonelle livsverden, impliserer metafysiske antagelser om virkeligheten av nettopp den typen Svendsen synes vi like gjerne kan slutte å debattere.

På det metateoretiske planet er det en vanlig forutsetning at ”jeg kunne ha handlet annerledes”, når det argumenteres for fri vilje og ansvarlighet: “And almost all philosophers agree that a necessary condition for holding an agent responsible for an act is believing that that agent could have refrained from performing that act” (Inwagen, 1975, s. 50). De fleste filosofer knytter dermed begrepet fri vilje til moralsk ansvarlighet. For å kunne holdes moralsk ansvarlig for en handling, kreves en type frihet som innebærer det å ”kunne ha handlet annerledes”. Men det å sammenholde disse er ikke nødvendigvis uproblematisk. Dersom fri vilje er en forutsetning for moralsk ansvarlighet, innebærer det at hvis vi ontologisk sett ikke har fri vilje, så er våre moralske sentimentaler illusjoner. Da er antagelsen om at vi har fri vilje kun en nyttig fiksjon, slik filosofen Vaihinger (1968) hevder. Det er imidlertid en rekke filosofer i dag som benekter at vi har fri vilje, men som samtidig ikke er villige til å konkludere med at moralsk ansvarlighet er en illusjon. Med hensyn til den kontrafaktiske tenkningen, framstår det lite hensiktsmessig å resonnerer kontrafaktisk over hendelsesforløp retrospektivt og tillegge bedømmelsene noen vekt, med mindre vi antar muligheten for at ting kunne vært annerledes, og at mennesket grunnleggende sett har den egenskapen at det kan virke *kausalt* eller *kontrakausalt* i verden. *Kontrakausalt* henviser til det å kunne virke kausalt eller starte en kausal kjede uten selv å være kausalt determinert av andre årsaker. Aristoteles omtaler det slik i Metafysikkens bok: “...there is something which moves without being moved...” (Aristoteles, 1935, XII , vii, 1072a).

Et annet poeng er at vi i mange sammenhenger er opptatt av å forklare hvorfor ting skjer, og noen ganger i den grad at det kan se ut som vi tar utgangspunkt i en idé om at det er mulig å forklare alt. Selv om det å undersøke kausale sammenhenger ikke forutsetter antagelsen om et deterministisk rammeverk, så kan det allikevel se ut som at det er nettopp det vi gjør i en del tilfeller. Når det kommer til handlingslivet vårt, handler de ulike antagelsene om hvilke egenskaper vi mener at mennesker har, og hva slags univers vi lever i.

4.2 Tesen om determinismen

Determinisme er en samlebetegnelse for ulike teser om virkelighetens beskaffenhet, som også har det til felles at de anses som potensielle trusler mot frihet. Begrepets flertydighet blir kanskje aller mest synlig når det settes opp mot eller sammenholdes med ulike begreper om viljens frihet. John Earman innleder boka *A Primer on Determinism* om begrepets tvetydighet:

Very little acquaintance with the philosophical literature is needed to reveal the Tower of Babel character of the discussion [...]; some take determinism to undermine human freedom and dignity, others see no conflict, and yet others think that determinism is necessary for free will; and on and on (Earman, 1986).

Tesen om determinismen sier at det i ethvert øyeblikk bare er én mulig fortsettelse av tingenes tilstand, som er konsistent med naturens lover. Det finnes dermed ingen alternative framtidige hendelsesforløp. Som William James uttalte: ”The future has no ambiguous possibilities hidden in its womb: the part we call the present is compatible with only one totality” (James, 1912, s. 150)⁷.

Determinismens røtter ligger i en utbredt filosofisk idé om at alt i prinsippet kan forklares, eller at alt som finnes har en tilstrekkelig årsak for å være nettopp slik det er, og ikke på en annen måte. Dermed er de ulike teoriene om determinisme nært knyttet til kausalitet og forklaring, og det er ikke uvanlig at determinisme og kausalitet assosieres med hverandre. Mange forstår determinisme som en form for sterk kausalitet. Men determinismen er ikke en tese om kausalitet. Ideen om at hendelser følger nødvendig fra deres forutgående uten rom for sannsynlighet eller usikkerhet, er en teori om virkelighetens beskaffenhet, mens ulike tilnærminger til kausalitet handler om definisjoner av et begrep. Det finnes en rekke ulike tilnærminger til kausalitet, men for dette tilfellet sammenfattes kausalitet generelt som påstanden om at hendelse *e* følger forutgående tilstander *c* med regelmessighet. Men hendelse *e* er ikke alltid resultatet av eller følger ikke med nødvendighet tilstand *c*; vi kan hermed ha kausalitet uten determinisme. Filosofen Elizabeth Anscombe utdyper en vesentlig forskjell mellom kausalitet og determinisme slik: ”...a thing hasn’t been caused until it has happened; but it may be determined before it happens” (Anscombe, 1981, s. 140).

⁷ Fra et foredrag på Harvard Divinity School i 1884 (James, 1912)

Fra gammelt av har det imidlertid vært vanlig å anta en sammenheng mellom kausalitet og nødvendighet, og selv om dette er i endring, kan tendensen til å assosiere eller blande sammen kausalitet og determinisme, forklares med at det er en sterk pragmatisk forbindelse mellom disse. Germund Hesslow hevder at vi impliserer og anvender et deterministisk rammeverk når vi karakteriserer kausalitet. Han mener at selv om det ikke er en logisk forbindelse mellom determinisme og kausalitet, kjennetegnes den nåtidige benyttelsen av kausalitetsbegrepet antagelsen av en form for determinisme (Hesslow, 1981). Filosofen Lars Fr. H. Svendsen omtaler noe av det samme når han sier at det å forutsette determinismen er et metodologisk grep; antagelsen om at x har tilstrekkelige kausale betingelser fordi man bare da kan gi en fullstendig forklaring. Han påpeker imidlertid det problematiske i å slutte fra en slik metodologisk antakelse til en ontologisk påstand. Svaret på hvordan verden faktisk *er*, ligger ikke i antagelsene om den (Svendsen, 2013, s. 40-41). Jeg tolker Hesslow dithen at han i større grad enn Svendsen påpeker et problem med at vi uten videre eller underbevisst antar determinismen når vi snakker og tenker kausalt. Et lignende analytisk rammeverk er også utbredt i vår dagligdagse tenkning. Bysted Möller (2011) beskriver hvordan den deterministiske tesen besitter en form for intuitiv evidens. Når det gjelder handlingene våre, synes det som selve muligheten for å gi en uttømmende begrunnelse for et valg, er en betingelse for at valget skal gi mening. Ved hendelser som er viktige for oss, er det sjelden vi nøyer oss med uforklarte handlinger. Da henvender vi oss gjerne til andre instanser som psykiatrien eller lignende, for å finne grunnene for handling. Det later til at vi forstår våre egne og andres valg ut fra et prinsipp som fordrer en tilstrekkelig grunn (Bysted Möller, 2011, s. 117-118). Vi synes å anta både i dagliglivet og innen vitenskapen et prinsipp om tilstrekkelig årsak. Det er i filosofisk praksis i dag vanlig å skille teorier om kausalitet fra konklusjoner om sannhetsgehalten ved determinismen. Man aksepterer dermed at det kan være sant at enhver hendelse har en årsak selv om determinismen er usann, noe som igjen innebærer at noen hendelser mangler tilstrekkelige årsaker. Allikevel synes tendensen til å sette likhetstegn mellom kausalitet og determinisme fremdeles å være tilstede.

4.3 Indeterminismen

Indeterminismen beskrives ofte som det motsatte av determinismen. William James beskriver tesen om indeterminismen med at delene til en viss grad er i “loose play” med hverandre:

“The laying down of one of them does not necessarily determine what the others shall be. It admits that possibilities may be in excess of actualities, and that things not yet revealed to our knowledge may really in themselves be ambiguous” (James, 1912, s. 150). Tesen om *indeterminismen* innebærer dermed at det finnes begivenheter som er *causa sui*⁸, dvs. ikke årsaksbestemte eller determinerte. Hendelser bare skjer, og det er ikke noe i verdens tidligere tilstand som forklarer dem. Teorier om indeterminisme er nært knyttet til idéer om usikkerhet. En sentral idé innen moderne kvantefysikk er at *indeterminerte hendelser* ikke er predikerbare fra tidligere hendelser, og at de i beste fall kan predikeres med sannsynlighetsberegninger.

Hvordan kan mennesket ha fri vilje hvis verden er deterministisk styrt? Dilemmaet med determinismen kan uttrykkes slik: Hvis man antar at en handling er slutten på en kausalt determinert kjede, og denne strekker seg tilbake i tid til hendelser aktøren ikke har noe oppfattet ansvar for, så kan ikke agenten sies å ha handlet fritt og være ansvarlig for handlingen. Indeterminismen synes heller ikke å kunne gi det vi trenger for å ha fri vilje; for hvis handlingen ikke er enden på en slik kjede, så har den oppstått på grunn av tilfeldigheter, og dermed kan ingen være ansvarlig for den.

Når filosofer argumenterer for at tesen om determinismen ikke er forenlig med fri vilje, kalles dette for *inkompatibilisme*. Hevder de tvert imot at tesen om determinismen er kompatibel med fri vilje kalles dette for *kompatibilisme*.

	Determinismen = sann	Determinismen = usann
Kompatibilisme (Fri vilje er kompatibelt med determinisme)	Myke determinister	
Inkompatibilisme (Fri vilje er ikke kompatibelt med determinisme)	Harde determinister	Libertarianere (Tror på fri vilje)

Tabell 3. Kompatibilisme/inkompatibilisme: Ulike posisjoner om fri vilje og determinisme.

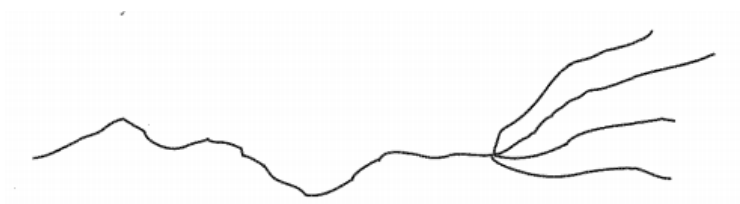
⁸ *Causa sui* (latin): Sin egen årsak, (noe) som årsaker seg selv.

Mange filosofer er enige om at fri vilje krever at mer enn én framtid står åpen for oss. Ideen om alternative muligheter blir gjerne illustrert med modellen ”The garden of Forking Paths”⁹. Dette er en hage med en sti som forgrener seg i ulike retninger som en metafor for tiden og hvordan livet vårt forløper med hensyn til mulige veivalg. Modellen benyttes ofte som et argument for inkompatibilistene, som hevder at determinisme ikke er forenlig med fri vilje. Jeg vil gå litt nærmere inn på en bestemt framstilling av argumentet, nemlig den av filosofen Peter van Inwagen (2002) i artikkelen *The Powers of Rational Beings: Freedom of The Will*. Deretter vil jeg sammenligne modellen med beskrivelsen av fenomenet ”snikende determinisme”, slik Fischhoff (1974) introduserer det i artikkelen *Hindsight ≠ Foresight: The effect of Outcome knowledge on Judgement Under Uncertainty*.

I tabell 3 så vi at inkompatibilister omfatter både *harde determinister* som ikke gir noe rom for eksistensen av fri vilje, og *libertarianere* som postulerer fri vilje. Selv tilhører Inwagen den siste gruppen. Det er ikke rom her for å yte rettferdighet til kompleksiteten i Inwagens teori, men kort oppsummert konkluderer han med at fri vilje er et mysterium som består uløst. Han velger derfor å konkludere med det han oppfatter som det *minst mystiske alternativet*: Nemlig at vi velger fritt, riktignok innenfor avgrensede områder. Dette gjør at Inwagen plasserer seg innenfor kategorien av en mykere variant av inkompatibilisme (libertarianisme). Kjerneargument hans er anerkjennelsen av at våre rasjonelle overveielser er betinget av *troen* på at vi har et reelt valg mellom ulike handlingsalternativer (Inwagen, 2002)

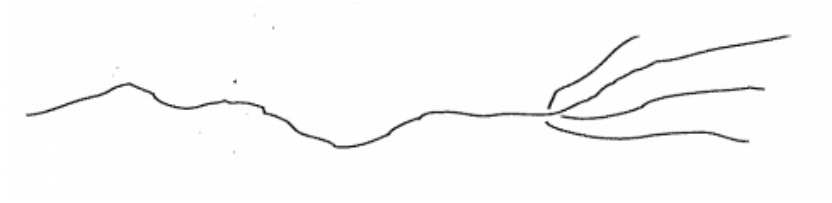
For å få en intuitiv forståelse av problemet med fri vilje og determinisme, henviser Inwagen til modellen ”garden of forking paths”, der vi ser for oss en hage med en sti som forgrener seg i ulike retninger (Inwagen, 2002, s. 202). Dette tjener da som en metafor for det kausale landskapet vi står overfor når vi (tilsynelatende) velger mellom ulike handlingsalternativer, og dermed ulike alternative framtider. Se figuren under:

⁹ Ideen og uttrykket ”The Garden of Forking Paths” kommer fra novellen med samme tittel skrevet av den argentinske forfatteren Jorge Luis Borge i 1941.



Figur 1 (Inwagen, 2002, s. 202)

Ifølge denne modellen har man fri vilje de gangene det er mer enn én vei i tidens vei som er åpen for en. Derimot har man ikke fri vilje hvis det er tilfelle at det bare finnes en vei å ta, og ingen alternativer. Determinismen er dermed tesen om at det er sant i ethvert øyeblikk at bare en av de alternative framtidene (à la forgreningene på modellen) er en *fysisk* mulig fortsettelse av tingenes tilstand i det øyeblikket. Inwagen forklarer videre at selv om det ser ut som vi ofte står overfor en rekke mulig framtider, slik det ser ut i figur 1, så er det ifølge tesen om determinismen kun denne realiteten vi står innenfor: (Se figuren under).



Figur 2 (Inwagen, 2002, s. 203)

Figur 2 er til forveksling til figur 1, men allikevel ikke helt, da tre av de fire grenene som leder vekk fra skillet, ikke lenger er forbundet med den originale veien, selv om de er svært nær den. De er derfor ikke egentlig forgreininger på veien, og de stedene hvor de nesten er i forbindelse med veien er det allikevel ikke et reelt veiskille. Inwagen fortsetter med å forklare at hvis vi ser denne figuren litt på avstand, for eksempel fra andre siden av rommet, så ser det ut som en sti som deler seg opp i flere stier som i figur 1. Men dersom vi studerer figuren nærmere, vil vi se at de tre grenene som på avstand synes å henge sammen med hovedstien, allikevel ikke er forbundet med den lange linjen eller med hverandre som i figur 2. Han forklarer: "The gaps between the long line and the unconnected lines represent causal discontinuities, violations of the laws of nature-in a word, miracles" (Inwagen, 2002, s. 204). Den første figuren viser en indeterministisk situasjon, mens den andre modellen viser en deterministisk situasjon. Inwagen er påpasselig med å påpeke at det er kun hvis universet er indeterministisk at tiden virkelig er en "hage med en sti som forgrener seg". Et annet sentralt poeng er at selv i et deterministisk univers, så kan tiden se ut som en hage med en sti som forgrener seg, hvis vi ser figuren litt på avstand.

Nå vil jeg sammenligne hagemetaforen til Inwagen med hvordan Fischhoff (1974) beskriver fenomenet med ”snikende determinisme”. Mens Inwagen forklarer en metafysisk tese om virkeligheten, så viser Fischhoff til og tolker empiriske funn om visse underliggende antagelser vi gjør ved retrospeksjon. Inwagen og Fischhoff representerer to forskjellige vitenskapelige disipliner, men det er allikevel noen berøringspunkter som kan være verdt å merke seg. Likheten er at vi har en noenlunde lik modell av tid og (tilsynelatende) alternative handlinger å velge mellom. Hagemetaforen begrenser seg til nåtiden og framtiden: Punktet der veien deler seg i flere retninger, viser ifølge Inwagen til nåtiden, og de forskjellige alternative veiene er framtiden (Inwagen, 2002, s. 204). Men ifølge tesen om determinismen er ikke den temporale dimensjonen relevant. I et deterministisk perspektiv er det likegyldig hvorvidt det dreier seg om retro- eller prospeksjon, da forholdet mellom disse er symmetrisk. Framtiden og fortiden stiller dermed likt med hensyn til det å være fastlagt.

Forskjellen ligger i at mens Inwagen benytter hagemetaforen for tid for å illustrere hvordan vi kan komme til å tro at et deterministisk univers er indeterministisk, viser Fischhoff hvordan vi er tilbøyelige til å oppfatte det motsatte, dvs. tendensen til det han kaller for ”snikende determinisme”. Fischhoff sin studie viser hvordan ”sikten” vår hemmes av den temporale avstanden når vi ser tilbake på hendelser, fordi vi ikke er klar over hvordan kunnskap om utfallet påvirker analysene våre av fortiden. Tendensen ved retrospeksjon er at vi antar at det kausale landskapet ligner på Inwagens figur 2, der det er et gap mellom de alternative stiene og hovedstien. Hvis vi derimot hadde kunnet frigjøre oss fra utfallskunnskapen og sette oss inn i det stedet vi befant oss i fortiden der framtiden var ukjent, så ville vi ifølge Fischhoff innse at det var figur 1 som var den riktige. Inwagen viser på den andre siden hvordan sikten vår hemmes, fordi vi aldri kommer nærme nok til å se at det vi oppfatter som veiskiller egentlig ikke er det.

Fischhoff uttrykker ikke eksplisitt hvilken oppfatning han har om virkelighetens beskaffenhet, dvs. hvorvidt den er deterministisk eller indeterministisk, men han er opptatt av hvordan vi ved retrospeksjon tillegger hendelser en slags indre nødvendighet og dermed utelukker alternative kausalkjeder. Men ettersom Fischhoff er opptatt av at vi potensielt sett kan lære av fortidens hendelser, er det grunn til å anta at han oppfatter virkeligheten som indeterministisk.” It is argued that this lack of awareness can seriously restrict one's ability to judge or learn from the past” (Fischhoff, 1974, s. 2). Det er vanskelig å se hvordan en (hard)

determinist ville hevde at læring kan finne sted, da et deterministisk verdensbilde ikke synes å gi rom for at individer kan utgjøre en årsak eller skape en forandring. Selv om dette må sies å avhenge av hva slags læringsbegrep man holder seg med. Det er uansett verdt å merke seg at det er ulike varianter av å kombinere en tese om determinismen med eksistensen av fri vilje, og at det dermed ikke er noen faste prinsipper i denne debatten.

Fischhoff påpeker at ”snikende determinisme” som fenomen står i kontrast til den typen filosofisk determinisme som Inwagen omtaler, der det handler om den bevisste oppfatningen at hva som enn skjer må skje. Fischhoff forklarer forskjellen mellom den ”filosofiske deterministen” og den ”snikende deterministen” slik: Den ”filosofiske deterministen” tror at et utfall var uunngåelig, fordi det som skjer uansett er uunngåelig. Han kan allikevel være overrasket over utfallet: ”Indeed, he may well set for himself the task of researching the situation until its inevitable character becomes apparent” (Fischhoff, 1974, s. 3). Om den ”snikende deterministen” sier han: ”The creeping determinist may well be totally unaware of the raging debate over historical inevitability and free will” (Fischhoff, 1974, p. 3). Det er en kjent sak at man ikke behøver å kjenne til den filosofiske tesen om determinismen for å anta at fortidens hendelser har skjedd på grunn av en indre nødvendighet.

4.4 Problemet med determinismen

Med tanke på oppgavens hovedtema, er det på sin plass å påpeke at tesen slik Inwagen (2002) legger den fram her, er en typisk kontrafaktisk betingelse, og at tesen om determinismen således innebærer en transcendental betingelse. Et problem med determinismen er derfor at det ikke er mulig for oss å kontrollere om det som skjer skyldes det som Inwagen kaller et mirakel eller om det er ufravikelige lover som utspiller seg, fordi det alltid kan finnes en mulighet for at det som skjer har sin årsak i noe vi ikke har oversikt over. Det er en ontologisk påstand om virkeligheten som ikke kan testes ut empirisk. Selv om vi en vakker dag skulle komme så langt at vi mener å ha identifisert alle de kausale lovene, kan vi ikke vite med sikkerhet om vi faktisk har det. Dette kan også kalles det epistemiske problemet med tesen, som innebærer at vi alltid synes å befinne oss mer eller mindre på den andre siden av rommet til Inwagen, dvs. at vi ikke kommer nærmere nok til at vi kan studere alle de kausale sammenhengene, og identifisere alle naturens lover som konstituerer virkeligheten. Determinisme lar seg dermed verken bevise eller motbevise. Det er imidlertid ikke bare tesen

om determinismen som kan beskyldes for dette. Patrick Suppes sier følgende: "Deterministic metaphysicians can comfortably hold to their view knowing they cannot be empirically refuted", og legger så til: "...but so can indeterministic ones as well" (Suppes, 1993, s. 254). Både determinismen og indeterminismen kan, som metafysiske teser, argumenteres å være transcendentale: "...any general thesis about the nature of the universe must transcend available scientific facts and theories by a very wide mark" (Suppes, 1993, s. 254).

Fischhoff studerte hva som skjer når vi ser vi tilbake på både den stien vi tok, og de alternative stiene vi ikke tok. Da er tendensen at vi sammenholder en rekke punkter eller hendelser som ledet til et visst utfall, og tillegger de en særskilt forbindelse, som om de er forbundet med en indre kausal nødvendighet. Enkelt sagt innebærer dette en oppfatning av at kun det som skjedde, kunne ha skjedd. Vi har altså en tilbøyelighet til å gi de faktiske hendelsene og dermed det empiriske forrang, framfor kontrafaktiske scenarioer som ikke kan undersøkes empirisk. Eller rettere sagt, så trekker vi konklusjoner om kausal nødvendighet ut fra hverdagsempirien. Når vi ser framover i tid, er tendensen at vi forestiller oss en rekke tilfeldig spredte punkter som potensielt sett kan utgjøre alternative og ulike kausale kjeder med mange mulige utfall. Fischhoff henviser til hvordan historikeren Georges Florovsky (1969) påpeker at tendensen om determinisme er implisert i den retrospektive metoden:

The tendency toward determinism is somehow implied in the method of retrospection itself. In retrospection we seem to perceive the logic of events, which unfold themselves in a regular order, according to a recognizable pattern, with an alleged inner necessity, so that we get the impression that it really could not have happened otherwise (Florovsky, 1969, s. 364).

Men det er ikke bare ved retrospeksjon at vi antar et deterministisk rammeverk. Selve prosessen med å forklare noe i termer av kausale sammenhenger synes å forutsette et deterministisk rammeverk. I kapittel 4.2 så jeg nærmere på assosiasjonen som ofte gjøres mellom kausalitet og determinisme. Den bakenforliggende antagelsen er at en hendelse har tilstrekkelige årsaker. Da er det ikke så merkelig at tendensen til å anta prinsippet om at en hendelse har tilstrekkelige årsaker også er implisert i den retrospektive metoden. Fortidens hendelser er nettopp de vi ønsker å forklare, mens framtidige hendelser er gjenstand for våre prediksjoner og overveielser. Som Fischhoff sier, så er selve studiet av fortiden predikert av troen på at hvis vi ser etter, vil vi kunne avdekke noen mønstre som kan fortolkes (Fischhoff,

1974). Dette ligner Inwagen sin redegjørelse for at, dersom tesen om determinismen er sann, så vil denne sannheten kunne bekreftes hvis vi kan ta hendelsenes kausale forløp i nærmere øyesyn. Vi synes dermed å anta en asymmetri mellom fortid og framtid, dvs. at fortiden karakteriseres ved nødvendighet, mens framtiden er åpen. Som tidligere nevnt er det tilsvarende forholdet mellom fortid og framtid i et *deterministisk* verdensbilde derimot fullstendig symmetrisk, da både fortid og fremtid karakteriseres ved nødvendighet.

Betyr dette at vi antar diametralt ulike teser om virkeligheten med hensyn til fortiden og framtiden? Dvs. at fortidens univers er deterministisk, mens framtidens univers er indeterministisk? Da lever vi i to inkommensurable univers, som ikke følger de samme naturlovene, i stedet for ett sammenhengende. Ut fra et epistemologisk perspektiv om hva vi kan vite om verden, virker det lønnløst å tro at det som fungerer i fortidens *deterministiske* univers, er relevant for hva som kan fungere i framtidens *indeterministiske* univers. Men ifølge de empiriske funnene til Fischhoff, er det delvis slik vi opererer. Hvis den snikende deterministiske tesen vi legger til grunn ved retrospeksjon er i nærheten av å se ut som den filosofiske tesen om determinismen, må vi enten forkaste den, eller også anlegge en deterministisk tese for framtiden. Det siste forslaget er som kjent svært kontraintuitivt, fordi det meste av den menneskelige virksomheten som betyr noe for oss, og som regulerer samlivet mellom oss, nettopp er idéene om frihet og ansvarlighet. Alternativt kan vi begynne å se fortiden som like åpen som framtiden, dvs. å ta innover oss at enhver tid vi måtte tenke tilbake på, stod vi overfor en ukjent fremtid. Det er dette som er Fischhoff sitt prosjekt; å vise hvordan en uoppdaget snikende determinisme på en alvorlig måte svekker evnen vår til å vurdere fortiden eller lære av den (Fischhoff, 1974). Men det er allikevel en sentral forskjell i og med at vi ikke kan gjøre fortiden ugjort. Som Aristoteles sier:

[...] what has happened cannot be made not happened. Hence Agathon is right in saying

This only is denied even to God,
The power to make what has been done undone (Aristotle, 1926, VI, 6, 1139b)

Allikevel er det et faktum at vi gjennom kontrafaktisk tenkning gjerne endrer på fortiden i det mentale. Det karakteristiske for den mentale omgjøringen er at noen ting synes det helt selvfølgelig å forandre på, mens andre rører vi ikke. Hofstadter spør: "What features do we let

slip? What ones do we not even consider letting slip?” (Hofstadter, 1979, s. 638). Hofstadter spør rett og slett hvorfor det er noen kontrafaktisiteter som: “...strike us as ‘less counterfactual’ than other counterfactuals?” For som han sier videre, så er det åpenbart at det som ikke skjedde ikke skjedde, og at ”there aren’t degrees of ‘didn’t-happen-ness’” (Hofstadter, 1979, s. 638). I kapittel 2.2.4 fulgte jeg opp dette spørsmålet i en vitenskapsteoretisk retning, der jeg drøftet noen ulike synspunkter på hva som skiller det faktiske fra det kontrafaktiske. I neste kapittel vil jeg ta spørsmålet videre i en pedagogisk og moralfilosofisk retning.

4.5 Ekstern kompatibilisme og pedagogisk filosofi

I dette kapittelet fortsetter jeg tråden fra forrige kapittel om fenomenet med at det er visse kontrafaktisiteter vi oppfatter som mindre kontrafaktiske enn andre. Det er vanskelig å si seg uenig med Hofstadter når han påpeker at det er åpenbart at det som ikke skjedde ikke skjedde, og at det ikke finnes grader av ”didn’t-happen-ness” (Hofstadter, 1979, s. 638). Hofstadter spør om hvilke spesifikke deler av virkeligheten vi med letthet omgjør mentalt, og hvilke vi absolutt ikke kan tenke oss annerledes. Det er som om vi, foruten å ha de fysiske naturlovene, har egne (”natur”-) lover for det imaginære. Han sier: ”We select from our fantasy a world which is close, in some internal mental sense, to the real world” (Hofstadter, 1979, s. 640). Vi ser hvordan dette ligner Lewis teori om *mulige verdener*, som jeg redegjorde for i kapittel 2.2.4. I fortsettelsen ønsker jeg å drøfte noen hypoteser om hvordan kompatibilistene og mange med dem fastholder at virkeligheten er plastisk i visse henseender, mens den ikke er det i andre. Spørsmålene om hvilke trekk ved virkeligheten vi synes det er greit å gjøre om på mentalt og hvilke trekk det ikke er greit å forandre på, griper rett inn i kjernen av metafysikkens fri vilje-problem, og overlapper samtidig med noen av de mest sentrale spørsmålene innen pedagogisk filosofi. Dette kommer til syne først og fremst i argumentene mellom inkompatibilistene og kompatibilistene. Tidligere så vi at kompatibilistene hevder at vi har fri vilje samtidig med at vi lever i et deterministisk univers. Derfor må de svare på hva denne friheten består i. Implisitt i dette, ligger også svaret på spørsmålet om hvorfor vi omgjør handlinger og hendelser mentalt. Før jeg kommer til det, ser jeg igjen på spørsmålene Hofstadter stiller: “What events are perceived on some deep level as being close relatives of ones which really happened?” (Hofstadter, 1979, s. 638).

Hofstadter tar utgangspunkt i at vi bygger den mentale representasjonen av en situasjon lag på lag. Han låner termene konstant, parameter og variabel fra matematikken for å belyse hvordan det er et hierarki ved ulike aspekter i en situasjon. K er konstanten, p er parameteren og v er variabelen. Alle kan variere, men det er et hierarki ved variabiliteten. I en situasjon representerer k den globale tilstanden og det dypeste laget, som vanligvis ikke forandrer seg i det hele tatt. Mens p representerer en mindre global tilstand, holdes k fast. Et eksempel på k er gravitasjonskreftene, som vi sjelden omgjør mentalt. Et annet eksempel er at vi oppfatter tilværelsen som tredimensjonal. Hofstadter kaller de dypeste aspektene ved en situasjon som konstante konstanter. Så finnes det faste konstanter k , som allikevel ikke er konstante på en permanent basis. Hofstadter foreslår at det kan være "background assumptions-things" som vi et sted vet kan forandre seg, men som vi mesteparten av tiden ikke stiller spørsmålstegn ved (Hofstadter, 1979, s. 641). Disse er dermed også en type konstanter (k). Hofstadter foreslår reglene i fotball, og jeg foreslår som et eksempel, bestemmelsen av antall år barn skal gå på grunnskolen, antall skoledager i året og antall timer per dag. Så er det parametere vi tenker på som variable, men som vi holder konstant. Det kan være hvilke lærere man har, eller oppgavene til eksamen. Til slutt er det variablene vi med letthet bytter ut. I og med temaet for oppgaven foreslår jeg at de tingene vi med letthet bytter ut, er små og store ting vi tenker at vi kunne ha gjort annerledes. Det kan også dreie seg om avgjørende valg vi tok, både der vi uventet gikk mot strømmen og situasjoner der det ikke var opplagt for noen hva valget skulle bli.

Hofstadter utvikler teorien videre som han kaller "nested frames", om hvordan vi for eksempel oppfatter personer kontekstuell. Dersom vi tenker på en person i lærerrollen, har vi lag-på-lag av informasjon eller såkalt kontekstuell kunnskap som skaper forventninger: "One of the main ideas about frames is that each frame comes with its own set of expectations". Men et annet like sentralt poeng han fremhever, er at en ramme inneholder, sammen med de fastsatte forventningene, også kunnskap om: "...its limits of applicability, and heuristics for switching to other frames in case it has been stretched beyond its limits of tolerance". Han sammenligner det med en zoom, der vi kan zoome ut og inn i en situasjon og undersøke ulike rammer, lag eller aspekter ved den. Jeg forslår eksempelet med en lærer som reflekterer over en situasjon med en elev som har slitt med å følge undervisningen. La oss si at læreren i utgangspunktet ikke har hatt høye forventninger til eleven på grunn av omstendigheter som oppfattes som konstante og dermed fastlåste. Ved den retrospektive metoden kan læreren

zoome ut og inn på ulike lag og rammer i situasjonen, og kanskje finne noen variabler *v* av den foranderlige typen, som tilsynelatende hadde kunne påvirke totalbildet. En slik undersøkelse eller tankeeksperiment har sin forutsetning i en kunnskap om at konteksten eller rammen vi forstår en situasjon innenfor kan utfordres. Men sier ikke dette bare at det faktiske vi i en del tilfeller tenker kontrafaktisk om, ikke er så faktisk som først antatt? Vi har dermed foreløpig ikke kommet videre med spørsmålet til Hofstadter. Læreren kunne ikke ha handlet på en annen måte uten den nye innsikten, og innsikten kom på bakgrunn av et utfall eller en faktisk situasjon. hadde ikke kunne handle annerledes før han eller hun kom til innsikt om det gjennom å undersøke de fortidige hendelsene. Så langt har Hofstadter sagt at når vi undersøker fortiden kontrafaktisk, så forutsetter vi at noen av rammene vi forstår den innenfor, ikke er tilstrekkelige, dvs. at dette er en del av den kunnskapen som ligger i fortolkningene våre. Jeg er fristet her til å stille spørsmålet om hvorfor vi ikke heller på en forenklet måte kan si at det er fortiden selv som ikke strekker til for oss, og ikke fortolkningen vår av den. Et viktig poeng er uansett at læreren ikke kan endre på fortiden som sådan, men at han eller hun kan ta innsikten med seg i fortsettelsen.

Videre ser jeg på hvordan kompatibilistene kan tenkes å svare på spørsmålet om hva den påståtte friheten består i, og herunder får vi kanskje svar på hvorfor vi omgjør handlinger og hendelser mentalt. I kapittel 4.3 så vi at forsvarerne av kompatilismen opprettholder at vi har fri vilje i et deterministisk univers. De mener dermed at en framtid, for å kunne være åpen, nødvendigvis må være fysisk mulig. Ifølge kompatibilistene er en framtid fysisk mulig: "...if everything that happens in it is permitted by the laws of nature" (Inwagen, 2002, s. 205). I og med at determinismen hevder at det kun finnes en framtid som har en mulig fysisk forbindelse med nåtiden, byr det på problemer å hevde at framtiden i noen forstand er åpen. Men kompatibilistene mener allikevel at det er noen framtider som er åpne for oss, selv om disse framtidene ikke er forbundet med nåtiden ved naturens lover. Inwagen påpeker at kjernen ved kompatibilistens forsvar, er et forsøk på å vise at grunnen til at vi er interessert i åpne eller tilgjengelige framtider, er at vi er interessert i å modifisere atferden til mennesker (Inwagen, 2002, s. 206) Det er nok andre kompatibilister som vil innvende at vi dypest sett *er* moralske vesener, og at interessen for å modifisere atferd heller bør betraktes som et biprodukt enn et hovedanliggende. Kompatibilistenes omforente løsning på at vi har fri vilje i et ellers deterministisk univers, kan sies å være at vi foretar valg og at disse er frie. Inwagen forsetter sitt resonnement med at en slik modifisering kan skje gjennom belønning og straff. Men

samtidig påpeker han at det er meningsløst å få mennesker til å handle på en måte som uansett ikke er åpne for dem. Det må være noen kriterier for hva som er karakteristisk for åpne framtider og for hva som ikke er det. I likhet med Hofstadter påpeker han at det finnes ting vi aldri ville drømme om å tenke at vi kan forandre på:

There are certain facts that no human being can do anything about~and that no human being in history could ever have done anything about. Among these are the fact that the earth is round, the fact that magnets attract iron, the fact that there were once dinosaurs, and the fact that 317 is a prime number (Inwagen, 2002, s. 207)

Det ligger i sakens natur at når man bedriver atferdsmodifikasjon, så antar man at den ønskede atferden ligger innenfor det som faktisk er mulig for en person å gjøre. Inwagen fortsetter med svare for kompatibilisten, med at man modifierer atferden til folk ved å modifisere valgene de tar. Dette prosjektet kan bare lykkes i den grad man oppnår den ønskede effekten. Det ville for eksempel være absurd å påvirke noen til å velge ikke å være gjenstand for gravitasjonskreftene, fordi de uansett ville mislykkes (Inwagen, 2002, s. 207). Dette prosjektet kan bare lykkes i den grad man oppnår den ønskede effekten. Det ville for eksempel være absurd å påvirke noen til å velge å ikke være gjenstand for gravitasjonskreftene, fordi de uansett ville mislykkes (Inwagen, 2002, s. 207).

Vi har nå kommet til det punktet der vi kan se hvordan dette området av metafysikkens fri vilje-problem, overlapper med noen av de mest sentrale spørsmålene innen pedagogisk filosofi. Dette ligger i argumentene mellom inkompatibilister og kompatibilister. Det kanskje viktigste argumentet for inkompatibilistene mot kompatibilistene, er det såkalte *manipulasjonsargumentet*. Inkompatibilistene viser til at kompatibilistene ikke skiller på ulike kausale historier. En kausal historie er ganske enkelt kjeden av hendelser og handlinger i en persons liv. Hvis en person har vært utsatt for manipulasjon eller indoktrinering kalles det i denne terminologien for en manipulert kausal historie. Jeg viser her til en drøfting Stefaan Cuypers gjør i bokkapittelet *Educating for authenticity: The paradox of moral education revisited* (Cuypers, 2009). En person med en såkalt manipulert kausal historie vil på den ene siden kunne være overbevist om at den har et større ansvar (og dermed større frihet) for ulike hendelser i sitt liv enn den egentlige hadde, og dermed trekke gale konklusjoner om kausale sammenhenger. Likeledes vil en person med en annen type manipulert kausal historie kunne trekke konklusjoner i motsatt retning. Da er personen overbevist om at han eller hun har mindre ansvar (og mindre frihet), enn han eller hun egentlig hadde. Det er imidlertid problematisk å henvise til noe ”egentlig” i denne sammenhengen, da det forutsetter at vi har

kriterier for å bestemme hvilke ting som var opp til oss og hvilke ting som ikke var det. Selv om det vanligvis ikke drøftes så direkte som her, må dette sies å være et av kjernetemaene innen pedagogisk filosofi. I den generelle læreplanen er det et gjennomgangstema at elevene skal ansføres til eller utvikle ulike former for ansvarlighet. Det innebærer kort sagt at de etter endt skolegang skal se seg selv som opphavet til en rekke handlinger. Dette kommer jeg tilbake til i kapittel 5.2.

La oss si at vi kan skille mellom kausale historier, og dele de inn i manipulerte kausale historier og vanlig determinerte kausale historier. Det er viktig å huske på at kompatibilistene opprettholder tesen om determinismen, og at vi er fri uavhengig av hvordan virkeligheten utspiller seg. En manipulert kausal historie som har kommet i stand gjennom for eksempel manipulasjon eller indoktrinering blir vanligvis sett på som å underminere fri vilje og ansvarlighet. Det er her inkompatibilistene har sitt hovedargument mot kompatibilistene. De hevder at det (for kompatibilister) ikke er noe som skiller en “vanlig” determinert kausal historie fra en såkalt manipulert historie: i begge tilfellene mangler aktøren kontroll over handlingen, hvis årsaker stammer fra fortiden og naturens lover (Cuypers, 2009, s. 123). Kompatibilistenes svar på dette har ført til en todeling blant kompatibilistene: På den ene siden har vi *internalistene*, som opprettholder at den kausale historien med de ulike psykologiske elementene uansett er irrelevant for fri vilje. På den andre siden har vi *eksternalistene* som bekrefter at det spiller en rolle om de kausale kildene til de psykologiske elementene har opphav i manipulasjon (Cuypers, 2009). De mener at man må skille mellom kausale kilder som undergraver autonomi og de som er ”naturlige” og bevarer autonomi. Jeg oppfatter at mange av diskusjonene innen pedagogisk filosofi kan kanskje plasseres innen *ekstern kompatibilisme*, der man drøfter virkningen av ulike kausale historier med hensyn til utviklingen av autonomi. Det er samtidig et paradoks, dersom noen av disse drøftingene bygger på antagelsen om et radikalt autonomibegrep, der autonomi anses å være en absolutt metafysisk og *ahistorisk* størrelse. En slik antagelse innebærer at vi er frie uavhengig av tid og rom, og dermed er den kausale historien ikke relevant. I det perspektivet synes det meningsløst hva som er viktig for barns utvikling med hensyn til autonomi. Men dette er samtidig noe av kjernen i det pedagogiske paradokset. Løvlie tar opp akkurat dette spørsmålet i tilknytning til det å anta at barnet er autonomt:

Det er nesten frekt å si at eleven er autonom når vi alle vet at barn er helt avhengige av voksnes omsorg, støtte og veiledning. På den andre side, hvis vi nekter å tenke kontrafaktisk og forkaster ideen om barnets autonomi, har vi svake barrierer mot å ta eleven som et produkt av disiplin, kunnskap og øvelse... (Løvlie, 2011, s. 784)

I denne delen har vi sett at det metafysiske temaet *fri vilje-determinisme* er stort og komplekst, og det er nærmest ingen grenser for hvor finmasket man kan gå til verks i debatten. Jeg har valgt meg ut visse deler av temaet for å belyse det jeg oppfatter som vesentlige aspekter ved kontrafaktisk tenkning i en pedagogisk kontekst. Vi har sett at fri vilje-determinisme kontroversen innen filosofien har sitt opphav i divergerende tolkninger om oss selv, og om hvordan verden rundt oss virker. Det at vi på den ene siden ser oss selv som frie og ansvarlige for handlingene våre, står tilsynelatende i konflikt med en oppfatning om at det som skjer er årsaksbestemt på en slik måte at alt skjer med nødvendighet. Det metafysiske problemet med den frie viljen, handler dermed om å lokalisere handlingene våre i strømmen av hendelser som resten av universet utgjøres av. Selv om det ikke kan settes et likhetstegn mellom determinisme og kausalitet, er det allikevel en tendens til å anta et deterministisk rammeverk når vi beskriver kausalitet.

Inwagen illustrerer hvordan vi kan komme til å tro at et deterministisk univers er indeterministisk, mens Fischhoff viser hvordan vi er tilbøyelige til å oppfatte et indeterministisk univers som deterministisk. Videre har jeg drøftet tendensen til å anta ulike teser om virkelighetens beskaffenhet med hensyn til fortid og framtid. Det ledet til en drøfting av hvilke ting vi omgjør mentalt, og hvilke ting vi finner det naturlig å holde konstant. Det finnes både en pedagogisk og moralfilosofisk gevinst i å se nærmere på dette spørsmålet. Innen kompatibilismen og i mange henseende *common sense*, er det vanlig å hevde at det er valgene våre og dermed friheten vår som utgjør de variable delene av virkeligheten. Inwagen hevder at dette kommer av et behov for å modifisere atferd. Herfra tok jeg veien via kompatiblistenes forsvar for sin posisjon, og viste hvordan en del av fri vilje-determinisme debatten overlapper med det pedagogiske paradokset som man drøfter innen pedagogisk filosofi.

5 Frihet

Frihet er et flertydig begrep som brukes i vidt forskjellige sammenhenger. Selv om de fleste har en relativt klar intuitiv forståelse av hva det vil si å være fri, lar det seg vanskelig gjøre å drøfte spørsmål vedrørende frihet uten først å foreta en rekke begrepsmessige kategoriseringer og avgrensninger. I det berømte essayet *Two Concepts of Liberty* forteller Isaiah Berlin (1969) at det er dokumentert over 200 betydninger av begrepet “liberty”. I fortsettelsen vil jeg redegjøre for noen sentrale dimensjoner av frihetsbegrepet ved hjelp av teoriene til Immanuel Kant, Jean-Paul Sartre og Maurice Merleau-Ponty. I forlengelsen av frihetsbegrepet redegjør jeg for autonomibegrepet og drøfter sentraliteten av kontrafaktisk tenkning for idealet om autonomi.

5.1 Frihetens ontologiske dimensjoner

Frihet er et flerdimensjonalt fenomen som kan omfatte for eksempel ontologisk, politisk og personlig frihet. *Moralsk frihet* er nært knyttet til et begrep om moralsk ansvar, da den moralske siden av denne type frihet innebærer at du kan stilles ansvarlig for handlingene dine og bli møtt med reaktive holdninger som ros og klander. *Moralsk frihet* kan også langt på vei hevdes å utgjøre en sentral kategori av frihet som er underliggende de andre dimensjonene, da den moralske siden ved menneskelivet har vært en viktig motivator for teorier om frihet. Men det er allikevel ikke en entydig oppfatning, da det finnes de som hevder at mens frihet er en metafysisk størrelse, så sorterer moralske spørsmål under etikken. Videre finnes det de som hevder at frihet ikke er en ontologisk størrelse, så tvetydigheten ved frihetsbegrepet må sies å være en alltid nærværende karakteristikk av det. *Politisk og personlig frihet* kan sies å omhandle frihetens konkrete dimensjoner og ulike hindre for frihet kan være både materielle og ikke-materielle størrelser. Det kan være ytre fysiske forhold som hindrer oss i å handle slik vi ønsker, eller det kan være “indre” forhold som for eksempel forpliktelser og vaner som virker betingende for handlingene våre. Spørsmål om *frihetens ontologiske dimensjon*, dvs. om vi *er* frie av natur, handler samtidig om grunnleggende vilkår for handlingslivet vårt. Men selv om frihet regnes for å være grunnleggende for hva det vil si å være menneske, er det delte meninger om hva dette grunnleggende består i. Innen den filosofiske diskusjonen om fri vilje og moralsk ansvarlighet, er det ulike syn på de ontologiske forutsetningene for dette. Jean-Paul Sartre hevder i *Being and Nothingness* fra et eksistensialistisk ståsted at mennesket

nødvendigvis alltid er fritt, og at det simpelthen ikke er mulig å feile i dette (Sartre, 1992). Immanuel Kant argumenterte også for en kategorisk frihet, men ut fra et annet ontologisk grunnsyn. Han gjorde rom for frihet og ansvarlighet ved å postulere muligheten av en noumenal verden bortenfor fenomenale framtredelelser. I *Critique of Practical Reason* (2002) tar Kant utgangspunkt i de intuitive moralske vurderingene som forutsetter kapasiteten til “å kunne ha handlet annerledes”, og legitimerte derigjennom den frie viljen. Gjennom å argumentere transcendentalt for fri vilje, kan Kant samtidig opprettholde en ubønnhørlig deterministisk tese om den naturlige verden. Den deterministiske tesen om at alt som skjer er bestemt av forutgående betingelser gjennom unntaksløse årsaksforhold, kommer ifølge Kant ikke i veien for at mennesket kan handle kausalt fritt, eller kontrakausalt som noen kaller det. Det innebærer at en handling ikke er årsaksbestemt av forutgående hendelser, men igangsetter en helt ny rekke av hendelser. Ved å trekke opp grensene for erfaringen, viste Kant at vi ikke kan ha sikker viten eller teoretisk kunnskap om frihet og andre transcendentale objekter som Gud, den udødelige sjelen. Løsningen til Kant var at ettersom dette samtidig innebærer at vi heller ikke kan ha viten om at transcendentale objekter *ikke* eksisterer, gir det åpning for en praktisk tro på at transcendentale objekter eksisterer. Blant de transcendentale objektene er det ifølge Kant allikevel bare den frie viljens aktualitet eller mulighet, vi kan sies å ha denne type innsikt i: “But freedom, among all ideas of speculative reason, is also the only one whose possibility we know a priori-though without having insight into it-because it is the condition of the moral law, which we do know” (Kant, 2002, s. 5:4). I denne sammenhengen er det verdt å legge merke til den kontrafaktiske naturen til transcendentale argumenter. I denne sammenhengen er det verdt å legge merke til den kontrafaktiske naturen til transcendentale argumenter (Kannisto, 2012, s. 243). Et transcendentalt argument tar utgangspunkt i et akseptert, universelt faktum og argumenterer at hvis noe ikke var tilfelle, ville det faktumet være umulig. Et direkte eksempel på dette hos Kant kan vi se i en fotnote til et avsnitt i *Critique of Practical Reason*: “But if there were no freedom, then the moral law could not be encountered in us at all” (Kant, 2002, s. 5:4). For en del moderne filosofer er det nettopp det moralske perspektivet som avgjør hvilken posisjon de inntar i debatten om fri vilje. Som en motreaksjon til nyere nevrofysiologisk forskning eller andre reduksjonistiske tilnærminger som ofte hevder at den frie viljen er en illusjon, så mener enkelte filosofer at den frie viljen må “reddes”. Dette var også den bakenforliggende motivasjonen til Kant da han satte seg fore å gi metafysikken en vitenskapelig status. Prosjektet til Kant var å danne metafysiske teser som var like velbegrunnede som den klassiske naturvitenskapens teser, slik at han kunne

redde menneskets frie og ansvarlige natur. At man mener at det er nødvendig å redde oppfatningen om mennesket som fritt og ansvarlig, er kanskje ikke så rart med tanke på hvor grunnleggende den normative og moralske dimensjonen ved menneskelivet synes å være. Og om det mot all formodning skulle være mulig å aktivt velge bort denne dimensjonen, ville en slik beslutning kreve en slags norm med henvisning til et bestemt syn på menneskets natur. Vi kan sammenligne dette med at man i et demokratisk samfunn avholdt valg for å komme fram til om man skulle oppløse det samme demokratiet. Hvis vi aktivt oppløser en “rådende” normgivende instans, må det nødvendigvis skje på bakgrunn av en tilslutning til en annen normgivende instans.

Hos Sartre er nettopp dette et sentralt argument for en radikal frihet, at det å fornekte frihet forutsetter friheten: “We said that freedom is not free not to be free and that it is not free not to exist. This is because the fact of not being able not to be free is the facticity of freedom, and the fact of not being able not to exist is its contingency” (Sartre, 1992, s. 625). Maurice Merleau-Ponty (1962) er på den andre siden et eksempel på en filosof som mener at det er et problem med å slutte seg til og hevde en ontologisk frihet bortenfor det vi erfarer. I *Phenomenology of Perception* påpeker han at kategorisk eller ontologisk frihet innebærer en begrepsmessig motsetning. Han viser til at dersom vi er like frie uansett hva vi føler eller gjør, ville ikke frihet som fenomen framtre for oss: “[...] since free action, in order to be discernible, has to stand out against a background of life from which it is entirely, or almost entirely, absent. We may say in this case that it is everywhere, but equally nowhere” (Merleau-Ponty, 1962, s. 437). Dette er en kritikk særskilt rettet mot Sartres påstand om radikal frihet, men også mot annen teoretisk tenkning med rasjonalistiske tendenser, der frihet i stedet for å bli behandlet som et konkret fenomen, blir gjort til en abstrakt størrelse som befinner seg utenfor det vi oppfatter som begrensinger i det kausale landskapet. Merleau-Ponty mener at det klassiske dilemmaet med frihet som enten total eller ikke-eksisterende, er et produkt av objektiv tenkning og analytisk refleksjon (Merleau-Ponty, 1962, s. 454). Jeg tolker han dithen at han mener at dette ikke er veien å gå, fordi det gir oss et problem med frihet som vi egentlig ikke har. I likhet med Kant går Merleau-Ponty gjennom den perseptuelle erfaringen for å begrunne frihet, men der Kants forklaring av persepsjon tar utgangspunkt i en analyse av hvordan *sinnet* prosesserer ulike representasjoner som intuisjoner og begreper, så analyserer Merleau-Ponty perseptuell erfaring som *kroppslig* kontakt med verden. Selv om Merleau-Ponty står langt fra Kant’s transcendentale idealisme, er han allikevel ikke helt fritatt for å stille kontrafaktiske betingelser. Det ser vi for eksempel

når han påpeker at dersom friheten var total eller ikke-eksisterende, ville den ikke framtre for oss (Merleau-Ponty, 1962, s. 437).

Sartre og Merleau-Ponty har til felles at de tar utgangspunkt i en eksistensiell fenomenologi, som søker å fortolke opprinnelige strukturer ved menneskets “væren-i-verden” som går forut for tankemessig refleksjon og vitenskapelige forklaringer. Begge er opptatt av situert frihet, dvs. at friheten framtrer i gitte begrensende betingelser. Sartres fortolkning av situasjonen knyttet til frihetsbegrepet, er imidlertid svært annerledes enn hvordan Merleau-Ponty ser på det. Sartre hevder at friheten kommer til syne for oss nettopp på grunn av det endelige målet om frihet. Han påpeker frihetens paradoks slik: “There is freedom only in a situation, and there is a situation only through freedom” (Sartre, 1992, s. 629). Han sier videre at: “The given in itself as resisting or as aid is revealed only in the light of the projected freedom” (Sartre, 1992, s. 629). Men der Sartre hevder at motstand og hindringer bare har mening i og gjennom det frie valget som er menneskets grunnleggende natur, ser Merleau-Ponty på forholdet vårt til situasjonen helt annerledes. Han sier at: “[...] our freedom does not destroy our situation, but gears itself to it” (Merleau-Ponty, 1962, s. 442). Ifølge Merleau-Ponty er vi allerede situerte i en fysisk og sosial verden, før eventuelle valg: “I am given to myself, which means that this situation is never hidden from me, it is never round about me as an alien necessity [...]” (Merleau-Ponty, 1962, s. 360). I motsetning til Sartre mener Merleau-Ponty at det er nettopp ideen om situasjonen som utelukker absolutt frihet: “The idea of situation rules out absolute freedom at the source of our commitments, and equally, indeed, at their terminus” (Merleau-Ponty, 1962, s. 454).

Å forklare hvordan hvordan Kant, Merleau-Ponty og Sartre konkluderer så ulikt om frihetens natur, vil kreve en langt mer inngående redegjørelse for teoriene deres enn det er rom for her. Men det er allikevel noen poenger jeg vil trekke ut av det vi har sett. Disse teoretikernes syn på frihet har det til felles at de på et eller annet vis tar utgangspunkt i erfaringene våre av frihet. Men derfra utleder Kant, Merleau-Ponty og Sartre svært ulikt om frihetens ontologiske dimensjon. Merleau-Ponty gir kroppen forrang framfor bevisstheten, og argumenterer heller for en fenomenologisk konkret frihet, der menneskenes resiprositet med omgivelsene står i sentrum for forståelsen av virkeligheten og handlingene våre. Han avviser dermed frihetens ontologiske dimensjon eller at den kan abstraheres slik Sartre og Kant gjør. Mens Kant har en tilnærming til frihet som kategorisk, dvs. at menneskets natur er å være fri, forholder Merleau-Ponty seg til frihet mer som en kraft og noe som til en viss grad karakteriserer

handlingene våre. For Sartre eksistere vi også innenfor en situasjon med hindringer, men utfra det som hindrer oss, utleder han i motsetning til Merleau-Ponty at vi er radikalt frie. Kant og Sartre kan sies å ha et sterkere frihetsbegrep enn Merleau-Ponty, der Kant argumenterer for handlingens absolutte spontanitet, og Sartre argumentere for en ontologi grunnlagt på ideen om radikal frihet.

Disse ulike perspektivene illustrerer samtidig hvordan spørsmål om frihet er sammenflettet med andre sentrale spørsmål om virkelighetens beskaffenhet. De fleste teorier om frihet forholder seg på en eller annen måte til virkeligheten som et kausalt landskap, og søker å si noe om hvilken rolle vi som mennesker spiller i dette. Frihet er gjerne forbundet med muligheten til å virke inn på eller å ha kontroll over deler av den virkeligheten vi befinner oss i. I forlengelsen av dette kan vi spørre om hva som karakteriserer den frie entitetens (eller selvets) natur, og hvor grensene går mellom det som er fritt og det som det er fritt i forhold til. Videre kan vi spørre om hva slags kausalitetsbegrep som på en tilfredsstillende måte kan beskrive denne frie entitetens aksjoner eller interaksjoner med omgivelsene. Her kan vi gå tilbake til Merleau-Ponty i sin kritikk av Sartre, som påpeker hvordan ideen om handling forsvinner dersom det er slik at friheten er vår grunnleggende natur: “Nothing can pass from us to the world, since we are nothing that can be specified, and since the non-being which constitutes us could not possibly find its way into the world’s plenum” (Merleau-Ponty, 1962, s. 437). Blir friheten total enten som totalt fravær eller som absolutt tilstedeværelse, slik at den ikke lenger framtrer for oss, kan det se ut som at framtredeisen av oss selv også forsvinner. Videre er frihetsbegrepet nært knyttet til et begrep om ansvarlighet, og moralske spørsmål er i mange tilfeller en viktig motivator for teorier om frihet. Hvis vi følger Sartre ser det ut som at jo mer frihet vi har, jo mer ansvarlige er vi. Men med Merleau-Ponty kan vi argumentere med at jo mer radikalt frihetsbegrep vi holder oss med, desto mer forsvinner framtredeisen av frihet, handlinger og den handlende bort for oss. Her må det legges til at selv om Sartre opprettholdt ideen om en radikal frihet på et ontologisk nivå, så innrømte han at vår frihet kunne være begrenset i det praktiske livet. En noe forenklet common sense slutning vi kan trekke ut av dette, er at jo mer frihet vi kan sies å ha, jo mer ansvar følger det med denne friheten. Vår betydning som handlende individer vokser i dette perspektivet opp til et visst punkt. Men når friheten blir absolutt, kan man med Merleau-Ponty si at framtredeisen av friheten og handlingene våre forsvinner, og dermed også oss selv. Er det ikke lenger noen som utøver noe, virker det vanskelig å snakke om ansvarlighet. For å være *noen* som kan utøve frie handlinger, må vi ha noe å bryne oss på, for å si det på en dagligdags måte.

I følge Merleau-Ponty ligger det i selve begrepet om frihet, at vår bestemmelse virker inn i framtiden og gjør noe:

The very notion of freedom demands that our decision should plunge into the future, that something should have been done by it, that the subsequent instant should benefit from its predecessor and, though not necessitated, should be at least required by it. If freedom is doing, it is necessary that what it does should not be immediately undone by a new freedom (Merleau-Ponty, 1962, s. 437).

Et annet sentralt spørsmål i teorier om frihet, er det epistemologiske spørsmålet om hva vi kan kjenne til. Kants oppfatning av grensene for hva vi kan erfare, står helt sentralt i hans tenkning. Neste spørsmål blir da hva slags slutninger vi kan trekke om frihet utfra det vi mener vi kjenner til. Her spiller ofte kontrafaktiske betingelser en viktig rolle. Selv ikke Merleau-Ponty kan forholde seg utelukkende til frihet som fenomen uten å sette det inn i et visst analytisk rammeverk, som også innebærer å forestille seg hvordan ting ville ha sett dersom premissene for erfaringen var annerledes. Viktige distinksjoner blir da å skille mellom det som er mulig og det som er virkelig, eller alternativt det klassiske begrepsparet potensialitet/aktualitet. En annen sentral distinksjon er mellom intensjon og handling, dvs. å skille mellom det som pågår internt hos en person forut for handlingen og selve handlingen. Hvordan man forholder seg til tid, er også et viktig moment ved frihetsbegrepet. Kant transcendente frihetsbegrep kan sies å være radikalt når det innebærer at årsaken til handlingen min ligger utenfor tiden og den fysiske verden. Men den kontrakausale årsaken han mener handlingene våre utgjør, er utledet fra de moralske intuisjonene våre, som gjerne legitimeres gjennom retrospektive vurderinger.

Vi har nå sett på noen oppfatninger om frihetens ontologiske dimensjoner, og også sett at disse er omstridte. Samtidig som det ikke synes å være et fullstendig skille mellom den friheten vi knytter til konkrete erfaringer og oppfatninger om frihetens ontologiske dimensjoner, er det allikevel ikke et symmetrisk forhold mellom den hverdagslige oppfatningen av fri vilje og den vitenskapelige/filosofiske litteraturen. Selv de mest hardnakkede determinister som benekter eksistensen av fri vilje, har nok vanskeligheter med basere seg på denne hypotesen i sitt eget dagligliv og språk.

5.1.1 Et moderne frihetsideal – eksempel fra skolen

I det demokratiske utdanningsprosjektet står frihetsbegrepet sentralt. Men det idealet om frihet som ligger til grunn den normative kjernetanken bak demokratiet er ikke nødvendigvis det samme som formidles til elevene ved ulike anledninger. Et egenopplevd eksempel er ved avslutningsseremonier og vitnemålsutdelinger i skolen. Samtidig med karakterene og vitnemålene som deles ut er avgjørende for framtidige muligheter med hensyn til videre utdanning og jobbmuligheter, formidles gjerne et budskap om et moderne frihetsideal som handler om muligheter. En gjenganger i talene fra samtlige representanter som rektor, lærere, elever og foreldre er ofte en variant av utsagnet: ”Husk at du har alle muligheter!”. Hele poenget med vitnemål er at de er ulike, og at de nettopp innebærer ulike begrensninger og muligheter for elevene i framtiden. Hvis alle hadde fått de samme karakterene i alle fag, hadde man ikke trengt å dele de ut. Elevene hadde heller ikke behøvd å anstrenge seg for å oppnå best mulig karakterer. Det å dele ut vitnemål er i seg selv en bekreftelse på at de ulike vitnemålene setter ulike begrensninger for elevene i framtiden. Når det ved samme anledning fortelles at ”ingen muligheter er utelukket”, er det en selvmotsigelse som er så opplagt at man kan spørre seg hvem som tror på det. Men hvorfor blir et slikt budskap presentert med så mye patos? Er det ment terapeutisk for å trøste de som i en forstand har mislykkes? De som ikke lengre kan drømme om å bli lege, advokat eller kanskje astronaut? Noen vil kanskje hevde at det ville være taktløst ikke å uttrykke optimisme på vegne av alle ved slike anledninger. Man kan også hevde at norske elever i en større grad enn elever i andre land faktisk har mange sjanser når det kommer til utdanning. Men i likhet med andre moderne samfunn, er kravene som stilles til elevene høyere i dag enn før. Den norske sosiologen Ivar Frønes hevder at økonomiske faktorer kun forklarer en liten del av barns risiko for marginalisering i samfunnet. Ifølge ham er det skolen som er den primære ekskluderingsfaktoren (Frønes, 2015, s. 43). Selv oppfatter jeg at ideen om ”ubegrensede muligheter” henviser mer til et moderne metafysisk ideal enn en pragmatisk virkelighet, og at budskapet reflekterer noen av hovedidéene som preger retorikken i dagens utdanningssystem.

Premisset for denne type påstander om ”ubegrensede muligheter” bli sjeldent utdypet eller drøftet. Men jeg antar at et premiss for mulighetene er at de er villet. Det som ikke er ønsket eller villet er vanligvis ikke forbundet med muligheter. Den indikative¹⁰ betingelsen kan

¹⁰ Indikativ: ”...modus som uttrykker at en ytring skal tolkes som et konstaterende eller fortellende utsagn...” (Bokmålsordboka, 2016)

dermed uttrykkes slik: ”Hvis du (bare) vil, så har du alle muligheter”. I hverdagsspråket kan dette enkelt oversettes (skjønt ikke nødvendigvis filosofisk uten noen ontologiske omkostninger) til den konjunktive¹¹ betingelsen og dermed den kontrafaktisk påstanden: ”hvis du (bare) hadde villet, hadde du hatt alle muligheter”. Dette budskapet er ikke på langt nær så motiverende for de som feilet i å leve ut sitt potensiale eller å gripe tilgjengelige muligheter. Tvert imot skaper det sannsynligvis negative følelser som tristhet, selvanklagelser, skam og kanskje sinne. Men i en viss utdanningsretorikk er det enkelt å skyve foran seg det indikative premisset og dermed skape positive forventninger og legitimere handlinger, mens den konjunktive betingelsen passerer i stillhet. Men ettersom vi som mennesker ikke bare engasjerer oss i prospeksjon, men også i retrospeksjon, vil indikative betingelser før eller senere erstattes av konjunktive versjoner. Det som *kan* skje, blir erstattet med det som *kunne ha* skjedd.

Dette uttalte frihetsidealet er et tidstypisk fenomen, og samtidig et radikalt uttrykk for et såkalt a-historisk frihetsbegrep. Et slikt begrep om frihet innebærer at fortidens betingelser og dermed de individuelle kausale historiene, er irrelevante for autonomi og de framtidige mulighetene. David Christopher Schindler (2012) tolker det moderne begrepet om frihet som varianter av et ”possibilistisk” begrep; dvs. frihet i termer av mulighet eller potensialitet som betyr kapasiteten eller evnen til å velge å gjøre X, Y eller Z (Schindler, 2012, s. xiv). Den uttalte ideen eller idealet om ”ubegrensede muligheter” ved slike anledninger er sannsynligvis en overdrevet og karikert versjon av en allmenn versjon av frihet i dag. Det vil også være feil å hevde at alle har den samme oppfatningen om frihet, da oppfatninger om frihet nok varierer fra person til person og fra situasjon til situasjon, selv for de samme individene. Det moderne idealet om frihet er et tveegget sverd, og kan oppfattes både som frigjørende og undertrykkende. De som holdt talene ved seremoniene kunne like gjerne ha sagt: ”Det kan se ut som at disse vitnemålene og det du har oppnådd gjennom ti år bestemmer framtiden din, men du har fremdeles friheten til å oppfylle drømmene dine”. Dette kan synes løfterikt da folk flest liker å få nye muligheter. Men i likhet med de fleste andre, engasjerer også elevene seg i retrospeksjon og kontrafaktisk tenkning; ikke minst når de blir vurdert ved prøver og eksamener, og når de vurderer seg selv opp mot gitte mål. Det moderne skolesystemet karakteriseres av kontinuerlige krav og kultivering av elevenes evne til å reflektere over og

¹¹ Konjunktiv: ”...verbalkategori, modus som uttrykker at et utsagn er tenkt, uvirkelig, mulig eller ønskelig” (SNL, 2016).

vurdere egne prestasjoner oppimot forhåndsbestemte mål og kriterier, gjennom retrospeksjon og introspeksjon. Hva skjer når elevene ikke lykkes i å innfri drømmene eller potensialene sine? Det er i det minste to svar på dette. Enten så var ikke premisset sant, eller så manglet eleven evnen til å ”benytte” friheten eller viljen sin, som han eller hun er forventet å besitte. I sistnevnte svar er det eleven som sitter igjen med ansvaret for ikke å lykkes.

Det moderne idealet om frihet som omtales her, der mulighetene i første instans synes å være knyttet til individets indre, minner om det klassiske idealet om autonomi, selv om det må påpekes at det er en overfladisk variant av dette. Det klassiske autonomibegrepet kan i enda større grad enn frihetsbegrepet sies å gå inn til kjernen av ideen om at vi er det ultimate opphavet til handlingen våre, og at det å handle innebærer å handle på en måte framfor en annen. Klassiske idealer om autonomi blir således særlig relevant for å belyse noen av de underliggende antagelsene ved kontrafaktisk tenkning. Samtidig kan innsikter om kontrafaktisk tenkning belyse idealer om autonomi, både klassiske og mer ”overfladiske” varianter. I neste kapittel vil jeg ta for meg autonomibegrepet og ulike betydninger av dette knyttet til kontrafaktisk tenkning.

5.2 Autonomi

Begrepet autonomi kommer fra gresk der *autos* betyr selv og *nomos* betyr lov. Opprinnelig ble begrepet brukt om bystaters selvstyre, men i senere tid har det også blitt vanlig å benytte autonomibegrepet om individets kapasitet for selvbestemmelse eller selvlovgivning. I *pedagogisk ordbok* defineres autonomi blant annet som “den egenskap hos en person å kunne foreta selvstendige handlinger og å kunne tenke uten å påvirkes av faktorer i omgivelsene” (Bø & Helle, 2008, s. 29). Akkurat hva det innebærer å handle selvstendig eller å være upåvirket av omgivelsene er det imidlertid ikke så enkelt å definere, og i likhet med frihetsbegrepet er autonomibegrepet ikke entydig. Men i enda større grad enn frihetsbegrepet, kan autonomibegrepet sies å innebære at vi er det ultimate opphavet til handlingene våre, og at vi derfor kunne ha handlet annerledes. Således kan man hevde at det å være autonom og det å vurdere kontrafaktiske scenarioer er nært forbundet. Men det innebærer naturligvis ikke at den kontrafaktiske tenkningen er helt vilkårlig, tvert imot stilles det strenge krav til det autonome individet. Som små sang vi ”for i bakvendtland, der kan alt gå an”, men slik er det ikke med den kontrafaktiske tenkningen knyttet til det autonome.

I en del sammenhenger brukes begrepene frihet og autonomi sammen eller om hverandre, men autonomibegrepet har i realiteten en snevrere betydning enn frihetsbegrepet. Samtidig anser mange at autonomi representerer selve kjernen i frihetsbegrepet. I boka *The Experience of Freedom* uttrykker Jean-Luc Nancy (1993) dette slik: ”Auto-nomy, which has represented the very regime of freedom, must be understood on this basis: as a legislation by the self in which the self does not pre-exist [...]” (Nancy, 1993, s. 107). Slik vi ser her hos Nancy, er begrepet om autonomi i tillegg nært knyttet til en oppfatning av selvet og da gjerne et dualistisk aspekt ved selvet. Kant, som regnes for å være opphavsmannen til autonomibegrepet, definerte autonomi i termer av en fri og fornuftig *selvlovgivning*. I forrige kapittel om frihet tok jeg opp problemet med at i en oppfatning om at friheten er absolutt eller total, enten i form av totalt fravær eller total tilstedeværelse, forsvinner tilsynelatende framtredeisen av oss selv. Dette gjelder om mulig i enda sterkere grad autonomibegrepet, og dermed vil oppfatninger om autonomi variere med oppfatninger om selvet. De som for eksempel betrakter selvet som sosialt konstituert, vil kanskje ha et annerledes syn på autonomi enn de som hevder at det finnes en sann kjerne av et selv, selv om disse ikke nødvendigvis står i motsetning til hverandre.

5.2.1 Frihet versus autonomi

For å synliggjøre hva som hovedsakelig skiller autonomibegrepet fra et begrep om frihet, tar jeg igjen turen innom Isaiah Berlin og det berømte essayet *Two Concepts of Liberty* (1969). Her tar Berlin for seg to sentrale typer frihet som han kaller *positiv frihet* og *negativ frihet*. Disse typene kan også formuleres som skillet mellom hva vi er *frie til* og hva vi er *frie fra*. Ideen med å skille disse to typene av frihet går for øvrig minst tilbake til Kant (Carter, 2016), mens Berlin utdyper og forsvaret inndelingen. Den negative varianten av frihet kaller Berlin det som er involvert når vi svarer på spørsmålet: “What is the area within which the subject - a person or group of persons - is or should be left to do or be what he is able to do or be, without interference by other persons?” (Berlin, 1969, s. 121). Negativ frihet forstås her som fravær av tvang eller ytre påvirkning. Positiv frihet innebærer på den andre siden svar på spørsmålet: “What, or who, is the source of control or interference that can determine someone to do, or be, this rather than that?” (Berlin, 1969, s. 122). Berlin påpeker at selv om spørsmålene er helt ulike, så kan svarene overlappe hverandre. Vi sies å ha negativ frihet dersom det ikke er noe ved de ytre omstendighetene, for eksempel tvang eller ytre påvirkning, som hindrer oss i å utøve en viss handling. Et eksempel kan være en elev som ønsker å gå ut

av klasserommet og spille fotball ute. Dersom det er friminutt eller en ”fritime” samtidig, og det heller ikke er andre omstendigheter som stopper eleven i å realisere den ønskede handlingen, så har eleven negativ frihet. Positiv frihet handler derimot om aktørers grunner for handling, dvs. interne faktorer som påvirker i hvilken grad individer eller grupper kan handle autonomt (Carter, 2016). Et eksempel på dette kan være den samme elevens kapasitet til å overveie og ta et valg mellom det å gå ut og spille fotball, eller å bli værende inne i friminuttet/fritimen og øve til en prøve i neste time. Diskusjonene om negativ og positiv frihet skiller seg fra den filosofiske fri vilje-debatten, men begrepet om positiv frihet overlapper tildels begrepet om autonomi. Hvis vi ser en gang til på spørsmålet til Berlin angående positiv frihet: ”What, or who, is the source of control or interference that can determine someone *to do, or be, this rather than that?*”¹², ser vi hvordan det å vurdere kontrafaktiske scenarioer er en sentral idé i definisjonen av positiv frihet og dermed autonomi. Berlin er eksplisitt med at positiv frihet handler om å sammenligne faktiske med kontrafaktiske scenarioer knyttet til både det å *gjøre* og å *være*.

5.2.2 Moralsk, politisk og personlig autonomi

I likhet med at det er flere dimensjoner av frihetsbegrepet, så skiller filosofene også autonomi gjerne inn kategorier som *moralsk*, *politisk* og *personlig autonomi*. Moralsk autonomi er først og fremst påvirket av Kant og handler om menneskets definisjon av seg selv som fornuftsvesen, med kapasitet til å foreta overveielser, og gi seg selv den allmenne lov det skal handle etter. Politisk autonomi handler om retten til å bli respektert for egne bestemmelser i en politisk kontekst, dvs. grunnlaget for “enkeltindividers politiske rettigheter mot kollektive interesser” (Sagdahl, 2014a). Med personlig autonomi menes gjerne kapasiteten til å ta selvstendige beslutninger og bestemme en handlingskurs uavhengig av religiøst eller moralsk innhold. Felles for disse er at for å kunne leve som *autonom* eller *selvstyrt* må ulike handlingsalternativer stå åpne for oss. Autonomibegrepet kan på den ene siden oppfattes som en reell, konkret og empirisk målbar kapasitet til å overveie og foreta beslutninger, mens det på den andre siden kan oppfattes som et abstrakt ideal og et grunnleggende syn på mennesket. Det er særlig autonomi som abstrakt ideal som er interessant i sammenheng med kontrafaktisk tenkning og de metafysiske antagelsene vi legger til grunn når vi tenker kontrafaktisk. Derfor vil jeg se nærmere på det i fortsettelsen.

¹² Min utheving.

5.2.3 Autonomi som abstrakt ideal

Da autonomi som abstrakt ideal ikke beskriver tingenes tilstand, men i stedet fungerer som en rettesnor for hvordan ting bør være, må vi i likhet med alle andre idealer anta det kontrafaktisk. Mark Coeckelbergh (2004) påpeker at idealet om å være autonom har både et indre og et ytre aspekt. Det å ville være autonom innebærer på den ene siden et ønske om å styre over ens egne ønsker og begjær (det indre aspektet), og på den andre siden å ikke bli styrt av noe utenfor en selv (det ytre aspektet). Joel Feinberg uttrykker noe tilsvarende og påpeker samtidig det dobbelte aspektet ved autonomi: "If we [...] call the "inner core self" the I and the rest of the comprehensive self over which it rules Me, then we can put the dual aspect of personal autonomy felicitously: *I am autonomous if I rule me and no one rules I*" (Feinberg, 2010, s. 117). At *jeg* kan styre over *meg* uttrykker en form for selvevaluering, og Coeckelbergh påpeker at det i tillegg innebærer at man har og kan bruke kapasiteten til fri vilje; "to be a space of freedom in a deterministic universe" (Coeckelbergh, 2004, s. 18). Det er imidlertid problematiske sider ved idealet om autonomi. Et problem er at det kan være vanskelig å skille ut og definere det "selvet" som styrer seg selv eller gir seg selv en lov å handle etter, dvs. hvor grensene går mellom det som styrer og det det styrer over. Et annet problem er vanskeligheten med å skille mellom det å være selvstyrt og det å være styrt utenfra. Jeg ser først litt problemet med å avgrense det autonome jeget. En forutsetning for autonome eller frie handlinger, er ifølge Coeckelbergh at de skjer i et deterministisk univers. I avsnittet om frihet påpekte jeg at frihet gjerne er forbundet med muligheten til å virke inn på eller å ha kontroll over deler av den virkeligheten vi befinner oss i. Hva er innholdet i, eller hva er inkludert i det deterministiske universet Coeckelbergh henviser til? Ettersom han snakker om hvordan det indre aspektet handler om å styre over ens egne ønsker og begjær, kan det se ut som han inkluderer ønsker og begjær som en del av det deterministiske universet vi ideelt sett kan styre over hvis vi er autonome. Dette er i så fall en vanlig, men samtidig interessant avgrensning av det autonome selvet, fordi det innebærer at *jeget* som styrer over *meget*, ikke inkluderer ønskene og følelsene våre. Men hva skal vi velge etter hvis det ikke er noen form for ønskemål som er styrende for valgene våre? Eller hvorfor i det hele tatt være opptatt av valg? Her kan vi vise til Kant som hevder at vi rett og slett ikke kan la være å ville følge den universelle loven som han kaller den. Det kan også kalles for et universelt ønske om å følge den – et ønske som overstyrer alle andre mer private ønsker vi måtte ha. Men hvor autonome er vi egentlig da, hvis vi uansett er forpliktet til den samme loven? Filosofen Derek Parfit (2011) retter kritikk nettopp mot dette aspektet av Kants filosofi. Han påpeker at Kant

postulerer en universell lov, samtidig med at enhver person har den moralske autonomien til å bestemme for seg selv. Da forsvinner alle valgmuligheter, og Parfit spør derfor hvordan man kan ha begge disse tingene samtidig:

I asked a Kantian, 'Does this mean that, if I don't give myself Kant's Imperative as a law, I am not subject to it?' 'No', I was told, 'you have to give yourself a law, and there's only one law.' This reply was maddening, like the propaganda of the so-called 'People's Democracies' of the old Soviet bloc, in which voting was compulsory and there was only one candidate. And when I said 'But I haven't given myself Kant's Imperative as a law', I was told 'Yes you have'. This reply was even worse (Parfit, 2011, s. 23).

Det finnes vektige argumenter på begge sider av denne debatten, og mange vil nok hevde at Parfit ikke forstår essensen av Kants kategoriske imperativ; noe Parfit åpent innrømmer at han ikke gjør. Her støter vi også på problemet med hvem som bestemmer kriteriene som ligger til grunn for å bedømme hva som er autonomt og hva som ikke er det. For gir det egentlig mening å kalle noe for autonomt dersom kriteriene for det autonome kommer utenfra? Et argument for at det finnes noe universelt som vi alltid tar utgangspunkt i, er at for eksempel Parfit deltar i en offentlig diskusjon, og dermed forutsetter at han kan uttrykke noe rasjonelt og meningsfullt om temaet selv om han ikke er enig i innholdet eller essensen av Kants teori.

For å kunne bruke friheten sin, fordrer det naturligvis reelle alternative handlingsmuligheter. Dernest fordrer det for at en handling skal være autonom, at den handlende selv har oversikt og kontroll over de ulike handlingsalternativene. Handlingsalternativer som ikke er mentalt tilgjengelig for den handlende, men som kun er "åpenbare" for andre personer, er ikke relevante i denne sammenhengen. Hvis dette stemmer kan man hevde at alle handlinger uansett er autonome, da det ikke er mulig å vurdere og velge handlingsalternativ man ikke kjenner til. Men dersom full autonomi virkelig lot seg realisere, ville det være overflødig å snakke om det. For at autonomi skal kunne være et ideal, noe å strekke seg etter, må det nødvendigvis innebære noe mer enn det som allerede og uansett skjer. Det paradoksale er at autonomi ifølge de ulike teoriene både er noe som vi allerede er i besittelse av, samtidig med at det er et ideal som må tenkes kontrafaktisk. Ifølge teoriene framstår dermed autonomi både som noe essensielt og faktisk, samtidig med at det er kontrafaktisk. Vi slipper ganske enkelt ikke unna det autonome i noen verdener.

5.2.4 Fornuftens allmakt

I kjølvannet av Parfit sin kritikk kan man også spørre seg om det i rasjonalitetens natur finnes noe som er så universelt at det står over subjektive interesser på en slik måte at det er en pålitelig kilde for moralske kriterier? At vi som mennesker uansett ikke kan unnsnippe et universelt gyldighetsteoretisk skille mellom hva som er sant og usant, eller rett og galt? Det er i alle fall en risiko for at myndigheter eller andre autoriteter kan hevde å inneha sannheten om disse skillene, og at de opptrer som en slags ”fornuftens allmakt”. Jeg begrenser meg her til å løfte fram denne kritikken for å se om den kan kaste lys over forståelsen av autonomi i oppdragelses- og utdanningssammenheng. For skolen med sitt samfunnsmandat er det et sentralt mål å få de unge til å styre over sine egne ønsker og begjær på skolens premisser, og aller helst er nok målet at skolen klarer å få de unge til å gjøre skolens mål til sine egne. På denne måten kan skolen opptre som ”fornuftens allmakt”, og dermed som en slags imitator for det universelle i dannelsen av det indre aspektet ved autonomi, som handler om å styre over ens egne ønsker og begjær. Det er kanskje litt pussig å hevde at en institusjon har ønsker og begjær, men til den grad man kan uttrykke noe slikt, så blir det skolens ønsker og begjær som blir styrende i denne dannelsesprosessen. Resultatet er at elevene skal kunne styre over seg selv på skolens premisser. Kan hende er det overdrevet å trekke sammenligningen helt til den gamle Sovjetblokken slik Parfit gjorde med tanke på den universelle loven, men jeg vil allikevel hevde at den typiske norske skolen formidler til elevene at de har en frihet som innebærer tilnærmet ubegrensede muligheter og dertil hørende ansvar. Jeg henviser her til det ”possibilistiske” frihetsbegrepet som jeg omtalte i kapittel 5.1.1. I realiteten er alternativene begrenset, i og med at skolen er obligatorisk, og at den i stor grad bestemmer hvilke erfaringer og refleksjoner elevene får gjøre seg. Et sentralt poeng er at skolens påvirkning pågår over lang tid og i den perioden av livet da et menneske er mest påvirkelig. Men det autonome individet skal ideelt sett bare styre over seg selv og ikke bli styrt av andre, og det er ikke nødvendigvis slik at målene og visjonene til et utdanningssystem samsvarer med den universelle loven slik Kant viser til den. Samtidig er det vanskelig å tenke seg noen form for oppdragelse eller utdanning som ikke setter grenser eller premisser for barnas opplevelser og refleksjoner, og herav kommer det pedagogiske paradokset. Vi er nå allerede langt inne på det ovennevnte beslektede problemet ved autonomi, som handler om å skille mellom det å være selvstyrt og det å være styrt utenfra.

5.2.5 Skillet mellom å være selvstyrt og styrt utenfra

Det å være fullstendig fri fra en ytre form for styring eller påvirkning, kan virke som en usannsynlig realitet, og Feinberg uttrykker følgende om det ettertraktede idealet om å være selvstyrt:

The absence of dictation from outside the total self is a much desired ideal situation (although *total* absence of external control is not a realistic hope), but since the fulfilment of this ideal of circumstance is largely a matter of luck, it would be odd to consider it also to be a virtue or ideal of character (Feinberg, 2010, s. 117)

Det er interessant å merke seg at Feinberg nevner betydningen av hell. Han bringer dermed inn et tilfeldighetsperspektiv på universet; dvs. at universet kan være indeterministisk. Men i et klassisk perspektiv, slik Coeckelbergh skisserer det, synes det å være en forutsetning at den autonome bruker friheten sin i et deterministisk univers (Coeckelbergh, 2004, s. 18).

Coeckelbergh er i tråd med Kant her, som postulerer at vi har fri vilje i en verden som følger ufravikelige lover.

Feinberg mener videre at en person bør berømmes for sin uavhengighet til andre, bare i den grad vedkommende ikke trenger ekstern kontroll, og dermed heller ikke tillater det (Feinberg, 2010). Med bakgrunn i denne påstanden, kan det synes som at autonomibegrepet ikke er så relevant knyttet til barn, som i mange henseende er avhengig av ekstern kontroll og ytre påvirkning. Lars Løvlie (2011) påpeker i denne sammenhengen viktigheten av å forstå barnets autonomi kontrafaktisk, nettopp fordi barnet i faktisiteten (eller realiteten) er avhengig av de voksnes innvirkning:

Det er nesten frekt å si at eleven er autonom når vi alle vet at barn er helt avhengige av voksnes omsorg, støtte og veiledning. På den andre side, hvis vi nekter å tenke kontrafaktisk og forkaster ideen om barnets autonomi, har vi svake barrierer mot å ta eleven som et produkt av disiplin, kunnskap og øvelse, og å se læring som innvielse og indoktrinering i samfunnets lokale normer [...] (Løvlie, 2011, s. 748).

Begrepet om *innvielse* ble introdusert av Richard Stanley Peters (1966) og betydningen av det innebærer en motvekt til indoktrinering. Jeg kommer tilbake til Peters innvielsesbegrep i kapittel 7 om *kontrafaktisk livsverden*.

Jeg påpekte for litt siden faren for at skolen opptrer som en imitator av fornuftens allmakt. Dersom den gjør det, kan det nettopp føre til at læring forblir det Løvlie kaller en ”innvielse og indoktrinering i samfunnets lokale normer” (Løvlie, 2011, s. 748). Autonomiforståelsen som ligger til grunn for en innvielse som kan kalles indoktrinering må sies både å være en fordreid versjon av Kants begrep om autonomi og helt i strid med det. Løvlie oppsummerer det pedagogiske paradokset med at ”selvbestemmelse eller autonomi både hører til elevens opprinnelige frihet og samtidig må formidles til henne” (Løvlie, 2011, s. 747). I tråd med paradoksets natur, så gjenstår problemet med at kriteriene for autonomi både er noe vi har i oss selv fra starten av, men at det også må formidles, og dermed utgjør en ekstern påvirkning. Løvlie hevder imidlertid ikke at paradokset er praktisk løsbart, men snarere at det er en oppfordring til ettertanke og handling.

5.2.6 Autonomihemmende versus autonomifremmende

I oppdragelses- og utdanningssammenheng pågår påvirkningen utenfra over lang tid, og dette øker kompleksiteten ved problemet med at kriteriene for det autonome må formidles. Selv om det kan hevdes at en person tilsynelatende tar selvstendige beslutninger (over seg selv), kan det på den andre siden anføres at den samme personen har vært utsatt for ytre påvirkning over lengre tid, som for eksempel har virket *autonomihemmende*. Dette kan være så integrert i personen at ikke engang vedkommende selv er klar over denne omstendigheten. Med *autonomihemmende* mener jeg en form for ytre påvirkning som hemmer utviklingen av autonomi. En annen måte å uttrykke det på er at barnet ikke får hjelp til å bli klar over dets egentlige autonome natur. Eller enda verre, så misledes barnet til å tro at det er en naturgitt ting at andre bestemmer over det. Men ved i det hele tatt å snakke om at autonomi er noe som utvikles, forutsetter jeg et visst syn på autonomi og griper dermed rett inn i debatten om hva autonomi er og når vi kan sies å være autonome. For å bedømme om noe er *autonomihemmende* kreves det kriterier som er utenfor den egne personen det handler om, og dette står delvis i strid med autonomibegrepet. Dersom man hevder at autonomi er en kapasitet som utvikles, bør man også kunne skille mellom virkninger av ulike kausale historier. En kausal historie som karakteriseres av for eksempel indoktrinering, blir vanligvis sett på som å underminere fri vilje og ansvarlighet, og dermed hemme utviklingen av autonomi. Men et syn på autonomi som innebærer at autonomi er noe som utvikles, står samtidig i strid med synet på frihet og autonomi som absolutte metafysiske og *ahistoriske* størrelser, dvs. at vi skulle være frie uavhengig av tid og rom. Med *ahistorisk* menes

uavhengig av historiske omstendigheter. Denne sentrale debatten innen pedagogisk filosofi overlapper dels med metafysikkens fri vilje-problem. Dette diskuterer jeg mer utførlig i kapittel 4.5 om *Ekstern kompatibilisme og pedagogisk filosofi*.

Med forbehold om at det er selvmotsigende å snakke om *autonomihemmende* påvirkning fortsetter jeg allikevel på dette sporet. Det ligger i sakens natur at påvirkningen og effektene av en indoktrinering er så integrert i personen at vedkommende selv ikke er klar over det, og utgjør dermed ikke en del av de parameterne vedkommende tar høyde for eller hensyn til ved overveielser. I slike tilfeller kan man hevde at refleksjons- og dermed handlingsrommet er betydelig redusert i forhold til hva det kunne ha vært eller burde ha vært. Men her støter vi igjen på kontrafaktiske idealer. Hva er kriteriene for det ideelle refleksjons- og handlingsrommet og hvem sitter inne med disse? Er det psykologene, politikerne, dommerne, filosofene eller hver enkelt av oss? Det kan se ut som at dette spørsmålet må løses før vi kan skille mellom hva som kan kalles *autonomihemmende* påvirkning og det motsatte – la oss kalle det *autonomifremmende* påvirkning? Det hører også til sakens natur at en person som i tilfellet over ikke selv er klar over at den har vært utsatt for en såkalt *autonomihemmende* påvirkning. Vi har da for eksempel den situasjonen at noen utenfra hevder nettopp dette, og dermed utgjør nok en ytre påvirkning. Personen det gjelder har nå flere ”lag” av ytre påvirkninger å ta stilling til, samtidig med at han eller hun skal gjøre en kvalifisert (eller autonom) bedømmelse av gyldigheten ved disse. Disse problemstillingene kan sies å være ulike aspekter ved det klassiske pedagogiske paradokset. Her er en fremstilling av paradokset slik Kant selv formulerte det i introduksjonen til forelesningsnotatene hans som ble publisert i 1803 i *Om Pedagogikk*:

“Et av oppdragelsens største problemer er hvordan man kan forene underkastelsen under lovens tvang med evnen til å benytte seg av sin frihet. For tvang er nødvendig! Hvordan kultiverer jeg friheten der det finnes tvang? Jeg skal vende min elev til å tåle tvang mot hans frihet og skal samtidig veilede ham i å bruke sin frihet godt” (Kant, 2016, s. 58)

5.2.7 Oppdragelse og utdanning som iscenesettelse

Den ”nødvendige tvangen” framstår som en del av et deterministisk univers, og som ifølge Kant synes å være en forutsetning for at det autonome individet skal ha noe å utøve sin frihet i. Med andre ord kan det se ut som en av oppgavene til de voksne er å kontrollere at visse

deler av virkeligheten er ”deterministisk” mens andre ikke er det, slik at de unge blir tilvent både å tåle en tvang mot friheten, og blir ledet til å bruke den. Dette er selvfølgelig ikke det vi vanligvis legger i begrepet om et deterministisk univers, der vi normalt henviser til ufravikelige naturlover. Men for illustrasjonens skyld velger jeg å formulere det slik; at det å oppdra og utdanne i en viss forstand dreier seg om å iscenesette eller imitere både et deterministisk og et indeterministisk univers gjennom de rammene og begrensningene for eksempel skolen setter for elevens refleksjons- og handlingsrom. Et poeng som bare indirekte berører dette med elevens autonomi, er at virkningen av de rammene og betingelsene skolen setter, går langt utenfor elevene selv. Utover det å virke inn på elevene, påvirker skolen i større eller mindre grad handlingsrommet til foreldrene og familien for øvrig, og deretter andre instanser knyttet til eleven og familien. Et eksempel er den siste fraværsgrensen som ble innført 1. august 2016 med hensikt å motivere elevene til jevn innsats og forhindre skulk. Foruten at den setter strenge grenser for elevens mulighet til både å ha fravær og ta en godkjent utdanning, virker den inn på fastlegenes valg, prioriteringer og arbeidsdag.

Kant snakker både om evnen til å betjene seg av sin frihet og det å kultivere frihet ved tvang, og hos Coeckelbergh så vi at en betingelse for å være autonom er det å kunne *bruke* friheten sin. I kapittel 5.1 om frihet påpekte jeg at de fleste teorier om frihet forholder seg til virkeligheten som et kausalt landskap. Så søker de å si noe om hvilken rolle vi som mennesker spiller i dette landskapet. Knyttet til autonomibegrepet synes det som det kausale landskapet vi orienterer oss i, først og fremst handler om vårt eget indre. Slik Feinberg definerer autonomi dobbelt, der *jeg* styrer over *meg* og ingen annen styrer over *jeg*, henviser han i første ledd til en intern hendelse, der utenforliggende aspekter ved virkeligheten ikke regnes med. Men poenget med det autonome individet faller vel bort hvis det i ytterste instans ikke finnes en virkelighet å navigere i utenfor *jeget* og *meget*. Ønsker og begjær må nødvendigvis også stå i forhold til noe utenfor subjektet. Men i definisjonene av autonomi er fokuset først og fremst på at det autonome individet ikke er styrt av noe utenfor seg selv, og ikke det som det styrer i.

For ytterligere å problematisere temaet om ytre påvirkning og kausal historie, skisserer jeg en enkel modell for hva det kan sies å innbefatte å være autonom. Fokuset mitt ligger på kapasiteten knyttet til valg og kontrafaktiske resonnementer. Jeg tar utgangspunkt i at den autonome forholder seg til to hoveddimensjoner av virkeligheten. Den første er slik virkeligheten til enhver tid er, eller det vi kaller *tingenes tilstand*. Den andre dimensjonen er

hvordan virkeligheten *bør* se ut. Valgene og handlingsalternativene våre kan sies å befinne seg innenfor den andre og *normative* dimensjonen. Vi handler ifølge denne modellen alltid fra og inn i tingenes tilstand. Det autonome individet forventes å kunne redegjøre for og legitimere handlingene sine, og må derfor kunne beskrive tingenes tilstand. Men ettersom det å beskrive krever utvalg og dermed valg, så er ikke tingenes tilstand tilgjengelig for nøytrale eller objektive beskrivelser. Derfor er vi over i den normative dimensjonen så fort vi søker å beskrive tingenes tilstand. Her må det innvendes at det allikevel synes å være en stor forskjell i å 1) ha et mer eller mindre ubevisst utgangspunkt i en fortolkning om tingenes tilstand, og 2) det å gjøre et bevisst og legitimt valg om framtiden. Poenget med at vi på samme måte ikke velger fortolkningene våre av tingenes tilstand, er sentralt. Men ettersom vi ikke kan unngå å handle, er det som om vi alltid befinner oss i et skjæringspunkt, eller at vi til og med er selve skjæringspunktet i en totalitet vi ikke kan overskue. Med sete i tingenes tilstand (som vi ikke kan beskrive nøytralt), kan vi ikke unngå å virke inn i den framtidige totaliteten, og vi gjør dette gjennom å forestille oss ulike framtidige versjoner av tingenes tilstand. Det å leve kan sammenlignes med alltid å ”skjære” inn i virkeligheten og ”dele den opp” i alternative virkeligheter. Vi kan ikke velge blant ulike *totale* versjoner av tingenes tilstand, da vi ikke besitter en absolutt kraft, men vi kan velge mellom ulike måter å påvirke den framtidige tilstanden. Det å være autonom og rasjonell innebærer i denne enkle modellen at man forholder seg til en noenlunde felles oppfatning eller fortolkning av: 1) tingenes tilstand og 2) mulige framtidige og dermed kontrafaktiske tilstander. Deretter kreves det en kapasitet til å sammenligne disse, og å kunne overveie og velge blant handlingsalternativer som på ulike måter vil virke inn på framtidige tilstander. Samtidig kreves det en kapasitet til å gå retrospektivt til verks og evaluere utfallet mot kontrafaktiske scenarioer. Det autonome individet skal beherske å være dette ”skjæringspunktet” i tilværelsen på en selvstendig måte, men skal samtidig oppfylle kriterier for å utføre dette på en selvstendig måte. Men hvis man holder seg til oppfatningen om at autonomi utvikles gjennom en kausal historie, er det grenser for hvor selvstyrte vi kan være. Vi styrer ikke over det totale bildet til enhver tid. Når det gjelder mulig *autonomihemmende* påvirkning, kan en vesentlig ytre påvirkning synes å ligge i det at ”beskrivelsesapparatet” vårt er prentet inn i oss så tidlig i livet, at det ikke er innenfor rekkevidden av verken hukommelsen eller vår rasjonelle tenkning å gripe opphavet til denne kunnskapen. Dermed er vi igjen over på diskusjonen om hva som er autonomihemmende og autonomifremmende påvirkning, og paradokset er som seg hør og bør uløst.

5.2.8 Frihet og autonomi i den generelle delen av læreplanen

Idealet om autonomi går hånd i hånd med idealet om demokrati. Ettersom autonomi er selve grunnpilaren i det demokratiske utdanningsprosjektet, vil det være interessant å se hvilket fokus frihet og autonomi gis i den generelle delen av læreplanen. Begrepet ”autonomi” forekommer ikke i denne delen av læreplanen. Begrepet frihet forekommer to ganger, men ingen av dem omhandler direkte idealet om det frie individet. Om betydningen av frihet for demokratiet velger jeg å sitere Lars Fr. H. Svendsen: ”Det mest grunnleggende trekket ved et liberalt demokrati er uansett at alle borgere tilkjennes et ukrenkelig frihetsrom” (Svendsen, 2013, s. 118). Forestillingen om frie og likestilte individer er selve kjernen i demokratiet, og man bør kunne si at det er en forutsetning for demokratiet. I læreplanen er demokratibegrepet nevnt et par ganger, men det viktigste står i hvert fall: ”Opplæringen [...] må fremme demokrati...” (LK06, 1993, s. 2). Men begrepene autonomi og frihet er altså nærmest utelatt i læreplanen. ”Ansvar” forekommer derimot en rekke ganger (28), og i mange ulike sammenhenger. Noen av formuleringene er i tråd med de oppfatningene jeg har drøftet over; om en felles fornuft og det å velge og vurdere blant alternative scenarioer: ”Opplæringen [...] må lære de unge å se framover og øve evnen til å treffe valg med fornuft. Den må venne dem til å ta ansvar - til å vurdere virkningene for andre av egne handlinger og å bedømme dem med etisk bevissthet” (LK06, 1993, s. 2). Det å vurdere virkningene (for andre) av egne handlinger innebærer å sammenligne scenarioer og dermed det å resonnerer kontrafaktisk. Det står riktignok ikke noe direkte om verken det å ”se bakover” eller å tenke kontrafaktisk, men det står for eksempel: ”Den (skolen) må organiseres slik at elevenes virke får konsekvenser for andre, og slik at de kan lære av konsekvensene av egne avgjørelser” (LK06, 1993, s. 17). Å stå for en slik organisering er et enormt ansvar, da det griper inn i selve dannelsen av elevenes rasjonelle livsverden og kontrafaktiske livsverden. Foruten at det sjelden er enighet om de tilsiktede effektene av enhver form for organisering av skolen, så er det et faktum at også utilsiktede effekter hersker. En legitim forutsetning må være at de som organiserer og styrer skolen krever det samme av seg selv som de gjør av elevene, dvs. at de lærer av konsekvensene av egne avgjørelser.

Det er allikevel påfallende at læreplanen er langt mer eksplisitt med at elevene skal læres opp i eller utvikle evnen til å ta ansvar, enn at det skal skje en kultivering av frihet. Opplæring i det å ta ansvar krever også, om ikke opplæring i, så bevisstgjøring av frihet. Men dette synes ikke å være et tema som gis særlig rom i den generelle delen. Selv ikke der det handler om

mennesket som moralsk vesen, synes det vesentlig å nevne frihetsaspektet ved det å opptre ansvarlig. En forklaring av at frihetsaspektet er så lite kommunisert, kan være en tendens til å tenke om barn at de i utgangspunktet har en iboende villskap eller frihetsfølelse som må temmes snarere enn å kultiveres gjennom oppfostringen. I norsk sammenheng er dessuten frihet noe som vi nærmest tar for gitt, og i en rekke sammenhenger kritiseres det heller at folk har og/eller tar seg for mye frihet og at de dermed utviser for lite ansvar. Dette er i så fall en helt annen type frihet enn den jeg drøfter i dette kapittelet. Men man kan spørre seg hvorfor det er så sentralt at de unge vennenes til å ta ansvar? Kort sagt er kjernetanken bak demokratiet frihet og likhet. For at demokratiet med frie og likestilte mennesker som en grunnleggende forutsetning tilnærmelsesvis skal fungere, kreves det borgere som opptrer ansvarsfullt. Men ved lesning av læreplanens generelle del, kan det se ut som at frihet heller oppstår som et resultat av ansvaret som tas, enn som en grunnleggende forutsetning for det samme. Men hvis det virkelig er slik at frihet er et resultat av ansvaret som skal tas, er det betimelig å spørre hvem det i så fall er som kan disponere denne friheten?

5.2.9 Autonomi og grunner for handling

I dette siste avsnittet om autonomi vil jeg se på betydningen av grunner knyttet til det å være autonom. I tråd med Berlin sin definisjon av positiv frihet, så handler positiv frihet (og autonomi) om aktørers grunner for handling. Begrepet grunn kan brukes i to betydninger: *Forklarende og normative grunner*. Forklarende grunner er en omstendighet som forklarer hvorfor en hendelse har skjedd, og er i mange tilfeller det samme som en årsak (Sagdahl, 2014b). Det er imidlertid en viktig forskjell mellom grunner for handling og årsaker til handling, uten at denne er entydig. Jeg utdyper dette i kapittel 6.5 om *handlingsforklaringer*. Men vi har ytterligere en kategori, som kalles *motiverende grunner*. Ifølge Sagdahl befinner disse seg mellom forklarende og normative grunner (Sagdahl, 2014b). Jonathan Dancy definerer skillet mellom motiverende og normative grunner: “Motivating reasons are supposed to be the ones that actually made the difference to how he acted; they constitute the light in which he chose to do what he did (Dancy, 2002, s. 2). Motiverende grunner dreier seg da om de hensyn en person selv legger til grunn for handlinger. De normative grunnene er, forklarer Dancy: “...the ones that should show that the action was sensible or right or whatever”. Det å oppgi grunner for handling impliserer kausale og kontrafaktiske resonnementer: Når vi angir grunner for handlingene våre, er det i form av en fortelling som direkte eller indirekte beretter om hva som ville ha skjedd dersom vi ikke hadde gjort slik vi

gjorde. Men det er ikke alle *motiverende grunner* som regnes for å være normative. Dancy uttrykker det slik: "...if an agent does something extremely silly, we might say he had a motivating reason for doing what he did but little normative reason" (Dancy, 2002, s. 2). Dancy påpeker imidlertid at dette ikke innebærer at det finnes to separate typer grunner. En motiverende grunn angir personens grunn for handling, mens det normative aspektet ved grunnen dreier som om hvorvidt den motiverende grunnen kan sies å være en god grunn: "...one that favours acting in the way proposed" (Dancy, 2002, s. 3). Når vi benytter uttrykket "god grunn" er det dermed et signal om at vi tenker normativt. Motiverende grunner uttrykker også hva slags oppfatning personer har om normative grunner. Vi kan ta eksemplet med gutten som velger mellom å spille fotball i "fritimen" eller å bli inne og øve til prøven i neste time. Sett at han velger å spille fotball og i stedet forsøker å jukse på prøven, for uansett å oppnå best mulig resultat. Det å spille fotball, og allikevel oppnå et godt resultat på prøven, kan i dette tilfellet sies å være guttens motiverende grunn. Men læreren som eventuelt oppdager at han jukser, er ikke nødvendigvis interessert i å høre hva grunnen til juksingen var; fordi det strider mot reglementet. Sett med "normative briller" finnes det ikke en god grunn til å jukse.

Det å utvikle "kapasiteten" til å begrunne handlinger på en slik måte at de rettferdiggjør handlingen, kan sies å skje gjennom innvielse i en rasjonell livsverden. Gjennom opplæringen kan vi si at elevene anspores til å foreta normativt "riktige" eller "gode" handlingsvalg. Dette gjelder ikke bare atferdsmessig, men også faglig. Som elev berømmes du både med ros og gode karakterer hvis du løser oppgavene i overensstemmelse med kriteriene for det høyeste nivået av måloppnåelsene. Her vil jeg løfte fram noen mulige spørsmål og fallgruver i pedagogiske sammenhenger. Den første er det jeg allerede har sett på, nemlig hvordan en innvielse i det allmenngyldige kan skje uten at det samtidig pågår en stor grad av påvirkning som bikker over i å hemme utviklingen av autonomi snarere enn å fremme den. Et eksempel på hvordan det kan skje i dagens skole, er at elevene blir så opptatt av å unngå sanksjoner og/eller få ros (for eksempel "bra-kort" gjennom PALS)¹³, at de ikke venner seg til å gjøre selvstendige bedømmelser og dermed styre over egne ønsker og begjær, men i stedet forsøker å unngå eller oppnå en utvendig respons. De er motivert at noe utenfor seg selv, og det er lite rom for eller oppmuntring til egne refleksjoner om hvordan man ønsker at verden og en selv

¹³ Bra-kort er en del av verktøykassen til PALS-programmet (Positiv atferd, støttende læringsmiljø og samhandling). Elevene får utdelt "bra-kort" når de gjør noe som læreren definerer som bra eller positivt. Bra-kortene telles og samles inn i slutten av uken og utløser belønning til klassen.

skal være. En kritikk her er at skolens atferdsprogrammer ikke fremmer utviklingen av autonomi, men snarere hemmer den.

Et annet problem er tendensen til å anta et deterministisk rammeverk når vi angir grunner for eller forklarer handlinger. Vi synes ofte å anta at det er like mye i årsaken som virkningen, med den konsekvensen at vi ikke tar høyde for tilfeldigheter og at universet kan være indeterministisk. Kravet til at vi skal kunne overskue og ta ansvar for konsekvensene av handlingene våre, kan lede til at vi prøver å overskue og kontrollere mer enn det som er mulig. Jeg har tidligere (2.2.3, 4.3) redegjort for *hindsight bias* og tendensen *tilsnikende determinisme*; det at vi ved retrospeksjon har en tendens til å overdrive hva som kunne vært antatt prospektivt, samtidig med at vi oppfatter det som skjedde som relativt uunngåelig (Fischhoff, 1974). Tendensen til ”snikende determinisme” kan også bidra til at vi oppfatter at alle handlinger må kunne forklares. Det kan oppstå en form for skyldighet til alltid å kunne forklare sine egne beveggrunner. Dette leder til at vi i retrospekt trekker konklusjoner om årsaker som er overdrevne og/eller utelukker andre alternativer. Men mange hendelser og handlinger lar seg ikke fullt ut forklare. Og selv om man antar et relativt deterministisk univers, innebærer ikke det at vi som mennesker kan overskue alle årsakssammenhenger. Våre egne beveggrunner er ikke alltid engang kjent for oss selv. I oppdragelses- og utdanningssammenheng forventes elevene i stigende grad å forstå seg selv, være bevisst handlingsalternativer og kunne redegjøre for valgene sine. Man kan spørre seg om mengden standardiserte mål og kriterier som er knyttet til nær sagt alle aktiviteter og oppgaver elevene beskjeftiger seg med i nåtidens skole, er med på å forsterke ideen om at valg i høy grad kan og skal kontrolleres. Det å ”begrunne valg” har blitt litt av et mantra i beskrivelser av måloppnåelser, der elevene nesten uavhengig av type fag eller oppgave, skal kunne redegjøre for hvorfor de gjør på en måte i stedet for en annen. Svært mange kriterier formuleres rundt dette, for eksempel ”å kunne begrunne valgene på en meget god og gjennomtenkt måte”, og det å ”jobbe selvstendig”. Eleven skal for eksempel kunne begrunne valget av én skjønnlitterær bok framfor en annen, eller hvorfor de valgte én oppgave eller én teknikk framfor en annen. Det er mange valg som ikke er så gjennomtenkte, og en del er rent ut tilfeldige. For eksempel : ”Jeg valgte den romanen fordi det var den først jeg så på biblioteket”, eller ”fordi den hadde et fint omslag”. Eller at ”jeg valgte å sy fast lappen i stedet for å lime den på, fordi jeg ikke fant limet den timen”. I stedet må elevene finne opp de riktige begrunnelsene som stemmer med kriteriene for faget. Et enkelt skjema for måloppnåelse kan gi inntrykk av at det finnes en uendelig mengde alternativer å velge mellom i en

arbeidsprosess. De utallige målbeskrivelsene kan gi inntrykk av at det faktisk er mulig å oppnå en bevisst kontroll over de utallige valgene man tar, og at ingenting av det vi *gjør* skjer på bakgrunn av tilfeldigheter, eller for eksempel på bakgrunn av (det noe vage begrepet) intuisjon. Det er riktignok ingen som har hevdet at disse skjemaene skal gjenspeile virkelighetens kausale natur. De angir bare hva elevene skal gjøre og ikke hva som skjer, og på denne måten angir de kun handlinger og ikke hendelser. Dette er for øvrig ikke helt korrekt beskrevet av meg, da en del kriterier også angir uønskede *hendelser* som gjerne leder til en lavere karakter. Dette tar jeg opp bredere i kapittel 6.4 om *Handlinger versus hendelser*. Det er riktig at vi foretar en mengde valg hele tiden, men de fleste av disse er automatiserte. Jeg tror ikke jeg tar feil i å hevde at elever aldri før i så stor grad har måttet redegjøre for sin kontrafaktiske tekning og begrunne alle valgene de tar. Samtidig er ikke dette i seg selv noe de blir gjort oppmerksomme på. Kritikken her går ut på at mengden valg som skal begrunnes er så stor, at elevene forledes til å anta at alt kan og skal forklares. Skolen ansporer til en ekstrem kontroll fra elevenes side. I årevis måler de seg opp mot idealer, med den bakenforliggende ideen at utfallet av deres aktørskap utelukkende er resultatet av deres egne bevisste handlinger. Det er et spørsmål om dette kan gi et fordreid bilde av hva det vil si å handle og av virkelighetens beskaffenhet. Faren er at eleven sitter igjen med opplevelsen av et totalt ansvar for utfallet. I neste omgang kan dette lede til ansvarsfraskrivelse for de som ikke lykkes, fordi de har ansvaret og dermed skylden for dette selv. Eleven forledes inn i ontologiske og epistemiske antagelser om virkeligheten uten å ha noen forutsetning for å vurdere om de deler disse antagelsene.

Den neste kritikken handler om at det å kunne begrunne valg etter gitte kriterier, ofte oppgis som et mål i seg selv. Derimot er det sjeldnere å se at kriteriene for selve begrunnelsen av valg er uttalte; som regel står det bare at man skal kunne begrunne valget. Men erfaringen min tilsier at det ikke er irrelevant hvordan valgene begrunnes, og at selve begrunnelsene er en del av vurderingen. Et av hovedmålene og det som ofte gir best uttelling karaktermessig, er at eleven er bevisst alle og enhver av sine beveggrunner, og er i stand til å redegjøre for den indre prosessen med å velge blant disse på den ”riktige” måten. Det faktum at denne type selvevalueringer overvåkes av læreren tilsier at det ikke bare er elevens motiverende grunner skolen er ute etter. Elevens egentlige motiverende grunner for handling kan dermed bli noe de holder for seg selv, mens han eller hun i stedet forsøker å finne ut hva som er den normativt ”riktige” grunnen for valget. Faren er at elevene sorterer ut sine egne motiverende grunner, uten å reflektere nøyere over de. Dette kan lede til at læring hemmes, fordi elevene ikke får

en reell mulighet til å lære av ”feilene” sine i retrospekt, og må manipulere sine egne begrunnelser allerede i det de tas. Min hypotese er at disse formene for evaluering og selvevaluering kan være mer hemmende enn fremmende for læring.

Tendensen til ”snikende determinisme” kan også bidra til at vi oppfatter at alle handlinger må kunne forklares. Det kan oppstå en form for skyldighet til alltid å kunne forklare sine egne beveggrunner. Dette kan lede til vi at i retrospekt trekker konklusjoner om årsaker som er overdrevne og/eller utelukker andre alternativer. Men mange hendelser og handlinger lar seg ikke fullt ut forklare. Og selv om man antar et relativt deterministisk univers, innebærer ikke det at vi som mennesker kan overskue alle årsakssammenhenger. Våre egne beveggrunner er ikke alltid engang kjent for oss selv. I oppdragelses- og utdanningssammenheng forventes elevene i stigende grad å forstå seg selv, være bevisst handlingsalternativer og kunne redegjøre for valgene sine. Man kan spørre seg om mengden standardiserte mål og kriterier som er knyttet til nær sagt alle aktiviteter og oppgaver elevene beskjeftiger seg med i nåtidens skole, er med på å forsterke ideen om at valg i høy grad kan og skal kontrolleres. Det å ”begrunne valg” har blitt litt av et mantra i beskrivelser av måloppnåelser, der elevene nesten uavhengig av type fag eller oppgave, skal kunne redegjøre for hvorfor de gjør på en måte i stedet for en annen. Svært mange kriterier formuleres rundt dette, for eksempel ”å kunne begrunne valgene på en meget god og gjennomtenkt måte”, og det å ”jobbe selvstendig”. Eleven skal for eksempel kunne begrunne valget av én skjønnlitterær bok framfor en annen, eller hvorfor de valgte én oppgave eller én teknikk framfor en annen. Det er mange valg som ikke er så gjennomtenkte, og en del er rent ut tilfeldige. For eksempel : ”Jeg valgte den roman fordi det var den først jeg så på biblioteket”, eller ”fordi den hadde et fint omslag”. Eller at ”jeg valgte å sy fast lappen i stedet for å lime den på, fordi jeg ikke fant limet den timen”. I stedet må elevene finne opp de riktige begrunnelsene som stemmer med kriteriene for faget. Et enkelt skjema for måloppnåelse kan gi inntrykk av at det finnes en uendelig mengde alternativer å velge mellom i en arbeidsprosess. De utallige målbeskrivelsene kan gi inntrykk av at det faktisk er mulig å oppnå en bevisst kontroll over de utallige valgene man tar, og at ingenting av det vi *gjør* skjer på bakgrunn av tilfeldigheter eller for eksempel på bakgrunn av (det noe vage begrepet) intuisjon. Det er riktignok ingen som har hevdet at disse skjemaene skal gjenspeile virkelighetens kausale natur. De angir bare hva elevene skal gjøre og ikke hva som skjer, og på denne måten angir de kun handlinger og ikke hendelser. Dette er for øvrig ikke helt korrekt beskrevet av meg, da en del kriterier også angir uønskede *hendelser* som gjerne leder til en lavere karakter. Dette tar jeg opp bredere i kapittel 6.4 om *Handlinger*

versus hendelser. Det er riktig at vi foretar en mengde valg hele tiden, men de fleste av disse er automatiserte. Jeg tror ikke jeg tar feil i å hevde at elever aldri før i så stor grad har måttet redegjøre for sin kontrafaktisk tekning og begrunne alle valgene de tar. Samtidig er ikke dette i seg selv noe de blir gjort oppmerksomme på. Kritikken her går ut på at mengden valg som skal begrunnes er så stor, at elevene forledes til å anta at alt kan og skal forklares. Skolen ansporer til en ekstrem kontroll fra elevenes side. I årevis måler de seg opp mot idealer, med den bakenforliggende ideen at utfallet av deres aktørskap utelukkende er resultatet av deres egne bevisste handlinger. Det er et spørsmål om dette kan gi et fordreid bilde av hva det vil si å handle og av virkelighetens beskaffenhet. Faren er at eleven sitter igjen med opplevelsen av et totalt ansvar for utfallet. I neste omgang kan dette lede til ansvarsfraskrivelse for de som ikke lykkes, fordi de har ansvaret og dermed skylden for dette selv. Eleven forledes inn i ontologiske og epistemiske antagelser om virkeligheten uten å ha noen forutsetning for å vurdere disse selv.

Den neste kritikken handler om at det å kunne begrunne valg etter gitte kriterier, ofte oppgis som et mål i seg selv. Derimot er det sjeldnere å se at kriteriene for selve begrunnelsen av valg er uttalte; som regel står det bare at man skal kunne begrunne valget. Men erfaringen min tilsier at det ikke er irrelevant hvordan valgene begrunnes, og at selve begrunnelsene er en del av vurderingen. Et av hovedmålene og det som ofte gir best uttelling karaktermessig, er at eleven er bevisst alle og enhver av sine beveggrunner, og er i stand til å redegjøre for den indre prosessen med å velge blant disse på den ”riktige” måten. Det faktum at denne type selvevalueringer overvåkes av læreren tilsier at det ikke bare er elevens motiverende grunner skolen er ute etter. Elevens egentlige motiverende grunner for handling kan dermed bli noe de holder for seg selv, mens han eller hun i stedet forsøker å finne ut hva som er den normativt ”riktige” grunnen for valget. Faren er at elevene sorterer ut sine egne motiverende grunner, uten å reflektere nøyere over de. Dette kan lede til at læring hemmes, fordi elevene ikke får en reell mulighet til å lære av ”feilene” sine i retrospekt, og må manipulere sine egne begrunnelser allerede i det de tas. Min hypotese er at disse formene for evaluering og selvevaluering kan være mer hemmende enn fremmende for læring.

Den siste mulige konsekvensen jeg vil trekke fram, er at målstyringen i dagens skole synes å gi lite rom for annet enn introspeksjon og retrospeksjon hos elevene. I denne virksomheten er det skolens kontrafaktiske ideal som setter premissene. Elevene gis ikke hverken tid eller rom for å undersøke de sammenhengene de inngår i, men i stedet ledes de kontinuerlig til å

undersøke og vurdere seg selv og egne prestasjoner opp mot gitte standarder. Det skjer dermed ingen bevisstgjøring hos elevene gjennom at de får sammenligne den nåværende virksomheten med kontrafaktiske scenarioer, dvs. med mulige verdener der skolen styres på en annen måte. Jeg viser her tilbake til det mulig fordreide bildet elevene kan utvikle om aktørskap og årsakssammenhenger.

I en artikkel i Klassekampen retter Gunn Imsen en sterk kritikk av staten, og hevder at den bedriver detaljstyring av tankegangen hos den oppvoksende generasjonen:

I troen på at et perfekt regelverk og perfekte mål vil føre til høy kvalitet, har myndighetene glemt at vi er et demokratisk land hvor enkeltmenneskets rett til å tenke, mene og uttrykke seg fritt står sentralt. Vi står derfor overfor det viktige spørsmålet om hvor langt staten kan gå i å detaljstyre tankegangen hos den oppvoksende generasjon gjennom skolens kompetansemål, uten å ende opp som totalitær (Imsen, 2016).

Jeg oppfatter denne kritikken som legitim. Mye av det jeg har drøftet over, handler om en detaljstyring av tankegangen. Samtidig har jeg gått mer i detalj om hva jeg mener at en slik detaljstyring handler om, og noen mulige konsekvenser av den.

I denne delen om autonomi har vi sett at ideen om å ”kunne ha handlet annerledes” står sentralt i idealet om det autonome individet. Det å tenke kontrafaktisk og sammenligne scenarioer utgjør dermed en av kjernekapasitetene til det ideelle autonome individet. Samtidig innebærer det å være autonom at man ikke er styrt av noe utenfor seg selv. Jeg har sett på noen problemer med dette idealet Det første dreide seg om det doble aspektet (Feinberg, 2010), der *jeg* styrer over *meg*, og ingen styrer over *jeg*. Oppfatningen om at det autonome individet bruker friheten sin i et deterministisk univers, leder til konklusjonen at individets egne ønsker og begjær ikke er en del av det autonome, men en del av det deterministiske universet det styrer over. Dersom kriteriene for det autonome samtidig må formidles utenfra, er det en fare for at myndigheter og autoriteter opptrer som ”fornuftens allmakt”. Da er det skolens ønskemål som blir bestemmende. Deretter benyttet jeg begrepsparet autonomihemmende og autonomifremmende om ytre påvirkning. I denne sammenhengen drøftet jeg blant annet problemet med at kriterier for det autonome kommer utenfra i oppdragelses- og utdanningssammenheng.

Vi så at den generelle læreplanen synes å vektlegge elevens ansvar framfor å omtale frihet eller autonomi. Akkurat denne delen av læreplanen synes skolen å følge opp. Kravet til elevene om å stå ansvarlige for valgene sine, har kanskje aldri vært så høye i skolehistorien. I det siste avsnittet så jeg på betydningen av å kunne angi grunner for handling for den autonome. I dagens skole med utallige mål og kriterier, er det å begrunne de egne valgene nærmest blitt et slags mantra. Jeg mener at en berettiget kritikk er at skolen inngir en illusjon om at det er mulig å forklare og kontrollere seg selv ned til minste detalj. Dette kan skape en virkelighetsforståelse om at alt kan og skal forklares. Når fokuset samtidig ligger ensidig på elevenes indre og handlingene deres, uten mulighet for kritisk refleksjon over hvem som setter premissene, er dette en form for manipulasjon.

6 Handlingsbegrepet

Handlingsbegrepet står sentralt i oppgavens problemstilling. Kontrafaktisk tenkning baserer seg på antagelsen om at en situasjon eller hendelse kunne ha resultert i flere mulige utfall, og når denne situasjonen inkluderer menneskers aktiviteter, innebærer det også antagelser om et visst handlingsbegrep. Når vi i dagliglivet gjennom retrospeksjon sammenligner det vi mener skjedde med ulike tenkte scenarioer, utfører vi en form for tankeeksperimenter der vi vurderer våre egne og andres handlinger. Konklusjonene som trekkes er gjerne basert på en type betingelse på formen: “hvis, så”. Ettersom kontrafaktisk tenkning gjennomsyrrer tankene og språket vårt, er eksemplene på kontrafaktiske resonnementer uendelige. Eksempler kan være en elev som tenker: “Hvis jeg hadde forberedt meg bedre på diktanalyse, så hadde jeg fått et bedre resultat på eksamen”, eller “hvis jeg ikke hadde hørt på rådet fra foreldrene mine om hvilken studieretning jeg skulle gå, men i stedet hadde hørt på meg selv, så hadde jeg vært mer motivert for skolearbeidet nå”. Det kan også dreie seg om forskjellen på liv og død. En historie fra virkeligheten handler om en elev som etter en medelevs selvmord i mange år var plaget av tanker hva på han kunne ha gjort annerledes: “Hvis jeg hadde forstått hvor alvorlig det var og tatt initiativ til å være venner med x, så hadde han kanskje ikke tatt livet av seg”. En lærer kan tenke: “Hvis jeg hadde tatt tak i problemene med uroen i klassen fra dag en, så hadde klassemiljøet fått en bedre utvikling”, eller “hvis jeg ikke hadde avfeid elev x sine kritiske poenger i klassesdebatten, så hadde flere elever turt å engasjere seg i den videre debatten”. Et annet eksempel kan være en lærer som reflekterer over et pedagogisk opplegg som feilet, og hva som kunne vært gjort annerledes. Når vi lurer på *om* utfallet hadde blitt annerledes dersom vi hadde gjort noe på en annen måte enn vi faktisk gjorde, ser vi oss selv som opphav til i det minste noe av det som skjedde. Kontrafaktiske tankeprosesser er dermed en type mental representasjon av oss selv og andre som aktører i verden; vi ser oss selv og andre som aktører som kan skape forandring i verden, årsake noe eller forhindre noe i å skje. Kausalitetsbegrepet knyttet til at mennesket kan virke som en årsak, skiller seg på visse måter fra det kausalitetsbegrepet vi anvender om naturfenomener. I en verden hvor vi ellers forklarer hendelser med såkalte naturlover, forklarer vi gjerne handlingene våre med grunner. Et sentralt tema i nyere filosofi har vært om handlinger kan forklares på samme måte som naturfenomener. Uansett så står kontrafaktiske refleksjonsprosesser sentralt i hypotesedannelser om årsakssammenhenger og aktørskap, og jeg vil dermed hevde at aktørskap manifesterer seg blant annet gjennom kontrafaktisk tenkning.

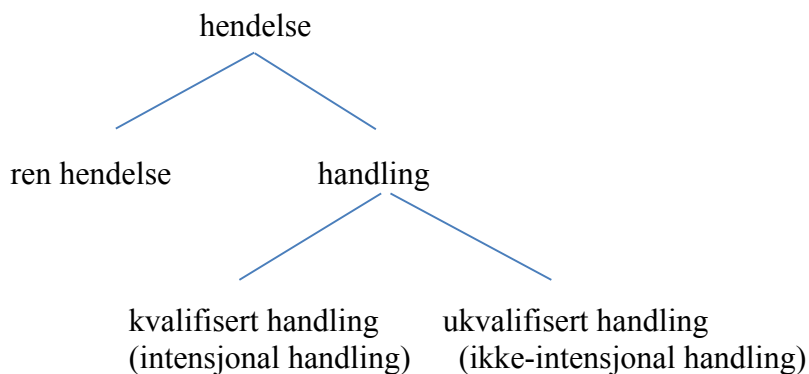
Men hva er egentlig en handling? Hva vil det si å handle? Hva innebærer det å snakke om eller beskrive handlinger? Har handlinger identitet på en slik måte at de hører hjemme i metafysikken? Innen handlingsfilosofien behandles slike og andre temaer knyttet til handlingers natur og identitet. I dette kapitlet vil jeg se nærmere på noen sentrale filosofiske tilnærminger til handling og aktørskap, for å se om det kan kaste lys over hva som skjer når vi tenker kontrafaktisk om oss selv og handlingene våre. Først redegjør jeg for hovedbegreper og distinksjoner knyttet til handlingsteori. Deretter ser jeg på hvordan handlingsbegrepet behandles innenfor ulike disipliner. Så redegjør jeg nærmere for hvordan handlings- og aktørbegrepet skilles ut og defineres, og setter dette i sammenheng med kontrafaktisk tenkning. Videre ser jeg på hvordan skillet mellom handlinger og hendelser behandles innen litteraturen, og drøfter eksempler fra skolen i lys av dette. Deretter ser jeg på hva som karakteriserer handlingsforklaringer.

6.1 Begrepsavklaring

Innen handlingsfilosofien er det et knippe hovedbegreper man forholder seg til. Jeg redegjør først for noen kjernebegreper, som jeg vil utdype videre i det påfølgende. Sentralt står de to størrelsene; aktører og handlinger. Hverken aktør- eller handlingsbegrepet er entydige størrelser, og innen den vitenskapelige litteraturen er det stor variasjon i hvordan begrepene avgrenses og defineres. Felles for ulike teorier er allikevel at man forsøker å skille ut og bestemme noe særegent ved menneskelige aktiviteter. Med *aktivitet* henviser jeg til prosesser som kjennetegnes av bevegelse eller virksomhet, dvs. forløp mennesker på ulike måter går gjennom. Eksempler kan være prosessen med å lage en tegning, løse et mattestykke, fortelle en løgn eller å lytte.

Begrepet *handling* benyttes til å beskrive personers atferd, med henvisning til den delen av atferden som personen *gjør* i motsetning til det som bare *hender* eller *skjer* med personen uavhengig av dennes vilje. Hva som skiller handlinger fra hendelser er et sentralt spørsmål innen filosofien.

Anton Ford (2011) gjør en skjematisk inndeling for å illustrere den standardiserte måten å oppfatte menneskelige handlinger:



Først skilles det mellom rene hendelser og handlinger - de tingene som skjer på grunn av menneskers vilje. Deretter skiller man ut handlinger, som selv om de kommer fra viljen bare kan sies å gjøre det på en indirekte måte. I skjemaet til Ford sorterer disse under ukvalifisert handling, fordi de er “unintentional, or idle, or addictive, or subconscious, or weak-willed, or compelled by force” (Ford, 2011, s. 76). De ukvalifiserte handlingene står dermed i motsetning til handlinger som regnes som handling “par excellence”. Ford benytter seg av begrepet “kvalifisert handling” som en felles betegnelse for ulike handlingsbegrep filosofer bruker i denne sammenhengen. Noen kaller det for *intensjonal handling*, mens andre bruker benevnninger som *fullblods*, *helhjertet*, *autonom*, *formålsrettet* og *fri handling*. Jeg vil heretter benytte begrepet *intensjonal handling* i stedet for kvalifisert handling, både fordi det er et vanlig begrep innen handlingsteori, og fordi det har god resonans med pedagogisk terminologi. Eksempler på intensjonale handlinger kan være det å stjele, tegne, danse, spise etc. Eksempler på ikke-intensjonale handlinger kan være det at jeg lager spor i snøen når jeg går ute om vinteren, eller at jeg uforvarende klemmer fingeren til en annen person når jeg lukker en dør. Andre eksempler på ikke-intensjonale handlinger er det å blunke, snuble, eldes, sove etc. Ut fra denne kategoriseringen kan vi si at alle menneskelige aktiviteter befinner seg på et kontinuum mellom intensjonale og ikke-intensjonale handlinger.

En viktig bemerkning Ford gjør til denne klassiske divisjonsmetoden, er at *handling* betraktes som en type *hendelse* eller underkategori til *hendelser*. Det har vært en vanlig antagelse at filosofenes viktigste oppgave er å fortelle hva en hendelse må være utover det å være en hendelse, for å kunne regnes som en handling (Ford, 2011, s. 77). Denne antagelsen

har også implikasjoner for hvordan man oppfatter at handlinger kan forklares. Det har lenge vært et stridstema blant filosofene hvorvidt handlinger kan forklares på lik linje med fysiske hendelser, eller om handlinger er av en slik natur at de ikke kan la seg forklare på samme måte. Bidragene fra Donald Davidson (1963) og Elisabeth Anscombe (1963), to av pionerene innen handlingsfilosofien, har vært med å legge grunnlaget for det som regnes for *standardoppfatningen av handling*. Denne oppfatningen bygger på to sentrale påstander. For det første regnes begrepet om intensjonal handling som mer fundamentalt enn begrepet om handling (eller ikke-intensjonal handling), og for det andre er det en nær forbindelse mellom intensjonal handling og det å handle for en grunn. Den først påstanden har sin bakgrunn i at en og samme handling kan være intensjonal under én beskrivelse og ikke-intensjonal under andre beskrivelser (Anscombe, 1963; Davidson, 1963). Et eksempel på en intensjonal handling er at jeg åpner vinduet for å si noe til en som står utenfor. At jeg samtidig slipper kald luft inn i rommet, er en ikke-intensjonal handling. Den ikke-intensjonale handlingene med å slippe inn kald luft, er underordnet og avhengige av den intensjonale handlingen som er å åpne vinduet for å snakke til noen. Selv om Anscombe og Davidson delte synet på at handlinger kan forklares i termer av intensjonaliteten ved handlinger, så skilte de seg i synet på hva slags forhold det er mellom hendelser og handlinger, og hvorvidt grunner kunne være årsaker til handlinger. Davidson begynte på 1960-tallet sitt berømte forsvar for posisjonen at forholdet mellom grunner og handlinger er en type kausal forbindelse. I denne “kausalteorien” om handlinger, gikk Davidson ut fra at handlinger er en type hendelser med en bestemt kausal historie, hvor de er handlinger nettopp på grunn av denne historien. Anscombe (1963) delte Davidson sitt syn på at handlinger kan forklares i termer av intensjonaliteten ved intensjonale handlinger. Derimot avviste hun antagelsen om at grunner kan være årsaker, og at handlinger kunne være en type hendelse slik Davidson hevdet. I og med dette videreførte og utdypet hun Wittgensteins tanker om intensjonsbegrepet. Taylor (1992), som også sto i den Wittgensteinske retningen, mener for eksempel at oppfatningen av mennesket som kausal faktor - en faktor som kan forårsake hendelser uten å være en hendelse selv, er så ulikt den vanlige filosofiske oppfatningen av en årsak, at det burde ikke engang hete det samme. Han begrunner dette med at det å være en årsak vanligvis innebærer: “being an antecedent sufficient condition or set of conditions” (Taylor, 1992, s. 52). Motsetningen mellom disse retningene kan synes fundamental, ettersom det dreier seg om handlingers ontologiske status, og om synet på kausalitet. Hvorvidt man mente at grunner kunne være årsaker eller ikke, var da også lenge et essensielt spørsmål i kontroversen mellom

naturvitenskapen og humanvitenskapen. Jeg kommer tilbake til handlingsforklaringer i kapittel 6.5.

En *aktør* er en som kan skape forandring i verden; årsake noe eller forhindre noe i å skje. En aktør er dermed en som har kapasitet til å handle, og begrepet *aktørskap* betegner utøvelsen av denne kapasiteten. I motsetning til ting som skjer på grunn av tilfeldigheter eller andre forutgående årsaker som ikke er knyttet til personers gjøren og laden, så er det aktøren som gjør *handlingen*. Vanligvis benyttes begrepet aktørskap om *intensjonale handlinger*. Denne oppfatningen om aktørskap har røtter langt tilbake til for eksempel Aristoteles og Hume. Innen nyere handlingsfilosofi er det særlig Anscombe (1963) og Davidson (1963) som la grunnlaget for dagens forståelse av aktørskap og intensjonale handlinger. Selv om de to altså skiller seg i spørsmålet om grunner kan være årsaker til handlinger, så deler de oppfatningen at handling kan forklares gjennom handlingers intensjonalitet.

Som personer oppfatter vi oss ikke bare som aktører, men også som subjekter for opplevelsene våre. Rowland Stout (2005) påpeker hvordan vi på den ene siden oppfatter oss som subjekter som er reseptive til hvordan ting er, og på den andre siden opptrer vi som aktører i verden (Stout, 2005, s. 1). Her vil jeg legge til at vi i en forstand også er subjekter for eget aktørskap, da vi er reseptive ikke bare til det som skjer oss eller hva andre gjør, men også til det vi gjør selv.

Begrepet *atferd* kan ha flere betydninger. Atferd benyttes gjerne om ytre observerbar virksomhet, og inkluderer både intensjonale og ikke-intensjonale handlinger. Begrepet kan også ha en bredere betydning som innbefatter alt det en person foretar seg både på det ytre og det indre planet.

Det er noen kategorier av handlinger jeg vil trekke fram, fordi de er spesielt relevante med tanke på kontrafaktisk tenkning og betydningen av hypotetiske scenarioer. En av disse kategoriene er det Gilbert Ryle (1973) kaller for “opphevende handlinger”¹⁴. Dette betegner en rekke intensjonale handlinger vi utfører for å forhindre andre ting, som ellers ville eller kunne ha skjedd. Eksempler kan være det å *motsette seg*, *hindre*, *frigjøre*, *gjemme seg* etc. (Ryle, 1973, s. 81). En annen kategori er det Ryle (1973) kaller for “negative handlinger”.

¹⁴ Gilbert Ryle kaller de for “nullifying acts” (Ryle, 1973, s. 81).

Disse kan også kalles *unnlatelser*, og karakteriseres av aktørens intensjonale ikke-opptreden av spesifikke handlinger. Eksempler er å *utsette*, *ignorere*, *unngå* etc. Det finnes også kategorier av *unnlatelser* som er ikke-intensjonale, for eksempel å *glemme*. Men selv om de regnes for ikke-intensjonale, er det ikke uvanlig at man i en del sammenhenger kan gjøres ansvarlig for disse. Den ontologiske statusen til negative handlinger er omdiskutert, fordi fravær av handling ikke kan være en faktisk eller fysisk årsak til noe, i hvert ikke slik vi vanligvis tenker om kausale sammenhenger. Men i *common sense* opererer vi med letthet med unnlatelser som årsaker, da en unnlatelse forstås i lys av den handlingen som unnlates og dens hypotetiske resultat. I dagliglivets språk omgås vi med mange “negative handlinger”, og evaluerer det som skjedde i lys av det som kunne eller burde ha skjedd, slik at vi kunne ha fått et ønsket utfall. Det er mange sammenhenger der man kan stilles til ansvar for hvordan verden ble, sammenlignet med hvordan den ideelt kunne ha blitt dersom man ikke hadde unnlatt å handle på en spesifikk måte. Michael S. Moore (2010) sier at unnlatelser er simpelthen fraværende handlinger (Moore, 2010, s. 28). Han påpeker at unngåelser ikke kan være årsak til noe, og argumenterer videre med at når jeg unnlater å forhindre skade, så gjør jeg ikke verden verre; “I only fail to make it better” (Moore, 2010, s. 28-29). Det innebærer at hvis jeg unnlater å redde livet til en person, så gjør jeg ikke verden verre, jeg unnlater bare å gjøre den bedre. Det er et kontroversielt standpunkt, som taler sterkt imot våre moralske intuisjoner. En del vil nok mene at det ikke å redde liv er å gjøre verden verre. Samtidig griper problemstillingen inn i kjernen av noen av dilemmaene knyttet til kontrafaktisk tenkning, der nettopp hypotetiske scenarioer og sammenligninger mellom scenarioer står sentralt. I et kontrafaktisk perspektiv finnes det uendelig mange måter verden kunne ha vært bedre på, og implisitt i dette, uendelig mange måter å unnlate å gjøre verden bedre på. Selv i eksemplet med det å redde liv, som er en grunnleggende handlingsregel for de aller fleste, kan det allikevel bli store kontroverser, jf. båtflyktningene i Middelhavet. Når det gjelder hvordan vi bedømmer unnlatelser eller fravær av handling, så er ifølge Byrne (2005) de fleste mer bekymret for skade som er forårsaket av en aktiv handling, enn skade forårsaket av å unnlate å handle. For eksempel anses det som verre at noen dør på grunn av en vaksine, enn at de dør på grunn av en sykdom som kunne vært unngått ved vaksinerings. Mange vurderer for eksempel at det er verre å fortelle en løgn enn å holde tilbake informasjon, selv om intensjonen om å bedra er lik i begge tilfellene (Byrne, 2005, s. 44).

Før jeg går videre ønsker jeg å kommentere en tilbakevendende utfordring når jeg skriver om og drøfter meg gjennom det begrepsmessige landskapet knyttet til aktørskap og handlinger. Begreper som *handling* og *hendelse*, eller uttrykk som det å *gjøre* noe eller at ting *skjer*, har ulike betydninger i forskjellige kontekster både i dagligspråket og i det vitenskapelige språket. Tvetydigheten til disse begrepene og uttrykkene er noe av bakgrunnen for at det er nyttig å analysere nærmere hvordan vi tenker og kommuniserer rundt disse fenomenene. For å si det med Lars Løvlie: “Vi bør i særlig grad vokte oss for våre ord og passe oss for våre begreper, for de kan sløve og forføre” (Løvlie, 2008, s. 62). Men selv det å redegjøre for og drøfte begreper og fenomener knyttet til handlinger og hendelser, er en utfordring i seg selv, fordi jeg lett går tom for begreper og uttrykk hvis jeg ikke skal benytte de samme begrepene i prosessen med å redegjøre for de. Dette har sin bakgrunn i at de utgjør en så essensiell del av språket og virkelighetsforståelsen vår. Løsningen min har vært i en del tilfeller å benytte begrepet *aktiviteter* i stedet for *handling*, for å unngå å implisere en idé om for eksempel *intensjonale handlinger*. Men andre ganger benytter jeg allikevel begrepet *handling* om situasjoner der det fremdeles ikke er klart hvorvidt noen handlet intensjonalt eller ikke, rett og slett fordi det er slik vi snakker. Jeg tror og håper at det fremgår av konteksten hvilken type det dreier seg om. Begrepet *hendelse* benytter jeg først og fremst for å betegne ting som vi ikke har kontroll over, dvs. ting som typisk *skjer* oss. Men i noen tilfeller benytter jeg det for å beskrive totaliteten av en situasjon, der også intensjonale handlinger inngår, slik vi gjør i en del tilfeller i dagligspråket. Et eksempel fra hverdagen er spørsmålet foreldre ofte stiller til barna sine: “Skjedde det noe på skolen i dag?”. Det inkluderer naturlig nok også barnets egne handlinger og ikke bare hva som pågikk uten at av barnet selv var involvert i viljemessig forstand.

En annen utfordring i dette begrepsmessige landskapet om aktører og handlinger, er at flere av begrepene benyttes i pedagogiske sammenhenger med en implisitt normativ dimensjon. For eksempel har aktørbegrepet innenfor konstruktivistisk læringsteori blitt satt opp mot begrepet *brikke*, der en aktørlignende selvforståelse er å foretrekke framfor en brikkelignende. Begrepsparet passer godt sammen med begrepene om fri vilje og determinisme, der brikken er determinert, mens aktøren handler fritt og er årsaksbringende. Det foretrukne aktør-synet er ikke bare noe som fremheves innenfor pedagogikkfaget, men er en del av en utbredt tenkning i den moderne vestlige kulturen. I *Pedagogisk ordbok* (2008) står det følgende under “aktør”:

Handlende og målsøkende person. Gir uttrykk for den oppfatning at mennesket ikke bare er en mottakende (reseptiv) skapning som passivt lar seg forme av krefter i familie, lokalmiljø og samfunn, men at det først og fremst er aktivt, villende og målsøkende (Bø & Helle, 2008, s. 14)

Definisjonen av aktør i ordboka er til dels den samme som vi finner i handlingsfilosofien, men etter min mening har den også undertoner av et foretrukket menneskesyn (eller elevsyn) i en pedagogisk kontekst.

Atferd er også et begrep som benyttes hyppig i pedagogikk- og skolesammenheng, og da gjerne sammen med andre begreper som *problem*atferd, *uønsket* atferd, *utfordrende* atferd, *antisosial* atferd, *prososial* atferd etc. Atferdsbegrepet fungerer nærmest som et “buzzord” i skolepolitiske sammenhenger, og det finnes i dag et stort marked for de som tilbyr atferdsmodifiserende programmer til etatene og skolene. I denne sammenhengen er begrepet ofte forbundet med noe som både kan og bør kontrolleres og reguleres utenfra, gjennom ulike tiltak og intervensjoner. Ikke minst er man i dag opptatt av å forebygge ”problematferd”; det snakkes mye om å være pro-aktiv i forhold til mulige framtidige atferdsproblemer, dvs. kontrafaktiske eller pro-faktiske atferdsproblemer. Selv om atferdsbegrepet i pedagogisk og skolepolitisk sammenheng inkluderer både *ønsket*, *mer nøytral* og *uønsket atferd*, og er ment å gi uttrykk for nøytralitet, benyttes det først og fremst med referanse til noe uønsket og problematisk som man kan forbygge og redusere. Uttrykket å få “nedsatt i atferd”, dvs. nedsatt karakter i oppførsel/atferd, er noe mange elever forholder seg til, og fremsettes av skolen som et utfall de kan forvente seg hvis de ikke følger skolens normer og regler. På samme måte som begrepet “klimaendring”, som i seg selv ikke sier noe om årsakene til endringer i forventet værmønster, i en del tilfeller oppfattes å implisere at endringene er menneskeskapte, gir nok begrepet om atferd i skolesammenheng lett assosiasjoner til problematisk og uønsket atferd. Når jeg benytter begrepet atferd, er det imidlertid for å beskrive menneskelige fenomener og ulike oppfatninger og definisjoner av disse, på en mest mulig nøytral og deskriptiv måte.

6.2 Handlingsbegrepet innenfor ulike disipliner

Menneskelig handling er et komplekst tema som behandles på forskjellige måter innen for ulike disipliner. Man kan sette et grovt skille mellom de spørsmålene om handling som reises innen *etikken* og de som reises innen *handlingsfilosofien*. Spørsmålene innen etikken tar utgangspunkt i moralpsykologiske aspekter og tillater en rikere oppfatning av menneskelivet der motivasjon og forpliktelser står i sentrum (Hornsby, 2013). Innen handlingsfilosofien har man vært mer opptatt av forholdet mellom kropp og sinn, samt mental kausalitet. Med mental kausalitet menes at mentale hendelser kan forårsake andre hendelser, mentale eller fysiske (Hansen, 2016). Dette er et grovt skille, og det finnes teorier som bygger på aspekter som gjenfinnes i begge områdene. Den moralske siden ved menneskelivet har vært en vesentlig motivator for handlingsteori. Et vanlig synspunkt er at jo mer man kan si om handlingers natur, desto større synes muligheten for å kunne utlede moralske prinsipper, eller å kunne avsløre feilaktige sådanne. Christine Korsgaard (1996) er et eksempel på en moralfilosof som søker å begrunne moralens normative autoritet ved å ta utgangspunkt i vår kapasitet til å handle rasjonelt. Korsgaard har særlig markert seg ved å utvikle Kants filosofi i nye retninger, og drøfte grunnlaget for moralsk forpliktelse for det moderne mennesket. Korsgaard stiller seg i en tradisjon med etiske rasjonalister når hun argumenterer for at det finnes uavhengige rasjonelle standarder for handling, og at disse standardene peker ut det moralske liv som det meste rasjonelle. Ikke alle er enige i at teorier om mennesket som aktør og handlende vesen, kan danne grunnlag for en moralfilosofi. Filosofen Kieran Setiya (2007) hevder for eksempel at en handlingsteori ikke kan gi oss de uavhengige normative standardene som en teori om praktisk rasjonalitet og moral trenger. Heine Alexander Holmen (2011) påpeker at det allikevel er en dyp forbindelse mellom disse to domenene, da handlingsfilosofiens primære spørsmål handler om den typen handlinger vi forbinder med rasjonalitet, dvs. handlinger som er knyttet til valg. Ettersom valg er knyttet til verdier, er de også gjenstand for moralsk ros eller klander. Dermed kan man si at handlingsfilosofiens fokus “is the performance that grounds moral accountability” (Holmen, 2011, s. 10).

Kort oppsummert er likheten mellom etikken og handlingsfilosofien at de har menneskelige handlinger som sitt undersøkelsesobjekt. De skiller seg imidlertid fra hverandre ved at moralfilosofene først og fremst er opptatt av normative spørsmål om hvordan vi bør handle og leve, samt det å drøfte det ontologiske grunnlaget for moralske vurderinger, mens handlingsfilosofene er opptatt av hvordan handling oppstår og hva forklaring av handling er.

6.3 Handling og aktør

Uløselig knyttet til et begrep om handling er aktørbegrepet. *Aktøren* er som tidligere nevnt den kausale faktoren, den som *gjør* handlingen, i motsetning til hendelser som bare *skjer*. En klassisk måte å karakterisere handling har vært å knytte handling til en unik kapasitet hos mennesket. Mens resten av naturen er styrt av ufravikelige kausale lover og dermed underlagt hendelser, har mennesket den kapasiteten at det selv kan initiere og frambringe handling. Ifølge dette klassiske synet kan mennesket være en kausal faktor som ikke er kausalt nødvendiggjort av forutgående hendelser. Denne oppfatningen er omstridt, særlig med tanke på nyere forskning om kausalitet. Tidligere var kausalitet nærmest ensbetydende med determinisme, mens det i dag er vanlig å formulere kausalitet i probabilistiske termer. Dette innebærer at kausalitet anses som kompatibelt med indeterminisme. Indeterminisme som metafysisk rammeverk gir nye muligheter til å drøfte kausalitet og aktørskap. Vi så imidlertid i kapittel 4.3 at indeterminismen ikke er svaret på alle problemene. Indeterministene har problemer med å begrunne at handlinger som oppstår på bakgrunn av tilfeldigheter, kan være frie og moralsk ansvarlige.

Jeg innledet denne delen med et klassisk perspektiv på aktørskapet, der *aktøren* er den kausale faktoren, den som *gjør* handlingen, i motsetning til hendelser som bare *skjer*. På samme er tid er handlingsbegrepet innen litteraturen ikke entydig. Hvor vi trekker opp grensen mellom hendelser og handlinger, dvs. mellom det som *skjer* oss og det vi *gjør*, er ikke uproblematisk. Variasjonene i grenseoppgangene mellom hendelser og handlinger gjenspeiler til dels de polare spenningsforholdene i den metafysiske ”fri vilje-determinisme” debatten. Handlingsfilosofene er imidlertid opptatt av å bestemme handlingers natur, og ikke hvorvidt vi i det hele tatt handler fritt.

Det finnes en *hypotetisk* og en *kategorisk variant* av spørsmålet: “Kunne jeg ha handlet annerledes?”. Det *hypotetiske* spørsmålet stilles ofte i det praktiske livet, med den underliggende ontologiske antagelsen at det er mulig for oss å kunne ha handlet annerledes enn vi faktisk gjorde; dvs. antagelsen om alternative utfall. Innen filosofien og metafysikken er det å *kunne ha handlet annerledes*, et vanlig premiss når det argumenteres for fri vilje og ansvarlighet. Den *kategoriske* varianten av spørsmålet “kunne jeg ha handlet annerledes?” er en ontologisk problemstilling. Da spør vi om det i det hele tatt er mulig for oss å kunne ha handlet annerledes enn akkurat slik vi gjorde, om spør samtidig om prinsippet om alternative

muligheter er sant. I denne sammenhengen er handlingsbegrepet noe misvisende, fordi et begrep om handling nærmest utelukkende, både i *common sense* og vitenskap, impliserer en bevisst vilje bak handlingene, og at ulike handlingsalternativer står åpne for oss. Frankfurt viser til hvor utbredt oppfatningen om alternative muligheter er:

Practically no one, however, seems inclined to deny or even to question that the principle of alternate possibilities (construed in some way or other) is true. It has generally seemed so overwhelmingly plausible that some philosophers have even characterized it as an a priori truth (Frankfurt, 1969, s. 829).

Innen den metafysiske fri vilje-debatten, er denne common sense-antagelsen om handlingers natur under press. I kapittel 4.3 skrev jeg om inkompatibilistene som hevder at tesen om determinismen ikke er forenlig med fri vilje. Dette innebærer et annerledes syn på aktørskap. Et sentralt moment ved aktørbegrepet innebærer det å ha en bevisst kontroll over aktiviteter og bevegelser, der en handling skjer med viten og vilje. Ifølge tesen om determinismen er enhver begivenhet årsaksbestemt på en slik måte at hver etterfølgende tilstand ved universet er uunngåelig. Ifølge inkompatibilistene som hevder at tesen om determinismen er sann, er dermed alle handlinger uunngåelige. Mennesker agerer ikke på bakgrunn av valg mellom reelle alternativer og har ingen egentlig kontroll over det de gjør, og opplevelsen av å velge mellom handlingsalternativer er en illusjon. De kontrafaktiske scenarioene vi forestiller oss *kunne ha skjedd*, er ifølge tesen om determinismen ikke reelle alternativer eller muligheter. Dette gjelder uansett hvor nært de måtte ligge opptil virkeligheten. En tese om virkelighetens beskaffenhet der faktorer bortenfor vår kontroll produserer alle handlingene våre, undergraver den gjengse oppfatningen om oss selv som frie og ansvarlige aktører. De av inkompatibilistene som mener at determinismen er sann, hevder derfor at det er en metafysisk umulighet for mennesker å være frie aktører. Tilhengerne av indeterminismen støter også på utfordringer når de skal argumentere for hvordan handlinger som er oppstått på grunn av tilfeldigheter kan tilskrives frie og ansvarlige aktører. Det finnes imidlertid mange varianter av sofistikerte løsninger som forsøker å forhindre at den frie aktøren forsvinner i et deterministisk eller indeterministisk univers. Mens kompatibilistene argumenterer for at fri vilje er kompatibel med determinisme, argumenterer for eksempel libertarianere for at indeterminismen er nødvendig for fri vilje. Det er også filosofer som mener at det metafysiske rammeverket ikke er avgjørende for om vi har fri vilje.

De fleste teoretikere er imidlertid enige om at mennesker *opplever* at de har fri vilje når de overveier og velger mellom flere handlingsalternativer, selv om inkompatibilister hevder at denne opplevelsen er en illusjon. Ved kontrafaktiske tankeprosesser er det vanlig å gjøre om på hendelser som faller innenfor en *common sense* oppfatning knyttet til menneskers handlingskapasitet. Det dreier seg oftest om *intensjonale handlinger*, og de kontrafaktiske scenarioene ligger gjerne tett opptil en oppfatning av den faktiske virkeligheten der naturens lover holdes konstant. Men vi tenker også kontrafaktisk om ikke-intensjonale handlinger. Det er ikke uvanlig å gremme seg over *ikke-intensjonale handlinger* som kanskje oppstod som en utilsiktet konsekvens av en *intensjonal handling*. I en del tilfeller er det heller ikke klart hvorvidt noen handlet intensjonalt eller ikke, og en del handlinger forblir i grenselandet mellom disse uansett hvor mye det diskuteres og analyseres i etterkant. Innen rettssystemet er det for eksempel et sentralt moment å avgjøre om en persons handlinger ble utført i god tro, eller om det er grunn til å mene at vedkommende ikke var uvitende om mulige utilsiktede konsekvenser av handlingen.

For å oppsummere så langt, så ligger antagelsen om oss selv og andre som årsaksbringende og dermed som aktører, til grunn for den kontrafaktiske tenkningen. Samtidig handler en del av de kontrafaktiske tankeprosessene om å utlede hva som var opptil oss, og hva som bare skjedde og som vi ikke hadde kontroll over. Idet vi setter i gang med kontrafaktiske resonnementer etter en hendelse, er vi ikke nødvendigvis klar over hva som var handlinger og hva som var hendelser, selv om vi kan ha noen hypoteser om saken. De kontrafaktiske tankeprosessene fungerer altså som en metode for å teste ut hypotesene våre om virkelighetens beskaffenhet knyttet til årsakssammenhenger og aktørskap. Påliteligheten ved denne metoden avhenger av en rekke faktorer. Jeg har tidligere omtalt problemet med “hindsight bias”¹⁵ og ”snikende determinisme”, som innebærer at kunnskapen vår om utfallet ved hendelser, påvirker minnet og oppfatningen av hvilken kunnskap vi hadde på det tidspunktet vi tenker tilbake på (Fischhoff, 1974). Med et visst forbehold knyttet til ulike definisjoner av handlingsbegrepet, vil jeg hevde at vi egentlig ikke kan kalle noe for en handling før det har fått utspille seg i tid og rom. Dette henger sammen med at vi ikke fullt ut kjenner omfanget av aktivitetene våre før de har utspilt seg i en kontekst. I forlengelsen av tendensen med “snikende determinisme”, er det derfor rimelig å anta at vi også er “biased”

¹⁵ Bias: “benyttes om utsagn, valg og vurderinger som på systematisk måte avviker fra det som stemmer overens med virkeligheten” (Svartdal & Teigen, 2014). Jeg velger å benytte det engelske ordet, fordi det ikke finnes et norsk ord som fanger inn hele betydningen.

når det gjelder oppfatningene våre av handlinger og tilskrivelser av aktørskap. Resultatene av våre egne og andres aktiviteter kan påvirke de tilskrivelsene av aktørskap vi gjør i ettertid, dvs. hvordan vi bedømmer grader av intensjonalitet ved handlinger. Tendensen kan være at vi tilskriver oss selv eller andre aktørskap der hensiktene ikke var så tilsiktet som de kan framstå i ettertid, eller motsatt, at vi tilskriver en mindre grad av aktørskap enn det som var tilfelle. Med “mindre grad av aktørskap” refererer jeg til at en persons handlinger befinner seg på den ikke-intensjonale siden av handlings-kontinuumet. Det sosiale samspillet og kommunikasjonen mellom mennesker er komplekse fenomener, og selv om vi ofte langt på vei kjenner våre egne intensjoner og hva som var opptil oss ved en bestemt anledning, kan det allikevel være tilfeller der det er vanskelig å avgjøre hvor skillet går mellom det vi utførte med en bestemt intensjon og det som kalles uforsettlig. Når det gjelder hypotesene våre om andres handlinger, der vi ikke har samme tilgang til hvilke prosesser personen gikk gjennom rundt tidspunktet for aktiviteten, er disse naturlig nok enda mindre pålitelige.

Jeg vil også trekke fram en annen faktor som kan påvirke påliteligheten i denne typen hypotesetesting, og de slutningene vi trekker gjennom kontrafaktisk tenkning. Vi har normalt forventninger til hvordan ulike mennesker kan komme til å opptre i ulike situasjoner. I klasserommet kan dette ta seg form av roller eller “merkelapper” elevene har blitt tildelt, alt fra å være såkalte “englebarn” til “bråkmakere”; merkelapper som kan farge lærernes felles fortolkningshorisont av hendelser. Slike merkelapper uttrykker oppfatninger om hva slags handlinger elevene tenderer å utøve, og kan fungere som forenklende rettesnorer for hvordan elevene skal oppfattes og håndteres. Dette kan lede til at hendelser eller situasjoner ikke undersøkes nøye nok, slik at elever tilskrives handlinger de ikke gjorde. Det motsatte kan også skje i tilfeller der elever ikke forventes å ville utføre noe de faktisk har gjort. Dette kan like gjerne handle om faglige prestasjoner som generell atferd. I forbindelse med en alvorlig mobbesak uttrykte en mor seg til en journalist om hvordan hun opplevde at hennes sønn på ni år ble behandlet på skolen: “Jeg sier ikke at han er noe englebarn, men at han har blitt gjort til syndebukk. Automatisk har han fått skyld i episoder uten at skolen har kartlagt hva som egentlig har skjedd” (Nygård & Samdal, 2013). Uavhengig av riktigheten ved disse påstandene, så er nok dette en opplevelse mange kan kjenne seg igjen i. Det hun anklager skolen for, er at skolen har antatt sønnens aktørskap knyttet til hendelser, dvs. at de konkluderte med at han var ansvarlig for episoder, uten at de grundig undersøkte hva som faktisk hadde skjedd. Indirekte anklager hun de for å ha gitt sønnen en type “merkelapp” som jeg beskriver over.

6.3.1 Hvem kan handle?

Der noen skiller mellom levende vesener inkludert mennesket, og objekter uten liv, trekker andre opp grensen mellom levende vesener med bevissthet og et sinn på den ene siden, og alle andre skapninger og objekter på den andre siden (Pitkin, 1972). For at noe skal kvalifisere som en handling, virker det vesentlig at det finnes en bevissthet, og en kapasitet til å ville og velge å *gjøre* noe. Men mennesker er også underlagt en rekke hendelser, og de fleste teoretikere anerkjenner dermed at mennesker på samme tid er fysiske objekter og dyr, slik at bare noe av adferden deres er handling (Pitkin, 1972, s. 158). I denne sammenhengen er det verdt å nevne at de ulike måtene vi i dag oppfatter menneskets kapasitet med hensyn til å handle fritt, har en lang historisk utvikling.

Hvordan kan vi så skille på handlinger og hendelser? Mange menneskelige aktiviteter som det å nyse, blunke, svette eller hoste kan ikke sies å være aktive handlinger på den måten vi tenker om ansvarlige eller moralske handlinger. Vi vil neppe dømme noen for å få et hosteanfall under en konsert. Hvis derimot anfaller ikke tar slutt etter en tid, vil vi kanskje dømme vedkommende for ikke å forlate konsertlokalet og spare resten av publikummet for forstyrrelsen. Knyttet til denne typen av “ufrivillige” kroppsuttrykk, beskriver for eksempel George Wilson og Samuel Shpall (2012) at personer riktignok *gjør* noe på en minimal måte, men at de allikevel forblir passive gjennom disse “gjøremålene” (Wilson & Shpall, 2012, s. 3). Begrepet passiv her tilsvarer det jeg også kaller ikke-intensjonal. Samtidig med at mange menneskelige aktiviteter kan oppfattes som passive og ikke formålsrettede, kan mange av dyrenes aktiviteter oppfattes både som aktive og formålsrettede, og at dyrene til en viss grad opptrer som aktører. Filosofen Harry Frankfurt (1978) advarer mot en tendens han mener finnes blant filosofer til å drøfte handlingers natur som om aktørskap forutsetter “characteristics which cannot plausibly be attributed to members of species other than our own” (Frankfurt, 1978, s. 161-162). Han mener at vi kan skjelne mellom handlinger og hendelser også hos andre arter enn vår egen. Han bruker som eksempel en edderkopp som går bortover en flate: I den ene situasjonen gjør den det “fritt”, og bevegelsene dens kan tilskrives edderkoppen. I det andre eksempelet har noen lykker med å binde tråder til føttene på den, og får den til å gjøre nøyaktig samme bevegelser. Men i det siste eksempelet kan ikke aktiviteten tilskrives edderkoppen, den er en hendelse som edderkoppen gjennomgår. Ifølge Frankfurt (1978) utøver edderkoppen i det første eksempelet en formålsrettet aktivitet. Han viser til hvordan kontrasten mellom handling og hendelse hos edderkoppen er parallell med kontrasten

mellom det at jeg løfter armen og at den løftes uten at jeg gjør det selv. Men han medgir også at selv om betingelsene for aktørskap er uklare, “it may well be that the satisfaction of these conditions by human beings depends upon the occurrence of events or states which do not occur in the histories of other creatures” (Frankfurt, 1978, s. 162).

6.4 Handlinger versus hendelser

Let us not forget this: when 'I raise my arm', my arm goes up. And the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm? (Wittgenstein, 1997, s. §621)

Det at jeg nå sitter og skriver på masteroppgaven kaller jeg for en handling. Men det at jeg samtidig blunker med øynene eller at jeg eldes, er noe som skjer meg. Hvis en elev holder en presentasjon for klassen, kaller vi det en handling. Men hvis hjertet samtidig slår hardt i brystet på eleven, blikket flakker og hendene skjelver vil nok mange mene at disse tingene *skjer* eleven, og dermed er en type hendelser. Det er ikke noe eleven med viten og vilje *gjør*, snarere er slike symptomer noe en del elever frykter skal skje i lignende sammenhenger. Samtidig vil kanskje noen mene at eleven indirekte kunne påvirket disse i noen grad ved å være bedre forberedt eller lignende. Forfatteren Siri Hustvedt (2012) skrev boka, “Den skjelvende kvinnen”, om prosessen med å analysere egenopplevde erfaringer av voldsomme og ufrivillige skjelvetokter, som oppsto helt ut av det blå av og til når hun skulle tale i forsamlinger. Hun beskriver opplevelsen slik: “Den skjelvende kvinnen følte som meg, og på samme tid ikke som meg. Fra haken og opp, var jeg mitt vanlige jeg. Fra halsen og ned var jeg en skakende fremmed” (Hustvedt, 2012, s. 22). På tross av den møysommelige prosessen med å analysere mulige årsaker til anfallene, og etter hvert å forsone seg med skjelvingene, der hun refererer til seg selv som “den skjelvende kvinnen”, kommer hun aldri fram til at skjelvingene er noe hun *gjør*. Skjelvingene *hender* henne.

De ulike eksemplene over kan plasseres på et kontinuum mellom ytterpunktene *rene hendelser* og *intensjonale handlinger*. I en del tilfeller er det enkelt for oss å skille mellom det noen *gjør*, og det som *skjer* dem, i andre tilfeller er det mer komplisert. Språket i seg selv er ikke en god indikator når det gjelder å skille mellom intensjonale handlinger og hendelser. Donald Davidson illustrerer dette poenget gjennom eksempelet med noen som faller: “In the case of bodily movements we sometimes have a brief way of mentioning a person and an event and yet leaving open the question of whether he was the agent, as: Smith fell down”

(Davidson, 2001, s. 54). Stout skiller mellom tilfeller ved menneskelige aktiviteter der intensjoner er involvert, og tilfeller der de ikke er det, eksempelvis som det å puste eller svette. Han bemerker videre at aktiviteter vi vanligvis ikke oppfatter som intensjonale allikevel kan bringes under en viss kontroll. Han benytter eksemplet med det å gå: "...the right knee bends and the left foot moves forwards in a controlled way, but we do not have to think about it" (Stout, 2005, s. 4). Videre argumenterer han for at denne type automatisk atferd bør regnes som intensjonale handlinger: "...most of what you are doing is done without any thought. If we ruled out all this behaviour out of the category of intentional action, we might not be left with very much" (Stout, 2005, s. 4). Men interessant nok, legger han til, at i tilknytning til det å gå og den påfølgende bøyingen av kneet og bevegelsen av foten: "indeed it is better not to think about these movements as they happen or else you start to limp" (Stout, 2005, s. 4). Det vi skal merke oss her er at selv om han påpeker at bevegelsene (i kneet og foten) er noe som skjer, og at de vanligvis utføres uten noen tanker på de, så argumenterer han allikevel for at de skal regnes som intensjonale. Jeg tolker han dithen at han like gjerne kunne ha sagt: "Jeg gjør det som skjer". Anscombe har en annen tilnærming til intensjonalitet. Anscombe blir betraktet som en av de store pionerene innen moderne handlingsfilosofi. I den banebrytende bok *Intention* som utkom i 1957, påpeker hun det tvilsomme med påstanden: "...when the description of what happens is the very thing which I should say I was doing" (Anscombe, 1963, s. 53). Hun argumenterer for at kunnskapen om hva man gjør ikke handler om observasjon, men at en handling er intensjonal dersom spørsmålet om "hvorfor" har applikasjon. Argumentasjonen hennes er langt mer sofistikert enn jeg kan få gitt uttrykk for her. Imidlertid er det et vesentlig poeng jeg vil trekke ut av dette: ifølge Anscombe kan den eneste beskrivelsen av det man gjør, skje på avstand fra en selv: "It is not the case that I clearly know the movements I make, and the intention is just a result which I calculate and hope will follow on these movements" (Anscombe, 1963, s. 53).

6.4.1 Eksempler fra skolen

Nyansene beskrevet over kan synes som ubetydelige. Men området eller gråsonen mellom rene hendelser og rene handlinger, dvs. de kroppsuttrykk eller aktiviteter som mer eller mindre indirekte kan påvirkes, blir i skolen ofte benyttet som kjennetegn på måloppnåelse for å beskrive kvaliteten på elevens mestring. Det skilles i en del tilfeller lite på *passive* og *aktive* måter å *gjøre* noe på. Man kan lett få inntrykk av at eleven har en kontroll eller bør ha en kontroll over seg selv og sin egen kropp, som man kan diskutere om overstiger det som er

mulig uten at man samtidig reduserer muligheten for faktisk å være *aktiv* i en handling. Dette åpner opp for å drøfte hva slags idealer skolen har om autonomi og aktørskap, og hva slags oppfatning elever danner seg av hva det vil si å handle og være ansvarlig.

Eksempelene under er hentet fra et vurderingsskjema for muntlige presentasjoner, lokalt utformet/ videreutviklet av norsklærere (Melbye & Harboe, 2009). Under rubrikker som “stemmebruk”, “tempo/uttrykk”, “språkføring”, “kroppsspråk”, “blikk” og “engasjement” oppgis en rekke typisk ønskede og uønskede kroppsuttrykk som noe eleven *gjør*: For eksempel: “du snakker for lavt”, “du står urolig”, “du fikler med hendene”, “du har liten øyekontakt med publikum” etc. Dette er ytre observerbar atferd som eleven kanskje ikke er seg bevisst, før læreren eller andre påpeker det. Deretter beskrives mer og mer såkalte ønskede kroppsuttrykk, og disse graderes for eksempel i *under middels*, *middels* og *over middels*. For å oppnå *over middels* og karakteren 5-6 står det for eksempel: “du unngår pauselyder”, “du varierer stemmen”, “du fester blikket flere steder i rommet”, “ansiktsuttrykket stemmer overens med innholdet” og “du bruker kroppsspråket for å understreke det du sier”. For å oppnå høy karakter er det også oppgitt ting du ikke må gjøre: “Du gjør ikke noe som kan avlede interessen fra presentasjonen (f.eks. fikler med penner, håret e.a.)”. Å fikle med penner, håret eller andre ting vil nok ifølge Anscombe regnes som ikke-intensjonale handlinger, mens Stout vil nok hevde det motsatte. Selv om disse sannsynligvis kan bringes under kontroll ved øvelse eller en bestemt oppmerksomhet, mener jeg at dette ikke er ting elever gjør med hensikt. De fleste forstår nok det “å fikle” i denne sammenhengen som noe som skjer relativt ufrivillig, utløst av uro og nervøsitet eller lignende. Oppå det hele skal eleven i tillegg være i stand til å gi uttrykk for noe de fleste forbinder med å være naturlig, levende og spontan: “du står ledig og avslappet”. Til en viss grad kan vi gå inn for å kontrollere vår egen kropps *passive* aktiviteter, men ikke nødvendigvis uten at det går på bekostning av å være *aktiv* i forhold til hovedaktiviteten, som i dette tilfellet er å presentere et bestemt innhold for andre mennesker. På samme måte som vi indirekte kan påvirke pusten vår, blunkingen med øynene, så *kan* vi ta kontroll over mye av det som skjer i kroppen eller hvordan vi uttrykker oss. Allikevel kan resultatet bli at det som til syvende og sist uttrykkes gjennom kroppen nettopp er *kontroll*, og ikke engasjement og spontanitet.

I takt med økt institusjonalisering, mål- og resultatstyring virker det som om det er en økning av mengden krav knyttet til det å vurdere og å kontrollere seg selv. Kriterier for måloppnåelse

med detaljerte beskrivelser er ikke bare laget som hjelpemiddel for lærerne eller den som skal gjøre vurderingene, de deles ut til alle elevene som en instruksjon for å oppnå de ulike målene. En annen utvikling er at beskrivelsene synes å bli mer omfattende. Det er som om reservoaret av beskrivelser blir større og større, der det også vektlegges mange mulige varianter av måter ting ikke skal gjøres på. Beskrivelsene av kroppen og hva den skal uttrykke kan være svært detaljerte; og i visse tilfeller helt konkrete om hvordan blikk, hender, pust, stemme og holdning skal være. I andre tilfeller skal elever vurdere seg selv etter kriterier som “positivt kroppsspråk”. Et eksempel er hentet fra en ukeplan for 4. klassinger, der det står øverst: “Jeg har husket på at positivt kroppsspråk er viktig for klassemiljøet”. Her skal eleven mot slutten av uka vurdere seg selv ved å sette ring rundt ulike smilefjessymboler: ☺ ☹ ☹. Beskjeden og den underliggende antagelsen, synes å være at eleven skal, og dermed bør kunne, *gjøre seg selv* i motsetning til å *være seg selv*. Ifølge Anscombe sitt syn på intensjonalitet innebærer denne instruksjonen at elevene skal gjøre det som skjer dem. Ludwig Wittgenstein sa dette om viljen og verden utenfor:

The world is independent of my will. Even if everything we wished were to happen, this would only be, so to speak, a favour of fate, for there is no logical connexion between will and world, which would guarantee this, and the assumed physical connexion itself we could not again will (Wittgenstein, 2001, 6.373, 6.374)

Jeg mener det er viktig å spørre seg hva slags oppfatning elever danner seg av aktørskap knyttet til hva det vil si å handle og være ansvarlig, når de kontinuerlig skal vurdere seg selv og blir vurdert av andre etter denne type instruksjoner. I Hustvedt sitt tilfelle med skjelvingene er det irrasjonelt å hevde at hun *gjør* det som *rammer* henne ufrivillig, selv om det handler om hennes egen kropp. De nevnte beskrivelsene av hva elevene skal gjøre og ikke gjøre som kriterier og kjennetegn på måloppnåelse, omhandler aktiviteter eller kroppsuttrykk som indirekte kan påvirkes, men ikke uten omkostninger. Forventningene og kravene til elevene framstår som en form for disiplinering som er i strid med idealene om autonomi. Elevene forventes å kontrollere seg selv på en svært detaljert måte, og det blir også formidlet at de blir observert: Det observeres både om de lever opptil målene, og om de evaluerer seg selv riktig etter de. Alt sammen skjer på skolens premisser.

6.5 Handlingsforklaringer

En annen måte å skille mellom handlinger og hendelser, er hvordan disse forklares. Fysiske fenomener forklares vitenskapelig ut fra universelle lover og kausal nødvendighet, mens menneskelige aktiviteter gis kontingente¹⁶ forklaringer ved å vise til “grunner”. Disse kan for eksempel være formål, motiv, hensikter eller resonnementer. Innen handlingsfilosofien er det de såkalte *kausalteoriene* om handlinger som har fått størst gjennomslagskraft hos teoretikerne. Ifølge kausalteorier om handlingers natur er det da også de kausale årsakenes beskaffenhet, som utgjør det essensielle skillet mellom handlinger og hendelser. Genuine handlinger er hendelser som forutgår av et visst mentalt innhold. Ramberg (2004) redegjør for Donald Davidsons¹⁷ syn på forholdet mellom det mentale og det fysiske: “Dette er antagelsen om at våre mentale tilstander – hva vi tenker, føler, ønsker, intenderer, og så videre – har årsaksmessig sammenheng med hendelser i den fysiske verden” (Ramberg, 2004, s. 92). Frankfurt (1978) anerkjenner hvor utbredt kausalteorier om handlinger er, men hevder samtidig at den kausale tilnærmingen ikke er troverdig, og at den ikke kan tilby en tilfredsstillende analyse av handlingers natur (Frankfurt, 1978, s. 157). “In asserting that the essential difference between actions and mere happenings lies in their prior causal histories, causal theories imply that actions and mere happenings do not differ essentially in themselves at all” (Frankfurt, 1978, s. 157). Dette underbygger han med at kausale tilnærminger impliserer at handlinger og hendelser kun differensieres ved noe som ikke eksisterer eller skjer på selve tidspunktet hendelsene pågår, men ved noe som pågår forut for dem; “...a difference at an earlier time among another set of events entirely” (Frankfurt, 1978, s. 157). Frankfurt mener at i stedet for å undersøke hva som skjer *før* handlingen, er det mer plausibelt å undersøke bevegelsene til en person mens de pågår. “What is not merely pertinent but decisive, indeed, is to consider whether or not the movements are occurring under the person’s guidance”. Frankfurt henviser til at en person er i et bestemt forhold til sine egne handlinger, på en måte som skiller seg fra det forholdet personen har til det som bare *skjer* den. En annen filosof, Elisabeth Anscombe (1963), som på mange måter divergerer med Frankfurt i synet på handlingers natur, deler allikevel med Frankfurt dette at hvorvidt det kan sies at en person utfører en handling, avhenger av hva annet som er sant for den personen på

¹⁶ Kontingent: “Det som er mulig uten å være nødvendig. Det faller da sammen med det som verken er nødvendig eller umulig. [...] I nyere filosofi betraktes alle naturlovmessige årsaksforhold som kontingente (ikke logisk nødvendige), men de er ikke derfor tilfeldige i hverdagslivets betydning av dette ordet” (Tranøy, 2011).

¹⁷ Ramberg refererer til *Mental Events* der Davidson (2001) diskuterer forholdet mellom det mentale og det fysiske.

tidspunktet for handlingen, og ikke de forutgående årsakene til den. Anscombe tar utgangspunkt i at vi har en spontan kunnskap om hva vi gjør; en kunnskap som ikke baserer seg på observasjon eller slutninger. Når en person har ikke-observerbar kjennskap til hva den gjør, og samtidig kan besvare spørsmålet ”hvorfor”, så kan man si at hendelsen er en intensjonal handling. Hun hevder også at intensjonal handling og forklaringer med henvisning til grunner eller som svar på ”hvorfor”-spørsmål, ikke kan reduseres til forklaringer gjennom årsaks- og naturlover slik man opererer innen naturvitenskapen. Denne kjennskapen eller kunnskapen kan kalles praktisk, og Anscombe anfører med dette at forholdet mellom personen og dennes handlinger er en epistemisk relasjon. Hun setter dermed to demarkasjonskriterier for intensjonal handling; det første er at vi kan svare på spørsmålet ”hvorfor” og angi grunner for handlingen, og det andre er at vi har en ikke-observerbar kjennskap til det vi gjør.

I denne delen har jeg sett på noen filosofiske tilnærminger til handlingsbegrepet, for å se om det kunne kaste lys over kontrafaktisk tenkning om handlinger. Ettersom kontrafaktiske refleksjonsprosesser står sentralt i hypotesedannelser om årsakssammenhenger og aktørskap, startet jeg med å hevde at aktørskap manifesterer seg blant annet gjennom kontrafaktisk tenkning. Samtidig handler en del av de kontrafaktiske tankeprosessene om å utlede hva som var opptil oss, og hva som rent av skjedde og som vi ikke hadde kontroll over.

Både handlingsfilosofien og moralfilosofien har vist seg å kunne kaste lys over betydninger av kontrafaktisk tenkning. Den modale tenkningen og de modale påstandene er ofte nært forbundet med sosiale normer og perspektiver på moralitet. Vi har sett at det hersker uenighet blant filosofene om hva som regnes som intensjonal handling. Anscombe skiller seg ut ved å reservere intensjonal handling til de tilfellene der spørsmålet ”hvorfor” har applikasjon. Et sentralt poeng hos Anscombe som jeg ønsker å ta med meg videre er følgende; den eneste beskrivelsen av det man gjør, kan skje på avstand fra en selv. Når elevene blir bedt om å vurdere seg selv, må de ta denne avstanden. Jeg viste også til at mange av kjennetegnene på måloppnåelse som skolen opererer med, befinner seg i gråsonen mellom rene hendelser og rene handlinger. Dette innebærer at elevene vurderes etter kroppsuttrykk eller aktiviteter som mer eller mindre indirekte kan påvirkes. Jeg mener at dette går på bekostning av en verdifull tilstedeværelse. I kapittel 5.2.9 så vi at de utallige målbeskrivelsene kan gi inntrykk av at det faktisk er mulig å oppnå en bevisst kontroll over de utallige valgene man tar. I denne delen om handlingsbegrepet har vi sett at en del av handlingene elevene må velge mellom uansett

dreier seg om å kontrollere kroppens ikke-intensjonale og passive aktiviteter. Anscombe sin teori kan med fordel utforskes mer for å belyse den formen for objektivisering som pågår i skolen i dag. Den spontane kunnskapen elevene har om hva de gjør blir helt fortrent, hvis de hele tiden skal trekke slutninger om hva de gjør og/eller observere hva de gjør.

7 Kontrafaktisk livsverden

Kontrafaktisk tenkning og ”kontrafaktiske” emosjoner spiller en fundamental rolle både for det mentale landskapet og for følelseslivet vårt. Som nevnt tidligere er kontrafaktiske tankeprosesser så integrert i tenkningen og kommunikasjonen vår, at det ikke er noe vi vanligvis tenker over at vi holder på med. I tabell 2 viste jeg til hvordan utdanningsfeltet i større grad enn andre samfunnsområder, som normativ sosial praksis, er et pre-kontrafaktisk prosjekt. Selv om tilgrensende temaer som transcendentale idéer, utopisk tenkning, forestillingsevnen eller moralsk dannelse er typiske pedagogiske temaer, vil jeg allikevel hevde at kontrafaktisk tenkning som sådan, er et nærmest utforsket område innenfor det pedagogiske feltet. Noe av målet med oppgaven min er å argumentere for at det kontrafaktiske, som griper dypt inn i så mange kroker og krokar av virkelighetsforståelsen vår, og samtidig er et uunnværlig analytisk verktøy ved legitimering av handlinger, fortjener en større plass i den pedagogiske bevisstheten. Derfor har jeg tatt meg friheten å utvide den typisk pedagogisk- filosofiske definisjonen av utdanning som ”innvielse inn i en rasjonell livsverden” til å implisere utdanning som ”innvielse i en *kontrafaktisk livsverden*”. Videre vil jeg begrunne og utdype denne versjonen av det klassiske begrepet og uttrykket. Deretter vil jeg belyse maktperspektiver ved kontrafaktisk tenkning i skolen, gjennom begrepsparet *instrumentelle* og *utopiske* kontrafaktisiteter. Til slutt skisserer jeg en hypotese om at skolens kontrollordninger og vurderingssystemer innebærer en sterk sosial kontroll, i form av å fungere skamproduserende.

7.1 Det ”utematisk fungerande”

En lang tradisjon innen pedagogisk filosofi har vært å identifisere utdannelsens overordnede mål og ideal om rasjonell tenkning og handling, her formulert av Harvey Siegel (2012) som: ”...that of the fostering or cultivation of rationality” (Siegel, 2012, s. 191). Et sentralt begrep for denne fostringen eller kultiveringen har etter Richard Peters klassiske dannelsesfilosofi, vært analysen av utdanning (education) som *innvielse* (Peters, 1966). Erling Lars Dale sammenfatter innvielsesbegrepet til Peters slik: ”...en innvielse i verdifulle prosesser og aktiviteter på en slik måte at verdiinternalisering finner sted og bevissthetsstrukturer dannes” (Dale, 1992, s. 9). Her ser vi at selve innvielsesbegrepet innebærer det motsatte av indoktrinering. I kapittel 5.2.5 siterte jeg Løvlie: ”... å se læring som innvielse og

indoktrinering i samfunnets lokale normer ” (Løvlie, 2011, s. 748). I og med den betydningen Peters legger i innvielsesbegrepet, mener jeg det blir misvisende å snakke om innvielse og indoktrinering samtidig, slik Løvlie gjør her.

I ettertid har det tilkommet ulike beskrivelser og varianter av utdanning som innvielse. Siegel selv drøfter betydningen av innvielse og rasjonalitet i lys av blant annet Wilfrid Sellars (1997) begrep om ”the logical space of reasons” (Sellars, 1997, s. 76). Inspirert av Sellars formulerer da Siegel utdanning som: ”...initiation into the space of reasons” (Siegel, 2012, s. 195). Et annet sted henviser Siegel til ”...initiation into the rational traditions; and democratic living” (Siegel, 1997, s. 4). Andre varianter på formuleringer kan være ”utdanning som innvielse inn i sosiale praksiser” eller ”innvielse i en rasjonell livsverden”. Ettersom den kontrafaktiske tenkningen utgjør en grunnleggende bestanddel ved rasjonaliteten vår (Byrne, 2005; Hofstadter, 1979) og dermed vår *rasjonelle livsverden*, mener jeg det er rimelig å hevde følgende: Utdanning som innvielse i en rasjonell livsverden, impliserer innvielse i en *kontrafaktisk livsverden*. Vedrørende uttrykket: ”logical space of reasons”, forklarer Sellars at: ”The essential point is that in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says” (Sellars, 1997, s. 76). Med utgangspunkt i dette kan man også si at det å plassere en episode eller en tilstand inn i det logiske rommet av begrunnelser, nødvendigvis innebærer å resonnere kontrafaktisk. Begrunnelser inneholder ofte beskrivelser av kausale sammenhenger hvor vi redegjør for hvorfor noe skjedde på en viss måte og ikke en annen. Ifølge Sellars befinner våre mentale tilstander seg som kunnskap eller observasjoner av ting i et logisk rom av begrunnelser, der ting står i et normativ og dermed rettferdiggjørende forhold til hverandre. Det er alltid en grunn til at noe er en grunn til noe – og det er alltid en grunn til at det er noe bestemt i stedet for noe annet, som er en grunn til noe.

Et eksempel på dette er karakteristikken av det som går under ”problemet med frafallet i videregående skole”. Jeg siterer her fra Utdanningsforbundets nettside: ”Skal en **reducere frafall i videregående opplæring**, må en **betydelig innsats** settes inn **tidlig i utdanningsløpet**” (Utdanningsforbundet, 2016)¹⁸. Det at elever avbryter skolegangen under videregående skole, tolkes gjerne i konteksten av det kontrafaktiske idealet om at de unge skal gjennomføre utdannelse, bli til gangs mennesker og delta i arbeidslivet. Den skolen det ”falles

¹⁸ Uthevelse med fet skrift er fra utdanningsforbundet.

fra” har med sitt samfunnsmandat så fundamentale og normative målsetninger for unge menneskers liv, at valget om ikke å fullføre videregående skole, sjelden betraktes som et ”riktig” veivalg. Det betraktes gjerne ikke engang som et veivalg, men mer som en desperat handling fra de unge. (Jeg bestrider ikke at dette kan være tilfelle). Innenfor de samme normative rammene for oppfatningen og forståelsen av betydningen av utdanning, er det få alternative handlinger som betraktes som fornuftige. Det at ungdom avslutter videregående skole før tiden karakteriseres oftest som en ”kontrafaktisk hendelse” som ikke burde ha skjedd, selv om det i realiteten er slik at mange slutter. Det er også et faktum at de unge slutter på skolen når den ikke lenger er obligatorisk. De har altså for første gang muligheten til å velge ikke lenger å tilbringe dagene sine på skolen. Det blir dermed en bekreftelse på at skolen ikke har lyktes med å få elevene til å gjøre skolens mål til sine egne, slik jeg diskuterte i kapittel 5.2.4 om *Fornuftens allmakt*. Det er grunner til at elever avbryter skolegangen, og det er grunner til at fenomenet omtales og foreslås løst slik det gjør. En mulig hypotese er at det er en sammenheng mellom disse grunnene, men jeg går ikke videre inn på å begrunne en slik hypotese her. Ifølge Sellars teori omtales og karakteriseres hele problematikken innenfor et logisk rom av begrunnelser, der ting står i et normativt og dermed rettferdiggjørende forhold til hverandre. Begrunnelsen for den foreslåtte hypotesen min ville (forhåpentligvis) også inngå i et logisk rom av begrunnelser, ”of justifying and being able to justify what one says” (Sellars, 1997, s. 76).

For at uttrykket *kontrafaktisk livsverden* ikke skal bli altfor løst, vil jeg definere livsverden i denne sammenhengen. Begrepet *livsverden* har en omfattende historie innen fenomenologien, men jeg begrenser meg her kort til en av Habermas mange utlegninger om begrepet: ”...the lifeworld—as the horizon within which communicative actions are "always already" moving” (Habermas, 1987, s. 119). Ifølge Habermas er livsverdenen den horisonten som all kommunikativ handling skjer innenfor, og i dette oppfatter jeg at han snakker om en felles og implisitt forståelseshorisont. Så vender jeg meg til Jon Hellesnes (1988), der han forklarer Habermas sitt begrep om livsverden (”Lebenswelt”), med at livsverdenen er den sammenhengen vi lever ureflektert innenfor: ”Livsverda er ’det utematisk fungerande’, bakgrunnen, det vi til ei kvar tid tek for gitt” (Hellesnes, 1988, s. 90). Hellesnes forklarer videre: ”Vi kan ikkje reise spørsmål ved alt samstundes, og det vi til ei kvar tid ikkje reiser spørsmål ved, men tek for gitt, det som temaet er problematisk på bakgrunn av, dét er livsverda vår” (Hellesnes, 1988, s. 90). Det setter jeg så inn i et kontrafaktisk perspektiv: Når vi i det daglige drøfter og bedømmer hendelser og handlinger gjennom å sammenligne de med

kontrafaktiske scenarioer, skjer det på bakgrunn av ”utematisk fungerende” metafysiske antagelser om virkelighetens beskaffenhet, som for eksempel fri vilje, autonomi, determinisme, indeterminisme, kausalitet, *selvet* og *jeget* med mer. Den kontrafaktiske livsverdenen er dermed underliggende oppfatninger om virkelighetens modale struktur. Dette innebærer ikke nødvendigvis ”rene” metafysiske antagelser koblet vekk fra innholdet i livene våre, men antagelser og oppfatninger om hvem vi kan være eller kunne ha vært, og om hvordan verden kan eller kunne ha sett ut. Dette gjelder naturligvis ikke bare oss selv, men også de andre. Vi kan også kalle det for det antatte mulighetsrommet og handlingsrommet som vanligvis ikke er under debatt. I den kontrafaktiske livsverdenen er ofte overgangen mellom fakta og verdier flytende eller tilslørt. I språket vårt skiller vi ikke alltid mellom ulike betydninger av å kunne noe. At jeg for eksempel ”kan gå ut av døra” eller at ”jeg kan ikke gå ut av døra”, kan ha ulike betydninger avhengig av konteksten. Slike utsagn kan vise til det som er fysisk mulig med hensyn til naturlovene, for eksempel at det er fysisk mulig for meg å gå ut av døra. Men det kan også vise til at jeg får lov til å gå ut av døra, for eksempel hvis jeg er elev ved en skole der det er friminutt. En lærer kan for eksempel si til en elev at ”du kan ikke gå ut av døra”, med den betydningen at eleven ikke har lov til å gå ut av døra. Som regel er det uproblematisk at vi ikke skiller på ulike betydninger av *kan*. De fleste er innforstått med hva slags ”mulighetslandskap” vi snakker om; enten det gjelder der de fysiske naturlovene gjelder, eller om det handler om sosiale normer. Barn er ofte klar over og noen ganger raske til å påpeke at det er mange ting de ifølge naturlovene *kan* gjøre, selv om det ikke er lov eller sosialt akseptabelt. Et eksempel på dette er læreren min på barneskolen som irettesatte en gutt da han midt i timen sto oppe på pulten: ”Du kan ikke stå på pulten!”, ropte læreren til gutten, hvorpå eleven kvitterte med: ”Jo, jeg *kan* det” med tyngde på *kan*. Både lærer og elev var sannsynligvis innforståtte med hva det handlet om; at eleven ikke fikk lov til å stå på pulten. Eleven gjorde et poeng ut av forskjellen mellom hva som er mulig i henhold til naturlovene og hva som er mulig i forhold til sosiale normer, og påpekte dermed at lærerens påstand om at han ikke *kunne*, ikke stemte. I dette eksempel var det eleven som løftet opp til overflaten det ”utematisk fungerende”. Skolens håndtering av åpenbare brudd på skolens regler og sosiale normer, er heller ikke det som nødvendigvis er problematiske i denne sammenhengen. Alvorlighetsgraden stiger når det over tid skjer en form for disiplinering der eleven innvies i en måte å tenke på, hvor de grunnleggende forutsetningene for tankegangen ikke er synlige for eleven, og dermed heller ikke oppe til debatt. Det faktum at skolen har et mandat som innebærer at den besitter en makt over eleven, er vel noe de fleste

oppfatter uten at det debatteres; de fleste tar for gitt at det nødvendigvis er slik det må være. Men innholdet i makten og hvordan denne utøves av skolen (politikere, skoleledere, lærere etc.) er et større politisk tema, selv om elevene sjeldnere får ta del i denne debatten.

Flere av temaene i kapittel 5.2 knyttet til autonomibegrepet, er eksempler på ting som er ”utematisk fungerende” i skolesammenheng. Når skolen har definisjonsmakten over de kontrafaktiske idealene eleven skal måle seg opp mot, uten at elevene bevisstgjøres dette, kan det føre til en fordreid forståelse av autonomi, aktørskap og årsakssammenhenger hos elevene. Både selve oppfatningen om at alt kan og skal forklares, inkludert det egne aktørskapet, og det at skolen setter premissene for denne virkelighetsforståelsen uten at det er synlig for elevene, kan sies å være eksempler på det ”utematisk fungerende”.

Et annet eksempel for det ”utematisk fungerende”, er en form for ansvarlighet i forhold til noe elever ikke med viten og vilje gjør. En spesiell type handling som går under kategorien *unnnlatelser*, dvs. å unnlate å gjøre et eller annet, er det å *glemme*. I ordets rette forstand er dette en ikke-intensjonal handling, fordi det ikke er noe vi gjør med viten og vilje – det bare skjer oss. Vi kommer i ettertid på at det er noe vi skulle ha gjort, men at det var helt borte fra bevisstheten på det tidspunktet vi skulle ha gjort det. Jeg snakker nå bare om denne type av det å *glemme*, og ikke der det av en eller annen grunn benyttes bevisst som unnskyldning. At elever glemmer ulike ting som bøker, datamaskin, gymnastikktøy, å gjøre lekser e.l., er et velkjent tema fra skolen, og skolen gjør sitt for at elevene skal ta ansvar for sine forglemmelser slik at det ikke skjer ofte eller helst aldri. Men hvordan skal eleven ta ansvar for noe han eller hun ikke har gjort, men som bare har skjedd? I atferdsmodifiserende programmer, som for eksempel PALS¹⁹, kommer det å glemme ting som bøker, lekser og annet, under hovedkategorien problematisk atferd. Helt fra skolestart som 5-6 åringer, blir elevene gjort bevisste på ulike grader av negative konsekvenser av for eksempel det å *glemme* sammen med en rekke andre typer av problematisk atferd. Det er litt ulikt hva slags konsekvenser ulike PALS-skoler opererer med, men en fellesnevner må sies å være at konsekvensene naturlig nok innebærer noe negativt for eleven, og at det i hvert fall ikke berømmes å glemme noe. I og med at negative konsekvenser kan oppleves som en slags

¹⁹ PALS står for: Positiv atferd, støttende læringsmiljø og samhandling. Fra atferdssenterets hjemmeside: ”PALS er en norsk tilpasning av de skoleomfattende modellene for SW_PBS (School-Wide Positive Behavioral Intervention and Support) som bl.a. er utviklet ved Universitetet i Oregon. Modellen er videreutviklet og tilpasset norske forhold av Atferdssenteret” (Atferdssenteret).

straff, og straff er noe de fleste (og kanskje særlig barn) forbinder med at man har gjort noe galt, blir denne formen for ikke-intensjonal handling noe man er ansvarlig for. Meningen er at elevene skal lære seg å gjøre de grepene som skal til for å huske ulike ting, og det *ikke å glemme* føyer seg da inn i rekken av de mange forventningene skolen stiller til elevene. I tillegg kan for eksempel elever glemme at de skal rekke opp hånden før de snakker. Dette er ikke en argumentasjon for at skolen ikke skal gå fram slik at elever lærer seg å ha med det nødvendige på skolen, eller følge skolens regler etc., men poenget med eksempelet er å trekke frem muligheten for at det ”utematiserte fungerende” som elevene innvies i, er en form for ansvarlighet i forhold til noe man ikke med viten og vilje har gjort. Barn i 5-6 års alderen er avhengige av at foreldrene hjelper til så de har med seg det de skal og kommer til skolen i tide etc. Det betyr at for denne aldersgruppen er det i stor grad foreldrene som har det reelle ansvaret for at elevene skal kunne oppfylle mange av de forventningene og kravene som stilles med påfølgende sanksjoner. Men det er elevene som bærer den fremste konsekvensen av ikke å oppfylle de, og det viktigste; det er elevene som innvies i en kontrafaktisk livsverden der ideen er at de er opphavet til hendelser de ikke har kontroll over, og at de burde og dermed skulle kunne ha vært annerledes. Det er verdt å legge merke til at retningen ofte går fra et normativt krav til en empirisk mulighet; fra bør til kan, eller fra burde til skulle kunne. I denne logikken gjøres elevene ansvarlige for ikke å ha gjort det som påstås å være mulig for de. I og med dette har de blitt tildelt en ”frihet” og en ansvarlighet de i utgangspunktet ikke hadde. I denne sammenhengen er det nok riktigere å snakke om indoktrinering enn innvielse, og jeg bør heller omtale det som at elevene disiplineres inn i én bestemt type kontrafaktisk livsverden”, i stedet for at de innvies i en, jf. innvielsesbegrepet til Peters (1966).

7.2 Instrumentelle og utopiske kontrafaktisiteter

I kapittel 2.2.4 viste jeg til sosiologen Roland Paulsen (2014), som med utgangspunkt i de to motstridende perspektivene hos filosofene Vaihinger (1968) og Bloch (1986), deler kontrafaktisiteter inn i to grupper: *Utopiske* og *instrumentelle*. Paulsen argumenterer for at kontrafaktiske forestillinger er nødvendige for å overskride makt. Instrumentelle kontrafaktisiteter er løsninger som gjør at maktstrukturer forblir uendrede, mens utopiske kontrafaktisiteter er løsninger som utfordrer dominerende maktstrukturer (Paulsen, 2014, s.

160). I fortsettelsen vil jeg benytte meg av Paulsens argumentasjon, men jeg utvider perspektivet til å inkludere at maktstrukturer også kan utfordres ”ovenfra og ned”.

I oppgaven har jeg vist til en rekke eksempler på kontrafaktisiteter i pedagogisk tenkning og praksis, og innen utdanningspolitikken. Jeg vil argumentere for at mange av disse er av den utopiske typen som Paulsen henviser til. Med utopiske mener jeg idéer om idealtilstander i samfunnet som ikke er realiserbare, for eksempel frihet og likhet. (Det betyr ikke at de ikke er verdt å strekke seg mot eller kjempe for.) De er ikke av den utopiske typen som innebærer en ”nedenfra og opp”-utfordring til dominerende strukturer slik Paulsen snakker om. I mitt perspektiv kan dominerende strukturer også utfordres ”ovenfra og ned”. Uavhengig av hva som ligger bak en viss utdanningspolitikk, det være seg økonomiske insentiver eller sosiale behov, så er målsetningen å oppnå en forandring eller forbedring. Elevenes og lærernes måte å agere på kan for eksempel fra myndighetenes side oppfattes som en dominerende struktur. Liberale krefter ønsker gjerne å utfordre statens innflytelse og makt over skolen og vise versa. Med dette ønsker jeg ikke så mye å poengtere graden av utopi i målene til de som styrer skolen, men snarere hvor lite rom det gis for at elevene får fungere slik at de blir sensitive til og bevisste sine egne utopiske kontrafaktisiteter, som igjen kan utfordre de dominerende maktstrukturene. Mens de som styrer skolen nærmest har som mandat å styre skolen etter kontrafaktiske utopier, så får elevene ingen øvelse eller erfaring i å tenke seg eller forestille seg hvordan ting kunne ha sett annerledes ut rundt dem. Det elevene i hovedsak øver seg i, og den *kontrafaktiske livsverdenen* de innvies i, synes å ha ensidig fokus på den kontrafaktiske versjonen av dem selv: Hvordan de selv kunne og burde ha vært annerledes, og hvem de bør bli. Således er det riktigere å si at elevene disiplineres inn i en kontrafaktisk livsverden, enn at de innvies i en, jf. innvielsesbegrepet til Peters (1966). Allerede fra tidlig alder har barn kontrafaktiske tanker og følelser, og insentiver til å skape forandring i omgivelsene. Men min påstand er at dagens skolesystem disiplinerer og tvinger elevene til å tenke på en måte som først og fremst handler om instrumentelle kontrafaktisiteter; hvordan de skal og skulle ha handlet annerledes ifølge eksisterende standarder. Ifølge Paulsens terminologi er instrumentelle kontrafaktisiteter løsninger som gjør at maktstrukturer forblir uendrede. Så langt jeg kan se er det få eller ingen mål som har til hensikt å bidra til at elevene blir bevisste sitt eget abstrakte forhold til institusjonen som organiserer livet deres. Et unntak er kanskje det ”possibilistiske” frihetsidealet jeg omtale i kapittel 5.1.1: ”Du har alle muligheter!”. Men denne overfladiske varianten av det moderne frihetsidealet, synes for det meste å handle om ulike måter å kontrollere seg selv. De alternativene elevene skal sammenligne virkeligheten

med, synes å lede til et ensidig fokus på observasjon, evaluering, selvobservasjon og selvkontroll. Oppfordringen ser ut til å være noe lignende dette: ”Du har alle muligheter til å kontrollere deg selv; til kontinuerlig å observere dine kroppslige uttrykk og være sikker på at du gjør det som forventes å skje”. Samtidig med skiftet i den temporale dimensjonen, forvandles denne idéen og formaningen i indikativ form, før eller senere til elevens betingelser i konjunktiv form: Det som *kan* og *bør* skje, erstattes med det som *kunne ha* og *burde ha* skjedd. I konjunktiv form skulle elevene ha kunnet utøve denne type kontroll over seg selv: De skulle ha kunnet gjøre det som var forventet å skje. Men det spillet er alle dømt til å tape, fordi vi aldri kan gjøre det som er forventet å skje. Jeg viser til kapittel 6.4 hvor jeg drøfter skillet mellom hendelser og handlinger; mellom det som *skjer* oss og det vi *gjør*. I Anscombe sin forståelse av intensjonalitet gir det ikke mening å snakke om at man gjør det som skjer:

I myself formerly, in considering these problems, came out with the formula: I *do* what *happens*. That is to say, when the description of what happens is the very thing which I should say I was doing, then there is no distinction between my doing and the thing's happening. But everyone who heard this formula found it extremely paradoxical and obscure. And I think the reason is this: What happens must be given by observation; but I have argued that my knowledge of what I do is not by observation (Anscombe, 1975, s. 52-53, §29).

Det som *skjer* er ifølge Anscombe gitt gjennom observasjon. Man behøver ikke nødvendigvis å dele synet til Anscombe på hva det vil si å handle intensjonalt. Men skillet mellom det vi kan slutte oss til eller observere når vi gjør noe, og det vi svarer på ”hvorfor”-spørsmålet, framstår som vesentlig. Dette er særlig med tanke på den sterke vektleggingen av vurderinger og observasjoner i dagens skole. Det kan være en fare for at svarene på ”hvorfor”-spørsmålet etter hvert tar formen av ferdig uttenkte normative grunner tilpasset skolens forventninger.

7.3 Kontrafaktiske følelser og sosial kontroll

Et tema jeg i liten grad har berørt direkte i oppgaven, er de ”kontrafaktiske følelsene”. På tross av at de har fått liten oppmerksomhet i undersøkelsen, utgjør de et sentralt element i kontrafaktisk tenkning, og er i mange tilfeller tett sammenvevd med den. Et fokus i oppgaven har vært å undersøke mulige konsekvenser av den kontrafaktiske tenkningen, som dagens skoleelever uunngåelig engasjerer seg i. Det er et faktum at mange unge i dag sliter med ulike

former for psykiske problemer. Begreper som depresjon, utmattelse og utbrenthet har blitt en del av de unges vokabular. Landets barneavdelinger ved sykehusene rapporterer om en radikal økning i alvorlige stressrelaterte tilstander hos barn; lidelser som antas å være utløst av vedvarende stress og prestasjonskrav. Når det gjelder ”kontrafaktiske følelser” er det spesielt én type jeg vil framheve som en mulig konsekvens av dagens skolesystem. Jeg har en hypotese om at skolens kontrollordninger og vurderingssystemer innebærer en sterk sosial kontroll, som representerer den mørke siden av det nesten euforiske possibilistiske frihetsidealet. Jeg tenker på den kanskje sterkeste kontrafaktiske følelsen, nemlig følelsen av skam. Selvet begrepet skam anvendes ikke så mye i dagligspråket, kanskje bortsett fra etter at den umåtelig populære norske ungdomsserien SKAM startet i 2015. I tilknytning til skolen handler det om skammen over ikke å lykkes i henhold til alle kravene, men kanskje enda verre; skammen over ikke å ha evnet å bruke friheten sin og utnyttet alle mulighetene. Jon Elster har påpekt at forventning om skam fungerer som en kraftfull regulator av atferd (Elster, 1999, s. 154). Skam innebærer at både selvfølelsen og aktelsen er rokket ved, og vi gjør det meste for å unngå skamfølelser. Skam er svært negativt definert i vår kultur; den fornektes og får dermed gå upåaktet hen. Med den tilsynelatende skamfrie kulturen vi lever i, er det kanskje slik at noen i god avstand fra klasserommene skamløst setter målene, mens det er de unge som eventuelt må bære skammen over ikke å leve opptil de. Hvis det er noe i denne hypotesen, er det allikevel ingen påstand fra min side om at en slik form for sosial kontroll skjer på bakgrunn av at noen *gjør* det med viten og vilje.

I det foregående har jeg foreslått å utvide den klassiske definisjonen av utdanning som ”innvielse i en rasjonell livsverden”, til å implisere utdanning som ”innvielse i en *kontrafaktisk livsverden*”. Hovedargumentet for dette er at den kontrafaktiske tenkningen utgjør en grunnleggende bestanddel ved rasjonaliteten vår, med de implikasjonene det måtte ha. Deretter lånte jeg begrepsparet *utopiske* og *instrumentelle* kontrafaktisiteter fra Paulsen (2014) for å belyse maktperspektiver ved kontrafaktisk tenkning. Konklusjonen min er at skolen har et tilnærmet formelt monopol på utopiske kontrafaktisiteter. Dagens skolesystem gir lite rom til at elevene blir sensitive til og bevisste sine egne utopiske kontrafaktisiteter, slik at også de kan utfordre de dominerende maktstrukturene. Den kontrafaktisk tenkningen elevene uunngåelig må engasjere seg i, karakteriseres av instrumentelle kontrafaktisiteter. I lys av teorien til Anscombe (1963) om intensjonal handling, stilles elevene overfor det

umulige kravet å skulle gjøre det som er forventet å skje. De vurderes etter hvor godt de klarer å bli et objekt for seg selv, helt på skolens premisser.

Til slutt foreslo jeg en hypotese om at skolens kontrollordninger og vurderingssystemer innebærer en sterk sosial kontroll, i form av å fungere skamproduserende. Tesen går ut på at det moderne possibilistiske frihetsidealet sammen med en sterkt kontrollerende skole, har en bakside og en omkostning som elevene bærer mer eller mindre i stillhet.

8 Avslutning

I denne oppgaven har jeg undersøkt hva som grunnleggende er involvert i kontrafaktisk tenkning, og hva vi antar om virkelighetens beskaffenhet når vi tenker at den kunne ha vært annerledes. Med spørsmålet ”kunne jeg ha handlet annerledes?” som inngang til tematikken, har jeg gjort en teoretisk analyse av kontrafaktisk tenkning. Oppgaven har i hovedsak hatt to innramminger. Den ene har vært de metafysiske rammene, og den andre har vært den pedagogiske konteksten. Spillerommet har dermed vært stort og det har vært mange stemmer inne i bildet. Den første konklusjonen er at den hverdagslige kontrafaktiske tenkningen er umåtelig kompleks—hvis vi undersøker den nærmere. Det er bare å stille noen enkle spørsmål, og så forgrener de seg ut i en rekke epistemologiske og ontologiske temaer. I hverdagen håndterer vi uten anstrengelse alle de mulige og umulige fortidene – og framtidene. Sjelden eller aldri spør vi hverandre hvilke metafysiske teser vi legger til grunn, når vi mener at noe burde ha vært annerledes enn det er. Så hvorfor undersøke det som tilsynelatende fungerer utmerket uten vår innblanding? Jeg mener naturlig nok, siden jeg har skrevet denne oppgaven, at vi har mye å hente på å være mer bevisste hva som skjer i når vi tenker kontrafaktisk. Er vi ikke bevisste på det, er det kanskje tjenlig for andre, men ikke nødvendigvis for oss selv og det vi bryr oss om.

En kontrafaktisk påstand er at uten å tenke kontrafaktisk, ville vi ikke kunne trekke kausale slutninger om selv de enkleste ting, og vi ville nok heller ikke oppfatte oss selv som frie eller bundne. Men slik er det ikke, tvert imot synes vi å leve i alternativenes univers. Betydningen av *kunne* er sjelden nøytral; det ligger for det meste et *burde* bak. Ikke nødvendigvis bare som et strengt normativt krav, men i betydningen av at det er en målestokk bak det meste av det vi foretar oss. Finnes det alternative måter noe kan være på, så finnes det som regel bedre og verre måter det kan være på. Når vi tenker kontrafaktisk om fortiden, så handler det ikke om mindre enn forestillingen om fraværet av virkelighetsbiter som hadde kunnet endre livsløpet vårt og ledet oss til en annen posisjon enn den vi inntar i dag. Når vi spør om vi kunne ha handlet annerledes, henvender vi oss innover i oss selv, inn i vårt eget sinn, og forsøker å finne ut av universets ytterste grenser i form av hva som er mulig, umulig og nødvendig. Vi gjør det sammen med andre og vi gjøre det alene, og vi gjør det i det uendelige.

Ettersom vi antar metafysiske antagelser om virkeligheten når vi tenker kontrafaktisk, undersøker jeg problemstillingen i lys av de metafysiske rammene *determinisme* og

indeterminisme, og herunder de moralfilosofiske begrepene *fri vilje* og *ansvarlighet*. Den tematiske analysen har samtidig fått utspille seg i konteksten av en påstand jeg presenterer i oppgavens begynnelse. Her hevder jeg at skolens mål i sin natur er *pre-kontrafaktiske*, fordi de handler om å realisere mål og verdier jeg vil kalle ideelle tilstander og størrelser. Jeg hevder samtidig at utdanningsfeltet og pedagogikkfaget stiller i særklasse sammenlignet med andre institusjonelle områder og fagdisipliner, i det å være et *pre-kontrafaktisk* prosjekt som *normativ sosial praksis*.

Jeg foreslå å utvide Peters (1966) klassiske danningsbegrep om *innvielse*. ”Utdanning som innvielse i en rasjonell livsverden” impliserer dermed ”utdanning som innvielse i en kontrafaktisk livsverden”.

I undersøkelsen har jeg vektlagt de metafysiske aspektene ved kontrafaktisk tenkning, og i mindre grad vektlagt de funksjonelle og emosjonelle sidene. Den kontrafaktiske tenkningens betydning i lærings- og beslutningsprosesser har jeg i liten grad berørt. Det har det blitt forsket mye på innen psykologien.

Et tilbakevendende tema har vært spørsmålet om hva eller hvem som bestemmer hvilke deler av virkeligheten vi kan og bør bytte ut? Hvilken rolle spiller skolesamfunnet og hva slags kontrafaktisk tenkning engasjerer dagens skoleelever seg i? Hva slags tenkelige implikasjoner har det for de unges utvikling, læring og identitetskonstruksjon stadig å måtte sammenligne seg med kontrafaktiske motparter som noen andre har definert? Både frihetsbegrepet og autonomibegrepet har blitt viet mye plass. Jeg viser til et moderne frihetsbegrep, det ”possibilistiske”, som må sies å være et tveegget sverd. Den indikative betingelsen av begrepet kan uttrykkes slik: ”Hvis du (bare) vil, så har du alle muligheter”. I hverdagsspråket kan dette enkelt oversettes til den konjunktive betingelsen og dermed den kontrafaktiske påstanden: ”hvis du (bare) hadde villet, hadde du hatt alle muligheter”. Da skyves alt ansvaret over på individet. I tilknytning til kontrafaktiske følelser foreslår jeg en hypotese om at skolens kontroll- og vurderingssystemer innebærer en sterk sosial kontroll, som representerer den mørke siden av det possibilistiske frihetsidealet. Dette er følelsen av skam.

Jeg har undersøkt hva slags oppfatning elevene kan tenkes å utvikle om for eksempel aktørskap, når de i årevis sammenligner seg selv med pre-kontrafaktiske idealer. Dette gjelder særlig når den bakenforliggende idéen er at utfallet av handlingene deres utelukkende er resultatet av deres egne bevisste valg. Her har jeg kritisert at mengden valg som elevene

forventes å kunne begrunne er så stor, at den inngir et bilde av at alt kan og skal forklares, også elevenes egne beveggrunner. I dagens målstyrte skole synes det å være lite rom for annet enn drilling i introspeksjon og retrospeksjon hos elevene. Jeg benyttet meg av Paulsens (2014) kategorisering av kontrafaktisiterer knyttet til det å utfordre maktstrukturer. Vi synes å tenke kontrafaktisk om fortiden for å utfordre de begrensningene den tilsynelatende legger på oss, i form av at den ikke kan gjøres om. På samme måte kreves det kontrafaktisk tenkning i form av utopiske kontrafaktisiteter, for å utfordre det vi mener er urettmessige begrensninger i for eksempel handlingsfriheten vår. Men en slik tenkning krever at man ser utover seg selv for å forstå de kausale sammenhengene man inngår i.

Med denne oppgaven ville jeg også undersøke om “stien” via det *kontrafaktiske landskapet* kunne representere en fruktbar åpning inn det pedagogisk-filosofiske feltet. Min konklusjon er på den positive siden. Via det kontrafaktiske er veien kort til mange av pedagogikkens grunnlagsproblem som autonomi, identitetskonstruksjon, indoktrinering og handlingslegitimering for å nevne noen.

Litteraturliste

- Anscombe, G.E.M. (1963). *Intention* (2nd ed. utg.). Oxford: Blackwell.
- Anscombe, G.E.M. (1975). *Causality and Determination i E. Sosa, ed., Causation and Conditionals*. London: Oxford University Press.
- Anscombe, G.E.M. (1981). *The collected philosophical papers of G.E.M. Anscombe : Vol. 2 : Metaphysics and the philosophy of mind*. Oxford: Blackwell.
- Aristoteles. (1935). *Aristotle : in twenty-three volumes : 18 : Metaphysics : Books X-XIV* (Hugh Tredennick & G. Cyril Armstrong, Overs. Vol. 287). London: Heinemann.
- Aristotle. (1926). *Nicomachean ethics* (H. Rackham, Overs.): Harvard University Press.
- Atferdssenteret. (2016, 17.06). Positiv atferd, støttende læringsmiljø og samhandling (PALS). Lastet ned fra <http://www.atferdssenteret.no/informasjon-om-pals-modellen/>
- Bentham, J., & Ogden, C.K. (1932). *Bentham's Theory of fictions*. London: Kegan Paul, Trench, Trubner & Co.
- Berlin, I. (1969). *Four essays on liberty*. Oxford: Oxford University Press.
- Bloch, E. (1986). *The principle of hope*: Blackwell.
- Bokmålsordboka. (2016). Indikativ.
- Byrne, R.M.J. (2005). *The rational imagination: how people create alternatives to reality*. Cambridge, Mass.: MIT Press.
- Bysted Möller, R. (2011). Moral og Frihed - Skitse til en ny teori om viljens frihed. *Res cogitans: Journal of philosophy*, 2(8), 115-166.
- Bø, I., & Helle, L. (2008). *Pedagogisk ordbok : praktisk oppslagsverk i pedagogikk, psykologi og sosiologi* (2. utg. utg.). Oslo: Universitetsforl.
- Carter, I. (2016). Positive and Negative Liberty. Retrieved 2016, 25.04, fra The Stanford Encyclopedia of Philosophy <http://plato.stanford.edu/entries/liberty-positive-negative/> - Aca
- Coeckelbergh, M. (2004). *The Metaphysics of Autonomy: The Reconciliation of Ancient and Modern Ideals of the Person*: Palgrave Macmillan UK.
- Crane, T. (2003). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*: Routledge.
- Cuypers, E.S. (2009). *Educating for authenticity: The paradox of moral education revisited. I H. Siegel (Red.), The Oxford handbook of philosophy of education*. Oxford: Oxford University Press.
- Dale, E.L. (1992). *Pedagogisk filosofi*. Oslo: Ad Notam Gyldendal.

- Dancy, J. (2002). *Practical reality*. New York: Oxford University Press.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23), 685-700.
- Davidson, D. (2001). *Essays on actions and events* (2nd ed. utg.). Oxford: Clarendon Press.
- de Champs, E. (2011). The Place of Jeremy Bentham's Theory of Fictions in Eighteenth-Century Linguistic Thought. 2011, 2.
- Demsetz, H. (1969). Information and Efficiency: Another Viewpoint. *The Journal of Law and Economics*, 12(1), 1-22. doi: 10.1086/466657
- Earman, J. (1986). *A primer on determinism*. Dordrecht: Reidel.
- Elster, J. (1978). *Logic and society : contradictions and possible worlds*. Chichester: Wiley.
- Elster, J. (1999). *Alchemies of the mind : rationality and the emotions*. Cambridge: Cambridge University Press.
- Epstude, K., & Roese, N.J. (2008). The Functional Theory of Counterfactual Thinking. *Personality and Social Psychology Review*, 12(2), 168-192. doi: 10.1177/1088868308316091
- Epstude, K., Scholl, A., & Roese, N.J. (2016). Prefactual Thoughts: Mental Simulations About What Might Happen. *Review of General Psychology*. doi: 10.1037/gpr0000064
- Feinberg, J. (2010). The idea of a free man. I F. Doyle James (Red.), *Educational Judgments(International Library of the Philosophy of Education Volume 9): Papers in the Philosophy of Education* (s. 104-124): Taylor & Francis.
- Fischhoff, B. (1974). Hindsight \neq Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty. *Research Bulletin. Oregon Research Institute*, 14(13), 1-37.
- Florovsky, G. (1969). The study of the past. I Ronald H. Nash (Red.), *Ideas of history : Vol 2 : The Critical philosophy of history* (s. 351-369). New York: E.P. Dutton.
- Ford, A. (2011). Action and generality. I Anton Ford, Jennifer Hornsby & Frederick Stoutland (Red.), *Essays on Anscombe's Intention*: Harvard University Press.
- Frankfurt, H.G. (1969). Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy*, 66(23), 829-839. doi: 10.2307/2023833
- Frankfurt, H.G. (1978). The Problem of Action. *American Philosophical Quarterly*, 15(2), 157-162. doi: 10.2307/20009708
- Frønes, I. (2015). Storverk om barn. *Apollon: tidsskrift fra Universitetet i Oslo*, 4, 40-44.
- Goodman, N. (1973). *Fact, fiction, and forecast*. Indianapolis: Bobbs-Merrill.
- Habermas, J. (1987). *The theory of communicative action : 2 : Lifeworld and system : a critique of functionalist reason* (Vol. 2). Boston, Mass: Beacon Press.

Hansen, M.K. (2016). Anomal Monisme. Retrieved 14. mars 2016, fra Store norske leksikon https://snl.no/anomal_monisme

Hanson, N.R. (1958). *Patterns of discovery : an inquiry into the conceptual foundations of science*. Cambridge: University Press.

Hellesnes, J. (1988). *Hermeneutikk og kultur : filosofiske stubbar*. Oslo: Samlaget.

Hesslow, G. (1981). Causality and Determinism. *Philosophy of Science*, 48(4), 591-605. doi: 10.2307/186838

Hoch, S.J., & Roediger, H.L. (1985). Counterfactual Reasoning and Accuracy in Predicting Personal Events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 719-731. doi: 10.1037/0278-7393.11.1-4.719

Hofstadter, D.R. (1979). *Gödel, Escher, Bach : an eternal golden braid* (Vol. 12). Hassocks: Harvester Press.

Holmen, H.A. (2011). *Ethics and the nature of action*. Department of Philosophy, Classics, History of Arts and Ideas, Faculty of Humanitits, University of Oslo, Oslo.

Hornsby, J. (2013). Basic Activity. *Aristotelian Society Supplementary Volume*, 87(1), 1-18. doi: 10.1111/j.1467-8349.2013.00217.x

Hume, D. (2004). *An Enquiry Concerning Human Understanding*: Dover Publications.

Hustvedt, S. (2012). *Den skjelvende kvinnen, eller Historien om nervene mine*. Oslo: Aschehoug.

Imsen, G. (2016, 09.03). Den nasjonale målfesten, *Klassekampen*, s. 23.

Inwagen, P.v. (1975). The incompatibility of free will and determinism. *Philosophical Studies*, 27(3), 185-199. doi: 10.1007/BF01624156

Inwagen, P.v. (2002). *Metaphysics* (2nd ed. utg.). Cambridge, Mass: Westview Press.

James, W. (1912). *The will to believe: and other essays in popular philosophy*: New York etc.: Longmans, Green, and co.

Kahneman, D., Tversky, A., & Slovic, P. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.

Kannisto, T. (2012). *From thinking to being: Kant's modal critique of metaphysics*. [Oslo]: University of Oslo, Faculty of Humanities.

Kant, I. (2002). *Critique of Practical Reason* (W.S. Pluhar, Overs.): Hackett Publishing Company.

Kant, I. (2016). *Om pedagogikk*. Oslo: Aschehoug i samarbeid med Thorleif Dahls kulturbibliotek.

- Kirk, G.S., Raven, J.E., & Schofield, M. (1983). *The presocratic philosophers : a critical history with a selection of texts* (2nd ed. utg.). Cambridge: Cambridge University Press.
- Kofta, M., Weary, G., & Sedek, G. (1998). *Personal Control in Action: Cognitive and Motivational Mechanisms*: Springer US.
- Korsgaard, C.M. (1996). *The sources of normativity*. Cambridge: Cambridge University Press.
- Kvernbekk, T. (2011). *Humaniorastudier i pedagogikk: pedagogisk filosofi og historie*. Oslo: Abstrakt forl.
- Kvernbekk, T. (2016). *Evidence-based practice in education : functions of evidence and causal presuppositions* (Vol. 147). London: Routledge.
- Lebow, R.N. (2000). What's So Different about a Counterfactual? *World Politics*, 52(4), 550-585. doi: 10.2307/25054129
- Lewis, D. (1973). *Counterfactuals*. Cambridge, Mass: Harvard University Press.
- Generell del av læreplanen (1993).
- LK06. (2005). *Kunnskapsløftet: læreplaner for gjennomgående fag i grunnskolen og videregående opplæring : læreplaner for grunnskolen*. Oslo: Utdanningsdirektoratet.
- Løvlie, L. (2003). Den pedagogiske filosofiens veier. *Utbildning och demokrati : tidskrift för didaktik och utbildningspolitik*, 12(1), 81-96.
- Løvlie, L. (2008). Kampen om individet. *Bedre skole*.
- Løvlie, L. (2011). Dannelse og profesjonell tenkning. I Gorana Ognjenovic & Bernt Hagtvat (Red.), *Dannelse : tenkning, modning, refleksjon : nordiske perspektiver på allmenndannelsens nødvendighet i høyere utdanning og forskning*. Oslo: Dreyers forl.
- Markman, K., & Miller, A. (2006). Depression, control, and counterfactual thinking: Functional for whom? *Journal of Social and Clinical Psychology*, 25(2), 210-227.
- Melbye, H., & Harboe, L. (2009). Kriterier for karaktersetting ved muntlige presentasjoner. Retrieved 2016, 09.09 <https://docs.google.com/document/d/1h5-cEZdm5zw4IxoohcUYXD3PwuJQzaVMu9LpJwK-vps/edit>
- Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge.
- Moore, M.S. (2010). *Act and Crime: The Philosophy of Action and Its Implications for Criminal Law*: Oxford University Press.
- Nancy, J.L. (1993). *The Experience of Freedom*: Stanford University Press.
- Nygård, M.Y., & Samdal, M. (2013, 08.05). Skolen mobbet sønnen min, *Adressa.no*. Lastet ned fra <http://www.adressa.no/nyheter/trondheim/article7641091.ece>
- Næss, A. (1963). *Logikk og metodelære: en innføring* (Vol. nr 14). Oslo: Klubben.

- Parfit, D. (2011). *On What Matters, 1*. Oxford: OUP Oxford.
- Paulsen, R. (2014). *The Counterfactual Imagination*: Stanford University Press.
- Peters, R.S. (1966). *Ethics and education*. London: George Allen & Unwin.
- Pitkin, H.F. (1972). *Wittgenstein and justice: on the significance of Ludwig Wittgenstein for social and political thought*. Berkeley, Calif.: University of California Press.
- Ramberg, B. (2004). Davidson og sjelsvitenskapene. *Norsk Filosofisk tidsskrift*, 39 ER(01-02).
- Ryle, G. (1973). Negative 'Actions'. *Hermathena*(115), 81-93.
- Sagdahl, M. (2014a). Autonomi. Retrieved 2016, 25.04, fra Store norske leksikon <https://snl.no/autonomi>
- Sagdahl, M. (2014b). Grunn Store norske leksikon.
- Sartre, J.P. (1992). *Being and Nothingness* (H.E. Barnes, Overs.): Washington Square Press.
- Schindler, D.C. (2012). *The Perfection of Freedom: Schiller, Schelling, and Hegel between the Ancients and the Moderns*: Cascade Books.
- Sellars, W. (1997). *Empiricism and the philosophy of mind*. Cambridge, Mass: Harvard University Press.
- Setiya, K. (2007). *Reasons without rationalism*. Princeton, N.J: Princeton University Press.
- Siegel, H. (1997). *Rationality Redeemed?: Further Dialogues on an Educational Ideal*: Routledge.
- Siegel, H. (2012). Education as initiation into the space of reasons. *Theory and Research in Education*, 10(2), 191-202. doi: 10.1177/1477878512446542
- Sloman, S. (2005). *Causal Models: How People Think about the World and Its Alternatives*. New York: Oxford University Press.
- SNL. (2016). Konjunktiv Store norske leksikon.
- Stalnaker, R. (1999). *Context and content: essays on intentionality in speech and thought*. Oxford: Oxford University Press.
- Stout, R. (2005). *Action : Rowland Stout*. Chesham: Acumen.
- Strassfeld, R.N. (1992). If . . . : Counterfactuals in the Law. *Faculty Publications, Paper 373*.
- Stray, J.H. (2010). Demokratisk medborgerskap i norsk skole? : en kritisk analyse.
- Suppes, P. (1993). The Transcendental Character of Determinism. *Midwest Studies In Philosophy*, 18(1), 242-257. doi: 10.1111/j.1475-4975.1993.tb00266.x

- Svartdal, F., & Teigen, K.H. (Forfattere). (2014). Bias i psykologi, *Store norske leksikon*.
- Svendsen, L.F.H. (2013). *Frihetens filosofi*. Oslo: Universitetsforl.
- Svetlova, E. (2009). "Do I See What the Market Does Not See?": Counterfactual Thinking in Financial Markets. *Historical Social Research / Historische Sozialforschung*, 34(2 (128)), 147-157.
- Taylor, R. (1992). *Metaphysics*. Englewood Cliffs, N.J.: Prentice Hall.
- Tranøy, K.E. (2011). Kontingent: filosofi *Store norske leksikon*. Hentet fra: <https://snl.no/kontingent%2Ffilosofi>.
- Tufte, P.A. (2013). Forståelser av kausalbegrepet i samfunnsvitenskapene. *Tidsskrift for samfunnsforskning*, 54(03), 341-354.
- Utdanningsforbundet. (2016). Frafall i videregående. Retrieved 2016, 23.10 <https://www.utdanningsforbundet.no/Hovedmeny/Vi-mener/Frafall-i-videregaende/>
- Vaihinger, H. (1968). *The philosophy of "As if"*. London: Routledge & Kegan Paul.
- Wilson, G., & Shpall, S. (2012). Action. *The stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/sum2012/entries/action/>
- Wittgenstein, L. (1997). *Philosophische Untersuchungen = Philosophical investigations* (2. Aufl., [Neudruck]. utg.). Oxford: Blackwell.
- Wittgenstein, L. (2001). *Tractatus logico-philosophicus* (David Pears & Brian McGuinness, Overs.). London; New York: Routledge.