

“The heart of effective teaching”?

Kan en rettferdig elevvurdering oppnås?

Synnøve Bjørn



Masteroppgave

Pedagogisk forskningsinstitutt

Utdanningsvitenskaplig fakultet

UNIVERSITETET I OSLO

1. Desember 2009

SAMMENDRAG AV MASTEROPPGAVEN I PEDAGOGIKK**TITTEL:**

“The heart of effective teaching”?

Kan en rettferdig elevvurdering oppnås?

AV:

Bjørn, Synnøve

EKSAMEN:

Masteroppgave i pedagogikk, allmenn
studieretning
Ped 5300

SEMESTER:

Høsten 2009

STIKKORD:

Elevvurdering

Formativ

Summativ

Rettferdighet

Didaktikk

Pedagogisk filosofi

Egalitarisme

Differanseprinsippet

Desert-prinsippet

Sammendrag:

Oppgaven du nå holder foran deg diskuterer hva det vil si å gi en rettferdig elevvurdering og om det i det hele tatt kan oppnås. Dette blir belyst igjennom teori på vurdering og filosofiske perspektiver på rettferdighet. Dette skjer ved en gjennomgang av begrepene evaluering, elevvurdering, rettferdighet og en drøfting av temaene i forhold til hverandre.

Problemstilling:

Problemstillingen min er ” Kan en rettferdig elevvurdering oppnås?”

Metode og kilde:

Denne masteroppgaven er skrevet i spenningsfeltet mellom pedagogisk filosofi og pedagogisk teori. Oppgaven er teoretisk og tar ikke i bruk empiri for å belyse problemstillingen. Metoden som er blitt benyttet er hermeneutisk og bygger på analyser og tekstfortolkning. Analysen skjer ved å diskutere enkeltbegrepene elevvurdering og rettferdighet. Hva ligger i disse begrepene? Deretter drøftes begrepene i forhold til hverandre og i forhold til problemstillingen *kan en rettferdig elevvurdering oppnås?* Kildene som anvendes er varierte, men sentrale kilder er Scriven (1991), Black m. fler (1998, 2001, 2003, 2004), Lamont og Favor (2007) og House (1977,1980, 1993). Øvrige kilder supplerer disse der det er nødvendig. En utfordring har vært å knytte elevvurdering og rettferdighet sammen, da det finnes svært begrenset litteratur som behandler dette temaet.

Resultat/hovedfunn:

For å belyse hva som er en rettferdig elevvurdering, skisserer oppgaven først ut evalueringsfeltet. Begrepet evaluering blir analysert ved å drøfte ulike definisjoner og vise til begrepsmessige distinksjoner som skaper spenningsforhold i evalueringsfeltet. Etter å ha satt ett rammeverk på oppgaven, blir den snevret inn mot elevvurdering. Elevvurderingsfeltet anses som snevrere enn evalueringsfeltet og tar sikte på å si noe om hvilke forhold som forekommer i en elevvurdering. Sentrale begreper her er formativ og summativ vurdering. Videre blir rettferdighetsbegrepet analysert. Det blir lagt fram ulike tilnærminger til hva et slikt innebærer. Noen av disse tilnærmingene er egalitarismen, differanseprinsippet og fortjenesteprinsippet. Videre drøftes elevvurdering og rettferdighet igjennom metode, validitet og konkrete former for elevvurdering, som karaktergivning og tilbakemeldinger. Oppgaven viser at en rettferdig elevvurdering avhenger av to forhold. På den ene siden handler en

rettferdig elevvurdering om vurderingskompetanse og fagkompetanse hos lærer. På den andre siden handler en rettferdig elevvurdering om hvilke premisser for rettferdighet som ligger bak vurderingen. Mens elevvurdering er avhengig av lærers kompetanse både innenfor vurderingsmetode og innenfor fagfeltet, sprer rettferdighetsbegrepet seg utover en rekke ulike prinsipper. Hver for seg utgjør disse prinsippene ulike resultater for hva som kan anses for å være en rettferdig elevvurdering.

Forord

Det er mange årsaker til at jeg valgte å skrive om en rettferdig elevvurdering. Min interesse for elevvurdering har vært med meg opp igjennom grunnskolen, ut videregående og inn på universitetet. Våren 2007 skrev jeg bachelorgraden min om elevvurdering. På mastergraden fikk jeg også øynene opp for utdanningsfilosofi og bestemte meg for å kombinere elevvurdering med filosofien. Det var med stor glede jeg oppdaget Utdanningsdirektoratets mål om en mer rettferdig elevvurdering. Jeg begynte å gruble over hva som lå bak rettferdighetsbegrepet og hvordan ulike forståelser av hva som er rettferdig kunne påvirke elevvurderingen. Oppgaven du nå har forran deg er resultatet av all grublingen min.

En riktig stor takk rettes til min veileder, Professor Tone Kverbekk, for å alltid ha døren åpen og å ha inkludert meg i et utdanningsfilosofisk miljø. Takk for at du har hjulpet meg finne veien igjennom denne prosessen og for dine rettferdige og konstruktive tilbakemeldinger.

En stor takk rettes til Astrid Øksnevad, Sandra Iren Kottum og Benedikte Årebrot Madsen for uvurderlig vennskap, konstruktive tilbakemeldinger og hjelp med språk.

En stor takk går også til Jens Christian Dinges for at du har pushet meg i tøffe perioder og motivert meg til videre arbeid. Jeg har satt stor pris på dine refleksjoner, gode humør og vennskap i denne perioden.

Jeg vil også takke min tyske vennefamilie og min pappa for å ha latt meg skrive og bo hos dem i perioder hvor jeg har trengt et avbrekk fra den hektiske hverdagen i Oslo.

Helt til slutt vil jeg takke alle dere andre som har holt ut med meg i denne perioden. Uten dere hadde ikke dette gått.

Oslo 28.09.2009

Synnøve Bjørn

Inholdsfortegnelse

INHOLDSFORTEGNELSE	9
1. INNLEDNING	11
2. EVALUERING – EN SKISSE AV FELTET	17
2.1 HVA ER EVALUERING?.....	17
2.1.1 Begrepsmessige distinksjoner.....	22
2.1.2 Hvilke standarder evalueres det opp mot?.....	26
2.2 HVORDAN EVALUERE?	28
2.2.1 Kvantitativ metode i evalueringen.....	29
2.2.2 Kvalitativ metode i evalueringen.....	31
2.3 EVALUERING OG NIVÅER FOR LÆREPLANTOLKNING.....	33
2.4 OPPSUMMERING OG KONKLUSJONER	37
3. ELEVVURDERING.....	39
3.1 HVA ER ELEVVURDERING?	39
3.1.1 Formativ og summativ elevvurdering.....	42
3.2 HVORFOR ELEVVURDERING ER AKTUELT – EN INNSIKT I DEBATTENE.....	45
3.3 ELEVVURDERINGENS HVA, HVORFOR, HVORDAN	48
3.4 OPPSUMMERING	50
4. RETTFERDIGHET	52
4.1 HVA ER RETTFERDIGHET?.....	52
4.1.1 <i>The Original Position</i>	53
4.1.2 <i>Distributive justice</i>	55
4.1.3 <i>Et spenningsforhold: likhet eller ulikhet?</i>	60
4.2 HVEM BESTEMMER HVA SOM ER RETTFERDIGHET?.....	63

4.2.1	<i>Rettferdighet på bakgrunn av ulike styringsformer</i>	63
4.2.2	<i>Nivåer for et rettferdig samfunn: hvem bestemmer?</i>	67
4.3	AVSLUTNING	71
5.	RETTFERDIGHET OG METODE FOR ELEVVURDERING	73
5.1	KVALITATIV OG KVANTITATIV METODE	74
5.1.1	<i>Kvalitative og kvantitative metoder for prøveform og vurdering – hva preger dem?</i> 74	
5.1.2	<i>En karakterdebatt</i>	77
5.2	KLARER VURDERINGSMETODEN Å FÅ FRAM ELEVENS KUNNSKAP?	81
5.2.1	<i>Taus kunnskap og metodetriangulering</i>	83
5.3	AVSLUTNING	86
6.	RETTFERDIG ELEVVURDERING OG BEGREPSVALIDITET	88
6.1	BEGREPSVALIDITET.....	88
6.2	SYSTEMATISKE OG TILFELDIGE MÅLINGSFEIL	92
6.3	VURDERERRELIABILITET	94
6.4	AVSLUTNING	99
7.	RETTFERDIGHET GJENNOM ULIKE VURDERINGSFORMER?	101
7.1	FORMER FOR VURDERING	101
7.1.1	<i>Tilbakemeldingen</i>	102
7.1.2	<i>Sluttvurdering</i>	106
7.2	AVSLUTNING	114
8.	AVSLUTNING	115
9.	LITTERATURLISTE	121
9.1	INTERNETKILDER.....	128

1. Innledning

Evaluering har blitt et stadig hetere tema i det norske samfunn. Temaet for denne oppgaven ligger innenfor evalueringsfeltet og fokuserer på elevvurdering. Mer presist fokuserer oppgaven på en *rettferdig elevvurdering*. Inspirasjon til temaet er hentet fra prosjektet Bedre Vurderingspraksis. Prosjektet tar utgangspunkt i og: ”Undersøke hva som skal til for at elevene i grunnskolen og i videregående opplæring skal få bedre faglige og *mer rettferdige vurderinger* [min kursiv] av arbeidet sitt” (Bedre vurderingspraksis 2007:3). Målet med prosjektet er å bruke elevvurdering som et redskap for å oppnå et høyere kunnskapsnivå i skolen. Dette mener utdanningsdirektoratet kan skje igjennom rettferdige elevvurderinger, men hva vil det si at en elevvurdering er rettferdig? Temaet rettferdig elevvurdering innebærer en rekke utfordringer. For det første hva slags vurderinger er det snakk om? Er det lærers vurderinger, elevens vurderinger eller samfunnets vurderinger? Er det formelle eller uformelle vurderinger? Er det formative eller summative vurderinger? Og hva er egentlig en vurdering? Hva vil det si at noen blir vurdert og hvilke konsekvenser har vurderingen for eleven? En annen utfordring ligger i forhold til rettferdighet. Hva er rettferdighet? Er rettferdighet når alle behandles likt eller er rettferdighet når det tas hensyn til individets forutsetninger og behov? Hva med innsats? Bør innsats telle som et perspektiv på rettferdighet?

Oppgavens tittel, ”**The heart of effective teaching**” [min utheving] (Black 2001:1), er hentet fra Paul Black and Dylan Williams artikkel ”Inside the black box” og henspiller på en oppfatning av elevvurdering som noe positivt og som et redskap for elevens faglige utvikling. Utført på riktig måte kan elevvurdering være et fantastisk redskap for utvikling av elevens kunnskapsnivå. Utført på feil måte kan elevvurderingen virke mot sin hensikt og få umotiverte og lite kunnskapslystne elever (Cullingford 1997:1). I første omgang kan prinsippet om en rettferdig elevvurdering tolkes dit hen at det innebærer en positiv utvikling av elevens læringsprosess. Ser en nærmere på konseptet ser en at ikke all elevvurdering nødvendigvis fører i positiv retning for elevens læringsprosess. Spørsmålet som reises i forlengelsen av dette, er om en rettferdig elevvurdering nødvendigvis innebærer at eleven får noe ut av vurderingen?

Denne oppgaven tar sikte på å fortelle noe om hva som ligger bak rettferdighetsbegrepet, og om dette kan forstås i forhold til elevvurdering. Problemstillingen er: *kan en rettferdig elevvurdering oppnås?* For og belyse dette vil jeg først se på evalueringsfeltet og vurderingsfeltet, før jeg går over og ser på ulike teorier for rettferdighet. De ulike teoriene for rettferdighet vil så bli drøftet i forhold til metode for elevvurdering, begrepsvaliditet og i forhold til ulike vurderingsformer. Intensjonen med oppgaven, er å gjennomføre en analyse av konseptet ”rettferdig elevvurdering”. I analysen blir teori fra både evalueringsfeltet og filosofien hentet inn. Målet mitt er å vise at rettferdighet må ses i forhold til de kriteriene som ligger til grunn, og at en rettferdig elevvurdering kan eller kan ikke, ligge i hjertet av en effektiv læring.

Denne oppgaven er en pedagogisk filosofisk, oppgave og ligger i spenningen mellom evalueringsteori og moralfilosofi. Evalueringsteorien er hentet fra pedagogikken og befinner seg både på systemnivå (Scriven 1991) og individnivå (Black 2001,2004). Da det ikke er noen pedagogisk filosofi eller teori som omhandler rettferdighet, har det vært nødvendig å hente inn perspektiver på rettferdighet fra moralfilosofien (se Lamont 2007). Dette viser at oppgaven ligger i krysningen mellom et filosofisk, abstrakt perspektiv og et teoretisk, empirisk perspektiv. Det er altså et spørsmål om å knytte pedagogisk teori sammen med moralfilosofi. At oppgaven er sentrert innenfor pedagogisk filosofi innebærer i dette tilfellet at det blir foretatt en analyse (se Løvlie 1992) av begrepene evaluering, elevvurdering og rettferdighet. Metoden for analysen er hermeneutisk og kjennetegnes igjennom sin vekselvirkning mellom enkeltdelene og helheten til oppgaven (Alvesson og Sköldberg 1994:115).

En hermeneutisk tekst vil alltid preges av skribentens forforståelse på den ene siden og leserens forforståelse på den andre siden. En tekst konstrueres når man ser enkeltdelene i teksten i lys av helheten og helheten i lys av enkeltdelene. Den hermeneutiske sirkelen tar utgangspunkt i at man bare kan forstå helheten ut fra delene og delene ut ifra helheten (Alvesson og Sköldberg 1994:116). Denne sirkulære vekselvirkningen kan oppfattes som at teksten ikke kommer noe sted, så for og ikke å bli stående fast i den hermeneutiske sirkelen, benyttes også begrepet hermeneutisk spiral (Hjardemaal 2002:41). Den hermeneutiske spiralen sikter på å gi en stadig en dypere forståelse for temaet ved og kontinuerlig dra inn nye perspektiver. Disse perspektivene utvikler forståelsen for temaet og gir nye perspektiver til diskusjonen.

I denne besvarelsen viser hermeneutikken seg ved at begrepene evaluering, elevvurdering og rettferdighet blir analysert. Dette skjer først ved å vise til ulike tolkninger av begrepene, for så å se tolkningene i forhold til hverandre og til slutt sette tolkningene sammen til en mer helhetlig forståelse. Tolkningene som denne teksten benytter seg av, representerer delene i den hermeneutiske sirkelen. Sammen danner begrepene evaluering, elevvurdering og rettferdighet en forforståelse og en plattform for problemstillingen: *kan en rettferdig elevvurdering oppnås?* Når problemstillingen skal drøftes, drar jeg inn yterligere momenter for å gi en dypere forståelse av temaet. Disse momentene er hentet fra vitenskapsmetoden og elevvurderingen. Ut ifra dette ser vi at den hermeneutiske sirkelen utvikler til en spiral.

Før jeg starter min analyse av rettferdig elevvurdering er det noen avgrensninger som må gjøres. Rettferdig elevvurdering kan analyseres på mange måter, den første skillelinjen som skal trekkes er mellom en normativ og en deskriptiv analyse. Når en rettferdig elevvurdering skal analyseres, har deskriptive og normative analyser ulike formål. En deskriptiv analyse er empirisk (Hjardemaal 2002:57) og ville tatt utgangspunkt i å spørre hvordan rettferdig elevvurdering kommer til uttrykk i praksis. Min intensjon med denne oppgaven er for øvrig ikke deskriptiv, men normativ. Kvernbekk skriver at: "Normative, målrettete teorier beskriver utviklingen mot en *ønsket måltilstand*" (Kvernbekk 2005:74). Denne oppgaven ønsker å si noe om rettferdig elevvurdering som en "ønsket måltilstand". Dette vil igjen si at jeg ikke kommer til å ta for meg hvordan rettferdig elevvurdering kommer til uttrykk i praksis, men fokusere på ulike vinklinger til om en rettferdig elevvurdering kan oppnås.

Den neste avgrensningen som gjøres i oppgaven er en avgrensning i forhold til elevvurdering. Jeg avgrenser elevvurderingen til et forhold mellom lærere, sensorer og eleven. Med dette fokuserer jeg ikke på elevvurdering som vurdering mellom elever og elevens vurdering av seg selv. Årsaken til at elevvurdering mellom elever og elevens selvutvurdering ikke er behandlet i oppgaven, ligger i at dette er uformelle vurderingsformer som ikke stiller noe spesifikt krav til rettferdighet. Jeg anser det heller ikke som relevant å diskutere rettferdighet i disse vurderingsformene, da elevenes kompetanse for vurdering og innad i faget gjerne er begrenset.

Når det kommer til rettferdighetsteorier er det også blitt gjort en del avgrensninger. Jeg har her valgt ut de rettferdighetsteoriene jeg har ansett som mest relevante for å få fram en diskusjon av hva en rettferdig elevvurdering er. Perspektiver jeg har valgt å fokusere på ligger

innenfor hva Lamont og Favor (2007) kaller *distributive justice*. Perspektivene som blir behandlet er egalitarismen, differanseprinsippet, libertarianismen, resursprinsippet og fortjenesteprinsippet. Perspektiver som faller bort er prinsipper om velferd og feminisme. Velferdsprinsippet tar utgangspunkt i fordeling av økonomiske resurser. At jeg ikke tar med dette perspektivet, betyr at jeg heller ikke får noen diskusjon ut ifra hva læreren får igjen økonomisk for å vurdere elevene. Feministprinsippet diskuterer kvinners rettigheter. I norsk elevvurdering anser jeg ikke dette som en relevant diskusjon, da jeg har tro på at dagens lærere vurderer gutter og jenter mer eller mindre likt. Det er faktisk mulig å stille spørsmål om den norske skolen er mer tilrettelagt for jenter enn for gutter. Noe som har resultert i at guttene får dårligere resultater enn jentene (se f. eks: Viksmo-slettan 2007, og Aftenposten Nettutgave 09.02.09). Dette er for øvrig en helt annen debatt og kommer ikke til å bli behandlet i denne oppgaven.

Helt til slutt i innledningen skal jeg nå ta for meg oppgavens kapitler. Oppgaven har til sammen 8 kapitler inkludert innledning og avslutning.

Kapittel 2 tar for seg evalueringbegrepet. Det er ulike oppfatninger rundt hvordan evaluering og vurdering skal ses i forhold til hverandre. I denne oppgaven anser jeg evalueringsteori som et rammeverk for elevvurderingen. En vandlig oppfatning er at evalueringbegrepet brukes på systemnivå, mens vurdering brukes i forhold til vurderinger av enkeltpersoner. Før jeg begynner å skrive om vurderingsbegrepet anser jeg det som viktig å skissere ut feltet elevvurdering ligger i. Kapitlet tar først for seg hva evaluering er. Dette blir gjort ved å se på ulike definisjoner av evalueringbegrepet hvor spenningene mellom systemnivå og individnivå kommer fram. I tillegg ser jeg på begrepsmessige distinksjoner og hvilke standarder det evalueres opp mot. Deretter tar kapitlet opp spørsmålet "Hvordan evaluere?". Dette blir belyst igjennom kvantitativ og kvalitativ metode. Til slutt i kapitlet blir evaluering diskutert opp mot nivåer for læreplantolkninger. Den siste delen er ment til å bygge bro inn i kapittel 3 som tar for seg elevvurdering.

I kapittel 3 avgrenses oppgaven inn mot elevvurdering. Elevvurdering handler grovt sett om den evalueringen elevene utsettes for i forbindelse med deres arbeid i skolen. I kapittel 3 blir det først stilt et spørsmål om hva elevvurdering er. For å gi en bred forståelse av elevvurdering, blir også begrepsparet formativ og summativ elevvurdering diskutert under dette spørsmålet. Videre stiller kapitlet spørsmålet: "Hvorfor er elevvurdering aktuelt?". Her

blir det vist til PISA undersøkelsene, debatter fra avisene og Utdanningsdirektoratets prosjekt "En bedre vurderingspraksis". Det siste som tas opp i forbindelse med elevvurderingskapittelet er: "Elevvurderingens hva, hvorfor, hvordan". I denne delen blir spørsmålene "Hva vurderes eleven i?", "Hvorfor vurderes eleven?" og "Hvordan vurderes eleven?" diskutert.

I kapittel 4 blir rettferdighetsbegrepet tatt opp for første gang. Spørsmålet om hva rettferdighet er kan ikke svares med en gitt definisjon. Det blir derfor i kapittel fire tatt opp ulike tilnærminger til hvordan forstå rettferdighetsbegrepet. Disse er blant annet original position, egalitarisme, differanseprinsippet og desert-/ fortjenestep prinsippet. I tillegg blir ulike samfunnsnivåer og rettferdighet diskutert. Er det slik at ulike samfunnsnivåer også gir ulikt grunnlag for å diskutere rettferdighet på? Hvilket samfunnsnivå diskuteres en rettferdig elevvurdering på?

I de tre siste kapitlene drøftes oppnåelsen av en rettferdig elevvurdering. Dette blir gjort igjennom metode, begrepsvaliditet og ulike former for vurdering, som tilbakemelding og sluttvurdering.

Vurdering kan på mange måter forstås som metode. En metode sier noe om hvordan eleven kan videreutvikle seg og måler hvilket nivå eleven ligger på. Fordi metode kan forstås som et redskap for elevvurderingen, har jeg i kapittel 5 og 6 valgt å fokusere på om riktig metodikk kan skape en rettferdig elevvurdering.

Kapittel 5 er det første kapittelet som drøfter om en rettferdig elevvurdering kan oppnås. Dette blir gjort med utgangspunkt i et spørsmål om forskningsmetode, en mer sikker elevvurdering, vil skape en mer rettferdig elevvurdering. Kan metode gjøre elevvurderingen mer rettferdig? Det blir her diskutert om ulike metoder kan gi rettferdige elevvurdering. Overordnet blir det sett på kvalitativ og kvantitativ metode for elevvurdering, som igjen utformer seg til en debatt om karakterer kan anses for å være rettferdige eller ikke. Et annet spørsmål er om metodetriangulering kan fungere som et ideelt redskap for vurdering og slik gi en mer rettferdig vurdering.

Kapittel 6 spinner videre på tankegangen i kapittel 5 og drøfter begrepsvaliditet i forhold til om en rettferdig elevvurdering kan oppnås. Hvilken betydning spiller validiteten i selve vurderingen for om vurderingen kan oppfattes som rettferdig eller ikke. Er fokus på

begrepsvaliditet og felles forståelse mellom lærer og elev en nøkkel til å oppnå rettferdige elevvurderinger?

Kapittel 7 er det siste hovedkapittelet i oppgaven og tar opp rettferdighet igjennom ulike vurderingsformer. Teori fra elevvurderingskapitlene og rettferdighetskapittelet blir her drøftet opp mot hverandre i en større helhet. Det vil i dette kapittelet komme tydelig fram at rettferdighet er relevant i forhold til hvilken teori som benyttes for å måle hva som er rettferdig og hvilke prinsipper som ligger til grunn for elevvurderingen.

Til slutt i oppgaven kommer en avslutning som har til hensikt å summere opp poengene og konklusjonene i oppgaven. Ulike sider ved om en rettferdig elevvurdering kan oppnås, blir her poengtert.

2. Evaluering – En skisse av feltet

Hensikten med dette kapitlet er å gi en skisse av evalueringsfeltet som elevvurdering ligger innenfor. Evaluering blir her betraktet som et overordnet rammeverk for elevvurdering. Det vil si at trekk som viser seg i evalueringsteorien også kan benyttes i elevvurderingen. Selv om evaluering og elevvurdering ikke er det samme, henter elevvurderingen store deler av begrepsapparatet sitt fra evalueringsteorien. Kapitlet er delt inn i tre underkapitler som alle tar utgangspunkt i evaluering. Det første kapitlet tar opp spørsmålet ”Hva er evaluering?”. I denne delen blir evalueringsbegrepets innhold, distinksjoner og standarder for måling diskutert. Det andre underkapitlet stiller spørsmålet ”Hvordan evaluere?”. Her blir evaluering diskutert gjennom kvantitativ og kvalitativ metode. Til slutt blir evaluering sett i forhold til nivåer for læreplantolkning. Dette siste kapitlet er ment som en innsnevring inn mot skolesystemet og elevvurderingen. Innsnevringen av evalueringsfeltet er ment til å skape en bro mellom evaluering som felt og elevvurdering i skolen. I kapitlet for øvrig blir det brukt en del eksempler fra skolen, dette med tanke på at resten av oppgaven har skolen som handlingsarena.

2.1 Hva er evaluering?

Selv om svaret på dette spørsmålet er kontekstavhengig, er målsetningen for delkapitlet å vise ulike måter begrepet kan forstås på og å gi en definisjon som dekker feltet i sin helhet. Delkapitlet begynner med en skisse av bredden på vurderingsfeltet. Deretter følger en drøfting av ulike definisjoner av evalueringsbegrepet. Etter det er kommet fram en hensiktsmessig definisjon av evaluering, vises det til begrepsmessige distinksjoner som muliggjør en drøfting av ulike sider ved evalueringsbegrepet. Når dette er gjort fokuserer kapitlet på hvilke standarder det evalueres opp mot. Å se distinksjonene og standardene for målinger i evalueringsbegrepet gir en forståelse av hva evaluering er og viser samtidig spennvidden til begrepet.

Evaluering er med på å prege menneskers liv på flere måter, dette gjelder både privat og på arbeidsplassen. Evalueringen kan påvirke enkeltpersoner så vel som grupper, samfunn, teknologisk utvikling og organisasjoner. Hva er så dette fenomenet som kalles evaluering? Gunn Insen skriver: ”Evaluering stammer fra det latinske verbet *valere*, som betyr å være verdt eller å sette verdi på noe” (Insen 1999:306). Verdier gir utgangspunkt for hvordan mennesker evaluerer. Hva som er av verdi kan være en individuell avgjørelse så vel som en samfunnsmessig avgjørelse. Før jeg går nærmere inn på hva evaluering er, anser jeg det som viktig å diskutere hvor evaluering skjer.

Det kan hevdes at verdier finnes over alt, noe som vil si at evalueringer finnes over alt. Dette er en tese som støttes av blant annet Michal Scriven. Scriven skriver om “*the ubiquity of evaluation*”. I et sitat hentet fra boka *Evaluation Thesaurus (1991)* skriver Scriven: “The process of disciplined evaluation permeates all areas of thought and practice, evading primary location in any one. It is found in scholarly book reviews, in engineering’s quality control procedures, in the Socratic dialogues, in serious social and moral criticism, in mathematics, and in the opinions handed down by appellate courts” (Scriven 1991:4). I følge det Scriven skriver her, er evaluering en del av hvordan mennesker tenker og handler. Det vil si at før et menneske handler, foretar det også en evaluering av hvilke muligheter det har. For eksempel: når du går tur i skogen og kommer til et veikryss må du foreta en evaluering av hvilken vei du skal gå. Din evaluering vil da basere seg på tre forhold: fortid, framtid og nåtid (Scriven 1991:361). Dette legger et mangslungent evalueringsgrunnlag. For eksempel kan turen i skogen baseres på følgende evalueringer: Kanskje vet du om en koselig hytte der de serverer vafler og kakao. Da er dine kunnskaper fra fortiden med på å informere deg om hytta. Du vurderer din egen form og motivasjon for å gå til hytta i nåtiden, og du vurderer vafler og kakao som en fremtidig motivasjon. Eksempelet viser hvordan fortid, nåtid og framtid kan prege valgene dine. Forkunnskap, dagsform og mål er alle faktorer som vil påvirke en evaluering. I tillegg kan en ta hensyn til påvirkning utenifra, og konteksten en befinner seg i.

Ikke alle er enige med Scriven i at alt kan evalueres. Reidar Almås skriver: ”Det er heller ikke alle tema som er *evaluerbare*. Sosiale intervensjonar kan vera så dårleg planlagte, eller evaluatoren bli kopla så seint inn, at evalueringsoppdraget ikkje kan gjennomførast” (Almås 1990:16). Spørsmålet er om dårlig planlegging og/eller at den som foretar evalueringen kobles sent inn i evalueringsoppdraget, kan lede til konklusjonen at ikke alle temaer er evaluerbare? På den ene siden kan årsakene som dårlig planlegging og sen tiltredelse av evalueringsperson

ha ført til at det ikke ble framlagt noen rapporter i evalueringsoppdraget. På den andre siden betyr ikke dette at det ikke ble foretatt noen evaluering. Evalueringer trenger ikke være formelle rapporter, de kan like gjerne foregå i et menneskes hode, som i eksempelet med skogturen ovenfor. Almås sitt sitat gir derfor ikke grunnlag til å hevde at ikke alt er evaluerbart. At et formelt evalueringsoppdrag ikke ender i en formell rapport, behøver ikke å bety at det ikke er blitt foretatt noen evalueringer.

Evaluering kan forstås som et transdisiplinært begrep. Det vil si at evaluering er relevant på flere felter. Scriven hevder at "Transdisciplines such as statistics, logic, or evaluating, are disciplines whose subject matter is the study and improvement of certain tools for other disciplines" (Scriven 1991:364). Det er allikevel ikke slik at dersom du har kunnskap om evaluering at du kan evaluere hva som helst. Evaluering handler også om kontekst og kompetanse i feltet som blir evaluert. I eksempelet om skogturen ovenfor ble det vist til forkunnskaper. Forkunnskaper i evaluering er kjekt å ha, men alene gir det et heller spakt grunnlag for evaluering. For at en evaluering skal kunne oppfattes som gyldig, er det vel så viktig å ha forkunnskaper eller kompetanse innenfor disiplinen som evalueres. Almås påpeker i denne sammenhengen at evalueringens brukergrupper vil ha ulike interesser (Almås 1993:21). Dersom en ser på skogturen i eksempelet ovenfor, ble det vist at ulike formål og ulike bakgrunnsforståelse vil kunne gi grunnlag til å velge ulike stier. Med andre ord, ulike turgåere har ulike interesser og evaluerer derfor ulikt. Dette vil også gjelde dersom en ser på eksempler hentet fra skolen, bedrifter eller i hjemmet.

Det ble ovenfor nevnt at evaluering handler om verdier. Verdier kan være individuelle og de kan være felles i et samfunn. Dette gir et utgangspunkt for en svært vid definisjon av evalueringbegrepet. Videre skal ulike definisjoner av evaluering drøftes. Selv om alle definisjonene handler om verdier er det forskjeller i hvordan definisjonene oppfatter verdiene og hvilket nivå verdiene ligger på. Det er mange forskjellige tilnærminger til hva evaluering er. Jeg begynner med definisjoner hentet fra Reidar Almås og Evert Vedung. Både Almås og Vedung har definisjoner på evaluering på systemnivå og ikke et personnivå. Kritikken jeg retter mot Almås og Vedung sine definisjoner av evaluering gjelder ikke deres bruk av begrepet, men mot hvordan definisjonene står i forhold til et mer helhetlig evalueringbegrep. Etter Almås og Vedungs definisjoner av evaluering har blitt drøftet, står Anders Lysne for tur. Til slutt ender jeg med en definisjon av Michal Scriven som jeg mener dekker evalueringbegrepet. Scrivens definisjon vil legge grunnlaget for oppgavens videre gang.

Reidar Almås forstår evaluering som ”*systematisk innsamling av data for å skilja og analysere verknaden av eit forsøk på å skapa endring på eit gitt område*”(Almås 1993:13). Almås sin definisjon på hva evaluering er, kan anses som snever av flere grunner. For det første omhandler definisjonen kun evaluering som er systematisk. Kan ikke evaluering også være usystematisk? For det andre setter Almås fokus på at evaluering skal føre til endring på et gitt område. Ikke all evaluering har til hensikt å endre. Det kan blant annet skilles mellom deskriptiv og normative evalueringer (Scriven 1991). Mens den normative evalueringen kan ha et mål om endring (Scriven 1991: 245), har den deskriptive evalueringen et mål om å beskrive (Scriven 1991:122). Et eksempel på en deskriptiv evaluering kan være karakterer i skolen. Karakteren har som mål å beskrive nivået eleven ligger på og karakteren gir ikke føringer som har til hensikt å endre elevens læringsprosess. Senere i oppgaven blir deskriptiv og normativ evaluering framstilt under begrepene formativ og summativ evaluering/vurdering.

Evert Vedung har en lignende definisjon som Almås. Vedung definerer begrepet evaluering som ”*en mekanism för systematisk kartläggning och bedömning av pågående eller avslutade offentliga interventioners implementering och effekter på omgivningen, vilken syftar till att producera information som tänkes komma til använding i framtida beslutsprocesser*” (Vedung 2002:5). I likhet med Almås vektlegger Vedung systematikk i evalueringen. Som nevnt ovenfor behøver ikke evaluering å være systematisk, den kan like gjerne være usystematisk. Et annet moment Vedung trekker fram som Almås ikke har lagt vekt på i sin definisjon, er bedømming i pågående eller avsluttende offentlige intervensjoner. Et viktig poeng som Vedung her får fram er at evaluering både kan skje under prosessen og etter prosessen. Det som derimot er kritikkverdig er at Vedung avgrenser evaluering til å handle om det offentlige. Dersom en skal tolke evaluering som en verdi, er det verdt å nevne at verdier ikke kun er offentlige, men også i høyeste grad private. Derfor behøver ikke evaluering å avgrenses til det offentlige. Dette ble også vist i eksempelet om skogturen ovenfor. Helt til slutt er det verdt å nevne at i likhet med Almås sin definisjon av evaluering, vektlegger også Vedung forbedringsmomentet i evalueringen. Selv om evalueringens muligheter for forbedring er viktig, har Vedungs definisjon en mangelfull forståelse av hva evaluering er.

Både Almås og Vedungs poenger er viktige når en skal avgjøre hva evaluering er. Almås og Vedungs momenter om endringer, prosess og produkt vil bli tatt med videre i oppgaven.

Anders Lysne definerer evaluering i et didaktisk perspektiv. I følge Lysne kan evaluering ”kjennetegnes ved at den alltid er et element i den veiledende kommunikasjon mellom deltakende aktører underveis i en styrt læreprosess, som regel kombinert med innslag av oppsummerende resultatkontraster på flere stadier i prosessen og ved avslutningen av et læreprogram for å se i hvilken grad de oppsatte mål er nådd” (Lysne 1999:10). Forskjellen mellom Lysnes definisjon av evaluering og Almås og Vedung sine definisjoner, er at Lysne knytter evaluering opp mot personer og ikke systemer. Likhetene mellom Almås, Vedung og Lysne er at alle tre definisjonene omhandler evaluering på offentlige arenaer. I tillegg vektlegger alle veiledning som et viktig moment i evalueringen. Kritikk som kan rettes mot Lysne er at ikke all evaluering skjer i en styrt læreprosess. Eleven lærer mye utenfor skolen også.

Det er mye læring som ikke er styrt, som også blir evaluert. For eksempel blir kan man lære innsats, flid og å møte opp i tide. Slik atferd kan ikke styres av lærer på samme måte som læring av fag. Lysnes definisjon av evaluering er avgrenset til et klasseromsperspektiv og tar derfor ikke hensyn til faktorer utenfor klasserommet. Selv om skole og klasserom er arenaer for en type evaluering, gjør ikke dette Lysne sin definisjon av evaluering, generaliserbart til andre arenaer.

Definisjonene ovenfor beskriver evaluering brukt som statsforvaltning. Den siste definisjonen er hentet fra Michal Scriven og vil danne utgangspunktet for hvordan evaluering forstås i resten av oppgaven. Scriven skriver: “Evaluation is the process of determining the merit, worth and value of things, and evaluations are the products of that process” (Scriven 1991:1). Scriven skiller her mellom evaluering på to stadier. På den ene siden er den kontinuerlige evalueringen en prosess og på den andre siden er evalueringen av produktet. Prosess og produkt kan også ses i forhold til formativ og summativ evaluering (Engelsen 2006:121). Jeg kommer i del 2.1.1 nærmere inn på hva som ligger bak formativ og summativ evaluering. I likhet med det Imsen (1999) gjorde innledningsvis i dette kapittelet, vektlegger Scriven verdier.

Vi har til nå sett ulike definisjoner av hva evaluering er. Ofte avgrenses definisjoner av evaluering til et relevant bruksområde. I denne oppgaven er hensikten med evaluering å skissere feltet som elevvurdering ligger i. Feltet i seg selv er stort og krever derfor en bred definisjon. Etter å ha forkastet en rekke definisjoner av evaluering fordi de var for rettet mot

statsforvaltning, har jeg valgt en definisjon av Scriven som vektlegger verdier, prosess og produkt. Scrivens definisjon viser bredden i evalueringsfeltet og kan derfor også ses i forhold til enkeltelever, som oppgaven fra kapittel 3 og utover retter fokuset mot. Videre vil jeg studere begrepsmessige distinksjoner som evalueringsbegrepet preges av og hvilke standarder evaluering kan settes opp mot.

2.1.1 Begrepsmessige distinksjoner

Bredden i evalueringsbegrepet kan belyses gjennom flere begrepsmessige distinksjoner. I dette kapitlet vil det redegjøres for fire begrepspar som bidrar til å forklare evalueringsbegrepet. Det første er subjektivitet og objektivitet i evalueringen. Det andre er formell og uformell vurdering. Det tredje er formativ og summativ vurdering. Det fjerde er intern og ekstern vurdering. Begrepene har til hensikt å forklare ulike sider av evalueringen. Begrepene rommer en rekke tolkninger. Som Gregersen skriver kan dette ”føre til mengder af misforståelser og skin-præcision til følge, fordi de av brugerne opfattes på forskjellig måde” (Gregersen 1977:262). Et godt eksempel på dette er når begrepsparene formativ og summativ og prosess og sluttprodukt brukes om evaluering.

Det første begrepsparet som blir belyst er subjektivitet og objektivitet. Hva evaluering er og hvilket utfall evalueringen har, kan preges av subjektiviteten og objektiviteten i evalueringen. Det finnes forskjellige meninger rundt subjektivitet og objektivitet i evalueringsteorien. Hva vil det si at en evaluering er subjektiv? Engelsen skriver: ”Subjektiv vurdering blir brukt om vurderingsformer som gir den som vurderer, stor mulighet til å bruke sin egen tolkning og sitt eget skjønn” (Engelsen 2006:124). Hva vil dette si? Er evalueringen subjektiv gjennom prosessen og til produktet? Vil ikke subjektivitet være bakgrunnen for alle evalueringer?

Som motpol til evalueringens subjektive side står objektivitet. Hva vil det si at evalueringen er objektiv? Kan objektivitet oppnås i en evaluering? Engelsen hevder at: ”Ved en objektiv vurdering er resultatet uavhengig av hvem det er som vurderer” (Engelsen 2006:124). Er dette realistisk? Hvordan kan en evaluering være uavhengig av hvem som vurderer? Evaluering kan forstås som et sosialt fenomen. Max Weber skriver ”Det finnes ingen rett og lett «objektiv» vitenskapelig analyse av kulturlivet eller av «sosiale fenomener» uavhengig av spesielle og «ensidige» synspunkter” (Weber 2000:181). Objektiviteten kan i et slikt perspektiv kun

fremstå som et ideal i evalueringsteorien. Om objektivitet skriver Scriven: "The technical sense of this term refers to a fairly specific description of an intended outcome; the more general description, under which this and possibly other objectives are subsumed, is referred to as a goal" (Scriven 1991:248). Hva evaluering er, preges altså først og fremst av individets syn på hvordan evaluere. Dette kan variere fra person til person. Vedung tar opp tre spørsmål til objektivitet i evalueringsbegrepet. Det første spørsmålet han skisserer er: "Is there a match between what actually happens in the real world and what might be expected or desired from some evaluative point of view?" (Vedung 2003:41). Det pekes her på evalueringens subjektive side. Hvem som evaluerer og bakgrunnsforståelsen til den som foretar evalueringen er av betydning for hvordan det evalueres. Det andre spørsmålet Vedung stiller er: "Is what actually transpires in the world in any way a product of intervention?" (Vedung 2003:41). Hvordan påvirker andres innblanding evalueringen? Scriven hevder at andres innblanding kan påvirke vurderingsevnen til den som foretar evalueringen (Scriven 1991:199). Det tredje spørsmålet til Vedung er: "Are the answers to the first two questions useful for or used by some stakeholder in the intervention?" (Vedung 2003:41). Vedung sine spørsmål er viktige. Hvordan vet vi at evalueringen dekker hva som faktisk har skjedd? Hvilken grad av samsvar er det? Hvordan påvirker andre menneskers innspill evalueringen? Evaluering handler om menneskers tolkninger. Dette er tolkninger av verdier. Problematikken rundt objektivitet blir drøftet videre i kapittel 5 og 6.

Et annet begrepspar som er med på å belyse hva evaluering er, er begrepene formell og uformell evaluering. Disse begrepene kan ses i forhold til hvor stor grad av objektivitet som er ønsket i evalueringen.

Hva evaluering er, kan preges av om evalueringsteorien beror på formell eller uformell praksis. Skillet mellom formell og uformell vurdering er viktig fordi det kan ha konsekvenser for hvordan evalueringen utføres og hvordan mottaker tar evalueringen til etterretning. I følge Engelsen skjer en formell evaluering innenfor et fast system (Engelsen 2006:123). Karakterer er et eksempel på en formell vurdering fordi karakterer gjerne følger formelle retningslinjer for hva karakteren betyr. Den uformelle vurderingen tar derimot for seg daglige evalueringer mellom mennesker. Brady definerer det slik: "informal assessment is when a teacher assesses a pupil performance (perhaps through some simple exercise or test) and the result is communicated only to the learner; it is used to remedy some immediate problem or to provide encouragement" (Brady 1997:15). Selv om Brady skriver om et lærer –elevforhold kan synet

på uformell vurdering også benyttes for/i et større evalueringsfelt. Den uformelle evalueringen handler om ustrukturert kommunikasjon som har til hensikt å veilede en annen part. Det kan også hevdes at selvevaluering er en form for uformell evaluering, men det er ikke tema for denne oppgaven. Også Engelsen bruker dialogen mellom lærer og elev som eksempel på uformell vurdering (Engelsen 2006:123). At dialogen mellom lærer og elev er uformell begrunner hun med at dialogen ikke følger formelle kriterier for evaluering. Evalueringen er allikevel en viktig del av tilbakemeldingene. Dette gjelder spesielt under prosessen. Prosessevaluering er et av momentene som kommer fram under formativ evaluering.

Det tredje begrepsparet er formativ og summativ evaluering. I følge Peder Haug henger disse to sammen. Haug skriver ”Det som av ein person er opplevd som formativ evaluering, vil for ein annan vere summativ” (Haug 1994:71). Om personen ser seg selv i prosessen for evaluering kan ha betydning for om evalueringen oppleves som formativ eller summativ. Det behøver for øvrig ikke bety at hensikten med evalueringen var ment å være slik personen oppfattet den. Når det videre redegjøres for begrepene formativ og summativ evaluering, er det derfor viktig å være seg bevisst at hensikten med evalueringen og opplevelsen av evalueringen ikke alltid stemmer overens og at overgangen mellom hva som er formativt og hva som er summativt kan være noe diffus.

Michal Scriven var en av de første til å benytte begrepene formativ og summativ evaluering (Engelsen 2006:121). Om formativ evaluering skriver han: ”It is typically conducted during the development or improvement of a program or product (or person, and so on) and it is conducted, often more than once, for the in-house staff of the program with the intent to improve (Scriven 1991:169). Scriven viser at en formativ evaluering har to vesentlige kjennetegn. For det første skjer det under prosessen og for det andre har den til hensikt å forbedre. Gunn Imsen bruker begrepet prosessvurdering istedenfor formativ evaluering. Imsen knytter prosessvurderingen opp mot veiledning og korreksjon av kursen i opplæringens forløp (Imsen 1999:325). I Imsens sitt perspektiv blir formativ evaluering en form for veiledning som tar sikte på å forbedre læringsprosessen. Om motstykket, summativ evaluering, skriver Scriven:

Summative evaluation of a program is conducted after completion of the program (for ongoing programs, that means after stabilization) and for the benefit of some external audience or decision-

maker (for example, funding agency, oversight office, historian, or future possible users), though it may be done by either internal or external evaluators or a mixture” (Scriven 1991:340).

I kontrast til den formative evalueringen legger Scriven i sitatet ovenfor vekt på at den summative evalueringen skal ha en hensikt utover å opplyse den personen eller saken som blir evaluert. Det vil si at en summativ evaluering også kan ha betydning for et eksternt publikum. Et eksempel på dette er når elever får eksamenskarakterer eller standpunktkarakterer i den norske skolen. Disse karakterene er summative karakterer og vil ha betydning når eleven skal søke seg videre inn i arbeidsliv eller skole (Lauvås 2007).

Skillet mellom de to evalueringsformene ligger i hva som vektlegges og når evalueringen foregår. Scriven viser til Robert Stake når han sammenligner formativ og summativ vurdering. ”When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative” (Scriven 1991:169). I kapittel 3 blir formativ og summativ vurdering tatt opp på nytt i forbindelse med elevvurdering.

Et fjerde begrepspar i evalueringsteorien er intern og ekstern evaluering. Det er her et spørsmål om hvem som evaluerer. Hva er internevaluering og hva er eksternevaluering? ”The internal evaluator knows the program better and so avoids mistakes due to ignorance, knows the people better and hence can talk to them more easily, will be there after the evaluation is finished and hence can facilitate implementation, probably knows the subject matter better, cost less, and is sure to *know of* some other comparable projects for comparison” (Scriven 1991:160). I det Scriven skriver om intern evaluering vises det at det interne evalueringsprinsippet bygger på en rekke subjektive forhold. Som det vil bli drøftet i kapittel 5 og 6 kan subjektivitet få både positive og negative konsekvenser for evalueringen. Også eksternevaluering har et subjektivt grunnlag, men fordi den som foretar evalueringen ikke har noen direkte tilknytning til det som evalueres, er det grunnlag for å tro at eksternevaluering er mer objektiv. Scriven skriver om eksternevaluering: “The external evaluator is less likely to be affected by personal or job-benefit considerations, is often better at evaluation; has often *looked closely at* comparable programs, can speak more frankly because there is less risk of job loss or personal retribution/ dislike, and carries some cachet from externality and as Freud observed, cost” (Scriven 1991:160). Ut ifra det Scriven skriver her, vises det at eksternevaluering har et mer nøytralt utgangspunkt for å foreta evalueringen enn det

internevalueringen har. Det behøver for øvrig ikke bety at kvaliteten på eksternevaluering er bedre enn på en internevaluering. Kvaliteten på evalueringen vil også la seg påvirke av fag- og evalueringskompetanse hos intern og ekstern evaluatør.

Begrepene som har blitt skissert ovenfor belyser ulike sider av evalueringsbegrepet. Det ble først hevdet at all evaluering er subjektiv, og at objektivitet er et ideal evalueringsteoretikerne ønsker å strekke seg etter. Deretter ble det vist at den formelle og uformelle evalueringen beskriver måten evalueringen uttrykkes på. Blir evalueringen gitt etter et visst regelverk, eller blir den gitt som et spontant innslag? Det tredje begrepsparet var formativ og summativ evaluering. Dette beskrev blant annet tidspunktet for evalueringen og måten den blir gitt på. En formativ tilbakemelding har mer innhold og er ment til å veilede, mens en summativ skal oppsummere resultatet. Det fjerde begrepsparet er et spørsmål om hvem som skal evaluere. Er det en intern eller ekstern sensor? Kvaliteten på intern og eksternevaluering ble også kort drøftet. Perspektivene som har blitt belyst i denne delen vil bli benyttet videre i avhandlingen. Jeg skal nå drøfte hvilke standarder det evalueres opp mot.

2.1.2 Hvilke standarder evalueres det opp mot?

Bredden i evalueringsbegrepet kan vises gjennom hvilke standarder evaluering måles opp mot. En standard kan ses på som en målestokk for hvilke premisser som ligger til grunn for evalueringen. Evalueres det i forhold til individets standarder, i forhold til en gruppe eller i forhold til et sett gitte kriterier? Det skilles her mellom individbasert evaluering, grupperelatert evaluering og kriteriebasert evaluering. Jeg kommer til å bruke eksempler hentet fra virksomhet i skolen.

Som det har blitt vist tidligere i dette kapittelet kan evaluering foregå på individnivå. Når evaluering foregår på individnivå blir det ofte betegnet som vurdering. Et perspektiv for evaluering av individer er individrelatert vurdering. Engelsen skriver: ”Ved en *individrelatert vurdering* blir den enkelte eleven vurdert mot sine egne forutsetninger” (Engelsen 2006:121). Dette fordrer en rekke spørsmål. Hvordan vet man hva eleven eller individets forutsetninger er? Individets forutsetninger og behov kan best ses ut ifra Lev Vygotskys modell om den nære utviklingssonen. Den nære utviklingssonen er en teori om hva barnet kan klare på egenhånd og hva barnet kan klare med hjelp fra en mer kompetent støttespiller (i Tetzchner 2003:199).

Sett i lys av den nære utviklingssonen, kan individrelatert vurdering forstås som et redskap som den mer kompetente personen kan benytte for å hjelpe individet til å mestre ferdigheter det i utgangspunktet ikke hadde forutsetning til å klare på egenhånd. Også i arbeidslivet benyttes individrelatert vurdering. Eksempler kan hentes fra praksisopplæring på sykehus, opplæring av snekkere, rørleggere, psykologer etc. Dette er alle yrker hvor individrelaterte vurderinger blir benyttet for å kunne tilpasse opplæringen til individets forutsetninger.

Tidligere ble det i den norske skole benyttet en såkalt grupperelatert evalueringsform. I grupperelatert evaluering ”blir den enkelte eleven vurdert mot prestasjonene i den gruppen han hører til (Engelsen 2006:121)”. Dette tar gjerne form av en normalfordelingskurve. Engelsen skriver: ”Når man [i] for store grupper av elever har anbefalt en normalfordeling av karakterene deres, er det ut fra tanken om en grupperelatert evaluering.” (Engelsen 2006:121). Scriven beskriver normalfordelingen som ”... an ideal distribution that results in the familiar bell-shaped curve” (Scriven 1991:244). Men hvor ideell er normalfordelingen egentlig? Et eksempel på normalfordeling er når klassens karakterer er jevnt fordelt blant elevene. Lysne beskriver normalfordeling av skolens tidligere karaktersystem slik: Særs godt 4%, Meget godt 24%, Godt 44%, Nokså godt 24% og Lite godt 4% (Lysne 1999:247). Det vil si at uavhengig av hvilket nivå elevene presterer på, vil det alltid være noen som får toppkarakter og noen som får bunnkarakter. Dette skyldes at elevenes prestasjoner sammenlignes med hverandre framfor å sammenlignes med kunnskapsnivået de ligger på. Normalfordeling i evalueringen ble benyttet i Normalplanen av 1939 og videreført inn i Mønsterplanen av 1974 (Lysne 1999:247). At evalueringssystemet fra N39 ble videreført inn i M74 skapte en debatt så heftig at normalfordelingen etter hvert ble avskaffet (Lysne 1999:247).

En tredje evalueringsform er kriterierelatert evaluering. Hva vil det si at evalueringen er kriterierelatert? Gregersen vektlegger kriterier i sin definisjon av evaluering. ”Evaluering er en iagttagelse eller observation baseret, norm- eller kriterierelateret vurdering af pædagogisk relevante faktorer, processer og produkter med henblikk på planlægning, revision, information, rådgivning og/eller kontrol” (Gregersen 1977:261). Kvernbekk gir eksempler på en kriteriebasert evaluering. Kvernbekk skriver: ”Kriterier for vurdering er for eksempel en klart formulert problemstilling, en «rød tråd» gjennom oppgaven, en metode som er i samsvar med problemstillingen, oversikt over og forståelse av temaet, bruk av relevant teori, redelighet i framstillingen, korrekte formalia, selvstendig tenkning og kritisk evne, forståelse for konklusjonenes gyldighetsområde” (Kvernbekk 2005:117). Kriterierelatert evaluering kan ses

i forhold til målrelatert vurdering hvor målene også setter kriteriene for hva som skal evalueres. Engelsen beskriver målrelatert vurdering: ”Målrelatert vurdering kan bidra til at eleven får et realistisk selvbilde. Da vurderer man hvorvidt elevene har nådd de målene som er satt for opplæringen. En slik vurdering krever relativt presise opplæringsmål.” (Engelsen 2006:121). Kriterierelatert evaluering er ment til å legge fokuset utenfor individets eller gruppens ferdigheter og behov, og fokusere på individets kompetanse. Hvilken kompetanse har individet oppnådd? Kriterier for evaluering vil bli drøftet i kapittel 5,6 og 7.

I denne delen har tre standarder for evaluering blitt drøftet: individevaluering, gruppeevaluering og kriterieevaluering. Individevalueringen har et evalueringsgrunnlag som tar utgangspunkt i hva individet kan oppnå. I den norske skole fokuseres det på at vurdering kan benyttes som hjelpemiddel til å tilpasse opplæringen etter individets forutsetninger og behov. En individrelatert evaluering står i kontrast til den grupperelaterte evalueringen. I grupperelatert evaluering blir individet evaluert i forhold til gruppens kompetansenivå. Det vil alltid være noen som blir evaluert til å ligge på topp eller bunn i en slik evalueringsform. Fordi evalueringen skjer i forhold til en gitt gruppes kompetanse, behøver ikke den beste karakteren nødvendigvis bety at eleven har en høy kompetanse innenfor feltet, det betyr bare at han ligger best an i den gruppen han ble evaluert i. Den siste evalueringsstandard som ble tatt opp var kriterierelatert evaluering. Kriterierelatert evaluering fokuserer på prestasjoner. Her spiller verken individets forutsetninger og behov eller gruppens prestasjoner noen rolle. Det er oppnåelsen av kriteriene som teller. I oppgavens videre gang blir det ikke fokusert på gruppeevaluering, da det ikke lenger er tema i den norske skole. Derimot vil individevaluering og kriterieevaluering bli drøftet i kapittel 5,6 og 7.

2.2 Hvordan evaluere?

I dette kapitlet skal spørsmålet ”Hvordan evaluere?” besvares gjennom kvalitativ og kvantitativ metode. Metodebegrepet er et stort begrep som dekker langt mer enn bare evaluering. I vitenskapsteorien kan metode forstås som ”en rasjonell aktivitet, sentrert rundt

framlegging av argumenter og evidens” (Kvernbekk 2002:22) Michal Scriven ser på evaluering som en metodisk aktivitet. Scriven skriver:

”a methodological activity which is essentially similar whether we are trying to evaluate coffee machines or teaching machines, plans for a house or plans for a curriculum. The activity consists simply in the gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings, and in the justification of (a) the data-gathering instruments, (b) the weightings, and (c) the selection of goals (Scriven 1967:40).

Den metodiske aktiviteten kommer til uttrykk igjennom kvalitative og kvantitative evalueringsmetoder. Eksempler på hvordan evaluering manifesteres kan være: karakterer, utregninger av en prosentandel, skriftlige tilbakemeldinger, muntlige tilbakemeldinger osv. (se for eksempel: Almås 1993, Lund 2002, Mehlbye m.fl. 1993, Tveit 2007). Både kvalitative og kvantitative metoder representerer et perspektiv på hvordan evaluere. Mehlbye skiller mellom metoder for datainnsamling og analyse og fortolkning (Mehlbye m.fl. 1993). Selv om metode for datainnsamling er av betydning for analysen og fortolkningen, er det analyse og fortolkning som er i fokus for denne oppgavens tolkning av evalueringsbegrepet. I argumentasjonen nedenfor kommer det av hensyn til metodebegrepets bredde til å bli tatt hensyn til både metoden for måling av resultater og selve evalueringen.

2.2.1 Kvantitativ metode i evalueringen

Spørsmålet om hvordan evaluere kan svares på ved hjelp av kvantitativ metode. Kvantitativ metode handler om ”*innsamling og bearbeiding av opplysningar etter visse reglar eller kriterium som gjer det mogleg å talfeste forekomsten av ulike fenomen*” (Almås 1993:49). Det kan her skilles mellom hva som er innsamling og bearbeiding av opplysninger (etter visse regler) og tallfesting av forekomsten eller resultatene. Evalueringen kommer i kvantitativ metode til uttrykk gjennom tallfestingen. Tallet er en såkalt deskriptiv evaluering (Scriven 1991:123), som prosentvis beskriver resultatene i en undersøkelse. I skolen gjøres tallet som regel om til en karakter. Karakterer kan i teorien være både kvalitative og kvantitative (se for eksempel Kvaale 1970, Scriven 1991). I dette delkapittelet blir karakterer kun belyst gjennom kvantitativ metode. Karakterer som kvalitativ evalueringsmetode vil bli tatt opp i del 2.2.2 og 5.1.2. Fordi denne oppgaven er rettet mot vurdering av elever, avgrensers jeg meg til å drøfte

forholdet mellom variabler (indre validitet) og overføring av variabler (ytre validitet) i kvantitative undersøkelser.

For at evaluering kvantitativt skal komme til uttrykk som et tall eller karakter, må også metoden i undersøkelsen tilpasses dette. Et eksempel på en kvantitativ metode som kan gi grunnlag for en tallfestet karakter er *multiple choice* tester. I *multiple choice* testene regnes karakteren ut ifra hvor stor prosentandel eleven har svart riktig på testen. Vi skal videre se på *multiple choice* som et eksempel på hvordan kvantitativ metode kan gi grunnlag for en tallfestet evaluering.

Scriven beskriver *multiple choice* som: "A test – or a question in a test – in which each question (the 'stem') is followed by several alternative answers, one and only one of which is to be selected as the correct or nearest to correct answer" (Scriven 1991:236). Det er delte meninger rundt hvilket utbytte elever og studenter har av *multiple choice* testing. Universitetet i Victoria i Canada, skriver følgende om *multiple choice*: "Multiple choice questions, (...) require fine distinctions between correct and nearly-correct statements" (University of Victoria). I følge Universitetet i Victoria er *multiple choice* testing et redskap til å trene studenten i et høyere nivå for tenking. Med høyere nivå for tenking mener Universitetet i Victoria: "(...) distinctions that involve the thinking for **Synthesis, Analysis, and Application**" (University of Victoria). Ikke alle er enige med Universitetet i Victoria sitt perspektiv på *multiple choice* testing. Caroline Gipps og Patricia Murphy stiller spørsmål rundt validiteten dens (Gipps m fl. 1994:112). Validitet blir her forstått som høy grad av sikkerhet (Lund 2002:85). Gipps og Patricia hevder å ha funnet forskningsbaserte resultater på at *multiple choice* favoriserer gutter framfor jenter og derfor ikke er til å stole på (Gipps m fl. 1994:112). Evalueringen ved *multiple choice* tester er avhengig av at testmetoden er til å stole på. Dersom det er sikkert at *multiple choice* viser seg og favoriserer gutter, er dette en lite sikker evaluering metode å bruke i undersøkelser der gutter og jenter blir satt i samme bås. Ytterligere kritikk av *multiple choice* tester er blant annet at de kan være for enkle for eleven eller for studenten (Bjørkeng 2008) og at *multiple choice* ofte er en økonomisk løsning i trange tider (se for eksempel: Gipps 1994, eller Aalen 2008). Dette er momenter jeg ikke ønsker å gå videre inn på.

Multiple choice legger som nevnt grunnlag for en kvantitativ evaluering ved at spørsmålene som stilles har rette og gale svar. Den som utfører testen belønnes som regel med poeng for

hvert riktig svar (Scriven 1991:236). Poengene blir så gjort om til en karakter eller prosentandel som viser hvor mye eleven har gjort rett og galt. Scriven skriver ”The advantage is ease and speed of correcting (e.g., machine-scoring is possible), and the reduction of the judgmental error involved in correcting essay answers” (Scriven 1991:236). Fordi svarene i en *multiple choice* test er rette eller gale, er grunnlaget for evalueringen sikrere enn i tester hvor svaret ikke er gitt. Med andre ord, når en elev får karakteren 6 på en *multiple choice* test, er det fordi eleven har svart alt riktig. I en kvalitativ metode vil karakteren 6 bero på andre kriterier enn i den kvantitative som fokuserer på om svaret er riktig eller galt.

Oppsummert kan en si at kvantitativ metode i skolen gir grunnlag for en tallfestet karakter. Vi har ovenfor diskutert positive og negative sider ved *multiple choice* som evalueringsmetode i skolen. Det ble vist at *multiple choice* på den ene side kan trigge refleksjon hos elever dersom spørsmålene og svaralternativene er utfordrende for studenten. På den andre siden kan *multiple choice* trues ved at spørsmål og svar er for enkle, at testen favoriserer kjønn, eller at testen kun er en økonomisk løsning. Evalueringen ved slike tester lar seg true av forholdene innad i testen. For øvrig gir evalueringen ved en *multiple choice* test en beskrivelse av hvor stor prosentandel eleven har svart riktig og galt. Som Scriven skriver er denne formen for evaluering både kjapp og lett og rette.

2.2.2 Kvalitativ metode i evalueringen

Kvalitativ metode er et alternativ til kvantitativ metode for hvordan evaluering kan foregå. Kvalitative metoder kan beskrives som ”*innsamling og bearbeiding av opplysningar med sikte på å få fram intensjonar og mønster i sosiale samanhengar*” (Almås 1993:57). Den kvalitative metoden har til hensikt å gi en evaluering utover det deskriptive. Scriven skriver om kvalitativ metode at det er: ”The part of evaluation that can’t be usefully reduced to quantitative measures. A substantial part of good evaluation (of personnel and products as well as programs) is wholly or chiefly qualitative, meaning that description and interpretation make up all or most of it (Scriven 1991:293). Skal en tro Scriven må kvalitativ metode til for å skape en god evaluering. Hvorfor er dette tilfellet? Den kvalitative metoden gir noe mer til evalueringen enn det den kvantitative metoden gjør. I følge Elliot W. Eisner har kvalitative metoder en tendens til å ”emphasize the importance of context in understanding, they tend to place great emphasis on the historical conditions within which events and situations occur,

and they tend to argue that pieces cannot be understood aside from their relationship to the whole in which they participate” (Eisner 1985:136). Med andre ord ses kvalitativ evalueringsmetode i forhold til konteksten den befinner seg i. Dette er også et kjennetegn ved den hermeneutiske sirkelen (se for eksempel: Ramberg 2005). En kvalitativ evaluering handler ikke lenger om kun å beskrive tingenes tilstand, men også forklare hvorfor de er slik og hvordan de kan endre seg.

Kvalitativ metode kan deles opp i flere metoder. Noen eksempler på kvalitative metoder er observasjon, for eksempel i klasserommet (se: Klette:1998), fritt formulerte spørreskjema (se: Kleven 2002), og samtalen (strukturer/ustrukturert) (se: Almås 1993). Det er i kvalitativ metode viktig å skille mellom hva som er metode for datainnsamling og hva som er metode for evaluering (Kleven 2002). Det som kjennetegner evalueringen i kvalitativ metode er som Mehlbye skriver: ”at resultatene af en evaluering begrundes i tekst og ikke tal” (Mehlbye m.fl 1993:132). Når jeg i kapittelet om kvantitativ metode hevdet at karakterer (som ofte kommer i tallform) også kan være kvalitative, var dette med tanke på betydningen bak karakteren. For eksempel uttrykker karakteren 5 ”at eleven har mykje god kompetanse i faget” (Forskrift til opplæringsloven § 3-4). Denne betydningen er kvalitativ og kan gis ved å evaluere kvalitative data. Med andre ord kan eleven ut ifra metoder som observasjon, spørreskjemaer og gjennom samtaler bli evaluert til en kvalitativ, tallfestet karakter. I min videre argumentasjon vil jeg gå nærmere inn på samtalen som grunnlag for evaluering.

I kvalitativ metode er samtalen et eksempel på hvordan data kan samles inn (Kleven 2002:73) og samtidig evalueres. Det er allikevel ikke alle samtaler som fungerer som evaluering. Astrid B. Eggen skriver i forbindelse med elevsamtalen at ”Hvorvidt de skal kunne betegnes som vurderingsredskaper er avhengig av vurderingselementet i samtalene” (Eggen 2009:65). Dette gjelder også samtaler generelt. Samtale eller dialog stammer fra gresk og kan deles inn i ”dia” og ”logos”. ”Dia” betyr ”gjennom” og ”logos” betyr ”ord” (Valvatne m.fl 2002:126). Altså ”gjennom ordet”. Samtalen kan i følge Valvatne ”beskrives ved at partene veksler på å ha ordet, de veksler mellom de to dialogrollene taler og lytter” (Valvatne m.fl 2002:126). Når samtalen skal ses på i en evalueringskontekst, er det først og fremst som tilbakemelding hvor begge parter får ytre sine meninger om et eller flere temaer. I følge Black og William kan samtaler som evaluering fremme læring (Black m.fl. 1998:28). Taler- og lytterrollen er sentrale aspekter i samtalen som bør bygge på en felles forståelse rundt et gitt tema. Dersom

forståelsen mellom partene i samtalen ikke er tilstede, kan dette anses å være en trussel i den kvalitative evalueringsmetoden (Kleven 2002).

Kvalitative metoder ligger til grunn for mange av skolens evalueringer, også karakterer. De tre metodene observasjon, spørreskjema og samtalen er blitt nevnt. Både observasjon og spørreskjema kan gi lærer et grunnlag for å evaluere elever, men er ikke kvalitative metoder for evaluering i seg selv. Evalueringer ut ifra observasjon kan finne sted enten som en karakter, et skriftlig dokument eller en samtale. Samtalen kan både ses på som metode for datainnsamling og metode for evaluering. Gjennom samtalen kan lærer gi elever kvalitative evalueringer som i beste fall fremmer elevens læringsprosess.

Kort oppsummert har spørsmålet om hvordan evaluere i dette delkapittelet blitt diskutert gjennom kvantitativ og kvalitativ metode. I både kvantitativ og kvalitativ metode er det blitt skilt mellom metoder for datainnsamling og metode for evaluering. I den kvantitative evalueringsmetoden kommer evalueringen til uttrykk som et tall. Jeg valgte å fokusere på tall slik de kommer til uttrykk etter kvantitative prøveformer i skolen. Disse tallene representerer som regel en prosentandel for hvor mye rett og galt eleven har svart på prøven. Den kvalitative metoden står i kontrast til den kvantitative ved at den enten kommer som en skrevet tekst eller i form av en samtale. Jeg har her ikke drøftet validitet i forhold til kvantitative og kvalitative prøveformer. Jeg kommer tilbake til spørsmålet om validitet og gyldighet i kvalitative og kvantitative metode når jeg i kapittel fem drøfter en rettferdig elevvurdering.

2.3 Evaluering og nivåer for læreplantolkning

I dette siste delkapittelet vil evaluering og nivåer for læreplantolkning drøftes. Kapittelet er ment som en innsnevring fra evalueringsfeltet som en helhet og til evaluering i skolen. Jeg skal videre fokusere på hvordan evaluering kan vise seg på ulike nivåer for læreplantolkning. Hensikten med dette er todelt. For det første viser nivåene at evaluering kan spesifiseres til ulike områder i skolesystemet, og for det andre legger nivåene for læreplantolkning en forståelse for evalueringens prosess fra teori til praksis. For å vise hvordan nivåene for

evaluering både skiller seg fra hverandre og jobber sammen benytter jeg John Goodlads modell om nivåer for læreplantolkning.

Goodlad opererer med fem nivåer for læreplantolkning. De fem nivåene representerer et spenn fra teori til praksis. Engelsen har valgt å kalle nivåene for: ideenes læreplan, den formelle læreplan, den oppfattede læreplan, den operasjonaliserte læreplan og den erfarte læreplan (Engelsen 2006:28). Evaluering fungerer i nivåene for læreplantolkning både som et redskap til å utforme og tolke styringsdokumentet. I tillegg kan nivåene for læreplantolkning brukes til å belyse hvordan evaluering som redskap i skolen kommer til uttrykk på ulike nivåer. Det er det siste momentet jeg her skal belyse.

Det første nivået til Goodlad tar for seg ideene bak skolens utforming (Engelsen 2006:28). Hvilke tanker og verdier ligger bak læreplanen? "Ideene bak læreplanen" er et overordnet nivå som sier noe om hvilke evalueringer som ligger til grunn for selve dokumentet (Engelsen 2006:28). Evalueringer av skolesystemet er med på å forme nye ideer til læreplanen. Eksempler på evalueringer som ble gjort i forkant av Læreplanverket for kunnskapsløftet er blant annet PISA (PISA 2003) TIMSS (Grønmo m. flere 2003) og PIRLS (PIRLS 2001). Evalueringene av PISA, TIMSS og PIRLS har vært av betydning for skolens og samfunnets ønske om et kunnskapsløft. Ønsket om et kunnskapsløft og evalueringene av PISA, TIMSS og PIRLS kan derfor sies å ha vært med på å danne ideen bak Læreplanverket for kunnskapsløftet.

Hvilke ideer ligger så bak Læreplanverket for kunnskapsløftet? Svarene på dette er flerfoldige. Den ideen det blir fokusert på i denne oppgaven, er ideen om en bedret evalueringspraksis. Ideen er at dersom evalueringspraksisen bedres, vil elevene få et kunnskapsløft. Heftet "Bedre vurderingspraksis" er et resultat av dette (Bedre vurderingspraksis 2007). Ideen om at evaluering kan bedre elevenes kompetanse er en forutsetning for de fire underliggende nivåene i Goodlads modell. Spørsmålet er om ideen bak Læreplanverket for kunnskapsløftet og dets retningslinjer for evaluering stemmer overens med hva som skjer i virkeligheten? Goodlad selv er kritisk til nivået for ideenes læreplan. Han skriver "Although good ideational work anticipates the problems of adoption and use, products usually are designed to serve a varied marketplace of decisions and actions and stop short of the give and take of sociopolitical processes" (Goodlad 1979:60). Spriket mellom ideen bak læreplanen og hva som faktisk står i læreplanen kan altså diskuteres. Ideene vil

allikevel være med på å forme hva som står i den formelle læreplanen. Den formelle læreplanen bringer oss videre til det andre nivået Goodlad presenterer.

Det andre nivået Goodlad legger fram er det Engelsen referer til som ”den formelle læreplanen” (Engelsen 2006:28). Den formelle læreplanen handler om selve styringsdokumentene. Eksempler på slike dokumenter i Norge er Læreplanen for kunnskapsløftet 2006, Stortingsmelding 16 og 30 og opplæringsloven. Goodlad beskriver den formelle læreplanen som:”

Formal curricula are those which gain official approval by state and local school boards and adoption, by choice or fiat, by an institution and/or teachers. For such approval to be secured or granted, there must almost necessarily be some sort of written documents: curriculum guides, state or local syllabi, adopted texts, units of study set forth by a curriculum committee, etc” (Goodlad 1979:61).

I den formelle læreplanen handler det ikke lenger om evalueringene bak styringsdokumentet, men om et offentlig dokument, Læreplanverket for kunnskapsløftet, med retningslinjer for evaluering i skolen. Et eksempel på hvordan evaluering formelt kommer til uttrykk i læreplanen er når det i Læreplanverket for kunnskapsløftet står skrevet at: ”Vurdering og veiledning skal bidra til å styrke deres [elevenes] motivasjon for læring” (LK06:33). Dette sitatet sier noe om vurdering som middel til å fremme elevenes motivasjon for læring, men det sier lite om hvordan lærer skal vurdere eleven. Dette må lærer selv tolke seg fram til. Dette bringer oss over på Goodlads tredje nivå for læreplantolkning.

Det tredje nivået til Goodlad beskriver Engelsen som ”den oppfattede læreplanen” (Engelsen 2006:28). Engelsen skriver om den oppfattede læreplanen: ”Når lærere (og andre) leser læreplandokumentet, tolker de råd og retningslinjer i planen. Det er denne tolkningen som blir utgangspunktet for deres planlegging, tilrettelegging, gjennomføring og vurdering av opplæringen” (Engelsen 2006:28). Det handler her om hvordan lærere evaluerer og oppfatter læreplandokumentene. Goodlad skriver om *perceived curricula* :”Perceived curricula are curricula of the mind. What has been officially approved for instruction and learning is not necessarily what various interested persons and groups perceive in their minds to be the curriculum” (Goodlad 1979:62). Med andre ord betyr dette at evaluering slik det står uttrykt i læreplanen eller at ideen bak hvordan lærere skal evaluere potensielt kan tolkes ulikt i skolen. Ulike tolkninger av retningslinjer for evaluering er med på å skape forskjeller i evalueringspraksisen. Dette kan igjen føre til at samme produkt blir evaluert ulikt av ulike

lærere. Dette bringer oss over på det fjerde nivået til Goodlad. Hvordan utføres læreplanen i praksis?

Det er til nå blitt vist at ideen bak læreplanen, selve læreplandokumentet og lærernes tolkninger av læreplandokumentet ikke nødvendigvis vil stemme overens. Det er allikevel slik at ideene, utformingen og tolkningen av evaluering i læreplandokumentet vil få konsekvenser når lærer praktiserer evaluering i opplæringen. Videre vil det skilles mellom hvordan evaluering blir operasjonalisert av eleven og hvilke erfaringer eleven sitter igjen med fra evalueringene.

Det fjerde nivået til Goodlad er hvordan de formelle retningslinjene operasjonaliseres i praksis (Engelsen 2006:28). Goodlad skriver: "What teachers perceive the curriculum of their classrooms to be and what they actually are teaching may be quite different things"(Goodlad 1979:62). Hvordan påvirker læreres tolkninger av evalueringens retningslinjer deres handling i timen? Gunn Imsen begrunner hvorfor læreres tolkninger og lærerens praksis ikke nødvendigvis alltid stemmer overens. Hun skriver "Undervisning går aldri helt etter oppskrifta, ganske enkelt fordi det er så mye som virker inn når en står overfor en sprell levende elevflokk" (Imsen 1999:174). Fordi elever er mer eller mindre forutsigbare, bør også evalueringer av elevene kunne preges av uforutsigbarhet. Som det ble vist i del 2.1.1 kan evaluering både være formell og uformell. Mens den formelle evalueringen gir rom for planlegging og struktur, er den uformelle evalueringen mer spontan og kan benyttes i uforutsette situasjoner som i klasserommet. Slik kan evaluering operasjonaliseres i opplæringen både formelt og uformelt. Det fjerde nivået til Goodlad kan oppsummeres med et sitat Imsen har hentet fra Peer Gynt "Tenke det, ønske det, ville det med, - men gjøre det! Nei det skjønner jeg ikke!" (Peer Gynt, sitert i Imsen 1999:174).

Det femte nivået til Goodlad beskriver Engelsen som den erfarte læreplanen. Dette handler om lærerens og elevens erfaringer fra opplæringen (Engelsen 2006:28). Goodlad beskriver det som det *experiential Curricula* (Goodlad 1979:63). Sett i et evalueringsperspektiv er den erfarte læreplanen et spørsmål om hvilke erfaringer lærer og elev har gjort seg etter gjennomført evalueringen. Hvordan har evalueringer av skolesystemet og evalueringer av eleven påvirket lærer og elev? Hvilke erfaringer sitter lærer og elev igjen med? Goodlad tar opp utfordringene i forskning på dette nivået. Goodlad skriver: "Watching students tells us

very little of what is going on in their minds. Asking them raises many questions of validity. Dare they trust us to care seriously about their answers? Dare we trust their answers?" (Goodlad 1979:63-64). Problematiseringen av hvordan den erfarte læreplanen kan måles, viser at det er vanskelig å oppfatte elevens erfaringer av skolens og lærerens evalueringer. Som Imsen skriver "Teoretisk vil en kunne få like mange svar på dette som det finnes elever i klassen" (Imsen 1999:175).

Det er gjennom Goodlads fem nivåer for læreplan nå blitt vist hvordan evaluering kan komme til uttrykk på ulike nivåer i skolesystemet. Det er ideen bak evaluering, det er evaluering slik den blir uttrykt i det formelle læreplanverket, det er lærerens tolkning av evaluering i det formelle læreplanverket, det er hvordan lærer praktiserer sin tolkning av evaluering fra de formelle dokumentene og til slutt – det er erfaringene lærer og elev sitter igjen med etter evalueringen har blitt gjennomført i skole og klasserom.

2.4 Oppsummering og konklusjoner

I dette kapitlet har det blitt satt fokus på ulike perspektiver ved evalueringsfeltet. Det har vist seg at evaluering er et stort felt som eksisterer og benyttes både i privat og offentlig kontekst. Jeg begynte kapitlet med å stille spørsmålet: Hva er evaluering? Å svare konkret på hva evaluering er, har vist seg som en utfordring. Det er først og fremst utfordrende fordi evalueringsbegrepet gjerne har ulik betydning i ulike kontekster. For eksempel avgrenses evalueringsbegrepet på systemnivå og så ofte deretter. Det ble tatt opp flere forslag til hvordan evaluering blir definert på systemnivå, dette er imidlertid ikke perspektiver jeg ønsket å ta med meg videre i oppgaven. Etter å ha drøftet ulike definisjoner av evalueringsbegrepet, endte jeg opp med en definisjon fra Michal Scriven: "Evaluation is the process of determining the merit, worth and value of things, and evaluations are the products of that process" (Scriven 1991:1). Dette perspektivet forstår evaluering som verdisetting både under prosessen og etter produktet er ferdig. Videre ble spørsmålet om hva evaluering er, belyst i forhold til begrepsmessige distinksjoner og ulike standarder for evaluering. De

begrepsmessige distinksjonene tok for seg perspektiver på subjektivitet og objektivitet, formell og uformell, formativ og summativ og til slutt intern og ekstern evaluering. Dette er alle distinksjoner som er med på å beskrive evalueringsfeltet, og de vil bli brukt videre i diskusjonen om elevvurdering. De ulike standardene tok for seg tre eksempler på ulike måleenheter i evalueringen. Det ble vist at en evaluering kan måles i forhold til individuelle ferdigheter og behov, den kan måles i forhold til en gruppe – en såkalt normalfordeling, og evaluering kan også måles i forhold til gitte kriterier. Dette er også viktige momenter som vil bli drøftet videre i oppgaven.

Del 2.2 stilte spørsmålet: Hvordan evaluere? I dette kapittelet ble det lagt vekt på metode for evalueringen. Dette ble analysert gjennom kvantitative og kvalitative metoder. I kvantitativ metode viste det seg at evaluering kommer til uttrykk gjennom tall. Dersom en skal se på kvantitativ metode i elevvurderingen kommer den til uttrykk gjennom karakterer. Karakterer er for øvrig ikke bare et eksempel på hvordan kvantitativ metode kan komme til uttrykk, men også kvalitativ evalueringsmetode. Den kvalitative evalueringsmetoden kommer gjerne til uttrykk igjennom tekst eller tale. Årsaken til at en karakter kan bli ansett som kvalitativ er meningen bak karakteren. Jeg kommer tilbake til kvantitativ og kvalitativ metode i kapittel 5.

I den siste delen av kapittelet tok jeg for meg evaluering på ulike nivåer for læreplantolkning. Her ble evalueringsbegrepet avgrenset i forhold til skolen. Det ble vist hvordan evaluering kommer til uttrykk på ulike nivåer og stadier i skolesystemet. Det startet med ideen bak evaluering, gikk ned til hva som faktisk står i læreplanen om evaluering, hvordan lærer tolker det som står i læreplanen, hvordan lærer praktiserer sine tolkninger av evalueringen og til slutt hvilken erfaring eleven sitter igjen med etter evalueringen. Det ble vist at det gjerne er et stort sprik mellom ideen bak evaluering og hva som skjer i praksis. Når vi i neste kapittel skal se nærmere på hva elevvurdering er, er det viktig å ta hensyn til dette skillet mellom tanken bak vurderingen og hva lærer og elev sitter igjen med.

3. Elevvurdering

I dette kapitlet vil det bli tatt en avgrensning fra evaluering til elevvurdering. Elevvurderingen har mange utfordringer den møter. Dette kapitlet har som hensikt å peke på en del av disse utfordringene og problematisere dem i forhold til den norske skolen.

Spørsmål om hva elevvurdering er, hvorfor det er et aktuelt tema og hvilken rolle elevvurderingen har blir henholdsvis tatt opp i kapittel 4.1, 4.2 og 4.3. Helt til slutt i kapitlet kommer en oppsummering av kapitlets poenger.

3.1 Hva er elevvurdering?

Evaluering blir i denne oppgaven behandlet som feltet elevvurderingen ligger i. Slik blir elevvurdering kun en bit av det store evalueringsfeltet. Det er mulig å argumentere for at evalueringsteori også er relevant i elevvurderingen. Elevvurdering kan vinkles fra tre perspektiver innenfor en klasseromskontekst. Det er lærerens vurdering av eleven (Imsen 1999), det er elevenes vurdering av andre elever og det er elevens vurdering av seg selv (Engh 2007:107). I tillegg vil miljø og samfunn spille inn på elevvurderingen. I dette avsnittet presenteres først elevvurderingsbegrepet som en spesialisering av evalueringsbegrepet. Deretter vil metoder for elevvurdering bli tatt opp.

Terje Ogden bruker evalueringsbegrepet når han skriver om elevvurdering. ”Evaluering bør handle om i hvilken grad elevene lærer det som de skal lære, ifølge læreplanen og lærersynspunkter (Ogden 2005:17). Et annet perspektiv på elevvurdering er i følge Lauvås ”å sammenholde noe forekommende med en eller annen målestokk (standarder, normer, krav, verdier) med tanke på å komme fram til en eller annen konklusjon – Det kan dreie seg om å konkludere med at noe er godt nok (bestått, godkjent, antatt)”(Lauvås 2007:56). Ogden og Lauvås sine perspektiver retter seg her mot elevens sluttprodukt. Hvorvidt målet er oppnådd. Lyngsnes og Rismark vil si seg uenig i dette. De skriver: ”Skal læring og undervisning

forbedres, er det ikke tilstrekkelig bare å måle resultatene i etterkant, en må også gå inn i prosessen underveis for å finne ut hva som fungerer bra og mindre bra” (Lyngsnes og Rismark 2007:118).

Elevvurdering er viktig av to årsaker. For det første er det viktig å kunne gi et speilbilde på hva eleven kan. For det andre så er elevvurdering viktig for å kunne veilede elevens arbeidsprosess og føre til videre læring.

Elevvurdering er et komplekst fenomen som preges av en rekke forhold ved opplæringen. Britt Ulstrup Engelsen påpeker i forhold til modellen didaktisk relasjonstenkning, at elevvurdering preges ikke bare av fenomener ved vurderingen, som formativ og summativ vurdering, men også av øvrige didaktiske kategorier som mål, innhold, arbeidsmetode, rammer, elev (Engelsen 2006:46) og motivasjon. Endringer i disse forholdene vil prege elevvurderingen og lærerens rolle. Erling Lars Dale argumenterer for økning av lærernes vurderingskompetanse. Han hevder at deres kompetanse om elevvurdering vil endres i samsvar med endring i fagstoff, metoder osv. Han skriver: ”Lærerne vil alltid måtte bli involvert dersom forslag om forandringer krever at de trenger å forandre sine kunnskaper og ferdigheter i det stoffet de underviser i og måten de underviser på (Dale & Wærnes 2006:140). I og med at informasjonen rundt elevvurdering endrer seg hyppig, er det god grunn til å tro at lærere bør få mer kompetanse innenfor området.

I følge Georg Matthiesen består hele vurderingsarbeidet av fem punkter som til sammen utgjør en prosess. Det første elementet er et spørsmål om hvilke mål eleven vurderes opp i mot. I kunnskapsløftet settes disse målene igjennom kompetansemål. Det andre er en ”beskrivelse av en arbeidssituasjon der elevene skal vise og dokumentere hva de kan og mestrer” (Matthiesen 2007:49). Her kan en trekke linjer til Piaget sin stadieteori og hva eleven mestrer på ulike nivåer (Miller 2002:38). Det tredje punktet til Matthiesen er kartlegging av elevenes dokumenterte læring. Dette bringer med seg noen utfordringer. Hvordan dokumentere hva eleven har lært på en valid måte? Det fjerde punktet er vurdering på grunnlag av kartleggingen. Dette tar utgangspunkt i nasjonalt fastsatte krav og individuelle forutsetninger. Det femte og siste punktet er tilbakemeldingen. Tilbakemeldingen skal skje ”i form av nasjonalt fastsatte vurderingsuttrykk (karakterer) eller i form av individuelt tilpassede beskrivelser” (Matthiesen 2007:50).

I det femte leddet legger Matthiesen vekt på at evaluering kan gis som både formativ og summative tilbakemeldinger. Jeg vil allikevel si meg uenig i at dette er hele vurderingsarbeidet. Hva skjer med lærerens vurderingsarbeid når det kommer fram til eleven? Blir tilbakemeldingen tatt til betraktning eller legges den i glemmeboka? Bransford et. Al vektlegger at:

”The use for frequent formative assessment helps make students’ thinking visible to themselves, their peers, and their teacher. This provides feedback that can guide modification and refinement in thinking. Given the goal of learning with understanding, assessments must tap understanding rather than merely the ability to repeat facts or perform isolated skills” (Bransford et. al.2004:19).

Mens Bransford vektlegger læringsutbyttet eleven har av vurderingen, vektlegger Deci og Ryan motivasjon. Deci og Ryan skriver om hvordan ytre belønninger som vurdering, kan påvirke indre motivasjon (1999). Der Matthiesen sin vurderingsarbeid slutter, begynner Deci og Ryan sin teori om hvordan belønning påvirker indre motivasjon hos eleven. Et spørsmål som kan stilles er hvilken motivasjon vurderingen gir eleven? Hvordan påvirker ytre vurderinger individets motivasjon og vurdering av seg selv?

Deci og Ryan skiller mellom vurderinger i form av ”Tangible rewards” og ”positive feedback”. Slik jeg ser ”Tangible rewards” handler det om ytre belønning som ikke er direkte knyttet opp mot området som mestres. For eksempel hvis jeg sier at dersom du får A på eksamen gir jeg deg en tur til Hellas. Positiv feedback på den andre siden er en tilbakemelding på det du har gjort. Noe du kan jobbe videre med. I en undersøkelse av Deci og Ryan kom de fram til at: ”Positive feedback enhanced both free-choice behavior and self-reported interests. Tangible rewards tended to be more detrimental for children than college students, and verbal rewards tended to be less enhancing for children than college students” (Deci et al. 1999:627). Både Deci og Bransford legger fokuset på hva eleven sitter igjen med etter elevvurderingen har funnet sted. Slik sett går vurderingstematikken fra den didaktiske disiplin og over til den psykologiske. Begge disiplinene er viktige å ha med når elevvurdering skal bedømmes. Dette utgjør også en utfordring for læreren, som bør ha kompetanse både rundt generelle didaktiske prinsipper og elevens individuelle ferdigheter.

Oppsummert ligger elevvurdering i et grensesnitt mellom disiplinene psykologi og didaktikk. Elevvurderingen har en beskrivende funksjon og en veiledende funksjon. Definisjonen som vil bli brukt videre i oppgaven er derfor denne: *Elevvurdering kan forstås som et redskap, der har til hensikt å beskrive og/eller utvikle elevens læringsprosess.* Jeg skal videre gå inn å se på

elevvurderingens beskrivende og utviklende funksjon. Dette skjer igjennom begrepene formativ og summativ elevvurdering.

3.1.1 Formativ og summativ elevvurdering

I kapittel to ble formativ og summativ evaluering beskrevet. I dette kapittelet skal formativ og summativ elevvurdering drøftes. Hva har formativ elevvurdering å si for elevens læringsprosess? Hva er hensikten med summativ elevvurdering? Hvordan kan begrepene ses i forhold til hverandre?

Det har blitt lagt ulike betydninger til begrepet formativ vurdering (Gregersen 1977:263). Her har formativ elevvurdering til hensikt å: ”provide feedback and to facilitate reflection on learning by the student” (Nixon 1990:108). Formativ vurdering finner som oftest sted i prosessen (Scriven 1991:168), men kan også komme i form av sluttvurdering. Eksempler på formativ elevvurdering er tilbakemeldinger med og uten karakter, egenvurdering, elevsamtale, samtaler med foreldre, diagnostiske prøver, kartleggingsprøver, mappevurdering osv. (St.meld.nr.30:37-39). Den formative elevvurderingen kan skje ved hjelp av interne og eksterne sensorer (Scriven 1991:169). I tillegg til disse kan eleven selv gjøre en formativ vurdering av læringsprosessen (Nixon 1990:108).

Summativ elevvurdering står i kontrast til formativ (Scriven 1991:168). Med summativ elevvurdering forstås det: ”provide a profile of competencies of the student” (Nixon 1990:108). Scriven poengterer også at den summative vurderingen skjer etter prosessen og ikke i prosessen (Scriven 1991:340). Eksempler på summativ vurdering kan være standpunktvurdering og eksamensvurdering (St. Melding Nr 30: 37-39). I likhet med formativ vurdering kan summativ vurdering være intern og ekstern (Nixon 1990:108).

Hva har formativ og summativ elevvurdering å si for elevens læringsprosess? I L97 het det: ”Hovedformålet med vurdering, er å fremje læring og utvikling” (Sitert i St.meld.30:37). Et spørsmål er til hvilken grad en summativ elevvurdering er i stand til å gjøre dette? Et annet spørsmål er om kvaliteten på den formative elevvurderingen holder mål?

Det er mulig å anta at både summativ og formativ elevvurdering har en nytteverdi. Summativ vurdering har til hensikt å gi en summerende vurdering av elevens måloppnåelse ut fra

læreplanen. Dette kan for eksempel skje igjennom en karakter. Spørsmålet er hva betyr denne karakteren? Hva legger en lærer i karakteren "B"? Ville eleven fått samme karakter dersom det var en annen lærer som satt karakteren?

En annen utfordring ved den summative vurderingen er om den på noen måte kan virke læringsfremmende? Paul Black m. fl. skriver: "Feedback by grades focuses students' attention on their 'ability' rather than on the importance of effort, damaging the self-esteem of low attainers" (Black m.fler 2003:46). Per Lauvås støtter opp rundt Black sin påstand. Lauvås skriver "Det er min påstand at vi bare lurer oss selv når vi holder myten om at karakterer er feedback til eleven ved like"(Lauvås 2007:64). Det er allikevel ikke slik at den summative vurderingen ikke har noen hensikt. Den summative vurderingen har en funksjon når eleven skal søke seg inn på blant annet universitet og høyskoler. Den summative vurderingen gir muligheter for kvantifisering av kunnskapen og et tallsnitt på elevens forståelse.

Til motsetning fra den summative vurderingen har formativ vurdering først og fremst hensikt å veilede i prosessen. Paul Black gjør her et skille mellom vurdering for læring og vurdering av læring. (Black m.fler2003:2). Det er den formative vurderingens arbeidsoppgave å fremme læring. Til hvilken grad kan lærer og elev samsvare i forståelsen rundt den formative vurderingen? Som Cedric Cullingford påpeker kan formativ vurdering like gjerne virke mot sin hensikt (Cullingford 1997:1). Hvorfor er det slik?

Det finnes en rekke teorier som kan forklare hvorfor formativ vurdering ikke nødvendigvis alltid fungerer. Et eksempel er Basil Bernsteins kommunikasjonskoder. Bernstein hevder at at "Hvis barnet er modtageligt for skolens kommunikationssystem og dermed for dens indkæring og relationer, så vil dette barns skoleerfaringer betyde en symbolsk og social udvikling. Hvis barnet derimod ikke er modtageligt for skolens kommunikationssystem, så vil dette barns skoleerfaringer kendetenes ved en symbolsk og social forandring" (Bernstein 1974:35). Hvordan den formative vurderingen blir uttrykt kan altså være et spørsmål om elevens forståelsesramme. Forstår eleven hva som blir sagt i tilbakemeldingen? Ulikheter i språkbruk kommer fra barnets sosiale identitet hevder Bernstein (1974:35). Hvilken bakgrunn barnet har, kan altså være av betydning for hvordan den formative vurderingen tolkes. Bernstein skiller videre mellom en *elaborative* og restringert språkkode. Med *elaborative* språkkode mener han at den talende vil "vælge mellom et stort udvalg af syntaktiske alternativer, der vil blive sammensat på en fleksibel måde"(Bernstein 1974:37). I kontrast står

den restringerte språkkoden. Bernstein skriver: ”Omvendt er det ved en restringeret kode langt nemmere at forudsige de syntaktiske valg eller alternativer du fra et representativt udsnit, idet de her hentes fra et begrænset udvalg” (Bernstein 1974:37). Hvilken språkkode barnet har, tilsvarer deres forutsetninger for å tolke formative tilbakemeldinger. En utfordring for læreren blir å bruke et språk som eleven forstår når han gir tilbakemeldinger.

Derimot, med riktig bruk kan formativ vurdering få funksjon som et stillas rundt eleven. Jerome Bruner m. fler skriver: “This scaffolding consists essentially of the adult "controlling" those elements of the task that are initially beyond the learner's capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence”(Bruner m. fler 1976:90). Igjen blir formativ vurdering sett i lys av elevens forståelsesramme. Hvilken kapasitet har eleven til å forstå tilbakemeldingen?

Både Bruner og Bernstein ser på elevens forståelse i opplæringen (Bernstein 1974, Bruner 1976). Selv om artiklene stammer fra 70-tallet er poengene rundt elevens forståelse like viktig. Black m. flere poengterer i tråd med Bruner og Bernstein at lærers kunnskap om elevens forståelse er av betydning for elevens læringsprosess. Black m. fler skriver: ”As the teachers came more attentively to the students’ responses, they began to appreciate more fully that learning was not a process of passive reception of knowledge, but one in which the learners were active in creating their own understandings” (Black m. fler 2003:59). Konklusjonen for dette resonnementet er at formativ vurdering ikke nødvendigvis gir eleven et utbytte. For at eleven skal få et utbytte må vurderingen tilpasses elevens individuelle forutsetninger.

Hvordan står formativ og summativ vurdering i forhold til hverandre? Ovenfor er det blitt vist at både formativ og summativ vurderingsformer har positive og negative sider. I den norske skole benyttes både formative og summative vurderingsformer. Mens den formative er ment til å videreutvikle elevens kunnskaper (Black m. fler 2003) er den summative ment til å oppsummere kunnskapene (Scriven 1991). De to vurderingsformene har slik sett to ulike formål. Mens en formativ vurdering først og fremst er for eleven, blir den summative vurderingen også sett på av et samfunn. Eleven trenger den summative vurderingen når de skal inn i arbeidsliv og når de skal søke seg inn på skoler. Slik sett kan en si at formativ og summativ vurdering står i både et spenningsforhold og et avhengighetsforhold. Jeg kommer tilbake til formativ og summativ vurdering i kapittel 5 og 7.

I likhet med evaluering har elevvurderingen mange sider som det kan stilles spørsmål til. I debatten rundt formativ eller summativ vurdering, er det egentlig ikke et spørsmål om enten eller, men både og. For at elevvurderingen skal være fullverdig for alle parter bør både formativ og summativ elevvurdering finne sted. Jeg skal videre drøfte hvorfor elevvurdering er et dagsaktuelt tema, blant annet igjennom PISA og kunnskapsløftet.

3.2 Hvorfor elevvurdering er aktuelt – en innsikt i debattene

Elevvurdering er aktuelt i dagens skole av flere årsaker. Jeg skal i dette delkapittelet først belyse elevvurderingens aktualisering igjennom internasjonale undersøkelser (PISA, TIMSS) og Kunnskapsløftet. Deretter tar jeg kort for meg fire debatter. Den første går på standpunktkarakter, den andre handler om eksamen som vurderingsform, den tredje tar opp tilbakemeldinger og den fjerde ser på tilpasset opplæring i forhold til elevvurdering. De fire innleggene er ulike perspektiver på hvorfor elevvurdering er aktuelt i dagens skole. Perspektivene er ikke ment som en helhetlig forklaring til elevvurderingens aktualitet, snarere gir perspektivene kun et lite innblikk til hvorfor elevvurdering er viktig.

Elevvurderingen har igjennom Kunnskapsløftet og blant annet Stortingsmelding nr. 16, 2006-07 fått et nytt løft. De siste årene har det vært debatter om internasjonale undersøkelser som TIMSS, PISA og nasjonale prøver (Dobson & Smith 2007:91). I Stortingsmelding nr. 16 står det følgende: ”I evalueringen av de skolene som startet tidlig med gjennomføring av læreplanene for Kunnskapsløftet, etterlyser lærerne tiltak som kan styrke lærernes og skolens vurderingskompetanse og evne til å vurdere elevenes måloppnåelse” (St. Melding nr. 16:77). Resultatet av dette har blitt prosjektet ”Bedre vurderingspraksis”. Den første rapporten kom i juni 2008. Målet med prosjektet ”er å se om kjennetegn på måloppnåelse i fag på trinn med kompetansemål kan gi en felles nasjonal retning for vurderingsarbeidet og en mer faglig relevant og rettferdig vurdering” (Skolerapport 1 2008:11).

Bak vurderingsprosjektene ligger en årsak om nivåsenkning i den norske skole. Terje Ogden (2005) skriver om hvorfor nivået har senket seg. En av de årsakene Ogden vektlegger er ”inflasjon i karaktergivningen”. Dette begrunner han med ”at elevene oppnår gode karakterer

på grunn av at oppgavene har blitt enklere og evalueringskravene lavere. Det fører til lavere gjennomsnittsprestasjoner som kommer til syne i internasjonale sammenligninger, eller når en sammenligner nivået over tid” (Ogden 2005:18). Året etter Ogden skrev denne kritikken, kom kunnskapsløftet. Kunnskapsløftet kan ses på som et resultat av kritikk rettet mot L 97 og R 94. Spørsmålet er om Ogdens kritikk fortsatt er gjeldende eller om kunnskapsløftet har klart å skape en bedre og mer læringsfremmende vurderingspraksis? I følge skolerapport 1 har lærere som deltar i undersøkelsen et ønske om enda bedre kjennskap til prosjektet (Skolerapport 1 2008:51). Det vil si et ønske om bedre kompetanse på elevvurdering. Hvorvidt Ogdens kritikk om inflasjon i karaktergivningen fortsatt er gjeldende gir rapporten ikke noe sikkert svar på.

Boka ”Elevvurdering i Skolen”, problematiserer en rekke områder som aktualiserer hvorfor elevvurdering er så aktuelt i den norske skole (Tveit 2007). utfordringer med standpunktvurderingen (Lauvås 2007), eksamen (Raaheim og Hauge 2007) så vel som tilbakemeldinger (Viddal 2007) står sentralt. I tillegg vektlegges elevvurdering som grunnlag for tilpasset opplæring i opplæringen (Dale og Wærnes 2007). Jeg skal videre ta for meg disse fire perspektivene som indikatorer på hvorfor elevvurdering er aktuelt i dagens skole.

Per Lauvås går sterkt til verk når han kritiserer standpunktvurderingen i norsk skole. Lauvås skriver: ”min påstand er at slik standpunktvurderingen foregår i norsk skole i dag, er det den som bidrar mer enn noe annet til dårlig kunnskapsnivå” (Lauvås 2007:63). Lauvås vil med dette fjerne det store fokuset på karakterer og heller legge fokus på formative tilbakemeldinger under elevens læringsprosess. Dette er et perspektiv han deler med Boud som skriver:

“One purpose has come to dominate all others. Assessment for certification has received an overwhelming amount of attention in the literature. Most papers on assessment either explicitly or implicitly focus on the aspect of assessment associated with grading, ranking and the portrayal of achievement to others. Assessment to aid learning is either ignored, regarded as a minor issue or relegated to a less technically sophisticated literature on teaching” (Boud 1995:14).

Lauvås poengterer videre at han ikke ønsker å fjerne karakterene, men at de bør suppleres med formative tilbakemeldinger som gir eleven grunnlag for utvikling (Lauvås 2007:57).

Per Raaheim og Helge Hauge fokuserer på problematikken rundt eksamen som vurderingsform (2007). Eksamen skiller seg fra standpunktvurdering ved at den kun baserer seg på en prestasjon, mens standpunktvurderingen tar utgangspunkt i mange vurderinger

igjennom et semester. Raaheim og Hauge peker på at ”skoleeksamen har blitt kritisert for å ha et nokså ensidig fokus på kontroll av hva elevene og studentene til bestemte tider er i stand til å reproducere” (Raaheim & Hauge 2007:66). De hevder videre at eksamen som vurderingsform står ovenfor et validitetsproblem (Raaheim & Hauge 2007:67). Dette skyldes blant annet spørsmålet om eksamen faktisk måler det den faktisk skal måle, eller om den først og fremst måler elevens evne til å prestere i en slik situasjon? Hvilke sider ved eleven er det som blir målt? Som det ble vist i kapittel 2.2.1 er all evaluering subjektiv. Raaheim og Hauge poengterer i denne forbindelse at ”Den karakteren som eleven/studenten oppnår, vil dermed være «urettferdig» i den forstand at dersom en annen sensor hadde vurdert eksamensbesvarelsen, kunne resultatet blitt et annet”(Raaheim og Hauge 2007:67). Det er derfor grunnlag til å kunne drøfte hvorvidt elevvurdering i det hele tatt kan framstå som rettferdig. Dette skal jeg komme tilbake til i kapittel 5.

Både Lauvås, Raaheim og Hauge presenterer en kritikk av to ulike sluttvurderinger i skolen. Linda Marie Viddal legger på sin side fokus på tilbakemeldinger. Tilbakemeldinger er først og fremst aktuelle under elevens læringsprosess. Viddal stiller spørsmålet ”Hvorfor har ikke alle elever utbytte av elevsamtalen?” (Viddal 2007:93). Hvilken rolle spiller tilbakemeldingene i elevvurderingen? Alle elever har en formell rett til elev- eller utviklingssamtaler minst én gang i halvåret (Forskrift til opplæringsloven § 4-5). Dette gjør at også tilbakemeldingene aktualiserer elevvurdering.

Tilpasset opplæring er også med på å aktualisere elevvurdering. I Kunnskapsløftet står det: ”Tilpasset opplæring for den enkelte elev kjennetegnes ved variasjon i bruk av lærestoff, arbeidsmåter, læremidler samt variasjon i organisering av og intensitet i opplæringen. Elevene har ulike utgangspunkt, bruker ulike læringsstrategier og har ulik progresjon i forhold til nasjonalt fastsatte kompetansemål” (LK 06:34). Elevvurdering blir aktuelt i den forstand at det kan være et virkemiddel for å oppnå tilpasset opplæring. Dale og Wærness argumenterer for ”Høy måloppnåelse langs hele karakterskalaen” (Dale og Wærness 2007:106). De hevder at: ”det bør utvikles et vurderingssystem som vektlegger hva eleven skal mestre på de forskjellige karakternivåene. Vi [Dale og Wærness] går altså ikke mot en gradering, forstått som beskrivelse av mestringsnivå. Vi kritiserer et vurderingssystem som beskriver *mangel på mestringsnivå* for alle elever som ikke får karakteren 5 eller 6” (Dale og Wærness 2007:107). I en slik forstand blir tilpasset opplæring å vise hva eleven mestrer, framfor å vektlegge hva eleven

ikke mestrer. ”Ideene om måloppnåelse som uttrykker mestring langs karakterskalaen knytter tilpasset opplæring nærmere til elevens opplæring i fag” (Dale og Wærness 2007:107).

Elevvurdering er helt klart et tema i det norske samfunnet. De internasjonale undersøkelsene ga ringvirkninger i debatten rundt den norske skole. Debattene førte til et nytt læreplanverk, som igjen har ført til mange nye diskusjoner. Fordi elevvurdering vil endre seg i takt med endringene i skolen og samfunnet, vil debatten aldri kunne dø helt ut. Det vil alltid være et spørsmål om hvordan elevvurdering kan bedre elevenes læringsprosess og kompetansenivå. I kapittel 4.3 blir det stilt en rekke spørsmål som problematiserer elevvurderingens rolle.

3.3 Elevvurderingens hva, hvorfor, hvordan

I dette delkapittelet skal elevvurderingens hva, hvorfor og hvordan problematiseres. Det blir tatt utgangspunkt i tre spørsmål. Disse tre spørsmålene er: ”Hva vurderes eleven i?”, ”Hvorfor vurderes eleven?” og ”Hvordan vurderes eleven?” vil bli diskutert. Hensikten er å vise ulike perspektiver som preger elevvurderingen. Bør innsats telle i elevvurderingen? Hvem er egentlig elevvurderingen for? Det er mange svar på slike spørsmål, og perspektivene som her blir lagt fram er bare et knippe av mulige forklaringer. Teksten begynner med spørsmålet ”Hva vurderes eleven i?”.

Hva vurderes eleven i? Hvilke mål setter læreplanverket? Som vist ovenfor er Kunnskapsløftet et produkt av dårlige resultater i den norske skole (se blant annet PISA-undersøkelsene). I større grad enn L 97 ønsker LK 06 å vektlegge elevenes faglige kompetanse. Hvilke faglige ferdigheter har eleven oppnådd? Kompetansemålene i læreplanene for fag er formulert innenfor hvert hovedområde og er i hovedsak angitt etter 2., 4., 7., og 10.årstrinn i grunnskolen. I videregående opplæring er målene formulert for hvert årstrinn i de fleste fagene (LK06:39). Eleven skal vurderes etter disse kompetansemålene. Mange lærere har kritisert dette punktet (Skolerapport 1:7). Lærerne spør: ”Hva med elevens innsats og entusiasme i faget?”. I følge skolerapport 1 har Norge en gammel kultur på å vurdere innsats i faget (Skolerapport 1:7).

Marilyn Osborne m.fl. sammenligner ulike kulturer for evaluering i Europa. Landene de sammenligner er England, Frankrike og Danmark. I følge Osborne skal du ikke langt utenfor Norges grenser for å finne en annen vurderingskultur. Om den franske vurderingskulturen skriver de: "in the French context, where assessment is restricted to the academic aspects of the pupil role and not aimed at the whole person, negative assessment could motivate pupils" (Osborne m. fler 2003:117). Hva elevene vurderes etter er i stor grad avhengig av den rådende vurderingskulturen. Et annet spørsmål er hvorfor vurdere elevene?

Hvorfor vurdere elevene? Elevvurdering kan ha mange hensikter. I følge Inge Bø m.fl. kan hensikter med vurderingen være å motivere, informere, foreta seleksjon og sammenligne ulike elevbesvarelser (Bø m.fl. 2002:88). Her er det på den ene siden en hensikt rettet mot eleven. På den annen side er det en interesse rettet mot samfunnet rundt eleven. To eksempler på at vurdering er for eleven kan hentes fra kunnskapsløftet og stortingsmelding nr. 30. I Kunnskapsløftet står det: "Vurdering og veiledning skal bidra til å styrke deres [elevenes] motivasjon for videre læring" (LK 06:33, min parentes). Stortingsmelding nr. 30 viser til L 97 nå de skriver om formålet med elevvurdering: "Hovedformålet med vurdering er å fremme læring og utvikling" (St.meld. nr 30: 37). Når Stortingsmelding nr. 30 viser til samfunnsmessige hensikter med elevvurderingen viser de til Reform 94. Reform 94 var i følge Stortingsmelding nr. 30 også opptatt av å få fram vurderingens betydning for å informere samfunnet, arbeidslivet og aktuelle utdanningsinstitusjoner om den kompetansen eleven har oppnådd (St.meld. nr 30: 37). PISA – undersøkelsen er også en undersøkelse med utgangspunkt i en elevvurdering rettet mot samfunnets interesser. PISA undersøkelsen vurderer hvordan norske elever ligger an i forhold til andre OECD land. PISA-undersøkelsen tar i hovedsak sikte på å "måle elevenes evne til aktivt å bruke kunnskaper og erfaringer i konkrete situasjoner" (Kjernsli m.fl. 2007:11). Hvorfor eleven vurderes og hva eleven vurderes i danner grunnlag for hvordan eleven vurderes.

Hvordan vurderes eleven? Inge Bø m.fl. skiller blant annet mellom skriftlige tilbakemeldinger, dialog, og muntlig tilbakemelding (Bø m.fl. 2002:89). Disse formene for vurdering kan skje formativ og/eller summativt. De kan også være formelle eller uformelle. Vurderinger kan også gis som belønning eller som en karakter. Til dette stiller Black seg kritisk. Black skriver: "In general, feedback given as rewards or grades enhances ego rather than task involvement – that is, it leads students to compare themselves with others and focus on their image and status rather than encourages them to think about the work itself and how

they can improve it” (Black m.fl.: 2004:46). Hvordan eleven vurderes kan også være opp til den enkelte elev. I følge LK 06 skal elevene også være med på ”å planlegge, gjennomføre og vurdere eget arbeid for å nå nasjonalt fastsatte kompetansemål” (LK 06:33). Metodene dette skjer på vil avhenge av elevens egne strategier. Black viser hvordan lærere kan være med å hjelpe eleven til selvvurdering. Black viser til et forsøk gjort av Sadler om selvvurdering. Black skriver:

”At the start of the project, initial attempts at self-assessment and target setting by students were unsuccessful. The teachers saw that the source of the problem was that their students lacked the necessary skills both to judge specific problems in understanding and to set realistic targets to remedy problems within reasonable time frames. However, those teachers who introduced feedback by comments, and thereby created the classroom environments where students worked together on understanding teacher comments about their work, were providing the training that students needed to judge their own learning and to begin to take action to improve”(Black m.fl.2004:50).

Hvordan eleven vurderes avhenger også delvis av hvem som vurderer eleven og kompetansen til den som vurderer. Engelsen skiller mellom interne og eksterne sensorer (2006:124). En annen mulighet, er som vist ovenfor, at eleven vurderer seg selv eller lar medelever vurdere (Black m.fl.: 2004:29).

Det kan stilles mange spørsmål rundt elevvurdering. Dette delkapittelet har sett på ulike vinklinger og roller til elevvurderingen. Det har blitt tatt utgangspunkt i tre spørsmål. Hva vurderes eleven i? Hvorfor vurderes eleven? Og hvordan vurdere eleven? Elevvurderingens roller er komplekse og har sjeldent entydige svar. Dette gjør at elevvurdering må ses i forhold til den konteksten den kommer fram i.

3.4 Oppsummering

Elevvurdering kan forstås som et redskap, der har til hensikt å beskrive og utvikle elevens læringsprosess. Igjennom formative og summative vurdering får lærer og elev muligheten til å forme elevens læringsprosesser. Endringer i samfunnet gjør at hva eleven vurderes i og hvilke strategier som blir brukt i vurderingen vil endre seg. Dette gjør elevvurdering til et aktuelt tema i dagens skoledebatt. Rollen som elevvurderingen har må ses i forhold til den enkelte

elev. Den vil variere i forhold til hva som vurderes, hvorfor det blir vurdert og hvordan det blir vurdert. Selv om elevvurderingsbegrepet ligger under evalueringsbegrepet er også elevvurdering et stort felt, med mange faktorer som påvirker. Disse faktorene vil ha mye å si for hvorvidt elevvurdering kan oppfattes som rettferdig. I neste kapittel skal jeg se på teori rettferdighet, før jeg i kapittel 5 begynner å drøfte en rettferdig vurderingspraksis.

4. Rettferdighet

I dette kapitlet skal jeg se på begrepet rettferdighet. Rettferdighet er et omstridt begrep med mange ulike synspunkter. Det er mitt hovedformål med dette kapitlet å introdusere ulike tilnærminger til rettferdighetsbegrepet og diskutere disse opp mot en samfunnskontekst. Dette gir en todelt struktur til kapitlet som belyses igjennom to spørsmål. Det første spørsmålet er: ”Hva er rettferdighet?”. Under dette spørsmålet blir det lagt fram teori om ulike tilnærminger til rettferdighetsbegrepet. Det andre spørsmålet jeg stiller er: ”Hvem bestemmer hva som er rettferdig?”. I denne delen diskuteres rettferdighet i forhold til ulike styringsformer og på ulike samfunnsnivåer.

Det er for øvrig viktig å gjøre oppmerksom på at rettferdighetsteorien som det vises til nedenfor er hentet fra økonomiske teorier om hva som er rettferdig. Fordi det finnes lite litteratur på om rettferdighet i forhold til elevvurdering, ser jeg det som nødvendig å hente inn teori fra økonomisk filosofi.

4.1 Hva er rettferdighet?

Spørsmålet om hva rettferdighet er kan svares på fra mange ståsteder. Jeg skal i denne delen holde meg til to teorier om rettferdighet, før jeg til slutt går inn og diskuterer spenninger mellom likhet og ulikhet i rettferdighetsbegrepet. Den første teorien som blir presentert er ”*The Original Position*” av John Rawls. Den andre teorien er ”*Distributive Justice*”. Jeg velger her å gjøre leser bevisst på at verken teorien slik den fremstilles her om ”*The Original Position*” eller teorien om ”*Distributive Justice*” gir noe helhetlig bilde av teoriene. Ei heller er teoriene noe fullkommet svar på spørsmålet ”Hva er rettferdighet?”. Teoriene representerer/gir kun et innblikk i hvordan rettferdighet kan legges fram. Jeg kommer tilbake med ytterligere avgrensninger underveis i kapitlet.

4.1.1 The Original Position

"*The Original Position*" er en idealistisk teori om hvordan man kan oppnå rettferdighet. I 1971 kom John Rawls ut med boken "*A Theory of Justice*" (Rawls 1971). I denne boken ble teorien om "*Justice as Fairness*" lansert for første gang (Wenar 2008). Rawls fortsatte å utvikle teorien om "*Justice as Fairness*" fram til sin død i 2002 (Wenar 2008). "*The Original Position*" er en del av teorien "*Justice as Fairness*". Jeg skal i det følgende kort se på hva Rawls legger i teorien "*Justice as Fairness*" før jeg ser nærmere på "*the Original Position*" og "*a Veil of Ignorance*". Jeg gjør oppmerksom på at verken "*Justice as Fairness*" eller "*The Original Position*" gir noe helhetlig bilde av Rawls sin teori. Dersom leser er ytterligere interessert i temaet, henviser jeg til Rawls (1971,1991) og Stanford Encyclopedia hvor mange av de kommende kildene er hentet fra.

I *Justice as fairness* uttrykker John Rawls at: "According to the provisional aim of moral philosophy one might say that justice as fairness is the hypothesis that the principles which would be chosen in the original position are identical with those that match our considered judgments and so these principles describe our sense of justice" (Rawls 1999:42). Dette viser at *justice as fairness* er en hypotese om hvilke prinsipper som vil bli valgt i *the original position*. Leif Wenar beskriver Rawls sin teori om "*Justice as Fairness*" som følger: "Rawls constructs justice as fairness around specific interpretations of the defining liberal ideas that citizens are free and equal and that society should be fair" (Wenar 2008). "*Justice as Fairness*" må altså ses i forhold til moralsk resonnering. Idealet er at en skal handle kun etter å ha vurdert alle mulige perspektiver, slik en ville gjort i "*the original position*". Wenar skriver "Justice is the maximal moral standard: the full description of how a society's main institutions should be ordered (Wenar 2008).

Underordnet "*Justice as Fairness*" ligger "*the Original Position*" og "*a Veil of Ignorance*". Samuel Freeman skriver om *the Original Position*: "It is designed to be a fair and impartial point of view that is to be adopted in our reasoning about fundamental principles of justice" (Freeman 2008). "*The Original Position*" beskrives som et tenkt samfunn der regler og retningslinjer for et rettferdig samfunn skal skapes. Wenar skriver: "The original position is a thought experiment: an imaginary situation in which each real citizen has a representative, and all of these representatives come to an agreement on which principles of justice should order the political institutions of the real citizens" (Wenar 2008). Troy Jollimore utdyper dette ytterligere når han beskriver Rawls "*Original Position*": "— a position in which agents

possess broad knowledge about human history and the nature of the world they live in, but are denied specific information regarding their own particular identities or prospects in the society in question, the nature of that society, and, crucially, the nature of their own particular conception of the good” (Jollimore 2006). “*The Original Position*” er ikke på noen måte et reelt perspektiv på hvordan rettferdighet kan bestemmes. Poenget til Rawls med å vise til “*the Original Position*” er å få individer og samfunn til å tenke over hvordan de ville sett på rettferdighet dersom de var helt nøytrale i en sak.

The original position” henger nært sammen med “*a Veil of Ignorance*”, på norsk ofte omtalt som uvitenhetens slør. Wenar beskriver “*a Veil of Ignorance*”: “the veil of ignorance, (...) prevents other arbitrary facts about citizens from influencing the agreement among their representatives. (...) Rawls holds that the fact that a citizen is for example of a certain race, class, and gender is no reason for social institutions to favor or disfavor him (Wenar 2008). Hvordan mennesker ville handlet under uvitenhetens slør, kan forstås som Rawls svar på hva som er rettferdig. Hilde Bojer hevder at “Valg bak uvitenhetens slør er et hjelpemiddel til å tenke upartisk; til å arbeide ut en teori for rettferdighet hvor alle teller likt og hvor alles interesser er like viktige”(Bojer: s.4). Oppsummert handler “*the Original Position*” og “*a Veil of Ignorance*” om et tenkt samfunn der mennesker avgjør hva som er rettferdig uten å vite hvilke konsekvenser det vil få for dem selv i den virkelige verden.

Hva kan Rawls sin teori om “*the Original Position*” og uvitenhetens slør gjøre for oss som lever i den virkelige verden? Slik jeg ser teorien gir den oss et tilbud om en tenkt nullstilling i situasjoner hvor vi må avgjøre hva som er rettferdig. Hva ville jeg gjort dersom jeg var helt nøytral i denne saken? “*A Veil of Ignorance*” er ikke ment til å være realistisk, det er ment som et tankeeksperiment. Ideen om at “*a Veil of Ignorance*” er et tankeeksperiment, støttes også av Freeman, som skriver “For as a thought experiment, the original position is not supposed to be a realistic situation. Instead, the veil of ignorance is a vivid representation of the kinds of reasons and information that are relevant to a decision on principles of justice for the basic structure of a society in which persons regard themselves as free and equal” (Freeman 2008). “*A Veil of Ignorance*” er i likhet med andre idealer, noe mennesket kan strekke seg etter for å oppnå et mer rettferdig samfunn. For eksempel kan “*a Veil of Ignorance*” fungere som et ideal for lærere når de skal gi en rettferdig elevvurdering.

John Rawls gir ikke noen fasit på hva rettferdighet er. Istedenfor legger han fram en teori om hvordan det kan oppnås. Denne teorien åpner for mange mulige perspektiver på hva

rettferdighet kan være. Jeg skal videre gå inn å se på noen av disse perspektivene igjennom ”*Distributive Justice*”. ”*Distributive Justice*” er en teori som tar utgangspunkt i at rettferdighet kan fordeles utover ulike prinsipper for hva rettferdighet er.

4.1.2 Distributive justice

I denne delen skal vi se på *Distributive Justice*. *Distributive Justice* kan forstås som ” (...) the moral assessment of distributions, or as the moral assessment of individual or collective decisions in light of how they affect distributions” (Barry 2004). Ifølge Lamont og Favor kan ”Distributive principles (...) vary in numerous dimensions”(Lamont og Favor 2007). Lamont og Favor fordeler *Distributive Justice* ut over sju prinsipper. Disse prinsippene er: *Strict Egalitarianism, the Difference Principle, Resource-Based Principles, Welfare-Based Principles, Desert-Based Principles, Libertarian Principles, og Feminist Principles* (Lamont og Favor 2007). Jeg kommer i det følgende til å se på *Strict Egalitarianism, the Difference Principle, Resource-Based Principles, Desert-Based Principles og Libertarian Principle*. Med dette kommer jeg ikke til å ta for meg: *Welfare Based Principle og Feminist Principle*, da jeg ikke regner disse av prioritert for oppgavens diskusjon om rettferdig elevvurdering.

Det første prinsippet som jeg skal se på i teorien *Distributive Justice* er *Strict Egalitarianism/ Egalitarianism*. Jeg kommer her ikke til å trekke noe konkret skille mellom *Strict Egalitarianism og Egalitarianism* i seg selv. Dette skyldes for det første at *Strict Egalitarianism og Egalitarianism* bygger på samme prinsipper og for det andre at forskjellen kun ligger i hvordan de velger å handle i forhold til prinsippene. *Strict Egalitarianism* tar i følge Lamont og Favor utgangspunkt i en likhetstankegang (Lamont og Favor 2007). Richard Arneson beskriver egalitarismen som følger: “An egalitarian favors equality of some sort: People should get the same, or be treated the same, or be treated as equals, in some respect. Egalitarian doctrines tend to express the idea that all human persons are equal in fundamental worth or moral status” (Arneson 2002). Likhet er altså hovedbudskapet til egalitarismen. Gosepath deler egalitarismen inn i tre underkategorier. Disse er *Intrinsic, Instrumental og Constitutive egalitarianism* (Gosepath 2007). Jeg skal videre se litt nærmere på disse kategoriene.

Det første perspektivet er *Intrinsic egalitarisme*. Om *Intrinsic egalitarianism* skriver Gosepath:”*Intrinsic egalitarians view equality as an intrinsic good in itself. As pure egalitarians, they are concerned solely with equality, most of them with equality of social*

circumstances, according to which it is intrinsically bad if some people are worse off than others through no fault of their own”(Gosepath 2007). Dette kan anses for å være en form for *strict egalitarianism*. Poenget i *intrinsic egalitarianism* er at alt skal være likt. Gosepath er selv kritisk til *intrinsic egalitarianism*, når han argumenterer: ” But would it be morally good if, in a group consisting of both blind and seeing persons, those with sight were rendered blind because the blind could not be offered sight? That would in fact be morally perverse” (Gosepath 2007). Likheter er altså ikke alltid en god løsning på hva som er rettferdig, men som nevnt ovenfor finnes det grader av hva som skal være likt i egalitarismen, dette er *Instrumental* og *konstitutive egalitarianism* eksempler på.

Det andre perspektivet på egalitarisme var *instrumental*. Arneson skriver om *instrumental* egalitarisme: “The instrumental egalitarian values equality as a means to some independently specifiable goal” (Arneson 2002). Det handler om å bruke likhet til å fremme et mål. Larry Temkin er inne på noe av det samme, når han skriver: ” “*instrumental* egalitarianism, where equality is valuable only insofar as it promotes some *other* valuable ideal” (Temkin 2009:156). Forskjellen mellom Arneson og Temkin er at Arneson ser på fremming av et mål, mens Temkin ser på fremming av et ideal. Mål og ideal skiller seg fra hverandre ved at mål kan som regel oppnås, men et ideal kan ikke oppnås.

Arneson gir et eksempel på *instrumental* egalitarisme som kan ses både i forhold til mål og ideal. Arneson skriver: “For example, someone who believes that the maintenance of equality across a group of people fosters relations of solidarity and community among them, and is desirable for that reason, qualifies as an instrumental egalitarian” (Arneson 2002). Eksempelet viser at likhet blir brukt til et formål om å skape fellesskap i en gruppe. Det er ikke likhet i seg selv som er målet, men fellesskapet. *Instrumental* egalitarisme handler altså om å bruke likhet som et instrument eller redskap for å oppnå et mål eller ideal. Målet eller idealet behøver i den forstand ikke være likhet i seg selv, slik som det er i *intrinsic* egalitarisme.

Det tredje perspektivet Gosepath er inne på, er *constitutive* egalitarisme. Gosepath skriver i forhold til *constitutive egalitarianism*: ”According to this approach, we aspire to equality on other moral grounds — namely, because certain inequalities are unjust. Equality has value, but this is an extrinsic value, since it derives from another, higher moral principle of equal dignity and respect”(Gosepath 2007). Rettferdighet som *constitutive* egalitarisme baserer seg på perspektiver om at ulikhet kan være urettferdig og at likhet derfor må ses som rettferdig.

Oppsummert er det blitt vist at det egalitaristiske perspektivet på rettferdighet legger ikke til rette for noen entydig forståelse av rettferdighetsbegrepet. Dette betyr at to egalitarister kan oppleve å være uenige om hva rettferdighet er innenfor en og samme retning. Når egalitarismen blir drøftet videre i kapittel 5, 6 og 7 er det først og fremst *intrinsic egalitarianism* som står i fokus. Jeg kommer i drøftingskapitlene ikke til å gå inn og drøfte ulike retninger innenfor egalitarismen.

Den andre tilnærmingen Lamont og Favor legger fram i teorien om *Distributive Justice* er *Difference Principle* (fra nå av omtalt som differanseprinsippet)(Lamont og Favor 2007). Ifølge Gordon så kan differanseprinsippet anses for å være ” the proposition that, in the “initial situation” rational self-interested persons would accept as just whatever inequalities might occur in real society so long as the absolute position of the least advantaged person is improved” (Scott Gordon 1973:275). Scott sin tolkning av differanseprinsippet kan spores tilbake til John Rawls. John Rawls ser på differanseprinsippet igjennom to punkter.

“Social and economic inequalities are to satisfy two conditions: (a) They are to be attached to positions and offices open to all under conditions of fair equality of opportunity; and (b), they are to be to the greatest benefit of the least advantaged members of society” (Rawls 1993, pp. 5-6. The principles are numbered as they were in Rawls' original *A Theory of Justice* I)” (Lamont og Favor 2007).

Differanseprinsippet kan ut ifra dette tolkes som at samfunnets ulikheter skal være et resultat av ulike posisjoner. Disse posisjonene skal være oppnådd igjennom en rettferdig konkurranse hvor alle har lik mulighet til å innta posisjonen. Differanseprinsippet sier også i sitatet fra Rawls, at det må tas hensyn til de mindre heldige i samfunnet. Dette gjør differanseprinsippet todelt. På den ene siden vektlegger differanseprinsippet at samfunnet består av forskjellige posisjoner, mens på den andre siden skal det være lik mulighet for alle til å oppnå disse posisjonene.

Tanken bak differanseprinsippet beskriver Lamont og Favor som:”More wealth can be produced and indeed this has been the experience of industrialized countries over the last few centuries. The most common way of producing more wealth is to have a system where those who are more productive earn greater incomes” (Lamont & Favor 2007). Differanseprinsippet bygger her på tanker om innsats og fortjeneste, men er det nødvendigvis rettferdig? Owen McLeod gir et eksempel på innsats og fortjeneste i forhold til vurdering av studenter:

”For example, suppose you've worked hard toward getting an A on your term paper. You spent hours in the library doing research, you composed several drafts, you went hours without sleep, and so on. Your end is morally okay, since there's nothing wrong with trying (in these ways!) to get an A. Even so, it's still possible that, in spite of your tremendous efforts, your paper is terrible. In that case, you probably don't deserve an A” (McLeod 2008).

Dette eksempelet viser hvordan innsats ikke er lik karakter, men hva er det en potensiell arbeidsgiver ønsker å vite om en mulig ansatt? Er det innsats eller karakteren til personen som spillers størst rolle for arbeidsgiver? Er det rettferdig at innsats ikke teller i skolen? Innsats er med på å skape ulikheter mellom både grupper og mennesker, dette betyr allikevel ikke at det er urettferdig. Egalitarismen og differanseprinsippet skaper et spenningsforhold mellom likhet og ulikhet i rettferdighetsbegrepet. Videre skal vi inn og se på det tredje perspektivet som Lamont og Favor presenterer i forhold til *distributive justice*, nemlig det *Resource-based principle*.

Det resursbaserte prinsippet er det tredje prinsippet som ligger under *distributive justice*. Lamont og Favor viser skriver om det resursbaserte prinsippet:” proposes that people begin with equal resources but end up with unequal economic benefits as a result of their own choices. What constitutes a just material distribution is to be determined by the result of a thought experiment designed to model fair distribution”(Lamont og Favor2007). Utgangspunktet for det resursbaserte prinsippet er at mennesker får tildelt like resurser fra fødselen av og at rettferdighet etter det kan ses på som et resultat av hvilke valg man tar. Dette kan kritiseres.

Har mennesker lik tilgang til resurser? Lamont og Favor kritiserer det resursbaserte prinsippet når de skriver:”It seems impossible to measure differences in people’s natural talents – unfortunately, people’s talents do not neatly divide into the natural and developed categories” (Lamont og Favor 2007). Sosial og genetisk bakgrunn skaper forskjeller som det må tas hensyn til når rettferdighet vurderes. Et eksempel på dette kan være elever med lærevansker. Det resursbaserte prinsippet tar ikke hensyn til at disse elevene har et annet behov enn flertallet. På den annen side så har prinsippet et poeng i at noe bør være likt. For eksempel lik rett på skolegang, men selv om skolegangen er lik betyr ikke det at elevenes kunnskaper er de samme når de går ut fra skolen. Slik sett befinner rettferdighetsprinsippet seg i et spenningsfelt mellom hva som bør være likt og hva som bør være ulikt.

Det fjerde prinsippet som står på dagsorden under *distributive justice* er *desert principle*, eller fortjenesteprinsippet. Dette er en viktig måte å tenke rettferdighet på. Hva har fortjeneste å si for hva som er rettferdig? Er det slik at dersom du har fortjent noe, får du det? Er det rettferdighet mellom hva du fortjener og hva du faktisk oppnår? I fortjenesteprinsippet ligger det tre perspektiver på hvordan rettferdighet kan fortjenes. Disse er "contribution", "effort" og "compensation".

Det skal nå vises hvordan "contribution", "effort" og "compensation" er viktige begreper i diskusjonen om rettferdighet. Begrepene kan ses i forhold til enkeltindivider og er derfor av særskilt interesse for elevvurderingen som skal drøftes i kapittel 7. Om *contribution* skriver Lamont og Favor: "People should be rewarded for their work activity according to the value of their contribution to the social product" (Lamont og Favor 2007). Rettferdighet må i et slikt perspektiv ses i forhold til belønningen og bidraget. Rettferdiggjør belønningen/straffen bidraget som ble lagt til grunn, eller kan denne oppleves som urettferdig? Det andre perspektivet som er relevant for fortjeneste er "effort". Om effort skriver Lamont og Favor: "People should be rewarded to the effort they expend in their work activity" (Lamont og Favor 2007). Hvordan måle en innsats rettferdig? Rettferdighet ses i dette perspektivet i forhold til innsatsen som gis og belønningen som oppnås. Hvordan kan lærer på en rettferdig måte vurdere elevens innsats? Det tredje perspektivet Lamont og Favor presenterer er "compensation". Om compensation skriver de: "People should be rewarded according to the costs they incur in their work activity" (Lamont og Favor 2007). Kan en kompensasjon være rettferdig? Igjen blir rettferdighetsbegrepet relativt. Hva som kompenserer for kostnadene et menneske har lagt i et arbeid, vil variere fra person til person.

Perspektivene Lamont og favor legger til fortjenesteprinsippet representerer en del av hva som må til for at noe skal oppleves som rettferdig. Rettferdighet må ses i forhold til hva som blir lagt i arbeidsoppgaven og hva som gis tilbake. I skolesammenheng kan dette ses i kontekst av elevens innsats og hva eleven får ut av det. Jeg kommer i kapittel 7 tilbake til fortjenesteprinsippet når jeg diskuterer en rettferdighet igjennom ulike vurderingsformer.

Det siste perspektivet som vi skal se på innenfor *distributive justice* er libertarianismen. Mens prinsippene ovenfor har kunnet bli sett i en sosial sammenheng, ligger libertarianismen tett opp mot individet og individets frihet. Lamont og Favor vektlegger forskjellene mellom libertarianismen og andre prinsipper når de skriver: "advocates of Libertarian distributive

principles rarely see the market as means to some desired pattern, since the principle(s) they advocate do not ostensibly propose a ‘pattern’ at all, but instead describe the sorts of acquisitions or exchanges which are themselves just” (Lamont og Favor 2007). Rettferdighet handler om hva som er best for individet og hva individet opplever som rettferdig. Lamont og Favor beskriver libertarianismens funksjon for *distributive justice* som et prinsipp som kan hjelpe dem med minst fordeler (Lamont og Favor 2007). Barry skriver om redistribuering i forhold til individualisme og hevder at: “redistributive policies often claim that while individuals may have *positive* ethical duties to aid poor or unwell persons, it is morally impermissible to compel them to do so through state-administered tax and transfer or other means, unless universal consent for these policies can be secured” (Barry 2004). Libertarianismeperspektivet tar avstand fra distributive mønstre, holdninger og verdier. Nozick argumenterer for at:”because people own themselves and hence their talents, they own whatever they can produce with these talents” (Favor og Lamont 2007). Problemet med et slikt perspektiv på rettferdighet er hvordan resten av samfunnet står i forhold til enkeltindividet. Det kan gis mange eksempler på hvorfor libertarianismen er et ekstremt perspektiv på rettferdighet. Dersom en person har et talent i å stjele, betyr ikke dette at stjeling er en rettferdig handling.

Ovenfor har rettferdighetsbegrepet blitt belyst fra fem perspektiver. Disse har vært rettferdighet i forhold til egalitarisme, *difference principle*, *resource based principles*, *desert principles* og *libertarianisme*. Ved å bruke Lamont og Favor sitt distribuerte syn på rettigheter har det blitt vist at rettferdighet er et sammensatt begrep. Det kan ta utgangspunkt i hva som er av fortjeneste, hvordan resursene har blitt fordelt, eller i forhold til individets ferdigheter og behov.

I en klargjøring av hva rettferdighet innebærer ligger det et spenningsforhold mellom individuelle rettigheter og samfunnets perspektiv på hva som er rettferdig. I rettferdighetsbegrepet bør det derfor være rom for tolkning, både til individets og til samfunnets beste.

4.1.3 Et spenningsforhold: likhet eller ulikhet?

I denne delen skal jeg diskutere rettferdighet i et spenningsforhold mellom likhet og ulikhet. Dette spenningsforholdet ble skissert ovenfor igjennom egalitarismen, differanseprinsippet og

som ytterste motpol libertarianismen. Det kommer i denne delen ikke til å bli vist hensyn til *resource* og *desert* prinsippene. Jeg kommer i det følgende først og fremst til å referere til egalitarismen som en likhetstankegang og differanseprinsippet som ulikhet. I dette delkapittelet skal det begrunnes hvorfor det er viktig å drøfte rettferdighetsbegrepets innhold. Det vil bli vist at likhet ikke nødvendigvis skaper rettferdighet. En slik debatt er ikke bare relevant for tilnærmingen til en rettferdig vurderingspraksis, men kan også belyse andre samfunnsorienterte problemstillinger.

Hva betyr det egentlig at elever blir vurdert ut fra et likt grunnlag? Skal alt være likt eller bare deler av vurderingen? Lucas skriver: "Equal treatment is often unfair, we have seen, because it fails to do justice to the individual's case" (Lucas 1989:175). Er likhet nødvendigvis rettferdighet?

Når rettferdighet blir definert, følger det gjerne en drøfting av spenningen mellom likhet og ulikhet. Stephen David Ross beskriver spenningen mellom likhet og ulikhet som: "...the pursuit of two incompatible ideas: equality among human beings and things when they are unequal; and promotion of their unique identities, building the differences among them" (Ross 1998:399). Ross får her fram at det ikke er mulig å se rettferdighet som både likehet og ulikhet under ett og samme tak. Tanken om frihet og tanken om fellesskap i rettferdighetsbegrepet er sentrale momenter. Problematismen av disse er viktige for å forstå hva rettferdighet er. Ross utdyper sitt perspektiv på rettferdighet: "Justice is equality under the Same, but the Same imposes injustice on heterogeneous plurality. Similarly, justice is recognition of uniqueness and individuality, but heterogeneity imposes injustice as inequality. Only with a measure that bridges homogeneity and heterogeneity can even God distribute justice without injustice." (Ross 1998:399). Men er det realistisk å bygge en bro mellom likhet og ulikheter? Vil rettferdigheten da bli rettferdig for alle?

Rettferdighet kan altså ikke tolkes som enten likhet eller ulikhet. Det vil nesten alltid finnes parter som kan argumentere for at fordelingen er urettferdig. For eksempel vil det i et multikulturelt samfunn være problematisk rundt hva som skal være likt og hva som skal være ulikt. Dette skyldes forskjellige kulturer som stiller med forskjellige oppfatninger på hva som er rettferdig. Dette fører til at det er konstante spenninger i rettferdighetsbegrepet. Dette bekrefter Lucas når han skriver: "Instead of seeing justice as a simple static assignment of benefits, responsibilities and burdens, we should see it as a dynamic equilibrium under tension, wanting to treat the individual as tenderly as possible, yet being prepared, for

sufficiently compelling reasons, to take a tough line” (Lucas 1989:18). Lucas får her fram de flerfoldige spenningene i rettferdighetsbegrepet og utdyper spenningene til å være mer enn forholdet mellom likhet og ulikhet. Lucas vektlegger også forholdet mellom samfunnet og det individuelle. I del 4.2.2 kommer jeg tilbake til flere eksempler på kulturkonflikter som kan skape uenighet i forhold til hva som er rettferdig.

Michal Blake viser problematikken mellom det individuelle og det universelle i et eksempel om skolen. Skolen Blake viser til er bygd opp på bakgrunn av kulturelle forskjeller. Blake skriver:

”For example, one of the implications of the right to culture discussed earlier was that different cultural groups have a right to establish schools that cater to their specific needs. It is justified for these schools to favor children who belong to the group they are meant to serve, but it would be unjustified for them to discriminate in any way among these children” (Blake 2005).

Skaper legitimering rettferdighet? Innholdet i rettferdighetsbegrepet er vanskelig og legitimere nettopp på grunn av at spenningen mellom fellesskapets likhetstrekk og individuelle ulikheter. John Raws vektlegger at i rettferdighetsbegrepet bør det være rom for tolkninger. Rawls skriver ”Men can agree to this description of just institutions since the notions of an arbitrary distinction and of a proper balance, which are included in the concept of justice, are left open for each to interpret according to the principles of justice that he accepts” (Rawls 1999:5). Spørsmålet blir hvor mye rom for tolkning er det gunstig å gi? For lite tolkningsrom vil gi likhet for alle og bli opplevd som urettferdig. For stort tolkningsrom vil gi for mye frihet og være lite gunstig for samfunnet som en helhet.

Hvorvidt rettferdighet skal basere seg på likhet eller ulikhet finnes det ingen fasit på. Det er mulig å tenke seg en gylden middelvei mellom likhet og ulikhet. Poenget er at prinsippet om rettferdighet bør ha rom for både likhet på samfunnsnivå og ulikhet på individnivå. I neste delkappittel blir det stilt spørsmål om hvem som bestemmer hva som er rettferdig.

4.2 Hvem bestemmer hva som er rettferdighet?

Spørsmålet ”Hvem bestemmer rettferdighet?” er et grunnleggende spørsmål for å kunne si noe om hva rettferdighet er. Ulike styringsformer i ulike samfunn vil prege hvem som bestemmer rettferdigheten. Det vil i dette delkapittelet bli skilt mellom ulike styringsformer. Deretter blir det gjort en drøfting av rettferdighet på ulike samfunnsnivåer. Hensikten er å vise at rettferdighetsbegrepet vil variere med ulike styringsformer, i ulike samfunn og på ulike nivåer i samfunnene. Det vil også bli tatt hensyn til individuell forståelse for rettferdighet i del 4.3.2. Hva skiller prosessen som definerer rettferdighet på verdensbasis fra prosessen som definerer rettferdighet innad i en familie?

4.2.1 Rettferdighet på bakgrunn av ulike styringsformer

Det vil i det følgende bli drøftet rettferdighet i lys av tre styringsformer: demokrati, anarki og paternalisme. Rettferdighetsbegrepet vil her ikke bli drøftet på et individnivå. Det er allikevel viktig å ta med seg at rettferdighet på individnivå ikke nødvendigvis henger nært sammen med det som skjer på samfunnsnivå.

Hvem bestemmer rettferdighet i et demokratisk samfunn? Sveinung Vaage vektlegger frigjøring i sin definisjon av demokratiet. Om demokrati skriver Vaage ”demokrati betyr frigjøring av intelligensen til selvstendig virksomhet – en frigjøring av bevisstheten som selvstendig organ til å fungere på egen hånd” (Vaage 2000:85). Tom Christiano utdyper Vaage sitt perspektiv på demokratiet. Christiano beskriver demokratiet som “a method of group decision making characterized by a kind of equality among the participants at an essential stage of the collective decision making” (Christiano 2006). Et demokrati er tosidig. På den ene siden handler det om individets frihet til å uttrykke seg og få fram sine meninger. På den andre siden handler det om å respektere andres meninger og flertallets avgjørelser. Demokratiet representerer altså en oppfatning om at det er flertallet i folket som bestemmer hva som er rettferdig. Det kan stilles spørsmål i forhold til hvor stor grad av medbestemmelse folket har. I norske rettssaker er det en generell oppfatning om at domstolene ikke ønsker å komme i konflikt med folkets rettsfølelse eller hva folket anser som rettferdig. Dette er for å

kunne behandle saker på et mest mulig nøytralt grunnlag. Hva er det med demokratiet som gjør det så bra? Christiano skriver: "Many have noted with Mill and Rousseau that democracy tends to make people stand up for themselves more than other forms of rule do because it makes collective decisions depend on them more than monarchy or aristocracy do. Hence, in democratic societies individuals are encouraged to be more autonomous" (Christiano 2006). Christiano peker her på viktige demokratiske verdier som nettopp at mennesker er autonome, selvstendige individer. Igjennom demokratiet får individet mulighet til å uttrykke sitt syn på rettferdighet. Det rettferdighetssynet som dominerer i demokratiet, kan da anses som denne styreformens syn på hva rettferdighet er. Dette vil for øvrig variere mellom ulike demokratiske samfunn.

I likhet med rettferdighetsbegrepet ligger demokratibegrepet i et spenningsforhold mellom individets interesser og samfunnets interesser. Et demokratisk rettferdighetsbegrep må ses i forhold til nivået demokratiet blir utført på. For eksempel så har skolen i Norge som oppgave å utdanne elever til å bli demokratiske aktører. I den generelle delen av læreplanen står det uttrykt:

"Opplæringen må spore den enkelte til driftighet og til tett samvirke for felles mål. Den må lære elevene framferd som gjør det lettere for dem sammen å nå resultatene de sikter mot. Den må fremme demokrati, nasjonal identitet og internasjonal bevissthet. Den skal utvikle samhørighet med andre folk og menneskenes felles livsmiljø, slik at vårt land blir et skapende medlem av verdenssamfunnet" (Den generelle delen:2).

Hvor frie elevene skal være og hvilke konsekvenser elevens handlinger skal ha, er noe som bør bli avgjort på en demokratisk måte. Det kan stilles spørsmål ved hvor stor medbestemmelse elevene skal ha. Er det greit å gi dem økende medbestemmelse ved økende alder?

Demokratiet som styringsform kan variere igjennom ulike samfunn. Et demokrati kan eksistere på statlig nivå så vel som i en familie eller på skolen. Demokratiske verdier er de verdiene som flertallet stemmer fram. Dette synet gjelder også rettferdighet. Det er demokratiet som avgjør hva som skal anses for å være rettferdig og hva som ikke er rettferdig. Jeg skal nå forlate demokratiet som styringsform og se på hvordan anarkiet legger til rette for rettferdighet.

Anarkiet står som kontrast til demokratiet. Anarki kan forstås som ” meaning the lack of a higher authority that makes claims on lower authorities”(Philpott 2009). Philpott viser her hvor fritt et anarki kan være. Vil rettferdighet i det hele tatt være et begrep i et slikt samfunn? I 2008 skrev dagsavisen om hvordan Sv ved kommunevalget i 2003 etterlot et inntrykk av at partiet nærmest ville innføre anarki i grunnskolen med ekstrem valgfrihet (Strand 2008). Hva skjer når ingen bestemmer hva rettferdighet er? Catherine Lu skriver om anarki som kritikk av et verdensherredømme. I kritikken bruker hun anarkiet som en pekepinn på hvor lite realistisk et verdensherredømme er. Lu skriver:

“Assuming that world government would lead to desirable outcomes such as perpetual peace, realists are skeptical that world government will ever materialize as an institutional reality, given the problems of egoistic or corrupted human nature, or the logic of international anarchy that characterizes a world of states, all jealously guarding their own sovereignty or claims to supreme authority” (Lu: 2006).

Lu får her fram hvordan et anarkisk samfunn ikke vil fungere på grunn av menneskers egoistiske og korrupte natur. Anarki kan oppleves på flere nivåer i samfunnet. Det kan finne sted innenfor landegrenser eller i mer lokale strøk. I 2003 skrev aftenposten om anarki og lovløshet i Bagdad (Seierstad 2003). Tilstanden i Bagdad beskrives som: ”De har levd i tiår under et regime som har tatt alt fra dem, friheten, kjøpekraften, valgmuligheten. Nå mener de at de kan stjele noe tilbake. Som om det gjør opp for alt de har tapt. Men de stjeler fra sitt eget folk, ikke fra regimet”(Seierstad 2003). Folket er frie til å gjøre hva de vil, det får ingen konsekvenser. I lys av dette kan det stilles spørsmål om frihet fører til rettferdighet? Kanskje frihet i et slikt tilfelle virker mot sin hensikt og skaper et urettferdig samfunn framfor et rettferdig? Dersom ingen setter grensene i forhold til hva som er rettferdig, vil rettferdighetsbegrepet bli visket ut. Anarki representerer i et slikt tilfelle et åpent syn på hva som er rettferdig. Individene kan få gjøre akkurat som de vil og slipper å se rettferdighet som noe mellommenneskelig. Dette er allikevel neppe det mest gunstige for å få samfunn til å fungere.

Paternalismen er det tredje perspektivet på hvem som kan bestemme rettferdighet. Tom Beauchamp definerer paternalismen som “the intentional overriding of one person's known preferences or actions by another person, where the person who overrides justifies the action by the goal of benefiting or avoiding harm to the person whose preferences or actions are overridden” (Beauchamp 2008). Ut ifra det Beauchamp skriver om paternalisme, blir

rettferdighet i et paternalistisk perspektiv bestemt av en person som står over en annen. Et slikt perspektiv på hvem som bestemmer kan få konsekvenser for den som blir bestemt over. Gerald Dworkin reflekterer rundt dette poenget. Dworkin skriver:

”Paternalism is the interference of a state of an individual with another person, against their will, and justified by a claim that the person interfered with will be better off or protected from harm. The issue of paternalism arises with respect to restrictions by the law such as antidrug legislation, the compulsory wearing of seatbelts, and in medical contexts by the withholding of relevant information concerning a patient’s condition by physicians. At the theoretical level it raises questions of how persons should be treated when they are less than fully rational” (Gerald Dworkin 2005).

Paternalisme kan ses i flere typer samfunn. For eksempel at legen bestemmer hva som er rett behandling for en syk pasient, kan argumenteres for i den forstand at legen mest sannsynlig har mer kompetanse enn pasienten til å avgjøre den beste behandlingen. Et slikt eksempel kan også gis i skolekontekst, hvor læreren har mulighet til å bestemme over eleven. Alan Wertheimer gir et lignende eksempel: ”A paternalistic argument could not justify interfering with exploitative transactions if the exploitative transaction is advantageous to the exploitee and if interference is not likely to result in a transaction that is more beneficial to *B*. For paternalism justifies interfering for someone's good, and this interference would not be to the target's benefit (Wertheimer 2008). Spørsmålet er allikevel om dette er en rettferdig avgjørelse? Hvilken grad er legitimering av en handling en rettferdig handling? Legitimering kan begrunne hvorfor en handling er riktig, men det rettferdiggjør ikke nødvendigvis handlingen.

I det norske samfunnet er det først og fremst demokratiet som avgjør hva som er rettferdig. Rettferdighetstanken dannes her igjennom et folkestyre. Anarki og paternalisme er andre måter å tenke rettferdighet på. I anarkiet er det ingen prinsipper om hva som er rettferdig. Slik sett blir anarkiet en representant for ekstrem individualisme. I hvilken grad disse avgjørelsene oppfattes som rettferdige, vil variere. Er det nødvendigvis slik at lærer kan bestemme hva som er best for eleven og vil eleven og miljøet rundt oppleve lærerens avgjørelse som rettferdig? Det vil videre bli satt fokus på hvilke samfunnsnivåer rettferdighet kan forstås på. Dette blir gjort igjennom en modell av Urie Bronfenbrenner.

4.2.2 Nivåer for et rettferdig samfunn: hvem bestemmer?

I det følgende skal jeg bruke Urie Bronfenbrenner sitt system for økologisk utvikling til å drøfte hvem som bestemmer rettferdighet på ulike samfunnsnivåer. Det er viktig å ha i minne at ”Det fins *mange* samfunn, ikke *ett* samfunn, unntatt i rent abstrakt eller ideell betydning” (Vaage 2000:69). Dette gjelder også modellen til Bronfenbrenner. Hvem som bestemmer rettferdighet, vil reflekteres igjennom ulike samfunnsnivåer. Disse nivåene er av betydning for hva rettferdighet kan forstås som.

Bronfenbrenners system består av fire relaterte systemer. Disse systemene er Mikrosystemet, mesosystemet, eksosystemet og makrosystemet. Stephen von Tetzchner påpeker at: ”Det som skjer innenfor et system, henger sammen med det som skjer i de andre. Mikrosystemer påvirker hverandre og danner mesosystemet” (Tetzchner 2003:19). Hva kan Bronfenbrenners økologiske system ha og si for hvem som bestemmer rettferdighet? Fordi systemet ikke har noen klare avgrensinger mellom nivåene, vil det heller ikke være en entydig fasit på hvem som bestemmer hva rettferdighet er. Det vil bli lagt større vekt på de to minste nivåene enn de to største, da mikro og mesosystemene er av større relevans for problematiseringen av en rettferdig elevvurdering.

”*Mikrosystemene* består av et barns helt nære, daglige relasjoner. I vestlige land er det som regel foreldrene, søsken og andre slektninger som bor med familien, barnehage, fritidsklubb, vennene i nabolaget også lignende”(Tetzchner 2003:19) Dette systemet ses i lys av hva som skjer innad i et miljø og hvilke styringsformer som blir benyttet. Rettferdighet blir her bestemt i forhold til gitte kontekster og relasjoner. I relasjonen mellom voksen og småbarn er det naturlig å benytte seg av en paternalistisk tankegang (se f. eks Dworkin 2005). Den voksne personen vet gjerne hva som er best for barnet og kan derfor også si hva som er en rettferdig for barnet. Etersom barnet blir eldre endrer forholdet mellom voksen og barn seg (Evenshaug og Hallen 2002:45). Barnet får mer å si og kan uttrykke egne meninger på hva som er rettferdig. I en slik situasjon kan det være naturlig med et mer demokratisk forhold mellom voksen og barn. Ved tilfeller av omsorgssvikt kan det stilles spørsmål om det er noen som veileder barnet til hva som er rettferdig. Ved ekstreme tilfeller kan det diskuteres hvorvidt disse barna vokser opp under anarkiske omstendigheter. Det kan også stilles spørsmål ved hvordan disse barna lærer seg å se på rettferdighetsbegrepet og hvordan dette påvirker deres oppførsel blant venner og i skolen.

Edna Ullmann-Margalit skriver om rettferdighetsbegrepet i en familiekontekst. Hun skriver "Members of the "good" or exemplary family will want to be fair to each other and will attempt to ensure that their family deal, as it evolves and changes, is guided by considerations of fairness"(Ullmann 2006:2). Ullmann problematiserer videre hvordan rettferdighetsbegrepet i familien ikke nødvendigvis reflekteres på andre mikronivåer. Ullmann skriver:"However, family members' idea of family fairness is likely to reflect an understanding of fairness that is different from the way fairness is understood in the context of other social institutions. In particular, "fair within the family tends not to be equated with "equal."" (Ullmann 2006:2). Dette betyr at forståelsen for rettferdighet i et mikrosystem ikke nødvendigvis gir forståelse for rettferdighet i andre mikrosystemer.

Spørsmålet er om et slikt rettferdighetssyn på mikronivå også gjelder de andre kulturene på dette nivået? Hva med skolen? Hvem bestemmer rettferdighet her? I læreplanverket for Kunnskapsløftet står det at "Skolen og lærebedriften skal forberede elevene på deltakelse i demokratiske beslutningsprosesser og stimulere til samfunnsengasjement både nasjonalt og internasjonalt" (2006:33). Tanken er at demokratiske verdier skal stå i sentrum i skolen. Elevene skal ha medbestemmelse i sin egen hverdag. Scriven skriver om en fordelt rettferdighet i skolen. "Distributional justice requires that the distribution of goods or taxes – or praise or blame – be done according to merit on the relevant dimensions" (Scriven 1991:162). For eksempel: "When students' academic work is graded on the basis of effort alone, rather than quality, or on the basis of gender or favoritism, they are treated unjustly and a better system of evaluation is needed" (Scriven 1991:162). Eksempelet Scriven gir, viser en diskusjon av rettferdighet og vurdering i et mikrosystem. Hva er mest rettferdig: innsats eller kvalitet på produktet? Jeg kommer tilbake til vurdering som rettferdig ut ifra innsats eller kvalitet på produktet i kapittel 7.

Rettferdighet på mikronivå må ses i forhold til barnets helt nære og daglige relasjoner. I mesosystemet er det forholdet mellom disse minimiljøene som står i fokus. "*Mesosystemet* består av forbindelsene mellom to eller flere mikrosystemer som barnet er aktivt deltaker i" (Tetzchner 2003:19). Som vist ovenfor er ikke alltid rettferdigheten det samme innad i to mikrosystemer. I det følgende skal rettferdigheten i mikrosystemene drøftes. Hvem kan si noe om hvilken rettferdighet som bør prioriteres?

Spørsmålet som stilles ovenfor lar seg vanskelig svare. Det er ingen fasit som gir et mikromiljø rett til å bestemme over et annet. Dersom mikromiljøene er enige om hva som er rettferdig kan man si at miljøene samarbeider om å bestemme rettferdighet. Dersom miljøene ikke er enige, blir spørsmålet vanskeligere å besvare (Om konflikt innenfor nivåene se Tetzchner 2003:19). Det er vanskelig å generalisere rettferdighet i mesosystemet til å gjelde alle mesosystemer. I det kommende eksempelet tas det derfor forbehold om at dette ikke gjelder alle hjem.

Debatten om bruk av Hijab i norsk skole kan være et eksempel på en slik konflikt. Thomas Hylland Eriksen viser til Hege Storhaug som foreslo ” et forbud mot hijab i skolen, ettersom hun mente at plagget symboliserte ’støtte til politisk islam, sharia og en totalitær ideologi’ ” (Eriksen 2004). Dette forslaget går imot mange muslimske hjems tradisjoner. Hylland Eriksen problematiserer bruken av hijab i forhold til prinsipper om likestilling, likehet, like rettigheter og individuell frihet (Eriksen 2004). Prinsippene drøfter for og mot bruken av hijab. Skal alle ha lik rett til å kle seg i det de vil eller skal skolen kunne bestemme dette? Er det individuell frihet å ikke få lov til å gå med hijab. Hvem bestemmer hva som er rettferdig?

Et annet eksempel på konflikter i mesosystemet er leksedebatten hvor det argumenteres mot skolens bruk av hjemmelekser. Er det rettferdig at skolen bestemmer at eleven skal ha hjemmelekser? I følge Krister Clausen Hoas kan elever droppe lekser med loven i hånd. Hoas hevder at ”Det finnes ikke hjemmel for å gi hjemmelekser i et eneste formelt styringsdokument, i loven eller forskriftene” (Hoas 2009). Hoas poengterer videre at ” elever og foreldre med loven på sin side ganske enkelt kan nekte å gjøre hjemmelekser” (Hoas 2009). Foreldre som hevder at elever skal slippe lekser begrunner dette ofte med at deres egne resurser ikke strekker til. Kanskje har ikke foreldrene tid til å hjelpe barna med leksene eller kanskje har de ikke kompetanse nok. Rita Gjørven ved institutt for lærerutdanning og skoleutvikling skriver i en artikkel for Dagsavisen at:

”Dersom elever skal gjøre skolearbeid utenfor undervisningstiden, må det være elevenes arbeid, og ikke foreldrenes. Det er selvsagt viktig at foreldrene er interessert i hva barna lærer på skolen og ikke minst i hvordan undervisningen legges opp. Men de færreste foreldre er pedagoger, og selv om de er aldri så interesserte i sine barns skolearbeid, har de ikke den faglige og pedagogiske innsikt som må til for å legge til rette for læring av skolefag” (Gjørven 2009).

Ut ifra dette vises det at foreldrenes argumentasjon ikke holder mål. Lekser er noe elevene skal klare på egenhånd. Så, hvem skal bestemme hvor rettferdig det er at eleven må gjøre lekser?

Det er vanskelig å gi noe konkret svar på rettferdighetsproblemer i mesosystemet. Dette skyldes at det er to eller flere sider i saken og fordi sakene må ses i forhold til en større kontekst, kan det vanskelig la seg avgjøre hvilken part som bestemmer rettferdighet. De to neste systemene beveger seg fra individnivå og opp på samfunnsnivå. Fordi systemene ikke er av større relevans for oppgaven, blir det en begrenset gjennomgang av disse.

”Eksosystemet er (...) sosiale sammenhenger barnet ikke selv er en del av, men som er viktige for familiens og barnets fungering. Det omfatter foreldrenes arbeidsplass og sosiale nettverk, fjerne slektninger, foreldre til venner og lignende” (Tetzchner 2003:19). Disse systemene er med på å sette rammer for mikro- og mesosystemets oppfattelse av rettferdighet. For eksempel kan foreldres arbeidstid redusere tiden de har til å hjelpe barnet med lekser (Gjørven 2009). Slik sett, blir lekser ansett som en urettferdig fra familiens side. I tillegg kommer makrosystemet. *”Makrosystemet”* er samfunnets institusjoner, som regjeringen, offentlige tjenester og lignende. Det omfatter samfunnets lover, vaner og verdier av betydning for det som skjer i de andre systemene” (Tetzchner 2003:19). Hvem som i makrosystemet bestemmer rettferdighet blir i det norske samfunn i avgjort av demokratiet. Det kan nevnes at prinsipper om en fordelt rettferdighet også kan drøftes i makrosystemet (Lamont og Favor 2007). Det er regjeringens ansvar å fordele rettferdighet ned til eksosystemet, mesosystemet og mikrosystemet.

Bronfenbrenners system for utvikling på ulike samfunnsnivåer kan vise forskjellen i hvem som bestemmer rettferdighet. Jeg har ovenfor diskutert hvem som bestemmer rettferdighet på macro-, exo-, meso-, og micronivå i samfunnet. Alle samfunnene gjenspeiler et bilde av hva rettferdighet er og hvem som bestemmer det, men forståelsen av rettferdighetsbegrepet er noe ulikt. Mens det i mikrosystemet hersker en stor individuell frihet for hvem som bestemmer rettferdighet, blir dette mer og mer samfunnsstyrt desto lenger opp i systemet du kommer. Eksosystemet er i større grad enn mesosystemet ansvarlig for et samfunnssyn på rettferdighet, mens chronosystemet tar det enda et hakk oppover. I kapittel 5,6 og 7 om en rettferdig vurderingspraksis vil vurderingspraksisen først og fremst bli drøftet i forhold til rettferdighet på mikro og mesonivået. De andre nivåene vil kun spille en indirekte rolle videre i oppgaven.

4.3 Avslutning

I dette kapittelet har det blitt vist at rettferdighet kan belyses fra mange perspektiver. Kapittelet er todelt igjennom to spørsmål. Det første spørsmålet er ”Hva er rettferdighet?” og det andre spørsmålet er ”Hvem bestemmer hva som er rettferdig?”.

Spørsmålet ”Hva er rettferdighet?” blir først besvart som et idealistisk bilde ved hjelp av John Rawls *Original Position*. Deretter besvares spørsmålet igjennom *Distributive Justice*, som representerer en rekke perspektiver på hva rettferdighet kan være. Igjennom John Rawls ble det vist at et ideal for rettferdighet er å kunne stille seg hundre prosent nøytralitet aldri vil være oppnåelig ble det også vist ulike tilnærminger til rettferdighetsbegrepet. De ulike tilnærmingene til rettferdighet var egalitarismen, differanseprinsippet, fordeling av resurser, fortjeneste og libertarianismen. Mens egalitarismen og fordelingsperspektivet ser på rettferdighet som likhet, la differanseprinsippet og libertarianismen fokus på at alle individer skal ha like muligheter til å kunne oppnå en posisjon i samfunnet uavhengig den konteksten de ble født inn i. Når rettferdighet blir drøftet videre i oppgaven, er spenningen mellom likhet og individuelle forutsetninger et sentralt poeng. Som et alternativ til debatten om rettferdighet som likhet kontra ulikhet, står fortjenesteprikket. Det er her et spørsmål om hva individet har fortjent i den konteksten eller samfunnet det befinner seg i. Ut ifra dette vises det at spørsmålet om hva rettferdighet er ikke har noen entydig svar. Hva som er rettferdig avhenger av perspektivet du ser rettferdighet fra.

Det andre spørsmålet ”Hvem bestemmer hva som er rettferdig?”, tar utgangspunkt i ulike styringsformer og ulike nivåer som et samfunn består av. Diskusjonen rundt de ulike styringsformene viser at demokratiet, anarkiet og paternalisme har ulike måter å bestemme rettferdighet på. I demokratiet er det folket som bestemmer rettferdighetssyn igjennom stemmerett og i anarkiet er rettferdighetsbegrepet individuelt. Hver og en av disse styreformene kan ses i kontekst av ulike samfunnsnivåer. I delen som diskuterer rettferdighet i forhold til samfunnsnivåer, blir det vist at hvem som bestemmer rettferdighet også avhenger

av typen samfunnsnivå du befinner deg i på. For eksempel kan skole og hjem ha ulike syn på hva som er rettferdig.

Oppsummert viser kapitlet at rettferdighet er et komplekst fenomen, med mange teorier og perspektiver å ta hensyn til. Perspektivene som her har blitt presentert og diskutert blir videre i oppgaven benyttet for å drøfte rettferdigheten i elevvurderingen.

5. Rettferdighet og metode for elevvurdering

"If (...) methods apply to all programs evaluated, and errors are randomly distributed across populations, then these errors are simply errors of method, not injustice. If injustices exist, it must be that the interests of certain groups of people are systematically abused"(House 1993:137).

Hvilken metode skal benyttes for å oppnå en rettferdig elevvurdering? I følge Kvernbekk er "en av de viktigste begrunnelsene for bruk av metode(...)å sikre holdbarheten av de påstander vi kommer med" (Kvernbekk 2002:19). Når det snakkes om metode i elevvurderingen, snakkes det derfor om hvilken holdbarhet lærerens vurderinger har for eleven. Holdbarheten i elevvurderingen kan også ses på som et kriterium for rettferdighet, i den forstand at en rettferdig elevvurdering også bør være en holdbar elevvurdering. Det kan stilles flere spørsmål rundt holdbarhet. Hva vil det si at noe er holdbart? Hvordan oppnås holdbarhet? Og sikrer holdbarhet rettferdighet?

I dette kapittelet vil metode bli drøftet som et redskap for å oppnå en rettferdig elevvurdering. Kapittelet består av to delkapitler og en avslutning. I det første delkapittelet blir kvalitativ og kvantitativ metode drøftet. I det andre delkapittelet stilles det spørsmål om vurderingsmetoden klarer å få fram elevens kunnskap.

Det finnes en rekke ulike metoder for å sikre holdbarhet. I det første delkapittelet skjer drøftingen på et generelt nivå i forhold til kvalitative og kvantitative metoder. Kvantitativ metode forstås som tilnærminger i vurderingen hvor fokuset er på data som kan telles og måles (Scriven 1991:296) og kvalitativ metode forstås som vurderinger gjort på grunnlag av forståelse og beskrivelse (Scriven 1991:293) av elevens læringsprosess. Både kvantitative og kvalitative metoder for vurdering, kan forstås som redskaper som preger elevens læringsprosess. Kvalitative og kvantitative metoder har ulike fremgangsmåter for å sikre metodens holdbarhet. Når dette skjer i forhold til elevvurdering er det viktig å ta hensyn til hvilken metode som blir benyttet under målingen av eleven, samtidig som det er viktig å se på hvilken metode læreren bruker i selve vurderingen av eleven. Metode blir her et redskap for rettferdig elevvurdering.

5.1 Kvalitativ og kvantitativ metode

Metodebegrepet kan forstås i forhold til kvalitativ og kvantitativ metode. Ulike metoder ligger til grunn for ulike resultater. I dette delkapittelet vil det først kort bli repetert hva som har blitt sagt om kvalitativ og kvantitativ metode tidligere i oppgaven, så blir det vist hvordan kvalitativ og kvantitativ metode påvirker til rettferdig elevvurdering. Til slutt vil det skje en drøfting om bruken av karakterer som kvalitativ og kvantitativ metode.

5.1.1 Kvalitative og kvantitative metoder for prøveform og vurdering – hva preger dem?

Kvalitet og kvantitet ble sist tatt opp i kapittel 2.2. Der ble det vist at det er viktig å skille mellom kvalitative og kvantitative *prøveformer* og kvalitative og kvantitative *vurderinger*. Jeg begynner med å studere kvalitative prøveformer og kvalitative vurderinger, kommer deretter til kvantitative prøveformer og kvantitative vurderinger. Helt til slutt vil jeg drøfte hvordan kvalitative og kvantitative prøveformer krysses med kvalitative og kvantitative vurderinger. Rettferdighet blir her drøftet i forhold til prinsipper om egalitarisme og differanse.

I det følgende vil to spørsmål om kvalitativ metode bli vektlagt. Hva er en kvalitativ prøveform og hva er en kvalitativ vurdering? En kvalitativ prøveform kan være alt fra essays, noveller, masteroppgave i pedagogikk til muntlig eksamen. Denne prøveformen kjennetegnes ved at spørsmålene ikke har noe gitt svar og at eleven selv må konstruere og kommunisere forståelsen for emnet han eller hun blir målt i (Lund 2002). En kvalitativ prøveform ber om kvalitative svar. En kvalitativ vurdering er i likhet med prøveformen ikke gitt. Dette er i den forstand at den som vurderer eleven ikke kan regne seg fram til vurderingen, men må ta utgangspunkt i egne tolkninger av besvarelsen. Tilbakemeldingen her er i stor grad subjektiv. Den kvalitative vurderingsmetoden baseres på hva læreren synes om elevens prestasjoner og den kvalitative vurderingen kan gis muntlig så vel som skriftlig.

Et eksempel på kvalitativ vurderingsform kan være ”tilbakemeldinger”. Tilbakemeldinger kan gis både skriftlig og muntlig. Hvilke kriterier bør ligge til grunn for at en kvalitativ

tilbakemelding er god? I heftet til prosjektet *Bedre Vurderingspraksis* står det: ”Det er flere oppfatninger av hva som er en god faglig tilbakemelding, men ingen er uenig i at det handler om å være så faglig konkret og ærlig at elevene forstår hvor godt eller dårlig de ligger an i faget”(Bedre Vurderingspraksis 2007:3). Dette poenget kan også ses i forhold til hva som er en rettferdig elevvurdering. I likhet med en god faglig tilbakemelding, bør også en rettferdig elevvurdering ta utgangspunkt i at eleven skal forstå hvordan de ligger an i faget. Men en kvalitativ vurderingsform har gjerne mer for seg enn å fortelle eleven hvordan de ligger an i faget. Den kvalitative vurderingsmetoden har også til hensikt å hjelpe eleven utvikle seg i faget. Utdanningsdirektoratet skriver ”Tilbakemeldinger er viktig for å fremme læring” (Larsen 2009:18). Dette står også uttrykt i Stortingsmelding nr 16 (2006-2007): ”Raske og konstruktive tilbakemeldinger underveis i arbeidet er en forutsetning for at læreprosessen skal være fruktbar”. Selv om tilbakemeldingene har til hensikt å fortelle eleven hvor han står faglig sett og videreutvikle hans forståelse og kompetanse for videre arbeid, kan også en kvalitativ vurderingsmetode virke mot sin hensikt. En ”feil” tilbakemelding kan gi umotiverte elever som ikke har lyst til å lære (Cullingford 1997:1). Cullingford påpeker faren ved at tilbakemeldinger kan gi umotiverte elever, men er en tilbakemelding som ikke motiverer eleven nødvendigvis en urettferdig elevvurdering? Svaret på dette spørsmålet avhenger av prinsippet som belyser hva rettferdighet er. Dersom rettferdighet ses fra egalitarismens ståsted og alle elever skal vurderes likt, kan ikke hensyn til elevens motivasjon være av betydning for om elevvurderingen kan regnes for å være rettferdig eller ikke. På den annen side ser en på differanseprinsippet og fortjeneste prinsippet kan eleven ut ifra prinsipper om tilpasset opplæring og egeninnsats hevde at en tilbakemelding ikke er rettferdig dersom den ikke fremmer motivasjon.

Metoden som blir benyttet i vurderingen av eleven kan ses på som et verktøy læreren kan benytte for å gi en objektiv og rettferdig tilbakemelding (House 1980:215-224). Hvordan denne metoden blir brukt kan få konsekvenser for elevens læringsprosess. En kvalitativ tilbakemelding kan til fremme elevens læring, men den kan også føre til demotiverte elever som ikke ser gleden i å lære. Rettferdighetsdiskusjonen blir her et spørsmål om hvilken metode som kan benyttes for at eleven skal få en rettferdig vurdering. I en slik diskusjon kan rettferdighetsprinsippene om egalitarisme og differanse være relevante. Jeg kommer tilbake til dette etter jeg har tatt for meg kvantitativ metode.

Kvantitativ metode handler om resultater som kan telles eller måles (Kleven 2002). I skolen benyttes også kvantitative metoder for å vurdere elevene. Disse metodene baserer seg på et samsvar mellom kvantitative prøveformer og kvantitative vurderinger. Det som kjennetegner kvantitative prøveformer er at svarene enten er riktige eller gale. Dette gir grunnlag for en kvantitativ vurderingsform hvor en karakter settes på grunnlag av utregninger som tar utgangspunkt i de kvantitative prøveformene.

To eksempler på kvantitative prøveformer i skolen er *multiple-choice* (Almås 1993:49) og matteprøver som krever presise utregninger. I disse to prøveformene er svarene gitt. En *multiple-choice* prøve kan minne om et spørreskjema brukt i forskning. Spørsmålene som stilles har gitte svaralternativer som eleven krysser ut i henhold til om det er riktig eller galt. Det er delte meninger rundt *multiple-choice* som prøveform. Prøveformen har positive sider i at den er enkel å vurdere. Det er bare å telle opp de riktige svarene og regne ut karakteren. Striden rundt *multiple choice* er derimot ikke basert på hvor enkel den er å vurdere, men blant annet hvorvidt den viser elevens kunnskap og legger grunnlag for videreutvikling (Scriven 1991: 236). En kvantitativ prøvemethode gir et utmerket utgangspunkt for en kvantitativ vurderingsmetode. En kvantitativ vurdering bør ta utgangspunkt i kvantitative prøveformer hvor det regnes ut en prosentandel som tilsvarende tilbakemeldingen på hva eleven har gjort riktig og galt. Jeg skal videre drøfte kvalitativ og kvantitativ metode i forhold til rettferdighet, sett under egalitarisme og differanse prinsippet.

I kapittel 4 ble egalitarismen og differanse prinsippene presentert som motpoler. Egalitarismen bygger på en tanke om at alle har lik rett på å få tildelt like mye (Lamont og Favor 2007), mens differanseprinsippet tar utgangspunkt i at alle har rett på å få tildelt resurser ut i fra hvilke behov de har (Lamont og Favor 2007). I utgangspunktet er både egalitarismen og differanseprinsippet rettferdighetsteorier konstruert for økonomisk fordeling (Lamont og Favor 2007). Når vi her drøfter kvalitativ og kvantitativ metode som redskap for en rettferdig elevvurdering, blir spørsmålet om rettferdighet et spørsmål om hvordan vurderingsmetoden skal fordeles blant elever. Skal alle elevene vurderes under samme metode, eller skal metodene som benyttes tilpasses elevens ferdigheter og behov? Kanskje passer kvantitative metoder bedre enn kvalitative metoder for noen elever. Egalitarismen vil etter prinsipper om likhet (Lamont og Favor 2007), hevde at det er rettferdig at alle elever skal bli vurdert under samme metode. I den *constitutive* egalitarismen som er et litt mer åpent syn

på egalitarisme, så ville argumentet for at alle elever skulle vurderes med samme metode, bli argumentert med at "certain inequalities are unjust" (Gosepath 2007). Det er altså urettferdig å forskjellsbehandle og hjelpe noen elever mer enn andre. Differanseprinsippet ville sagt seg uenig i den egalitaristiske argumentasjonen og hevdet at den kvalitative metoden måtte tilpasses elevens ferdigheter og behov, slik at alle elevene til slutt lå på et likt nivå.

Argumentasjonen ovenfor viser at rettferdighet ikke nødvendigvis avgjøres i forhold til hvilken metode som benytter, men i forhold til hvilket rettferdighetsprinsipp vi baserer metoden på. Egalitarismen og differanseprinsippet var bare to av en rekke prinsipper som kunne vært diskutert. Videre skal kvantitative prøver og kvantitative vurderings metoder drøftes.

Det har her blitt vist at kvantitative og kvalitative metoder kan benyttes i elevvurderingen. Bruken av kvantitative og kvalitative metoder er viktig for hvordan eleven blir vurdert og hvorvidt vurderingen kan oppfattes som rettferdig. Når rettferdighet skal diskuteres i forhold til metode, er det allikevel ikke nødvendigvis kvaliteten på metode som ligger til grunn for å avgjøre om metoden er rettferdig (selv om kvaliteten også er viktig). Vurderingsmetode kan drøftes igjennom ulike perspektiver for rettferdighet. Jeg valgte her å drøfte vurderingsmetodene i forhold til egalitarismen og differanseprinsippet. Det ble vist at egalitarismen anser metoden som rettferdig dersom alle elever blir vurdert igjennom en lik vurderingsmetode, mens differanseprinsippet ønsker å tilpasse metoden etter elevens ferdigheter og behov, slik at alle kan oppnå et likt resultat. Videre skal karakterer bli drøftet som en kvalitativ og kvantitativ vurderingsform. Er karakteren rettferdig?

5.1.2 En karakterdebatt

Karakterer kan være både kvalitative og kvantitative. I denne delen skal jeg vise hvorfor karakterer både har kvalitative og kvantitative egenskaper. Jeg ønsker å problematisere synet på karakterer som en kvantitativ vurderingsform og drøfte dette opp mot rettferdighet både med tanke på individet og samfunnet. En karakter kan forstås som urettferdig i de tilfellene den ikke stemmer overens med elevens prestasjoner. Dette gir en hypotese om at en kvantitativ prøveform gir mer grunnlag for en riktig karakter enn en kvalitativ prøveform.

Dette behøver derimot ikke å bety at den kvantitative prøveformen gir et mer rettferdig grunnlag for en karakter enn den kvalitative prøveformen.

I diskusjonen mellom kvalitativ og kvantitativ vurderingsmetode, er det et spørsmål om hva en karakter er. Er karakteren kvalitativ eller kvantitativ? Steinar Kvale og Michal Scriven har noe ulike perspektiver på hva som betegner en karakter. Steinar Kvale skriver: "Å karakterisere vil si å beskrive de egenskaper som anses vesentlig ved noe" (Kvale 1970:127). Ut ifra Kvale sin definisjon knyttes ikke karakteren betingelsesløst opp mot kvantitative vurderingsformer, tvert om ser det ut til at karakteren er kvalitativ. Scriven knytter karakterer opp mot symbolbruk. Han beskriver karakter som "Allocating individuals to an ordered (usually small) set of named categories, the order corresponding to merit – for example, A-F for "letter grading"(Scriven 1991:183). Scriven sitt perspektiv skiller seg fra Kvale sitt i den forstand at karakterer knyttes opp mot symboler eller tegn som beskriver hvilket nivå eleven ligger på. Fokuset Scriven legger på verdier, viser allikevel at det ligger en form for vurdering bak. Denne vurderingen kan ha bakgrunn i kvalitative og kvantitative metoder. Scriven og Kvale viser en forskjell på hvilke karaktertrekk en kan beskrive elevens prestasjoner igjennom og karakterer slik vi kjenner dem i skolen. Synene til Kvale og Scriven legger vekt på at det kan være et kvalitativt utgangspunkt for karakteren, videre skal karakterens kvantitative funksjon diskuteres.

Jeg ønsker å skille mellom en ekte kvantitativ karakter og en kvantifisering av kvalitative resultater. En ekte kvantitativ karakter tar utgangspunkt i kvantitative metoder hvor svarene kan regnes for å være rette eller gale. Kvale skriver om kvantifisering av kunnskap(Kvale 1970). Dette innebærer noe tvetydighet. Et relevant spørsmål er hva slags kunnskap gjelder det? I den norske skole gis det karakterer med utgangspunkt i både kvalitative og kvantitative metoder. Kvale skriver at "Kvantifisering av kunnskap og tenking gjør de enkelte kandidaters prestasjoner sammenliknbare" (Kvale 1970:129). Karakteren kan altså ha en kvantitativ funksjon, men gjør det nødvendigvis karakteren kvantitativ? Det åpenbare svaret på dette er nei. En karakter kan kun være kvantitativ dersom den tar utgangspunkt i kvantitative metoder. Når metoden er kvalitativ vil også karakteren ha en kvalitativ betydning. Det vil si at karakteren betyr noe mer enn en svarprosent på en kvantitativ prøve.

Black m. fl. kritiserer bruken av karakterer. De hevder at karakterer ikke gir læringsfremmende tilbakemeldinger (Black 2003). Dette støttes også av Tveit m. fl. i boka

”Elevvurdering i skolen. Grunnlag for kulturendring”(Tveit 2007). I følge boka bør fokuset i skolen ligge på konstruktive tilbakemeldinger framfor karakterer (Tveit m.fl. 2007). Konstruktive tilbakemeldinger baserer seg på kvalitative metoder og sier langt mer enn en karakter gjør om hvordan eleven ligger an. Det er allikevel fordeler ved å bruke karakterer. En positiv side ved karakterer er at de representerer et felles uttrykk for elevens kompetansenivå. At uttrykket er felles, innebærer at det er det samme karaktersystemet som blir brukt på gitte klasstrinn. Kvantifiseringen av karakterer gjør at elevs prestasjoner kan sammenlignes på kryss og tvers. Eleven kan selv sammenligne karakterer med tidligere karakterer, lærer kan sammenligne elevens karakterer med andre elevs karakterer og på samfunnsnivå kan elevenes karakterer gi uttrykk for nasjonalt kunnskapsnivå (Kvale 1970:129). Dette er svært anvendelig for både arbeidsgivere og når eleven søker på videreutdanning. Utfordringen ved karakterer er at de gjerne baserer seg på lærerens subjektive kriterier og meninger. De subjektive meningene er en trussel både for rettferdigheten og metoden til elevvurderingen.

Spørsmålet om hva som skiller karakterer fra hverandre er en annen utfordring, spesielt i forhold til kvalitative prøveformer. I en kvantitativ prøveform, kan man som vist ovenfor, bruke en metode for å regne seg fram til en riktig svarprosent og gi karakter ut ifra dette. I en kvalitativ prøveform er ikke dette mulig. Hva skiller en firer fra en femmer og en treer? I følge opplæringsloven § 4-8 uttrykker karakteren 3 at ”eleven har nokså god kompetanse i faget”, mens karakteren 4 uttrykker at ”eleven har god kompetanse i faget” og karakteren 5 uttrykker at ”eleven har mykje god kompetanse i faget” (Opplæringsloven § 4-8). Hvordan vet læreren hva som er god kompetanse i et kvalitativt arbeid? Ville en annen lærer vurdert oppgaven likt eller ville eleven fått et annet resultat? Dette er viktige spørsmål også i forhold til rettferdighet. Rettferdighetsteori som bygger på ideer om likhet og ulikhet (se Lamont og Favor 2007) står sentralt i forhold til karaktersetting. Karakterer bør gis på likt grunnlag, men også dels ta hensyn til individets ferdigheter og behov. Er det rettferdig at en og samme oppgave ville fått to ulike karakterer, dersom den ble gitt til to forskjellige sensorer? Det åpenbare svaret her vil være nei. Nei, det er ikke rettferdig at en besvarelse blir vurdert ulikt av forskjellige sensorer.

Jeg skal i det følgende drøfte rettferdighet i kvantifisering av kvalitative prøvemetoder. I mange tilfeller blir det gitt kvantitative vurderinger på bakgrunn av kvalitative prøveformer. Hvor sann eller rettferdig kan en kvalitativ vurdering bli? Eisner stiller seg i boka *“The art of*

educational evaluation. A personal view” kritisk til hvordan et minimalistiske uttrykk, som en karakter kan representere en sannhet. Eisner skriver: “The less expressive a description is, the truer it appears to be – at least in our culture. Yet if there is an absence of emotional or qualitative content, the description risks leaving out more of what is important in the classroom” (Eisner 1985:158-159). Bak en karakter ligger det som regel en kvalitativ vurdering. Dette gjør at karakterens innhold blir diffust, hva gjør en treer til en treer og en firer til en firer? Eisners poeng viser at karakterer har mer enn en kvantitativ funksjon, det ligger også en kvalitativ vurdering bak den. Karakterens betydning kan diskuteres i forhold til egalitaristiske prinsipper. Egalitarismen som verdsetter likhet (Lamont og Favor 2007), vil også hevde at for at en karakter skal være rettferdig bør karakteren ha lik betydning på tvers av klasser og skoler. Dette er allikevel ikke alltid tilfelle. Liv Berit Tessem kritiserer i en artikkel i aftenposten, karakterer for ikke å ha lik betydning på tvers av skoler (Tessem 2009). En annen artikkel skrevet av Astrid Falck Olsen, påpeker at standpunkt karakterer og eksamenskarakterer kan variere som resultat av juks og ordninger for vurdering som ikke vektlegger lik over hele landet (Olsen 2009). Olsen skriver: ”Vi trenger tydelige regler for gjennomføring av muntlig eksamen, der det er elevens og ikke omgivelsenes kunnskaper som skal testes, og sist, men ikke minst, vi trenger et direktorat med vilje og evne til å sørge for at ordningene gjennomføres likt over hele landet” (Olsen 2009). Ut ifra artiklene i aftenposten virker det som at det norske karaktersystemet har et egalitaristisk ideal, men at det overhode ikke lever opp til dette. At karakterer ikke har lik betydning mellom skoler og på tvers av klassetrinn, fører til at karakterer ut ifra en kvalitativ vurderingsform ofte blir opplevd som urettferdig.

Helt til slutt vil jeg diskutere en den kvantitative vurderingsformen som grunnlag for karakter. I motsetning til den kvalitative vurderingen baserer kvantitative vurderingsformer seg på utregning (Kleven 2002). Denne utregningen er lik for alle lærere. Noe som igjen gir likt resultat mellom skoler og på tvers av klasser. Det er allikevel ikke nødvendigvis slik at den kvantitative prøve og vurderingsformen er å anbefale ovenfor den kvalitative. Selv om en kvantitativ prøveform møter premissene egalitarismen stiller for likhet (Lamont og Favor 2007), møter den kvantitative prøveformen andre perspektiver som truer dens rettferdighet. Eksempler på dette er prinsipper om differanse og fortjeneste. Jeg kommer tilbake til dette i del 7.1.2 som tar for seg sluttvurdering.

Det har nå blitt vist at karakterer kan gis ut ifra kvalitative og kvantitative metoder i tillegg til at karakterer kan ha en kvalitativ og kvantitativ funksjon. Egalitarismen ser ut til å være et rådende syn når karakterer skal vurderes som rettferdige eller ikke. I avisene diskuteres det hvorvidt karakterer har lik betydning mellom skoler og på tvers av klasser. Svaret avisene kommer fram til er nei. Det er allikevel ikke slik at ingen karakterer er rettferdig. Dersom en legger andre kriterier til grunn for rettferdighet, er det meget mulig at vi vil få andre svar. For eksempel kan en stille spørsmål om karakteren var fortjent eller ikke? Jeg kommer tilbake til spørsmålet om karakterer og fortjeneste i kapittel 7.1.2.

5.2 Klarer vurderingsmetoden å få fram elevens kunnskap?

Når en rettferdig elevvurdering skal drøftes, må det stilles spørsmål både i forhold til metoden for prøveform og metoden for vurdering. Det første som vil bli drøftet i dette kapitlet gjelder hvilken vurderingsmetode læreren bruker. Klarer vurderingsmetoden å få fram elevens kunnskap? Dersom vurderingsmetoden ikke klarer dette, legges det da et lite rettferdig grunnlag for elevvurderingen? Det andre som skal drøftes er kvalitet og kvantitet i elevvurderingen. Er det rettferdig at kvantitative vurderinger gis på grunnlag av kvalitative prøveformer og vice versa?

Både kvalitative og kvantitative prøveformer har til hensikt å vise hva eleven kan. Dette fordrer spørsmål om hva eleven faktisk vet og hvilke deler av denne kunnskapen som lar seg vurdere. Er det rettferdig å måle eleven etter en prøveform som ikke viser hva eleven kan? Bør ikke en rettferdig prøveform ha som mål å vise elevens kunnskap innenfor det området den er ment å måle? Og kan ikke elevvurderingen kun bli rettferdig i den grad prøveformen måler elevens faktiske kunnskap?

I diskusjonen nedenfor vil det bli vist at ulike elever har behov for ulike prøve og vurderingsmetoder. Denne diskusjonen tar utgangspunkt i en vurdering basert på differanseprinsippet for rettferdighet. Differanseprinsippet baserer seg på at mennesker har rett på en fordeling som gjør dem likeverdige andre mennesker (Lamont og Favor 2007).

For noen elever vil en kvantitativ prøveform (i form av *multiple-choice*) være mer gunstig enn en kvalitativprøveform (i form av stil). Det er viktig å påpeke at det også er store forskjeller i hvordan lærere utformer prøvene sine. Ulike metoder for prøver kan påvirke elevens opplevelse av rettferdigheten i både prøveformen og elevvurderingen. Både Eisner og House er opptatt av at evalueringen skal tilpasses den tenkte målgruppen (Eisner 1985, House 1977). Eisner skriver: "Since different audiences have different needs and different aims, and since the goal of educational criticism is to contribute to the improvement of practice, the criticism's is to contribute to the improvement of practice, the criticism's relation to the audience is a crucial consideration" (Eisner 1985:184). Dette kan også ses i forhold til rettferdig elevvurdering. I en rettferdig elevvurdering bør tilbakemeldingen, enten den er kvalitativ eller kvantitativ, tilpasses eleven. Det kan her stilles spørsmål ved om det er mulig å tilpasse en kvantitativ vurdering til eleven, da den kvantitative vurderingen kun sier noe om hvordan forholdene er, framfor hvordan de bør være (Kleven 2002). Det handler om å kommunisere vurderingen på en måte som appellerer til elevens utviklingszone (Miller 2002). Med dette mener jeg at kvalitative vurderinger bør tilpasses elevens forutsetninger og behov. Dersom eleven ikke forstår hva vurderingen innebærer er tilbakemeldingen mer eller mindre verdiløs. House støtter Eisner i sitt perspektiv på å tilpasse evalueringen til publikummet når han skriver: "Thus the situation the evaluator faces is almost always an appeal to particular audiences which he can define with some precision. If he cannot define his audiences, the evaluation is indeterminate"(House 1977:9). Eisner og House er begge opptatt av hva publikummet har å si for evalueringen. En rettferdig elevvurdering ble her sett i forhold til forskjellene blant ulikt publikum eller ulike elever. Dette er som nevnt ovenfor, et utgangspunkt hentet fra et differanseprinsipp om rettferdighet.

Tilpasset opplæring (Dale og Wærnes 2006) blir et viktig prinsipp for en rettferdig elevvurdering i forhold til hver enkelt elev. På den annen side er det viktig å huske at for mye tilpassing også kan skape store forskjeller blant elevene som igjen vil kunne oppleves som urettferdig blant mange elever. Eisner og House sine utsagn ovenfor om at evaluering må tilpasses sitt publikum er derfor kun gyldig til en viss grad. Dette kan anses som en kritikk av differanseprinsippet. Dersom lærer blir for engasjert i den enkelte elev, er det fare for at rettferdighet kun ses fra læreren og elevens ståsted og ikke samfunnets slik som det egalitærrettslige rettferdighetsbegrepet legger opp til (Lamont og Favor 2007). Rettferdighet i elevvurderingen bør ses i en posisjon mellom individ og samfunn, likhet og ulikhet.

Lucas ser på rettferdighetsbegrepet som dynamisk. I følge Lucas står rettferdighetsbegrepet i et spenningsforhold mellom det individuelle perspektivet og samfunnets perspektiv (Lucas 1989:18). Lucas argument kan forstås som at rettferdighet innebærer å ta hensyn til både individet og samfunnet. Spørsmålet om ulike vurderingsmetoder klarer skape rettferdig elevvurdering, kan også ses i lys av denne debatten mellom samfunn og individ. For at kvalitative og kvantitative metoder skal framstå som rettferdige, bør de ta hensyn til elevens ferdigheter, behov og faktorer som taus kunnskap. Dette er allikevel ikke nok. Som det ble vist ovenfor bør en rettferdig elevvurdering også være rettferdig i en samfunnskontekst. Rettferdighet i elevvurderingen er ikke alene basert på individets forutsetninger, men et samspill mellom hva som er rett for eleven og hva som er rett for samfunnet (Lamont 2007). Et eksempel på hvordan vurderingsmetodene varieres med tanke på rettferdighet hos samfunn og individ, er bruken av interne versus eksterne sensorer. Mens interne sensorer har mulighet til å ta hensyn til elevens ferdigheter og behov, har den eksterne sensoren ingen forutsetninger som gir grunnlag for en slik vurdering. Ekstern sensor handler på et såkalt nøytralt grunnlag og er derfor kun forpliktet til å vurdere hva eleven har prestert, framfor faktorer som innsats. I det neste avsnittet vil taus kunnskap bli problematisert. Jeg anser taus kunnskap som en trussel for rettferdighetsbegrepet, samtidig som det vanskelig lar seg vurdere metodisk sett.

5.2.1 Taus kunnskap og metodetriangulering

I dette delkapittelet blir taus kunnskap sett på som en trussel for en rettferdig elevvurdering, mens metodetriangulering blir drøftet som et redskap for å redusere den trusselen taus kunnskap utgjør i elevvurderingen.

Taus kunnskap er på mange måter et kontroversielt tema. Britt Ulstrup Engelsen skriver om taus kunnskap. Taus kunnskap kan forstås som kunnskap eleven sitter inne med, men ikke klarer å uttrykke skriftlig eller verbalt (Engelsen 2006). I elevvurderingen er taus kunnskap problematisk i den forstand at den vanskelig lar seg måle og gir lite grunnlag for en rettferdig vurderingspraksis. Vurderinger som gjøres med hensyn til taus kunnskap, baserer seg ofte på synsing og lærerens kjennskap til eleven (Engelsen 2006). Slik sett har vurderinger av taus kunnskap dårlig validitet og gir et svakt grunnlag for en rettferdig elevvurdering. En lærer som ikke har kjennskap til eleven, for eksempel en ekstern sensor, vil ikke ha forutsetninger

til å vurdere elevens tause kunnskap og blir derfor ikke tatt med som et alternativ i den følgende diskusjonen.

Metodetriangulering beskriver hvordan ulike metoder kan benyttes for å måle begreper eller situasjoner fra ulike vinkler (Lund 2006). Trianguleringsbegrepet benyttes ofte i forskning, men kanskje ikke så mye i forhold til elevvurdering. Dette betyr allikevel ikke at denne metodiske tankegangen ikke kan gjøre seg gjeldende i elevvurderingen. Å måle elevens kunnskap ved hjelp av flere metoder kan gi et mer helhetlig bilde på hva eleven faktisk kan (House1980:215). Triangulering kan slik brukes som et redskap for vise elevens tause kunnskap og legger grunnlag for en mer rettferdig vurdering. Videre vil jeg vurdere to eksempler på hvordan taus kunnskap har blitt målt på ulike arenaer.

Eisner og Kryger gir begge eksempler på hvordan taus kunnskap ikke kommer til uttrykk. I det følgende blir først Eisners perspektiv på hvordan visuelle bilder ikke kan gjenspeiles i teori tatt opp. Perspektivet blir drøftet i forhold til triangulering og en rettferdig elevvurdering i form av egalitarisme og differanseprinsippet. Deretter blir Krygers ”machogutter”/ ”*the lads*” tatt opp som et eksempel på hvordan taus kunnskap kan komme til uttrykk igjennom alternative metoder.

Eisner skiller mellom kunnskap som visuelt og kunnskap som kan ordlegges. I et eksempel på hva som lar seg beskrive og hva som ikke lar seg beskrive, skildrer Eisner denne utfordringen. “Your knowledge of your mother’s face and my knowledge of mine are represented in visual images. What we know in this form is made possible through our sense of sight” (Eisner 1985:149). Eksempelet er ment til å vise en forskjell på hva vi ser og hvordan vi klarer å artikulere det. Hvordan kan kvalitative og/eller kvantitative vurderingsmetoder vise elevens kunnskap? Eisner legger selv fram forslag i form av kvalitative metoder. Han foreslår bruk av modellering, tegning, verbale beskrivelser osv. Disse metodene for å vise kunnskap om et ansikt kan være virkningsfulle dersom eleven er god til å modellere, tegne eller uttrykke seg verbalt. Men selv om eleven har disse ferdighetene vil skissen de lager eller beskrivelsen de artikulerer kun være et mer eller mindre godt speilbilde på den kunnskapen og det minnet han faktisk har av sitt mors ansikt (Eisner 1985). Dersom eleven mangler disse ferdighetene vil kunnskapen forbli taus kunnskap. I et slikt tilfelle blir rettferdigheten vanskelig å definere. Hvordan kan en vurdering vise noe som læreren ikke får målt og på en rettferdig måte bli et uttrykt i elevvurderingen?

Rettferdighetsprinsippet kan som vist i kapittel 4 ses som en rekke prinsipper. Under rettferdighetsprinsippet ligger blant annet forskjeller mellom differanseprinsippet og egalitarismen. Når taus kunnskap skal vurderes, må det tas hensyn til at dette er vurderinger som ikke har noen rot i annet enn hva læreren tror eleven kan. Differanseprinsippet spiller her en kontroversiell rolle. Gosepath hevder at: “The difference principle only allows the talented to earn more to the extent that this raises the lowest incomes” (Gosepath 2007). Når det gjelder den gjennomsnittlige elev blir det et spørsmål om det i det hele tatt er relevant å ta hensyn til at eleven også har taus kunnskap. Men hva med de svake elevene? Kan det være fordelaktig for en rettferdig elevvurdering at lærer vurderer taus kunnskap, i den grad taus kunnskap lar seg vurdere?

Machoguttene til Kryger er enda et eksempel på taus kunnskap. I et forsøk gjort av Kryger blir en gruppe gutter, kalt ”the lads”, tatt med ut for å teste kunnskapene sine på gokartbanen. På banen viser det seg at guttene har utmerket kunnskap om gokart. Når de derimot kommer tilbake til klasserommet, virker det som om guttene ikke klarer å uttrykke kunnskapen de viste på gokart banen (Ulriksen 2003:189). Dette viser at kunnskap guttene i bunn og grunn satt inne med, ble vist igjennom en alternativ metode for vurdering. En såkalt metodetriangulering. Taus kunnskap er en utfordring for en rettferdig elevvurdering fordi den vanskelig lar seg måles. I eksempelet til Eisner og i eksempelet til Kryger vises det hvordan metodetriangulering kan gjøre at taus kunnskap lar seg vurdere. Dersom rettferdighet skal oppfattes som å vise elevens kunnskap bør forskjellige metoder benyttes for å fremme kunnskapen. Eisner poengterer: “We must find a way of transforming the qualitative aspects of our experience into a form that others can “read” (Eisner 1985:154). Ulike metoder for elevvurdering handler om å synliggjøre kunnskap som kanskje ikke kommer fram igjennom andre vurderinger.

Det har til nå blitt vist at kvalitative og kvantitative prøveformer igjennom triangulering kan bidra til å variere vurderingspraksisen. En kvalitativ prøveform kan også skje i form av modellering eller observasjon på gokartbanen. Kvantitative vurderinger kan komme fra matteprøver og *multiple-choice* (Almås 1993). En rettferdig elevvurdering handler også om tilrettelegging av selve testen og å finne hva som appellerer til eleven (Tveit m flere 2007) Det kan tenkes at *multiple-choice* prøver kan hjelpe eleven til å få uttrykt en kunnskap de ellers ikke ville klart å skrive på egenhånd. Dette er et spørsmål om å finne metoden som viser

elevens kunnskap mest nøyaktig. Hvis metoden er riktig, får den rettferdige elevvurderingen et langt bedre grunnlag. I alle tilfeller bør en rettferdig elevvurdering gi uttrykk for hva eleven har prestert og ikke hva lærer tror eleven kunne ha prestert. Verken kvalitative eller kvantitative metoder kan ta fullstendig høyde for taus kunnskap og blir derfor kun verktøy som kan brukes i et forsøk på å få fram denne kunnskapen.

Spørsmålet om vurderingsmetoden klarer å få fram elevens kunnskap avhenger av metoden som benyttes, den som vurderer eleven og eleven selv. Når metode benyttes som redskap for å fremme rettferdig elevvurdering kan det være gunstig å benytte seg av metodetriangulering slik at eleven vurderes ved hjelp av ulike redskaper. Redskapene gjør det mulig å vise mer av elevens kunnskaper. Selv om vurderingsmetoder kan benyttes for å fremme læring og metodetriangulering kan være nyttig for å vise taus kunnskap, gir ikke dette alene et svar på om en rettferdig elevvurdering kan oppnås. Elevvurderingen og metoden må ses i forhold til ulike prinsipper for rettferdighet. Jeg har i dette kapitlet kun valgt å drøfte egalitarisme og differanseprinsippet som mulige løsninger. Mens egalitarismen vektlegger likhet i bruken av metodene, ser differanseprinsippet metode i forhold til elvers forskjeller og behov. Dersom metode triangulering skulle bli belyst fra et egalitaristisk ståsted ville alle elever bli nødt til å ta de samme prøvene, noe som igjen ville ført til en ekstrem bruk av resurser. I et differanseprinsipp ville metodetriangulering kun blitt tatt i bruk dersom det ble ansett som nødvendig.

5.3 Avslutning

I dette kapitlet har det blitt satt fokus på kvalitative og kvantitative metoder som hjelpemidler for å oppnå en rettferdig elevvurdering. Det har blitt vist at selv om kvalitative og kvantitative metoder kan fungere som hjelpemidler til å oppnå en rettferdig elevvurdering, behøver ikke disse metodene være rettferdige. Rettferdighet må ses i forhold til hvilke prinsipper for rettferdighet som ligger til grunn. Jeg har i dette kapitlet tatt utgangspunkt i rettferdighetsprinsipper for egalitarismen og differanseprinsippet.

Igjennom egalitarismen har jeg vist at en rettferdig metode for elevvurdering kan ses på som en metode som er felles og lik for alle. Dette er et rådende syn, spesielt når det kommer til karakterer. I karakterdebatten er det allikevel slik at selv om egalitarismen setter idealet for hva som er en rettferdig karakter, stemmer ikke dette overens med hva som skjer i praksis. Dette gjør at karakter ikke alltid oppleves som rettferdige. Som motpol til egalitarismen har jeg drøftet differanseprinsippet. Mens egalitarismen har sett på rettferdighet igjennom like vurderingsmetoder, ser differanseprinsippet på hvordan vurderingsmetodene kan tilpasses den enkelte elev og gi en vurdering i forhold til elevens ferdigheter og behov. Egalitarismen og differanseprinsippet kommer begge til uttrykk i dagens skole. Karakterdebatten ovenfor viser et ønske om likhet og en egalitaristisk tankegang som skolen så langt ikke har klart å oppfylle. På den annen side kommer differanseprinsippet til uttrykk igjennom retningslinjer for tilpasset opplæring. Svaret på hva som er en rettferdig metode for elevvurderingen står altså i forhold til hvilket prinsipp du legger til grunn for elevvurderingen.

I neste kapittel fortsetter jeg metodediskusjonen, men setter fokuset på gyldighet i elevvurderingen. Er en gyldig vurdering nødvendigvis en rettferdig vurdering?

6. Rettferdig elevvurdering og begrepsvaliditet

*‘Valid’ assessments are thought to capture what they are supposed to measure, namely the actual knowledge and skills which pupils possess” (Davis 1998:149[*min kursiv*]).*

Tanken bak dette kapitlet er å diskutere om en valid elevvurdering også gir en rettferdig elevvurdering. Kapitlet tar for seg begrepsvaliditet som tema og drøfter begrepsvaliditet i forhold til om en rettferdig elevvurdering kan oppnås. Det er tre delkapitler som hver tar utgangspunkt i begrepsvaliditet. I det første delkapitlet går jeg inn på hva begrepsvaliditet innebærer. I det andre delkapitlet blir systematiske og tilfeldige målingsfeil diskutert som trusler for begrepsvaliditet og en rettferdig elevvurdering. Det siste delkapitlet diskuterer vurdererrelabilitet i forhold til tilfeldige målingsfeil og en rettferdig elevvurdering.

6.1 Begrepsvaliditet

Jeg begynner her med å ta for meg begrepet validitet før jeg går nærmere inn og ser på hva som ligger i begrepsvaliditet. Deretter gir jeg et eksempel på hvordan begrepsvaliditet kan være en trussel for en rettferdig elevvurdering. Helt til slutt blir dette drøftet i forhold til libertarianismen og samfunnsrettede rettferdighetsprinsipper.

Validitet kan forstås som gyldigheten til de slutningene vi trekker (Kleven 2002:27). Slutninger kan forstås som den vurderingen lærer gjør av eleven ut ifra prøver eller i klasserommet. Med andre ord handler validitet i denne oppgaven om å gi gyldige vurderinger av eleven. Validitet inngår i både kvalitative og kvantitative metoder og kan deles opp i begrepsvaliditet, indre validitet, ytre validitet og statistisk validitet (Lund 2002:104). I denne

oppgaven blir det satt et særlig fokus på begrepsvaliditet, som er relevant for både kvalitative og kvantitative metoder når en rettfærdig elevvurdering skal drøftes.

Scriven ser på validitet i vurdering: "Valid evaluations are ones that take into account all relevant factors, given the whole context of the evaluation (particularly including the client's needs) and weight them appropriately in the synthesis process. One may or may not build issues of credibility and the appropriateness of the reports into the validity of the evaluation (Scriven 1991:372-373). Scriven får her fram at validitet i vurderingen handler om hele prosessen, ikke bare sluttresultatet. Validitet blir i forhold til elevvurdering et spørsmål om samsvar og gyldighet mellom hva eleven prøves i, produktet eleven leverer og tilbakemeldingen eleven får. I følge Popeham kan ikke en test i seg selv være valid. Popeham skriver: "All validity analysis should center on the test-based inference, rather than on the test itself" (Popeham 2003:43). Fokuset ligger i følge Popeham på elevens besvarelse og lærerens vurdering av besvarelsen framfor på hvordan selve testen er formulert. Er det samsvar mellom dem? Popeham skiller mellom presise og upresise fortolkninger som en del av validitetsproblematikken. Han skriver: "If the interpretation is *accurate*, we say that the teacher has arrived at a *valid* test-based inference. If the interpretation is *inaccurate*, then the teacher's test-based inference is *invalid*" (Popeham 2003:43). Med andre ord blir begrepsvaliditet i elevvurderingen et spørsmål om samsvar mellom det som blir vurdert og tilbakemeldingen. Skal man tro Popeham baserer en valid vurdering seg på en presis beskrivelse av hva som er blitt gjort. Dersom fortolkningen av elevens svar ikke stemmer overens med hva eleven har ment blir vurderingen "*invalid*". Hva har dette å si for rettfærdigheten i vurderingen? Jeg skal videre gå inn å se på begrepsvaliditet i forhold til om en rettfærdig elevvurdering kan oppnås.

Hva er begrepsvaliditet? Begrepsvaliditet handler på den ene siden om hvordan begreper konstrueres (Norris 1983) og på den andre siden om hvordan de operasjonaliseres (Lund 2002:120). Norris skriver om konstruksjon: "Construct validations are conducted to justify the techniques that are used to make inductive inferences about the unobservable attributes of people" (Norris 1983:54). I elevvurderingen vises dette blant annet igjennom hvordan lærer konstruerer begrepene i elevvurderingen. Hva legger læreren i begrepene som blir brukt i elevvurderingen? Operasjonalisering har en litt annen funksjon for begrepsvaliditet. Kleven skriver "Operasjonalisering er illustrert som forbindelseslinjer mellom det teoretiske ikke-observerbare og det empiriske observerbare plan" (Kleven 2002:143). Med dette mener

Kleven at det kan være en forskjell mellom hvilken betydning som ligger i begrepet og hvilke sider av dette begrepet som faktisk blir målt empirisk. Sidene som blir målt kan både underrepresentere og overrepresentere innholdet som begrepet i utgangspunktet var ment til å måle. I forbindelse med elevvurdering kan et teoretisk ikke-observerbart plan kan forstås som hva lærer ønsker å vurdere eleven i. Det empirisk observerbare plan representerer så hva læreren faktisk vurderer eleven i. Elevens og lærerens forståelse av begrepene er her den største trusselen for om en rettferdig elevvurdering kan oppnås. Jeg skal forklare dette litt nærmere igjennom et eksempel hentet fra Ivar Bjørgen. Eksempelet tar utgangspunkt i en oppgavetekst Bjørgen selv fikk i oppgave å besvare. Jeg har tenkt til å bruke eksempelet for å vise hvordan begrepsvaliditet er viktig for en rettferdig elevvurdering.

Eksempelet Bjørgen viser til tar utgangspunkt i en oppgave tekst han fikk i sine yngre skoledager. Oppgaveteksten lød "Skriv om tømmerstokkens historie" (Bjørgen 2007:121). I et eksempel viser Ivar Bjørgen hvordan han som barn misforstod prøveteksten og på grunnlag av det strøk. Han forteller hvordan han selv skrev om frøet som begynte å gro, spiren som vokste opp, treet som sto der så praktfullt og hvordan det til slutt ble hogget ned og ble til en tømmerstokk. På dette resonnementet strøk han. Læreren begrunnet det med at han skulle ha skrevet om hvordan tømmerstokken var blitt benyttet i industrien opp igjennom årene (Bjørgen 2007:121).

Eksempelet til Bjørgen (2007) viser at det var liten grad av samsvar mellom begrepene slik de var definert teoretisk av læreren, og begrepene slik Bjørgen lyktes å operasjonalisere dem. Begrepene i teksten ble tolket ulikt av ulike parter og førte til at elevvurderingen ikke kunne reflektere elevens kunnskap på området den var ment til å måle. Vil en slik målingsfeil medføre at elevvurderingen ble oppfattet som urettferdig? Svaret på dette vil avhenge av hvilket perspektiv du ser det ifra.

For å oppnå begrepsvaliditet bør det være en felles forståelse av det konstruerte begrepet og begrepet slik det operasjonaliseres (Kleven 2002). I eksempelet fra Bjørgen var det begrepet "historie" som ble tolket ulikt av lærer og Bjørgen selv. Jeg skal videre drøfte begrepsvaliditet i forhold til et libertarianistisk og et samfunnsbasert syn på rettferdighet. Lamont og Favor skriver om libertarianismen: For Libertarians, just outcomes are those arrived at by the separate just actions of individuals; a particular distributive pattern is not required for justice"(Lamont & Favor 2007). Libertarianismen er altså et ekstremt individsentrert

rettferdighetsperspektiv. Ut ifra dette perspektivet er det et spørsmål om Bjørgen oppfattet resultatet som rettferdig. Flere årsaker kan være relevante for om Bjørgen selv så denne vurderingen som rettferdig eller ikke. På den ene side kan det tenkes at Bjørgen ikke så på vurderingen som rettferdig, fordi missforståelsen av oppgaveteksten førte til at han ikke fikk vist den kunnskapen læreren var ute etter å vurdere. Det kan også tenkes at Bjørgen mente besvarelsen var god ut ifra oppgavens egne premisser. På den andre side kan det tenkes at Bjørgen anså vurderingen som rettferdig fordi han innså at besvarelsen hans ikke svarte på det læreren var ute etter når han ga oppgaven. Det individuelle perspektivet lar det være opp til den enkelte elev å avgjøre om vurderingen de får er rettferdig eller ikke. Det individuelle perspektivet står som motpol til samfunnsperspektivet på rettferdighet. Samfunnsperspektivet kan hentes fra egalitarismen. Egalitarismen tar utgangspunkt i spørsmål om hva som er rettferdig for samfunnet og baserer seg ofte på trekk om likhet og fellesskap (Lamont og Favor 2007). Dersom en skal drøfte vurderingen av Bjørgens besvarelse ut ifra egalitarismen er det ikke lenger Bjørgen selv som skal svare på om vurderingen var rettferdig eller ikke, men samfunnet. På den ene siden kan samfunnet si at vurderingen var rettferdig fordi alle elevene ble målt etter samme målestokk og Bjørgen ikke fortjente noen særskilt behandling. På den annen kan samfunnet si at vurderingen var urettferdig med hensyn til at oppgaveteksten var uklar. Dersom teksten ble ansett som uklar ville det kun være rettferdig å gi alle elevene en ny sjanse til å gi en besvarelse på tømmerstokkens historie.

Et libertarianistisk rettferdighetsperspektiv begrenser bruken av felles kriterier som vurderingsgrunnlag. Begrepsvaliditet tar avstand fra libertarianistisk holdninger. Dette skjer ved å legge vekt på at det bør være intersubjektiv enighet rundt tilbakemeldingene (Kleven 2002). Egalitarismen forsvaret i sin tur fellesskapet og da også et intersubjektivt perspektiv (Lamont og Favor 2007). Allikevel er det slik at intersubjektiv enighet ikke nødvendigvis innebærer at alle mener det samme og at alle er enige. Dette gjør at dersom begrepsvaliditet skal forsvares fra et egalitaristisk rettferdighetsperspektiv må alle være enige i den vurderingen som blir gjort.

Ernest R. House vektlegger fellesskapets forståelse når han skriver at "Evaluations should serve the interests not only of the sponsor but also of the larger society, and of various groups within society, particularly those most affected by the program under review"(House 1993:128). Selv om House sitt utsagn gjelder på systemnivå, kan dette også brukes i forhold

til elevvurdering på individnivå. Elevvurderingens betydning, enten den kommer som karakter eller som formative tilbakemeldinger bør også være forståelig for det øvrige samfunn. Spesielt viktig blir da betydningen av karakterer når eleven skal søke videre utdanning eller jobb. For å oppnå en rettferdig og valid elevvurdering, bør begrepene som benyttes i elevvurderingen bygge på et felles sett kriterier for vurderingen. Slik vet både lærer, elev og det øvrige samfunn hva som ligger bak vurderingen.

Begrepsvaliditet handler om gyldigheten til elevvurderingen. Er vurderingen gyldig og på hvilket grunnlag er den gyldig. Jeg har her vist at det ikke alltid er samsvar mellom begrepene lærer ønsker å vurdere eleven i og elevens forståelse av begrepene. Missforståelser og begreper som måler perspektiver de ikke er ment å måle, skaper en trussel for både gyldigheten i elevvurderingen så vel som en rettferdig elevvurdering. Til slutt i denne delen har jeg drøftet libertarianistisk rettferdighetsperspektiver opp mot samfunnsrettede perspektiver. Det ble vist at begrepsvaliditet ikke stemmer så godt overens med libertarianistisk rettferdighets perspektiver. Dette skyldes at mens begrepsvaliditet har krav om en eller annen form for intersubjektiv gyldighet og rettferdighet er libertarianismen på sitt mest ekstreme opptatt av hvordan individet opplever vurderingen. Dersom eleven oppfatter vurderingen som rettferdig, kan vurderingen anses for rettferdig. Samfunnsperspektivet på rettferdighet vil her stå som motpol til individet syn på rettferdighet og si seg enig i begrepsvaliditetens krav om intersubjektivitet. Jeg skal videre gå nærmere inn på systematiske og tilfeldige målingsfeil som trusler for begrepsvaliditet og oppnåelsen av en rettferdig elevvurdering.

6.2 Systematiske og tilfeldige målingsfeil

Jeg vil her betrakte systematiske og tilfeldige målingsfeil som to trusler for begrepsvaliditet i en rettferdig elevvurdering. Fordi rettferdigheten lar seg påvirke av hvordan begrepene er konstruert og operasjonalisert er det viktig å diskutere hvilke trusler som kan reduseres, slik at elevvurderingen kan bli mer rettferdig. I elevvurderingen er det viktig å være bevisst på denne typen trusler fordi de blant annet kan true oppfattelsen av en rettferdig elevvurdering. Jeg

begynner med å studere systematiske målingsfeil. Hva er de og hvordan påvirker de rettferdighet i elevvurderingen? Deretter vil jeg drøfte tilfeldige målingsfeil. Dette gjøres ved å først presisere hva som ligger i en tilfeldig målingsfeil, for så å drøfte hvordan tilfeldige målingsfeil truer en rettferdig elevvurdering.

Når en elev vurderes kan det oppstå systematiske målingsfeil som gjør at eleven blir vurdert på et skjevt eller feil grunnlag. En systematisk målingsfeil beskrives av Kleven som ”et mer eller mindre skjevt bilde av det begrepet vi skal måle, dels fordi vi ikke får med oss hele fylde i begrepet, og dels i at det sniker seg inn noe som ikke hører til begrepet i det hele tatt” (Kleven 2002:122). Det er typisk for de systematiske målingsfeilene at de ikke jevner seg ut i det lange løp og derfor gir et skjevt bilde på målingen (Kleven 2002). Et skjevt bilde av målingen kan oppstå når eleven legger mer eller mindre i vurderingen enn det den som vurderingen hadde til hensikt.

Et annet eksempel kan være at eleven har et språkproblem som gjør det vanskelig å uttrykke seg skriftlig. Dersom læreren ikke er seg bevisst elevens skriftlige språkproblem kan dette føre til en systematisk målingsfeil. Elevens språkvansker kan ses på som en systematisk målingsfeil fordi dette vil gjenta seg hver gang eleven vurderes ut ifra et skriftlig arbeid. Elevvurderingen blir urettferdig i den forstand at den ikke måler elevens kunnskapsnivå slik intensjonene med elevvurdering er. Andre eksempler på systematiske målingsfeil kan være eksamensangst og veldig gode eller svake skriftlige ferdigheter (Kleven 2002:152). Dette er altså ”feil” ved målingen som gjentar seg gang på gang og gjør at eleven ikke får vist sin egentlige kunnskap.

Kleven skriver at det er viktig å være oppmerksom på at validitet ikke knyttes til selve testen, men til tolkningen av resultatene (Kleven 2002:173). Dette kan ses i forhold til Popehams perspektiv som ble nevnt i innledningen. I overført betydning kan dette bli sett på som vurderingen av eleven. Når systematiske målingsfeil skal undersøkes, ligger fokuset på selve vurderingen eller tilbakemeldingen og ikke på hvordan prøven alene er utformet. For å oppnå rettferdighet i elevvurderingen er en felles forståelse mellom lærer, elev og samfunn avgjørende. Det er viktig at alle forstår hva som faktisk menes med vurderingen og unngå misforståelser. Det er også avgjørende at det er samsvar mellom lærers tilbakemelding og elevens produkt, enten det er en prøve, en muntlig tilbakemelding i klasserommet eller et praktisk arbeid eleven har gjort i kunst og håndverk.

Om tilfeldige målingsfeil skriver Kleven: ”Når vi snakker om *tilfeldige* målingsfeil, så betyr ikke det nødvendigvis at feilene skyldes tilfeldigheter. Det betyr at feilene oppfører seg tilfeldig, de følger det som vi av og til kaller tilfeldighetenes lov eller de store talls lov. Slike feil jevner seg ut i det lange løp” (Kleven 2002:121). Eksempler på tilfeldige målingsfeil kan være at eleven er heldig med oppgavene og får oppgaver han har spesialisert seg på eller dagsformen til eleven. Reliabiliteten til læreres vurdering er også en faktor innenfor tilfeldige målingsfeil. Dette kommer jeg tilbake til i delen om vurdererrelabilitet. For at en vurdering skal være rettferdig bør den oppfattes som gyldig i forhold til elevens besvarelse. Det bør også være enighet om hvilke kriterier besvarelsen måles etter.

I denne delen har vi sett at begrepsvaliditet spiller en rolle i en rettferdig elevvurdering. Metoden som benyttes gjør det mulig å forklare ulike former for målingsfeil. Bevissthet hos lærer på ulike metodiske trusler kan være med på å skape en felles forståelse for hvorfor tilbakemeldingen er som den er, og derav være med på å gjøre vurdering rettferdig. I neste del vil diskusjonen om vurdererrelabilitet tatt videre fra tilfeldige målingsfeil.

6.3 Vurdererrelabilitet

I dette delkapittelet skal jeg ta for meg vurdererrelabilitet. Jeg begynner med å se på reliabilitetsbegrepet. Så går jeg inn på vurdererrelabilitet som begrep. Til slutt bli vurdererrelabilitet drøftet i forhold til intersubjektivitet og objektivitet i en rettferdig elevvurdering. Spørsmålet blir om intersubjektivitet og objektivitet kan gjøre at en rettferdig elevvurdering oppnås?

Reliabilitet er viktig for en rettferdig vurdering for å begrense tilfeldige målingsfeil. Feil i elevvurderingen er med på å skape urettferdig vurderingsgrunnlag. Det er derfor viktig at personen som foretar elevvurderingen er til å stole på. I denne delen begynner jeg med å undersøke hva vurdererrelabilitet er. Deretter drøfter jeg momentene subjektivitet og objektivitet i forhold til vurdererrelabilitet og en rettferdig elevvurdering. Til slutt blir det gjort en avrunding og oppsummering.

Reliabilitet handler om i hvilken grad data er fri for tilfeldige målingsfeil (Kleven 2002:154). I følge Kleven er det klassiske reliabilitetsbegrepet knyttet til påliteligheten i målingen hos enkeltpersoner på målingstidspunktet (Kleven 2002:124), mens det i forskningsmetodelitteraturen blir sett på som et ”uttrykk for i hvilken grad data er fri for tilfeldige målingsfeil” (Kleven i Lund 2002:154). Eisner beskriver hvordan reliabilitet kan oppnås. “Reliability, or perhaps a better term, believeability, is achieved through a process of structural corroboration, a process something like ‘triangulation’ in sociological methodology, or circumstantial evidence in law” (Eisner 1985:185). Dersom reliabilitet representerer en tilfeldig målingsfeil hos enkeltpersoner, kan graden av tilpasning framstå som en slik feil. Andrew Davis problematiserer i artikkelen ”Assessment and Fairness” tilpasning som en reliabilitetstrussel for en rettferdig elevvurdering. Davis skriver:

”It is arguably ‘unfair’ to give a severely dyslexic student the same amount of extra time as another who is only mildly afflicted. Some dyslexic students need ‘more’ accommodation than others. Needless to say, it is even *more* unfair to highlight disabilities such as dyslexia if we cannot in principle be clear about that counts as suffering from such a condition, or if related existence claims lack adequate empirical evidence” (Davis 2008:10).

Tilpasning blir fort en trussel for reliabilitet fordi det er vanskelig å vite hvilken grad av tilpasning eleven har rett på. Hvor mye tilpasning skal til for at elevvurderingen blir rettferdig? Popeham skriver: ”*An unreliable test will rarely yield scores from which valid inferences can be drawn*” (Popeham 2003:54). Hvordan kan man sikre seg mot reliabilitetssvikt og vil dette føre til en mer rettferdig elevvurdering? Reliabilitet er viktig for rettferdighet fordi reliabilitet er knyttet sammen med tilfeldige målingsfeil. Selv om tilfeldige målingsfeil jevner seg ut over lenger tid (Kleven 2002), bør ikke slike feil finne sted i elevvurderingen. Tilfeldige målingsfeil kan fort føre til at vurderingen blir sett på som urettferdig. Som eksempel kan vi bruke ”Kristian” som ble dumpet av kjæresten rett før en prøve i samfunnsfag. At Kristian ble dumpet blir her en tilfeldig målingsfeil. Kristian som ellers er en kløpper i samfunnsfag er helt ute av seg og gjør sin dårligste prestasjon i samfunnsfag noensinne. Vil en vurdering etter disse omstendighetene være rettferdig? Ut ifra hva Kristian presterer på prøven kan det tenkes at vurderingen er rettferdig, men ut ifra Kristian sitt faktiske kunnskapsnivå behøver ikke vurderingen være rettferdig. Det kan stilles spørsmål om vurderingen er reliabel i forhold til Kristian sine kunnskaper. Mangel på reliabilitet kan også føre til mangel på rettferdighet, samtidig som rettferdighet må ses i

forhold til den konteksten den befinner seg i. På den ene siden var nok vurderingen av Kristian rettferdig. Kristian presterte dårlig og prøven ble vurdert deretter. På den annen side var ikke prøven reliabel i forhold til Kristian sitt egentlige kunnskapsnivå og kan derfor ikke ses på som en rettferdig indikator på hva Kristian faktisk kan. Det er altså dårlig reliabilitet i vurderingen av Kristian.

Et annet perspektiv på reliabilitet er å rette fokuset mot den som vurderer, framfor den som blir vurdert. Et slikt perspektiv kalles for vurdererrelabilitet. Vurdererrelabilitet handler om ”hvem som ser, tolker og vurderer”(Kleven 2002:126). Det er et spørsmål om den som foretar vurderingen er til å stole på. Har læreren kompetanse til å kunne gi en pålitelig elevvurdering? Rettferdigheten trues her av vurdererrelabiliteten i den forstand at den som vurderer i mer eller mindre grad står ansvarlig for hvorvidt elevvurderingen blir rettferdig. House skriver at “Both formulation and interpretation are subjective processes” (House 1977:14). Dette kan forstås som at begrepsvaliditet i stor grad baserer seg på en subjektiv vurdererrelabilitet. Formuleringen og fortolkningen henger sammen med forklaring og forståelse. House skriver: “Explanation and understanding are functions of the way information is coded in the mind. Explanation implies a person who is understanding the explanation. It does not exist by itself. The understanding is ultimately reducible to something familiar in the mind of the audience doing the understanding – or else it is not an explanation”(House 1977:42). Forståelse er et viktig redskap for å kunne diskutere rettferdighet i en elevvurdering. For at en vurdering skal kunne fremstå som rettferdig er det viktig at både den som foretar vurderingen og den som blir vurdert forstår hva som blir lagt i vurderingen. Dersom det skjer misforståelser kan det hende at en vurdering som i utgangspunktet var rettferdig blir oppfattet som urettferdig, eller omvendt. Vurdererrelabilitet blir i et slikt perspektiv et tosidig spørsmål: Hvordan vurderingen blir utført fra lærers side og hvordan vurderingen blir tolket av eleven. Som House skriver:”In the fullest sense, then, an evaluation is dependent both on the person who makes the evaluative statement and the person who receives it” (House 1977:42). Sett i forhold til om en rettferdig elevvurdering kan oppnås, blir det her viktig å få fram at det bør være en felles forståelse mellom den som gir vurderingen og den/de som mottar vurderingen. Rettferdighet blir her vurdert ut ifra om det er enighet rundt premissene som ligger til grunn, både i forhold til hva rettferdighet er og i forhold til hvilke kriterier elevvurderingen gis på bakgrunn av.

Vurdererrelabilitet er en trussel for rettferdig elevvurdering fordi hvordan en person vurderer ikke alltid samsvarer med virkeligheten. Vurdererrelabilitet kan vise seg ved at ulike subjekter velger å benytte seg av forskjellige kriterier for vurderingen, og vurderingen vil derfor bli opplevd som urettferdig. House skiller mellom perspektivet til den som fortar en evaluering og de som blir evaluert. House skriver: “The audience themselves have images, memories, and theories of the program under evaluation. In using evaluation as a perspective (in this case a verbal model), the audience matches its conception of the program to evaluation” (House 1977:18). Hensikten med sitatet til House, er å vise at ulike subjektive bakgrunnsforståelser for vurderingen gir ulikt grunnlag for å tolke vurderingen som rettferdig eller urettferdig. Spørsmålet er om et subjektivt perspektiv kan tilfredsstillende vurdererrelabiliteten og sikre en rettferdig elevvurdering?

Jeg skal videre diskutere intersubjektiv enighet som kriterium for vurdererrelabilitet. Gjennom Scriven viser House til et skille mellom kvalitative og kvantitative tilnærminger til objektivitet. “In the quantitative sense of the term, one person’s opinion about something is regarded as being subjective – the disposition of one individual” (House 1980:86). Med andre ord er det opp til individet hvordan han eller hun forstår elevvurdering. Dette gjør at hvorvidt eleven opplever en vurdering som rettferdig eller ikke, vil avhenge av eleven og hans bakgrunnsforståelse for å forstå vurderingen. House fortsetter å sitere Scriven når han tar for seg objektivitet som en kvantitativ form. House skriver: “Objectivity is achieved through the experiences of a number of subjects or observers. The common experiencing makes the observation public through intersubjective agreement”(House 1980:86). Intersubjektivitet kan ses på som det nærmeste en kan komme enighet om hva rettferdighet er. For eksempel kan en ta en elev som får karakteren 3 på eksamen. Eleven opplever karakteren som urettferdig og klager. Det har vært to sensorer som har kommet fram til at eleven skulle få karakteren 3 i utgangspunktet og etter klagen blir ytterligere to sensorer enige om at eleven fortjener karakteren 3. En kan si at sensorene har en intersubjektiv enighet rundt at karakteren 3 er en rettferdig karakter for denne eleven. Det kan selvfølgelig tenkes at dersom du dro inn ytterligere sensorer ville karakteren kanskje ikke ses på som rettferdig av de nye sensorene. Det kan også tenkes at elevens oppfattelse av karakteren som urettferdig skyldtes at eleven la til grunn andre kriterier enn det sensorene hadde når de vurderte besvarelsen. Det er allikevel rasjonelt å argumentere for at intersubjektiv enighet rundt hva som er rettferdig bør kunne foretrekkes over et individs personlige oppfattelse av hva som er rettferdig. I kapittel 3 om

rettferdighet ble det vist til ulike rettferdighetsteorier som tok utgangspunkt i spenninger mellom individ og samfunn. Dersom en skulle drøftet problemstillingen om eleven som fikk 3 på eksamen gjennom perspektivene som ble tatt opp i rettferdighetskapittelet, ville en sett at ut ifra et libertarianistisk perspektiv at karakteren 3 ville vært urettferdig nettopp på grunn av at eleven som ble vurdert ikke opplevde karakteren som rettferdig. Det libertarianistisk perspektivet ville sett på intersubjektivitet som et irrelevant moment for individets opplevelse av karakteren som rettferdig. Derimot ville de samfunnsbaserte prinsippene *strict egalitarianism, the difference principle, Resource-Based Principles, Welfare-Based Principles, og Feminist Principles* kunne støtte opp rundt en intersubjektiv tilnærming og derfor hevde at karakteren 3 var den mest rettferdige karakteren.

Intersubjektivitet representerer som vist ovenfor en kvantitativ tilnærming til objektivitet. På den andre siden står den kvalitative formen for objektivitet. Om dette skriver House: "The qualitative sense of objectivity is quite different. It refers to the quality of the observation regardless of the number of people making it. Being objective means that the observation is factual, while being subjective means that the observation is biased in some way"(House 1980:86). Vurdererrelabiliteten i en kvalitativ tilnærming for objektivitet handler om å finne fakta. Hva kan regnes som sant, hva er ikke sant? Hvem som helst kan i prinsippet innta en stilling hvor han eller hun bestemmer seg for hva som er sant. Dette er en trussel både for objektiviteten i seg selv, vurdererrelabiliteten og rettferdigheten. Spørsmålet er om det er mulig å få en objektiv elevvurdering ut fra dette og om det fremmer rettferdighet? Dersom det er intersubjektiv enighet mellom en rekke ufaglærte lærere i en vurdering, behøver dette langt ifra å bety at elevvurderingen tilfredsstiller det kvalitative objektivitetsidealet, eller for den saks skyld prinsippet om rettferdighet. Som House skriver: "The individual's qualitative objectivity can be assessed by his previous track record on such matters and by his current self interests"(House 1977:40). House påpeker her tidligere erfaring, og med dette også kompetansens betydning for objektivitet. Objektivitet er utfordrene, men objektivitet behøver ikke være ekvivalent til rettferdighet. Scriven understreker utfordringene med objektivitetsidealet i evaluering. "It is hard to maintain objectivity in the face of caused pain or joy and to decline bribe and threat combinations of various degrees of severity. The avoidance of evaluation thus achieves considerable support from many of those who would be obliged to do it, as well as those who would be subject to it" (Scriven 1991: 140). Det er viktig å være seg bevisst at objektivitet og rettferdighet ikke er det samme, men at krav om

objektivitet kan forhindre at elevvurderingen blir subjektiv. Subjektivitet kan være en trussel for elevvurderingen ved at det bygger på personlige meninger uten noen faktisk tilhørighet til det som skal vurderes. Ved å benytte seg av krav for objektivitet kan en forhindre at elevvurderingen blir gjennomsyret av subjektive meninger og på denne måten skape en mer rettferdig elevvurdering.

Objektivitet er utfordrende og kan derfor ses på mer som et ideal i elevvurderingen, enn en realitet. Hvorvidt objektivitet gir rettferdighet, er et spørsmål om hva sannheten er og om sannheten er rettferdig. Uansett hvordan det vris og vendes på, så vil evaluering ha et subjektivt fundament. Tønnessen og Tønnessen problematiserer dette ved å stille spørsmålet ”Skal vi vurdere bare det vi kan vurdere rimelig objektivt, eller skal vi vurdere alt, sjøl om en god del av det må bli temmelig subjektivt?” (Tønnessen og Tønnessen 2007:275). Dersom elevvurderingen blir truet av dårlig vurdererrelabilitet vil dette rettferdigheten stå på spill. Spørsmålene som blir stilt videre tar utgangspunkt i at en objektiv elevvurdering også vil gi grunnlag for en mer rettferdig elevvurdering. Kompetanse og intersubjektive forhold blir brukt som kriterier for å sikre vurdererrelabiliteten så vel som rettferdigheten i elevvurderingen.

En drøfting av vurdererrelabilitet har vist at vurdererrelabilitet kan påvirke, men ikke sikre rettferdighet. Rettferdighet i elevvurderingen lar seg påvirke fra mange hold. Å minske forekomsten av tilfeldige målingsfeil ved er bare et tiltak for å gjøre elevvurderingen mer rettferdig. Fordi det er så mange som har meninger om hva rettferdighet er, vil det alltid kunne diskuteres om en elevvurdering er rettferdig.

6.4 Avslutning

I dette kapitlet har det blitt satt fokus på begrepsvaliditet. Begrepsvaliditet handler på den ene siden om gyldighet i elevvurderingen og på den andre siden om en felles forståelse mellom lærer, elev og samfunn for hva eleven vurderes i. For å oppnå en rettferdig elevvurdering bør det være et vist samsvar mellom vurderingens gyldighet og kravene som

stilles. Spørsmålet er om en gyldig elevvurdering nødvendigvis er en rettferdig elevvurdering? At en elevvurdering er gyldig betyr ikke nødvendigvis at den er rettferdig, men at den må kunne forsvare de kravene som stilles til den. Et sentralt poeng i kapitlet har vært at dersom en elevvurdering kunne anses for å være rettferdig, bør det også være en felles forståelse rundt hvilke kriterier som ligger til grunn for rettferdighet og elevvurderingen. Det er viktig at både lærer og elev er bevist på hvilke krav som stilles til vurderingen, slik at deres oppfattelse av hva som blir vurdert er noenlunde den samme. En rettferdig elevvurdering bør bygge på både felles kriterier for vurderingen og gi rom for elevenes forskjeller.

Spørsmålet om begrepsvaliditet skaper en rettferdig elevvurdering, må i et slikt bilde ses etter hvilke kriterier som ligger til grunn for rettferdighet. Dersom libertarianistisk perspektiver ligger til grunn, bryter dette med retningslinjer som begrepsvaliditeten har om objektivitet og intersubjektivitet. På den annen side kan begrepsvaliditeten igjennom egalitaristiske prinsipper fort bli for generelle og glemme individets ferdigheter og behov. Rettferdighet som bygger på individuelle forutsetninger, viser seg ofte som urettferdig for alle andre. Når eleven skal ut i arbeidsmarkedet er det viktig for arbeidsgiver at en 5'er er en 5'er. Dette kan begrepsvaliditet igjennom en felles forståelse bidra med.

7. Rettferdighet gjennom ulike vurderingsformer?

I elevvurderingen finnes det ulike vurderingsformer. I dette kapitlet vil ulike former for vurdering drøftes opp mot hva som kan være rettferdig. Jeg vil se på ulike utfordringer innenfor hver av vurderingsformene i kapitlet. Med utfordring menes det i denne sammenheng de forhold og problemer læreren møter under vurdering av elevene, som også kan skape ubalanse i den rettferdige elevvurderingen. Jeg deler vurderingsformene inn i to måter å vurdere eleven på. Disse er tilbakemelding og sluttvurdering. Sluttvurderingen blir senere delt inn i både standpunkt karakter og eksamens karakter.

7.1 Former for vurdering

Det finnes mange måter å vurdere elever på. Det er for eksempel skriftlige tilbakemeldinger, karakterer og muntlige tilbakemeldinger. Det er skille mellom formelle vurderinger og uformelle vurderinger og det er vurderinger av fag og det er vurderinger av elevens orden og oppførsel. Måtene for vurdering er som sagt flertallige. I dette kapitlet skal kun to måter å vurdere eleven på tas opp. Den første måten er tilbakemeldingen. Den andre er sluttvurdering. Sluttvurderingen deles opp i standpunkt vurdering og eksamens vurdering. Ulike måter å vurdere eleven på stiller ulike krav for elevvurderingen. Eksempel på et krav er at formative tilbakemeldinger skal fungere som en veiledning i elevens læringsprosess. De ulike kravene gir også ulikt grunnlag for å bedømme hva som er rettferdig. Det vil si at en formativ vurdering har et annet utgangspunkt for rettferdighet enn det en summativ vurdering har. Jeg skal videre fokusere på formative tilbakemeldinger og dens veiledende funksjon. Samt sluttvurderingen og dens oppsummerende funksjon.

7.1.1 Tilbakemeldingen

Jeg ser i denne delen på tilbakemelding som en formativ vurderingsform og avgrensner meg dermed fra debatten om karakterer som tilbakemelding. I del 7.1.2 kommer jeg tilbake til karakterer som vurderingsform. Tilbakemeldingen er først og fremst et redskap lærer og elev sammen kan benytte for og videreutvikle elevens kompetanse. I kapittel 3.2 ble tilbakemeldingen kort problematisert i forhold til hvorfor elevvurdering er et aktuelt tema. Jeg kom frem til at det ønskes å legge et økt fokus på denne vurderingsformen for å bedre elevens læringsutbytte. Fordi tilbakemeldingene har elevens faglige utvikling som mål, stiller denne vurderingsformen andre krav for hva som er rettferdig, enn det en standpunktkarakter eller eksamenskarakter gjør. Eksempel på krav for tilbakemeldingen kan være at den skal gi eleven utfordringer å jobbe videre med. Forskriften til opplæringsloven skriver at Undervegsvurderinga skal innehalde grunnleggjande informasjon om kompetansen til eleven, lærlingen og lære kandidat og skal givast som meldingar med sikte på fagleg utvikling (Forskrift til opplæringsloven 2006:§ 3-11). Standpunktkarakteren derimot ”må baserast på eit breitt vurderingsgrunnlag som samla viser kompetansen eleven har i faget” (Forskrift til opplæringsloven 2006:§ 3-18). Det vil si at rettferdigheten må dømmes i forhold til elevens prosess. I denne delen av oppgaven skal det diskuteres hva tilbakemeldinger er, og dens rolle for en rettferdig elevvurdering. Videre vil jeg først se på hva som ligger i begrepet tilbakemelding. Deretter blir det drøftet hvordan en tilbakemelding kan framstå som rettferdig. Et viktig moment i denne drøftingen blir hvem tilbakemeldingen er rettet mot.

Tilbakemeldingen blir her sett på som en formativ vurderingsform, som kan sies å ligge ”at the heart of effective teaching” (Black m. flere 2001:1). En tilbakemelding åpner mulighetene for å hjelpe elevene i deres læringsprosess. Stortingsmelding nummer 31 2007-2008 skriver at hovedintensjonen ved tilbakemeldingen er ”å øke elevens kunnskap, ferdigheter og forståelse innenfor et spesifikt område eller ferdighet” (st. meld. nr 31 2007-2008:32). Tilbakemeldingen kan ta utgangspunkt både i prøver eleven har gjort, og utfordringer eleven møter i klasserommet. Stortingsmelding nr 31 fokuserer på at tilbakemeldingen også kan gis i når klassen arbeider seg igjennom et problem i fellesskap (St.meld.nr.31 2007-2008:32). Tilbakemeldingen kan finne sted i formelle så vel som uformelle vurderingssituasjoner. Elevsamtalen er en formel tilbakemeldingsform, mens lærers spontane tilbakemeldinger i klasserommet kan forstås som uformelle tilbakemeldinger (Viddal 2007, Eggen 2009). Viddal definerer en tilbakemelding som:

”dialoger som fremmer læring, anerkjennelse, engasjement, refleksjon og potensiell kompetanse hos den lærende. Dette realiseres når den profesjonelle part trer inn i den lærendes perspektiv på hans eller hennes arbeid uten å miste sitt eget, er seg sin egen posisjon bevisst og i øyeblikk lar være å sammenligne og vurdere den andre, slik at den lærendes opplevelse av tilbakemeldinger er i kontakt med vedkommendes selvbevissthet knyttet til eget arbeid og virke” (Viddal 2007:93).

Viddals definisjon på tilbakemelding kan nok anses for å være mye ønsketenkning. Betyr definisjonen at dersom en tilbakemelding ikke fremmer læring, anerkjennelse, engasjement, refleksjon og potensiell kompetanse hos den lærende, at det ikke er en tilbakemelding? Cedric Cullingford skriver: ”Assessment of some kind or another is a necessary tool. It is a starting point to manufacture something good, to improve and refine good practice or work of art. Used in a different way it is destructive. It destroys what it ostensibly seeks to improve” (Cullingford 1997:1). Tilbakemeldingen bør også ses i et slikt perspektiv og selv om definisjonen til Viddal er godt ment, kommer man ikke utenom at også tilbakemeldingen kan virke mot sin hensikt. Det vil si at istedenfor å gi eleven motivasjon til å øke kunnskapsnivået sitt, kan tilbakemeldingen også drepe elevens motivasjon.

Feedback er et annet begrep som benyttes om tilbakemeldingen. Både Deci og Ryan (1999) og Black og William (1998,2001,2004) er opptatt av hvordan feedback kan gi positive effekter i læringsprosessen. Deci og Ryan viser til forskningsresultater der: ”Positive feedback enhanced both free-choice behavior and self-reported interest” (Deci og Ryan 1999:627). At feedback har positiv effekt i læringsprosessen behøver for øvrig ikke bety at det førte til større læringsutbytte. Resultatene Deci og Ryan viser til vektlegger først og fremst motivasjon og ikke læringsutbytte som sådan. Black og William retter feedback mot læringsresultater når de skriver “Several studies show firm evidence that innovations designed to strengthen the frequent feedback that students receive about their learning yield substantial learning gains” (1998:7). Forskningen Black og William referer til, har som utgangspunkt å vise hvordan feedback kan være positivt for eleven, og viser derfor ikke de tilfellene hvor feedback virker mot sin hensikt. Black og William hevder også at "Feedback can only serve learning if it involves both the *evoking* of evidence and a response to that evidence by *using* it in some way to improve the learning” (Black m.fler 2003:122). I dette sitatet får Black og William fram at feedback ikke nødvendigvis fremmer læring. For at læring skal kunne fremmes må tilbakemeldingen involvere en vekking hos eleven, som gjør at eleven får lyst til å jobbe videre i læringsprosessen.

Tilbakemeldingen er som nevnt ovenfor formativ. En av hensiktene med tilbakemeldingen er å fremme elevenes mestringsfølelse og derfor også deres læringsprosess. St. meld. Nr 31 skriver ”For at elevene skal oppleve faglig fremgang og mestring, må de møtes med høye og realistiske forventninger til hva de kan få til, og lærerne må kunne formidle til elevene hva de trenger å jobbe mer med for å bli bedre”(st.meld.nr31 2007-2008:4). Stortingsmelding Nr. 31 retter seg mot den enkelte eleven. Hvordan kan lærer bidra til å hjelpe den enkelte eleven i hans eller hennes læringsprosess og hvordan kan en tilbakemelding ses på som rettfærdig? Jeg skal videre se på ulike rettfærdighetsteorier for å belyse dette spørsmålet.

I kapittel fire ble et rettfærdighetsprinsipp om fortjeneste, eller *desert*, tatt opp. Fortjenesteprinsippet skal her brukes til å argumentere for hvordan tilbakemeldingen kan rettfærdiggjøres. Fortjenesteprinsippet for rettfærdighet tok utgangspunkt i rettfærdighet som at mennesker bør bli belønnet for det arbeidet de har gjort, i tråd med verdien av deres bidrag til det sosiale produktet (Lamont og Favor 2007). Lamont og Favor skriver “(...) people freely apply their abilities and talents, in varying degrees, to socially productive work. People come to deserve varying levels of income by providing goods and services desired by others (Lamont og Favor 2007). I skolen får ikke eleven som fortjent i form av en inntekt, men i form av den vurderingen lærer gjør på bakgrunn av elevens prestasjoner. Det vil si at når en elev utfører et arbeid i klasserommet, vil det i følge fortjenesteprinsippet være gunstig å belønne elevens innsats. En måte å gjøre dette på er igjennom tilbakemeldingen. Utfordringen med fortjenesteprinsippet er at selv om eleven får en tilbakemelding som han eller hun i utgangspunktet fortjener, behøver ikke dette bety at det fremmer elevens læring. For eksempel kan ros av et godt arbeid føre til at eleven sier seg fornøyd med arbeidet og velger og ikke å videreutvikle seg. Dette kan ses i forhold til Vygotskys nære utviklingssone. Den nære utviklingssonen beskriver hvordan barn trenger utfordringer som ligger utenfor det nivået de mestrer på egenhånd (Miller 2002:377). Elevens tilbakemelding vil i eksempelet ovenfor være rettfærdig i den forstand at det er fortjent, men dette betyr ikke at tilbakemeldingen nødvendigvis vil føre til læring. Dersom en tilbakemelding blir gitt på et ufortjent grunnlag, som at lærer gir mer ros eller kritikk enn nødvendig, kan dette gjøre at tilbakemeldingen kan betraktes som urettfærdig. Et interessant spørsmål i denne sammenhengen er om en vurdering som ikke er fortjent kan fremme elevens læring, og at vurderingen ut ifra andre prinsipper enn fortjenesteprinsippet da kan anses som rettfærdig?

I følge Black m. fler er det dessverre ikke alltid slik at eleven ser det positive med en tilbakemelding (2003). Dersom eleven kun ser på tilbakemeldingen som noe negativt, kan et tiltak som i utgangspunktet er ment til å være en resurs i elevens læringsprosess bli til en trussel og et hinder (Black m. fler 2003:14). Når tilbakemeldingen fungerer mot sin hensikt, er det også grunn til å tro at det kan være fordi eleven opplever tilbakemeldingen som urettferdig. Spørsmålet om en rettferdig elevvurdering kan oppnås må her stilles til både lærer, elev og samfunn.

I kapittel 5 og 6 ble det lagt fokus på at for å oppnå en rettferdig elevvurdering bør det også være en felles forståelse mellom lærer, elev og samfunnet på hva som ligger i en rettferdig elevvurdering. Black m. fl. legger dette fokuset på hva feedback skal være: "The priority in giving feedback is to challenge students to tease out their assumptions and to help them be critical about the quality of arguments" (Black m. fler 2004:16). Et slikt perspektiv viser at lærers oppgave er å få fram lærelyst hos alle elever. Selv de elevene som har bestemt seg for og ikke lære. Dette er ikke alltid like lett. Hva som er rettferdig i slike situasjoner kan man vanskelig bli enig om. På den ene siden kan det hevdes at lærer ikke gjør jobben sin ordentlig og tilbakemeldings derfor ikke er rettferdig. På den annen side kan en beskyldte eleven for å være lat og ikke gjøre noe innsats og derfor ikke gjøre seg tilbakemeldingen fortjent. Det er ikke noe fasitsvar.

Jeg skal videre gå inn og diskutere tilbakemeldingen igjennom egalitarismen. Jeg viste i kapittel 3 at egalitarismen bygger på rettferdighet igjennom lik fordeling. Jeg skal i det følgende avgrense meg og holde diskusjonen innenfor *strict egalitarianism*. Dette betyr at det er den rene likhetstanken som blir diskutert og at jeg avgrensner meg vekk fra perspektivene om *instrumental* og *constitutive* egalitarisme (Lamont og Favor 2007). Hvordan kan en rettferdig tilbakemelding oppnås igjennom egalitarismen?

Egalitarisme prinsippet tar utgangspunkt i en lik fordeling av materielle goder og tjenester for å oppnå rettferdighet (Lamont og Favor 2007). En lik fordeling vil bety like mye tilbakemelding per elev. Lik fordeling fører til at tilbakemeldingen kan vanskelig la seg rettferdiggjøre under egalitarismen. Dette skyldes blant annet at tilbakemeldingen bygger på individuell tilpassing, noe som vil være vanskelig å utføre helt likt. Elever i likhet med andre mennesker er forskjellige og har derfor forskjellige utgangspunkter og behov i en

tilbakemeldingen. Hvordan lærer skal fordele tilbakemeldingen som en læringsressurs kan være et spørsmål om hvem som har behov for hjelp.

En rettfærdig tilbakemelding bør også kunne si noe om hvilke krav som stilles til et godt arbeid. Dette kan være krav basert på likhet eller individuelle forskjeller. Black m. fler skriver "Effective feedback should make more explicit to students what is involved in a high-quality piece of work and what steps they need to take to improve"(Black m. fler 2004: 19). Black og William forutsetter her at for å oppnå en virkningsfull tilbakemelding, må det ligge gitte steg som kan benyttes for å få eleven til å forbedre seg. Det er her et spørsmål om det er rettfærdig å gi noen elever mer hjelp for at de skal oppnå gode resultater, mens andre elever får mindre hjelp til å oppnå de samme resultatene. Egalitarismen strider imot en tankegang der noen elever har rett på mer hjelp enn andre. Når en elev får mer hjelp enn en annen vil dette være urettfærdig for de andre elevene.

Avslutningsvis kan en si om tilbakemeldingen at det er ingen fasit på hva som kan være en rettfærdig tilbakemelding. Rettfærdigheten i tilbakemeldingen kan på den ene siden sies å bli rettfærdiggjort ut ifra hva som fortjenes. Samtidig er innsatsen til lærer og elev med på å rettfærdiggjøre tilbakemeldingen. Det handler om hva de gjør ut av tilbakemeldingen og hvordan de jobber med den videre. Egalitarismen poengterer at alle har lik rett til like mye hjelp. Igjennom egalitarismen vil dette også gjelde tilbakemeldingen. Det er allikevel slik at elever er forskjellige. Noen elever er mer selvstendige og trenger mindre hjelp, andre elever er mindre selvstendige og trenger mer hjelp. Hvorvidt det kan argumenteres for at en tilbakemelding er rettfærdig eller ikke, avhenger av prinsippene man legger til grunn for hva som er rettfærdig. Jeg har ovenfor drøftet prinsippene fortjeneste og likhet. Jeg skal videre gå over og drøfte standpunkt karakteren og dens grad av rettfærdighet.

7.1.2 Sluttvurdering

På slutten av vært skoleår får elever i den videregående skolen en sluttvurdering. I forskriften til opplæringsloven står det at "Sluttvurderinga skal gi informasjon om kompetansen til eleven, lærlingen og lære kandidaten ved avslutninga av opplæringa i fag i læreplanverket" (Forskrift til opplæringsloven:§3-18). Målet med sluttvurderingen er derfor å gi en beskrivelse av den læringen eleven har nådd. I de fleste tilfeller består sluttvurderingen av karakterer og er derfor en form for summativ vurdering. En sluttvurdering i videregående skole kan forstås

som standpunktkarakter, eksamenskarakter og karakterer til fag-/svenneprøve og kompetanseprøve (forskrift til opplæringsloven: §3-18). I denne delen av oppgaven skal det handle om standpunktkarakterer og eksamenskarakterer.

I likhet med tilbakemeldingen som ble diskutert ovenfor, møter sluttvurderingen utfordringer i hva som kan sies å være rettferdig. På samme måte som all annen vurdering må sluttvurderingen kunne forsvare gitte krav for å regnes som en rettferdig vurdering. Når sluttvurdering gis som karakterer bør det være en generell enighet og forståelse for hva som ligger i disse karakterene. Er en 5'er en 5'er, eller varierer dette fra skole til skole og fra lærer til lærer?

Jeg skal i min videre drøfting av sluttvurdering skille mellom standpunktkarakterer og eksamenskarakterer. Årsaken til at en rettferdig sluttvurdering ikke blir drøftet under ett, er de ulike vurderingsprosessene som ligger til grunn for standpunktkarakteren og eksamenskarakteren. For eksempel er standpunktkarakteren satt av en intern vurderer mens eksamenskarakteren er satt av en intern og ekstern vurderer. Intern- og eksternvurdering kan føre til fordeler og ulemper ved en rettferdig slutt karakter. Jeg begynner med å se på standpunktkarakteren, deretter tar jeg for meg eksamenskarakteren, før jeg til slutt avslutter kapittelet.

Standpunktkarakteren er som alt nevnt en form for internvurdering. Det vil si at vurderingen blir foretatt av den læreren som er ansvarlig for den gitte eleven i et gitt fag. I tillegg til at standpunktkarakteren er intern, er den som oftest også summativ. At karakteren er summativ betyr i praksis at den kommer til uttrykk i form av et tall mellom 1 og 6, hvor 6 uttrykker høyst mulig oppnåelse i faget og 1 uttrykker lav/ingen oppnåelse i faget (forskrift til opplæringsloven § 3-4). I Forskriften for opplæringsloven står det følgende om standpunktkarakterer i fag:

”Standpunktkarakterer er karakterer som blir gitt ved avslutninga av opplæringa i fag, jf. læreplanverket, og som skal førast på vitnemålet.

Standpunktkarakteren må baserast på eit breitt vurderingsgrunnlag som samla viser kompetansen eleven har i faget, jf. § 3-3. Eleven skal ha høve til å forbetre kompetansen sin i faget inntil standpunktkarakteren er fastsett. Eleven skal bli gjort kjend med kva det er lagt vekt på i fastsetjinga av hennar eller hans standpunktkarakter” (forskrift til opplæringsloven §3-18).

Ut ifra sitatet fra formålsparagrafen ovenfor kan en standpunktkarakter settes på flere måter. Lauvås legger fram to varianter for hvordan gi standpunktkarakter som han mener forekommer i den norske skole. Lauvås skriver: ”I den ene uttrykker standpunktkarakteren faglæreres totale vurdering av den enkelte students nivå innenfor et gitt tidsspenn. Den andre består av et antall prøver, innleveringer og tester som hver blir karactersatt, og der standpunktkarakterer settes til gjennomsnittet av alle elementer som inngår” (Lauvås 2007:59). Mens det første perspektivet til Lauvås åpner mulighet for å gi standpunktkarakter på bakgrunn av elevens progresjon i løpet av et år, er det andre perspektivet kun et spørsmål om elevens snittkarakter. Hvordan skal en standpunktkarakter framstå som rettferdig ut ifra så ulike grunnlag? Dette kan problematiseres ytterligere med henvisninger til lærers subjektive rolle i elevvurderingen.

I likhet med enkeltkarakteren som ble drøftet i kapittel 5, trues standpunktkarakterens rettferdighet av lærers subjektive mening. Lauvås skriver: ”Standpunktvurderingen er en privatisert form for summativ vurdering med svært begrenset innsyn fra andre. Prøvene som gis, er bestemt av faglærer, enten de er laget av læreren selv eller hentet annetstedsfra. Det er ingen kvalitetskontroll av dem utover det faglærer selv står for” (Lauvås 2007:59). Dette kan forstå som at lærers subjektive meninger kan føre til en skjev fordeling i standpunktkarakteren. I praksis vil dette kunne bety at elever med ulike lærere blir vurdert på forskjellige grunnlag, og at karakterene elevene får derfor ikke kan oppfattes som rettferdige dersom en sammenligner grunnlaget lærere benytter for vurdering. Hvor store disse forskjellene kan være vet vi ikke, da det som Lauvås skriver, ikke skjer noen kvalitetskontroll av standpunktkarakterene.

Som nevnt tidligere i oppgaven, jobber Utdanningsdirektoratet med å skape en felles vurderingsform mellom lærerne (Bedre vurderingspraksis 2006). Den felles forståelsen for vurdering er et tiltak for å kunne oppnå en mer rettferdig elevvurdering. Spørsmålet er bare hvorvidt en felles forståelse for vurdering kan bli felles, når lærerne har ansvar for å velge ut og lage elevenes prøver. Selv om lærerne har en felles forståelse av hva og hvordan vurderingen skal finne sted, behøver ikke dette bety at prøvene de vurderer ut ifra gir samme grunnlag. Fordi lærerne fortsatt selv velger prøvene, velger de også vanskelighetsgraden på disse prøvene, noe som gjør at de felles kriteriene for vurdering blir gitt i forhold til ulike prøver.

Fortjeneste er et viktig rettferdighetsprinsipp også når det gjelder standpunkt karakterer. Har eleven fortjent den standpunkt karakteren som settes? Andrew Mason skriver ”the basis of desert- the characteristics in virtue of which people are said to deserve this or that – appears to change according to the kind of benefit in question” (Mason 2003:55). Hvilke karaktertrekk som vurderes, kan ut ifra Mason avgjøre om det er fortjent. Dersom det er mange karaktertrekk som skal vurderes, vil dette gjøre etterprøvbare mer utfordrende. For eksempel kan det tenkes at kriterier for en 4’er kan basere seg på både innsats, motivasjon, elevens refleksjon, struktur, innhold osv. Karaktertrekk som innsats og motivasjon er vanskelig å etterprøve for en eksternsensor som ikke er tilstede under elevens læringsprosess og på selve prøven. Ut ifra bare en oppgave kan eksternsensor vanskelig vite hvor motivert eleven var. Dersom eleven skulle vurderes ut ifra motivasjon, ville dette være en lite rettferdig vurdering.

Hvem som drar nytte av en rettferdig standpunkt karakter kan sies å være både elev og samfunn. Problematikken rundt *desert* – prinsippet oppstår blant annet når standpunkt karakteren blir urettferdig tildelt. Dette kan ses på bakgrunn av karaktertrekk, som jeg var inne på ovenfor. For eksempel dersom en elev som skulle fått en 5’er får en 3’er fordi læreren ikke har kompetanse i hvilke karaktertrekk han skal vurdere og hvordan standpunkt karakterer settes. Elevens standpunkt karakter vil da verken være rettferdig eller fortjent. Et annet eksempel kan være dersom eleven jukser seg fra en 3 til en 5’er i standpunkt karakter. Denne karakteren vil heller ikke være fortjent eller rettferdig. I det første tilfellet vil konsekvensene av standpunkt vurderingen være ufortjente og urettferdige. En slik kompetanseglipp kan føre til at en elev som i utgangspunktet skulle kunne komme inn på ønsket studie ikke kommer inn. Dette vil igjen kunne føre til konsekvenser på samfunnsnivå. I det andre eksempelet hvor eleven har jukset seg opp til karakteren 5, vil dette kanskje føre til at denne eleven får en fordel når han søker seg på videre studium, men at samfunnet kanskje får en ukompetent deltager. Det er altså ikke uten grunn at Havneis og Raaheim siterer Knite og Yorke med:”... the everyday trust we put in grades, marks and classes is not shared among experts in assessment” (Havnes, Raaheim 2006:74).

Standpunkt karakteren har som vist i første omgang konsekvenser for eleven, men den har også konsekvenser for samfunnet. Samfunnet har behov for å vite hva som ligger i en karakter og det bør derfor både for elev og samfunns skyld jobbes mot en mer felles forståelse for all vurderings praksis av elevene. Dersom lærere er enige om hva som skal vurderes, hvordan

eleven skal vurderes, hvilke kriterier eleven skal vurderes etter – kan dette være med på å påvirke standpunktkarakteren til å bli mer rettferdig. I kapittel 5 viste jeg til Houses oppfattelse av at enkeltpersoners kompetanse innenfor et felt, kan være med på å gjøre deres vurderinger mer objektive innenfor det samme feltet (House 1980:86). Når det gjelder standpunktkarakter er det derfor viktig at lærer har god kompetanse i både faget som skal vurderes og ha bevissthet om hvilken måte man tenker på rettferdighet.

Oppsummert kan en si at standpunktkarakteren slik den settes i norsk skole i dag, på mange måter bidrar til å skape urettferdighet i vurderingen av norske elever. Selv om Kunnskapsdirektoratet jobber mot en felles forståelse av hvilke kriterier eleven skal måles etter, er det fortsatt mye som kan bli sagt i forhold til det subjektive innslaget på hvem som lager prøvene, hvem som tolker hvordan eleven skal vurderes og hvem som setter karakterene. I tillegg til lærerens subjektive innslag i standpunktkarakterene, møter også standpunktkarakterene trusler fra eleven side, som for eksempel juks. Jeg skal videre ta for meg en annen form for sluttvurdering vurdering, nemlig eksamen.

Det finnes flere former for eksamen i videregående skole. Rammene som er satt for eksamen på eksamensdagen kan i følge forskriften til opplæringsloven § 3-30, enten være skriftlig eksamen på inntil 5 timer, muntlig eksamen på inntil 30 minutter per elev, muntlig praktisk eksamen på inntil 45 minutter per elev eller en praktisk eksamen på inntil 5 timer (forskrift til opplæringsloven: § 3-30). Eksamen skiller seg fra standpunktkarakteren på tre vesentlige områder. For det første er det både en intern og en ekstern sensor som vurderer det skriftlige arbeidet. For det andre er eleven mer eller mindre anonym i sin besvarelse, forutsatt at besvarelsen leveres anonymt, elektronisk og at faglærer ikke klarer å skille mellom elevenes skrivestiler. For det tredje er det på eksamen bare prestasjonen som skal vurderes og ikke prosessen. I følge Raaheim og Hauge skal: ”Alle i prinsippet gis de samme betingelsene for dermed å få like muligheter for å vise hva de kan” (2007:67). At elevene blir anonymisert er et tiltak for å gjøre vurderingen mest mulig lik. Det er allikevel slik at også vurderingen av skriftlig eksamen møter sine utfordringer. For å løse disse har forskriften til opplæringsloven bestemt at dersom det er uenighet mellom intern og ekstern sensor vil den eksterne sensoren ha det avgjørende ordet for elevens vurdering (Forskrift til opplæringsloven: § 3-30).

Før eksamenstidene er det i avisene fullt av eksempler på hvordan en rettferdig eksamen lar seg true for eksempel av elevenes eksamensangst eller elevens fusk. Jeg kommer til å bruke artikler fra Aftenposten for å aktualisere truslene til dagens vurderingsdebatter.

Aftenposten har under overskriften ”Hold styr på eksamensnervene” (Kaur 2009) en rekke tips til hvordan eleven kan unngå eksamensnerv. Eksamensangst er i seg selv ikke et unormalt fenomen og i de fleste tilfeller vil eksamensangst heller ikke true rettferdigheten i eksamenskarakteren. Raaheim og Hauge hevder: ”at en viss mengde spenning kan være nødvendig for å øke prestasjoner, men at prestasjonene svekkes dersom aktiveringen blir for sterk” (Raaheim og Hauge 2007:70). Det er allikevel slik at noen elever opplever mer eksamensangst enn andre og at dette går ut over deres prestasjonsevner på eksamen. Dersom en skal se en rettferdig eksamen ut ifra et kriterium om at elevens kunnskap skal fram, blir en elevens eksamensangst en trussel som må tas i betraktning dersom eleven skal vurderes rettferdig. Dette er et vanskelig problem, da sensorene i bunn og grunn ikke har noe utgangspunkt for å vite noe om hvilke besvarelser som har latt seg påvirke av elevens eksamensangst

Det er vanskelig å gi en fasit på hvordan eksamensangst kan viskes ut som en trussel for rettferdighet i eksamenskarakteren. Eksamensangst er heller ikke en faktor det er lett for sensorene å ta hensyn til når de vurderer elevens prestasjoner. Det beste tiltaket som kan gjøres for å hindre eksamensangst er at eleven selv jobber med angsten og lærer seg teknikker som gjør at angsten minskes. Raaheim og Hauge foreslår: ”For studenter som blir så stresset under eksamen at de presterer dårligere enn sitt optimale, er løsningen å benytte motsatt strategi av det de vanligvis gjør. I stedet for å skyve tanken på eksamenssituasjonen fra seg, skal de trene mentalt på å være i eksamenssituasjon hver dag i lang tid før eksamen kommer” (Raaheim og Hauge 2007:71). Et tilbud fra skolen om hjelp til å takle eksamensangst kan være et positivt tiltak for å minske eksamensangst som trussel for en rettferdig elevvurdering. Hjelp til å takle eksamensnerv kan også hjelpe en del elever til å ta ansvar for og selv skaffe seg en mer rettferdig karakter på eksamen.

Den andre trusselen for en rettferdig eksamenskarakter som avisene gjerne tar opp før og i eksamensperioden er fusk. Den 17.12.06 har nettutgaven til aftenposten følgende overskrift: ”Forvirring rundt eksamensfusk” (Sætre 2006). Aftenposten viser her hvordan studenter ved BI ikke vet hva som kan regnes for fusk og hva som ikke gjør det. På videregående bør ikke

dette være noe problem da eksamen arrangeres enten med eller uten hjelpemidler. Tar en elev med seg hjelpemidler eller benytter seg av hjelpemidler som internett eller mobiltelefon, blir dette sett på som fusk. Forskriften til opplæringsloven skriver: ”Eksamene kan organiserast med eller utan hjelpemiddel. Utdanningsdirektoratet fastset kva for hjelpemiddel som er tillatne i kvart fag ved sentralt gitt eksamen” (Forskrift til opplæringsloven: § 3-31). Om bruken av hjelpemidler er tillatt eller ikke er altså klart definert i forbindelse med eksamen, all overtreddelse av disse reglene kan bli tolket som fusk. I likhet med eksamensangst er eksamensfusk en trussel for en rettferdig eksamensvurdering. Lauvås skriver i artikkelen ”Fusk – et lite eller stort problem?” at det er: ”forbausende lite forskning om fusk i Norge” (Lauvås 2007:111). Dette gjør det vanskelig å uttrykke hvor stor trussel fusk er for en rettferdig eksamenskarakter i norsk skole. Det som er klart er at når elever blir tatt i fusk får det konsekvenser. I forskriften til opplæringsloven står det:

”Eksamen i eit fag kan annullerast dersom eleven eller privatisten har fuska eller forsøkt å fuske ved eksamen. Dette gjeld også dersom det er sensorane som oppdagar fusk eller forsøk på fusk. Spørsmålet om annullering av eksamen i faget skal avgjerast av rektor sjølv for elevar. Eleven skal ha høve til å uttale seg munnleg før rektor fattar enkeltvedtak om annullering av eksamen [...] For elevar som får eksamen annullert på grunn av fusk eller forsøk på fusk, fell standpunktarakteren bort.” (forskrift til opplæringsloven :§ 3-37).

Ut ifra forskriften til opplæringsloven § 3-37 vises det at det skal mye til for å bli tatt i fusk og få besvarelsen annullert. Det viser også at fusk får alvorlige konsekvenser dersom det blir oppdaget. Dette gjør ikke at fusk blir mer rettferdig. Det kan diskuteres om straffen til elever som blir tatt i fusk er rettferdig i forhold til handlingen de har gjort.

Aftenposten har 16.04.2008 overskriften: ”Jukserne straffes for strengt” (Kemp 2008). I artikkelen står det forslag om at studenter som jukser skal kunne slippe unna med en advarsel. En kan stille spørsmål ved om en advarsel for fusk, rettferdiggjør at en elev får eksamenskarakteren satt på et falskt grunnlag. Et poeng er at enten eleven blir tatt i fusk eller ikke, så vil dette fortsatt gi et urettferdig grunnlag for vurdering. Den eneste forskjellen er at dersom eleven blir tatt i fusk, får fusket i dag negative konsekvenser. Dersom eleven ikke blir tatt i fusk vil eksamenskarakteren settes på et ugyldig grunnlag og derfor også kunne oppfattes som en urettferdig karakter. Jeg skal videre gå inn og se på rettferdigheten i selve vurderingene som gis på eksamen. Det vil si kvaliteten i eksamensvurderingen. Dette kommer bare kort til å bli drøftet her, da det også var tema i de to foregående kapitlene.

Når eksamen er over, går avisene fra å fokusere på elevens utgangspunkt før og under eksamen, til å se på kvaliteten i vurderingen som intern og ekstern sensor gir. Et eksempel kan hentes fra Aftenposten som viser til en elev fra Sogn og Fjordane med standpunkt karakter seks i norsk sidemål, som fikk karakteren to på eksamen. Etter en klage gikk eleven opp to karakterer og fikk fire på eksamen (Moe 2009). Et slikt eksempel viser kanskje ikke at den ene karakteren er mer rettferdig enn den andre, men at denne eleven muligens har fått en eller flere karakterer tildelt som ikke kan oppfattes som rettferdig i en intersubjektiv kontekst. Men ettersom eleven gikk opp to karakterer ved å klage er det mulig å tro at karakteren som ble gitt på eksamen var gitt på et urettferdig grunnlag. Eksempelet viser hvordan lærernes meninger rundt karakterer kan bli skjevt fordelt. Steinar Kvale sier i et intervju med Kari Smith ”Ved eksamen skal [...] lærerne komme fram til en felles slutt karakter for en eksamensprestasjon. Det at sensor og eksaminatorer i noen tilfeller kan møte med vidt forskjellige karakterforslag, kan fra et psykometrisk synspunkt oppfattes som at karaktergivingen er tilfeldig og lite reliabel” (Kvale i intervju med Smith 2007:96). I kapittel 6 viste jeg til at en dårlig vurdererrelabilitet truer rettferdigheten i elevvurderingen. Sett mer spesifikt, kan man også si at dårlig reliabilitet truer rettferdigheten i eksamens karakteren.

Eksamen har to sensorer, en intern og en ekstern, som avgjør karakteren, der ekstern sensor har mest å si (forskrift til opplæringslova § 3-30). Allikevel kan det stilles spørsmål om hva som skjer når intern og ekstern sensor blir uenige om karakteren og ekstern sensor setter ned foten. I kapittel 6 ble subjektivitet, intersubjektivitet og objektivitet drøftet. Det ble vist enkeltpersoners vurderinger kan være objektive dersom lærer har tilstrekkelig kompetanse i både det målte feltet og i hvordan vurdere elever. Kompetanse i vurdering er også et viktig nøkkelord når intern og ekstern sensor setter eksamens karakterer. Hvilken kompetanse sensor sitter inne med kan ses i forhold til karaktertrekkene sensor skal vurdere. Dersom vurderingen tar utgangspunkt i struktur, innhold og refleksjon, er dette karaktertrekk sensor bør ha kompetanse på for at elevvurderingen skal bli rettferdig.

Oppsummert har jeg tatt for meg ulike oppfatninger av hva som er en rettferdig eksamens karakter. En kan ha lett for å tro at en rettferdig elevvurdering baserer seg på subjektive meninger og hvorvidt man er enige i disse meningene. Det kan problematiseres i forhold til hvorvidt elevens opplevelse av hva som er rettferdig, stemmer overens med et samfunnssyn på hva som er rettferdig. Rettferdighetsspørsmålet handler om hvilke

karaktertrekk som blir lagt til grunn for vurderingen av de ulike partene. Dersom eleven ser på karaktertrekk som innsats og motivasjon mens sensor ser på struktur og innhold, er det mulig at vurderingen kan anses som rettferdig ut ifra sensors karaktertrekk, men ikke fra elevens.

7.2 Avslutning

I dette kapitlet har jeg drøftet ulike vurderingsformer. Det har særskilt blitt sett på tre måter å vurdere eleven på. Disse var tilbakemelding, standpunktkarakter og sluttvurdering. Mens tilbakemeldingen er en vurderingsform med mål om videre læring, er standpunktkarakteren og eksamenskarakteren måling av oppnådd læring. Hvilke mål vurderingen har er en viktig faktor når rettferdighet skal drøftes. Rettferdighet må alltid ses i forhold til målet. Men at vurderingen har nådd målet sitt, er ingen garanti for at den er rettferdig. Uansett er subjektivitet en stor trussel både i tilbakemeldingen og sluttvurderingen som gis. Konsekvensene for dette kan vise seg både i forhold til individet som opplever vurderingen som urettferdig og i forhold til samfunnet som tar i mot kandidater som er mer eller mindre kompetente for jobb eller studier. Hvorvidt en elevvurdering kan oppfattes som rettferdig eller ikke, avhenger av hvilket rettferdighetssyn som legges til grunn. Dersom en vurderer eleven ut ifra fortjeneste vil man få andre resultater enn om man vurderer eleven ut ifra egalitaristiske prinsipper. Derfor vil jeg argumentere for mer kompetanse om hvordan gi gyldige vurderinger av elevene. Dette gjelder da ikke bare i selve vurderingspraksisen, men også alt som foregår rundt vurderingene. Som alt fra kunnskap om eksamensangst og fusk til kunnskap om hva som skal vurderes og hvordan det skal vurderes.

8. Avslutning

Hvorvidt en rettferdig elevvurdering kan sies å ligge i hjertet av effektiv læring eller ikke, avhenger av flere forhold. Denne oppgaven har tatt sikte på å diskutere problemstillingen ”*Kan en rettferdig elevvurdering oppnås?*”. Konklusjonen på dette er at det ikke finnes noe universelt svar. Svaret på dette spørsmålet kan være både ja og nei. Ja i den forstand at en elevvurdering kan være rettferdig under ett sett premisser. Nei, i den forstand at det alltid vil være andre premisser som fungerer som motargumenter på hva som er rettferdig. For eksempel kan ikke en elevvurdering være rettferdig både i forhold til differanseprinsippet og egalitarismeprinsippet på en gang. For å svare på problemstillingen har rettferdig elevvurdering blitt redegjort og drøftet. Det har blitt vist at en rettferdig elevvurdering må ses i forhold metode og en rekke ulike perspektiver på hva som er rettferdig.

I kapittel 2 har fokuset ligget på en kartlegging og analyse av hva som ligger i begrepet evaluering. Evaluering ble i denne oppgaven forstått som “the process of determining the merit, worth and value of things, and evaluations are the products of that process” (Scriven 1991:1). Ut ifra dette har oppgaven tatt hensyn til evaluering som både prosess og som produkt. En slik definisjon er åpen for evaluering innenfor flere områder. Evalueringer skjer i skolen, hjemmet, mellom venner, på arbeidsplassen, samtidig som individer også evaluerer seg selv og sine egne forutsetninger og behov. Etter en skisse av evalueringsfeltet ble gitt, avgrenset oppgaven seg inn mot elevvurderingsfeltet.

Kapittel 3 tok for seg hva elevvurdering er. Jeg valgte her å holde meg til en forståelse av elevvurdering som *et redskap, der har til hensikt og/ eller beskrive elevens læringsprosess*. Den største forskjellen fra evalueringsfeltet til elevvurderingen er at elevvurderingen rettes mot eleven, det enkelte individ. Evaluering derimot er generelt og kan rettes mot systemer så vel som individer, samt samspillet mellom systemer og individer. Det evaluering og elevvurdering har til felles, er mye av metodikken. Både evaluerings teori og elevvurdering skriver om formative og summative evalueringer. Begge områdene diskuterer metoden som ligger til grunn og validiteten bak denne metoden.

I kapittel 4 blir spørsmålet om hva rettferdighet er diskutert. I min søken etter rettferdighet har jeg vist til ulike perspektiver på hva rettferdighet er. Disse har blant annet vært basert på

egalitarismen, differanseprinsippet og desert/fortjeneste prinsippet. Et egalitaristisk rettferdighetsprinsipp tar utgangspunkt i lik fordeling. Alle har rett på de samme godene i samfunnet. Alle skal ha lik rett til å få utdanning. Alle skal ha de samme mulighetene osv. Differanseprinsippet står som kontrast til egalitrismeprinsippet ved at det baserer seg på ulik fordeling. I differanseprinsippet skal fordelingen skje på bakgrunn av individers ferdigheter og behov. Dersom en person trenger mer hjelp på grunn av gitte forutsetninger vil differanseprinsippet ta hensyn til dette. Det tredje perspektivet jeg har vært inne på er fortjeneste prinsippet. Det vil si at rettferdighet blir sett ut ifra et perspektiv om fortjeneste. Har en person, gruppe eller samfunn gjort seg fortjent til den belønningen eller straffen de får? Å diskutere rettferdighet ut ifra disse tre ståstedene gir kun et lite perspektiv på hvordan rettferdighet kan forstås. Blant annet har perspektiver om velferd og feminisme bevist blitt utelatt, da jeg ikke vurderte dem til å være av større relevans for diskusjonen videre i oppgaven.

Når problemstillingen *kan en rettferdig elevvurdering oppnås?* blir drøftet, har jeg valgt å diskutere dette igjennom tre kapitler. Kapittel 5 og 6 problematiserer forholdet mellom en rettferdig elevvurdering og vitenskapsmetode. Kapittel 7 ser på mer generelle prinsipper for vurdering i form av tilbakemeldinger og sluttvurderinger.

Kapittel 5 ser på problemstillingen i forhold til vitenskapsmetode. Vitenskapsmetode blir her sett på som et redskap for å gi en sikker og gyldig elevvurdering. Spørsmålet er om en sikker elevvurdering også gir en rettferdig elevvurdering. Kapittel 5 viser at vitenskapsmetode kan benyttes som et redskap for å gi mer gyldige elevvurderinger, men at dette ikke nødvendigvis kan sidestilles med at vurderingen er rettferdig (House 1993). Jeg bruker egalitarismen, differanseprinsippet og fortjenestep prinsippet til å argumentere for at rettferdighet ikke nødvendigvis er et resultat av metodiske tilnærminger, men at rettferdighet må ses i forhold til de prinsippene som ligger til grunn for rettferdigheten. Dersom sannhet og sikkerhet i vurderingen er et kriterium for hva som er rettferdig og elevvurderingen oppfyller dette kriteriet, kan en si at elevvurderingen er rettferdig. Et spørsmål som kan stilles ut ifra dette og som oppgaven ikke har tatt stilling til er ”hva er sannhet?” og ”hva er sikkerhet?”. Kan en elevvurdering noen gang bli sikker og er sikkerhet nok for å hevde at vurderingen er rettferdig?

Kapittel 6 fortsetter metodediskusjonen i kapittel 5. Jeg så i dette kapitlet på hvordan begrepsvaliditet og en felles forståelse legger grunnlaget for om en elevvurdering oppfattes som rettferdig. Kapitlet legger spesielt to kriterier til grunn for at noe skal være rettferdig. For det første bør det være samsvar mellom vurderingen som gis og det som vurderes. For det andre bør det være en felles forståelse mellom den som foretar vurderingen, den som blir vurdert og resten av samfunnet. Gyldighet og felles forståelse blir her kriterier for hva som kan anses for å være rettferdig. Det er allikevel ikke grunnlag for å hevde at en kan sette likhetstegn mellom gyldighet og rettferdighet, da en gyldig elevvurdering kun vil være rettferdig i den forstand at gyldigheten samsvarer med premissene for rettferdighet.

Ut ifra kapittel 5 og 6 ser vi at metode alene ikke er nok for å kunne oppnå en rettferdig elevvurdering. Rettferdighet må ses i forhold til de kriteriene som ligger til grunn for elevvurderingen og ut ifra hvilke kriterier for hva rettferdighet er. Selv om metode kan fungere som et redskap til å få en gyldigere elevvurdering, er ikke dette inneforstått med at elevvurderingen er rettferdig. Rettferdighet inngår slik sett i en større kontekst, hvor metode, kriterier for rettferdighet, samfunnets forståelse, elevens forståelse og elevens forutsetninger og behov bør tas hensyn til.

I kapittel 7 blir ulike vurderingsformer diskutert. Det skilles mellom tilbakemeldinger og sluttvurderinger. Om en rettferdig elevvurdering kan oppnås avhenger også her av prinsippene som ligger til grunn for hva som er rettferdig. Det vil nedenfor bli oppsummert hvordan egalitarismeprinsippet, differanseprinsippet og fortjenestepriksippet gir ulike svar på om en rettferdig elevvurdering kan oppnås.

En egalitaristisk tilnærming til problemstillingen ”om en rettferdig elevvurdering kan oppnås” innebærer at elevvurderingen skjer på et likt grunnlag for alle elevene. Det vil si at så lenge vurderingen skjer likt i forhold til alle elever, vil vurderingen også kunne anses for å være rettferdig. Det vil si at så lenge alle elevene vurderes på samme måte og etter de samme kriteriene vil vurderingen være rettferdig. Dette er derimot ikke spesielt lett å praktisere. Både lærere og elever er forskjellige. Det vil derfor være veldig ambisiøst og tro at elevvurdering utført av ulike lærere og sensorer vil forekomme på en lik måte. Selv om like kriterier ligger til grunn for vurderingen, innebærer ikke dette nødvendigvis at forskjellige lærere tolker kriteriene likt. Likhet kan etterstrebes i vurderingspraksisen, men likhetsidealet for en rettferdig elevvurdering vil aldri fullkomment oppnås.

Som motpol til egalitarismen står differanseprinsippet. Det har i oppgaven blitt vist at differanseprinsippet representerer et syn på rettferdighet sett ut ifra individuelle ferdigheter og behov. I et slikt prinsipp må en rettferdig elevvurdering ses i forhold til elevens ferdigheter og behov. Differanseprinsippet møter en speilvendt argumentasjon på hvorfor dette ikke er et prinsipp som kan oppnås i spørsmålet om en rettferdig elevvurdering. For selv om vurderingene kan fremstå som rettferdige i forhold til det enkelte individ, blir det vanskelig å vite hva som ligger i en elevvurdering på samfunnsnivå. Når eleven skal søke seg videre opp i utdanningssystemet eller søke jobb etter endt utdanning, er det en fordel at vurderingen ikke kun ser på rettferdighet i forhold til individet, men også på rettferdighet i forhold til andre mennesker.

Om en rettferdig elevvurdering kan oppnås, kan også ses i forhold til fortjeneste. Det har i oppgaven blitt vist at faktorer som elevens innsats, vilje og motivasjon kan være grunnlag for elevvurderingen. Når en elevvurdering tar hensyn til disse faktorene kan det argumenteres for at elevvurderingen er rettferdig fordi eleven har fortjent den. utfordringene for fortjenesteperspektivet på rettferdighet er at å vurdere fortjeneste. Hvordan kan lærer vurdere en elevs fortjeneste dersom vurderingsgrunnlaget ikke strekker til? Å vurdere fortjeneste innebærer å vurdere mer enn prestasjonen til eleven. Faktorer som innsats og motivasjon bør også tas med. Dette er faktorer som ofte går utenfor hva en lærer kan vurdere eleven etter.

Min konklusjon om det ikke finnes et universelt svar på om en rettferdig elevvurdering kan oppnås, har flere konsekvenser. Ovenfor ble det vist hvordan en rettferdig elevvurdering kan oppnås dersom en benytter seg av gitte prinsipper for rettferdighet. Nedenfor vil jeg vise at også denne konklusjonen møter utfordringer.

Denne oppgaven har tatt stilling til problemstillingen *Kan en rettferdig elevvurdering oppnås?* Denne problemstillingen kan ses i lys av utallige teorier for vurdering og rettferdighet. Det du har lest til nå har vært mitt bidrag til debatten. Som vist i innledningen har den hermeneutiske metode til hensikt å skape en helhet ut ifra enkeltdelene. Helheten fortsetter og utvikle seg som en spiral ved å dra inn nye perspektiver som kan belyse temaet og problemstillingen. Dette innebærer at den diskusjonen jeg til nå har lagt fram kun er et lite bidrag til hvordan problemstillingen kan tolkes.

I innledningen tok jeg for meg tre avgrensninger. Den første avgrensningen tok sikte på å vise at oppgaven er av normativ karakter. At oppgaven har vært av normativ karakter, har gitt meg

et annet svar enn dersom jeg skulle gått ut og vurdert om en rettferdig elevvurdering kan oppnås empirisk. utfordringer ved å måle problemstillingen empirisk ligger blant annet i uklarhetene rundt rettferdighetsbegrepet. Hvordan skal man kunne måle rettferdighet empirisk? Rettferdighet er ikke noe man kan se eller ta på. Rettferdighet er et konsept med mange ulike tolkninger. Denne oppgaven har kun tatt stilling til tolkninger rundt likhet, ulikhet og innsats.

Den andre avgrensningen min skjedde i forhold til elevvurdering. Jeg skrev i innledningen at elevvurderingen kun ble belyst i forhold til forholdet mellom lærer, sensor og elev. Perspektiver som da faller vekk er elevens selvutvurdering og elevvurdering som skjer mellom elever. Fordi disse vurderingsformene ikke har noen formelle krav til hva som kan forventes av dem, blir rettferdighetsbegrepet enda mer utfordrende å definere. Det er mulig å anta at ulike elever vil ha ulike oppfatninger av hva som er rettferdig, noe som vil prege deres vurdering av dem selv og av andre.

Den tredje avgrensningen jeg gjorde var i forhold til hvilke perspektiver på rettferdighet jeg ville ha med i besvarelsen. Jeg utelot her blant annet perspektiver om velferd og feminisme. Det er selvfølgelig mulig å diskutere om en rettferdig elevvurdering kan oppnås ut ifra disse perspektivene også. Spørsmålet er bare hvor relevant det er? For min egen del hadde jeg liten interesse av å diskutere oppnåelsen av en rettferdig elevvurdering igjennom prinsipper for velferdsstaten og feministene. Dette overlater jeg til personer som har interesse av det.

Konklusjonen på at rettferdighet kan ses fra flere perspektiver, gir innsikt i ulike forståelser av rettferdighetsbegrepet. Helt avslutningsvis i oppgaven vil jeg gjøre meg noen refleksjoner på hva som kunne vært gjort annerledes i oppgaven.

Selv om oppgaven må forstås ut ifra de delene den består av og den helheten den presenterer, så vil det alltid finnes perspektiver som kunne gitt andre svar og vinklinger til oppgaven. Avgrensningene som ble gjort innledningsvis i forhold til elevvurdering, selvutvurdering, vurderinger gjort av medelever og rettferdighetsperspektiver er eksempler på faktorer som kunne gitt oppgaven et annet resultat.

Alle besvarelser foretar en eller annen form for avgrensninger. Det vil alltid være perspektiver som kan utfylle og gi mer bredde og dybde i en debatt. Slik ser vi at hermeneutikkens spiral fortsetter og utvikle seg også etter at en tekst er fullbyrdet. Hvorvidt en rettferdig

elevvurdering er nøkkelen til hjertet med effektiv læring, er jeg usikker på. Som J.R Lucas skriver, så er ikke rettferdighet alene nok. Jeg ønsker å ende denne oppgaven med dette sitatet:

“Justice by itself is not enough. It does not make a man happy or fulfilled, and is no guarantee of salvation in this world or the next. There is no justification by justice alone. Nevertheless, it is not to be despised. It may not be everything, but it is something. It is a virtue, an important virtue, one of the cardinal virtues; for it is the bond of peace, which enables the individual to identify with society, and brethren to dwell together in unity” (Lucas 1989:263).

9. Litteraturliste

Almås, Reidar (1993): *Evaluering på norsk*. Universitetsforlaget AS. Oslo

Alvesson, Mats og Sköldberg, Kaj (1994): *Tolkning och reflektion. Vetenskapsfilosofi och kvalitativ metod*. Studentlitteratur. Sverige.

Bernstein, Basil (1974): *Kodeteori: Et udvalg af hans artikler om sprog, socialisering og kontrol redigeret og med indledning af Jan Enggaard & Kirsten Poulsgaard*. Christian Ejlers Forlag, København. Oslo.

Black, Paul og William, Dylan (1998): *Assessment and classroom learning*. I *Assessment in Education*, Vol. 5, No. 1, 1998. UK.

Black, Paul and William, Dylan (2001): *Inside the black box. Raising standards through classroom assessment*.

Black Paul, Harrison Christine, Lee Clare, Marshall Bethan, Wiliam Dylan (2003): *Assessment for learning. Putting it into practice*. Open university press. Berkshire. England.

Black, Paul, Christine Harrison, Clare Lee, Bethan Marshall and William, Dylan (2004): *Working inside the black box. Assessment for learning in the classroom*. I *Phi Delta Kappan*, (2004) Vol 86 # 1, side 8-21.

Bjørgen, Ivar (2007): *Tømmerstokkens historie – eller hvorfor det er så vanskelig å si noe fornuftig om læring*. I: Tveit, Sverre (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*. Universitetsforlaget. Oslo.

Brady, David (1997): *Assessment and the Curriculum*. I Cullingford, Cedric (1997): *Assessment versus evaluation*. Biddles Ltd, Guildford.

Bransford m.fl. (2004): *How people learn. Brain, mind, experience, and school*. National academy press. Washington, D.C gikk og 1975): *Innføring i økopedagogikk*. Ascheoug.Oslo.

Bruner, Jerome, Ross, Gail, Wood, David (1976): *The role of tutoring in problem solving*. I *J. Child Psychol. Psychiat.*, Vol. 17, 1976, pp. 89 to 100. Pergamon Press. Printed in Great Britain.

Bruner, Jerome S. (1996): Verdien av struktur. Oslo: Gyldendal. I: Dale, Erling Lars (red.). *Skolens undervisning og barnets utvikling: klassiske tekster*. (Ny tittel: om utdanning), s. 125-141

Bruner, Jerome (2002): *The process of Education – A landmark in educational theory*. Harvard university Press. USA

Boud, D (1995): *Ensuring that assessment contributes to learning*. I: Proceedings from international conference on problem based learning in higher education. Inter-faculty program-oriented and general experiences, September 24-27, 1995, Linköping university. Linköping.

Bø, Inge og Helle, Lars (2002): *Pedagogisk ordbok: praktisk oppslagsverk i pedagogikk, psykologi og sosiologi*. Oslo. Universitetsforlag

Cullingford, Cedric (1997): *Introduction* i Cullingford, Cedric (red.)(1997): *Assessment versus evaluation*. London.

Dale, Erling Lars og Wærness, Jarl Inge (2007): *Tilpasset opplæring og elevvurdering*. I Tveit, Sverre (red.) (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*. Universitetsforlaget. Oslo.

Dale, Erling Lars & Wærness, Jarl Inge (2006): *Vurdering og læring i en elevaktiv skole*. Universitetsforlaget. Oslo.

Dale, Erling Lars (1999): *Utdanning med pedagogisk profesjonalitet*. Oslo: Ad Notam Gyldendal.

Davis, Andrew (1998): *The limits of educational assessment*. Blackwell publishers. Oxford.

Deci, Edward L, Ryan, Richard M og Koestner, Richard (1999): *A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation*. I *Psychological Bulletin* 1999. Vol.125, NO.6. 627-668. American Psychological association. Inc.

Dobson, Stephen & Smith Kari (2007): *Vurdering og vurderingsformer*. I Norsk pedagogisk tidsskrift – forum for pedagogikk og fagdidaktikk. Universitetsforlaget. Oslo.

Dweck, Carol S. (1999): *Self-theories. Their role in motivation, personality and development*. Psychology Press

Eggen, Astrid (2009): *Utviklingssamtalene som Vurdering*. I Tidsskrift for lærere og skoleledere: Bedre skole. NR. 2 – 2009.

Engelsen, Britt Ulstrup (2006): *Kan læring planlegges? Arbeid med læreplaner – hva, hvordan, hvorfor?* Gyldendal Norsk Forlag AS. Oslo.

Eisner, W. Elliot (1985): *The art of educational evaluation. A personal view*. The falmer Press. Great Britain.

Engh, Knut Roar (2007): *Elevmedvirkning i vurderingsarbeidet*. I Norsk pedagogisk tidsskrift – forum for pedagogisk fagdidaktikk. Universitetsforlaget. Oslo.

Evenshaug, Oddbjørn og Hallen, Dag (2002): *Familiepedagogikk. Oppdagelsens hva, hvordan og hvorfor*. Gyldendal akademisk. Oslo.

Gregersen, Jørgen (1977): *Evaluering – betraktet med utgangspunkt i en modell*. I Ålvik, Trond (1977) *Undervisningslære 1. Aktuelle synspunkter og problemer*. Copenhagen: Gyldendals Pædagogiske bibliotek.

Goodlad, John I. (1979): *Curriculum Inquiry. The study of Curriculum Practice*. McGraw-Hill Book Company. New York.

Hargreaves, David (1994): *Lærerarbeid og skolekultur – lærerurkets forandring i en postmoderne tidsalder*. Gyldendal Norsk Forlag, Oslo.

Haug, Peder (1994): *Evaluering – vitenskap eller politikk?* i Ålvik, Trond (red.): *Skolebasert Vurdering – en artikkelsamling*. Ad Notam Gyldendal AS. Norge

Havnes, Anton og Raaheim, Arild (2006): *Vurdering og eksamen*. I Strømø, Helge, Hofgaard Lycke, Kirsten og Lauvås, Per (red.). *Når læring er det viktigste. Undervisning i høyere utdanning*. Cappelen Akademisk Forlag. Oslo.

- Hjardemaal, Finn (2002): *Vitenskapsteori*. I Kleven, Thor Arnfinn (red.) (2002): *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolkning og vurdering*. Unipubforlag. Oslo
- House, R. Ernest (1993): *Professional evaluation*. Sage publications Inc. USA.
- House, R. Ernest (1980): *Evaluation with validity*. Sage publications. Inc. USA.
- House, R. Ernest (1977): *The logic of evaluative argument*. University of California, Los Angeles.
- Imsen, Gunn (1999): *Lærerens verden, innføring i generell didaktikk*. Tano Aschehoug.
- Kjernsli, Marit, Lie, Svein, Olsen, Rolf Vegar, Roe, Astrid (2007): *Tid for tunge løft. Norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006*. Universitetsforlaget
- Kleven, Thor Arnfinn (red.) (2002): *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolkning og vurdering*. Oslo: Unipub.
- Kleven, Thor Arnfinn (2002): *Begrepsoperasjonalisering*. I Lund, Thorleif (red.) (2002): *Innføring i forskningsmetodologi*. Unipub. Oslo.
- Klette, Kirsti (red.) (1998): *Klasseromsforskning – på norsk*. Ad Notam Gyldendal. Oslo.
- Kvale, Steinar (1970): *En eksaminasjon av universitetseksamener*. Universitetsforlaget. Norge.
- Kvale, Steinar (2007) I intervju med Smith, Kari (2007): *Om eksamensforskning – Intervju med Steinar Kvale*. I Norsk Pedagogisk Tidsskrift – Forum for pedagogikk og fagdidaktikk. NR 2. 2007: Vurdering og vurderingsformer. Side 94-99.
- Kvernbekk, Tone (2002) *Vitenskapsteoretiske perspektiver*. I Lund, Thorleif (2002): *Innføring i forskningsmetodologi*. Unipub AS. Oslo.
- Kvernbekk, Tone (2005): *Pedagogisk teoridannelse. Insidere, teoriformer og praksis*. Fagbokforlaget. Bergen.
- Lauvås, Per (2007): *Standpunktvurderingen – et problem eller selve problemet?* I Tveit, Sverre (2007): *Elevvurdering i skolen, grunnlag for kulturendring*. Universitetsforlaget. Oslo.
- Lucas J.R (1989): *On justice*. Clarendon press. Oxford

Lund, Thorleif (red.) (2002): *Innøfring I forskningsmetodologi*. Unipub. Oslo.

Lyngsnes, Kit og Rismark, Marit (2007): *Didaktisk arbeid*. Gyldendal Norsk Forlag AS. Oslo.

Lysne, Anders (1999): *Karakterer og kompetanse: Stridstema i norsk skolehistorie. Bind 1. Utviklingen frem til 1970-årene*. AVA forlag. Haslum.

Læringscenteret (2003): *Karakter – mer enn karakterer. Debatt om elevvurdering i grunnopplæringen*. Læringscenteret (LS).

Løvlie, Lars (1992): *Pedagogisk filosofi*. I Dale, Erling Lars (red.) (1992): *Pedagogisk filosofi*. Ad Notam Gyldental. Norway.

Mason, Andrew (2003): *Meritocracy, Desert, and the Moral Force of Intuitions*. I Bell, Daniel A. og de-Shalit, Avner (2003): *Forms of justice*. Rowman & Litterfield Publishes, inc. USA.

Matthiesen, Georg (2007): *Vurdering i skole og opplæring – problemstillinger og erfaringer*. I Tveit, Sverre (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*. Universitetsforlaget. Oslo.

Mehlbye, Jill, Riwpwe, Olav og Togeby, Miael (1993): *Håndbog i evaluering*. AKF Forlaget. København.

Miller (2002): *Theories of Developmental Psychology*. New York: Worth. 4. utgave.

Nixon (1990): *Assessment issues in relation to experienced-based learning on placements within courses*, C.Bell og D.Harris (red.): *World Yearbook of Education 1990: Assessment and Evaluation*, Great Britain.

Norris, Stephen P. (1983): *The Inconsistencies at the Foundation of Construct Validation Theory*. I: House, Ernest R (red.): *Philosophy of Evaluation*. Jossey-Bass Inc, Publishers. San Francisco.

Nygård, Roald (1993): *Aktør eller brikke: Om menneskers selvforståelse*. Oslo: Ad Notam Gyldendal.

- Ogden, Terje (2005): *Kvalitetsskolen*. Gyldendal Norsk Forlag AS. Oslo
- Pervine, Lawrence A., Cevrone, Daniel, John, Oliver P (2005): *Personality. Theory and research*. John Wilwy & Sons, Inc. USA
- Popham, James. W (2003): *Test better, teach better. The instructional role of assessment*. Association for supervision and curriculum development Alexandria, Virginia USA.
- Rawls, John (1971): *A Theory of Justice*. Cambrige, Harvard University Press.
- Rawls, John (1999): *A Theory of Justice*. Cambrige, Harvard University Press. Revidert utgave.
- Raaheim, Arild og Hauge, Helge (2007): *Eksamen og eksamensangst*. I Tveit, Sverre (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*. Universitetsforlaget. Oslo.
- Ross, Stephen David (1998): *Ideals and responsibilities. Ethical Judgment and Social Identity*. Wadsworth publishing company. USA.
- Scriven Michael (1991): *Evaluation Thesurus, Fourth Edition*. Sage Publicatoins. United States of America.
- Scriven, Michael (1967): *Perspectives of Curriculum Evaluation*. Rand mc.Nally & Company, Chicago.
- Skolerapport 1. Evaluering av modeller for kjennetegn på måloppnåelse i fag i prosjektet bedre vurderingspraksis.
- Stortingsmelding nr. 16 (2006-2007): *...og ingen stod igjen. Tidlig innsats for livslang læring*.
- Stortingsmelding nr. 30 (2003-2004): *Kultur for læring*.
- Stortingsmelding nr. 31 (2007-2008): *Kvalitet i skolen*.
- Tetzchner, Stephen von (2003): *Utviklingspsykologi. Barne- og ungdomsaldern*. Gyldendal Norsk Forlag AS. Gjøvik
- Temkin, Larry S. (2009): *Illuminating egalitarianism*. I Christiano, Thomas og Christman John (red.)(2009): *contemporary debates in political philosophy*. Wiley Blackwell, UK.

Tveit, Knut (2002) *Historisk forskningsmetode* i Kleven, Thor Arnfinn (red.): *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolking og vurdering*. Unipub AS. Oslo.

Tveit, Sverre (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*.

Universitetsforlaget. Oslo.

Tuck, Ralph (1997): *Is self-assessment a valid concept?* I red. Cullinford, Cedrik (1997): *Assessmen versus evaluation*. Biddles Ltd, Guildford & Kings's Lynn. London.

Tønnessen Rolf Th.og Tønnessen Marianne (2007): *Demokratisk dannelse, fagdidaktikk i samfunnskunnskap*. Bergen. Fagbokforlaget Vigmostad & Bjørke AS.

Ulriksen, Lars (2003): *Barnehage, skole og samfunn – rammer og muligheter*. Bearbeidet og kommentert av Lars Monsen.. I: Arneberg P og Overland, B: *Pedagogikk. Mangfold og muligheter*, s. 161- 284. Damm.. Oslo:

Utdanningsdirektoratet (2006): *Læreplanverket for kunnskapsløftet. Midlertidig utgave juni 2006*.

Utdanningsdirektoratet (2007): *Vurdering. Et felles løft for bedre vurderingspraksis – en veiledning*.

Utdanningsdirektoratet (2007): *Bedre vurderingspraksis. Utprøving av kjennetegn på måloppnåelse i fag – en veiledning*.

Vaage, Sveinung (2000): *Utdanning og demokrati, barnet, skolen og den nye pedagogikken: John Dewey i utvalg*. Abstrakt forlag. Oslo.

Valvatne, Helene og Sandvik, Margareth (2002): *Barn, språk og kultur. Språkutvikling fram til sjuårsalderen*. Cappelen Akademisk Forlag. Oslo.

Vedung, Evert, Furubo, Jan-Eric og Sandahl, Rolf (2002): *Utvärdering I det svenska politiska systemet. Åtta iakttagelser*. Uppsala Universitet, Institutet för bostads-och urbanforskning (IBF).

Vedung, Evert (2003): *Models of Evaluation*. I: Haug, Peder og Schwandt, Thomas A (red.) (2003): *Evaluating Educational Reforms*. Information age publishing. Usa. s.41-73.

Viddal, Linda Marie (2007): *Tilbakemeldingsdiskurser i elevsamtalen*. I Tveit, Sverre (2007): *Elevvurdering i skolen. Grunnlag for kulturendring*. Universitetsforlaget. Oslo.

von Tetzchner, Stephen: *Utviklingspsykologi. Barne og ungdomsskolealderen*, 2001. Oslo: Gyldendal Akademisk.

Weber, Max (2000): *Samfunnsvitenskapenes "objektivitet"*. Oslo: Gyldendal. I: *Makt og byråkrati*

9.1 Internetkilder

Aalen, Ida (2008): *Alle tjener på at ingen stryker. 9000 kroner i timelønn*.

<http://e24.no/oppogfrem/article2575463.ece> [Lesedato 01.06.2009]

Aftenposten nettutgave (09.02.09): *Skolen favoriserer jenter*.

<http://www.aftenposten.no/nyheter/iriks/article2914696.ece> [Lesedato 29.11.2009]

Barry, Christian (2004): *Redistribution*. <http://plato.stanford.edu/entries/redistribution/>

[Lesedato 22.04.2009]

Beauchamp, Tom (2008): *The Principle of Beneficence in Applied Ethics*.

<http://plato.stanford.edu/entries/principle-beneficence/#LibLimBenProBenPat> [Lesedato

17.03.2009]

Bedre Vurderingspraksis (2007): *Utprøving av kjenneteng på måloppnåelse i fag – en veiledning*. Utdanningsdirektoratet.

Bjørkeng, Per Kristian (2008): *Mener Pedagogikken kan bli bedre*.

http://www.aftenposten.no/kul_und/article2221495.ece [Lesedato 25.08.2009].

Blake, Michal (2005): *International Justice*. [http://plato.stanford.edu/entries/international-](http://plato.stanford.edu/entries/international-justice/)

[justice/](http://plato.stanford.edu/entries/international-justice/) [Lesedato 15.02.2009]

Bojer, Hilde: *En rettferdig verden?* <http://folk.uio.no/hbojer/rettferdigverden.PDF> [Lesedato 17.09.2009].

Christiano, Tom (2006): *Democracy*. <http://plato.stanford.edu/entries/democracy/> [Lesedato 15.02.2009]

Clausen Hoas, Krister (2009):- *Dropp leksene med loven i hånd*.
<http://www.aftenposten.no/nyheter/iriks/skole/article2908020.ece> [Lesedato 26.04.2009]

Dworkin, Gerald (2005): *Paternalism*. <http://plato.stanford.edu/entries/paternalism/>
[Lesedato 17.02.2009]

Forskrift til opplæringsloven (2006:§3-4):*Karakterar i fag mv*:
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-004.html#3-4> [Lesedato 15.08.2009]

Forskrift til opplæringsloven (2006:§3-11): *Undervegsvurdering*.
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-005.html#3-11> [Lesedato 27.11.2009]

Forskrift til opplæringsloven (2006:§3-18): *Standpunktkarakterar i fag*.
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-006.html#3-18> [Lesedato 02.08.2009]

Forskrift til opplæringsloven (2006: §3-30): *Lokalt gitt eksamen i vidaregåande opplæring*.
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-008.html#3-30> [Lesedato 09.09.2009]

Forskrift til opplæringsloven (2006: §3-31): *Hjelpemiddel til eksamen*.
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-008.html#3-31> [Lesedato 28.08.2009]

Forskrift til opplæringsloven (2006: §3-37): *Annullering av eksamen*
<http://www.lovdata.no/for/sf/kd/td-20060623-0724-008.html#3-37> [Lesedato 28.08-2009]

Freeman, Samuel (2008): *Original Position*. <http://plato.stanford.edu/entries/original-position/#ArgForDifPri> [Lesedato: 02.06.2009]

Gjørven, Rita (2009): *lekser og læring*. <http://www.dagsavisen.no/meninger/article400810.ece>
[Lesedato 26.04.2009]

Gordon, Scot (1973): *John Rawls, difference principle, utilitarianism and the optimum degree of inequality*. In *journal of philosophy*.70 (9) (1973): 275-280
http://books.google.no/books?id=Uqlf2P_9FqEC&pg=PA177&lpg=PA177&dq=John+Rawls+difference+principle&source=bl&ots=nwj4GT6QZl&sig=G_ZFNfdLtJ8FqjjsSKAjBSQjGuM&hl=no&ei=3Tm3SuuUOJX4mwPji_DUDw&sa=X&oi=book_result&ct=result&resnum=10#v=onepage&q=John%20Rawls%20difference%20principle&f=false [Lesedato 23.08.2009]

Grønmo, Liv Sissel med flere (2004): *TIMSS 2003 Med få ord*.
http://udir.no/upload/Rapporter/TIMMS_kortversjon.pdf [Lesedato 25.08.2009]

Haugfos, Tine (2004): *Dere er blitt narret*. <http://www.aftenposten.no/jobb/article930687.ece>
[Lesedato 10.02.2009]

Hylland Eriksen, Thomas (2004): *Hijaben og "de norske verdiene"*.
<http://folk.uio.no/geirthe/Hijab.html> [Lesedato 11.02.2009]

Kaur, Simmi og Kemp, Halvorsen Ida (2009): *Hold styr på eksamensnervene*.
<http://www.aftenposten.no/jobb/article1784282.ece> [Lesedato 20.08.2009]

Kemp, Halvorsen Ida (2008): *Jukserne straffes for strengt*
<http://www.aftenposten.no/nyheter/iriks/article2372440.ece> [Lesedato 20.08.2009]

Lamont, Julian og Favor, Chrsti (2007): *Distributive Justice*.
<http://plato.stanford.edu/entries/justice-distributive/> [Lesedato 14.02.2009]

Larsen, Botnan Kjersti (2009): *Utdanningsdirektoratets anbefalinger etter høring – endring av forskrift til opplæringsloven kapittel 3 og 4 og forskrift til privatskoleloven kapittel 4 – Individuell vurdering*.
http://www.utedningsdirektoratet.no/upload/Forsiden/Udirs_anbefaling_forskrift_vurdering.pdf [Lesedato 05.06.2009]

Lu, Catherine (2006): *World Government*. <http://plato.stanford.edu/entries/world-government/> [Lesedato 15.02.2009]

McLeod, Owen (2008): *Desert*. <http://plato.stanford.edu/entries/desert/> [Lesedato 24.04.2009]

Moe, Ingeborg (2009): *Klaget og gikk opp to karakterer*.
<http://www.aftenposten.no/jobb/article1417604.ece> [Lesedato 20.08.2009]

Mosbakk, Kjersti (2008): *Kritiske til evalueringen av lærere*
http://www.utdanningsnytt.no/templates/udf20_17978.aspx [Lesedato 06.01.2009]

Opplæringsloven § 4-8. <http://www.lovdata.no/for/sf/kd/td-20060623-0724-009.html#3-5>
Pisa <http://www.pisa.no/index.html> [Lesedato 14.04.2009]

Olsen, Falck Astrid (2009): *Skoleeksempelet på juks*.
<http://www.aftenposten.no/meninger/kronikker/article3271171.ece> [Lesedato 20.09.2009]

Philpott, Dan (2009): *Sovereignty*. <http://plato.stanford.edu/entries/sovereignty/> [Lesedato 15.02.2009]

PIRLS (2001): *Leseferdighet blant 10-åringer i Norge 2001*.
http://udir.no/upload/Rapporter/PIRLS_norsk_del_faktaark.pdf [Lesedato 10.10.2008]

PISA (2003): *Resultater fra PISA 2003*
http://udir.no/upload/Rapporter/PISA2003_presentasjon.pdf [Lesedato 10.10.2008]

Ramberg, Bjørn og Gjesdal, Kristin (2005): *Hermeneutics*:
<http://plato.stanford.edu/entries/hermeneutics/> [Lesedato 04.09.2009]

Scriven, Michal: <http://www.cgu.edu/pages/4745.asp> [Lesedato 06.01.2009]

Seierstad, Åsne (2003): *Anarki og lovløshet i Bagdad*.
<http://www.aftenposten.no/nyheter/uriks/irak/article526803.ece> [Lesedato 19.02.09]

Skolenettet.no (2009): *TIMSS, PISA, PIRLS og Nasjonale prøver*.
<http://www.skolenettet.no/templates/News.aspx?id=54771&epslanguage=NO> [Lesedato 13.10.2009]

Strand, Arne (2008): *skatt og skole enda en gang*.
<http://www.dagsavisen.no/meninger/article329683.ece> [Lesedato 20.03.2009]

Støme, Kristin (2008): "*Ekteskapslovens mange muligheter*".
<http://www.aftenposten.no/meninger/debatt/article2524291.ece> [Lesedato 10.02.2009]

Sætre, Eriksson Mari (2006): *Forvirring rundt eksamensfusk*.
<http://www.aftenposten.no/jobbs/article1572948.ece> [Lesedato 10.08.2009]

Ullmann-Margalit, Edna (2006): "*Family fairness*". <http://ratio.huji.ac.il/dp/dp427.pdf>
[Lesedato 12.02.2009]

University of Victoria (2004): *The multiple choice exam*
<http://www.coun.uvic.ca/learning/exams/multiple-choice.html> [Lesedato 20.05.2009]

Viksmo Slettan, Olav (2007): *Gutter er skoletapere*.
<http://www.aftenposten.no/meninger/debatt/article2077482.ece> [Lesedato 28.11.2009]

Wenar, Leif (2008): *John Rawls*. <http://plato.stanford.edu/entries/rawls/> [Lesedato 13.02.2009]

Wertheimer, Alan (2008): *Exploitation*. <http://plato.stanford.edu/entries/exploitation/>
[Lesedato 25.04.2009]