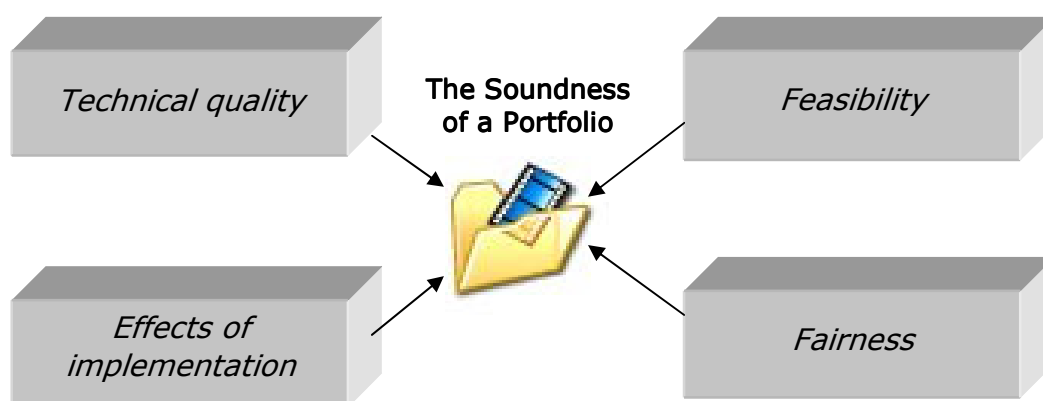


Mapper og kvalitet



Mappevurdering i lys av reliabilitet og validitet

Innhold:

Figurliste.....	iv
Forord.....	v
Sammendrag.....	vi
1 Innledning.....	1
1.1 Mappevurdering	1
1.2 Forskningsfelt og problemområde.....	2
1.3 Eget utgangspunkt og begrunnelse for valg.....	3
1.4 Undersøkelseskontekst.....	4
2 Teoretisk rammeverk.....	6
2.1 Å bære ark	6
2.1.1 Digitale mapper.....	8
2.1.2 Ulike mappetyper	10
2.1.3 Åpne og lukkede mapper	12
2.1.4 Mappens innhold – prosess og produkt	13
2.1.5 Paradigmer i konflikt	15
2.1.6 Grad av transparens - usynlighet.....	16
2.2 Vurdering for ulike formål	17
2.2.1 Summativ vurdering – vurdering av læring	19
2.2.2 Formativ vurdering – vurdering for læring	20
2.2.3 Formative vurderingsformer	24
2.2.4 Vurdering der vurderingsformer kombineres.....	25
2.2.5 Vurdering av digitale mapper?	27
2.2.5.1 Refleksjon i læring.....	28
2.2.5.2 Muligheter og utfordringer ved digitalisering av mapper.....	29
2.3 Kriterier i vurdering.....	31
2.3.1 Kriterier	32
2.3.2 Implisitte og eksplisitte kriterier	32
2.4 Et internasjonalt blikk på bruk av mapper	35
2.4.1 Internasjonal forskning.....	35
2.5 Rammeverk for sertifisering av mappekvalitet	37
2.5.1 Teknisk kvalitet “The technical quality of a portfolio”.....	37
2.5.2 Ærlighet og rettferdighet “The fairness of a portfolio”.....	38
2.5.3 Ringvirkning ved å ta i bruk “The effects of implementation”	39
2.5.4 Egnethet “The feasibility of a portfolio”	39
2.6 Oppsummering	40
3 Metode	41
3.1 Metode for dokumentasjon av mappekvalitet.....	42
3.2 Problemstilling.....	42
3.3 Forskningsdesign	43
3.3.1 Vitenskapsteori, forskningsperspektiv og paradigme	43
3.4 Utvalgsplan	45
3.4.1 Utvalgsstørrelse og metode	45
3.4.2 Utvalgsenhet.....	46
3.5 Innsamling av data.....	49
3.5.1 Forskerrolle	49
3.5.2 Det åpne kvalitative intervju	50
3.5.3 Bruk av fokusgrupper og intervju	51

3.5.4	Intervjuguide	53
3.6	Analyseplan	53
3.6.1	Analyse av dataene	53
3.7	Undersøkelsens gyldighet.....	54
3.7.1	Intern gyldighet	54
3.7.2	Forskningshorisont.....	55
3.7.3	Generalisering og overføring av resultatene til andre utvalg.....	56
3.8	Oppsummering og kommentarer.....	56
4	Empiri.....	57
4.1	Datamaterialet og utvalgte presentasjonskategorier	57
4.2	Forståelse og tolkning	58
4.2.1	Begrepene	58
4.2.1.1	Mappe	58
4.2.1.2	Formativ vurdering	61
4.2.1.3	Kriterier	65
4.2.2	Fenomenene.....	67
4.2.2.1	Formativ vurdering i lys av reliabilitet og validitet	68
4.2.2.2	Kriterier i lys av reliabilitet og validitet	71
4.2.2.3	Mapper i lys av kvalitet.....	73
4.3	Oppsummering	76
5	Sammenfatning og drøfting	77
5.1	Mappevurdering i lys av teoretisk rammeverk.....	77
5.1.1	Mapper.....	77
5.1.2	Formativ vurdering	79
5.1.3	Kriterier	81
5.2	Mappevurdering i lys av teknisk kvalitet	83
5.2.1	Reliabilitet og validitet	84
5.2.1.1	Formativ vurdering	84
5.2.1.2	Kriterier	87
5.3	Konklusjon i lys av problemstilling.....	87
5.3.1	Fenomenforståelse	88
5.3.2	Utvikling og utøving.....	89
5.3.3	Innpassning i Herman & Winters sertifiseringsmodell.....	89
5.4	Oppsummering	91
6	Konklusjon.....	94
6.1	Undersøkelsens resultater	94
6.2	Videre forskning	96
6.3	Avslutning	96
	Litteratur	98
	Vedlegg.....	107

Figurliste:

Figur 1: Husk de tre (Piet Hein)	5
Figur 2: Åpne og lukkede mapper (Otnes)	12
Figur 3: Vurderingsdimensjoner (National Research Council).....	17
Figur 4: Lineær og hypertextuell forståelse.....	30
Figur 5: Balanse i utforming av kriterier.....	34
Figur 6: The Soundness of a Portfolio	37
Figur 7: Rik mappekontekst.....	41
Figur 8: Kunnskapsstigen	43
Figur 9: Mappedvurdering ved sykepleierutdanningen.....	47
Figur 10: Mappedvurdering ved fotojournalistutdanningen.....	48
Figur 11: Den hermeneutiske spiral.....	49
Figur 12: Den som tvivler (Piet Hein).....	56
Figur 13: Forskningsdimensjoner	57
Figur 14: At vide hvad (Piet Hein).....	76
Figur 15: Teknisk kvalitet i lys av læring.....	90
Figur 16: Dokumentasjon av mappens tekniske kvalitet ved SU og FJU	91
Figur 17: Mappedvurdering i lys av kvalitet	92
Figur 18: Den virkelig vise (Piet Hein)	93
Figur 19: I evighetens perspektiv (Piet Hein).....	97

Forord

Nå nærmer tiden seg for avslutning av arbeidet med denne masteroppgave. Min mastergrad er i utgangspunktet bygget på ulike delkomponenter, der de første ble gjennomført ved IT-seksjonen (ikt og læring) på lærerutdanningen ved Høgskolen i Oslo (HiO), og de siste ved IT-seksjonen (ikt i læring) på lærerutdanningen ved Høgskolen Stord/Haugesund (HSH).

Generelt vil jeg takke IT-seksjonen/læringsmiljøet ved HSH, for den forståelse og tålmodighet de har vist meg som fjernstudent, i arbeid med oppgaven. Jeg har følt meg veldig godt ivaretatt, og er svært takknemlig for det. Å være fjernstudent kan av og til oppleves ganske isolert, og mange ganger har jeg nok savnet å være fysisk til stede, i et nærmiljø med mulighet for ansikt til ansikt diskusjoner. Men samhandling og samarbeid med IT-seksjonen ved HSH via e-post eller LMSet Classfrontier, har vært upåklagelig. I dette samarbeid vil jeg spesielt takke min veileder, førsteamanuensis Knut Steinar Engelsen, som har bidratt med innspill og uvurderlig konstruktiv veiledning i skriving av oppgaven. Oppmuntring og veiledning fra ham, har bidratt til at arbeidet sakte men sikkert har gått fremover, og nå nærmer seg en avslutning. Også respondentene som deltok i den empiriske undersøkelse må jeg takke. At både fotojournalist- og sykepleierutdanningen ved HiO sa seg villig til å stille opp, setter jeg stor pris på. Uten dem hadde det ikke vært så greit å finne representative studiemiljøer tilknyttet forskningsfeltet. Jeg takker derfor så mye for tanker og erfaring de har delt med meg underveis i prosessen, og for at de tok seg tid til å reflektere over de prioriterte forskningsspørsmål. Jeg sender også en stor takk til høgskolelektor Vibeke Bjarnø ved IT seksjonen på lærerutdanningen ved HiO, for den jobben hun har gjort med å korrekturlese oppgaven. Den jeg likevel må takke mest av alt, er min mann Erik Urdal, som har vært villig til å forsørge meg, mens jeg har forsket. I ettertid ser jeg at uten hans hjelp og støtte hadde jeg nok ikke kunnet prioritere å avsette tid og ressurser til en slik omfattende og krevende oppgave. Jeg håper han nå ser at arbeidet har båret frukter, og kan representere verdifull lærdom og erfaring for meg i fremtiden.

Å begynne på en masteroppgave når man er godt over 50, er i seg selv en utfordring. Når man i tillegg får sitt første barnebarn underveis i prosessen, er det nærliggende å stille spørsmål ved, om det nå ikke snart er på tide å slutte å studere. Erfaringen er at arbeid med denne oppgave har gitt meg verdifull innsikt, i en arbeidsmetode jeg overhodet ikke hadde kjennskap til på forhånd. Konklusjonen er derfor at man aldri blir for gammel til å lære!

Nesbru 15. mai 2006

Henni Sneibjerg Urdal

Sammendrag

Prosjektets overordnede mål har vært å belyse mappekvalitet. Masteroppgaven retter spesielt søkelys mot mappekvalitet knyttet til bruk av formativ vurdering og kriterier. Forskningsmessig har hensikten vært å undersøke utvalgte studiemiljøers forståelse av definerte begreper og fenomener, der den presise problemstillingen har tatt sikte på å forstå:

1. Hvilken forståelse knytter praksisfeltet til fenomenene kvalitet, reliabilitet og validitet?
2. Hvordan utvikles og utøves fenomenene i en mappevurderingskontekst, der formativ vurdering og bruk av kriterier inngår?
3. Hvordan passer fenomenforståelsen inn i sertifiseringsrammeverket utviklet av Herman & Winters?

Datainnsamlingen ble gjennomført senhøstes 2005, ved hjelp av fokusgruppe- og telefonintervju. Respondentene/informantene representerte henholdsvis sykepleier- og fotojournalistutdanningen ved Høgskolen i Oslo.

I oppgavens kapittel 1 reflekteres det kort over valgt forskningsfelt og problemområde. I tillegg grunngis valgene. Kapittelet beskriver hva som er undersøkt, hvilket studiemiljø undersøkelsen er gjennomført i, og hva som har vært motivet for undersøkelsen.

I kapittel 2 presenteres det teoretiske rammeverk. Her trekkes frem aktuell teori, som synes relevant og sentral i forhold til begrepene **mappe, vurdering og kriterier** og fenomenene **reliabilitet, validitet og kvalitet**. Spesielt Williams (1998, 2000b) uttrykte forståelse av reliabilitet og validitet, og Herman & Winters (1994) sertifiseringsrammeverk som utgangspunkt for dokumentasjon av kvalitet i arbeid med mapper, belyses i kapittelet. Sertifiseringsrammeverket er siden videreutviklet av Carney (2004). Også internasjonal forskning belyses kort i kapittelet. Dette inkluderes i det teoretiske rammeverk, som generelt grunnlag for forståelse av mappepraksis, i lys av internasjonal forskning.

Kapittel 3 beskriver undersøkelsens metodiske rammeverk. Med utgangspunkt i en undersøkende problemstilling, ble kvalitative data innhentet fra tre datakilder henholdsvis det åpne kvalitative forskningsintervju, sentrale skriftlige dokumenter og enkelte studentoppgaver med tilhørende respons/veiledning (datatriangulering). Hensikten med undersøkelsen var å studere, tolke og få kunnskap om, hvordan utvalgte studiemiljøer forsto, utviklet og utøvde begrepene og fenomenene nevnt i problemstillingen over. Studien var således rotfestet i hermeneutisk vitenskapsteori, der det subjektivt talte og nedskrevne språket ble utgangspunkt for forståelse. Respondentene ble sammensatt i fokusgrupper, og besto av personer som bekledd ulike roller innen studiemiljøet. Alle intervjuer ble tatt opp på

lydbånd, og siden i sin helhet ordrett nedskrevet. Empirien ble deretter innholdsanalysert, og knyttet til hensiktsmessige kategorier. Undersøkelsen ble til slutt sammenholdt med intern og ekstern gyldighet.

I kapittel 4 presenteres de empiriske funn innenfor en begreps- og fenomendimensjon. Dimensjonene sammenholdes med områdene tilknyttet problemstillingen, og beskriver en subjektiv forståelse av datakildene, som lå til grunn for undersøkelsen. Foruten tolkninger presenteres sitater fra intervjuene, ordrette nedtegninger fra skriftlige dokumenter, og nedtegninger fra studentoppgaver.

I kapittel 5 drøftes de empiriske funn i lys av teori presentert i kapittel 2. Spesielt tre drøftingsområder trekkes frem; mappevurdering i lys av a) teoretisk rammeverk, b) teknisk kvalitet og c) konklusjon i lys av prosjektets problemstilling. Under punkt **a** drøftes fokusgruppene forståelse av begrepene mappe, formativ vurdering og kriterier på generelt teoretisk grunnlag. Under punkt **b** drøftes gruppenes forståelse av formativ vurdering og bruk av kriterier i lys av fenomenene reliabilitet, validitet og kvalitet. Drøftingen knyttes her spesielt til Williams uttrykte syn på reliabilitet og validitet og dokumentasjon av mappekvalitet i lys av Herman & Winters sertifiseringsrammeverk. Hva gjelder sistnevnte, konsentrerer drøftingen seg spesielt om å belyse et av fire dokumentasjonsområder i sertifiseringsrammeverket, nemlig **teknisk kvalitet**. Denne drøftingsdelen anses å utgjøre hovedessensen i forskningsprosjektet. Under punkt **c** trekkes konklusjoner i lys av prosjektets problemstilling. Kapittelet avsluttes med en oppsummering.

Oppgavens avsluttende kapittel trekker konklusjoner, og antyder mulige fremtidige forskningsoppgaver. De empiriske funn i undersøkelsen, kan tyde på at pålitelig og gyldig formativ vurdering ved de to undersøkte læringsmiljøene, fant sin referanse i samhandling og fagforankring på studiet. Funnene kan dermed indikere at forståelse av pålitelig og gyldig formativ vurdering, avhenger av hvordan læringsmiljøer utvikler og utøver arbeidsmetodikken. Pålitelig og gyldig formativ vurdering blir på denne måten kontekstavhengig, og knyttes til ivaretagelse av funksjoner. De ønskede funksjoner fremstår således som styrende parametere for fenomenforståelse. Drøftingen antyder også, at når en mappe brukes som læringsredskap i formativ vurdering, er det mer naturlig å dokumentere mappens tekniske kvalitet, med utgangspunkt i mappens evne til å påvirke og fremvise dyp læring (Carney 2004), enn i mappens evne til å fremvise reliabilitet og validitet (Herman & Winters 1994).

Kapittel

1

Innledning

1 Innledning

1.1 Mappevurdering

Undersøkelsens overordnede problemstilling har vært å undersøke mappekvalitet, og har spesielt vært rettet mot å belyse formativ vurdering og bruk av kriterier. Vanligvis knyttes mapper til **arkivering** av en eller annen art, der hensikten som oftest er å etablere et lagringssystem, hvor man lett og fort finner igjen viktige skriftlige dokumenter.

I et utdanningsperspektiv har mapper for mange, fremstått som et hensiktsmessig sted å **oppbevare** studentarbeider, og vært samlende utgangspunkt for formell vurdering. Naturlig nok blir slike mapper kalt **vurderingsmapper** (Dysthe & Engelsen 2003), og innhold i mappen er fortrinnsvis laget med tanke på hva som formelt ønskes vurdert. Innhold i disse mapper styres da i stor grad av fastlagte føringer, der studenten ikke på fritt grunnlag kan velge mappeinnhold. Mappetypen knyttes lett til begrepet **mappevurdering**, et ord som kan gi inntrykk av at det nettopp kun dreier seg om vurdering. Stortingsmelding nr. 27 (2000) og Mjøsutvalget skisserer imidlertid et bredere perspektiv, når de anbefaler å bruke mappevurdering og deksamener, som en hensiktsmessig måte å kombinere undervisning og evaluering i ett studium. Læring og vurdering ses da som to sider av samme sak.

Departementet slutter seg til Mjøsutvalgets forslag om bruk av studentporteføljer/mappevurdering og deksamener som en hensiktsmessig måte for å kombinere undervisning og evaluering. (St.meld. nr. 27, 2000:5.3.5)

Mappevurdering knyttes altså fra politisk hold i Norge, til både undervisning og vurdering. Ut fra dette synet, involverer mappevurdering da både et lærings- og vurderingsaspekt. At mappevurdering omfatter begge deler, har også vært perspektivet i dette forskningsprosjekt. På 80-tallet innførte man for alvor prosessorientert skriving som metode i norsk skole og høyere utdanning. I den forbindelse var det hensiktsmessig å ta i bruk såkalte **læringsmapper** (jf. 2.1.2). Grunnprinsippet i prosessorientert skrivemetodikk, er utvikling av skrevet tekst underveis i en prosess, som involverer flere personer enn bare skriveren. Tekstutvikling skjer altså i samhandling, der skriverens tekstutkast gis tilbakemelding (respons) fra lærer eller medstudent. Underveis i arbeidsprosessen utfordres elevens/studentens/lærerens refleksjons- og vurderingsevne. Nettopp refleksjon og selvurdering knyttes internasjonalt til mappebegrepet, og begge anses som viktige læringsfremmende elementer. I prosessorientert skriving var det vanlig å bruke papirmapper og tekst utformet på ark. Ny digital teknologi legger imidlertid godt til rette for samhandling om og arbeid med tekst over tid, og kan således bidra til etablering av en ny læringsarena (Otnes 2003). Nye dokumen-

tasjonsformer og mediespesifikke virkemidler skaper andre muligheter til å strukturere tekst og formidle kunnskap (jf. 2.1.1). I undersøkelsen har nettopp digitale læringsmapper og samhandling om tekst/studieoppgaver, vært grunnleggende forutsetninger.

Ulikt læringsteoretisk syn vil bety ulik mappepraksis og bruk av ulike mappetyper (jf. 2.1.2), som igjen mest sannsynlig vil tjene ulike formål (jf. 2.1.5). Teoretisk læringsplattform og hensikt bestemmer da i tur og orden de aktiviteter som knyttes til mappen. Problemet oppstår når a) hensikten med mappen ikke blir fulgt og b) når mappen tar sikte på å tjene flere formål, som til dels er i konflikt. Brukes mappen til både å ivareta læreprosess underveis og som grunnlag for endelig formell vurdering, oppstår fort usikkerhet hos studenten, om hvilken hensikt mappen egentlig tjener. Er hensikten med mappen at den skal stimulere til læring, er det nettopp dette aspektet, som må ivaretas (jf. 2.2.2). Er derimot hensikten at den skal brukes til formell vurdering (jf. 2.2.1), må dette aspektet klart og tydelig fremstå. Generelt diskuteres det internasjonalt om det er klokt å bruke mapper både som lærings- og vurderingsredskap. Og dersom det er, hvordan kan det da dokumenteres at de ivaretar læring og vurdering på en **kvalitativ** god måte (jf. 2.5)? Også å kombinere både et lærings- og vurderingsaspekt innen samme mappe diskuteres. Oppgavens drøftingskapittel tar opp denne diskusjonen.

Når mappen brukes til vurdering oppstår også behovet for et vurderingsgrunnlag. Brukes mapper i formell vurdering, er det sannsynlig at klart definerte standarder (jf. 2.3), er den formelle vurderingsramme. Brukes derimot mappen til uformell vurdering, vil kriterier (jf. 2.3.1) ofte ligge til grunn for vurderingen. Hvor nødvendig er det da at man samarbeider om å utforme disse, hva skal de uttrykke, og hvordan skal man sikre at de forstås likt innenfor samme referanseramme? Det har vært viktige aspekter i undersøkelsen.

1.2 Forskningsfelt og problemområde

Det er vanlig å si at vurdering har tre hovedfunksjoner: a) å danne grunnlag for utvelging og en eventuell sertifisering (formell sluttvurdering jf. 2.2.1), b) å gi studenten informasjon om egen fremgang og stimulerer til bedre læring (formativ vurdering jf. 2.2.2), og c) å gi utdanningsinstitusjoner og lærere innsikt i nivå og eventuelle sterke og svake sider, ved utdanningens læringsmessige aspekter, slik at de eventuelt kan forbedres (Dysthe & Engelsen 2003). Men også eksamen har en relevant funksjon på høyere utdanningsnivå, som konsolidering av sentral og viktig kunnskap innen en disiplin (Kvale 2000). Generelt har utvelgingsfunksjonen hatt en dominerende plass i høyere utdanning og vært tett knyttet til kontroll av forventet lærdom på et gitt tidspunkt. Oppgavene som skulle vurderes ble gjerne utformet av lærere, og eksamensproduktet ble vurdert i henhold til et bestemt faginnhold. Her har det imidlertid skjedd store endringer. Både nasjonalt (Taasen, Havnes, & Lauvås 2004) og internasjonalt (Falchikov & Goldfinch 2000, Topping 2003) finnes flere eksempler på at selv- og medstudentvurdering nå legges inn i flere utdanninger, og dermed erstatter den

tradisjonelle evalueringen. Med nye vurderingsmåter følger også andre arbeidsformer. Her står samarbeid sentralt. Uformell vurdering av studiekraav/oppdaver underveis i et studium, skjer gjerne med det formål å hjelpe en student til å forbedre studiearbeidet, og utvikle seg i en læringsprosess over tid. Formell vurdering derimot knyttes til en endelig bedømming av studentens prestasjon på et bestemt tidspunkt. Alt etter hvilken vurderingsmåte som velges, er vurderingsparametrene knyttet til enten kriterier (formativ) eller fastsatte standarder (summativ). I mappevurdering forekommer ofte en kombinasjon av vurderingsmåter. Internasjonalt er det et omdiskutert tema.

Når mapper tas i bruk som arbeidsmetode i lærings- og vurderingsarbeid involveres studenter og lærere på en helt ny måte, og begge parter blir en aktiv del av vurderingsprosessen. Det er da naturlig at lærere og studenter samarbeider om kriterieutvikling. Utforming og synliggjøring av vurderingskriterier kan imidlertid oppleves vanskelig. Likevel når mappevurdering tas i bruk som arbeidsmetode, synes samarbeid om kriterier å være en nødvendighet. Ser man kvalitet i lys av internasjonale forskning, er kriterier et høyaktuelt forskningsområde.

Per dags dato er det ikke gjennomført mange studier, rettet mot å undersøke mappekvalitet, og hvilken konsekvens det har å bruke mapper som pedagogisk redskap (Zeichner & Wray 2000). Det gjelder både mapper brukt i læring og vurdering. Om hovedmålet med mappen er **læring**, må det være sterk dokumentasjon på at den virkelig fremmer læring. Er målet **vurdering**, trengs det dokumentasjon på at mappen virkelig er i stand til å vise mappeeierens viten og kompetanse innen et felt. Enda mer presserende blir forskning innen feltet, når digitale mapper tas i bruk (Carney 2004). Vår kunnskap knyttet til mapper som lærings-, samhandlings- og vurderingsredskap, er således ganske mangelfull. Her trengs mer forskning.

Hva er det så i den formative vurderingsprosess som oppleves som kvalitet, og kan dette knyttes til fenomener som reliabilitet og validitet? Slike spørsmål er ikke av ny dato, likevel viser nyere forskning, at det fortsatt er mange ubesvarte spørsmål på området. Nettopp belysning av kvalitetsaspektet har vært hovedessensen i dette forskningsprosjekt.

1.3 Eget utgangspunkt og begrunnelse for valg

Min vurderingserfaring har hovedsakelig vært knyttet til ivaretagelse av vurderingsfunksjonene: **a)** å danne grunnlag for utvelgning og en eventuell sertifisering, **b)** å gi utdanningsinstitusjoner og lærere innsikt i nivå og eventuelle sterke og svake sider ved utdanningens læringsmessige aspekter, slik at de eventuelt kan forbedres og **c)** eksamen som konsolidering av sentral og viktig kunnskap innen studiet (jf. 1.2). Disse funksjonene knyttes til formell vurdering. I valg av område for masteroppgave, var jeg i utgangspunktet interessert i å forske på digitale mapper. Ulike utdanningsnivåer satser i dag for fullt på dette, helt i tråd med føringer lagt i ulike utdanningsplaner og rammeverk, både på høyere og lavere

nivå (jf. nye læreplaner 2006, St.meld nr. 27 2000-2001, Program for digital kompetanse 2004). I planene knyttes mappevurdering til både et lærings- og vurderingsaspekt, og ses således som to sider av samme sak. Å forske på alle felt innen mappevurdering, ville imidlertid bli alt for omfattende i lys av prosjektets rammer. Valget falt derfor på å knytte forskningsprosjektet til vurderingsaspektet, da det alltid har interessert meg, både i rollen som student og lærer. Som beskrevet er min vurderingserfaring hovedsakelig knyttet til summative vurderingsmåter og formell vurdering. Behovet for innblikk i og kunnskap om en uformell vurderingsmåte syntes derfor innlysende. En uformell vurderingsmåte knyttes til **læring og læringsprosess**, der samarbeid og samhandling er viktige komponenter i forbedring av læringsarbeidet. I denne vurderingsmåte står altså læringsaspektet sentralt, i motsetning til formell vurdering, der ferdig produkt er det sentrale. I tillegg til interessen for digitale mapper som pedagogisk redskap, har altså behovet for innblikk i en alternativ vurderingsmåte (uformell vurdering), vært medvirkende årsak til valg av emne. Å få innblikk i og kunnskap om en vurderingsmåte, som først og fremst tar sikte på å hjelpe en student til forbedring underveis i en pågående læringsprosess, fremsto i lys av egen erfaring og føringer lagt i aktuelle planer og rammeverk, som et hensiktsmessig og meningsfullt forskningsområde. Valget falt til slutt på å undersøke hvordan utvalgte læringsmiljøer oppfattet, utviklet og utøvde kvalitet i arbeid med formativ vurdering, der det inngikk bruk av kriterier. I det inngikk belysning av pålitelighet og gyldighet.

1.4 Undersøkelseskontekst

En forutsetning for å forske på mappekvalitet, er at man får tilgang til læringsmiljøer, som er opptatt av nettopp å få til det. Å gjennomføre en undersøkelse knyttet til mappekvalitet, kunne i utgangspunktet vært gjort i ulike miljøer og på ulike nivåer. I lys av prosjektets problemstilling, måtte imidlertid et krav være oppfylt. Mappearbeidet måtte inkludere formativ vurdering og bruk av kriterier. Kravet knyttes gjerne til mapper brukt i læringsarbeid. Det ble derfor lett etter miljøer som aktivt arbeidet med læringsmapper.

Mitt eget yrkesfaglige utgangspunkt er HiO, der jeg har jobbet med ikt ved Avdeling for Lærerutdanning og Avdeling for Estetiske fag. Nettopp ved denne høgskolen pekte to utdanninger seg ut, hva gjelder mappevurdering. **Sykepleierutdanningen** (SU), har siden høsten 1999 arbeidet med mappemetodikk, og tatt utfordringene knyttet til endring i arbeids- og vurderingsform, på alvor. De siste årene har de i tillegg jobbet mye sammen med studentene, om å utvikle en felles forståelse relatert til formativ vurdering og bruk av kriterier. Ved **fotjournalistutdanningen** (FJU) har mapper vært brukt siden 1995. I studieprogrammet (2005) uttrykkes det at en students studiearbeid skal samles i en mappe, og det brukes mappevurdering i undervisning. Også på denne utdanningen er mappevurdering en gjennomført arbeidsmetodikk, der det inngår formativ vurdering og bruk

av kriterier. Studiemiljøene har således jobbet med mapper i henholdsvis 6 og 10 år, hvilket tilsier at de har tilegnet seg god erfaring med konseptet.

Argumentasjonen ved SU for å ta i bruk mappevurdering var, at de ønsket å bringe læring mer i sentrum, og få studenter til å bli mer bevisst og reflektert over egen læring. Likeledes ønsket man å styrke knytningen mellom teori og praksis, og dele vurderingsansvaret mellom student og lærer. Til det ønsket man å bruke medstudentrespons. Men å dele vurderingsansvar mellom student og lærer, medførte også krav til at vurderingskriterier måtte være kjent og synlig for alle delaktig i vurderingen. De synlige vurderingskriterier skulle da brukes til å vurdere om produktet/arbeidskravet var godt eller dårlig. Å samarbeide med studenter om utforming av vurderingskriterier, var kanskje et av de viktigste grep som ble gjort, for å trekke studentene med i vurderingsarbeidet. SU har således de siste årene jobbet mye sammen med studenter, om å utvikle en mappemetodikk, som inkluderte bruk av både formativ vurdering og kriterier. Utdanningen tilfredsstilte derfor kravene forskningsprosjektet hadde til praksisfeltet. Også FJU arbeidet med mappevurdering i henhold til forskningsprosjektets krav. Med erfaring følger også refleksjoner knyttet til hva som er bra eller mindre bra, og hva som anses som kvalitet/ikke kvalitet. Håpet var, at undersøkelsen skulle få frem erfaringene disse studiemiljøene hadde opparbeidet seg.



(Hein 1968:12)

Figur 1: Husk de tre (Piet Hein)

Kapittel

2

Teori

2 Teoretisk rammeverk

2.1 Å bære ark

Hensikten med en mappe/portfolio har gjerne, i et utdanningsperspektiv, vært å samle alt en student leverer av ferdige studiearbeider, på et for utdanningen hensiktsmessig sted. Studentarbeidene har i tillegg vært grunnlag for en formell vurdering. Ofte har mapper vært forbundet med noe av fysisk håndfast form. Eksempelvis kan nevnes Avdeling for Estetiske fag ved HiO, som i stor grad bruker håndfaste mapper der innholdet manuelt kan vises frem, tas på og vurderes. Mapper av denne typen refereres gjerne til som **kompetanse-mapper** (Dysthe & Engelsen 2003). Innholdet i disse er i første omgang ment håndfast å dokumentere hvilken kompetanse studenten besitter. På en utdanning knyttet til estetiske fag er slike mapper, fagets egenart tatt i betraktning, hensiktsmessige. Mappen brukes senere som kompetansefremvisningsorgan i egenprofilering ved jobbsøknad og liknende. Også i prosessorientert skriving ble det brukt mapper, gjerne referert til som **læringsmapper** (jf. 2.1.2). Det var innen denne tradisjonen mappevurderingsbegrepet oppsto på slutten av 80-tallet.

Mange teoretikere har bidratt i debatten rundt hva som kjennetegner en mappe. Stiggins (1994) definerer for eksempel en mappe som en **samling** av studiearbeider som viser fremgang og forbedring, og fungerer som **katalysator** knyttet til **kommunikasjon** rundt en students faglige utvikling. Allerede i 1991 understreket Paulson, Paulson & Meyer (1991:60) at mappemetodikk måtte omfatte studentmedvirkning når det gjaldt **valg** av *mappeinnhold, utvalgskriterier, vurderingskriterier* og *selvrefleksjoner*. Deres innspill tilførte allerede da en ny dimensjon i forståelsen av mapper, en dimensjon som i dag omfavnes av mange mappeetterkommere. Dysthe og Engelsen (2003) beskriver en mappe i pedagogisk sammenheng med fire kjennetegn:

For det første er ei "mappe" ei samling av fagleg arbeid gjort over ein lengre tidsperiode ...
For det andre er det viktig å presisere at ei mappe kan inneholde alle typar arbeid, ikke berre ark med tekst og/eller bilde ...
For det tredje blir portfolio-omgrepet internasjonalt knytta saman med refleksjon og sjølvvurdering ...
For det fjerde inneber bruk av mapper at både prosess og produkt er viktige ...
(Dysthe & Engelsen 2003:13/14)

Dysthe og Engelsen uttrykker her at en mappe i pedagogisk sammenheng kan kjenne- tegnes ved at den er en samling **faglige** studiearbeider gjort over **lengre tid**, som kan inneholde arbeider av **ulike sjangrer** og **type**, aktuelt å vurdere i et studium. Studie- arbeidene kan være utformet av studenten alene eller i samarbeid med andre medstu- denter. Internasjonalt knyttes mappebegrepet til **refleksjon** og **selvvurdering** av fag,

læreprosess og faglig utvikling. Underveis i prosessen forekommer en form for **metakognisjon**, en overordnet erkjennelse av hva som skal til for å lære seg faget, å lære seg å lære. I arbeid med mapper er både produkt og prosess viktig og omfatter således **både læring og vurdering**. På denne måten utfordres etablerte arbeids- og vurderingsformer. Læring skjer i en **prosess**, der studenten underveis **reflekterer, selv vurderer, og medvurderes** av andre medstudenter eller lærere. Dette resulterer gjerne i et endelig **produkt**, som **vurderes**. Dysthe (2002) supplerer Dysthe og Engelsens fire mappekjennetegnene, med vurdering av **ferdig produkt** og antyder at også sluttvurdering er en viktig side ved mapper. En sluttvurdering av mappens innhold, kan gjerne kombineres med andre former for vurdering, for eksempel muntlig eksamen.

Også The Northwest Evaluation Association definerer mappe som en hensiktsmessig samling av studentarbeider, der studentens innsats, fremskritt og tilegnelse på ett eller flere områder over en tidsperiode, fremvises. Mappesamlingen bør inkludere aktiv deltakelse fra studenten i valg av innhold, utvalgsriterier, vurderingsriterier og fremvise selvrefleksjon.

A purposeful collection of student work that illustrates efforts, progress, and achievement in one or more areas [over time]. The collection must include: student participation in selecting contents, the criteria for selection, the criteria for judging merit, and evidence of student self-reflection. (Barrett 2005:4)

Slik man ser av presentasjonen over fremhever alle at mappe er en hensiktsmessig og systematisk samling av studentarbeider, der studentens innsats og tilegnelse over tid på ett eller flere områder, fremvises og vurderes. Mappen kan inneholde alle typer arbeid og sjangrer, og kan involvere både prosess og produkt, samt knyttes til refleksjon og selv-vurdering. Arbeidsformen bør involvere aktiv studentmedvirkning med hensyn til valg av mappeinnhold, utvalgsriterier og vurderingsriterier. I noen tilfeller vil det være en styrke å kombinere vurdering av mappeinnhold, med andre former for vurdering, for eksempel muntlig eksamen.

Læringsteori har sitt utspring i psykologi og avhenger generelt av synsvinkel man velger å se læring ut fra. Gjennom årene har ulike teoretiske og pedagogiske retninger, i tur og orden, utviklet seg, og bidratt med en forklaring på læring som fenomen. Knytter man kjennetegnene for en mappe i pedagogisk sammenheng, slik de er beskrevet over, til læringsteori, får man assosiasjoner til en sosiokulturell læringstradisjon. Innen denne tradisjon oppstår læring i samspill mellom mennesker, og er noe langt mer enn det som skjer i hodet til den lærende. Læring påvirkes av både individet selv, omgivelsene og medmenneskene omkring. I prosessen står kommunikasjon og språk sentralt, likeså balansen mellom det individuelle og det kollektive. Læringstradisjonen er til dels presentert som et alternativ til, og dels en videreføring av kognitiv læringsteori, og har de siste femten årene fremstått, som retning innen psykologisk og pedagogisk forskning. Brukes mapper i læring med samhandling som dominerende prinsipp, synes derfor et naturlig valg for pedagogisk ståsted å være en sosiokulturell plattform.

2.1.1 Digitale mapper

Bruk av mapper har også fått fotfeste i norsk utdanning og fremstilles i dag av mange som problemløser for utfordringene knyttet til lærings- og vurderingsarbeid. Med slike tendenser blir det viktig å forske på hvilken betydning bruk av mapper egentlig har, både i norsk og internasjonal utdanning.

Hvis mapper og mappevurdering representerer noe nytt, blir det jo avgjørende å forstå hva det er nytt i forhold til, hva som er konteksten.
(Madsen 2003:159)

I litteraturen forekommer ordene **elektronisk** og **digital** ofte synonymt. Noen forskere (Kankaanranta 2002, Barrett 2000) har vært opptatt av å skille de to begrepene, og knytter forskjellen til mappeinnhold. Diskusjonen rundt dette går ikke denne oppgaven noe nærmere inn i. Begrepet elektronisk og digital vil derfor her være ensbetydende.

The National Learning Infrastructure Initiative (NLII) uttrykker om en elektronisk mappe:

... a collection of authentic and diverse evidence, drawn from a larger archive representing what a person or organization has learned over time on which the person or organization has reflected, and designed for presentation to one or more audiences for a particular rhetorical purpose.
(Barrett 2005:5)

NLII ser altså en elektronisk mappe, som en presentasjonsmappe beregnet på utvalgte personer, og har spesielt retorikk som mål. Mappen fremviser en samling av utvalgte og autentiske arbeider, hentet fra et større arkiv, for å vise hva en person/organisasjon over tid har lært, reflektert over og endelig utformet, for så å vise frem ferdig produkt til utvalgte personer. På denne måten tenderer NLII's mappedefinisjon i retning av det Dysthe og Engelsen kaller en vurderingsmappe. Det uttrykkes riktignok ikke i definisjonen at mappen er rettet mot endelig vurdering, kun at den har som formål å legge til rette for retorikk. Til å definere en digital mappe kombinerer Barrett en tradisjonell mappedefinisjon med digital teknologi. Barrett antyder at digitale mapper bruker elektronisk teknologi som en beholder/container, der studenten får mulighet til å samle og organisere sine digitale artefakter (digital video, lyd, grafikk, tekst og lignende). Teknologien øker mulighetene for arkivering, lenking/tenking, historiefortelling, samarbeid og publisering. I organisering og fremvising av materialet kan det benyttes hypertekstlenker på en hensiktsmessig måte. For Hildegunn Otnes representerer digitale mapper ikke kun en arkivskuff men også en læringsarena. Digitale mapper dokumenterer ikke kun at læring har funnet sted men også hvordan. Når digitale mapper dokumenterer hvordan læring har funnet sted, fremstår de som en læringsarena.

... blir de digitale mappene ikke bare en dokumentasjon på at læring har funnet sted, men også på hvordan den har foregått. ... slik blir de digitale mappene ikke bare en arkivskuff, men også en læringsarena.
(Otnes 2003:109)

Å bruke digitale mapper som læringsredskap og læringsarena har flere fordeler i arbeids- og læringssammenheng, antyder Otnes. Hun uttrykker at **mediespesifikke virkemidlene**

og **dokumentasjonsformer** skaper nye muligheter for å strukturere tekst og formidle kunnskap. Der tekstprosessen står i fokus kan digitale tekster stadig oppdateres, forbedres og utvides. På denne måten legger digitale mapper godt til rette for arbeid med tekst over tid, og medvirker til etablering av en digital læringsarena. Otnes fremhever spesielt **åpne** digitale mapper, og uttrykker at disse gir autentiske skrivesituasjoner og reelle mottakere. Slik legger de til rette for arbeid med kollektive tekster, og samarbeid om fellesdokumenter på en helt ny måte. Som redskap for dialog og samarbeid er hun opptatt av **den synkrone nettsamtale** og **den kollektive hypertekst**. Otnes antyder at kollektive digitale mappe-tekster viser hvordan studenter drar nytte av hverandres kompetanse. På denne måten blir mappen *ikke* bare et redskap i den enkelte students læringsprosess, men også i medstudentens. Digitale og analoge mapper skal ikke fylles med det samme. Bruk av teknologi skal heller ikke bare bety en effektivisering av tradisjonelle arbeidsformer, men representere noe nytt, som er kvalitativt forskjellig fra tradisjonelle mapper (Otnes 2003:98). Det er derfor viktig at teknologi synliggjør, hvordan den kan være en forutsetning for og støtte opp om læring. I den prosess blir valg av dokumentasjonsform et viktig aspekt. Andre antyder at forskjellen på en tradisjonell og digital mappe i prinsippet ikke er så stor. Geir Winje skriver:

En digital mappe er i prinsippet ikke annerledes enn en tradisjonell mappe, den lagres og organiseres – som begrepet tilsier – digitalt. Den er ideelt transparent, det vil si åpen og tilgjengelig for læringsfellesskapet.
(Winje 2004:66)

En digital mappe favner i prinsippet alt det den tradisjonelle mappen gjør, men tilføres i tillegg det digitale element. Ideelt er mappen transparent, slik at den er åpen og tilgjengelig for alle i læringsfellesskapet. Winjes definisjon er på denne måten ganske lik Barretts definisjon av digitale mapper. M. Øhra skriver om digitale mapper:

En digital mappe kan beskrives som: en digital læringsarena som søker å utvikle et datastøttet dynamisk samarbeidsnettverk med kunnskapsdeling og kunnskapsutvikling som mål.
(Øhra 2004:18)

Digitale mapper er en læringsarena, som gjennom datastøttet læringsmiljø forsøker å utvikle nettverk for samarbeid og kunnskapsdeling/utvikling. Målet i det dynamiske samarbeidet er å utvikle og dele kunnskap. Også denne definisjonen favnes i definisjonen gitt av Barrett og Otnes. Sistnevnte fremhever at det rundt mapper ofte brukes ulike definisjoner/begreper, som til dels kan virke både forklarende og forvirrende og understreker at begrepsbruken først og fremst avspeiler, **hvilket syn man har på mapper** (som en arkivskuff eller som en læringsarena), og hvilken **hensikt** de skal tjene. Flere kriterier bør oppfylles/tilstrebes for at man kan snakke om digitale mapper som fenomen, og som kvalitativt skille fra andre mappetyper. Otnes selv (2003:88) uttrykker at en digital mappe må lagres og organiseres digitalt, enten åpent eller i et lukket rom på nettet, og må gjøre bruk av mediespesifikke virkemidler i læringsprosess og dokumentasjon. Dessuten må mappen/arbeidet bygge på velkjente tradisjonelle pedagogiske prinsipper, og være mer

eller mindre åpen med tilgang for andre tilknyttet læringsfellesskapet. Også disse prinsippene synes å passe godt inn under Barretts definisjon av digitale mapper.

Et vesentlig punkt i den digitale læringsprosess og dokumentasjon av den, er bruk av **mediespesifikke virkemidler**. Otnes viser her til begrepene multimedialitet, interaktivitet og hypertekstualitet. Et sentralt aspekt ved digitale mapper som arbeidsform, er muligheten for et ikke lineært aspekt. Hypertekstualitet gjør det mulig for mottaker/leser å følge/velge sin egen vei gjennom stoffet. Organisering av hypertekst, åpner for digital fremvising av sammenhenger på ulike måter. Selve hypertekstprinsippet gir nye muligheter til samhandling, samskriving og gruppearbeid, der hyperteksten lett integrerer hyperlenker til tekster utformet av andre personer. Hypertekst synliggjør på denne måten hvordan ulike tekster kan bygge på og veves inn i hverandre.

I hypertekster blir det mer konkret og visualisert hvordan tekster kan bygge på hverandre og veves inn i hverandre.
(Otnes 2003:106)

Det primære målet med referanselenking og bruk av hypertekstuelle tekster er kollektiv kunnskapsbygging, der man gjennom lenking viser at man har lest, forstått og lært det andre har resonnert seg frem til. Det eneste som teknisk kreves er at studenten lærer seg å lage nettsider og hyperlenker. Referanselenking fra eget arbeid til andres, tjener formålet å vise interesse for andres arbeid og å kontrollere. Hyperlenkereferanser til andre studenters tekster/arbeider, vil vise hvordan studenten i sin egen tekst har reflektert over og eventuelt inkludert tanker formidlet i medstudenters tekster.

Etter Otnes' syn er altså en digital mappe mer enn en arkivskuff, og like aktuell som **læringsredskap** og **læringsarena**, der vanlige pedagogiske prinsipper knyttes til prosess, samarbeid, refleksjon og studentmedvirkning. Er teknologien så likevel blitt så krevende at de færreste lærere og studenter er i stand til å bruke den, som det læringsverktøy den kunne være? Og befester lukkede e-læringsystemer til en viss grad gamle arbeidsmåter, slik at systemer av denne art nødvendigvis ikke fører til bedre læring? Spørsmålene til Otnes berører her slik jeg ser det, inn på to av fire dokumentasjonsområder (The Fairness of a Portfolio og The Feasibility of a Portfolio) fremhevet av Herman & Winters (1994), i dokumentasjon av mappekvalitet (jf. 2.5). Hvor rettferdig fremvises innholdet i en digital mappe, om ikke lærere og studenter er teknisk parate, og hvor egnet er da mappen?

2.1.2 Ulike mappetyper

Bruk av mapper der det benyttes **refleksjon** og **selvvurdering** kan ha to veldig ulike læringsteoretiske utspring, en instrumentell eller humanistisk/konstruktivistisk tradisjon. Dysthe og Engelsen (2003) bygger på Brown, Bull & Pendlebury, og uttrykker at det i den instrumentelle tilnærmingen ofte vektlegges arbeidskrav, som oppfyller detaljerte læringsmål og standarder. En slik tilnærming faller naturlig inn under summativ vurdering.

Det eine utspringet er i "the competency movement" og er instrumentelt i si tilnærming. Her blir det lagt vekt på å dokumentere kompetanse i høve til detaljerte læringsmål, og formålet er ofte eit slag utvida curriculum vitae som kan danne grunnlag for vurdering ved jobbsøknad (Dysthe & Engelsen 2003:15)

Målet med mapper innen denne tradisjonen, er ofte å etablere en slags utvidet CV, til bruk ved jobbsøknad. Faren med slike mapper kan være at selvvurdering og refleksjon først og fremst knyttes til internalisering av andres krav/standarder og normer, og ikke oppstår av interessen for å forstå og lære. Innen humanistisk/konstruktivistisk tradisjon brukes selvvurdering og refleksjon til personlig utvikling, som oftest i **samarbeid** med lærere og medstudenter. Målet er utvikling av ny forståelse og innsikt, med liten vekt på kontroll og fastsatte normer. En slik tilnærming knyttes til formativ vurdering.

Slik situasjonen er i dag appellerer en humanistisk/konstruktivistisk tradisjon, mest til lærere i høgere utdanning, mens et instrumentelt lærings syn appellerer til læringsssammenhenger, der spesifikke kunnskaper og ferdigheter skal dokumenteres i forhold til spesifikke læringsmål. Ulikt lærings syn krever ulike mappetyper. Dysthe og Engelsen nevner tre reindyrket og en blandet, men uttrykker også at grensen mellom mappetyperne ikke er absolutt.

Kompetansemappe er laga med tanke på å søke jobb, oppdrag eller studieplass. Formålet er å vise fram og dokumentere sin eigen kompetanse.

Vurderingsmappe har summativ vurdering som hovudformål. Mappede oppgåvene blir da primært laga med tanke på kva ein ønskjer å evaluere, ikkje ut frå kva læringseffekt dei kan ha. ...

Vurderingsmappa har 'sensor' som primær mottakar, anten denne er intern eller ekstern.

Læringsmappe signaliserer at mappa blir brukt som eit arbeids- og læringsreiskap og ikkje inngår i grunnlaget for formell evaluering. ... Studenten sjølv og medstudentar er 'mottakar'.

Formålet er å vere eit hjelpemiddel i studentens eigen læringsprosess, og oftast inngår også samarbeid med medstudentar og lærarar som eit sentralt aspekt.

Lærings- og vurderingsmappe kombinerer dei to meir reindyrka formene.

(ibid:16/17)

Rendyrkede **kompetansemapper** har primært jobbtilyder som mottaker, og målet er å fremvise en håndfast kompetanse. Mappetyper har sitt utspring i kunsthøgskole arkitektur. Kompetansemapper knyttes *ikke* til fremvising av læreprosess, men kun til ferdig produkt og brukes til fremvising av egenkompetanse innen et felt.

Rendyrkede **vurderingsmapper** brukes som grunnlag for formell sluttevaluering/eksamen (summativ vurdering), og har sensor, eller den som formelt vurderer, som mottaker. Mappens innhold/produkt vurderes isolert fra læringsprosessen, der man først og fremst er interessert i å vurdere i henhold til et fastsatt vurderingsgrunnlag/standarder. Vurdering av læringsprosess og mappen som læringsartefakt er altså fraværende her.

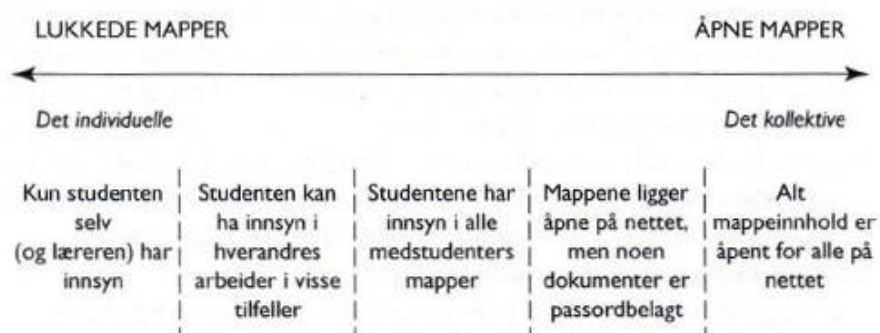
Rendyrkede **læringsmapper** brukes som redskap for kunnskapstilegnelse, i en læringsprosess støttet av uformelle vurderinger og tilbakemeldinger. Mappetyper har først og fremst som mål å være medierende (legge til rette for) redskap i samarbeid mellom lærere og studenter, og har begge parter som mottaker. Mappetyper knyttes til læringsprosess og

utvikling. Formell vurdering er fraværende her. Som fundamentale komponenter i arbeid med læringsmappe fremhever Zubizarreta (2004: 20) dimensjonene **refleksjon, dokumentasjon** og **samarbeid**.

En **kombinasjon** av vurderings- og læringsmappe forekommer også, der mappetypen ivaretar både læring og vurdering. Om hovedhensikten er å styrke sammenhengen mellom læring og vurdering, må en altså ha begge deler i tankene, når mappemodell utformes. Barrett fremhever at formålet med en elektronisk mappe bestemmes av dens innhold, og rettes enten mot **vurdering, markedsføring** av seg selv ved jobbsøking, eller **dokumentasjon** av læringsprosess/faglig utvikling (Barrett 2005:5). Formålet med en elektronisk mappe, styres av hensikt og funksjon den skal fylle. Barrets kjennetegn av en elektronisk mappe støtter på denne måten de rene mappetyperne beskrevet av Dysthe og Engelsen.

2.1.3 Åpne og lukkede mapper

Når digitale mapper brukes som arbeidsform, avhenger innsyn og samarbeidsmuligheter av i hvilken grad mappen er allment åpen. Ved å innskrenke eller åpne mappetilgang, bestemmes adgang til ulike deler av læringsmaterialet. I et lukket e-læringsystem av typen LMS¹, er det greit å opprette digitale mapper, der det lett lar seg gjøre å innskrenke eller utvide mappens åpenhet og innsynsrett, en egenskap som legger til rette for ulik grad av samarbeid. Det er et viktig poeng når for eksempel samskrivingstekster brukes som dokumentasjonsform. Ulik grad av mappens åpenhet/lukkethet skisseres av Otnes slik figur 2 viser.



Figur 3.1 Ulike grader av åpenhet i de digitale mappene

(Otnes 2003:90)

Figur 2: Åpne og lukkede mapper (Otnes)

Den mest åpne og kollektive digitale mappen i tabellen, plassert ytterst til høyre, er et lagringsområde på en Internettjener, der alle med tilgang til Internett, kan klikke seg inn og se mappens innhold. Mappen kan manifesteres i form av en hjemmeside, der man via hyperlenker, klikker seg videre til andre nettressurser. Slike åpne digitale mapper gir mulighet for autentiske skrivesituasjoner og varierte mottakergrupper, uttrykker Otnes. Til venstre finner man den individuelle og lukkede mappen, der kun den enkelte student og

¹ LMS – Learning Management System - for eksempel e-læringsplattformen Luvit og Classfrontier.

lærer har tilgang og innsynsrett. Skalaen illustrerer på en god måte ulike grader av åpen- og lukkethet i bruk av digitale mapper, der grad av åpen- og lukkethet vil avhenge av mappens funksjon, og pedagogiske plattform. Mappearbeid grunnlagt på et sosiokulturelt læringsperspektiv, vil involvere lærere og medstudenter som samarbeidspartnere og responsgivere. Det blir da nødvendig å bestemme grad av åpenhet, slik at mappen åpnes for dem som er tiltenkt en mottaker- og responsrolle.

2.1.4 Mappens innhold – prosess og produkt

Tradisjonelt har det på høyskoler og universiteter vært tendens til at studenter, for å bli vurdert, har måttet levere egenprodusert individuelle tekstoppgaver, da gjerne av arten få og store. Oppgaven(e) leveres som avslutning på et studium, og karaktersettes (summativ vurdering) ut fra bestemte standarder. Mottaker av oppgavene har vært lærere/sensorer, som fremfor alt har vurdert det enkelte produkt og ikke prosessen forut i utvikling av produktet. Otnes (2003) påpeker at bruk av digitale mapper som arbeidsform, legger til rette for et **skisma** på dette område. En digital mappe åpner for muligheter til større variasjon og mangfold i studentarbeider, noe som kan berike studerende og fagpersoner, både på lavere og høyere utdanningsnivå. Digitale mapper har et særlig potensial, hva gjelder mulighet for sjangervariasjon, oppbygging og sammenheng, men byr også på sentrale utfordringer. Mappemetodikk kan lett stille store krav til oppgaveproduksjon og dokumentasjon, som igjen krever stor aktivitet både foran og borte fra pc-skjermen. Utfordringen blir å møte dette dokumentpress, og finne arbeidsmåter som ivaretar både læringsprosess, refleksjon og varierte læringsformer, samtidig som det blir tilstrekkelig relevante og gode studiearbeider i mappa.

Som oftest er det vanlig med individuelle mapper, der arbeidene til den enkelte student står i fokus. Når læring bygges på et sosiokulturelt kunnskaps- og læringssyn trer imidlertid det **kollektive** aspektet frem, også hva gjelder dokumentasjonsform. Hypertekst (Otnes 2003) åpner for nye muligheter til kollektiv samhandling, samskriving og gruppearbeid, og inviterer til fremvisning av sammensatte tekstpresentasjoner og dokumentasjonsformer av digital art. Den synkrone nettsamtalen (chat) forutsetter i motsetning til den kollektive hypertekst at alle deltakere i læringsfellesskapet er til stede og pålogget systemet samtidig. Otnes har selv brukt synkrone nettsamtaler, som erstatning for muntlige samtaler ansikt til ansikt, og gir uttrykk for at fremdeles er f2f samtalen viktig og uerstattelig. Samtaler av digital karakter, har altså både sine fordeler og ulemper.

Hva gjelder dokumentasjonssjangrer i mappen, går det et klart skille mellom **saksprosa** og **skjønnlitteratur**. Saksprosa er gjerne dominerende dokumentasjonsform i høyere utdanning og inkluderer **argumenterende** og **informative** saksprosattekster (rapporter og analyser). Den såkalte "**brukssjanger**" (anmeldelser, leserbrev, intervju og diskusjonsinnlegg) er i den senere tid mye brukt som supplement til argumenterende og informative tekster. Otnes ser heller ikke bort fra at også skjønnlitterær skriving, kan brukes som redskap for

læring/formidling av kunnskap. Noen av de nevnte sjangrene passer like godt som dokumentasjon i en analog, som i en digital mappe, mens andre er typisk tilpasset det digitale mediet. I tekster tilpasset digitale medier, brukes mediespesifikke virkemidler. Et eksempel er hypertekstnovelle. For både å produsere og lese en hypertekstnovelle, er en forutsetning at man har tilgang til en skjerm. Skjermtekster har den mulighet at det i teksten kan kombineres både digital tekst, bilde og lyd. Teksten blir på denne måten multimedial, og leseren kan selv velge å gå i dialog, via aktive hyperlenker i teksten.

Mapper grunnlagt på en sosiokulturell læringstradisjon ønsker gjerne å dokumentere at læring skjer og hvordan. Spesielt å dokumentere **hvordan** læring har skjedd, byr på store utfordringer. Som dokumentasjon har det i stor grad vært brukt **logger** og **refleksjonstekster**, gjerne supplert med en form for egenvurdering, eventuelt vurdering fra medstudent. Otnes uttrykker om sjangeren:

Det er en litt særegen sjanger med uklare sjangerkriterier. Den kan ha ingredienser fra dagbok, rapporten, evalueringsskjemaet, (leser-)brevet og anmeldelsen, men utgjør en helt egen tekstkategori.
(Otnes 2003:97)

Sjangeren utgjør altså en helt egen tekstkategori, er litt særegen med uklare sjangerkriterier, og blir lett både subjektiv og personlig. For at studentene skal forstå hvordan refleksjonstekster blir faglig fundert, må de bevisstgjøres tekstinhold og tekstutforming. I tillegg må det være klart hvordan teksten vurderes, dersom den inngår som del i en vurderingsmappe. Omskriving av tekster er aktuelt i prosesskriving, der det åpnes for enkel redigering underveis. Mediespesifikke dokumentasjonsformer gjør det lett å forandre, slette, tilføye og flytte på tekst gjennom hele skriveprosessen. Teksten utvikles over tid og dynamisk. Men der tekst utvikles over tid og i stadig endring (dynamisk), skilles det **ikke** på utkast. Det er det første tekstutkastet som endres hele tiden. Her antyder Otnes at det med fordel kan utformes og arkiveres flere utkast i en mappe, for sammen med responsen å dokumentere læringsprosessen. En tekst blir egentlig aldri helt ferdig, og Otnes oppfordrer til å bruke begrepet "under construction", ikke bare på en uferdig nettside, men også i forbindelse med tekstutforming. Å vise så vel prosesstekst som presentasjonstekst i en digital mappe går også an. Digital strukturerings- og kommunikasjonsverktøy gjør det mulig å smelte sammen og gjengi både prosess- og presentasjonstekster (ferdige tekster). En mappetekst kan på denne måten gjengi både prosess og produkt (jf. hypertekstprinsippet). Dysthe og Engelsen antyder likevel (2003: 231) at mapelementer bare utgjør toppen av isfjellet, og kun **til dels** dokumenterer læreprosesser. Otnes gir sin støtte til det, men uttrykker også at visse digitale sjangrer dokumenterer læringsprosess bedre enn andre, og viser til den synkrone nettsamtale og den kollektive hypertekst (jf. kapittel 2.1.1). Å utvikle bred tekstkompetanse er viktig. Å utvikle tekst i digital sammenheng krever en enda mer allsidig tekstkompetanse enn vanlig, og favner både tekniske og tekstlige kompetansedeler (jf. multiliteracies Tyner 1998).

Som digital sjanger foreslår Barrett bruk av Weblogs (blogs), refleksive journaler, online diskusjoner, selvrappotering, kartlegging og digital historiefortelling og supplerer med dette Otnes. Blogs er av flere tatt i bruk som støtte i refleksjonsprosessen, og manifesteres i form av en Internettside, der innholdet organiseres etter dato. Mange uttrykker at denne dokumentasjonsformen er skreddersydd for digitale mapper.

A weblog is defined as any web page with content organized according to date.
(Barrett 2005:22)

Blogs kan være engasjerende og motiverende, når refleksjon og samarbeid foregår online. Digital historiefortelling, et moderne uttrykk av den gamle formen for historiefortelling, henter sin styrke i å kunne veve sammen digitale medier som bilder, musikk, tekst (fortelling) og lyd/stemme, slik at fortellingen fargelegges av digitalt liv til historiens innhold (personer, situasjoner, opplevelser og innblikk). Også for den digitale fortellingen er hypertekst hovedprinsippet. Det viktigste er til syvende og sist likevel å finne teknologiske verktøy, som engasjerer og motiverer studenter til refleksjon (Barrett 2005).

2.1.5 Paradigmer i konflikt

Allerede i 1994 påpekte Paulson & Paulson at det fantes ulike mappeparadigmer (syn på hvilken hensikt mappene skulle tjene), som i tur og orden oppfordret til mappeaktiviteter som var totalt forskjellig. Formålet med mapper i det positivistiske paradigmet var å vurdere et endelig læringsutbytte, som regel en standard bestemt av utenforstående.

The purpose of the portfolio is to assess learning outcomes and those outcomes are, generally, defined externally. Positivism assumes that meaning is constant across users, contexts, and purposes...
(Paulson & Paulson 1994:36)

Mappen fungerte som et gjemmede for studentarbeider, og ble brukt til å vise hvordan og hvor mye læring som hadde skjedd. Positivistene antok således at meningsforståelse var en konstant størrelse, både hva gjaldt studenter, sammenhenger og formål. Den digitale mappes primære formål var å formelt vurdere det rene læringsutbyttet. Slik premierte et positivistisk mappeparadigme, studentarbeider som oppfylte og fulgte eksternt oppsatte standarder. En students individuelle læringsprosess fremvist gjennom ulike arbeider, ble i liten grad premiert. På denne måten favnet paradigmet rene digitale vurderingsmapper, med fokus på endelig produkt. I det konstruktivistiske paradigmet var det den lærende selv som skapte mening og læring, og den digitale mappen fungerte som læringsarena.

The portfolio is a learning environment in which the learner constructs meaning. It assumes that meaning varies across individuals, over time, and with purpose.
(ibid)

Konstruktivismen antok at meningsforståelse varierte fra individ til individ, over tid og i forhold til målsetninger. Mappen synliggjorde læringsprosess, og gjenga aktiviteter som lå til grunn for læring. Et sammendrag av hver enkelt mappe ble for kompleks for en normativ

beskrivelse (vurdering etter norm). Læring skjedde prosessuelt i et lærende miljø, og kunne ikke beskrives ut fra en norm. Konstruktivistisk mappeparadigme hadde grunnleggende fokus på individuelle læreprosesser, og det var naturlig å bruke digitale mapper av typen læringsmappe (jf. 2.1.2).

Konflikten mellom det positivistiske og konstruktivistiske paradigme var knyttet til **hensikt** mappen skulle tjene, og hvilket pedagogisk grunnsyn som lå til grunn. For positivistene var hensikten å bruke mappen til fremvisning av ferdige normspesifiserte produkter, med liten fokus på individuelle forskjeller og læringsprosessen forut, mens hensikten for konstruktivistene var å fremvise individets læringsprosess, med liten fokus på produkt. Bruk av Internett teknologi og utvikling av en rekke kommersielle IT verktøy (jf LMS), de siste årene, åpner for forsterkning av denne konflikten. Mange av verktøyene lover støtte for både individuelle studentmapper og innsamling av formelle vurderingsdata, og tar sålede sikte på å møte læringsinstitusjoners behov for både individuell studentoppfølging, institusjonell godkjenning (vurdering) og rapportering i form av genererte rapporter.

Many of these systems promise support for student portfolios AND aggregated assessment data to meet reporting requirements.
(Barrett 2005:7)

Å love støtte til både studentmapper og institusjonell behandling av innsamlet vurderingsdata, representerer mappeparadigmer i konflikt og byr på store utfordringer, antyder Barrett. Jeg selv ser konflikten som en konsekvens av at ikke institusjonen er tro mot hensikten med mappene. Ved hjelp av IT verktøy er det fort gjort å hente ut data fra en læringsmappe, der hensikten kun er læring og ikke formell vurdering, og så bruke dataene i eventuell rapportering. Det blir feil! En digital mappe beregnet for å møte den lærendes behov, der det gis formativ vurdering i en støttende og forsterkende læreprosess, må brukes i henhold til det som var hensikten. Konflikten er et faktum om ikke mappehensikten overholdes 100 %, slik jeg ser det. For å møte konflikten foreslår Barrett å finne digitale løsninger, der det er mulig for studenter å vedlikeholde sitt eget digitale arbeidsarkiv, der de kan hente oppgaveeksempler og tilføre disse refleksjoner og notater. Fra samlingen kan det siden trekkes ut arbeider til andre mapper. Barrett foreslår her å skille klart både på offentlige og private mapper og lærings- og vurderingsmapper.

2.1.6 Grad av transparens - usynlighet

Et viktig aspekt som ofte nevnes ved omtale av digitale mapper er grad av transparens, noe som er til stede, men likevel usynlig. For M. Øhra (2004) innbefatter begrepet transparens både transparent teknologi og transparente læreprosesser. I dimensjonen transparent teknologi ligger en forventning om at teknologi skal være medierende redskap, og like usynlig som blyant og papir. Norman uttrykker om transparens:

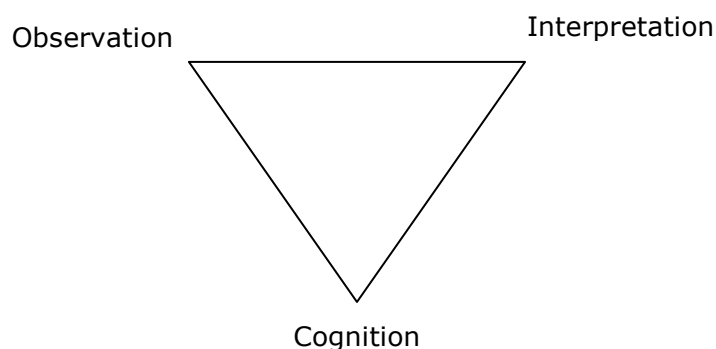
That is the end result, hiding the computer, hiding the technology so that it disappears from sight, disappears from consciousness, letting us concentrate upon our activities, upon learning, doing our jobs, and enjoying ourselves.
(Norman 1999:Vii-iX)

Teknologi skal usynlig tjene det menneskelige behov, og er en nødvendig faktor i det vi gjør. Men den skal i seg selv ikke være i fokus, kun et hjelpemiddel til å nå de mål vi setter oss. Hva gjelder dimensjonen transparente læreprosesser, vektlegger Øhra (2004) her det kollektive aspektet, og er opptatt av de usynlige læreprosesser, der studenter samarbeider og lærer i grupper, i en aktiv og kollektiv læringskontekst. Slik jeg tolker dimensjonen, er målet at kollaborasjon og fellesskap skal være katalysator og usynlig medierende redskap for enkeltindividets læring.

2.2 Vurdering for ulike formål

Hensikten med vurdering av læring fremstår som todelt, der den på den ene siden skal tjene som dokumentasjon av en students læring/faglige forbedring, og på den andre stimulere studenten til forbedring av studiearbeid. Vurdering bygger generelt på observasjon, tolkning og kognisjon (tanke/forståelse). De tre vurderingsdimensjonene/ pilarene er ubønhørlig knyttet sammen, danner en helhetlig metode for vurdering og må forstås i lys av hverandre.

I kognisjonspilaren legges rammer og strukturer for hvordan en student faglig kan vise innsikt og utvikle kunnskap og kompetanse innen sitt fagområde. Her bestemmes grunnlaget studenten skal bruke for å dokumentere sin viten og læring, oppnådd i en kognitiv læringsprosess innenfor gitte rammer. I observasjonspilaren bestemmes observasjonspraksis som gir observatøren tilgang til å observere studentdokumentasjonen, studentens kognitive prestasjoner/oppgaver. Tolkingspilaren sørger for en tolkningsmetode, til bruk for beslutningstaking knyttet til studentprestasjonene utformet i kognisjonspilaren og tilgjengelig i observasjonspilaren.



(National Research Council 2001:44)

Figur 3: Vurderingsdimensjoner (National Research Council)

Viktige faktorer i å forbedre vurdering er å utvide sin forståelse av hva som er viktig å vurdere, eventuelt finne bedre vurderingsmetoder eller forbedre sin evne til å tolke komplekse former av studentarbeider.

... advances in cognitive sciences, a broadened concept of what is important to assess, advances in measurement sciences, and expanded capability to interpret more complex forms of evidence. (Barrett 2005:8)

Vurdering må tilpasses situasjonen. En type vurdering passer ikke i alle situasjoner. Ofte brukes en enkel vurdering til mange formål. Jo flere formål en enkel vurdering tar sikte på å tjene, desto mer vil man måtte gå på akkord med hvert formål. Det trengs derfor en rikere og mer sammenhengende vurderingspraksis, slik jeg tolker Barrett. Hva som er viktigst å avklare, illustreres godt med dette sitatet hentet fra Otnes:

Det er ... viktig å ha en avklart forståelse av hva en mappe egentlig er, og hva den pedagogiske hensikten med en mappe er, før man diskuterer hva den kan inneholde. (Otnes 2003:109)

Før man i det hele tatt planlegger å ta i bruk digitale mapper, må det klart defineres hva hensikten med mappen skal være. En forutgående diskusjon anses som essensiell. Valg av mappetype vil avgjøre hvilke tekniske verktøy som trengs, og iverksetting av og system for ulike mappetyper velges etter behov. Barrett antyder at det kan være hensiktsmessig å supplere læringsmapper med weblogs (blogs), mens vurderingsmapper, der artefakter (studentarbeider) knyttes til et sett med standarder, best tas i bruk ved hjelp av en relasjonsdatabase. Til kompetansemapper trengs det kun et webområde, med autorisert adgang til et Internettområde. Generelt går Barrett i forsvar for digitale systemer, som støtter både det individuelle studentarbeid, administrasjon og personlig utfoldelse. Slik favner hun på en måte alle mappetyper i ett og samme system, slik jeg ser det.

Valgt vurderingsfilosofi og vurderingsprosedyrer virker grunnleggende inn på hva studenter deltar i, hvor hardt de jobber, hvor de gjennomfører sitt studiearbeid, og hva de tillater seg å bli interessert i under studiet. Slik representerer det også verdier og kvalitet knyttet til en utdanning, samtidig som den påvirker selve undervisningen, ikke bare hvordan den blir mottatt, men også hvordan den blir utført. Fra et studentperspektiv kan vurdering være forunderlig og motsetningsfullt, i den forstand at det ikke alltid er lett å forstå hvilke kriterier det vurderes etter. Fra et lærerperspektiv kan vurdering være vanskelig og arbeidsomt, i den forstand at studentarbeider må karaktersettes. Spørsmål som "Hva skal verdsettes?", "Hva skiller den gode og dårlige studentoppgaven?", "Hvilken standard skal man følge?", og "Hvordan skal standarden holdes over tid?" presser seg frem. Vurderingsprosessen beskrives av William og Black, som en syklus av ulike bevis.

The assessment process is characterised as a cycle involving elicitation of evidence, which when interpreted appropriately may lead to action, which in turn, can yield further evidence and so on. (William & Black 1996:537)

Vurderingsprosessen skjer i en syklus, der dokumentasjon bringes for dagen, tolkes hensiktsmessig og fører til handling. Handling kaster nytt lys over dokumentasjonen. Tre hovedtyper for vurdering beskrives av Wittek og Dysthe (2003) **normbasert**, **ipsativ** og **kriteriebasert**. I normbasert vurdering utgjør studentgruppen målestokken, og rangerer studentene i forhold til gruppens prestasjoner. Ipsativ vurdering har studenten selv som utgangspunkt og referanseramme, og både prosess og produkt er gjenstand for vurdering. Vurdering av **prosessen** er det sterkeste element. Ipsativ vurdering gir assosiasjoner til en formativ vurderingsform. Wittek og Dysthe (2003:131) antyder at Ipsativ vurdering er et naturlig valg, når det jobbes med formativ vurdering underveis i en læringsprosess, der det primære siktemål er å få frem det enkeltpersoner er i stand til å prestere. Arbeid med mapper legger på denne måten godt til rette for bruk av en ipsativ vurderingsform, både i formativ og summativ vurdering. Kriteriebasert vurdering innebærer at man på en eller annen måte identifiserer hva som karakteriserer en god prestasjon/produkt, for siden å vurdere studentprestasjonene opp mot disse. I vurderingsformen understreker Wittek og Dysthe det de kaller **gjennomsiktighet**. Med det menes at alle som involveres i vurderingen, har kjennskap til vurderingsrammeverket. Med slik kjennskap og tilgang kan også studenter aktivt være med og vurdere medstudenters og egne prestasjoner. Formativ vurdering, vurdering **for** læring, vil slik jeg ser det preges av ipsativ eller kriteriebasert (eller en kombinasjon av disse) vurdering, mens summativ vurdering ofte vil kombinere en norm- og kriteriebasert vurderingsform. I summativ vurdering vurderes studenter således både i forhold til prestasjoner underveis i studiet (kriterier), og i tilknytning til vedtatte faglige standarder, der de rangeres opp mot andre studenter (norm). Det frembringer nødvendigheten av å utarbeide kriterier, som kan brukes både til vurdering av prosess og produkt. Nøkkelen til vellykket vurdering er likevel å bygge fortrolighet, og få studenten til reelt å tro på at målet er innen rekkevidde (Sadler 2001).

2.2.1 Summativ vurdering – vurdering av læring

Uttrykkene formativ og summativ vurdering forklarer ikke vurderingsformen i seg selv, og en vurderingsform kan brukes både formativt og summativt. Uttrykkene knyttes **til hva** den informasjon som vurderingen gir, brukes til (William 1998).

Assessment **of** Learning (vurdering av læring) er ledet av spørsmålet "Hvor mye har en student lært på et bestemt tidspunkt?" (Sadler 2001). Hensikten med vurdering **av** læring, er å kontrollere kunnskaps- og kompetansenivå hos en student på et bestemt tidspunkt, og sjekke om nivået er godt nok i henhold til bestemte standarder/kriterier, slik at det kan gis et godkjenningsstempel/karakter (Lauvås).

Vurdering av læring ('summativ vurdering') = vurdering av studenter for å kontrollere at kunnskap og kompetanse er tilstrekkelig for godkjenning, evt også for å rangere dem som godkjennes. (Lauvås:1)

Summativ vurderingen fremstår på denne måten, som en sammenfattende statusrapportering om kompetanse/kvalitetsnivå hos en person, på et bestemt tidspunkt, etter avsluttet læringsforløp. I summativ vurdering vises gjerne de ”sterke” sider frem, mens de ”svake” holdes skjult. Det blir overlatt til den som skal vurdere, å finne de svake sidene, for så å sammenholde dem med de sterke. Vurdering **av** læring slik jeg tolker Barrett (2005) sertifiserer/kontrollerer hva en student har lært frem til et bestemt tidspunkt, og er tiltenkt folk som ikke direkte er involvert i den daglige læring og undervisning. Vurderingen vises i form av lettfordøyelig bevis som tall, karakterer eller grader, og presenteres i formelle rapporter. Studentens prestasjoner sammenlignes gjerne med andre studenters prestasjoner, eller med fastsatte og bestemte standarder. Vurderingsformen kan, men behøver ikke, involvere studenten.

Assessment of Learning
Checks what has been learned to date
Is designed for those not directly involved in daily learning and teaching
Is presented in a formal report
Usually gathers information into easily digestible numbers, scores and grades
Usually compares the student's learning with either other students or the 'standard' for a grade level
Does not need to involve the student

(Barrett 2005:17)

Elektroniske/digitale mapper brukt til vurdering av læring kjennetegnes ved at hensikten med mappen er foreskrevet av utdanningsinstitusjonen. Artefakter (arbeider) får sitt mandat fra institusjonen og bestemmes/vurderes ut fra generelle standarder. De kvantitative data samles inn til eksterne personer. Mappen utvikles vanligvis først mot slutten av terminen, er tidsbegrenset og skapes og struktureres med et fastbestemt antall arbeider, utarbeidet etter bestemte mål og standarder. Den brukes summativt til å bestemme/vurdere hva som er lært pr dags dato, og utvikling av mappeinnhold styres av ytre motivasjon. Studenten har liten kontroll med hvem (eksterne) som får tilgang til mappen. Vurderingsformen tar altså sikte på å være en avsluttende, kontrollerende vurdering og har godkjenning, sertifisering, eventuelt også en rangering av studentprestasjoner i forhold til andre studenter, i tankene. Studenten selv har liten innflytelse på vurderingen.

2.2.2 Formativ vurdering – vurdering for læring

Assessment **for** Learning (vurdering for læring) er ledet av spørsmålet ”Hvordan kan en vurdering brukes til å hjelpe en student å lære mer?” (Sadler, 2001). Formativ vurdering baserer seg på en annen logikk, der hensikten er å gi studenter vurdering og veiledning underveis i en læreprosess. Vurdering for læring er en pågående aktivitet som vises i form av tilbakemelding, der hensikten/målet er å støtte opp om studentens utviklings- og læreprosess, samt motivere for ny forståelse, kunnskapstilegnelse og forbedring.

Vurdering for læring ('formativ vurdering') = vurdering av studenter som hjelp i deres læringsprosess, veiledende vurdering. (Lauvås:1)

Begrepet formativ vurdering knyttes, i motsetning til summativ vurdering, til **prosess** og ikke til produkt, og har individuell forbedring som sitt primære mål. ARG (The Assessment Reform Group) definerer vurdering **for** læring, som prosessen de lærende og deres lærere går gjennom, i oppsøking og tolking av dokumentasjon, for å bestemme hvor den lærende befinner seg på **nåværende** tidspunkt i læreprosessen, **mot hvilke mål** den lærende bør bevege seg, og **hvordan** vedkommende best kommer seg dit.

Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

(The Assessment Reform Group 2002:2)

Vurdering for læring involverer altså **både** studenter og lærere, og karakteriseres ved å være sjekkpunkter for pågående og fremtidig læring. Vurderingsformen fungerer som kommunikasjonsredskap mellom de lærende og lærere, og bidrar til å skape refleksjon med hensyn til videre læreprosess. Forbedring er vurderingens hovedmål.

Assessment for Learning
Checks learning to decide what to do next
Is designed to assist teachers and students
Is used in conversation about learning
Usually detailed, specific and descriptive feedback in words (instead of numbers, scores and grades)
Usually focused on improvement, compared with the student's 'previous best' and progress toward a standard
Needs to involve the student -- the person most able to improve learning

(Barrett 2005:17)

ARG (2002) uttrykker at formativ vurdering bør være

1. En del av effektiv planlegging av læring og undervisning.
2. Ha fokus på læringsstrategier.
3. Sensitiv og konstruktiv, da all vurdering har en emosjonell side.
4. Legge vekt på å oppmuntre til læringsmotivasjon.
5. Fremme forpliktelse til å lære målsetninger og en delt forståelse av kriterier som ligger til grunn for vurdering.
6. Utvikle den lærendes evne til selvutvikling og fremme refleksjon og administrasjon av egen situasjon.
7. Anerkjenne hele spekteret av oppnåelige arbeider fra den lærende.
8. Gi konstruktiv veiledning til den lærende om hvordan man kan bli bedre.
9. Vurderingsformen anses som sentral i klasseromundervisning.
10. Er en nøkkelegenskap hos en lærer.

A. Davies (2000) definerer vurdering for læring som en kontinuerlig prosess, som krever dyp involvering av den lærende. Underveis i prosessen skal den lærende klargjøre resultater, kontrollere pågående læring, samle inn læringsbevis og presentere disse for andre.

Assessment for learning is ongoing, and requires deep involvement on the part of the learner in clarifying outcomes, monitoring on-going learning, collecting evidence and presenting evidence of learning to others."
(ibid)

Vurdering **for** læring karakteriseres ved at de lærende utvikler felles språk og forståelse, via selv vurdering og tilbakemelding fra medstudenter og lærere, gjennom hele læringsprosessen. Dokumentasjon samles, organiseres og fremvises som tegn på læring og prosess, og tilpasses som konsekvens av kontinuerlig tilbakemelding. Formativ vurdering bør sørge for et trygt læringsmiljø, som inviterer de lærende til å ta risiko, til å lære av sine feil, og gjøre det mulig å ha en klar fokusert målsetning, der tankefull læring støttes. For effektivt å støtte vurdering for læring, må en **elektronisk/digital** mappe ivareta en **kontinuerlig** læringsprosess (Barrett 2005). Mappen bør karakteriseres av at den utvikles over tid gjennom studiet, er tidsfleksibel, der mappestruktur og hensikt bestemmes av den lærende selv, eller forhandles frem med veileder/mentor/lærer. Arbeid med mappen bør stimulere til indre motivasjon og engasjere den lærende, som selv bør velger artefakter (studentprestasjoner) for fremvising av læringsprosess og utvikling. Artefaktene bør bedømmes både av den lærende selv, og brukes som grunnlag for tilbakemelding knyttet til fremtidig forbedring/læring. Mappen bør brukes formativt til å bestemme læringsbehov med utgangspunkt i læring fra nå og fremover. Den lærende velger selv hvem som skal ha tilgang til mappen, og brukes sjelden til å ta viktige beslutninger (karakterer). Også i st.meld. nr. 27 (kvalitetsreformen) understrekes formativ vurdering, som en viktig vurderingsform.

Departementet mener at jevnlige evalueringer som gir studentene hyppige tilbakemeldinger om utbyttet av læringsprosessen, skal inngå i studiene som en del av undervisningsarbeidet.[...] Vurderinger underveis i studiet kan organiseres på mange ulike måter, blant annet gjennom oppgaveinnleveringer, deleksamener, mappevurdering og annet.
(St.meld. nr. 27 2000)

Jevnlige evalueringer og hyppige tilbakemeldinger til studenter, knyttet til hvilket utbytte de har av læringsprosessen, bør altså inngå i undervisningsarbeidet. I organisering av jevnlige tilbakemeldinger, peker departementet som en av flere måter, på mappevurdering. Formativ vurdering er generelt læringsrettet og tar sikte på å være drivkraft i læringsprosessen. Forutsetninger er ærlighet og åpenhet, og at studenten har nytte av vurderingen. Underveis i den formative vurderingsprosess vil det være naturlig at studenten viser både svake og sterke sider. Skal den som vurderer kunne gjøre jobben sin, må alle sider ved studenten komme frem. En viktig forutsetning er at studenten forstår og aksepterer vurderingsformen, som en reell og læringsproduktiv vurderingsform, og ikke er i tvil, om at det virkelig er formativ vurdering som forekommer. Det er derfor viktig å gå nøye gjennom prosedyrer og andre detaljer, som eventuelt kan oppfattes av studenten som summativ vurdering.

... se nøye på prosedyrer og andre detaljer som gir studentene mistanke om skjult summativ vurdering, og vi skal sørge for at studentene etter hvert opplever den formative vurderingen som reell og læringsproduktiv. (Lauvås:4)

Elimineres mistanken om summativ vurdering, vil studentene etter hvert oppleve formativ vurdering som en reell og læringsproduktiv vurderingsform. Mange studenter ønsker nok underveis i læringsprosessen, å få en summativ vurdering på ett eller annet tidspunkt. Man ønsker å bli "ferdig med ting", og anser summativ vurdering for å være det avsluttende punkt. Men her er det grunn til å være på vakt, slik jeg tolker Lauvås.

Å oppfylle studenters ønske om summativ vurdering underveis, kan fort føre til en sammenblanding av summativ og formativ vurdering. Konsekvensen kan bli at summativ vurdering helt overtar og ødelegger for mulighetene til en sann formativ vurderingsprosess. Summativ vurdering er som gjøkungen, den dytter den andre ut av redet.

Høgere utdanning i Norge skal ikke lenger være et rent eksamensdrevet system. Studenter skal lære ved å jobbe aktivt med lærestoffet, og produsere og presentere på læringsarenaer, der de både får og gir tilbakemelding. Det som produseres underveis skal telle med i en endelig formell (summativ) vurdering. En forutsetning for at et slikt system skal virke, slik jeg tolker Lauvås, er et klart undervisnings- og læringsopplegg, der studentene ikke mistenker at det ligger skjult summativ vurdering bak.

For at et slikt system skal virke, må det lages et klart opplegg der det er liten grunn til å mistenke at det også vil foregå en eller annen skjult summativ vurdering.
(ibid:5)

Studenter skal kunne stole 100 % på at formativ vurdering er en veiledning som gis underveis i læringsprosessen, og sikrer læring frem mot en endelig avsluttende summativ vurdering. For å sikre formativ vurdering kan en **kvalifiseringssekvens** utarbeides. Her støtter Lauvås seg på en 6 trinnmodell beskrevet av Black & William (2001:13).

- (a) a design with formative/feedback opportunities built in;
 - (b) implementation in which students' responses are evoked;
 - (c) reception and interpretation of these responses by a teacher (or by peers)
 - (d) further teaching action based on the interpretation of the responses;
 - (e) reception and interpretation of these responses by the student;
 - (f) moving on to the next part of the design.
- (ibid)

Kvalifiseringssekvensen involverer både studenter og lærere. På trinn **(a)** utarbeides og forhandles rammene for den formative vurderingsdesign/form. I designen planlegges mulige måter å gi tilbakemelding/feedback. Rammene utarbeides av studenter og lærere i fellesskap, og det spesifiseres hvilken tilbakemelding/feedback, som er relevant når.

Trinn **(b)** og **(e)** i modellen favner aktivitet i studentgruppen. Trinn **(b)** iverksetter den formative vurderingsdesign og rammene utarbeidet under trinn (a). Studentene begynner nå å ta i bruk rammeverket og responsarbeidet /feedbacken tar form. Trinn **(e)** er fasen, der studenter tolker og oppfatter responser de får.

Tinn **(c)** og **(d)** rettes først og fremt mot lærergruppen og deres tilbakemelding. På trinn **(c)** fremkommer lærerens oppfatning og tolkning av respons fra studenter, og det gis en tilbakemelding/feedback på responsen. Tilbakemelding kan eventuelt også etableres som en peer to peer (medstudent) aktivitet, dersom det er hensiktsmessig. Trinn **d** fører til under-

visningshandling, som en konsekvens av responstolkning. Etter å ha tolket studentresponsene/feedbacken under trinn (c), blir det synlig hva studentene trenger mer undervisning i. Læreren iverksetter supplerende og klargjørende undervisning.

Trinn **(f)** avslutter den del man aktivt har jobbet med, og bringer aktørene videre til neste del i designen/vurderingsformen.

Et av de kritiske punktene i modellen slik jeg tolker Lauvås, er først og fremst hvor mye formativ vurdering, som skal komme fra lærere (jf. trinn c), og hvordan studenter på en kvalitativt god måte kan involveres i formativ vurdering. Overlates for mye av vurderingen til studentene alene, kan vurderingsformen fort miste sin legitimitet, og med det sin verdi. Det er heller ikke riktig at læreren sitter med hele tilbakemeldingsjobben. Det er derfor viktig å få en balansert fordeling i vurderingsprosessen. Det aller viktigste er at tilbakemelding blir gitt raskest mulig etter at vurderingsobjektet er produsert. Dernest må man avveie fordelingen av tilbakemelding fra henholdsvis student og lærer.

Hvorfor kaster så noen studenter ved studieslutt sine mapper i søplebøtta eller for den sak skyld brenner dem, og hvorfor utlover andre studenter dusør til dem som klarer å skaffe mappa deres til veie? Avhenger slike ekstreme mappeholdninger av om mappa har vært brukt i vurdering **av** eller **for** læring spør Barrett:

What are the variables that produce these extremes in attitudes toward ownership of portfolios?
(Barrett 2005:19)

Vil en mappe som brukes til vurdering **for** læring (formativ vurdering), bidra til styrking av eierskapsfølelse og verdifullhet, og vil en mappe som brukes til vurdering **av** læring (summativ vurdering), fremelske de motsatte holdninger? Dette vet man ennå for lite om. Det trengs fortsatt forskning på hvilke variabler, som er avgjørende for utvikling av holdninger til en mappe.

2.2.3 Formative vurderingsformer

Formativ vurdering omfatter ikke et bestemt sett av arbeidsmetoder, og mye av det læreren allerede gjør i undervisning har formative funksjoner. Lauvås målfører seks formative vurderingsformer, og mener at listen kan gjøres lengre, slik jeg tolker ham.

Vurderingsformen **gjensidig studentvurdering** bygger på prinsippet om feedback. Det er godt dokumentert at denne form for studentvurdering er en effektiv vurderingsform. Vurderingen bygger på **å få og gi respons** og gir dobbelt læringseffekt, i den betydning at studenten både lærer av å gi og få tilbakemelding. Tilbakemelding fra en medstudent vil ofte være mer klar og kontant enn den som kommer fra lærer. Når tekst skal vurderes av medstudenter, legges litt mer arbeid i teksten. Gjensidig studentvurdering legger til rette for et aktivt og produktivt læringsfellesskap, der studenter blir medaktører og medansvarlige i egen og medstudenters utvikling og læring. Å jobbe med samme fagstoff samtidig, øker i tillegg muligheten for kvalitativt gode faglige diskusjoner. Med dette utgangspunkt blir det

lett å rettferdiggjøre arbeidsformen for studentene. En forutsetning for at arbeidsformen skal lykkes er dog at studenter tar den seriøst. Om ikke, er faren for kollaps stor (Lauvås:8). For å kunne nyttiggjøre seg arbeidsmetoden trengs opplæring. Som oftest har studenter ingen erfaring med en arbeidsform, der tilbakemelding/feedback utgjør en essensiell del av lærings- og vurderingsprosessen. De må derfor lære seg til å levere produkter som er uferdige, for så å få tilbakemelding og utvikle sin fagkompetanse derfra.

En annen form for formativ vurdering er **selvvurdering**, der hensikten er å utvikle selvstendige studenter, som ikke følger undervisningen på en autoritetstro måte. Studenter må her ta ansvar for sin egen læringsprosess. Lauvås viser til David Boud (1995:18/19) når han karakteriserer arbeidsformen, og uttrykker at hensikten er å fremme god læringspraksis og gi kunnskap om/evne til å forstå hvordan læring skjer. Selvvurdering blir et substitutt for andre vurderingsformer, og bidrar til at studenten styrer sitt studiearbeid selv, gjennom innspill og inspirasjon fra litteratur og medstudenter. Studenten overvåker sitt arbeid og oppøver selvforståelse, selvstendighet og selvinnsett over tid. Erfaring innvirker på utvikling av innhold i mappen, der studenten dokumenterer egen læringsprosess. Læring og erfaring konsolideres og overføres til andre områder. Underveis i prosessen er det viktig med veiledning og støtte fra lærer. Problemet med selvvurdering slik jeg tolker Lauvås, er at studenter ofte ikke ser poenget med selvvurdering, og derfor saboterer formen.

Den formative vurderingsformen **synliggjøring av progresjon** og vurdering av egen læring og utvikling, kan være både motiverende og klargjørende og vel så lærerikt. I vurdering er fokus ofte rettet mot det man *ikke* får til og *ikke* kan, i stedet for hva man egentlig har fått til eller er blitt bedre i underveis. Det er relativt enkelt å gi studenter referansepunkter, slik at de med disse som utgangspunkt, kan sammenlikne sitt eget kompetansenivå fra ett tidspunkt til et annet.

Å gjøre **krav og kriterier kjent** gjennom formativ vurdering, kan også bidra til å forebygge usikkerhet og misnøye knyttet til summativ vurdering. Synliggjøring av krav og kriterier er noe som fremheves av Wittek og Dysthe (jf. 2.3).

2.2.4 Vurdering der vurderingsformer kombineres

Kan det så ha en hensikt å bruke en blandet vurderingsform der både formativ og summativ vurdering brukes til å støtte opp om hverandre? Ytterst sjeldent i følge Per Lauvås, som fremfører flere argumenter til støtte for sitt synspunkt.

Summativ vurdering er som gjøkungen, den presser den formative vurdering ut av redet. Kombineres vurderingsformene vil summativ vurdering overta og fortrenge formativ vurdering, selv om intensjonen var en likestilling mellom de to. En vellykket kombinasjon forekommer derfor ytterst sjeldent. En slik tankegang støttes av R. Sadler (2001). Studenter flest er så vant med summativ vurdering, at erfaringen de har med vurderingsformen, lett fører til at formativ vurdering oppleves som ureell og lite læringsproduktiv. Det er derfor uhyre viktig at det ikke foreligger noen som helst tvil om hvilken vurderingsform som

brukes. Når formativ vurdering er valgt, må studenter kunne stole på at det kun er det som forekommer. Blandes de to vil studenter lett komme i den situasjon, at de lurer på om det likevel ligger skjult summativ vurdering bak den formative vurdering. Det er derfor viktig at studenter får det overtydelig klart for seg, hvilken vurderingsform som er valgt, når.

Studenter skal ha det overtydelig klart for seg når det er formativ og når det er summativ vurdering som foregår.
(Lauvås:5)

Studenter må kunne stole på at de kan satse på å tilegne seg maksimalt læringsutbytte av den formative vurdering, inklusiv å blottlegge sine eventuelt manglende fagkunnskaper, uten å frykte at det de arbeider med, blir gjenstand for summativ vurdering. Åpnes det for summativ vurdering vil det umiddelbart bidra til å skape usikkerhet. For å unngå en assosiasjon til summativ vurdering, bør studieoppgavene likne minst mulig på standard skolepregete oppgaver, og bestå av korte poengterte ting, slik at det er lett å gi tilbakemelding uten å bruke for lang tid. Det studenten skriver skal i første omgang kun inngå i formativ vurdering, for eventuelt siden å inngå i en summativ vurdering, dersom det er hensiktsmessig.

D. William, forgrunnsfigur innen formativ vurdering, gir uttrykk for at man i pedagogisk sammenheng burde gå over til mer formativ vurdering, fordi det viktigste med undervisning er læring og ikke testing. Læringsaspektet ved formativ vurdering er det viktigste, ikke sertifisering som skolen egentlig bygger sin eksistens på.

Uansett hvordan man beskriver formativ vurdering som begrep, er ikke vurderingsformen av nyere dato (William & Black 1996). Lærere har til enhver tid tatt med seg erfaring fra tidligere undervisning, og på denne måten tilpasset sin undervisning i lys av suksessfaktorer fra tidligere perioder. Forklaring/veiledning kan umiddelbart tilpasses den enkelte student/situasjon, slik at den får mening. Hovedelement i vurderingsprosessen er **veiledning**, der målet er å fylle et hull, hos den studerende. For at formativ vurdering skal være **gyldig**, må den vurderes ut fra hvilken konsekvens den får, så vel som hvilken betydning den har.

Å vurdere læring støtter ulike hensikter. En hensikt kan være å kartlegge en hel nasjons standarder (jf, nasjonale prøver), der kartlegging brukes til å vise utvikling over tid, en annen å sammenlikne med andre land og deres standarder. Hensikten kan også være å skaffe informasjon for å holde aktuelle instanser (lærere, etater, myndigheter) ansvarlige. Mange av disse hensiktene kan lett føre til prioritering av summativ vurdering av læring. For å komme seg ut av en ond spiral foreslår William derfor å **kombinere** summativ og formativ vurdering i et læringsløp.

In place of this vicious spiral, I propose that developing a system that integrates summativ and formative assessment will improve both the quality of learning and the quality of the assessment.
(William 2000a)

Ved å utvikle et vurderingssystem som integrerer både formativ og summativ vurdering, vil det åpnes for muligheten til å vurdere selve læreprosessen. Det vil øke kvaliteten både på læring og vurdering.

Når læring skal vurderes er det i hovedsak to nøkkelspørsmål knyttet til reliabilitet og validitet. **Reliabilitet** defineres som et mål på hvor **pålitelig** eller hvor stabil noe måles. I praksis vil det si at en prøve som har god reliabilitet, måler det samme hver gang den brukes, uavhengig av hvem som måler, eller hvem som blir målt. William (2000b) definerer reliabiliteten på en test til å være et forhold mellom **signal** og **støy**.

The reliability of a test is defined as a kind of 'signal to noise' ratio. The mark of any individual on a particular test is assumed to be made up of a signal (the true score) and some noise (the error), and of course, we want the ratio of signal to noise to be as large as possible.
(ibid)

Vurdering en student får på en test, gis ut fra signal og støy studenten sender (det som er riktig, og det som er feil). Man ønsker da selvfølgelig at signalet studenten sender, skal være så godt som overhodet mulig, og at det sendes minst mulig støy. Et kraftig og klart signal gir god vurdering. For å forbedre signalet har man to innfallsvinkler, henholdsvis å redusere støyen (reducere antall feil), eller å forbedre signalene (jobbe med kunnskapsoppbygging, slik at antall riktig svar økes). Generelt gir William uttrykk for at tester som brukes i vurdering har liten læringsverdi, og ikke kan anses å være reliable, da de kun tester kunnskap som er lett å teste og i et begrenset omfang.

Validitet forteller om prøven/vurderingen virkelig måler/vurderer det den tar sikte på å måle/vurdere. En bør f.eks. kunne være sikker på at en prøve i leseferdighet, virkelig måler leseferdighet. Men i stedet for å fokusere på "Hvor nøyaktig man måler", burde man heller være opptatt av å spesifisere "**Hva er det nøyaktig man måler**", og da se på målingens betydning, slik jeg tolker William. Lenge har det vært vanlig å definere en vurdering å være valid, når den vurderer det den påberoper seg å vurdere.

An assessment is valid if it assesses what it purports to assess.
(William 2000b:112)

William antyder i midlertidig at en slik definisjon er utilfredsstillende, **da validitet ikke kan være en enhetlig vurderingsegenskap**. Validitet (vurderingsvaliditet) endres i forhold til hvem som testes/vurderes, og hvem som tester/vurderer. En vurdering påberoper seg derfor ingen ting, den vurderer kun det den vurderer, og forteller ingenting ut over det resultat som fremkommer av vurderingen. De siste 30 år har det blitt mer akseptert at validitet ikke er en egenskap ved en test/vurdering, men en egenskap av konklusjonene, som trekkes på basis av testresultatene.

2.2.5 Vurdering av digitale mapper?

Studenter på alle nivåer ser ofte på vurdering som noe andre gjør, der grunnlaget for vurdering er arbeidet i klasserommet. Generelt har studenter liten kjennskap til hva som

ligger til grunn for vurdering av deres arbeid (Office of Research 1993). Men studenter vil tjene på å bli oppmerksomme på prosess og strategi tilknyttet ulike aktiviteter som skriving, problemløsning, utforskning av et emne, analyse av informasjon og beskrivelse av egne observasjoner. Uten nærmere instruksjoner i hvordan man får til effektive og gode prestasjoner, vil studentene i liten grad lære seg strategiene og prosessen, som er grunnlaget for gode prestasjoner.

Når bruk av mapper beveger seg fra å være papirbasert til å være elektronisk og webbasert, må det diskuteres hvordan det digitale mediet støtter og påvirker formålet/hensikten med mappa. En mappe som etterlikner en papirbasert, og tilfeldig etableres i en elektronisk container, er veldig forskjellig fra mapper som brukes seriøst i læring og vurdering. Men for effektivt å kunne bruke mapper i vurdering, må læringsinstitusjoner ha opparbeidet en kultur, for hvilken **dokumentasjon** som skal telle/ligge til grunn for og ved vurdering.

To effectively use portfolios for assessment, a learning organization needs to establish a culture of evidence.
(Barrett 2005:7)

Vurderingsdokumentasjonen fremkommer ved å summere artefakter (studentarbeid), refleksjoner (rasjonale) og tilbakemelding på artefaktene (validering).

Evidence = Artefacts + Reflection (Rationale) + Validation (Feedback)
(Barrett 2003)

Bevisene i den elektroniske mappen består ikke kun av artefaktene (a) studenten har der, men også av studentens rasjonelle forklaringer (r) for hvorfor nettopp disse artefaktene er med, og hvordan de beviselig fyller de oppsatte målene. Det er dog ikke nok at en student hevder at deres artefakter dokumenterer faglig kunnskap. Vurderingsdokumentasjon trenger validering fra en øvet vurderingsperson. Til det må det brukes velutviklet forskrifter med identifiserbare spesifikke kriterier, slik jeg tolker Barrett.

2.2.5.1 Refleksjon i læring

Digitale mapper som virkelig støtter vurdering **for** læring (formativ vurdering), og som følger ARGs (2002) 10 vurderingsprinsipper for læring (jf. 2.2.2), må understøttes av **strategier**, som hjelper den lærende til å fortelle om egen læring. En strategi er å **sette seg egne mål** (Davies 2000, Barrett 2005).

Research is indicating that closing in on a goal, triggers a part of the brain linked to motivation (e.g. Csikszentmihalyi, 1990; Pert, 1997; Pinker, 1997). Setting goals is a powerful way to focus students' learning.
(Barrett 2005:19)

Forskning viser at å sette seg mål er en kraftfull måte å rette fokus mot det studenter skal lære. Målsetting synes å virke som trigger, for den del av hjernen som er knyttet til motivasjon. Dyp læring (Roger Schank 1991, Zull og Moon) forekommer når størsteparten av hjernen engasjeres i læringsprosessen. Mange (Pearl and Leon Paulson, John Dewey,

Jürgen Habermas, David Kolb, Donald Schön, James Zull) har derfor diskutert mulighetene digitale mapper gir til å skape mening og dyp læring. Allerede tidlig på 90-tallet skapte Pearl og Paulson (1991) metaforen "Mapper som meningsskapende verksted", der studenter kunne skape mening av sin akkumulerte viten. De pekte på at mapper forteller en historie.

The metaphor of "portfolio as story" emphasizes the richness of this tool to support reflection as an essential component of learning
(ibid)

Metaforen "Mappe som historieforteller", åpner for gode muligheter til å støtte **refleksjon**, en mental prosess, som brukers til å oppfylle en hensikt eller å forutse noe. Refleksjon utgjør en vesentlig komponent i læringsprosessen (Moon 1999).

Reflection is a form of mental processing – like a form of thinking – that we use to fulfil a purpose or to achieve some anticipated outcome. It is applied to relatively complicated or unstructured ideas for which there is not an obvious solution and is largely based on the further processing of knowledge and understanding and possibly emotions that we already possess
(ibid: 20)

Refleksjon er den mentale tankeprosess, som et menneske gjennomgår, for å utøve en handling, og som leder til fullførelse av et mål eller oppnåelse av forutsette resultater. Moon uttrykker at overflatelæring ikke involverer refleksjon og understreker at betingelsene for at refleksjon skal skje avhenger av tid og sted, gode hjelpere, støttende pensum og et emosjonelt støttende miljø. Refleksjon trigges av dårlig strukturerte, uryddige eller virkelighetsnære situasjoner, og av oppgaver som stiller de "riktige" spørsmålene og ikke har entydige fasitsvar. Også oppgaver som byr på utfordringer, integrerer ny læring i tidligere læring, krever tankegjennomgang og evaluering, frembringer refleksjon. Mapper representerer på denne måten et kraftfullt miljø, der studenter kan samle og organisere sine artefakter, som et resultat av refleksjon over tid. Når studenter forteller historier om sine erfaringer, hjelper det dem å integrere disse i langtidshukommelsen.



2.2.5.2 Muligheter og utfordringer ved digitalisering av mapper

I en forskningsstudie gjort av Linnakylä (2000), fant man at ved lesing av en digital mappe skiftet leseren fokus hver gang. Ved første gangs lesing var forsiden avgjørende for inntrykket av mappen. Ved annen gangs lesing var fokus mer rettet mot mappens struktur/oppbygning og innhold, og ved tredje gangs lesing var oppmerksomheten vendt mot å studere hvordan mappen reflekterte kritisk tenking. Undersøkelsen gir grunn til å være på vakt i arbeid med digitale mapper. Visuelt fremtredende virkemidler blir lett svært dominerende. Å lese digitale mapper krever derfor en ny leserkompetanse (Linnakyläs 2002, Blair & Takayoshi 1994). Å finne mening og sammenheng i en ikke lineær tekst, som inneholder bilder, lenker, tekst og digitale effekter, krever en annen leserkompetanse, enn det som er påkrevd ved lesing av tradisjonelle tekster/mapper.

Blir det så virkelig noen kvalitativ endring i læringsprosessen, når digitale mapper tas i bruk, ut over det at de rent praktisk hjelper til med å effektivisere allerede eksisterende læringsaktivitet?

... om digitalisering av mappene fører til kvalitative endringer i læringsprosessene utover rent praktiske og effektiviserende forbedringer av allerede eksisterende læringsaktivitet?
(Madsen 2003:174)

Og hva er det som definerer kvalitative endringer i en digital læringsprosess? Her finnes lite praksisdokumentasjon, men interessante innspill har kommet fra Blair & Takayoshi (1997), som peker på utfordringer knyttet til både lesing og vurdering av elektroniske mapper. I en digital mappekontekst, der det gjøres bruk av tekst, bilder, lyd, grafikk, video, hypertekst og interaktivitet, får studenten/mappeeieren mulighet til å ta kontroll over mappen, på en måte som ikke har vært mulig i en vanlig lineær mappe. I den digitale mappen kan studenten (mappeeieren) selv styre og lede leseren dit hun/han ønsker. Blair & Takayoshi snakker om at studenten gjennom digitale mapper, får mulighet til å ta kontroll over siden "taking control of the page". Begrepet har de hentet fra Sullivan (1991). Etter eget behov og ønske kan studenten lede leseren inn i et digitalt landskap, formet og tilrettelagt etter mappeeierens eget ønske. Skillet mellom prosess og produkt i en lineær mappe er vanligvis synlig, men i en digital mappe vil spesielt bruk av hypertekst bidra til å viske dette ut. Via hypertekst ilegges leseren en aktiv rolle, der leserens egen fortolkning av det digitale innhold blir et viktig aspekt. I motsetning til lineær papirbaserte mapper, der leserens forståelse utvikles ved å følge 1. utkast, til 2. utkast og så videre, kan mappeeieren (studenten) i en digital sammenheng nå selv bestemme presentasjonsstrukturen, og slik styre leserens vei til forståelse av mappens innhold. Her er det fullt mulig å starte den digitale mappepresentasjon med et endelig ferdig utkast, for derfra via hyperlenker å lede leseren til andre relaterte mappeelementer.

Lineær forståelse av mappeinnhold – Prosessuell tilnærming	Ikke lineær forståelse av mappeinnhold – Hypertekstuell tilnærming
<p>Mappens innhold bestemmes av mappeeier. Forståelse av mappens innhold etableres hos leseren ved lineært å følge utvikling av mappens elementer fra start til slutt, via 1. utkast, 2. utkast og frem til endelig resultat. Forståelse av innhold utvikles lineært.</p> 	<p>Mappens innhold bestemmes av mappeeier. Forståelse etableres hos leseren, ved å følge en digital hyperstruktur, bestemt av mappeeier. Leseren navigerer i et hyperaktivt spindelnev og kan bevege seg fritt, alt etter hva som passer.</p> 

Figur 4: Lineær og hypertekstuell forståelse

Leserens forståelse i en lineær mappekontekst, tar utgangspunkt i 1. element og følger utviklingen derfra. Forståelse avhenger altså først av innblikk i prosesselementer, for deretter å forstå det ferdige produkt. Forut for forståelse av det ferdige produktet, trengs ingen prosessforståelse. Skillet mellom prosess og produkt viskes ut. Når forståelse tar utgangspunkt i et ferdig element, via en hypertekstuell tilnærming, blir konsekvensen at

forholdet mellom leser og skriver endret. En slik endring får innvirkning på vurdering av digitale mapper (Blair & Takayoshi 1997). Forståelsen av begrepene **tekst**, **skrivning**, **lesing**, **leser** og **skriver**, må derfor endres i tilknytning til digitale tekster/mapper. Foruten en endring i forståelse av begrepene nevnt over, peker Blair & Takayoshi på den hypertextuelle mappevurderingens validitet, som et problemområde. Avhenger mappeforståelsen av valgene leseren gjør, vil det knyttes usikkerhet til hvilket grunnlag vurderingen er gjort ut fra. Er vurderingen et resultat av studentens faglige forståelse, eller knyttes den til sensors evne til å oppdage og skape mening i studentens hypertextlige presentasjon? Hva er gyldig (valid) ved vurderingen da? Etter min oppfatning blir begge aspekter viktige. Vurdering bør baseres på studentens evne til å vise faglig forståelse. Viser denne i en hypertextuell sammenheng, er det viktig at sensor innehar evnen til å oppdage og skape mening i konteksten. Å vurdere en hypertextuell presentasjon, skal dog ikke avhenge av sensors evne til å kunne oppdage og skape mening i den. Det er derfor viktig at studenten har en reflektert holdning, når den hypertextuelle presentasjonen utformes. Det skal ikke være slik at sensor skal lete etter presentasjonens mening og innhold. En god hypertextuell utforming ivaretar leseren på en god måte. I arbeid med digitale mapper er det imidlertid viktig å ta denne problemstillingen på alvor.

2.3 Kriterier i vurdering

R. Sadler (1996) viser til fire velkjente måter å karaktersette/ vurdere studentoppgaver, å telle antall riktige svar og gi poengsum ut fra det (Counting forward), å starte med maksimal poengsum og trekker underveis for feil eller manglende faglig innsikt i oppgaven (Counting backwards), å vurdere kvalitet av et studentarbeid og gi en tallverdi ut fra bedømmelsen (Direct assignment) eller ulike hybride vurderingsmåter.

Generelt er nøkkelelement i vurdering selve **vurderingsrammeverket**. De to mest alminnelige er normbasert eller kriteriebasert vurdering (jf. 2.2). Ved kriteriebasert vurdering skiller flere på begrepene kriterium og standard. Slik jeg tolker Sadler (1996) definerer han **kriterium** som en **egenskap** eller en karakteristikk, hvorved kvalitet av noe bedømmes. **Standard** defineres som et bestemt og **avgjørende nivå**, som det går an å oppnå. Sistnevnte knyttes for det meste til summativ vurdering.

By CRITERION is meant a property or characteristic by which the quality of something may be judged. By a STANDARD is meant a definite level of achievement aspired to or attained.
(Sadler 1996)

også Wittek og Dysthe (2003:134) diskuterer begrepene med rot i Dunn, Parry & Morgan. Som Sadler skiller også de på kriterier og standarder og uttrykker at **kriterium** er *egenskaper eller trekk, som brukes til å gjenkjenne kvalitet*, og **standard** knyttes til å *bedømme grad av kvalitet*. I internasjonal litteratur skiller det ofte **ikke** på disse begrepene, der oppfattes de rett og slett synonymt. Men både Sadler, Wittek og Dysthe, understreker altså at de ikke oppfatter begrepene til å representere det samme.

2.3.1 Kriterier

I vurderingsarbeid, der vurderingskriterier er styrende element, er **synliggjøring** og **forhandling** av kriterier viktig av flere grunner (Wittek & Dysthe, 2003). For det første skal studenter på ett eller annet tidspunkt til slutt i et studium ha en avsluttende vurdering. Det er da nødvendig for alle parter på forhånd å vite hvilke kriterier sluttvurderingen tar utgangspunkt i. For det andre har det i dag blitt mer og mer vanlig at studenter gir respons til hverandre underveis i studiet. Skal studenter kunne fylle veilederrollen, må de vite hvilke kriterier de skal respondere/vurdere i forhold til. I arbeid med mapper står ofte egenvurdering sentralt. Det er derfor essensielt at studenten vet hvilke kriterier de skal vurderes seg selv etter. Forståelse av vurderingskriterier for et fag/profesjon, vil som oftest være knyttet til kompetansen studenten skal utvikle gjennom studiet. Kjennskap til vurderingskriterier vil på denne måten også styrke fagkompetansen.

Wittek og Dysthe (2003) skiller på **generelle** og **spesifikke** kvalitets- og vurderingskriterier. Spesifikke kriterier forhandles gjerne frem og tilpasses en spesiell fagkontekst/læringssituasjon. Behovet for generelle kriterier bortfaller da. Selve arbeidet med utforming av vurderingskriterier er en komplisert prosess, som alltid må henge nært sammen med faglige og begrepsmessige målsetninger i et studium.

Det som ofte skiller mappevurdering fra andre vurderingsformer, er en mulig sammenblanding av summativ og formativ vurderingsform. Å utforme summative vurderingskriterier, rettet mot ferdig produkt, går for de fleste greit. Dilemmaet oppstår slik jeg tolker Wittek og Dysthe, når det er behov for å definere i hvilken grad og hvordan formativ vurdering underveis i læreprosessen, skal **integreres** i det summative vurderingsgrunnlag. For mange synes det å være vanskelig å finne/utarbeide gjennomførbare kriterier, som gir et godt grunnlag for vurdering av prosess. Nettopp et slikt utgangspunkt kan lett føre til vegring mot prosessvurdering. Et grundig forskningsstudie gjennomført i Norden, initiert av Skolverket i Sverige og ledet av professor Lars Lindstrøm (1999), taler imidlertid mot at det er vanskelig å vurdere prosess. Studien konkluderte dog også med at en nødvendig forutsetning for prosessvurdering, var at studenter ble vant med å skrive regelmessig logg underveis i prosessen. Loggene ble da en del av vurderingsgrunnlaget.

2.3.2 Implisitte og eksplisitte kriterier

Jeg kjenner igjen en god prestasjon når jeg ser den. (Dysthe & Breistein 1999)
I utsagnet over kjenner sikkert mange lærere seg igjen. De har gjennom flere års faglig erfaring, tilegnet seg kompetanse på hva som preger og karakteriserer en god studentoppgave. Kriteriene som ligger til grunn for vurdering av besvarelser, er som oftest ikke nedskrevet noe sted, men er en del av lærerens forståelse for hva som karakteriserer høy kvalitet innen vedkommendes fagfelt. Med et slikt utgangspunkt, blir det da gjerne avvist å utforme eksplisitt nedskrevne og synlige kriterier. Studenter derimot er slett ikke på langt nær så sikre. Tvert imot er de ofte usikre på hvilke vurderingskriterier som gjelder, det være

seg en vurderingsform tuftet på tradisjonell eksamen, eller på mapper. Mange gir uttrykk for at de gjerne så, at det fantes klare eksplisitt nedskrevne kriterier for vurdering. Undersøkelser antyder dog også at studenter generelt er uvitende om hvilke vurderingskriterier som gjelder, uansett om kriteriene er synlig nedskrevet eller ikke (Wittek, 2004). I Witteks forskningsstudie gjennomført på Universitetet i Oslo fra 2001-2003, ga de fleste studenter uttrykk for at de **ikke** visste hvilke kriterier de ble vurdert etter.

When the students were asked about the assessment criteria for their portfolios, most informants exclaimed: "I have no idea!" When asked what they thought the assessors would give credits for, the students mainly expressed what they thought should count towards the grade. (Dysthe et al. 2005:4)

Når studenter ble spurt om hva de trodde sensor ville premiere i deres mappe, uttrykte de det de **selv trodde** skulle telle for å få graden. Studenter ga også uttrykk for at de aktivt ønsket å planlegge sin egen vurderingsprosess, og delta i utvikling av vurderingskriterier. Så man vurdering i lys av den nasjonale rammeplan, var det naturlig. Wittek oppsummerte sin studie med:

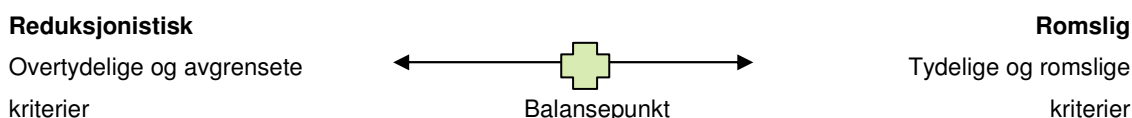
1. Studenter har rettmessig behov, for å vite hva som teller som kvalitet i deres arbeid.
2. Publisering av et sett med kriterier er ikke tilstrekkelig, om de ikke blir brukt aktivt som en ressurs for å involvere studentene. Aktiv bruk og fokus på kriterier er nødvendig. Læreren står i fokus for å få til det.
3. Kriterier må være lagt inn i praksis.
4. Veldig få studenter er i stand til på egen hånd å se sammenhengen mellom kriterier og praksis.

Ønsket om å ha uttrykte eksplisitt nedskrevne kriterier er dog sprikende. Mens noen studenter ønsker klart formulerte nedskrevne kriterier, slik som ved Universitetet i Oslo, ønsker andre det ikke. I en forskningsstudie ledet av K. S. Engelsen (Dysthe et al., 2005) ved Høgskolen Stord/Haugesund, ble det dokumentert at studenter i en læringskultur, der det var naturlig med forhandlinger mellom student og lærer, mente at eksplisitt nedskrevne kriterier var overflødige. Tett samarbeid og dialog mellom student og lærer, gjorde at det var unødvendig med nedskrevne og synlige (eksplisitte) vurderingskriterier. De ble forstått og utviklet i den gode samtalen/dialogen mellom student, medstudent og lærer, ved hjelp av ulike medier som LMS, web, msn, e-post og diskusjoner i klassen. På denne måten ble kriterier en integrert del av hele læringssituasjonen, uten å være spesifikt diskutert som eget tema. Studentene utviklet, litt etter litt gjennom dialog, en implisitt (egen) forståelse av hva som karakteriserte gode arbeider innen faget. Veien for læringsprosessen lå i utgangspunktet åpen og ble til gjennom kaos, og besto i stor grad av modning over tid, der fokus på karakterer var minimal. Oppmerksomheten var heller rettet mot ulike oppgaver, som skulle løses underveis, i samarbeid med medstudenter og lærere. Generelt var læringsmiljøet preget av åpenhet, der det var lov å ta opp spørsmål til diskusjon. For å se sammenhenger og skape egen læringsprosess, som ikke alltid på forhånd var staket ut, hadde noen studenter behov for å ta kontroll og utvikle egne kriterier. Det involverte studentene på en ny måte.

Kapittel 2.3.2 ble innledet med sitatet ”*Jeg kjenner igjen en god prestasjon når jeg ser den*”. Uttalelsen antyder at læreren helt intuitivt og i kraft av sin faglige kompetanse, vet hva som preger en god studentbesvarelse, og derfor ikke har behov for klart definerte eksplisitt nedskrevne kriterier til vurdering av en oppgave. Mange studenter derimot ønsker oppgavevurdering, som følger klare nedskrevne vurderingskriterier, og tror de da vil få mer presis og rettferdig vurdering. Dette fremheves også av Wittek og Dysthe (2003), som antyder at bruk av klart definerte kriterier øker **reliabiliteten** (pålitelighet) i kriteriebasert vurdering. Med utgangspunkt i felles og synlig nedskrevne kriterier, vil vurderingsforståelsen øke mellom dem som vurderer oppgaven. Slik jeg tolker Wittek og Dysthe er det generell enighet om at nøkkelen for å oppnå reliabel (pålitelig) og valid (gyldig) vurdering basert på kriterier, er at det på forhånd etableres klare og egnede vurderingskriterier knyttet til prestasjoner.

Det man kan enes om innenfor dette problemområdet er at nøkkelen til å oppnå en reliabel og valid vurdering basert på kriterier, ligger i å etablere klare og egnede kriterier for prestasjoner. (Wittek & Dysthe 2003:141)

Også Baume (2002) antyder at en vurderings **validitet** øker ved bruk av kriterier. Her er det dog ulike oppfatninger. Baume argumenterer generelt sterkt for kriteriebasert vurdering innen høgere utdanning, og gir uttrykk for at det **bidrar til å gi klare anvisninger** for hvilken kvalitet som belønnes og verdsettes. I tillegg **klarerer det retningslinjene** for undervisning, og hjelper til med å **sette standarder** som tydeliggjør for studenter, hva de må kunne for å oppnå en god karakter. Kriteriebasert vurdering bidrar også til at **vurdering oppfattes som meningsfull** og sammenhengende, med det som læres gjennom studiet. I tillegg **legger det rammer for tilbakemelding** på arbeid underveis i prosessen. Sist men ikke minst **etablerer det en felles referanseramme** for dem som involveres i vurderingen. Wittek og Dysthe peker selv på at det ikke alltid er så lett å etablere klart uttrykte vurderingskriterier, og antyder at man i kriterienes etableringsfase, må være på vakt for ikke å banalisere kompleks kunnskap. De uttrykker at alle parter må innse at selv om det utarbeides spesifikt nedskrevne vurderingskriterier, må det alltid være rom for faglig tolke. Ofte er det basert på innforstått fagkunnskap, hvilket mange lærere vegrer seg mot å uttrykke i nedskrevne kriterier. Utfordringen ligger i å finne balansen mellom **tydelige og romslige** kriterier på den ene side, og **overtydelige og avgrensede** kriterier på den andre. Det ideelle er å unngå ytterligheter, og finne balanserte vurderingskriterier som ivaretar både et romslig og et tydelig aspekt. Altså at de nedskrevne kriteriene befinner seg rundt et balansepunkt, slik modellen under viser.



Figur 5: Balanse i utforming av kriterier

Det må være rom for menneskelige og uforutsigbare kvaliteter. Tenking, kompetansenivå, kreativitet må få lov å vise seg i nyskapende kunnskapsutvikling. Rigide kriterier og standarder vil kunne dempe dette. Sentrale vurderingskriterier, må synliggjøres og ikke forveksles med en sjekkliste, som lett fører til instrumentell utførelse av vurderingsarbeidet. I tillegg er det viktig å få frem en felles oppfatning av hva kriteriene betyr. Å dokumentere kriterier som for eksempel forståelse, innsikt og kritisk vurdering vil alltid innebære tolkning. Om vurderingskriterier kan tolkes, er det viktig på forhånd å drøfte hva man tillegger dem, slik at alle forstår dem innen samme referanseramme.

Arbeidet med å skape felles forståelse av kriterier er en sammensatt, komplisert og viktig prosess, og gir stort potensial for forbedring av både læring og vurdering, slik jeg tolker Wittek og Dysthe. En naturlig arena for drøfting og bruk av vurderingskriterier er undervisning. Selve arbeidet med vurderingskriterier blir således en viktig del av både lærings- og utviklingsprosessen.

2.4 Et internasjonalt blikk på bruk av mapper

2.4.1 Internasjonal forskning

Mapper (portfolios) brukes i dag i utstrakt grad i USA. Carney (2004) viser til Salzman, Denner et al. (2002), når hun påstår at nesten 90 % av alle skoler/høgere utdanningsinstitusjoner i USA, bruker mapper til å ta beslutninger knyttet til en kandidat/student.

Nearly 90 % of schools, colleges, and departments of education (SCDE's) use portfolios to make decisions about candidates.
(Carney 2004:1)

Mange gir sterk teoretisk støtte til bruk av mapper som pedagogisk redskap, men ting tyder på at det fortsatt mangler empirisk bevis og dokumentasjon på effekten ved bruk av mappeverktøyet. Allerede i 1994 registrerte Herman & Winters de antatte fordelene med mapper, men viste også allerede da til manglende forskningsdokumentasjon på effekt ved bruk. Gjennom 89 forskningsartikler tilknyttet mappevurdering fant de at kun 7 enten rapporterte om tekniske data eller tilførte nye forskningsmetoder. I de fleste artiklene var derimot fokus enten rettet mot å forklare rasjonale for å ta i bruk mappevurdering, eller presentere ideer og modeller for hvordan implementeringen kunne foregå. Relativt lite oppmerksomhet var rettet mot områdene a) teknisk kvalitet, b) seriøse indikatorer av betydning eller, c) en for rigid (streng) uttesting av antakelser. Også T. G. Madsen (2003) uttrykker, at digitale mapper og mappevurdering internasjonalt ofte trekkes frem, som en løsning på utdanningens problemer og utfordringer.

Innføring av mapper og mappevurdering blir ofte fremstilt som løsningen på mange av utdanningssystemets utfordringer
(Madsen 2003:159)

Om mapper og mappevurdering fremstilles som løsning, blir det avgjørende å forstå hva det nye er, i forhold til det som var. Til det trengs det forskning. Er forskning på dette område

så blitt noe bedre siden 1994? Det ser ikke slik ut (Lyon 1998, Zeichner & Wray 2000). Til tross for mappers popularitet har det vært veldig få seriøse og systematiske forskningsstudier, rettet mot å undersøke mappers egenskaper, og hvilken konsekvens de har som pedagogisk redskap. Det gjelder både i vurdering og læring.

Despite the current popularity of teaching portfolios, there have been very few systematic studies of the nature and consequences of their use for either assessment or development purposes. (Zeichner & Wray 2000:615)

Carney stiller i den forbindelse spørsmål ved om mapper er gode nok som redskap for læring og vurdering, sånn som mappetilhengere hevder. Og om de er, under hvilke betingelser? Når mapper blir mer og mer viktig i vurdering og utvikling av en students kunnskap, betyr en rivende teknologisk utvikling, at mapper som pedagogisk redskap også må gjennomgå radikale endringer. Spørsmålet er da om man vet nok om hvordan Internett- og multimedieteknologi vil påvirke mappeprodukter og læringsprosesser.

Utvikling av programvare og databaseverktøy har i tur og orden bidratt til at mapper også brukes som verktøy, for gjennomføring av digital vurdering i stor skala. Dersom mapper skal brukes til viktige kartlegging og plassering, må man være veldig oppmerksom på at verktøyene bygger opp om den tekniske standarden, som gjelder for enhver god vurderingsprosedyre. Og selv om en kan hevde at mapper oppfyller disse kravene, er det da klokt å bruke mappeverktøyet til å ta viktige beslutninger, eller vil det kunne ødelegge for bruk av mapper som redskap i en læringsprosess? Finnes det empiriske bevis for at mapper er reliable (pålitelige), for på den måte å kunne tilføre dem som skal vurdere, et grunnlag for valid (gyldig) tolkning av en students prestasjoner? Fortsatt er det mange ubesvarte spørsmål knyttet til bruk av mapper, ikke minst i digital form, men få svar å hente fra empirisk forskning, slik jeg tolker Carney. Det finnes mange rapporter som beskriver implementasjon, og hvordan ulike mappeprogram er bygget opp. Det er bra og nyttig når et system skal designes eller redesignes, men det i seg selv er ikke i stand til å gjøre mapper til effektive enheter for vurdering eller læring. Om en mappe skal legge til rette for læring må det være sterke bevis på at mappa virkelig har fremmet innehaverens viten/kunnskap (Moss 1998). Likeså om hensikten med en mappe er vurdering. Er vurdering hovedmålet, trengs det bekreftelse på at mappa **virkelig** er i stand til å vise/representere innehaverens viten og kompetanse innen et gitt felt. Forskning på mapper bør derfor innbefatte mange datakilder, triangulering av dokumentasjon og systematisk analyse av mappeinnhold (Carney 2004). Empiri bør i tillegg hentes fra langtidstudier og inkludere informasjon om 6 kritiske dimensjoner (Zeichner & Wray 2000).

1. Hensikt med mappa!
2. Kontroll med mappa! Hvem bestemmer innhold og rammebetingelser?
3. Presentasjonsmodus! Hvordan er mappeorganisering og format – inkludert teknologien som er valgt for skriving?
4. Sosial interaksjon gjennom arbeidsprosess! Hvordan og av hvilken kvalitet?
5. Grad av involvering i mappa! Av og fra hvem?
6. Bruk av mappa! Spennvidde fra lavrisiko (uformell vurdering) til høyrisiko (formell vurdering).

Som diskusjonen viser, er det fortsatt mange ubesvarte spørsmål relatert til mapper som arbeidsform.

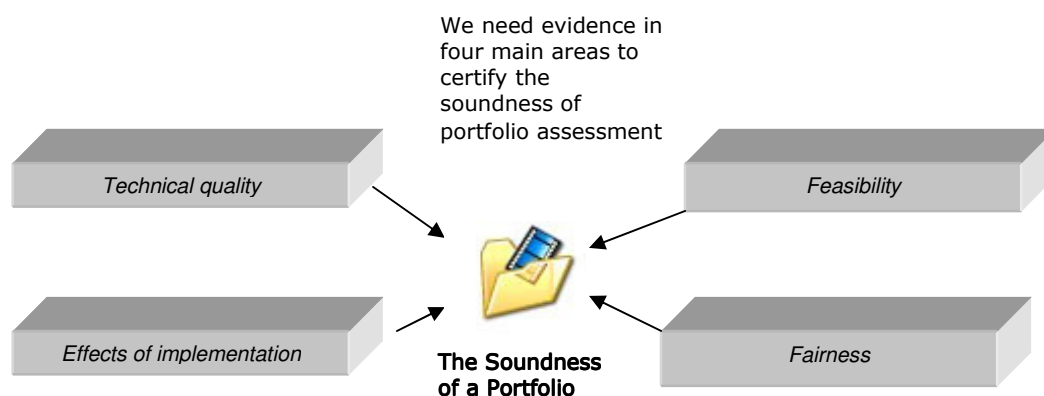
2.5 Rammeverk for sertifisering av mappekvalitet

Allerede i 1994 rettet Herman & Winters oppmerksomheten mot sertifisering/dokumentering av kvalitet i mappevurdering. Sentralt i sertifiseringsrammeverket de utviklet, sto på en helhetlig måte å sertifisere mappekvalitet ”**The soundness of a portfolio**”. I ordbøker oversettes ordet *sound* med ord som: Sikker, sunn, frisk, ekte, sterk, trygg, traust, fast, dyp, uforstyrret, gyldig, lovlig. I den digitale ordboken Clue knyttes ordet soundness til materialet betong, og grad av kvalitet på betongens ulike bestanddeler. For selv å forstå begrepet soundness velger jeg å oversette frasen ”The soundness of a portfolio” med ”**mappekvalitet**”. I frasen legger jeg en mappes evne til så ekte og rettferdig som mulig, å legge til rette for, ivareta og fremvise grunnlaget en studenten skal vurderes på, slik at alle kan stole på at mappen ivaretar studenten på en uklanderlig måte, i henhold til vedtatte føringer og rammer.

Til å sertifisere mappekvalitet identifiserte Herman & Winters fire aktuelle dokumentasjonsområder.

- Technical quality, • Fairness, • Effects of implementation, • Feasibility
(Carney 2004:7)

For å sertifisere en mappes kvalitet, anså Herman & Winters det altså nødvendig å innhente dokumentasjon for delområdene **teknisk kvalitet, ærlighet/rettferdighet, ringvirkning ved å ta i bruk og nytteverdi/egnethet**. Slik jeg tolker sertifiseringsrammeverket, utgjør de fire delområdene et helhetlig rammeverk for sertifisering av en mappes **totale** kvalitet. Om innhenting av dokumentasjon på et av områdene utelates, vil en total sertifisering være mangelfull, slik jeg ser det.



Figur 6: The Soundness of a Portfolio

2.5.1 Teknisk kvalitet “The technical quality of a portfolio”

Frasen ”The technical quality of a portfolio” velger jeg å oversette med ”**En mappes tekniske kvalitet**”. Som nøkkelidentifikatorer for innhenting av dokumentasjon til dette

område, pekte Herman & Winters på fenomenene **reliabilitet** og **validitet**. En mappes tekniske kvalitet ble da i utgangspunktet bestemt, ved innhenting av dokumentasjon på hvor gyldig og pålitelig mappen ble ansett å være. Herman & Winters var generelt mest opptatt av å dokumentere mappens tekniske kvalitet i lys av **vurdering**, og knyttet derfor dette til dokumentasjon av hvor pålitelig og gyldig vurdering var. I en slik sammenheng var det hensiktsmessig å fokusere på fenomener som reliabilitet og validitet. I 2004 videreutviklet Carney rammeverket, og var da opptatt av å dokumentere mappekvalitet i lys av **læring**. Hun foreslo derfor å dokumentere en mappes tekniske kvalitet, ved innhenting av dokumentasjon knyttet til mappens evne til **påvirkning** og **fremvisning** av dyp og ekte læring.

I would suggest technical quality would lie in its capacity for prompting and demonstrating deep learning.
(Carney 2004:8)

I lys av læring foreslo altså Carney å dokumentere en mappes tekniske kvalitet, ut fra andre nøkkelidentifikatorer enn dem Herman & Winters identifiserte. Der digitale mapper brukes som samhandlings- og læringsarena med læring som mål, kan hennes valgte nøkkelidentifikatorer være hensiktsmessige. Skal mappens tekniske kvalitet derfor dokumenteres i lys av læring, er det relevant å rette oppmerksomheten mot mappens evne til å påvirke og fremvise dyp og ekte læring. Å dokumentere teknisk kvalitet, vil da helt og holdent avhenge av mappens formål. Er hensikten å bruke mappen i vurdering, velges andre nøkkelidentifikatorer, enn om hensikten er læring. Ulik hensikt/formål medfører altså dokumentasjonsinnhenting knyttet til ulike nøkkelidentifikatorer.

2.5.2 Ærlighet og rettferdighet “The fairness of a portfolio”

Adjektivet fair oversettes i ordbøker med synonymene skjær, ren, klar, rettferdig, real, ærlig, riktig, rimelig. Substantivet *fairness* oversettes med klarhet, åpenhet, ærlighet, rettferdighet. Frasen “The fairness of a portfolio” velger jeg derfor å oversette til ”**En mappes ærlighet/rettferdighet**”. Generelt knytter Carney dette område til dokumentasjon av hvem som egentlig har produsert mappearbeidene.

A second category Herman & Winters identified as important for research on portfolios is that of fairness. By this, they mean at the outset, *whose work is it?*
(Carney 2004:14)

Hovedhensikten med innhenting av dokumentasjon på dette område, er å dokumentere hvem som egentlig har gjort jobben, med innholdet i mappen. **Grad av støtte** en student får, under arbeid med mappeinnhold, kan bidra til å så tvil om mappens ærlighet/rettferdighet, i den forstand at man spør om det virkelig er studentens eget arbeid, som vurderes. Skal man fastslå/vurdere hva en person virkelig vet og genuint har tilegnet seg av kunnskap, kan stor grad av hjelp/støtte fra medstudenter eller lærere være problematisk, spesielt om mappen brukes som utgangspunkt for viktige beslutninger (karakterer). Rett-

ferdighetsproblemet oppstår slik jeg ser det, når hensikten med mappen er uklar, eller brukes til andre formål, enn det den var tiltenkt. Læringsmapper har først og fremst som hovedmål å bidra til individuell læring uten formell vurdering. Knyttes læringsmappen likevel til formell vurdering, oppstår usikkerhet angående arbeidenes ærlighet.

Et annet område, der det også kan være aktuell å dokumentere ærlighet/rettferdighet, er bias (forskjellsbehandling). Fører for eksempel mapper til et mønster av høy eller lav prestasjon fra studenter av ulik kjønn, ulik sosialøkonomisk bakgrunn, ulik etnisk/kulturell bakgrunn eller ulik rasetilhørighet? Å dokumentere at mapper ikke fremmer **bias** favnes av Wilkerson & Lang (2003).

I arbeid med digitale mapper, må studentene gjøres kompetente til å håndtere teknologien, og ha tilgang til hensiktsmessig maskin- og programvare. Teknologisk parathet vil legge tilrette for arbeid med digitale mapper, men selv om teknologisk parathet blir lagt til rette på studiestedet, skal man til enhver tid være oppmerksom på, om den på noen måte som helst måte disfavoriserer studenter. Også her er et aktuelt dokumentasjonsområde tilknyttet rettferdighet/ærlighet.

2.5.3 Ringvirkning ved å ta i bruk ”The effects of implementation”

Som tredje dokumentasjonsområde til sertifisering av mappekvalitet, mente Herman & Winters at dokumentasjon måtte hentes fra det de kalte ”The effects of implementation”.

Implementation oversettes i ordbøker med ord som gjennomføring, utførelse, realisering, og **effect** med ord som virkning, følge, utførelse og fullbyrdelse. Frasen ”The effects of implementation” velger jeg derfor å oversette til ”**Ringvirkning ved å ta i bruk**”.

I et forskningsprosjekt gjennomført av Mumbi Kariuki (2001) hadde man som mål å kaste lys over ringvirkningene ved å ta i bruk mapper som læringsredskap. I prosjektet var man opptatt av å se på lærerstudenters bruk av laptop i et **ett** år langt feltstudie, der fokus var rettet mot utvikling og bruk av **elektroniske elevmapper** på grunnskolenivå (Elementary school). Dokumentasjonen på ringvirkningene ved å ta i bruk teknologi og elektroniske mapper ble mange. Blant annet viste elevene større entusiasme og motivasjon for læring, og fikk anledning til å innta en veilederrolle både over for lærere og medelever. I tillegg tilegnet de seg tekniske ferdigheter. Å legge til rette for bruk av teknologi og arbeid med elektroniske mapper i en trygg kontekst, her klasserommet, er **en** vei å gå, slik jeg tolker Carney. Det trengs likevel fortsatt mer forskning, for å kunne dokumentere ringvirkninger ved å ta i bruk mapper.

2.5.4 Egnethet “The feasibility of a portfolio”

Feasibility oversettes i ordbøker med ordene gjørlighet og mulighet. Frasen ”The feasibility of a portfolio” velger jeg å oversette til ”**En mappes egnethet**”. På dette område identifiserte Herman og Winters selv en mengde utfordringer, og pekte på at arbeid med mapper krevde vesentlig av både lærere og studenters tid. Dessuten måtte nye roller og ny vurde-

ringspraksis tilegnes, egnede mappeoppgaver og forskrifter tenkes ut, og en kultur for undersøkelse/forskning og refleksjon rundt egen læreprosess etableres.

Carney uttrykker at også elektroniske mapper åpner for nye utfordringer med hensyn til egnethet, og peker på ytterligere 6 dilemmaer (Carney 2001, Carney 2002, Carney 2003). Et dilemma knytter hun til antall formål en mappe skal tjene (en mappe for ett eller mange formål), et annet til personlig avsløring (samhandlingsarena for deling eller avsløring av personlig kunnskap). Også dilemmaene standardiserte systemer kan skape (hemsko kontra grunnlag for måloppnåelse, selvuttrykk og teknisk innsikt, vurdering tilknyttet viktige og mindre viktige beslutninger) peker Carney på og antyder at digitale mapper ikke vil være egnete enheter i lærings- og vurderingsarbeid, dersom ikke dilemmaene tas på alvor.

Å dokumentere en mappes egnethet blir enda mer presserende i et vurderingsperspektiv, enn i et læringsperspektiv. Wilkerson & Lang (2003) anbefaler sågar å **ikke** bruke elektroniske mapper i vurdering, uten kontinuerlig å overveie mappens legale og læringsmessige grunnlag. Brukes mapper som grunnlag for viktige beslutninger (karakterer), blir belysning av egnethet enda mer presserende. Om ikke, vil de bli gjenstand for et utall av utfordringer. Elektroniske mapper har potensial til å fange opp og utvikle både student- og lærerkunnskap, på en måte som gammel teknologi ikke er i stand til, men unnlater man kritisk å evaluere bruken, kan det vise seg at mulighetene feiloppfattes.

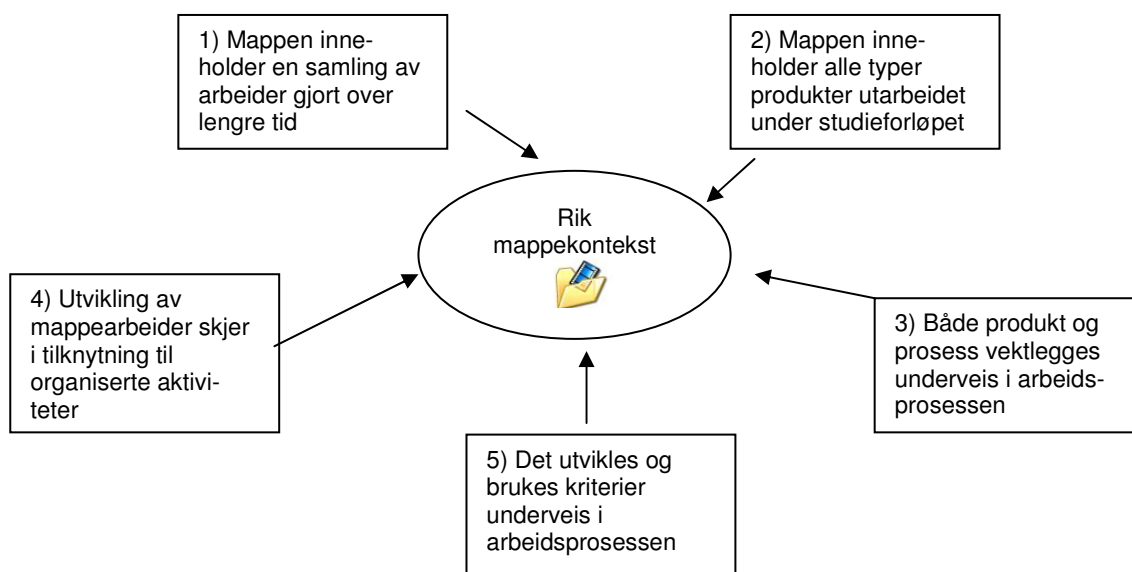
2.6 Oppsummering

I dette kapittel har jeg trukket frem og diskutert teorier som synes relevante og sentrale i forhold til prosjektets problemstilling. Det teoretiske rammeverk presenterer og diskuterer således aktuell teori tilknyttet områdene mappe, vurdering, kriterier, reliabilitet og validitet. I tillegg belyses spesielt sertifiseringsrammeverket utarbeidet av Herman & Winters (1994), et rammeverk som tar sikte på å dokumentere total kvalitet i arbeid med mapper. Med utgangspunkt i en problemstilling, knyttet til belysning av kvalitet, reliabilitet og validitet, synes dette sertifiseringsrammeverket svært relevant. I kapitlet belyses også kort internasjonal forskning tilknyttet bruk av mapper. Dette område er tatt med i det teoretiske rammeverket, som generelt grunnlag for forståelse av internasjonale læringsmiljøers mappepraksis. I kapittel 3 presenteres undersøkelsens metodiske rammeverk, der forskningsdesignen inkluderer belyningsområder tilknyttet problemstilling, datainnsamling, utvalgsplan, analyseplan og undersøkelsen gyldighet.

3 Metode

Prosjektets overordnede mål har vært å belyse kvalitet i arbeid med mapper. Den empiriske undersøkelse er gjennomført med utgangspunkt i to profesjonsutdanninger (jf. 3.4.2) ved HiO. Begge arbeider aktivt med mappevurdering og har etablert det som i litteraturen kalles rik mappekontekst, et læringsmiljø hvor arbeid med mapper står sentralt i organisert studiearbeid. Det brukes formativ vurdering, som preges av samarbeid, respons, bruk av kriterier og egenvurdering. Også ulike former for sluttvurdering knyttes til mappen. Med utgangspunkt i de fire mappekjennetegn nevnt av Dysthe og Engelsen (2003) (jf. 2.1), og et femte nevnt av Wittek og Havnes (2005) knyttet til bruk av kriterier, blir min definisjon på **rik mappekontekst**:

Rik mappekontekst representerer en læringsarena, der mapper brukes aktivt i arbeid med ulike typer oppgaver gjort over lengre tid, og både produkt og prosess vektlegges underveis i arbeidsprosessen. Utvikling av mappearbeider skjer i tilknytning til organiserte aktiviteter, og vurdering gjøres i henhold til kriterier.



Figur 7: Rik mappekontekst

Hovedhensikten med undersøkelsen har vært å identifisere hvilken forståelse praksisfeltet har **kvalitet**, og hvordan det utvikles og utøves i tilknytning til mappevurdering. Å tilføre mappevurdering kvalitet, avhenger i utgangspunktet av hvordan ulike komponenter ivaretas. I kvalitetsprosessen rettes fokus mot ulike områder av betydning, og omfatter således en rekke tiltak i en total helhet. Til å belyse dette prosjekts problemstilling er noen få områder valgt ut.

3.1 Metode for dokumentasjon av mappekvalitet

I 1994 var Herman & Winters opptatt av å finne gode metoder til å dokumentere kvalitet tilknyttet mappevurdering. For å sertifisere/godkjenne mappens totale kvalitet, antydte de allerede da at mappen måtte dokumenteres med hensyn til a) **teknisk kvalitet**, b) **ærlighet** og **rettferdighet**, c) **ringvirkning** ved å **ta i bruk** og d) **nytteverdi /egnethet** (jf. 2.5). Deres sertifiseringsrammeverk har jeg funnet hensiktsmessig å knytte til analysen og drøftingen av det empiriske datamateriale, da primært rettet mot dokumentasjon av en mappes **tekniske kvalitet** (jf. 2.5.1). Også Williams (2000a) uttrykte syn på reliabilitet og validitet ønsker jeg å benytte som metodisk bakteppe i analysen og drøftingen. Disse teoretikerne vil derfor være mine referanser i kapittel 5.

3.2 Problemstilling

Prosjektets overordnede mål har som sagt vært å undersøke hvordan kvalitetsaspektet i arbeid med digitale mapper forstås, utvikles og utøves innen høyere utdanning. Temaet i seg selv er vidt, og jeg måtte derfor gjøre valg med hensyn til relevante områder. Valget falt på å undersøke utvalgte læringsmiljøers forståelse av kvalitet, spesielt knyttet til utvalgte fenomener. Den overordnede problemstillingen belyser derfor følgende:

1. Hvilken forståelse knytter praksisfeltet til fenomenene kvalitet, reliabilitet og validitet?
2. Hvordan utvikles og utøves fenomenene i en mappevurderingskontekst, der formativ vurdering og bruk av kriterier inngår?
3. Hvordan passer fenomenforståelsen inn i sertifiseringsrammeverket utviklet av Herman & Winters?

Problemstillingen tatt i betraktning var det naturlig å se kvalitetsforståelsen i lys av **ett** av dokumentasjonsområdene identifisert av Herman & Winters, nemlig **teknisk kvalitet**. Her trekkes spesielt **reliabilitet** og **validitet** frem. Av aktuelle forskningsspørsmål tilknyttet problemstillingen, ble disse utledet:

- 1) Forståelse av fenomener og begreper:
a) Kvalitet, b) Reliabilitet og validitet, c) Mapper og digitale mapper, d) mappevurdering, e) Formativ vurdering, f) Kriterier.
- 2) Forståelse av utvikling og bruk av pålitelig og gyldig:
a) mappevurdering, b) formativ vurdering, c) kriterier
- 3) Total forståelse knyttet til:
a) Mappekvalitet

3.3 Forskningsdesign

Siktemålet for forskning er å produsere ny kunnskap, som siden gjerne manifesteres i en vitenskapelig teori. Formålet er langt på vei å trekke slutninger, som enten be- eller avkref-ter visse oppfatninger/påstander. Foruten valg av verktøy og teknikk, handler det om å forstå konsekvensen av ulike valg, og om å knytte sammen vitenskapsteori og empiri. I innhenting av data rundt fenomener brukes gjerne en kvalitativ og/eller kvantitativ metode. Det er først og fremst problemstillingen som avgjør, hvilken metodisk tilnærming som egner seg best. For den presenterte problemstillingen (jf. 3.2) ser det ut til å være hensiktsmessig å benytte en kvalitativ metode, fordi jeg ønsker å gå i **dybden** innen et felt, snarere enn bredden (kvantitativ tilnærming). Holter & Kalleberg (1996:25) sier om **kvalitativ** forskning at den preges av kvalitativ tenking gjennom hele forskningsprosessen. Tenkemåter og utsagn formuleres i ord og ikke numerisk, slik det ofte er tilfelle innen kvantitativ metodikk. Ulike faser bindes sammen til en kronologisk og logisk helhet, der det i hver fase tenkes kvalitativt.

3.3.1 Vitenskapsteori, forskningsperspektiv og paradigme



Forskeren befinner seg på kunnskapsstigen vitenskapsnivå. På dette nivå arbeides det med teoridannelse og objekter studeres fra virkeligheten. Min plass som forsker har da naturlig vært plassert under vitenskapsnivået.

Figur 8: Kunnskapsstigen

Som Patel et. al. (1995:10) påpeker, er hovedhensikten med et forskningsarbeid å bidra til kunnskap, der utgangspunktet for forskningsarbeidet i mindre eller større grad er forankret i den kunnskap, som allerede finnes på området. I motsetning til utredning, er forskning vitenskapelig og teoretisk forankret. Teorien representerer et system av innbyrdes relaterte begreper, som til sammen gir et bilde av et fenomen. Til anskueliggjøring av sin teori kan forskeren bruke modeller. Modellene brukes under forskningens gang, som et bånd mellom virkelighet og teori, og kan for eksempel ha form av en skisse, der alle begreper inngår (jf. mine modeller i denne oppgave).

At "være i" et paradigme er ligesom at have taget nogle bestemte briller på.
(Launsø & Rieper 2000:43)

I synet på hvordan forskning og vitenskap skal bedrives har det over tid oppstått ulike skoler. To av disse plasseres innen det samfunnsvitenskapelige paradigme.

Den **positivistiske** tradisjonen, navngitt av den klassiske positivisten A. Comte på midten av 1800-tallet, tar utgangspunkt i en empirisk og naturvitenskapelig tradisjon, der hypoteser og teorier gjerne utformes i form av matematiske formler. Forskning skjer etter den hypotetisk-deduktive metoden, der det fra eksisterende teori utledes hypoteser som prøves empirisk. Teorivalget bestemmer hvilken informasjon, som skal samles inn. Her antar man at all vitenskap i bunn og grunn bygges opp på samme måte, der forskeren preges av objektivitet, er logisk og analytisk, og står i en **ytre** relasjon (usynlig) til forskningsobjektet, som er observer- og målbart. Tesen er at viten kan nås gjennom empirisk erfaring, der komplekse fenomener reduseres til enkelte bestanddeler. Ut fra den **enkelte** bestanddel sies så noe om helheten.

Som motsats til positivistene finner man **hermeneutene**, først og fremst plassert innen en human/kulturell tradisjon.

Hermeneutikk betyr noe slikt som tolkningslære, og er i våre dager en vitenskapelig retning der man studerer, tolker og forsøker å forstå grunnbetingelsene for den menneskelige eksistens. (Patel & Davidson 1995:25)

Hermeneutikk knyttes til tolkning, der ordet i sin opprinnelse er forbundet med den greske guden Hermes. I hermeneutikken rettes fokus mot å studere, tolke og forstå grunnbetingelsene for den menneskelige eksistens, der forskeren inntar en åpen, subjektiv og engasjert rolle. Gjennom det talte og skrevne språket prøver forskeren ved hjelp av tolkning å skaffe seg kunnskap om det ekte menneskelige. Den subjektive forskeren nærmer seg forskningsobjektet ut fra eget ståsted, og pendler mellom å se på helheten og delene (jf. den hermeneutiske spiralen), for å forstå så mye som mulig av den menneskelige eksistens. Forskning utføres etter induktiv metode (oppdagelsens vei), og formuleres til slutt i en teori. Forskningsprosessen i den hermeneutiske spiralen har utgangspunkt i en forståelse av fenomener, uttrykt gjennom menneskelig dialog, for siden å tolkes og føre til ny forståelse med eventuell ny tolkning. Hver runde i spiralen fører til et høyere forståelsesnivå. Mange av dagens samfunnsvitenskapelige forskningsparadigmer/forskningspraksiser tar utgangspunkt i en analytisk (positivistisk) eller systemisk (hermeneutisk) tilnæringsmåte. Faren ved en forskningspraksis innen det analytiske paradigme kan være at vesentlige fenomener i den virkelige sosiale verden overses, mens faren innen det tolkningsvitenskapelige paradigme kan være at forskeren reproducerer de utforskede illusjoner (Launsø & Rieper 2000).

Hvilket samfunnsvitenskapelig **forskningsparadigme** (briller) ønsket så jeg å legge til grunn for dette forskningsprosjekt? Jeg var interessert i å forstå helheten, og gjennom menneskelig skrevet tekst og samtale, prøve å forstå/tolke hvordan mappekvalitet ble forstått, utviklet og utøvd. Det var derfor hensiktsmessig å ta utgangspunkt i et hermeneu-

tisk vitenskapsparadigme. Målet med prosjektet var å lytte til, forstå og tolke menneskelige utsagn (skriftlig og muntlig), knyttet til subjektive opplevelser og erfaringer.

Den fortolkningvidenskabelige tilgang bygger på en antagelse om, at en udtømmende forståelse af menneskelige og sociale handlingsforløb kan opnås ved at tolke og sammenføje den præbevidsthed og bevidsthed, der allerede "ligger hos folk, til et meningsfuldt hele. (Launsø & Rieper 2000:54)

Når et systemisk paradigme er valgt, egner det seg å bruke et åpent, intensivt og beskrivende forskningsdesign.

3.4 Utvalgsplan

Å finne gode praksisfelt som tilfredsstillende kravet rik mappekontekst er utfordrende. Jeg vet om mange som bruker digitale mapper, men få av dem er kommet langt, når det gjelder å bruke mapper helhetlig både i lærings- og vurderingsarbeid. Det utvalget jeg ønsket skulle arbeide aktivt med digital mappevurdering, og gjerne ha gjort det en god stund. Dessuten skulle de aktivt bruke vurderingskriterier og formativ vurdering.

For å få til en meningsfylt undersøkelse var det viktig at det aktuelle utvalget hadde en mening om (forståelse av) fenomenene reliabilitet, validitet og kvalitet. Det var også viktig at personene som ble med i utvalget, var i stand til å formidle sin forståelse. Som aktuelt populasjon for undersøkelsen så jeg for meg enten høyskolenivå eller videregående. Jeg har selv jobbet på HiO i flere år, og visste at SU aktivt siden høsten 1999 hadde arbeidet med studentmapper i sin utdanning. Rapporten 'Vurderingsformer ved HiO' (Wittek & Havnes 2005) bekreftet at dette var et aktuell praksisfelt. Også FJU ved høgskolen arbeidet aktivt med mapper. De to utdanninger fremsto i utgangspunktet som veldig forskjellige, noe som kunne bidra til en bredere forståelse av forskningsspørsmålene.

3.4.1 Utvalgsstørrelse og metode

I litteraturen kan man lese at ved smågrufforsknig er den ideelle gruppestørrelse mellom 5 og 8 personer (Jacobsen 2003:95), alt etter hvor homogen intervjupersonene er som gruppe. Ved homogene grupper kan evt. gruppestørrelsen økes. En gruppestørrelse på 6 personer var formålstjenlig for undersøkelsen. Jeg ønsket i utgangspunktet en så **bred** utvalgssammensetning som mulig, i den forstand at jeg ønsket representanter fra ulike grupperinger (bredt utvalgskriterium) tilknyttet studiemiljøet. Fokusgruppen ble derfor søkt sammensatt av representanter fra disse grupperingene. Gruppestørrelsen ble valgt med det formål at den skulle representere 2 lærere, 2 studenter, 1 ekstern sensor og 1 fra studieledelsen, der målet var å få frem gruppens felles forståelse og eventuelle uenigheter knyttet til temaene i problemstillingen. Jeg ønsket i tillegg å operere med et **typisk** utvalg, der hensikten var å velge ut personer som var representative og typiske for enheten. Da jeg **ikke** selv var knyttet til forskningsfeltet, hjalp utdanningene til med å velge personer, som kunne bli med i utvalget. Et bredt sammensatt utvalg ville ved sin sammensetning

inkludere både respondenter (studenter/lærere/ekstern sensor med direkte kjennskap til fenomenene) og informanter (studieledelse - med kunnskap om fenomenene, men ikke direkte involvert), en sammensetning jeg håpet ville føre til en dynamisk diskusjon under intervjuet. Gruppesammensetningen ble dessverre ikke slik jeg i utgangspunktet hadde tenkt. Det viste seg vanskelig å få med informantene (ledelsen) i fokusgruppene. Den empiriske undersøkelse ble derfor grunnlagt på lærernes, studentenes og de eksterne sensorers forståelse.

Det begrensede utvalget på 2 fokusgrupper og 2 informanter, kan ha bidratt til at representativ og komplett fenomenforståelse ikke kom frem i undersøkelsen. Det er likevel sannsynlig å tro, at de empiriske funn i denne undersøkelse, også ville ha blitt resultatet, om andre enheter, som jobber med formativ vurdering slik disse to studiemiljøer gjør, hadde blitt undersøkt. En forutsetning er da at mappevurdering bygges på et sosiokulturelt læringsparadigme, der samarbeid og sosial samhandling er dominerende. Formålet med mappevurdering vil da være å ivareta funksjoner som den formative vurdering skal fylle. En slik bekreftelse av funnene, ville ha bidratt til å styrke undersøkelsens gyldighet.

Også den endelige utvalgssammensetning kan ha vært en svakhet. Grunnet ulike omstendigheter var ingen av lederne for studiene, i stand til å delta i intervjuene. Deres forståelse er derfor ikke representert, og viktig informantforståelse kan ha blitt utelatt. At lederne ikke var representert i fokusgruppene, gjorde også at utvalgskriteriet bredde og variasjon ikke ble helt innfridd. Undersøkelsens validitet kunne generelt ha blitt styrket ved å gjennomføre intervju med flere og andre grupperinger innen studiemiljøene, altså å styrke utvalget. I lys av vanskelighetene med å skaffe respondenter/informanter, ville det da ha vært en fordel å være kjent innen studiemiljøet, slik at man på forhånd hadde en form for goodwill.

3.4.2 Utvalgsenhet

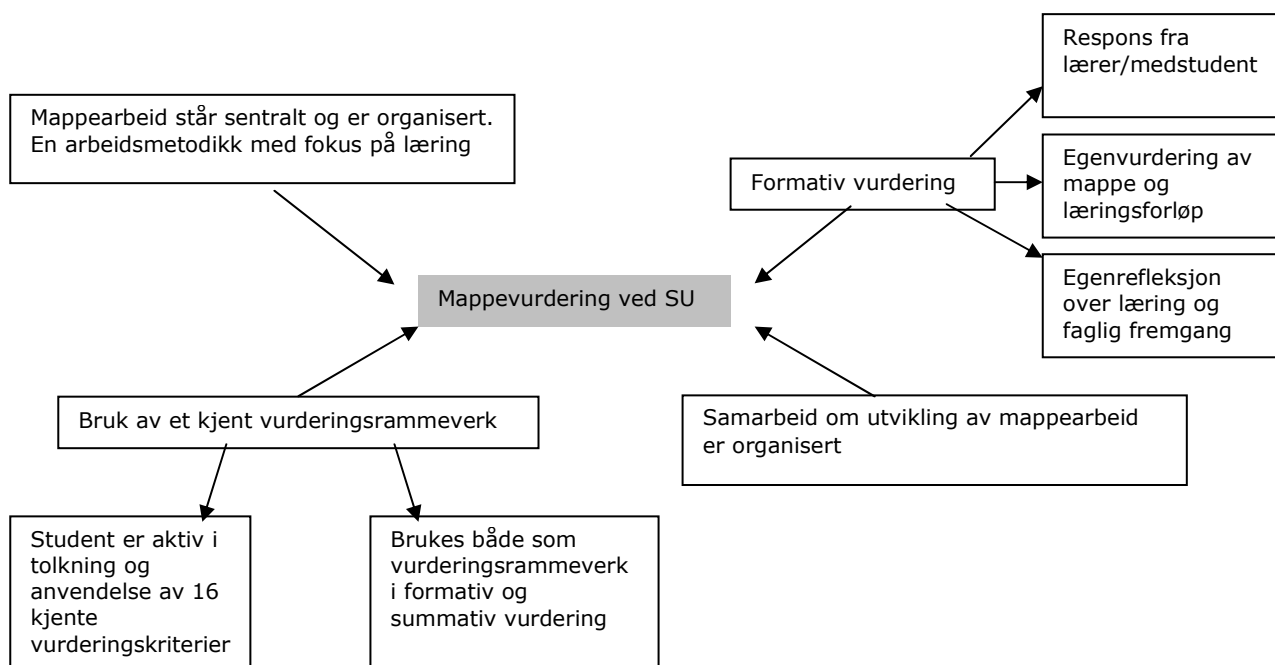
a) Avdeling for sykepleierutdanning (SU)

Digitale mapper har siden studieåret 1999/2000 vært brukt som arbeids- og vurderingsform ved SU på HiO. Utdanningen har således tilegnet seg god og bred erfaring med dette vurderingskonsept. Argumentasjonene for å ta i bruk vurderingsformen var knyttet til et ønske om å ta bort fokus fra pugg, slik det vanligvis praktiseres i gamle etablerte eksamensformer, og rette oppmerksomhet mot selve læringsprosessen. Også vurderingsansvar mellom student og lærer/sensor ønsket de å dele, gjennom bruk av medstudentrespons (Aae et al:2002). Ved å dele vurderingsansvar mellom student og lærer, oppsto kravet til synlige og kjente vurderingskriterier. Et kjent vurderingsrammeverk medvirket til å skape en felles plattform for vurdering av et produkt/studieoppgave.

Hvert studiekrav blir vurdert i forhold til ett eller flere av nedennevnte kriterier. I mange tilfelle velger studentene selv hvilke kriterier de vil ha vurdert sitt arbeid i forhold til. Deres valg må da begrunnes. (Aae et al:2002)

De obligatoriske studieoppgavene vurderes i forhold til ett eller flere vurderingskriterier. Ett av de viktigste grepene var derfor å trekke studenter med i utvikling av vurderingskriterier og involvere dem i vurderingsarbeidet.

Mappevurdering ved SU (Wittek & Havnes 2005), beskrives med at mappearbeid er sentralt og organisert. Det brukes formativ vurdering og eksplisitt nedskrevne vurderingskriterier som vurderingsrammeverk. Vurderingskriterier brukes systematisk, gjennom hele lærings- og arbeidsprosessen, og inngår i tillegg som del av summativ sluttvurdering. Kriteriene revideres årlig i samråd med studentene. Studenter tolker og anvender aktivt kriteriene i studieoppgavene, og har medinnflytelse på utforming av studieoppgavene (ikke 1. halvår). Studieoppgaver knyttes til ulike sjangrer, der oppgaveutvikling skjer i samhandling. Tilbakemelding gis i form av respons fra medstudent eller veileder/lærer. Innleveringsfrister for studieoppgaver bestemmes av lærer eller studenten selv.



Figur 9: Mappevurdering ved sykepleierutdanningen

Det etableres studiegrupper ved oppstart av hvert studieår. Disse jobber sammen hele studieåret, der planlegging og gjennomføring skjer i henhold til detaljerte retningslinjer. Et visst antall studiearbeider løses som gruppearbeid. Hensikt og formål med respons er utvikling og forbedring og fremkommer i form av medstudentrespons, lærerrespons og egenvurdering/refleksjon. På utdanningen opereres med flere mappetyper. I fagmappen (læringsmappen) finnes ulike studieoppgaver og refleksjoner, mens det i vurderingsmappen finnes enkelte arbeider valgt ut av lærer eller student/lærer i fellesskap. De utvalgte studentarbeidene inngår i summativ vurdering, der enten ett arbeidskrav karaktersettes (bokstav/bestått/ikke bestått), eller ett eller flere mappeelementer trekkes ut og benyttes som utgangspunkt for en hjemmeeksamen. Mappevurdering slik det praktiseres ved SU, kan da visuelt fremstilles slik modellen i figur 9 viser.

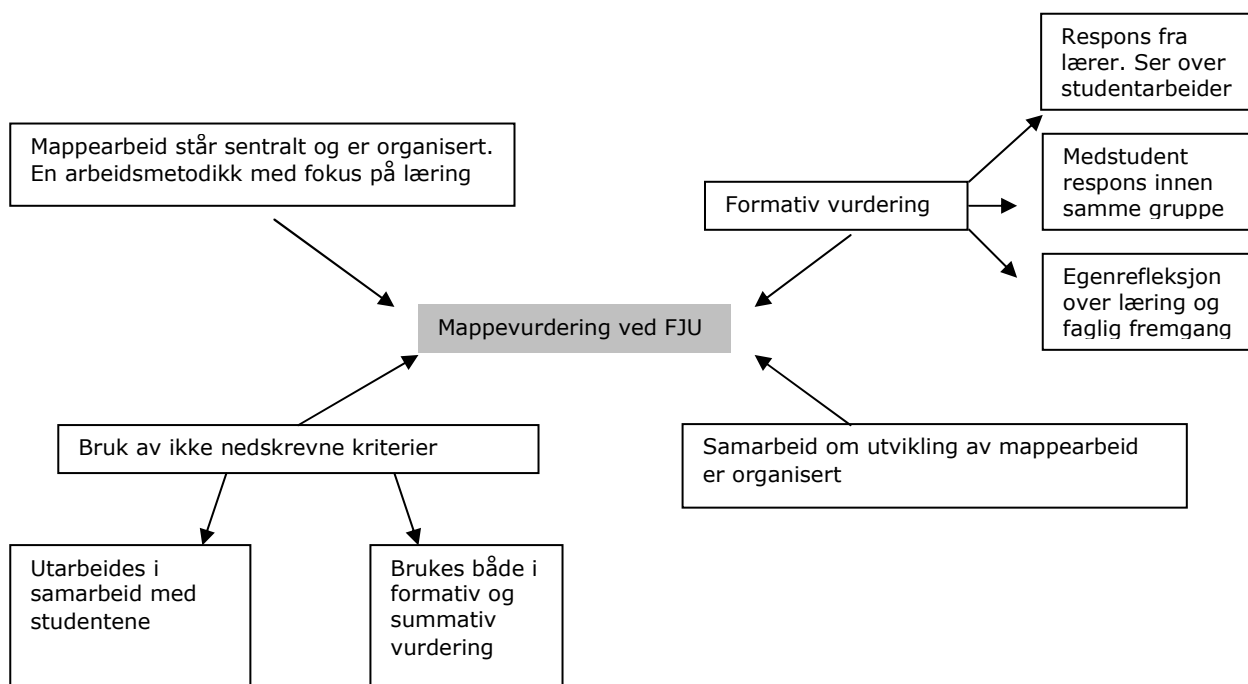
Mappevurdering ved SU plasserte seg innenfor min definisjon av rik mappekontekst, og fremsto som et ideelt praksisfelt for undersøkelsen.

b) Avdeling for journalistikk, bibliotek- og informasjonsfag (FJU)

I studieplanen for bachelorprogrammet i fotojournalistikk ved HiO står det å lese:

Studentene samler alle ferdige og godkjente oppgaveløsninger, utkast og underkjente oppgaveløsninger i ei mappe. Gjennomgående brukes mappevurdering i undervisning. Arbeidskravene (obligatoriske oppgaver) er helt sentrale i arbeidet med mappa. På Fotojournalistikkutdanninga legges det til rette for at mappa er elektronisk. (Avdeling for journalistikk, bibliotek- og informasjonsfag 2005:46)

Det understrekes i studieplanen at poenget med løpende mappevurdering er at både studenter og lærere effektivt skal kunne få et inntrykk av en students faglige utvikling. Vurdering skjer gjennom hele læringsprosessen, og er en integrert del av undervisningsforløpet. I summativ vurdering (muntlig eksamen) kan mappa brukes på ulike måter, enten i forbindelse med skriving av refleksjonsnotat tilknyttet utvalgte arbeider/pensum, som utgangspunkt for en muntlig samtale eller som grunnlag for en kvalitetsbedømmelse (Wittek & Havnes 2005). Mapper ved FJU er en integrert arbeidsform knyttet både til undervisning, læring og vurdering, og ukentlig samarbeides det om mappeoppgaver.



Figur 10: Mappevurdering ved fotojournalistutdanningen

Tilbakemelding (respons) fra lærer og medstudenter er systematisert, ved at det etableres faste responsgrupper, som møtes ukentlig for å diskutere mappearbeidene. Både lærere og studenter er til stede på disse møtene. Kriterier brukes både underveis i formativ vurdering og til slutt ved summativ vurdering, og utarbeides i samarbeid mellom lærer og student gjennom hele læringsprosessen.

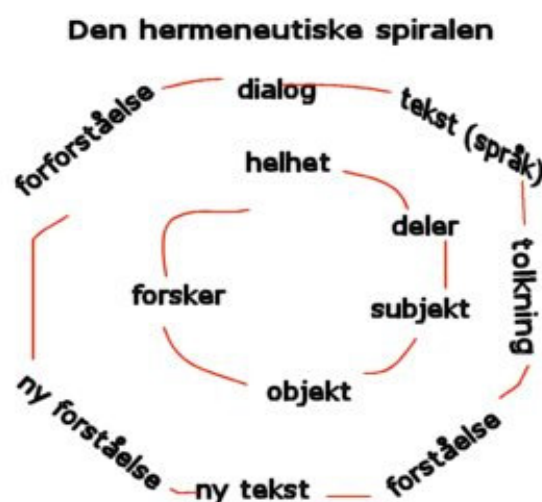
Studentene leverer både individuelle og gruppeoppgaver, som godkjennes av lærer. Produksjonsoppgavene formuleres av lærer og kan være arbeider av ulik sjanger og form. Responsgrupper etableres fra studiestart, og det gis systematisk respons på alle arbeider. Det arrangeres skrivekurs i hvordan man skal gi respons, og studentene arbeider med både en samlemappe og en vurderingsmappe. Sistnevnte mappe består av studentarbeider, valgt ut av lærer, som vurderes og karaktersettes, og tjener som utgangspunkt for muntlig eksamen. En justering av karakteren fastsettes etter skriftlig eksamen. Også FJU plasserte seg innenfor min definisjon av rik mappekontekst, og fremsto som aktuelt praksisfelt for undersøkelsen.

3.5 Innsamling av data

For prosjektet er det som sagt valgt en kvalitativ tilnæringsmetode og problemstillingen belyses ved det åpne kvalitative intervju, aktuelle dokumenter og utvalgte studentoppgaver. Min subjektive tolkning og forståelse ble således til både forut for intervjuene (lesing av dokumenter), under intervjusamtalene, og i etterkant av intervjuene (bearbeiding av innsamlet intervjudata og studentoppgaver). Resultatet ble en subjektiv beskrivelse og tolking, som siden ble drøftet og knyttet til sertifiseringsrammeverket utviklet av Herman & Winters (1994) og Williams (2000a) uttrykte forståelse av reliabilitet og validitet. Min involvering i læringsmiljøet var derfor tredelt, der en del var rettet mot å forstå rammene/faktorene som var med på å gjøre mappevurdering kvalitativ, en annen mot å tolke forståelsen forskningsfeltet hadde tilknyttet kvalitativ mappevurdering, og en tredje mot å forstå/tolke kvalitativ respons i studentoppgaver.

3.5.1 Forskerrolle

Min egen forskerrolle vil jeg beskrive som det Launsø & Rieper (2000) kaller 'Den forstående forskningstype'. Forskningstypen er en rendyrket form, med særegne trekk, der forskeren utvikler forståelse i språklig dialog. Under intervjuene ønsket jeg at min opptreden skulle domineres av en mellomting mellom å være aktiv og passiv, i en lyttende posisjon der forståelse ble vist ved nikk og bebreftelser. I tillegg ønsket jeg å være observerende og



Figur 11: Den hermeneutiske spiral

eksplorerende ved å stille spørsmål (grave og spørre). Hensikten med intervjuene var å rette fokus mot prosjektets prioriterte temaområder. Det var derfor viktig for meg å fungere som en slags ordstyrer under samtalen, med fokus på intervjuet agenda. Etter intervjuene

gikk jeg inn i en tolknings- og analysefase, der jeg subjektivt tolket og beskrev respondentenes forståelse av fenomenene. Prosessen førte til ny forståelse (jf. figur 11).

3.5.2 Det åpne kvalitative intervju

I selve interviewet må intervjueren nødvendigvis gå inn som deltager (Fog 2004:35)

Fog antyder at en forutsetning for et kvalitativt og ekte intervju, er at intervjueren involverer seg som deltaker i intervjusituasjonen og skaper likeverdig og gjensidig kontakt mellom seg og intervjuobjektet(ene). Intervjuer skal bli en vesentlig del av feltet som undersøkes. Andre uttrykker at forskerintervjuet er en samtale mellom to parter om et tema av felles interesse, men at det likevel ikke kan bli en likeverdig samtale, da det er forskeren som definerer og kontrollerer situasjonen (Kvale 2001). Jeg kan for så vidt være enig i begge synspunkter.

Det åpne intervju egner seg best under følgende forutsetninger:

- a) Når relativt få enheter undersøkes.
 - b) Når vi er interessert i hva mennesker sier.
 - c) Når vi er interessert i hvordan mennesker tolker og legger mening i et spesielt fenomen.
- (Jacobsen 2003:71/72)

I litteraturen nevnes spesielt det **åpne intervju** som datainnsamlingsmetode, når forskeren er interessert i hvordan mennesker tolker og legger mening i et spesielt fenomen (Jacobsen 2003). Det finnes mange ulike inndelinger av intervju som forskningsmetode; åpne eller strukturerte, retrospektive, gruppeintervjuer, forskningssamtaler og samtaler for å nevne noen. Felles for alle er å samtale med et utvalg av mennesker, for å få rede på hvordan de oppfatter viktige forhold i hverdagen (Mellin-Olsen 1996). Samtalen oppfattes som en arena, der sosial samhandling struktureres både av respondent/informant og forsker. Det skilles generelt mellom standardisert og reflektert intervju.

Undersøkelsens problemstilling tatt i betraktning egnet det åpne reflekterte intervju seg for belysning av temaene. Jeg ønsket dog ikke en 100 % åpen ustrukturert intervjusamtale, der intervjupersonene åpent og fritt kunne snakke uten noen form for temastyring, men mer en prestrukturert intervjusituasjon, der jeg som intervjuer ved hjelp av en intervjuguide, fikk frem gruppens forståelse, tilknyttet prosjektets problemstilling.

Den feilen vi ofte gjør, er at vi antar at mottakeren har det samme bildet av virkeligheten som vi selv har dannet oss. (Bøyesen et al. 1997)

Vi individer danner ofte vår egen forståelse av virkeligheten, og tror at andre individer har samme oppfattelse. Under intervjuene ønsket jeg i så stor grad som mulig, å få frem intervjupersonenes forståelse av hva som tilførte digital mappevurdering kvalitet. Egen forståelse, ønsket jeg først skulle komme til syne etter å ha tolket, analysert og drøftet datamaterialet. Både skriftlig i forkant og muntlig ved oppstart av samtalene la jeg derfor vekt på, at hovedintensjonen var å få vite mest mulig om hvordan respondentene jobbet med å tilføre mappevurdering kvalitet, og var opptatt av deres refleksjoner i tilknytning til dette.

Intervjuguiden ble brukt som en huskeliste under intervjuet, og hadde spørsmål tilknyttet problemstillingen. Noen fast agenda var ikke aktuell. Temaene ble tatt opp i den rekkefølgen, det var hensiktsmessig.

Under intervjuene merket jeg at det for meg som intervjuer av og til var krevende å følge samtalen/dialogen. Årsaken til det var nok at respondentene helt naturlig hadde et mer helhetlig syn på og erfaring med studiet og studiesituasjoner, enn det jeg som utenforstående intervjuer hadde. På tross av at jeg på forhånd hadde lest tilgjengelig litteratur og beskrivelse av studiemiljøene, var det til tider likevel som utenforstående vanskelig å forstå helheten. I ettertid erfarte jeg derfor hvor verdifulle lydbåndopptakene var, da de bidro vesentlig til utvikling av min forståelse og tolkning av respondentenes uttrykte forståelse tilknyttet de fokuserte begreper og fenomener. Uten lydbåndopptakene, ville en forståelse og tolkning ha vært ufullstendig og mangelfull. En nitid transkribering av lydbåndene, med etterfølgende bearbeiding og fortolking av respondentenes forståelse, var i den sammenheng påkrevd.

3.5.3 Bruk av fokusgrupper og intervju

Til å belyse problemstilling ønsket jeg å benytte fokusgrupper, hvor respondentenes ulike erfaringer knyttet til problemsstillingens temaer kunne fremføres.

Mange har påpekt at gruppeintervjuer først og fremst er gode hjelpemidler når det er et relativt avgrenset tema som skal diskuteres.
(Jacobsen 2003:72)

Da undersøkelsen var knyttet til å belyse et relativt avgrenset tema, syntes fokusgrupper hensiktsmessig. Hensikten med fokusgruppene var å få frem en bredt sammensatt gruppes forståelse, tilknyttet de utvalgte begreper og fenomener. Ved å samle respondenter fra ulike grupperinger i studiemiljø, håpet jeg å få til en dynamisk dialog/diskusjon rundt undersøkelsens problemstilling. Med utgangspunkt i en fenomenologisk tradisjon, bygget jeg således forståelsen på, at virkeligheten var slik fokusgruppene umiddelbart oppfattet og opplevde den (Kvale 2001). Med begrunnelse i et sosiokulturelt perspektiv (Dysthe 2001), er en gruppeprosess hvor dialog står i fokus, egnet til å få frem hvert enkelt individs forståelse, i tillegg til hva gruppen som helhet mener. Opprinnelig var tanken at utvalgssammensetningen skulle bestå av lærere, studenter, eksterne sensorer, samt en fra studieledelsen. Det viste seg imidlertid vanskelig å hente inn eksterne sensorer til HiO. I tillegg var ledelsen ved studiene opptatt i tidsrommet intervjuene var planlagt gjennomført. Med disse forutsetningene ble fokusgruppene til slutt satt sammen av representanter fra lærer- og studentgruppen. Ved SU besto gruppen av tre lærere og to studenter (2. og 3. studieår), alle tilknyttet deltidsstudiet. En av lærerne representerte i tillegg ledelsens synspunkt, da vedkommende hadde deltatt i arbeidet med mappevurdering ved utdanningen, helt fra metodikken ble innført i 1999. Gruppen var således bredt sammensatt av representanter fra deltidsut-

danningen. Fokusgruppen på FJU besto, i likhet med gruppen på SU, også av tre lærere og to studenter. Lærerne i denne gruppen representerte ulike fagområder tilknyttet studiet, mens studentene representerte 1. og 3. studieår. At fokusgruppene ved begge utdanningene til slutt besto av 3 lærere og 2 studenter var en ren tilfeldighet. Forfall og vanskeligheter med å få fatt i respondenter var avgjørende faktorer i den sammenheng.

På SU var det tett samarbeid mellom eksterne sensorer og studiet underveis i arbeidsprosessen. På FJU ble de kun koblet inn når summativ eksamen ble avviklet. Problemstilling tatt i betraktning, fant jeg det derfor kun relevant å ha telefonintervjuer med eksterne sensorer tilknyttet SU. Alle intervjuene ble tatt opp på lydbånd, og nøyaktig ordrett transkribert i etterkant.

Intervjuene med fokusgruppene, ble gjennomført tidlig på dagen, i et møtelokale lokalisert i bygningen der utdanningene holdt til. De forløp uten forstyrrelser, og ble gjennomført innen den stipulerte tiden. I intervjuet var håpet å få til en utveksling av synspunkter, mellom de ulike grupperinger, og at gruppeintervjuet ville fungere som et forum, der respondentene kunne reflektere over de utvalgte fokusområdene, og fremføre egne spørsmål og svar.

Ved gjennomføring av intervjuene informerte jeg på forhånd respondentene/informantene om hovedhensikten med undersøkelsen, og at jeg tenkte å gjøre lydopptak. Jeg orienterte også om at lydopptakene kun var tiltenkt brukt i analyse av datamaterialet. Begrunnelsen for å ta opp intervjuene på lydbånd var, at jeg i ettertid nøye kunne gjennomgå materialet, og analysere respondentenes ulike perspektiver. Å sette seg i en situasjon, der kun løse setninger nedskrevet under intervjuet var eneste håndfaste dokumentasjon, ville slik jeg så det medføre mangelfull og sårbar innsamling av primærdata. Et lydopptak ville bidra til at alt som ble sagt i intervjusituasjonen var tilgjengelig for meg i ettertid, og således legge til rette for sann forståelse og tolking også i etterkant av intervjuene. Jeg kunne da under intervjuene fokusere på å stimulere respondentene til å bidra med sine kunnskaper og erfaringer. Notater (vedlegg 2) jeg tok underveis i intervjuet, fungerte som en slags indeks i forhold til lydopptaket. Etter avsluttet intervju, ble opptaket nøye gjennomgått og knyttet til mine nedskrevne notater. Selve intervjutiden ble satt til 1½ time, og ble gjennomført i møtelokaler tilknyttet utdanningene.

Erfaringen var at intervjuene dessverre ikke ble preget av diskusjon, i så stor grad jeg hadde håpet. Respondentene var til dels mer opptatt av enkeltvis å gi svar på spørsmålene som ble stilt. Forståelse uttrykt under intervjuet fremkom derfor mest i lys av enkeltindividet. Min måte å oppmuntre til felles refleksjon/diskusjon var å stille spørsmål av typen, "Er dere andre enig i det X her sier?". Manglende diskusjon kan skyldes at respondentene stort sett var enig i det som ble sagt, at det rett og slett var områder som var vanskelige å diskutere, eller at jeg som intervjuer ikke var flink nok til å understreke diskusjonsaspektet.

3.5.4 Intervjuguide

I utarbeiding av intervjuguiden (vedlegg 1), anså jeg det som viktig å holde fokus på den uttrykte problemstilling. Spørsmålsutformingen ble derfor knyttet til områdene mapper, formativ vurdering og kriterier, da med spesiell fokus på forståelse av fenomenene reliabilitet, validitet og kvalitet. Jeg ønsket å utarbeide en intervjuguide som ivaretok prosjektets fokusområder, men som også med relativt åpent utformete kategorier, kunne få frem hvordan aktørene oppfattet, utviklet og utøvde prosjektets fokusområder i deres arbeids- og læringsmiljø. Samme intervjuguide ble brukt ved gjennomføring av alle samtaler. I ettertid var erfaringen at intervjuguiden fungerte som samlende agenda under intervjuene. Når respondentenes refleksjon kunne knyttes til ustilte spørsmål, unnlot jeg å stille spørsmål, som da allerede var besvart.

Ordet reliabilitet hadde jeg i intervju spørsmålene erstattet med ord som pålitelig, pålitelighet og noe man kan stole på, slik at fenomenet ble knyttet til lettforståelige norske ord og vendinger. Likeså brukte jeg for validitet, ord og vendinger som målrettet og noe som er gyldig. Kvalitet valgte jeg å bruke ordrett slik det var, da jeg regnet med at dette var et ord, som var lettere for respondentene å tillegge en umiddelbar mening. Jeg regnet også med at respondentene i samtalen løp ville spesifisere nærmere hva de tilla både begrepene og fenomenene. Erfaringen under intervjuene var at respondentene umiddelbart fant det vanskelig å uttrykke forståelse tilknyttet fenomenene reliabilitet og validitet, selv i en fornorskert versjon. Årsaken til det, slik jeg tolket respondentene, var at det var uvant å knytte slike fenomener til det de arbeidet med, innen sin mappevurderingskontekst. Det falt derfor hensiktsmessig at respondentene, etter å ha tenkt over spørsmålene jeg stilte, selv spesifiserte hva de knyttet til fenomenene. Ved å fortelle hvordan de utøvde mappevurdering, eksemplifiserte de indirekte hvilken forståelse de knyttet til reliabilitet og validitet. Å uttrykke forståelse tilknyttet begrepene mappe, formativ vurdering, kriterier og fenomenet kvalitet falt umiddelbart lettere for respondentene.

3.6 Analyseplan

En hermeneutisk analyse veksler mellom å analysere ulike deler og se delene i en større sammenheng/helhet, der forskeren underveis i analyseprosessen inntar ulike perspektiver (jf. 3.3.1 og figur 11). Med utgangspunkt i en subjektiv analyse, vil tolkning selvfølgelig også til enhver tid bære preg av subjektiv erfaringskunnskap og forståelse. Som grunnlag for tolkningene i undersøkelsen er det brukt datatriangulering. Her sammenstilles respondentenes forståelse, med opplysninger tilgjengelig i fagplaner/studieplaner og i utvalgte studentoppgaver.

3.6.1 Analyse av dataene

Rådataene valgte jeg å **innholdsanalysere**, der analysearbeidet besto av følgende faser:

1. Kategorisering. Utgangspunktet her var intervjuguiden.
2. Fylle kategoriene med innhold. Utgangspunktet her var sitater fra de ulike intervjupersonene, hentet fra lydbåndene.
3. Sammenlikning av intervjuene og lete etter forskjeller/likheter.
4. Innplassering av empiriske funn i teoretisk rammeverk.

Kategoriseringen i fase 1 fulgte temaene, som ble tatt opp under intervjuene. Det var derfor naturlig å velge kategorier knyttet til forståelse av fenomenene kvalitet, reliabilitet, validitet, og begrepene mappe/mappevurdering, formativ vurdering og kriterier. I fase 2 ble lydopptakene nøye gjennomgått og **kategoriene fylt** med sitater. Fase 3 besto i å **sammenlikne** forståelsen fremkommet under intervjuene og finne frem likheter og forskjeller dem imellom. I fase 4 ble funnene belyst og innplassert i teoretisk rammeverk.

I drøftingen knyttet jeg funnene til det teoretiske rammeverk, da med spesiell belysning av sertifiseringsrammeverket utarbeidet av Herman & Winters (1994), og Williams (2000a) uttrykte forståelse av reliabilitet og validitet. I Herman & Winters rammeverk, ble de empiriske funn knyttet til belysning av dokumentasjonsinnhenting tilknyttet en mappes tekniske kvalitet (jf. 2.5.1). Foruten området teknisk kvalitet var også Herman & Winters opptatt av å dokumentere mappekvalitet ved å belyse **ærlighet/rettferdighet, ringvirkning ved å ta i bruk** og **egnethet** (jf. 2.5). Disse dokumentasjonsområder inngikk ikke som del av prosjektet, da problemstillingen kun var rettet mot å belyse teknisk kvalitet.

3.7 Undersøkelsens gyldighet

3.7.1 Intern gyldighet

Test av **intern gyldighet** ble gjennomført vha respondentvalidering (gyldighet ved første blick - face validity), der respondentene vha et enkelt avkryssingsskjema, enkeltvis skulle uttale seg om undersøkelsens sentrale funn og konklusjoner. Skjemaet ble utformet med kategoriene **kjenner meg igjen** og **kjenner meg ikke igjen** (vedlegg 3), med mulighet for utdyping. Tilbakemeldinger uttrykker at man kjenner seg igjen i den empiriske beskrivelse. Det bemerkes også at betydningen av gruppe og gruppeaktiviteter ikke kommer så godt frem. Generelt var hovedhensikten under intervjuene først og fremst å belyse fokusgruppene forståelse av reliabilitet, validitet og kvalitet, men knytter man utvikling av fenomenforståelse til samhandling, spesielt fremhevet av studiemiljøene som et kvalitetselement, har gruppe og gruppeaktivitet absolutt betydning. Nærmere belysning av dette aspektet ville derfor ha styrket undersøkelsens interne validitet. I tillegg til respondentvalidering, ble den interne gyldighet testet ved en kritisk gjennomgang av kilder og informasjon fra kildene. Intervju ble benyttet som metode for innsamling av data tilknyttet samtaler med respondentene. Denne metoden har imidlertid sine svakheter. Forskeren/intervjueren kan påvirke

intervjusituasjonen ved sin måte å stille spørsmål, ved å fremstå som autoritær, ved ikke å oppfatte hva intervjupersonene uttrykker, eller ved ikke å skape tillit (Kvale 2001). Håpet er at ingen av disse gjorde seg gjeldende i intervjuene tilknyttet denne undersøkelsen. Man kan likevel ikke se bort fra at noen av respondentene kan ha vært påvirket av intervjuerens personlige fremtoning, og dermed ikke fikk uttrykt det de hadde på hjertet. Det kan ha ført til at ikke all forståelse kom frem.

Også intervjupersoner har betydning for intervjuets gyldighet. Respondentene var spesielt valgt ut i kraft av å representere kompetente kilder i studiemiljøet, i lys av prosjektets problemstilling. Lærerrepresentantene ble valgt av studieledelsen, fordi de var involvert i arbeidet med formativ vurdering, studentrepresentantene ble valgt av lærerrepresentantene, og eksterne sensorer anbefalt av lærerrepresentantene fordi de samarbeidet tett med læringsmiljøet. Under intervjuene opplevde jeg at respondentene var flinke til å reflektere over og uttrykke sin forståelse, tilknyttet de fokuserte emnene. Man kan likevel ikke se bort fra, at alle var like trygge på å ta frem sin forståelse, eller at kommunikasjonen mellom oss, som deltok i samtalen, var god nok. Faren for at respondenter opptrådte unormalt i intervju-situasjonen kan også ha vært til stede. Slik jeg observerte samhandlingen under samtalen, var inntrykket dog at alle i gruppen(e) så ut til å være trygge. Ingens adferd ble derfor fra min side oppfattet som unormal. Jeg viser ellers til andre feilkilder tilknyttet intern gyldighet, beskrevet underveis i kapitlene forut.

3.7.2 Forskningshorisont

Forskningsprosjektet var basert på hermeneutiske grunnprinsipper, og tok sikte på å studere, tolke og forstå grunnbetingelsene for pålitelig og gyldig formativ vurdering, uttrykt gjennom menneskelig dialog og i nedskrevne dokumenter. Hensikten var, gjennom det talte og skriftlige språket, å innhente kunnskap om forståelse av mappekvalitet, og gjennom det skrevne språket formidle mine tolkninger. Generelt nærmer den subjektive forskeren seg forskningsobjektet fra sitt eget ståsted, og pendler mellom å se helheten og delene (jf. den hermeneutiske spiralen), for å forstå så mye som mulig av den menneskelige eksistens. Mitt subjektive utgangspunkt, hva gjelder innsikt i formativ vurdering, var på ett absolutt nullpunkt, og veien til en helhetlig forståelse besto derfor av innsamling og lesing av skriftlige dokumenter, lytting og deltaking i intervjuer, skriving/transkribering og fortolkning. Forståelsen ble således til via oppdagelsens vei (induktiv metode).

Faren innen det tolkningsvitenskapelige paradigme (hermeneutikken) kan være at forskeren reproducerer de utforskede illusjoner (Launsø & Rieper 2000). I en undersøkelse, som tar sikte på å tolke og forstå omverden gjennom det talte og skrevne ord, er absolutt denne fare til stede. En sammenlikning av flere forskeres tolkning, ville derfor ha bidratt til å styrke undersøkelsens gyldighet. Det ville dog være svært ressurskrevende, og avhengig av at flere ønsket å forske på det samme. Generelt ble det skriftlige og muntlige datamaterialet innholdsanalysert flere ganger. En slik fremgangsmåte øker med all sannsynlighet sjansen

for en valid tolkning, slik jeg ser det. I undersøkelsen ble det også benyttet datatrigulering. Man kan imidlertid ikke se bort fra at et enda bredere datagrunnlag, kunne ha styrket undersøkelsens validitet.

3.7.3 Generalisering og overføring av resultatene til andre utvalg

Kvalitative studier kjennetegnes av at de gjennomføres innen et avgrenset forskningsfelt, der målet først og fremst er å si noe om konteksten som undersøkes. Metoden medfører dermed også begrenset gyldighet. For å styrke undersøkelsens allmenne gyldighet hadde det, foruten det nevnt i avsnittene over, vært naturlig å sammenholde undersøkelsens empiriske funn, med andre liknende undersøkelser. Det hadde også vært relevant å knytte de empiriske funn til andre aktuelle teorier. Hva gjelder sistnevnte er i det i denne oppgaven fokusert på teorien til Herman & Winters (1994) og William (2000a).

3.8 Oppsummering og kommentarer

Som utgangspunkt for forskningsdesignen hadde jeg en eksplorerende problemstilling. Undersøkelsen var basert på innsamling av kvalitative data, da med en intensiv design, og hadde en tverrsnittsundersøkelse som målsetning.

Innsamling av data ble gjort gjennom aktuelle skriftlige dokumenter, i det åpne reflekterte kvalitative forskningsintervju, og ved studie av studentoppgaver med tilhørende respons. Intervjuene ble gjennomført med utvalgte fokusgrupper, der samtalene ble tatt opp på lydbånd. Fokusgruppene besto av personer, som besatte ulike roller innen miljøet, og utvalget var således basert på bredde, variasjon, informasjon og de typiske. Den empiriske undersøkelse og analyse var rotfestet i hermeneutisk vitenskapsteori, der de empiriske dataene ble innholdsanalysert og tolket. Til å bestemme undersøkelsens gyldighet, ble det til slutt fokusert på intern gyldighet. Denne ble knyttet til refleksivitet (Hammersley 1987) intervjuereffekt, konteksteffekt og respondentvalidering.



(Hein 1968:6)

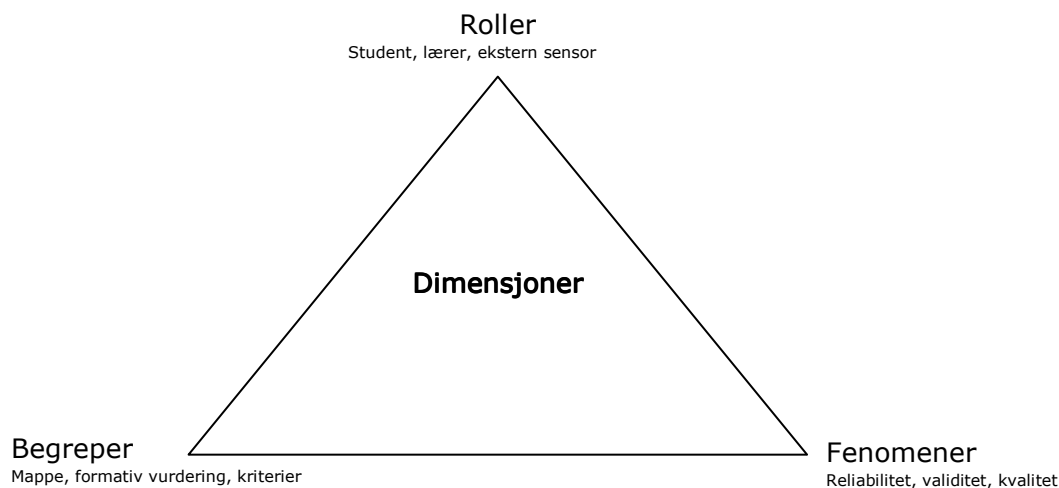
Figur 12: Den som tvivler (Piet Hein)

4 Empiri

Empirien i denne undersøkelse er hentet via skriftlige dokumenter og intervju med fokusgrupper og enkeltpersoner. Intervjuene ble gjennomført ansikt til ansikt, og på telefon senhøstes 2005. Under intervjuene ble det spesielt lagt vekt på å forstå, hva intervjupersonene tilla begreper og fenomener spesifisert i forskningsprosjektets problemstilling. I empirikapittelet er derfor respondentenes forståelse trukket frem og knyttet til begrepene og fenomenene. Sitater brukt i kapittel uttrykker dermed presist det respondentene sa. Fra intervjuene har jeg, så nøye som overhodet mulig, prøvd å tolke og danne en forståelse av respondentenes praksis, slik den muntlig ble uttrykt.

Alle respondenter som stilte opp i undersøkelsen ble lovet anonymitet. I presentasjon av det empiriske materiale, er dette søkt ivaretatt gjennom bruk av rollenavnene *student SU/FJU*, *lærer SU/FJU* eller *ekstern sensor SU/FJU*. Der det refereres til direkte sitater i *kapittelet*, angis derfor referansen med rollenavn utsagnet er knyttet til. For mer utfyllende om undersøkelsens forskningsdesign, henvises til kapittel 3.

4.1 Datamaterialet og utvalgte presentasjonskategorier



Figur 13: Forskningsdimensjoner

I analyse og presentasjon av funnene fra det empiriske materiale, var det problemstillingen tatt i betraktning, hensiktsmessig å rette søkelyset mot en **begreps-** og **fenomendimensjon**. Undersøkelsen har i tillegg vært sterkt styrt av en tredje dimensjon, **rolledimensjonen**, da fokusgruppene ble sammensatt ut fra respondentenes rolle (student, lærer,

sensor). Rolledimensjonen var ikke fokusert i forskningsstudien og blir da heller ikke presentert i denne oppgave.

Generelt har forståelse av de forskningsfokuserte begreper og fenomener vokst frem i en prosess. Først gjennom studie av tilgjengelig skriftlig materiale med beskrivelse av de utvalgte forskningsmiljøene, så under samtale og deltakelse i intervjuene, der jeg foruten lydbåndopptak også tok egne notater, og sist gjennom utforsking av studentoppgaver. Til å gjengi hva respondentene presist uttrykte muntlig under intervjuene, har jeg valgt ut spesielle sitater. Disse er spesifikt valgt for å illustrere hovedpoenger jeg anså som sentrale.

4.2 Forståelse og tolkning

I dokumentanalyse av studienes fagplaner og studieplaner (2005-06), samt Wittek & Havnnes (2005), var inntrykket at begge studiemiljøene arbeidet med mapper, i henhold til det jeg i metodekapittelet (jf.3) beskrev som **rik mappetekst**. Slik jeg forsto, brukte begge utdanningene digitale mapper, der bruk av formativ vurdering, og vurderingskriterier sto sentralt. Men én ting er å forstå og tolke **ved å lese** nedskrevne dokumenter, en annen å forstå og tolke **i samtale** ansikt til ansikt. Under intervjuene fant jeg det derfor hensiktsmessig, først å fokusere på respondentenes begrepsforståelse av mappe, formativ vurdering og kriterier, dernest på fokusgruppens forståelse av fenomenene reliabilitet, validitet og kvalitet. Forståelse og tolkning tilknyttet de ovenfor nevnte områdene, beskrives i kapitlene 4.2.1 og 4.2.2.

4.2.1 Begrepene

4.2.1.1 Mappe

I fagplan for SU (bachelorstudiet 2005) beskrives mappe med følgende:

En mappe kan beskrives som en systematisk samling av arbeider og andre former for dokumenter som til sammen gir et bilde av en students utvikling og læring innenfor et vist område. En mappe kan altså inneholde alt fra besvarelser av studiekraav til egne og andres vurderinger av studentens teoretiske og praktiske arbeider/ferdigheter. Selvvurdering knyttet til egne læringsmål inngår som en sentral faktor.

(Avd. for sykepleierutdanning 2005:8)

Mappe beskrives altså i denne fagplanen som en systematisk samling av dokumentasjon, som til sammen skal gi et bilde av en students læring og utvikling. Dokumentasjon kan være alt fra oppgavebesvarelser til tilbakemelding fra studierelevante personer. Mapper defineres som en egen arbeidsmetodikk, der studenten ikke kun har ansvar for egen læring, men også for å delta i medstudenters læringsprosess, gjennom medstudentrespons, diskusjoner, veiledning og undervisning av hverandre. I arbeidsmetodikken inngår bruk av ulike mappetyper, enten av fysisk eller elektronisk art. I fagplanen nevnes spesifikt arbeidsmappe, presentasjonsmappe og vurderingsmappe. I arbeidsmappen samler studenten alle sine arbeider. Her inngår notater og skriv som studenten anser viktige i sin læringsprosess. Presentasjonsmappen er den som registreres for å kunne gå opp til den summa-

tive vurderingen. I vurderingsmappen presenteres teksten som skal vurderes med eget vurderingsuttrykk i en summativ vurdering. For hvert studium ved SU er det utarbeidet en egen studieplan. I studieplanene er fagplanens intensjoner integrert. I studieplanen til deltid 1. studieår uttrykkes det for eksempel under avsnittet lærings- og vurderingsformer:

En mappe er en metafor for ulike typer arbeid og undervisningsformer, som inneholder og representerer ulike læringsaktiviteter i et studieforløp. Mappemetodikk er altså en studieform, hvor ulike mapper inngår. Mappedragene kan bestå av elektroniske dokumenter, video, ulike tekster, posters m.m.
(Avd. for Sykepleierutdanning 2005:13)

Mappe sidestilles altså i studieplanen for 1. studieår deltid 2005, med en metafor for **ulike typer arbeids og undervisningsformer**, der **ulike læringsaktiviteter** inngår. Mappemetodikk er en studieform, der studenter samler sine dokumenter i ulike mappetyper. Studieforken skal først og fremst legge til rette for læring, dernest for vurdering.

Mappe knyttet fokusgruppen ved **SU** slik jeg tolket den, til en **etablert arbeidsform**, der man jobbet med studieoppgaver **underveis** i en **prosess**.

Tanken har vært at det er ikke mapper i den forstand den mappa den mappa og den mappa, som står for seg sjøl, det er mer en tankegang, at man jobber med ting underveis (lærer SU).

Begrepet mapper var mer knyttet til en tankegang og ble i samhandling med studentene, mest brukt av lærerne som en **metafor**, der ordet fungerte som en slags overordnet samlebetegnelse, for der arbeids-/studieoppgavene skulle leveres inn. Slik jeg tolket SU fungerte den digitale mappen som en slags **container**, der ulike arbeids-/studieoppgaver ble lagret. På denne måten ble mappen et usynlig og nødvendig artefakt i lærings- og vurderingsprosessen, et redskap som støttet student og veileder/lærer i en pågående arbeidsprosess. Utvalgte mappeoppgaver ble senere brukt som underlag for student og sensor, i avsluttende summative vurdering.

Kvaliteten på studieoppgavene i (lærings)mappen, ble sikret ved at studenten, etter å ha fått tilbakemelding fra veileder eller medstudent, kunne jobbe dypere med oppgavens form og innhold, fram mot en ny og bedre tekst. Generelt skulle studieoppgavene gi innblikk i studentens faglige utvikling, vist i form av evne til **fagintegrering**, **strukturering** av oppgavene, skriftlig **formulering** og **anvendelse** av bestemte eksplisitte vurderingskriterier. Hensikten med mappen var å få studentene til å **jobbe jevnt** hele studieåret, der de over tid tilegnet seg og brukte fagkunnskapene de fikk gjennom studiet. Det var i stor grad semesterets faglige fokus som avgjorde hva som ble mappas innhold, der faglig fokus var definert i henhold til rammene skissert i plan for studiet. Kvaliteten ble søkt ivaretatt gjennom rammene i studieplanen. Fokusgruppens forståelse av mappe stemte slik godt overens med det som ble uttrykt i fag- og studieplan.

I studiehåndboka for **FJU** studieåret 2005/06 leser man om mappe under avsnittet organisering og arbeidsmåter, at studenter på FJU samler alle studieoppgaver gjort gjennom studieåret, i en mappe. Det legges til rette for at denne skal være elektronisk

(Avdeling for journalistikk, bibliotek- og informasjonsfag 2005:46). I undervisningsforløpet knyttet mappevurdering, til obligatoriske innleveringsoppgaver av praktisk eller teoretisk art. Oppgavene har innleveringsfrist og gjennomgås gruppevis, med obligatorisk fremmøteplikt for studentene (formativ vurdering). De fleste innleveringsoppgavene er produksjonsoppgaver, der studentene utfører journalistiske oppdrag. Å jobbe med mappevurdering understrekes i studieåndboka, der hensikten er at både studenter og lærere skal få et kontinuerlig inntrykk av studentens faglige utvikling. Mappa kan eventuelt inngå på ulike måter til slutteksamen. Generelt har studentenes oppfatning og erfaring med vurderingsformen hittil vært svært positiv (Sivertsen et al. 2004:31).

Fokusgruppen ved **FJU** uttrykte at man i 1995 knyttet forståelsen av mappe, til et sted der studenter kunne synliggjøre sine obligatoriske innleveringsoppgaver. Mappeorganisering bygget på digital katalogstruktur og dokumentarkivering, og var en nødvendighet, for å ha orden på besvarelsene/studieoppgavene som skulle vurderes. Mappens egenskap som digital samhandlingsartefakt (redskap) i organisering av besvarelsene, og muligheten for åpen og allmenn tilgang, var vesentlig for å ta dem i bruk.

Når vi begynte var mappe et sted å legge oppgavene/besvarelsene. En katalogstruktur hvor man kunne legge alle oppgavene, og som var tilgjengelig for studentene og alle lærerne. Tilgjengeligheten var jo veldig viktig (lærer FJU).

Det var altså et viktig poeng at det var mulig å **åpne** de digitale mapper for medstudenter og lærere, slik at alle fikk tilgang til studieoppgavene. På den måten bidro mappene til en viktig lærings- og samhandlingsarena. Mapper var i dag å anse som en motsetning til summativ vurdering, og en arbeidsform som gjorde det mulig for studenter og lærere å jobbe **proessorientert** med studieoppgaver. Slik representerte mappevurdering noe som fungerte bedre, både fra et student- og lærerperspektiv, og la til rette for jevn jobbing gjennom studieåret, i motsetning til summativ vurdering, som ble forbundet med hektisk arbeidsinnsats rett før eksamen. Mappevurdering var et godt hjelpemiddel i prosessen med synliggjøring og forbedring av arbeid underveis i studiet.

Jeg har alltid tenkt på mappe som en slags stor motsetning til en vanlig eksamensform og en form som gjør det mulig for studentene å hele tiden blir lagt under press for å jobbe (student FJU).

Studenter ved FJU leverte underveis i studiet, studieoppgaver som lærer godkjente/ikke godkjente. Gruppen mente, slik jeg tolket den, at det var greit å holde disse oppgavene utenfor en summativ vurdering. Generelt var mappene ikke knyttet til summativ vurdering, men det kunne gjennomføres hjemmeeksamen, der studentene fikk en ny oppgave og deler av mappa ble trukket inn. For noen i fokusgruppen fremsto mappe og mappevurdering som ensbetydende, mens andre mente at mappevurdering måtte være en vurdering av en større samling oppgaver på en gang, da knyttet til summativ vurdering. Veiledning av studieoppgaver underveis, falt da for noen i gruppen heller ikke naturlig inn under begrepet mappevurdering. Generelt ble mappen ved FJU, kun brukt som arbeidsredskap for organisering og synliggjøring av obligatoriske studieoppgaver underveis i studiet.

Men vi har det ikke som eksamensform. Det er en obligatorisk oppgave bare, hvor du går igjennom det du har gjort, og etterpå skal du trekke frem 2-3 eksempler du er veldig fornøyd med, og så skal du trekke frem to dårlige eksempler. Så skal du da analysere dem og så ta inn litt teori i forhold til det (lærer FJU).

Studentenes mandat besto i første omgang av å trekke frem gode og dårlige oppgaveeksempler og så analysere disse i lys av teori.

Sammenfatning mappe:

Slik jeg tolket respondentene knyttet begge utdanningene mapper til et sted, der studenter kunne samle og publisere studieoppgaver/innleveringsoppgaver, som de jobbet med underveis i studiet. Mapper var en nødvendig **container** (beholder), der studenter publiserte sine studieoppgaver/innleveringsoppgaver, og der lærere og medstudenter synliggjorde sin formative vurdering. Slik var containeren et **nødvendig artefakt** (medierende redskap) for oppgaveorganisering og oppgavesamhandling mellom studenter og lærere i en åpen formativ veilednings- og arbeidsprosess. Mapper som lærings- og samhandlingsarena trekkes frem av Otnes (2003). Spesielt **åpenhet** og mulighet for utvidet tilgang, var sentrale egenskaper, som ble trukket frem ved artefaktet. Egenskapene gjorde det mulig for studenter og lærere å samhandle underveis i prosessen. Den digitale containeren, oppfattet jeg som et usynlig og nødvendig artefakt i arbeid med formativ vurdering, der beholderen i seg selv ikke var så veldig interessant. Mer interessant var det å knytte mapper/mappevurdering til **en etablert arbeidsform**, der det ble lagt til rette for prosessorientert jobbing gjennom studieåret. Hensikten med mappen var først og fremst læring og å jobbe jevnt med studiet hele året. Mappens innhold ble styrt av faglig fokus og rammer i studieplanen, og var ikke gjenstand for formell vurdering underveis i prosessen. Arbeidsformen representerte slik en motsats til en summativ eksamensform. Fokusgruppens forståelse av mappe ivaretar dermed alle kjennetegn knyttet til en læringsmappe (jf. Dysthe & Engelsen 2003).

4.2.1.2 Formativ vurdering

I studieplanen til deltid 1. studieår ved **SU**, uttrykkes det under avsnittet lærings- og vurderingsformer, om formativ vurdering:

Skriftlig og muntlig respons fra veileder og medstudenter
(Avd for Sykepleierutdanning 2005:26)

Formativ vurdering defineres altså i denne studieplan, som en skriftlig og muntlig tilbakemelding fra studierelevante personer som medstudenter og veiledere. I planen nevnes at når mappemetodikk tas i bruk som studieform, er det vesentlig at studenten får respons underveis i studiet, både i praktiske og teoretiske fag. Responsen skal ta sikte på å veilede og støtte en student i læringsprosessen. Om prinsippene for respons skrives det i studieplanen: at den skal gis til oppgaver under utarbeiding, og skal fokusere både på oppgavens styrke og forbedringspotensial. Den gis så raskt som mulig og skal være tilgjengelig for

kullet. Prinsippene for respons på avd. SU er altså helt klart nedskrevet og spesifisert i studieplanen.

Formativ vurdering knyttet fokusgruppen ved **SU** til **en vurdering underveis** i arbeidsprosessen, der målet var **en vurdering å lære av**. En slik tankegang omfavnes av flere teoretikere (Barrett & Carney 2005, Wiliam 2000a)

Det er jo en vurdering underveis i prosessen. At du får en skisse - en innledning underveis. Du skal ikke være ferdig, du har tenkt å jobbe videre med produktet (lærer SU).

Underveis i prosessen var det et mål at studenter ikke leverte fiks ferdige produkter, men at de jobbet videre med oppgaver etter veiledning, der målet var å hjelpe studenten videre i arbeidet med å forbedre studieoppgavene. Studenten selv måtte klart spesifisere hva vedkommende ønsket tilbakemelding på, og formativ vurdering på noe studenten ikke hadde bedt om, var uhensiktsmessig. Dog var også fokusgruppens forståelse at det helt automatisk ble gitt respons, dersom man oppdaget noe som var helt feil i en medstudents studieoppgave. Formativ vurdering forekom både i form av respons fra veileder, respons fra medstudent og i noen grad i form av egenvurdering og egenrefleksjon.

Formativ vurdering ved SU kommer i ulike former: Som en respons fra medstudent og lærer, og i form av en egenvurdering av mappe og læringsforløp og refleksjoner og vurdering av egen læring og faglig fremgang (ekstern sensor SU).

Respons fra veileder/medstudent tok utgangspunkt i uferdig tekst og hadde til hensikt å motivere for faglig forbedring, ved å fokusere på både sterke og svake sider ved oppgaven.

De faglige momentene du har tatt med er et godt utgangspunkt for videre arbeid. Du kan styrke oppgaven din ved å få en mer helhet. Du har mange kapitler derfor kan teksten framstå litt oppstykket. Du har valgt vurderingskriteriene 3 og 5, disse krever at du trekker inn egen eller andres praksis. Kan du for eksempel lage en problemstilling knyttet til din pasient? Da blir det lettere å vise faglig innsikt, siden du diskuterer fagstoffet opp mot noe konkret. Få også inn litteraturhenvisninger (lærerrespons hentet fra en studentoppgave).

Prinsippene rundt formativ vurdering, var at studieoppgaver og formativ vurdering, ble **publisert synlig** for alle (lærere og studenter). Veiledningen skulle **tjene et formål** og ble gitt **underveis** i arbeidsprosessen, da gjerne når studentene sto litt fast og trengte hjelp til å komme seg videre. Forståelsen var også at jo mer studenten selv hadde jobbet med studieoppgavene sine, jo lettere var det å definere hva den formative vurdering skulle knyttes til. Dette medførte en mer presis og målrettet veiledning.

Det mer studenten har jobbet med spørsmålene, det bedre tilbakemelding får hun. Og det tror jeg og henger litt sammen med, innbilder jeg meg, at når de har jobbet mye med spørsmålene, så har de en erfaring sjølv, og da er de i farten, og da fungerer de godt i prosessen. Og da skjer det noe med oppgavene etterpå.
(Lærer SU)

Aktive studenter som i arbeid med studieoppgavene satte seg grundig inn i teorien og faget, bidro til erfaringslæring av stor verdi i tilknytning til veiledning. Inngående oppgavebearbeiding tilførte altså studentene stor grad av innsikt i og forståelse av fagstoffet. Utfordringene, slik jeg tolket fokusgruppen, besto for både lærer og student i å **treffe** med den

formative vurdering, og forståelse var at formativ vurdering traff best, når det var helt klart spesifisert hva studenten ønsket veiledning på.

Og jo mer konkrete spørsmålene er, jo bedre og mer konkret blir tilbakemeldingen fra oss lærere. (Lærer SU)

Jo mer konkret studenten var i sine veiledningsspørsmål, jo bedre ble tilbakemeldingen fra lærer/medstudent. Uklare og vide spørsmål førte ofte til dårlig veiledning, mens konkrete og klare ga god og konstruktiv tilbakemelding. Relevant og målrettet formativ vurdering forutsatte inngående kunnskap om det man skulle gi respons på, og bidro til at studenten fikk økt fagkunnskap, utviklet kritisk sans, evnet å være selektiv og bedret sin fagforståelse i et bredere perspektiv. Å blottlegge eventuelle svakheter i den formative vurderingsprosessen kunne oppleves skummelt (jf. ett av Carneys 6 dilemmaer i punkt 2.5.4), men etter hvert som man ble kjent, gikk dette lettere, og studentene ble både flinke til å gi og motta konstruktiv kritikk.

... men det blir jo lettere etter hvert. Du blir jo mer åpen, og mer kjent i hvert fall innad i gruppa. Man lærer jo ufattelig mye av det (student SU).

Å gi og få respons kunne av noen studenter oppfattes vanskelig og sårende, men bidro også sterkt til oppøving av kritisk fagsans og utvikling av evne til kommunikasjon. For ikke å såre en medstudent begynte studentene som regel veldig forsiktig, når de startet med å gi respons. Etter hvert som arbeidsformen ble kjent, forsto man at respons ikke var ment å såre, men heller en kritisk tilbakemelding, der målet var forbedring av studieoppgaven. Studentene sto uansett fritt til å nytte seg av tilbakemeldingen, slik var formativ vurdering ganske åpen.

Man starter veldig sånn forsiktig, og tenker oi sykepleier, jeg skal aldri såre noen. Og så etter hvert så får man jo mer kunnskap, og man blir mer kritisk, og man forstår på en måte både det at å få respons og gi respons, det er jo konstruktiv kritikk. Og det er ikke kritikk sånn at du har skrevet en dårlig oppgave, man kan velge om man vil endre på det selv. Det er en ganske åpen greie (student SU).

Med hensyn til å fange opp feil eller useriøs medstudentrespons, var gruppens forståelse at veileder spilte en viktig rolle her. Også studenten kunne trenes opp til selv å vurdere responsens seriøsitet. Generelt ble det ikke gitt formativ vurdering fra veileder, når det ble gitt medstudentrespons. Dog kunne veileder gi respons på medstudentresponsen, dersom det var nødvendig. Lærerrespons på medstudentresponser bidro til å forbedre studentens skriftlighet og evne til å være konkret og presis.

På **FJU** var forståelsen, at formativ vurdering ble knyttet til vurderinger/veiledninger underveis i studiet. Målet var at arbeid med studieoppgaver skulle foregå i en prosess, der formativ vurdering ble gitt fra start til slutt, og at jevnlig tilbakemelding skulle skape jevn progresjon.

Og jevn evaluering skaper også jevn progresjon, i motsetning til å gå og fylle et helt semester, og så kanskje ikke vite helt hva man driver med. Jeg har i alle fall følt at jeg har progresjon, når jeg

hele tiden får tilbakemelding og kan gå tilbake og rette på feila underveis. Det er den der underveistingen (student FJU).

Jevnlig tilbakemelding på studieoppgaver underveis i studiet, førte til at studenter fikk god oversikt over hva de drev med, og resulterte i forbedring av innholdet i studieoppgavene. Formativ vurdering var ment å stimulere til **kvalitetsheving** av studieoppgaver, og hadde **læring** og **prosess** som sine viktigste fokusområder. Studieoppgavene var ikke gjenstand for formell vurdering (karakter), kun godkjent/ikke godkjent, og både bilder og tekst måtte godkjennes. Godkjenningsprosedyren omfattet opp til flere veiledninger, der siktemålet var å hjelpe studenten videre i arbeid med oppgaven, en prosess som til tider kunne oppleves både god og frustrerende. Arbeid med oppgaven pågikk over flere uker, og inkluderte undervisning, utarbeiding av idéskisse, løsning av oppgave og felles gjennomgang. Gjennomgangen (formativ vurdering) var organisert og studentenes oppgave besto i å forberede gjennomgang av en medstudents oppgave.

Til gjennomgangen på onsdag 03.03 skal du forberede gjennomgang av en medstudent. Tirsdag 02.03 kl 0930 samles vi i P476 for å etablere en norm for gruppegjennomgangen. Etter lunsj ca kl 1230 starter gruppene arbeidet med å se gjennom leveringen og forberede gjennomgangen (utdrag fra en studieoppgave).

Forut for den organiserte gjennomgang ble det etablert en norm, slik at alle visste hvilke kriterier oppgaven skulle vurderes etter. Siden vurderte studentene først hverandre i gruppe, for så til slutt å ha en felles gjennomgang (formativ vurdering) med lærer.

Sammenfatning formativ vurdering:

Felles for begge utdanninger var at formativ vurdering (fv) ble knyttet til respons/veiledning underveis i studiet. Hensikten var en vurdering å lære av, og målet var å hjelpe studenten videre i forbedring av studieoppgavers innhold. Aktive studenter, som selv spesifiserte hva de ønsket veiledning på, bidro til å tilføre formativ vurdering kvalitet. Veiledning ble gitt i form av respons fra lærer, fra medstudenter eller som egenrefleksjon/egenvurdering, der grunnlaget for fv var publiserte studieoppgaver synlig og tilgjengelig for alle på kullet. Fv var uformell vurdering og i utgangspunktet isolert fra summativ vurdering. På FJU var det dog vanlig at lærere godkjente/ikke godkjente studieoppgaver, etter at formativ vurdering var gitt. På SU ble noen oppgavene valgt ut til formell summativ vurdering. Også de kjente vurderingskriteriene inngikk her som vurderingsgrunnlag i summativ vurdering.

Det som skilte utdanningene var om formativ vurdering tok utgangspunkt i et på **forhånd** kjent vurderingsgrunnlag eller ei. På SU var man av den oppfatning, at fv ble mest målrettet, konstruktiv og presis når studenter på **forhånd** kjente til vurderingskriteriene og ble formativt vurdert med utgangspunkt i disse. Erfaringen var at da traff fv best. I fotofag på FJU derimot ble formativ vurdering målrettet, konstruktiv og presis i diskusjon og samhandling i **etterkant** av oppgaveløsning, der lærere og studenter i fellesskap ble enig i vurderingskriterier aktuelle for studieoppgavene. Når studentene løste oppgaven valgte de selv egne kriterier, som siden ble tatt opp til felles diskusjon. Formativ vurdering ble på

denne måten knyttet til definering av aktuelle kriterier for oppgavene i etterkant av oppgaveløsning.

På tross av ulik praksis var begge utdanninger opptatt av å få til målrettet, konstruktiv og presis fv. Ulik praksis kom som en konsekvens av at den ene utdanningen brukte kjente vurderingskriterier som rettesnor for sin veiledning, mens den andre i egenskap av fagets egenart diskuterte seg frem til vurderingskriterier i etterkant av oppgaveløsning.

4.2.1.3 Kriterier

I studieplanen til deltid 1. studieår ved SU, uttrykkes det om kriterier under avsnittet lærings- og vurderingsformer, at hver studieoppgave gjør bruk av kjente vurderingskriterier (vedlegg 4). Hensikten med kjente vurderingskriterier er å gi studentene innsikt i hva som skal **vektlegges** i studieoppgaven, å uttrykke hva oppgaven skal **vurderes** etter, og tilkjennegi vurderingsgrunnlaget for den **summative** vurdering. Hensikten med de kjente vurderingskriteriene på SU defineres altså som tredelt (Avd. for Sykepleierutdanning, 2005:20). Fokusgruppen ved **SU** knyttet sin forståelse av kriterier til de 16 utvalgte vurderingskriterier vedtatt brukt ved utdanningen. Disse bidro til å fastsette rammer og føringer for **vektlegging** i studieoppgavene. Hensikten med vurderingskriteriene sto klart beskrevet i studieplanen.

Det står jo så vakkert der i planen. Altså rammer eller føringer for hva som på en måte skal vektlegges i akkurat den og den oppgaven (lærer SU).

Vurderingskriteriene ble ansett å være **rettesnoren**, og det som holdt ting på plass i mappevurdering. I tillegg til andre oppgaverammer var vurderingskriteriene det som spesifiserte hva som akkurat skulle vektlegges i den bestemte studieoppgave. Formativ vurdering ble gitt i forhold til vurderingskriteriene, valgt for studieoppgaven. Første semester utformet lærerne oppgavene, senere når studentene lagde sine egne studieoppgaver, falt det naturlig at de også selv valgte oppgavens vurderingskriterier. Også her antydte valg av vurderingskriterier vurderingsfokus for studieoppgaven. Vurderingskriteriene utgjorde således en essensiell del av rammene rundt den formative vurdering, og både lærer- og medstudentrespons ble gitt i henhold til disse. Kriteriene var vedtatt brukt ved utdanningen og var eksplisitt skrevet ned, slik at alle klart visste hva de uttrykte, og hvor de sto på trykk. De var med hensikt utformet såpass vidt, at det gikk an for studenter, selv å definere innholdet i dem. Likevel var vurderingskriteriene så klart utformet, at det var vanskelig å tillegge dem noe annet, enn det de tekstlig uttrykte. Så man på rammene for studieoppgaven, falt innhold i og forståelse av vurderingskriteriene naturlig på plass. Kriteriene fikk på denne måten både eksplisitt og implisitt karakter. Studentene tok utgangspunkt i de eksplisitt nedskrevne kriteriene, men sto og fritt til selv å definere og tolke (implisitt) hva de la i dem, da selvfølgelig med en faglig forankring. Generelt lå ansvaret for at kriteriene ble trukket inn i og forstått i studieoppgavene, hos veileder.

Ved **FJU** var forståelsen at kriterier var rettesnorer for det som skulle vektlegges i en oppgave, og de konkrete punktene man kunne følge. Kriteriene var utgangspunkt for handling.

Forutsetninger for hva som ligger til grunn for å gjøre noe, hva som helst. Altså rettesnorer (student FJU).

Kriteriepraksis varierte fra fag til fag. I fotofag **definerte studentene selv** i stor grad kriterier for oppgave(n)e. Oppgaveløsning var av mer subjektiv karakter, og forutsetningene var forskjellig fra oppgave til oppgave. I teoretiske fag derimot ble kriteriene (vedlegg 5) bestemt på forhånd ved studiestart og nedskrevet eksplisitt. Her begynte studenter og lærere alltid kurset, med å sette opp aktuelle vurderingskriterier, som da siden ble brukt til vurdering av studieoppgavene. Erfaringen var, at de fremdiskuterte vurderingskriteriene ofte ble forbausende like fra år til år.

Bildeteori som er et teoretisk fag. Vi begynner kurset med at studentene er med å sette kriteriene. Og det er forbausende hvor like kriteriene blir fra år til år. De legges da ut på nett og fungerer som rettesnor for oppgaveløsning, og er min og medstudenters rettesnor når vi retter og vurderer oppgavene (lærer FJU).

Generelt var studentmiljøet så lite, at det gikk greit å diskutere seg frem til en kriterieenighet. Ved oppgaveløsning i fotofagene var hensikten å få frem hver students egne kriteriepreferanser og tolkning, og målet var personlig utvikling og trening i å stole på valgene man gjorde. Det fantes ingen evige sannheter. Det var viktig å vite at kriterier fantes, og kjennskap til dem fikk komme enten ved at de var eksplisitt nedskrevet og publisert, eller man snakket om dem gang på gang. For fotofagene fungerte det altså bra å forholde seg til ikke nedskrevne kriterier (implisitte), mens det for teoretiske fag var hensiktsmessig å ha dem eksplisitt nedskrevet. Kriteriebegrepet ute i bransjen var generelt omstridt. Det ville variere fra redaksjon til redaksjon, alt etter hva som ble ansett som viktig. Siden studiet i stor grad forholdt seg til bransjen, var det naturlig å ha den som referanse, også i bruk av kriterier.

Sammenfatning kriterier:

På begge utdanninger definerte man kriterier som rettesnorer/føringer for vektlegging i en studieoppgave, og de faste holdepunkter for en oppgaves faglige fokus. I kriteriepraksis var de to utdanningene delte. På SU ble formativ vurdering knyttet til 16 kjente eksplisitt nedskrevne og vedtatte vurderingskriterier. Selv om kriteriene var eksplisitt nedskrevet og fantes på trykk i studieplanen/fagplanen, var de likevel såpass vidt formulert, at det fantes rom for studentens egen tolkning, innen en faglig avgrensning. Vurderingskriterier av en slik utforming understrekes av Wittek og Dysthe (2003). På FJU ble formativ vurdering knyttet til bruk av både eksplisitt nedskrevne og ikke nedskrevne (implisitte) vurderingskriterier. I praktiske fag var det naturlig å vurdere oppgaver etter ikke nedskrevne kriterier, mens det i teoretiske fag var hensiktsmessig å vurdere etter nedskrevne. I praktiske studieoppgaver

valgte studenter selv kriterier, der et viktig poeng var å knytte de valgte føringene til det rådende syn i bransjen. I diskusjoner/veiledningsøkter (formativ vurdering) ble lærere og studenter siden enig i kriterienes innhold og faglige substans. Ulik kriteriepraksis ved FJU, var hensiktsmessig i lys av fagenes egenart.

4.2.2 Fenomenene

Fenomenene reliabilitet og validitet forbindes i utgangspunktet med målinger/undersøkelser av ulike slag. **Reliabilitet** knyttes gjerne til **hvordan** (metoden) man måler, og i hvor stor grad metoden er til å stole på (pålitelighet), mens fenomenet **validitet** (gyldighet) knyttes til **hva** som måles (innhold), og om målingen måler det som er tiltenkt målt (Patel & Davidson 1995). Fenomenet **kvalitet** knyttes gjerne til en vares beskaffenhet og verdi, der verdien settes etter vedtatte normer for vedkommende vare. Kvalitet uttrykker således et mål for hvor god en vare er.

I undersøkelsen var fokus rettet mot å forstå og tolke hva respondenter, som arbeidet innen en rik mappekontekst, tilla de ovenfor nevnte fenomenene. Å undersøke fenomenene knyttet til bruk av kriterier og formativ vurdering, var i seg selv ikke så lett, da de fort ble oppfattet som både abstrakte og lite målbare i en slik sammenheng. Under intervjuene var målet å **ikke** legge personlige føringar for hvordan fenomenene skulle forstås. Det jeg heller ønsket var å la respondentene selv fortelle, hvordan de forsto fenomenene. For selv å tilegne meg en overordnet fenomenforståelse før intervjuene, landet jeg på at reliabilitet måttet knyttes til 'noe å stole på', validitet til 'noe som var gyldig og målrettet' og kvalitet til 'noe av verdi'. I intervjuguiden (vedlegg 1) stilte jeg for eksempel spørsmålene:

- Er det relevant å knytte formativ vurdering til fraser som "En vurderingsform vi kan stole på er riktig" og "En vurderingsform som gir en målrettet vurdering"?
- Hvis du skal forklare frasen "Et kriterium man kan stole på og et kriterium som har gyldighet" hva vil du trekke frem da?
- Hvis du skal si noe generelt om kvalitet i forbindelse med mapper hva vil du trekke frem da?

Erfaringen under intervjuene var at det var både vanskelig og uvant for respondentene umiddelbart og direkte å definere og knytte fenomenene reliabilitet og validitet til det de jobbet med. Også å skille fenomenene var vanskelig, slik jeg tolket utsagn og reaksjoner. En forklaring jeg ga meg selv var at å fokusere og skille disse fenomenene krevde inngående diskusjon og refleksjon. Det faktum at det er både krevende og vanskelig å diskutere slike abstrakte fenomener, og knytte dem til aktivitet i en studiesammenheng, var en annen. En tredje var at mappevurdering nå var prøvd ut ved utdanningene gjennom flere år (henholdsvis 1999 ved SU og 1995 ved FJU), og automatisk hadde tilegnet seg både reliabilitet og validitet. Det var derfor ikke påkrevd å ta opp slike fenomener til diskusjon i

studiet på nåværende tidspunkt. Som en konsekvens av vanskeligheten med å skille validitet og reliabilitet, har jeg derfor valgt å presentere en tolkning av respondentenes fenomenforståelse, sammenhengende og ikke atskilt. Fenomenet kvalitet anser jeg å omfattes av det som uttrykkes om fenomenene reliabilitet og validitet.

4.2.2.1 Formativ vurdering i lys av reliabilitet og validitet

Pålitelig formativ vurdering (fv) avhang til syvende og sist av hvem man fikk respons fra, og var **en vurderingssak** mellom den som fikk og den som ga respons, slik jeg tolket forståelsen ved **SU**. Alle responser ble vurdert for **relevans**, og knyttet til om **responsen traff mål**, og kunne brukes i videre bearbeiding av studieoppgaven. Formativ vurdering ble altså ansett å være pålitelig, når den var **måltrettet, relevant og kunne brukes i arbeidsprosessen videre**. Generelt anså studenter respons fra veileder, som mer pålitelig enn om den ble gitt av en medstudent. Veileder hadde i utgangspunktet større veileder- og fagkompetanse, og ble derfor ansett mer kompetent til å gi pålitelig formativ vurdering. Men også medstudenters evne til å se ting, som ikke veileder så, var viktig. Studentene bidro ofte med detaljkunnskaper tilknyttet studieoppgavene, mens veileder så oppgavene i et mer helhetlig og bredere fagperspektiv.

... hvis vi ser på fysiologi så er vi veldig på detaljkunnskapen da. Og er det noe studentene kan, så er det detaljkunnskapene (lærer SU).

Et fag som fysiologi var for eksempel rettet mot detaljer, der mange studenter hadde sin styrke. En ideell fv kombinerte både et smalt og bredt fagspekter.

Å gi og få respons hang unektelig sammen, og begge former bidro til både kunnskaps- og erfaringslæring. Men studenter lærte antakelig mest ved *å gi* respons, ikke *å få*.

Noe av det studenten kanskje lærer mest ved, er å gi responsen. Ikke å få den. Det vil jeg tro (Lærer SU).

Grunnen til det var at i prosessen med utforming av en respons, måtte studenten selv skaffe seg innsikt i faget, i tillegg til å sette seg inn i hva medstudenten, som skulle ha responsen, hadde tenkt faglig. Denne prosessen førte til mye læring, hvilket igjen forsterket vurderingens pålitelighet og gyldighet. Gjennom *å få* tilbakemeldinger tilegnet studenter seg fagkunnskap og erfaring i hva som var aktuelt å vektlegge i studieoppgaven, og opplevde personlig hvordan det var å få konstruktiv kritisk respons. Gjennom denne erfaring tok de lærdom av hvordan en respons eventuelt kunne oppfattes av andre, og hva de selv anså som god og dårlig respons. Disse erfaringene var viktige faktorer i utvikling av gyldig og pålitelig formativ vurdering. Fv var pålitelig og gyldig når den ga studenten såpass mye veiledning, at vedkommende student kom seg videre i arbeidet med studieoppgaven. Pålitelig og gyldig fv ble således knyttet til **faglige innspill** med anvisning av mulig oppgaveutvikling, som igjen ble oppfattet som **hensiktsmessig** og **meningsfylt**, av studenten som fikk responsen.

Det er jo mer om jeg har gitt studenten den pushen til å kunne jobbe videre. Altså hvis du da kan gjøre noe med oppgaven som er meningsfylt for henne. Hvis hun lærer noe av å gjøre noe med den. Så er den suksessfull (lærer SU).

Vurdering var vellykket når veiledning ble opplevd som meningsfull, førte til at studenten lærte noe og kom seg videre i arbeidet med å forbedre studieoppgaven. Å pushe, hjelpe og motivere studenten videre i prosessen med faglig forbedring av studieoppgavene, gjennom hensiktsmessig og meningsfylt veiledning, med det formål å lære, kunne betegnes som pålitelig og gyldig fv.

Med forankring i studieplanen var fv gyldig, når den rettet seg mot **studieårets fokus**, og tilstrebet å **veilede studenten i å nå de spesifiserte fagmål** underveis i studiet. Det kunne gjøres på ulike måter, da studieoppgavene ikke hadde faste fasitsvar. Om alle oppgaver ble løst likt, hadde noe sviktet i den formative vurderingen. Ulike løsningsforslag vitnet om selvstendige aktive studenter, som jobbet individuelt og ikke var avhengig av noen mal.

Det finnes mange veier å gå i forhold til det som er fokus. Det er mange måter å se dette på. Det er det som er hensikten også (lærer SU).

Styrken ved studieoppgavene var at det fantes mange mulige veier for studentene å gå, i forhold til løsning av disse. Gyldig (valid) fv, knyttet også fokusgruppen ved SU til de 16 eksplisitt nedskrevne vurderingskriterier. Det var disse som i stor grad definerte rammene for den formative vurderingen. Kriterievalg kunne variere, i den forstand at studentene selv aktivt valgte kriterier for oppgaven sin. Gyldig og pålitelig fv ble søkt ivaretatt i forhold til de valgte vurderingskriteriene. Hadde studenten for eksempel valgt vurderingskriteriet 'faglig innsikt', ble kriteriet ikke kun gyldig formativt vurdert, ut fra hvordan det teoretisk ble anvendt i oppgaven, men også i forhold til praktisk anvendelse.

I fotofagene på **FJU** var pålitelig og gyldig fv, fenomener som hele tiden måtte ses i sammenheng med studiets/fagets utvikling i bransjen. Pålitelig og gyldig formativ vurdering var **ikke konstante tilstander**, men hele tiden i endring, alt etter hvordan bransjen ute utviklet seg, og fant dermed sin gyldighet i det som var akseptert der. Faglærere som til enhver tid var **kompetente** og oppdatert i forhold til det som var akseptert ute, ble på denne måten en viktig faktor i utvikling av pålitelig og gyldig fv på studiet. Det var også viktig at lærerne var **tilgjengelige** og kunne gi studentene veiledning underveis i arbeidsprosessen.

Det er to deler. Den ene er den ekspertisen som ligger hos lærerne, som skal gi tilbakemeldinger til hver oppgave. I tillegg at de er tilgjengelige for spørsmål underveis dersom noen trenger det (lærer FJU).

I tillegg til fagkompetente og tilgjengelige lærere, spilte **klassemiljøet** en vesentlig faktor i utvikling av fv. Et lite, aktivt og åpent klassemiljø med meget kompetente studenter, bidro til pålitelig og gyldig vurdering.

Jeg synes jo mange ganger at studenter er de beste lærerne overfor hverandre. Hvis du ikke klarer å etablere et godt læringsmiljø blant studentene så er det fryktelig vanskelig å få til den formative læringen (lærer FJU).

Studentene var mange ganger de beste lærerne overfor hverandre, og uten et velfungerende klassemiljø, var det vanskelig å etablere et åpent klima for diskusjon, noe som i bunn og grunn la premissene for kriterieforståelse, og pålitelig og gyldig fv. Den **åpne mappestrukturen**, var et nødvendig artefakt til felles fremvisning av fotooppgaver. Den digitale synliggjøring muliggjorde en samlet oppgavediskusjon, og bidro til å utvikle læringsmiljøets forståelse av pålitelighet og gyldighet. **Faglige diskusjoner** (formativ vurdering), der studenter og lærere kom frem til felles løsningsforslag, forsterket en forståelse av pålitelighet og gyldighet, noe som da til enhver tid var under utvikling.

Og sånn sett tror jeg, som fotograf da, så er man i stadig utvikling. Man kan ikke stagnere, og man må på en måte søke etter sitt eget bildeuttrykk. Og sånn sett så synes jeg formativ vurdering kan være både positivt og kvalitetsutviklende (lærer FJU).

På FJU var målet at studenten skulle finne sitt eget fotouttrykk, og samtidig ikke stagnere i egen fotoutvikling. I den prosessen var fv et godt hjelpemiddel, og bidro til å skape konsensus om hva som var pålitelig og gyldig. Også **yrkeslivet** utgjorde en viktig faktor i den formative vurderingen. Bransjens veiledning og tilbakemeldinger til studenter både i og etter praksisperioder, bidro i stor grad til å sette kriterier for hva som var pålitelig og gyldig ute.

Så det er en 3. faktor altså, du har lærer- og studentmiljøet og så den 3. er yrkeslivet, som er veldig viktig i den formative. ...

Enkelt sagt så prøver vi å måle gyldigheten i forhold til hva bransjen leverer av varer selv, fordi vi i hovedsak utdanner folk til i all hovedsak avisverden, og det har gjort at gyldigheten vår har ligget der (lærer FJU).

Pålitelighet og gyldighet var imidlertid ikke entydige størrelser innen bransjen heller.

Altså Bergens Tidendes fotografer er veldig forskjellig fra fotografer i VG (lærer FJU).

Studerte man for eksempel hvordan ulike fotografer i ulike medier dekket samme sak gjennom fotomontasjer/reportasjer og lignende, så var det stor forskjell på hvordan saken ble vinklet og presentert. Dette **tvisyntet** var det viktig å holde fokus på i fv ved FJU.

Og fotografi i seg selv er subjektivt. Hvis du ser på gyldighet i bransjen i 1995, hva var det som var et godt portrett, hva var en god reportasje, hva var et godt nyhetsbilde. Det forandrer seg hele tiden (lærer FJU).

Profesjonen i seg selv var svært foranderlig, slik at fotografering som ble ansett å være gyldig i 1995, ikke var det i dag. Pålitelighet og gyldighet var derfor relative fenomener, og i stadig endring. Også veileder og veiledningssituasjonen var subjektiv og førte til subjektiv refleksjon. **Subjektiviteten** var dermed en uunngåelig og viktig faktor i utvikling av pålitelig og gyldig fv. Men den subjektive gyldigheten forandret seg over tid, hvilket var helt naturlig for profesjonen.

Det er ikke noen direkte gyldighet sånn sett. Den forandrer seg over tid og (lærer FJU).

Studerte man bransjens utvikling opp gjennom årene, ville man se at det stadig hadde skjedd endringer, og så man på studieoppgavene fantes det alltid flere løsningsforslag.

... det er klart når det er ganske stor enighet blant studenter og lærer i hva som er bra og dårlig her, da begynner man å få en viss form for gyldighet etter min mening for da prøver du på en måte å diskutere deg frem til en slags gyldighet (lærer FJU).

De formative fagdiskusjonene førte som regel til at det oppsto stor grad av enighet mellom lærere og studenter, om hva som var bra og hva som var dårlig i en studieoppgave. Slik bidro **enigheten** til å definere hva som var pålitelig og gyldig oppgaveløsning.

I og med at vi driver med foto, så er det veldig styrt av den situasjon du er i, når du tar bilder (student FJU).

Profesjonens egenart (fotojournalist), gjorde at man i løsning av praktiske studieoppgaver ble styrt av her og nå situasjoner. Man kunne på forhånd legge rammer og føringer for hvordan et fotografi skulle tas, men fotografering ble til syvende og sist likevel styrt av situasjonen, slik den utartet seg, når bildet ble tatt. Det var derfor naturlig at fv fant sin pålitelighet og gyldighet i diskusjoner i **etterkant** av fotograferingen.

Sammenfatning formativ vurdering i lys av reliabilitet og validitet:

Formativ vurdering involverte både lærere og studenter ved begge utdanninger og inkluderte å gi og få respons. Responsgiving bidro på denne måten til pålitelig og gyldig veiledning. Når veiledning ble oppfattet som målrettet, faglig relevant, motiverende, hensiktsmessig og kunne brukes videre i prosessen med forbedring av studieoppgavene var den pålitelig og gyldig. Forutsetningene for å få til pålitelig og gyldig fv var den åpne mappestruktur, som la til rette for samhandling og utvikling av læringsfellesskapets forståelse for pålitelighet og gyldighet. Fagoppdaterte kompetente og tilgjengelige lærere og faglig aktive studenter var viktige faktorer. I veiledningsprosessen støttet begge utdanninger seg på vurderingskriterier, knyttet til studieoppgavenes faglige innhold.

Det som skilte utdanningene var kriteriepraksis. På SU ble fv pålitelig og gyldig blant annet gjennom bruk av 16 vedtatte og eksplisitt nedskrevne vurderingskriterier. Både lærere og studenter ga veiledning i forhold til disse. I fotofagene på FJU rådet et mer subjektivt syn på bruk av kriterier. Her var det i utgangspunktet ingen fastlagte kriterier for studieoppgavene, og studentene bestemte selv i første omgang oppgavens rettesnor. Pålitelig og gyldig fv utviklet seg siden i fagrelaterte diskusjoner, der det oppsto en form for enighet om hva som var bra/dårlig. Fokusgruppen ved FJU var spesielt opptatt av å understreke det subjektive momentet i pålitelig og gyldig formativ vurdering, både hva gjaldt veileder, veiledning og refleksjon. Fenomenene var relative og forandret seg stadig i bransjen FJU assosiert seg med.

4.2.2.2 Kriterier i lys av reliabilitet og validitet

Pålitelighet og gyldighet var uvante ord å knytte til vurderingskriteriene som ble brukt ved **SU**.

Det er så rare ord du bruker. Mener du om de er anvendbare eller gjenkjennbare? (lærer SU)

Slik sitatet viser foreslo gruppen å bruke ord som anvendbare eller gjenkjennbare for å identifisere et kriteriums pålitelighet/gyldighet. Ansvar for å ivareta vurderingskriterienes pålitelighet og gyldighet, var tillagt andre enn studentene, og for så vidt også lærerne, og var mer plassert på ledelsesnivå.

Om disse her er valide eller gjeldende eller hva det var for noe, det må jeg si er opp til veilederne og de som setter opp den her planen, ikke sant (student SU).

De 16 vurderingskriteriene som ble brukt i studieoppgaver ved SU, fant således sin pålitelighet og gyldighet i kraft av at de var **vedtatt** anvendt ved studiet. Vurderingskriteriene hadde avdelingen med faglig forankring, utviklet i samarbeid med tidligere studentkull, og fant på det grunnlag sin pålitelighet og gyldighet. Pålitelighet og gyldighet ble ytterligere forsterket i formativ vurdering, der veileder via responsgivning sikret seg at studentene **forsto, tolket og konkret anvendte** vurderingskriteriene i løsning av studieoppgavene.

Og hvis jeg synes det virker litt vakt, med akkurat hva han mener med faglig innsikt, så går jeg og gir respons til det, og ber ham utdype ytterligere. Her kommer det ikke helt klart frem, hva du tenker med faglig innsikt. Kunne du ha konkretisert det ytterligere? (lærer SU).

Om veileder syntes det ikke kom helt frem i en oppgave hva studenten tilla vurderingskriteriene, ble studenten bedt om å klargjøre dette. Helt konkret kunne det uttrykkes ved at veileder i sin respons ba studenten tolke og klarere definere hvordan vurderingskriteriene skulle forstås og anvendes i oppgaven. Veileders jobb var altså å rette studentenes oppmerksomhet mot hvordan vurderingskriterienes pålitelighet og gyldighet kunne forstøkes ytterligere i oppgaveløsning.

Og det er en viktig jobb for meg som veileder å gi tilbakemelding til studentene på – har dere gjort dere noen tanker om, hvordan dere tenker å anvende akkurat dette vurderingskriterium? (lærer SU)

At studenter reflekterte over hvordan vurderingskriteriene ble tolket og anvendt, bidro således til mer pålitelighet og gyldighet. Om det å knytte kvalitet til et vurderingskriterium, var forståelsen ved SU, at kvalitet oppsto som følge av, at man ved oppgaveløsning hadde mulighet for å **gjøre aktive valg** og **kombinere** kjente vurderingskriterier. Det var nesten umulig å skrive dårlige oppgaver da. Med tanke på avsluttende summativ vurdering, var det et stort poeng at vurderingskriteriene var **kjent** for studentene. Kjennskap til og erfaring med vurderingskriteriene i oppgaveløsning gjennom studieåret, førte til kvalitet både i studiet og i formell summativ vurdering.

Ved **FJU** ble det brukt både synlig nedskrevne og ikke nedskrevne kriterier. Ved vurdering av oppgaveløsning i teoretiske fag falt det naturligere å bruke vurderingskriterier, som på forhånd var fremdiskutert og eksplisitt nedskrevet. I fotofag derimot var det mer naturlig at studenter i utgangspunktet selv valgte kriterier for studieoppgaven(e), for siden i felleskap å diskutere seg frem til hva som var gode eller dårlig kriterier. Enighet om pålitelige og gyldige rettesnorer for fotooppgavene, fremkom altså her i etterkant av oppgaveløsning.

Kriterier fremstår mer for meg som grep, som jeg vet jeg kan bruke når jeg er ute og fotograferer. Noen ganger fungerer det og noen ganger ikke. Disse grepene er i sum med på å skape kvalitet. Men det er ikke noen absolutte kriterier (student FJU).

Kriterier i fotooppgaver ble knyttet til praktisk handling ved oppgaveløsning, der studieoppgavene var preget av prøving og feiling i uforutsigbare situasjoner. Det egnet seg derfor ikke å fastsette kriterier på forhånd. Ved felles gjennomgang av studieoppgaven i etterkant, ble kriterienes pålitelighet og gyldighet diskutert og fremmet i lys av konsensus. Det faktum at kriteriene eksisterte i oppgavene, enten de nå var av eksplisitt eller implisitt art, gjorde også at de tilførte studieoppgavene en form for kvalitet.

At de eksisterer i det hele tatt. At de finnes. De er kanskje ikke alltid like uttalt, og de er kanskje mange, men de finnes og det tror jeg gir kvalitet (student FJU).

Kriteriene var ikke alltid like uttalt, men uten kriterier ble heller ikke oppgavene kvalitativt gode. Kriterier, eksplisitte eller implisitte, var de pålitelige og gyldige rettesnorer, og dem man begrunnet sin vurdering ut fra. Noen kriterier var det ingen tvil knyttet til (kildekritikk), mens andre var mye mer subjektive og relative (originalitet). Kriteriene ble i utgangspunktet kun knagger man kunne orientere seg etter. Det var gjennom diskusjonen de fikk sin pålitelighet og gyldighet. Denne tankegangen støttes av Dysthe et al. (2005), Rust et al. (2003) og Sadler (1983).

Det å arbeide journalistisk er et kriterium. Det å arbeide etisk holdbart er et annet. Det er veldig generelle ting, som bare blir knagger, som du kan orientere deg etter (lærer FJU).

Oppøving av subjektiv kritisk tenkning knyttet til bruk av kriterier, var en styrke for hele læringsmiljøet.

Sammenfatning kriterier i lys av reliabilitet og validitet:

Ved begge utdanningene var forståelsen, at vurderingskriteriene var de pålitelige og gyldige rettesnorer, for vektlegging i en oppgave. De to utdanningene hadde imidlertid ulik forståelse av, på hvilket grunnlag kriterier fant sin pålitelighet og gyldighet. Ved **SU** hentet vurderingskriteriene sin pålitelighet og gyldighet, i det faktum at de gjennom faglig forankring og samarbeid mellom lærere og tidligere studentkull, var vedtatt ved avdelingen. Pålitelighet og gyldighet til de 16 vedtatte vurderingskriteriene ble ytterligere forsterket i formativ vurdering, når veileder/medstudent ba studenter utdype sin egen faglige forståelse, tolkning og anvendelse av vurderingskriteriene. I fotofagene ved **FJU** hentet vurderingskriterier sin pålitelighet og gyldighet i fagdiskusjonen mellom lærere og studenter. Utvikling av pålitelige og gyldige kriterier ble knyttet til erfaringslæring fremvist og debattert i plenumsdiskusjoner, der subjektive kriterievalg var utgangspunktet. I diskusjon og veiledning oppsto så en form for enighet om kriterienes pålitelighet og gyldighet.

4.2.2.3 Mapper i lys av kvalitet

Ordet kvalitetsstempel knyttet fokusgruppen ved SU, til det å sette et endelig stempel på et produkt.

Jeg er litt enig i at summativ vurdering blir kvalitetsstempelen for det hele. Det er der vi skal vise hva vi har lært, det er der som det skal komme frem, om studiemåten har fungert på en måte. Det er det endelige stempelen (student SU)

Kvalitetsstempelen var den avsluttende summative eksamen/vurdering, der studentene helhetlig viste hva de hadde lært på studiet. Stempelen symboliserte således en endelig karakterfast vurdering av alt studentene hadde tilegnet seg av fagkunnskap og erfaring, gjennom den kontinuerlige læringsprosess i studiet, og ble uttrykt i en karakter på vitnemålet. Ordet kvalitet derimot knyttet SU til en arbeidsmetodikk, hvor veiledning og responsgiving inngikk som viktige prosessfaktorer, og en saksorientert formativ vurdering, der veiledning underveis i arbeidsprosessen ble gitt i henhold til valgte vurderingskriterier. Studenter sto i utgangspunktet fritt til å velge hvor mye de ville bruke av den formative vurdering, og kunne bruke hele, deler eller ingenting. Også kjente vurderingskriterier var i stor grad med på å sikre kvalitet, og avspeilte gjennom praktisk anvendelse studentens faglige og teoretiske innsikt.

Det er jo også det at kriteriene avspeiler jo den faglige innsikten studenten viser, altså vi gjør oss jo opp en mening om de har en ide om hva faget dreier seg om, og hvordan de ser sammenhenger, og hvordan de antyder i teksten at de kan anvende den teorien de fremstiller (lærer SU).

Spesielt prosessen og friheten til å gjøre individuelle valg ble trukket frem i forståelse av hva som gir mappearbeid kvalitet.

At det er tillatt å jobbe i en prosess. Sånn som jeg har vært vant til som student selv, så har det vært veldig produktfokusering. Her ligger det en mulighet til å jobbe individuelt (lærer SU).

Kjempeenig. I hvert fall som student. At det ikke er fastsatt. Du kan liksom, du har på en måte en ganske stor rekkevidde selv, som student hva du vil gjøre med mappa. Det er jo utrolig deilig. Å vite at du har frihet og frustrerende samtidig (student SU).

Å kunne jobbe med oppgaver i en prosess, og ikke måtte gjøre dem ferdig med en gang, var kvalitet. Det fjernet fokus fra produkt, og rettet oppmerksomhet mot kvalitativ fagforståelse i utvikling av studieoppgavene. Dessuten ga det stor grad av frihet, noe som dog også kunne oppleves frustrerende. Studenter flest var vant med at studiearbeider gjerne ble honorert med en stor R. Men å flytte fokus fra karakterer til utviklende arbeidsprosess, bidro til å gi mapper kvalitet. Ansvar for eget studie og egen læring kunne være svært skremmende. Til å forebygge dette var **god informasjonsformidling** en av de viktigste faktorer. Alle vurderinger gitt underveis i læreprosessen, var åpne og tilgjengelige for alle på kullet. Åpenhet gjorde at de som var interessert, kunne lese tilbakemeldingene om de ønsket. Dette også tilførte arbeid med mapper kvalitet.

Det å kunne se på hvilken tilbakemelding andre fikk, er verdifullt. Det er jo alltid noe som går igjen. Det skrives masse responser og det er mye felles problematikk (lærer SU).

Ved å lese tilbakemeldinger andre hadde fått, erfarte studenter at det i responsene ble uttrykt mye felles fagproblematikk. Åpenhet og tilgjengelighet la på denne måten til rette for felles kunnskaps- og erfaringslæring.

Læringsmappene ved **FJU** var heller ikke gjenstand for summativ vurdering. Før studieoppgavene ble lagt i mappen, måtte de imidlertid være godkjent av lærer. Spesielt åpenhet og grundig tilbakemelding på studieoppgavene ble fremhevet av FJU som kvalitet ved mapper.

Men og at det er åpent og at alle kan gå inn og se. At det ikke bare er en karakter du får sjøl, som du logger deg inn på, og ingen annen får vite om. Det skaper en sånn åpenhet rundt det (student FJU).

Et åpent system gjorde at følelsen av hemmelighet var fraværende. Det åpne digitale rommet bidro til et debatterende og aktivt studiemiljø, der studentene i tillegg til formelle tilbakemeldinger (formativ vurdering), opplevde å få uformelle spontane tilbakemeldinger fra medstudenter. Mulighet for å kunne studere medstudenters studiearbeid underveis i studiet, og en kontinuerlig og jevn formativ vurdering, skapte jevn læringsprogresjon hvilket ble fremhevet som kvalitet. Formativ vurdering åpnet for mulighetene til å rette feil underveis, og ga studentene en følelse av å ha kontinuerlig fremgang i studiet.

... jevn evaluering skaper også jevn progresjon, i motsetning til å gå og fylle et helt semester og så kanskje ikke vite helt hva man driver med (student FJU).

I motsetning til vanlig summativ vurdering, som ofte var forbundet med hektisk og lite målstyrt studiearbeid rett før eksamen, ga jevn tilbakemelding (formativ vurdering) jevn og målrettet studieprogresjon, uten stress før summativ eksamen. Et nært og åpent læringsmiljø med engasjerte og aktive lærere, som klarte å motivere studenter til prøving og feiling, bidro også til kvalitet i arbeid med mappen.

Sammenfatning mapper i lys av kvalitet:

Mappene var i utgangspunktet på ingen av utdanningene tilknyttet summativ vurdering. På SU valgte imidlertid studenten selv, eller i samarbeid med lærer, ut mappeelementer de ønsket å bli summativt vurdert på. I summativ vurdering inngikk også bruk av 16 kjente vurderingskriteriene. På FJU måtte studieoppgavene være godkjent før de ble lagt i mappene. Som mappekvalitet fremhevet begge utdanningene:

1. Åpenhet. Alle har tilgang til og kan lese alt som lagres i mappene.
2. Prosessjobbing med fokus på læring og utvikling og ikke karakterer.
3. Kontinuerlig formativ veiledning fra lærer eller medstudent.
4. Friheten til å gjøre individuelle valg i et nært og åpent læringsmiljø.
5. Bruk av kriterier.

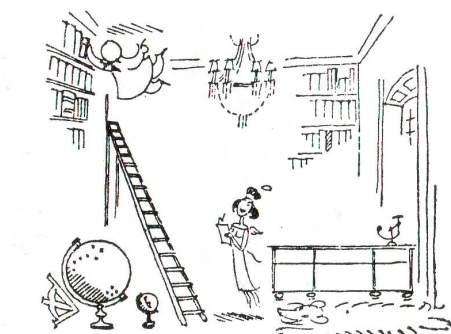
Det som skilte de to utdanningene var ulik praksis i forhold til kriterier. På SU ble knytningen mellom formativ vurdering og bruk av kjente vurderingskriterier, fremhevet som et kvalitetselement. Kjente vurderingskriterier fremelsket en saksorientert og kvalitativ formativ vurdering. På FJU var det i fotofagene mer naturlig at studentene i utgangspunktet selv valgte oppgavekriterier, for siden gjennom diskusjon å fremme konsensus om hva som var

pålitelig og gyldig i forhold til profesjonen. Her var altså den etterfølgende kriteriediskusjon et kvalitetselement.

4.3 Oppsummering

I presentasjon av empirien har det vært viktig for meg å holde fokus på forskningsområdene tilknyttet prosjektets problemstilling, derav oppbyggingen i kapittel 4. Tolkningene som fremkommer i kapittelet, og som representerer en subjektiv forståelse, tar generelt utgangspunkt i 3 datakilder; intervjuene med SU og FJU, analyse av tilgjengelig og aktuell skriftlig dokumentasjon tilknyttet de to studiemiljøene, og studentoppgaver med tilhørende tilbakemelding. Hva gjelder sistnevnte hadde det vært ønskelig med innsyn i flere, men akkurat det viste seg vanskelig.

I kapittelet 5 tas analysefunnene opp til nærmere drøfting i lys av generell teori presentert i kapittel 2, og forsøkes spesielt innpasset i sertifiseringsrammeverket utarbeidet av Herman & Winters.



*At vide hvad,
man ikke véd
er dog en slags
alvidenhed.*

(Hein 1968:28)

Figur 14: At vide hvad (Piet Hein)

Kapittel

5

Drøfting

5 Sammenfatning og drøfting

Drøftingen knyttes til tre drøftingsområder, der fokus er lagt på mappevurdering i lys av a) **teoretisk rammeverk** b) **teknisk kvalitet**, og c) **konklusjon** i lys av prosjektets problemstilling.

Under punkt 5.1 drøftes fokusgruppene forståelse av begrepene mappe, formativ vurdering og kriterier i lys av generelle teoretiske betraktninger. Dernest følger en drøfting av formativ vurdering og bruk av kriterier i lys av fenomenene reliabilitet, validitet og kvalitet (jf. punkt 5.2). Denne drøftingsdelen anses å være hovedessensen i prosjektet, og knyttes spesielt til Williams uttrykte forståelse av reliabilitet og validitet og Herman & Winters sertifiseringsrammeverk, med spesiell fokus på dokumentasjonsinnhenting tilknyttet en mappes tekniske kvalitet. Under punkt 5.3 trekkes konklusjoner knyttet til prosjektets problemstilling. Kapitlet avsluttes med en oppsummering under punkt 5.4.

5.1 Mappevurdering i lys av teoretisk rammeverk

5.1.1 Mapper

Fokusgruppene uttrykte mappeforståelse (jf. 4.2.1.1) følger mappekjennetegnene gitt av Stiggins (1994), Dysthe og Engelsen (2003), Barrett (2005) og Paulson, Paulson & Meyer (1991) (jf. 2.1) Den uttrykte mappeforståelse viser også at Zubizarretas (2004) fundamentale dimensjoner refleksjon, dokumentasjon og samarbeid ved arbeid med læringsmapper, ble fulgt. Læringsmappen tjente som medierende redskap, for personlig og felles kunnskapstilegnelse i en læringsprosess støttet av veiledende tilbakemelding og vurdering fra både medstudenter og lærere.

Læringen var **situert**, forankret i en bestemt kontekst, berørte både mappeinnhold, utvikling av kunnskap og vurdering, og var grunnleggende **sosial**. Mappeinnhold ble til i samhandling med andre. Studenter og lærere var en integrert del av et læringsfellesskap og læring oppsto som en konsekvens av samhandling. Den var **distribuert** i den forstand at kunnskap ble delt med fellesskapet/gruppene, der de **kommunikative** prosessene var sentrale i kunnskapsutvikling. Mappene syntes på det grunnlag å bygge på et sosiokulturelt lærings-syn (Dysthe 2001). Når mappens hensikt er å stimulere til læring, utviklet i et praksisfellesskap bygget på samhandling, synes dette læringsparadigme å passe inn.

Læringsmappen ble brukt som samhandlings- og læringsarena, og inngikk i utgangspunktet ikke i grunnlaget for formell vurdering. Å bruke digitale mappe som læringsredskap og læringsarena omfavnes av H. Otnes (2003). Når arbeid med utvikling av en tekst/studie-

oppgave står i fokus, gjør mediespesifikke virkemidler det enkelt å oppdatere, forbedre og utvikle tekst over tid. Åpne digitale mapper legger således godt til rette for arbeid med tekst over tid, og for kollektiv samhandling rundt en tekst. På det grunnlag synes bruk av mapper som lærings- og samhandlingsarena både på SU og FJU å være svært relevant. På FJU var det i praktiske fag mindre arbeid med tekst og mer arbeid med konkrete visuelle fotooppgaver. Akkurat disse studieoppgavene var det ikke så enkelt å endre i ettertid. Læringsarenaen fungerte på det grunnlag som fremvisningsarena, og la til rette for diskusjon og samarbeid om felles oppgavefortolkning og vurdering. De mediespesifikke dokumentasjonsformene inkluderte her digitale bilder, som en betydelig del av oppgaven. Ved SU kunne noe av mappene inngå som del av grunnlaget for formell sluttvurdering. Arbeidsformen involverte på denne måten både prosess og produkt. Studieoppgaver aktuelle for formell vurdering, ble lagret i egen vurderingsmappe. Slik opererte utdanningen med både lærings- og vurderingsmappe, knyttet til henholdsvis formativ og summativ vurdering. Selv om mappemetodikken ved SU medførte bruk av ulike mappetyper, var likevel fokusgruppen først og fremst opptatt av å beskrive **mappe i lys av læring**, med formativ vurdering som hoveddrivkraft. Det samme gjaldt på FJU.

Formativ vurdering hadde på begge utdanninger som mål å gi studenten informasjon om egen fremgang, og legge til rette for bedre læring og utvikling. En slik tankegang støtter bruk av **læringsmapper** (Dysthe & Engelsen 2003). Mappetypen synes adekvat, slik arbeidsformen ble praktisert ved både SU og FJU.

Ved FJU var arbeid med visuelle elementer vektlagt. En slik aktivitet peker i retning av fremvisning av en form for kompetanse, og med det bruk av **kompetansemapper** knyttet til en kunstfagtradisjon. På tross av det, var det likevel hensiktsmessig med læringsmapper ved FJU, da oppmerksomheten hovedsakelig var rettet mot å bruke mappen som samhandlings- og læringsarena, der tett samarbeid mellom lærere og medstudenter var en viktig faktor i faglig utvikling. Hensikten med mappen bestemte dermed også valg av mappetype. Viktigheten av å bestemme en mappes hensikt understrekes av mange i den pågående mappediskusjon (Zeichner & Wray 2000, Carney 2004, Barrett & Carney 2005). Med utgangspunkt i kjernepunktet **mappens hensikt**, har altså FJU valgt læringsmappe som arbeidsredskap, noe som synes hensiktsmessig i lys av deres læringsprosess.

Ved å kombinere vurderings- og læringsmappe har SU i stor grad fulgt Williams (2000a) teoretiske synsvinkel, og på den måten klart å utvikle et mappesystem, som ivaretar både vurdering **for** og **av** læring. Slik har de tatt sikte på å dekke ulike behov. I læringsforløpet er det primært fokus på læringsmappen og formativ vurdering, senere velges utvalgte studiearbeider fra læringsmappen ut til den formelle vurderingsmappen. Ved SU vurderes læringsprosessen ikke formelt, men summativ vurdering tar likevel utgangspunkt i 16 velkjente vurderingskriterier. Slik knyttes formativ og summativ vurdering sammen. Ser man denne praksis i lys av Lauvås, antyder han at faren for at summativ vurdering skyver

den formative ut, er stor. Andre (William 2000a), ser fordeler med å kombinere formativ og summativ vurdering. På FJU brukte man stort sett kun mapper som samhandlings- og læringsarena uten summativ vurdering. Slik støtter deres mappepraksis, synet som fremføres av Lauvås. Ved helt å unnlate bruk av mappeinnhold i summativ vurdering, unngås også konflikten, som kan oppstå ved å kombinere summativ og formativ vurdering. Mappe-systemet ved FJU tar på denne måten mindre hensyn til de ulike vurderingsbehov som skal dekkes, men studentene kan føle seg trygge på at ingen skjult summativ vurdering knyttes til læringsmappens innhold. Diskusjonen knyttet til formativ og summativ vurdering belyses nærmere under punkt 5.1.2.

Teknologien ble først og fremst brukt til å støtte opp om de funksjoner, som vurderingsmåten tok sikte på å inkludere. Slik bruk av digital teknologi, favnes av flere teoretikere (Barrett 2005, Otnes 2004). Teknologien var **transparent**, og i seg selv ikke interessant, men kun et usynlig verktøy i å nå de mål som var satt. En slik praksis understrekes av Norman (1999:Vii-iX). Mappene befant seg midt på åpenhetsskalaen, (jf. kapittel 2.1.3 (Otnes 2003)), og var en mellomting mellom lukket og åpen mappe. Grad av åpenhet må ses i sammenheng med mappens funksjon og hensikt. I læringsmiljøet til SU og FJU syntes en middels grad av åpenhet, å være egnet. Mappen åpnet for både autentiske skrive- og refleksjonssituasjoner, der mottakergruppene stort sett var de samme. De digitale læringsmapper ble først og fremst brukt som **samhandlings- og læringsarena** innen kullet. Å bruke digitale mapper som læringsmiljø for kunnskapsdeling og samarbeid, er en forståelse som Otnes (2003) understreker.

5.1.2 Formativ vurdering

På begge utdanninger var forståelsen at formativ vurdering var en underveisvurdering. Formålet var å hjelpe studenten til å reflektere over den videre læringsprosess. Denne forståelsen av formativ vurdering favnes av flere teoretikere (Barrett & Carney 2005, Barrett 2005, Lauvås, William & Black 1996, Sadler 2001, Black, Harrison, Lee, Marshall & William 2003). Ved SU bar formativ vurdering preg av å være en vurderingsprosess karakterisert av **syklustenking** (William og Black 1996), og ble knyttet til et kjent og synlig vurderingsrammeverk. Ved FJU fremkom formativ vurdering i felles diskusjoner og samtaler. Handling (oppgaveløsning med etterfølgende gruppe- og plenumsdiskusjon) kastet lys over bevisene (produksjonsoppgavene), og førte til kunnskaps- og erfaringslæring, som ble anvendt i arbeid med neste oppgave/utgave. Ved begge utdanninger ble formativ vurdering brukt som **uformell** vurdering, der individuell og kollektiv kunnskapsutvikling var det primære mål. Vurderingsformen hadde således til hensikt å være en vurdering **for** læring, og læringsmappen ble utviklet over tid (Lauvås, Davies 2000, William & Black 1996, Assessment Reform Group 2002).

Vurderingsformen bar ved begge utdanninger preg av å være en kombinasjon av ipsativ og kriteriebasert vurdering (jf. 2.2), der både den individuelle students og gruppes faglige ut-

viklingspotensial sto i fokus (Wittek & Dysthe 2003, Lauvås). Formativ vurdering ble knyttet til prosessen lærere og de lærende gjennomgikk, med å bestemme hvor den lærende befant seg på nåværende tidspunkt i læringsprosessen, for så å bestemme hvilke mål den lærende skulle bevege seg mot og hvordan - helt i tråd med tankegangen til Assessment Reform Group (2002). Det ble gitt **systematisk respons** på alle obligatoriske studieoppgaver. Tilbakemelding ble gitt i form av respons fra lærere og medstudenter i tillegg til egenvurdering og refleksjon. **Refleksjon** beskrives av flere som hjertet i mappevurdering (Dysthe 1999, Ellmin 2000). Det skrevne og talte språket fungerte som **kommunikasjonsredskap** mellom de lærende og lærere.

Ved SU ble studieoppgaver formativt vurdert på grunnlag av 16 kjente og synlige vurderingskriterier. Også summativ vurdering tok utgangspunkt i disse. Formativ og summativ vurdering ble dermed knyttet sammen. Ser man denne praksis i lys av Sadler (2001) og Lauvås, antyder de at faren for at summativ vurdering skyver den formative ut, er stor. Wiliam (2000a) derimot, kan se fordeler ved å kombinere formativ og summativ vurdering. Generelt er sammenknytningen mellom formativ (prosess) og summativ vurdering (produkt) et omdiskutert tema i den internasjonale mappe- og vurderingsdiskusjon. På den ene siden argumenteres det for å la den formative vurdering stå alene, for ikke å overskygges av den summative (Lauvås), mens det på den andre argumenteres for at også formell vurdering av læring er nødvendig og viktig (Wiliam 2000a). Et vurderingsopplegg som kun favner formativ vurdering, vil mottas positivt av studenter og lærere som er spesielt opptatt av læring og mindre av kontroll, men vil ignoreres av studenter og lærere som fortrinnsvis kun er opptatt av det som teller ved endelig summativ vurdering (kontroll og standarder). Fra et administrativt synspunkt vil formell vurdering av læring (summativ) være den viktigste faktor, og fra et læringssynspunkt uformell vurdering (formativ). Slik knyttes vurdering av læring til ulik hensikt og behov, som til en viss grad står mot hverandre. William & Black (1996) støtter i utgangspunktet ren formativ vurdering, og viser til at hovedhensikten med læring er selve læringsprosessen. Et faktum er dog at vurdering av læring, til enhver tid støtter ulike hensikter. En kombinasjon av summativ og formativ vurdering kan derfor være et godt alternativ, når ulike behov skal dekkes, og kan forbedre både kvalitet på læring og vurdering (Wiliam 2000a). Å utvikle et vurderingssystem som integrerer summativ og formativ vurdering på en god måte, vil da åpne for adekvat vurdering av både prosess og produkt. Barrett og Carney (2005) derimot antyder at det kan være problematisk å blande kontroll (summativ) og læring (formativ) og støtter dermed Lauvås' synspunkt. Generelt er Barrett og Carney bekymret for at for sterk kontroll (summativ), vil kunne føre til at studenten ikke på ekte vis, tør vise sitt sanne jeg, i arbeid med mappen (putting oneself in the portfolio), slik at læringsaspektet (formativ) forsvinner på bekostning av kontrollfunksjonen. En eventuell løsning for å unngå dette kan være å konstruere et balansert mappesystem, som dekker både behovet for **vurdering** og **læring**. Nøkkelen er å anerkjenne mappers konflikt-

fylte formål og paradigmer, og derfra utvikle et mappesystem, som balanserer og ivaretar behov både knyttet til læring (formativ vurdering) og evaluering (summativ vurdering). Digitale mapper kan være et kraftfullt verktøy i læring, hvis de utgjør en del av et balansert system rettet mot både vurdering **for** (formativ) og **av** (summativ) læring. Men da må det utvikles et mappesystem, som tar høyde for ulike **formål/hensikter**. Nettopp å bestemme en mappe hensikt, er et kjernepunkt i diskusjonen knyttet til mapper (jf. Carneys 6 dilemmaer nevnt i 2.5.4). Dysthe (2002) derimot støtter bruk av mapper, som favner både prosess (formativ) og produkt (summativ), men faren kan altså være at studenter opplever formativ vurdering som ureell og mindre viktig. Konsekvensen kan fort bli en ekskludering av den formative vurdering, og nedprioritering av selve læreprosessen.

Knytter man den internasjonale vurderingsdiskusjon til praksis slik den gjennomføres ved SU, antyder de empiriske funn at utdanningen på en god måte skiller formativ og summativ vurdering, i kraft av en bevisst holdning til og hensikt med separate lærings- og vurderingsmapper (jf. Barrett & Carney 2005). Å bruke samme kjente vurderingsrammeverk i summativ og formativ vurdering, bidrar også til å forsterke hensikten med formativ vurdering, og knytter summativ vurdering til læringsprosessen forut. Slik fremstår vurderingsmodellen ved SU til, på en god måte, å både ivareta et lærings- og vurderingsaspekt. Ved FJU hadde man valgt å kun knytte formativ vurdering til mapper. En slik praksis fremheves av Lauvås og Sadler. De ulike vurderingsbehov denne utdanningen har, vil da måtte ivaretas på annen måte. Oppstår eventuelt et ønske om å knytte læringsmappen ved FJU til summativ vurdering, må en inngående avklaring til, for å bestemme hvordan både lærings- og vurderingsaspektet skal imøtekommes.

5.1.3 Kriterier

På begge utdanningene definerte man kriterier som rettesnorer, og de faste kvalitative holdepunkter, for det som skulle vektlegges ved arbeid med en studieoppgave. En slik forståelse gir også Sadler (1996) og Wittek & Dysthe (2003) uttrykk for. Kriteriene ble utviklet og forstått gjennom samhandling, der samarbeid var faglig begrunnet, og faller på denne måten innenfor kategorien **spesifikke** kvalitets- og vurderingskriterier (Wittek og Dysthe 2003). Forskjellen i kriteriepraksis besto i om studenter og lærere på forhånd hadde kjennskap til vurderingskriteriene eller ikke.

Erfaringene med kriteriebaserte/ikke kriteriebaserte vurderingskriterier er imidlertid delte. Sadler (1989) antyder at det er en illusjon å tro at vurderingskriterier i formativ vurdering kan uttrykkes fullt ut eksplisitt. I vurdering vil det alltid være behov for faglig skjønn, som delvis forutsetter underforståtte, usynlige og tause kriterier. Å anvende vurderingskriterier er en form for praksis, som til enhver tid vil involvere underforståtte og ikke uttalte kriterier. Som regel er eksplisitte kriterier kun et nedskrevet uttrykk, for normer innen fellesskapet. Kriteriebasert vurdering støttes imidlertid av flere forskere (Baume 2002, Dunn, Parry, & Morgan 2002, Shephard 2001, Wittek & Dysthe 2003), og flere undersøkelser peker på at

mange studenter ønsker oppgavevurdering som følger klare eksplisitt nedskrevne vurderingskriterier. Begrunnelsen er at kriteriebasert vurdering vil gi en mer presis og rettferdig vurdering. Vurderingspraksis ved SU henter sin støtte i slike undersøkelser. Forskningsundersøkelser indikerer dog også, at der vurderingskriterier eksisterer, er det få studenter som vet hvilke kriterier som gjelder (Wittek 2004, Dysthe & Kjeldsen 1997, Dysthe & Breistein 1999, Dysthe & Lied 2001).

Andre forskningsundersøkelser (Dysthe et al. 2005) målfører derimot at studenter som studerer i en læringskultur preget av tett dialog mellom lærer og student, der det forhandles om vurderingsramme/kriterier, ikke er spesielt opptatt av å få eksplisitt nedskrevne kriterier. Kriteriene utvikles og forstås i den gode samtalen/dialogen, og blir en implisitt (ikke skriftlig uttrykt) del av hele læringssituasjonen. Slike forskningsresultater støtter FJUs vurderingspraksis, men kolliderer med synet til Shepard (2001). Å tilegne seg kriterieforståelse gjennom samarbeid med andre, fremheves dog av Dysthe et al. (2005), Rust et al. (2003) og Sadler (1983).

Om vurderingskriterier er styrende element i vurdering, er det et poeng at de er synlige og forhandlingsvennlige av flere grunner (Wittek & Dysthe 2003, Shephard 2001). Vurdering basert på klart definerte vurderingskriterier vil bidra til at studenten får en forståelse av hva som må til for å bli god, og kaster lys over hvordan dyktighet kan utvikles.

Sentrale grunner for **synlige og forhandlingsvennlige vurderingskriteriene** på **SU** var samhandling bygget på et sosiokulturelt læringssyn, egenvurdering, responsgivning, kjennskap til profesjonsrelevante kriterier, og bruk av kriteriene i summativ vurdering. Vurderingskriteriene var i utgangspunktet så vidt formulert, at en form for kriterieforhandling skjedde ved at studenten selv valgte ut og definerte aktuelle vurderingskriterier inn i studieoppgaven. Å utarbeide vurderingskriterier som er vidt formulert favnes av Wittek og Dysthe (2003). Bruk av kjente vurderingskriterier gir assosiasjoner til en kriteriebasert vurderingsform. Baume, (2002), Dunn et al. (2002) og Wittek & Dysthe (2003) antyder alle at en klar fordel med kriteriebasert vurdering er at vurderingsgrunnlaget blir **gjennomsiktig**, i den forstand at referansene for vurdering er synlig for alle involvert i vurderingen. Klart definerte kriterier øker med all sannsynlighet sjansen for at det oppstår en felles forståelse av vurderingsgrunnlaget. Det vil bidra til å forsterke vurderingens reliabilitet (pålitelighet), noe flere teoretikere legger vekt på. Baume antyder til og med at bruk av kriterier vil forsterke en vurderings validitet (gyldighet), men her er det ulik oppfatning.

Sentrale grunner for **ikke å ha synlige nedskrevne vurderingskriterier** ved **FJU** i praktiske fag, var at studentene **selv** i første omgang skulle definere og velge kriterier som passet til den praktiske studieoppgaven. Selvstendighet og ansvar for å ta egne valg var viktige i den innledende arbeidsprosess. Eksplisitt nedskrevne kriterier som utgangspunkt

for en praktisk studieoppgave, ville bidra til å stoppe selvstendig tenking og kreativitet i løsning av studieoppgaven. I etterfølgende dialog/samtale og fellesdiskusjon ble **synliggjøring** og **forhandling** av vurderingskriteriene en realitet. I noen tilfeller kan kriteriebasert vurdering altså oppleves som en hemske i en læringsprosess. Som i undersøkelsen gjennomført av Engelsen ved Høgskolen Stord/Haugesund (Dysthe et al. 2005), har også FJU et studentmiljø med få studenter, som hele tiden er i nær kontakt med lærere og medstudenter. Kriterier som gir rom for egen studenttolkning og faglig skjønn, fremheves av Sadler (1989) og Wittek & Dysthe (2003).

Samarbeid om kriterier var altså dominerende på begge studier. At samarbeid om kriterier gir stort potensial for forbedring av både læring og vurdering, er noe Wittek og Dysthe (2003) trekker frem. De antyder dog også at sentrale vurderingskriterier må synliggjøres. Nøkkelen til **reliabel** og **valid kriteriebasert vurdering** uttrykker de, er at det på **forhånd** er etablert klare og egnede vurderingskriterier knyttet til prestasjoner. Det bidrar både til å vise hvilken kvalitet som belønnes og verdsettes, retningslinjer for undervisning, og tydeliggjør for studenten hva som må til for å oppnå en god karakter. Også Baume fremhever at synlige kriterier øker sannsynligheten for reliabel og valid kriteriebasert vurdering. Et klart og egnet vurderingsrammeverk letter tilbakemelding på arbeid som er i produksjon, og gir mening til selve vurderingsarbeidet. Slike tanker støtter den eksplisitte kriteriebruk slik den praktiseres ved SU, og taler mot kriteriepraksis på FJU.

Om ikke kriterier defineres kan studenter utvikle sin egen oppfatning av hva som teller, og utvikle egne Rules of the Game (Norton et al. 1996). Det er nok det som i første omgang skjer, når studentene løser studieoppgaver i praktiske fag på FJU. Det er først i etterkant, i de formative diskusjonene, kriteriene formuleres, presiseres og synliggjøres.

5.2 Mappevurdering i lys av teknisk kvalitet

De siste 30 årene har det blitt allment akseptert at validitet ikke er en egenskap ved en vurdering, men **en egenskap av konklusjonene**, som trekkes på basis av vurderingen (William 2000a). Dette synet er i seg selv interessant i lys av prosjektets forskningsområde, når hovedhensikten har vært å undersøke forståelse av fenomenene reliabilitet og validitet knyttet til formativ vurdering og bruk av kriterier. Det synes derfor nærliggende å plassere en drøfting av reliabilitet og validitet, inn under dette paradigme.

En vurdering påberoper seg ingenting, den vurderer kun det den vurderer, og forteller ingenting ut over det resultat som fremkommer av vurderingen (William 2000a). Ser man formativ vurdering i lys av utsagnet over, antyder det at formativ vurdering da kun vurderer det den vurderer, og ikke forteller noe ut over det resultat som fremkommer av vurderingen. I det perspektiv vil formativ vurdering, slik jeg ser det, kun kunne påberope seg å uttrykke noe om arbeidsmåten innen konteksten den utøves i, og resultatene som fremkommer der.

Det er et perspektiv som også vil ligge til grunn i min drøfting av fagfeltets uttrykte fenomenforståelse.

5.2.1 Reliabilitet og validitet

Prosjektets problemstilling uttrykte ønske om å undersøke fagfeltets forståelse av fenomenene reliabilitet, validitet og kvalitet, og hvordan disse ble utviklet og utøvd i en digital mappevurderingskontekst, der formativ vurdering og bruk av kriterier inngikk. William (2000a) uttrykker om **reliabilitet** at fenomenet kan karakteriseres ved et forhold mellom **signal** og **støy**. Jo kraftigere signal jo bedre reliabilitet (pålitelighet), jo mer støy jo dårligere reliabilitet. For å forbedre reliabiliteten har man to innfallsvinkler, enten å redusere støyen eller øke signalene. Om **validitet** uttrykker William at det lenge har vært akseptert å definere en vurdering som valid, når den **vurderer det den påberoper seg å vurdere**. Men validitet er ikke en enhetlig vurderingsegenskap, den endres i forhold til hvem som vurderer og vurderes. Så i stedet for å være opptatt av å **hvor nøyaktig man vurderer** foreslår William heller å se på **hva er det nøyaktig man vurderer**. Williams uttrykte syn vil være referanse i den videre drøfting. Et av forskningsspørsmålene var også å innpasse fagfeltets forståelse i rammeverket utviklet av Herman & Winters(1994). Deres teori vil derfor også være hovedreferanse i drøftingen under punktene 5.2.1.1, 5.2.1.2 og 5.3.

5.2.1.1 Formativ vurdering

Hva er det da nøyaktig som vurderes, eller sagt på en annen måte hvilken **funksjon** er det nøyaktig formativ vurdering skal ha, og hvilke konklusjoner kan fremmes på grunnlag av respondenters uttrykte forståelse? Ved SU ble formativ vurdering oppfattet å være pålitelig og gyldig når:

- 1 Vurderingen var rettet mot studieårets rammeplan og den enkelte studieoppgaves faglige fokus
- 2 Ble knyttet til en vurdering basert på 16 kjente vurderingskriterier
- 3 Tilstrebet å veilede en student i å nå individuelle og spesifiserte fagmål
- 4 Var meningsfull og relevant for både den som mottok og ga vurdering
- 5 Var motiverende, stimulerte til forbedring av studieoppgaven og ga økt kunnskapsnivå
- 6 Ble gitt av fagkompetente personer som kjente godt til studenten
- 7 Var konkret og presis
- 8 Førte til økt kunnskapstilegnelse og fagforståelse

Ved FJU var formativ vurdering pålitelig og gyldig når den:

- 1 Rettet seg mot kriterier, som var akseptert å være pålitelige og gyldige i bransjen
- 2 Ble gitt av fagkompetente studenter og lærere underveis i læreprosessen

- 3 Fant sin aksept og forståelse i åpne fagrelaterte diskusjoner
- 4 Tok utgangspunkt i ikke nedskrevne kriterier i et åpent og diskuterende miljø
- 5 Var preget av subjektivitet
- 6 Bidro til å fremelske en students eget faguttrykk (foto)
- 7 Bidro til at studenten ikke stagnerte i sin kreative utvikling
- 8 Bidro til tvisyn i løsning av studieoppgaver

a) Reliabilitet og validitet i lys av William

Sammenholder man punktene nevnt over, med Williams uttrykte synspunkt, er ikke reliabilitet og validitet en egenskap ved formativ vurdering i seg selv, men en egenskap og karakteristikk av **konklusjonene**, som trekkes på basis av en vurdering. Forståelsen av pålitelig og gyldig formativ vurdering, fremkommer da som egenskaper/karakteristikker av konklusjonene som fremmes. Som empirien og punktene på punktlistene over viser, hadde fokusgruppene ved de to utdanningene noe ulik praksis hva gjelder forståelse, utvikling og synliggjøring av fenomenene. Med det som utgangspunkt vil det derfor være hensiktsmessig å knytte pålitelig og gyldig formativ vurdering ved henholdsvis SU og FJU, til ulike områder. Tolker man Sus forståelse av pålitelig og gyldig formativ vurdering (jf. 4.2.2.1), trer følgende konklusjoner frem:

SU:

1. Vurdering ivaretar studiets målsetninger skissert i rammeplan/studieplan/studieoppgaver
2. Vurdering tar utgangspunkt i 16 velkjente nedskrevne vurderingskriterier som forstås og anvendes
3. Vurdering ivaretar studentens individuelle (eller gruppes) og spesifikke læringsmål for faget
4. Vurdering oppleves meningsfull og relevant
5. Vurdering motiverer for videre arbeid og forbedring av studieoppgavene
6. Vurdering er konkret presis og faglig relevant
7. Vurdering gir økt kunnskap og fagforståelse
8. Vurdering gis fra fagkompetente personer

Med utgangspunkt i konklusjonene gjengitt i lista over vil reliabel og valid formativ vurdering ved SU dermed knyttes til egenskapene: 1) målbasert, 2) nedskrevet kriteriebasert, 3) individuell tilpasset, 4) meningsfull og relevant, 5) motivasjonsfremmende, 6) konkret, presis og faglig relevant, 7) fag- og kunnskapsberikende og 8) fagkompetent.

Tolker man forståelsen av reliabel og valid formativ vurdering, slik det ble uttrykt ved FJU (jf. 4.2.2.1), trer følgende konklusjoner frem.

FJU:

1. Vurdering bidrar til å ivareta bransjens kriterier for et godt arbeid
2. Vurdering gis fra fagkompetente personer underveis i læreprosessen
3. Vurdering gis i åpne fagrelaterte diskusjoner
4. Vurdering tar utgangspunkt i ikke nedskrevne kriterier, som diskuteres og fremmes i lys av konsensus
5. Vurdering er preget av subjektivitet
6. Vurdering bidrar til å fremme studentens eget faguttrykk
7. Vurdering bidrar til utvikling av en students kreativitet
8. Vurdering fremmer utvikling av et subjektivt fagvisyn

I lys av disse konklusjonene vil reliabel og valid formativ vurdering ved FJU, dermed beskrives ved karakteristikkene 1) bransjetilpasset, 2) fagkompetent, 3) diskusjonsbasert, 4) ikke nedskrevet kriteriebasert, 5) subjektiv, 6) individuell tilpasset, 7) kreativitetsutviklende, 8) fagvisynsfremmende.

b) Formativ vurdering i lys av Herman & Winters sertifiseringsrammeverk

Fenomenene **reliabilitet** og **validitet** ble i 1994 foreslått av Herman & Winters som nøkkelidentifikatorer til å dokumentere en **mappes tekniske kvalitet**. I lys av **formell** vurdering synes disse dokumentasjonsområder relevante. De empiriske funn viser imidlertid at mappene som ble brukt, i utgangspunktet var rene læringsmapper knyttet til områder som samhandling og læringsprosess og mindre til formell vurdering. Knytter man det til Herman & Winters sertifiseringsrammeverk, støtter det innhenting av dokumentasjon i lys av læring og nøkkelidentifikatorene foreslått av Carney (2004). De empiriske funn indikerer derfor klart at en mappens teknisk kvalitet ved de undersøkte utdanningene, bør dokumenteres ved mappens evne til å **påvise** og **fremvise** dyp læring og ikke reliabilitet og validitet.

Ved SU var forståelsen at formativ vurdering var pålitelig og gyldig når den var 1) målbasert, 2) nedskrevet kriteriebasert, 3) individuell tilpasset, 4) meningsfull og relevant, 5) motivasjonsfremmende, 6) konkret, presis og faglig relevant, 7) fag- og kunnskapsberikende og 8) fagkompetent. Ved FJU var formativ vurdering pålitelig og gyldig når den var 1) bransjetilpasset, 2) fagkompetent, 3) diskusjonsbasert, 4) ikke nedskrevet kriteriebasert, 5) subjektiv, 6) individuell tilpasset, 7) kreativitetsutviklende, 8) fagvisynsfremmende. Til å dokumentere læringsmappens **tekniske kvalitet** på disse utdanningene, vil det derfor være hensiktsmessig å innhente dokumentasjon tilknyttet de ovenfor uttrykte funksjonene. For at formativ vurdering skal oppfattes som pålitelig og gyldig, må det dokumenteres at de uttrykte egenskapene ivaretas. Viser dokumentasjonen at funksjonene ivaretas, kan man gå god for mappens tekniske kvalitet. I så tilfelle, er ett av dokumentasjonsområdene i sertifiseringsrammeverket tilfredsstillende. Da gjenstår det å dokumentere de tre andre, for å kunne si noe helhetlig om mappens totale kvalitet.

5.2.1.2 Kriterier

Kriterier dannet rammen for formativ vurdering. Det som skilte studiemiljøene var om kriteriene i utgangspunktet var eksplisitt nedskrevet, eller om de var underforståtte, usynlige og tause, for siden gjennom diskusjon å fremmes i lys av konsensus.

På SU fungerte de eksplisitt nedskrevne vurderingskriterier som støttestillas og vurderingsramme, for alle involvert i den formative vurdering. Det faktum at kriteriene var vedtatt brukt ved studiet, gjorde at de fant sin pålitelighet og gyldighet, som ble ytterlig forsterket, når studenten i studieoppgavene først viste, og senere i samhandling med medstudenter og lærere, tilegnet seg utvidet forståelse, tolkning og anvendelse av kriteriene. Vurderingskriteriene ved SU fant på denne måten sin pålitelighet og gyldighet, i **faget** og i **samarbeid** mellom studenter og lærere. Ved FJU ble kriteriene i utgangspunktet kun knagger man kunne orientere seg etter i første omgang. Det var i plenumsdiskusjonen de fikk sin pålitelighet og gyldighet. Men også vurderingskriteriene ved FJU fant sin pålitelighet og gyldighet i faget og i samarbeid.

Når er så vurderingskriterier pålitelige og gyldige? Ser man det i lys av William ville fokus være rette mot å forstå hva er det nøyaktig kriteriene vurderer, og ikke hvor nøyaktig vurderer kriteriene. Validitet er heller ikke en enhetlig kriterieegenskap, **den endres i forhold til hvem som bruker kriteriene, og i hvilken sammenheng de opptrer**. En slik beskrivelse støtter undersøkelsens empiriske funn. Ser man dette i lys av kriterieforståelse og praksis ved både SU og FJU blir konklusjonen at:

1. Kriteriene bidro til faglig forankret
2. Kriteriene ble knyttet til både teori og profesjon
3. Kriterieforståelse ble utviklet i samarbeid

Av konklusjonene følger egenskapene 1) faglig forankring 2) teori- og praksisrelatert og 3) samarbeidsbasert. Det som skilte kriteriepraksis var om kriterier på forhånd var nedskrevet og synlig eller ikke.

Ser man de to utdanningenes kriteriepraksis i lys av Herman & Winters sertifiseringsrammeverk, vil man da for å dokumentere kriteriene tekniske kvalitet, måtte innhente dokumentasjon, som viser at bruk av kriterier bidrar til faglig forankring, profesjonsforståelse og utvikles i samarbeid. Generelt vil dokumentasjon av teknisk kvalitet tilknyttet formativ vurdering integrere dokumentasjon av kriterier (jf. diskusjon i 5.2.1.1), men også spørsmål av typen "Hva er den nøyaktige funksjon til dette kriteriet?", vil åpne for diskusjon rundt kriteriers pålitelighet og gyldighet. Slik kan både uttalte, uuttalte, underforståtte og skjulte kriterier få sitt innhold, slik jeg ser det.

5.3 Konklusjon i lys av problemstilling

Prosjektets problemstilling tok sikte på å belyse:

1. Hvilken forståelse knytter praksisfeltet til fenomenene kvalitet, reliabilitet og validitet?
2. Hvordan utvikles og utøves fenomenene i en mappevurderingskontekst, der formativ vurdering og bruk av kriterier inngår?
3. Hvordan passer fenomenforståelsen inn i sertifiseringsrammeverket utviklet av Herman & Winters?

I de følgende kapitler, belyses problemstillingen med utgangspunkt i drøftingen under punkt 5.1 og 5.2.

5.3.1 Fenomenforståelse

Som både empirien og drøftingen viser var reliabilitet, validitet og kvalitet uvante fenomener å knytte til læringsarbeidet og læringsmiljøet fokusgruppene jobbet innenfor. Det ble da heller ikke under intervjuene, gitt lettvinde og enkle fasitsvar til forståelse av fenomenene. For å fylle dem med forståelse, viste fokusgruppene ved begge utdanninger til måten de jobbet på, og knyttet dermed reliabilitet, validitet og kvalitet til selve **arbeidsmetodikken**, og forståelse av **fagprofesjonen**. For å uttrykke fenomenforståelse, trakk fokusgruppene frem funksjonene den formative vurdering og kriteriene skulle fylle. Fenomenforståelse ble dermed uttrykt i **utøvelse**.

Fokusgruppene forståelse stemmer med Williams (2000a), som uttrykker at validitet bør ses i lys av spørsmålet: Hva er det nøyaktig formativ vurdering og kriterier vurderer? Knyttes det til undersøkelses empiriske funn, er det nærliggende å erstatte spørsmålet over med spørsmålet: **Hva er den nøyaktige funksjon til formativ vurdering og kriterier?** Funksjonene som fremkommer på grunnlag av spørsmålet, er da avgjørende for hva som fremstår som pålitelig og gyldig.

Empirien presentert i kapittel 4, og drøftingene ført i lys av William og Herman & Winters i dette kapittel (jf. 5.2), viser at de to fokusgruppene ga uttrykk for både like og ulike aspekter, når de beskrev hva de anså å være pålitelig, gyldig og kvalitet. **Ulike aspekter støtter utvikling av fenomenforståelse i lys av arbeidsmetode og fagprofesjon**. Funnene kan derfor indikere at for å forstå fenomenene kvalitet, pålitelighet og gyldighet knyttet til formativ vurdering og bruk av kriterier, er det relevant å forstå dem i lys av valgt arbeidsmetode og fagprofesjon. Fenomenforståelse vil på denne måten variere, alt etter hvilken profesjon/studie de knyttes til, og få innhold og forståelse gjennom vedtatte planer og rammeverk og arbeidsmetode. Følges Williams tese uttrykkes pålitelig og gyldig formativ vurdering og kriterier gjennom egenskaper/karakteristikk, som fremkommer på grunnlag av konklusjoner som trekkes. De empiriske funn indikerer at **funksjonene** fremheves ved formativ vurdering og kriterier, og slik blir **styrende parametere** for forståelse av pålitelighet, gyldighet og kvalitet.

5.3.2 Utvikling og utøving

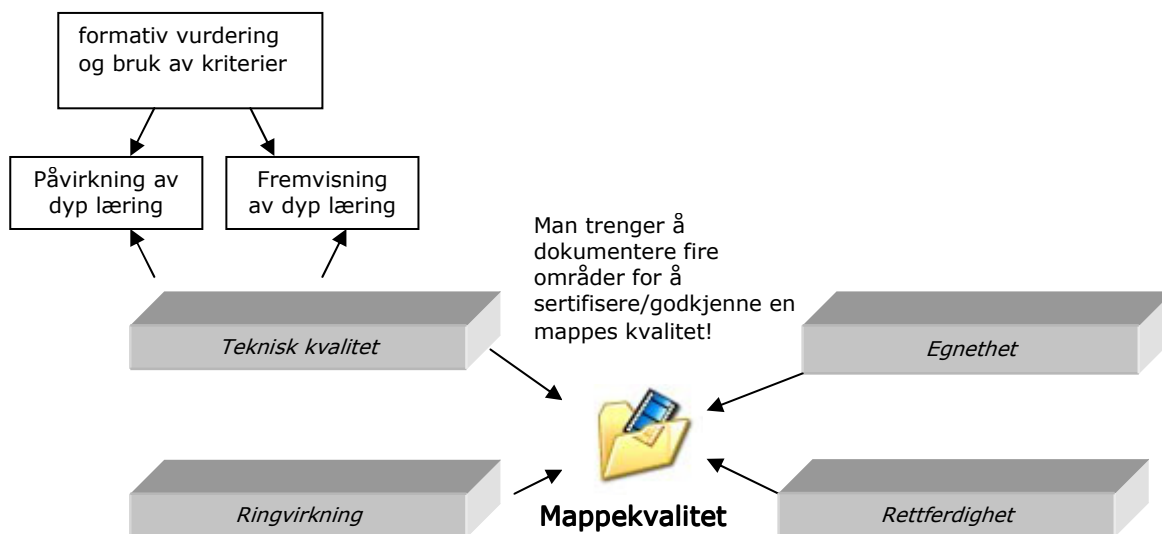
Utvikling og utøving av pålitelig, gyldig og kvalitativ formativ vurdering, skjedde både likt og ulikt i de to studiemiljøene. På begge utdanninger ble formativ vurdering utviklet og utøvd gjennom systematiske responser, enten fra enkeltpersoner eller fra grupper. Også egen-vurdering og refleksjon over egen læringsprosess, inngikk som viktig del av formativ vurdering.

Ved SU tok formativ vurdering utgangspunkt i 16 nedskrevne og vedtatte vurderingskriterier. Kriteriene var pålitelige og gyldige, i kraft av å være vedtatt brukt ved avdelingen. Pålitelighet og gyldighet ble ytterligere, utviklet og utøvd, når studentene i formativ vurdering ble bedt om å uttrykke egen faglig forståelse, fortolking og anvendelse av kriteriene i studieoppgavene. På FJU fulgte utvikling og utøving av formativ vurdering en annen praksis. Her var det på forhånd ikke nedskrevne, synlige og klart definerte vurderingskriterier knyttet til oppgaveløsning. Aktuelle vurderingskriterier ble først drøftet og utviklet i etterkant av oppgaveløsning. Pålitelige og gyldige vurderingskriterier fikk dermed sitt innhold gjennom fagdiskusjon/plenumsdiskusjon, og i samarbeid og samhandling mellom studenter og lærere. Utvikling og utøving ble her knyttet til **erfaringslæring**. De empiriske funn i denne undersøkelsen kan derfor indikere, at utvikling og utøving av formativ vurdering skjer best når samarbeid og samhandling er bærende elementer innenfor praksisfellesskapet, og at behovet for synliggjøring av vurderingskriterier helt vil avhenge av læringsmiljøet.

5.3.3 Innpasning i Herman & Winters sertifiseringsmodell

Med en problemstilling forankret i å undersøke forståelsen av reliabilitet, validitet og kvalitet, falt det i lys av Herman & Winters sertifiseringsrammeverk, hensiktsmessig å ta utgangspunkt i dokumentasjonsinnhenting knyttet til området teknisk kvalitet. Drøftingsdelens kapittel 5.2 beskriver aktuell dokumentasjonsinnhenting for de to undersøkte utdanningene. Funnene indikerer at fenomenforståelse ble uttrykt og knyttet til arbeidsmetode/funksjoner og fagforankring. Forståelse av teknisk kvalitet var således mer forankret i en lærings-, enn i en formell vurderingsprosess. Dette medførte at fenomenene ble mindre assosiert med det man normalt forbinder med pålitelighet (måle nøyaktig) og gyldighet (måle innhold), og mer assosiert med læring, arbeidsmåte, fagutvikling og funksjoner som skulle ivaretas. Når vurdering har til hensikt å være en vurdering **for** læring, synes det å være et godt prinsipp. Funnene kan derfor antyde at fenomenene reliabilitet og validitet passer bedre som nøkkel-identifikatorer for innhenting av dokumentasjon ved mapper brukt i formell vurdering, og ikke er de rette til innhenting av dokumentasjon, når man er opptatt av å dokumentere en mappes tekniske kvalitet i lys av læring og formativ vurdering. Mye av det fokusgruppene ga uttrykk for under intervjuene, var nettopp rettet mot å beskrive hvordan arbeidsmetoden/mappen kunne stimulere til påvirkning og fremvisning av dyp læring. En slik tankegang

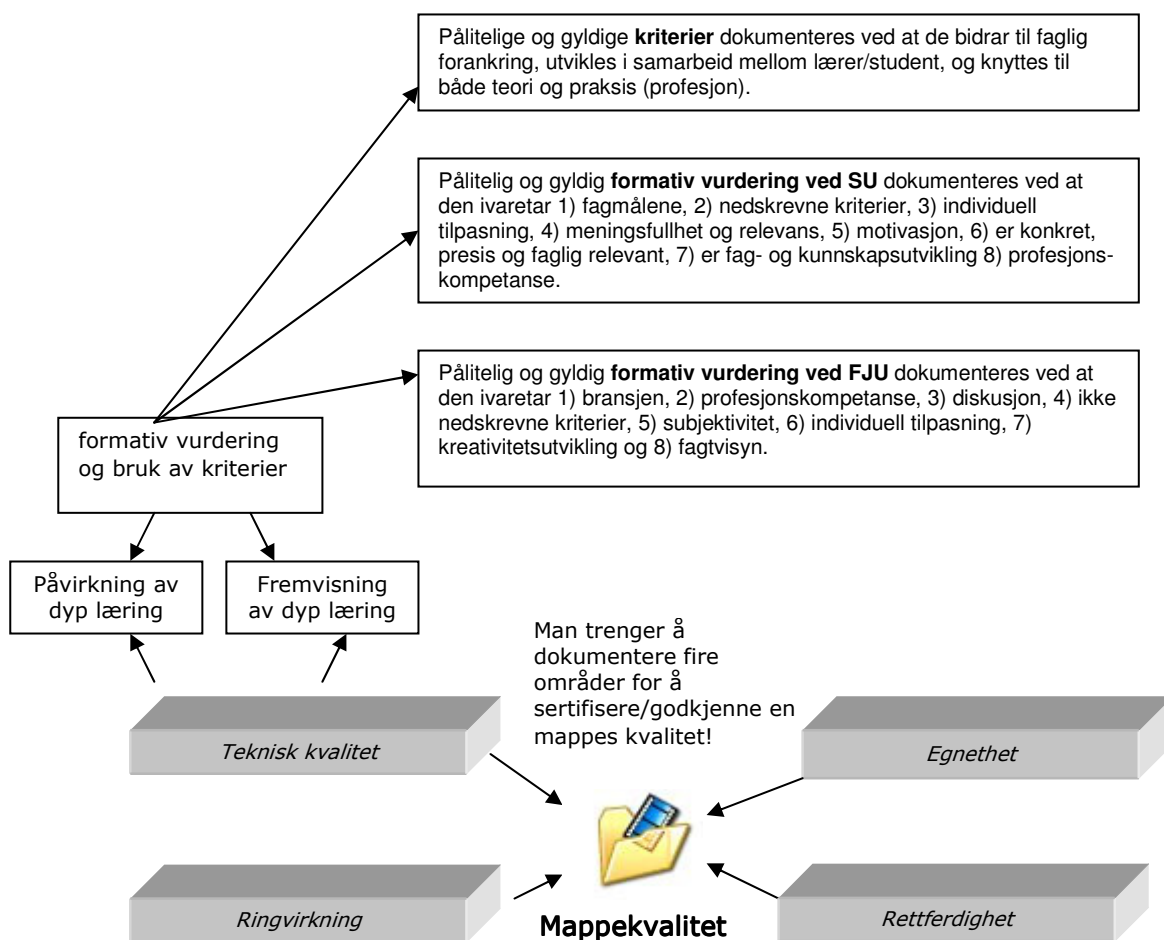
støtter innhenting av dokumentasjon i lys av mappens evne til **påvirkning** og **fremvisning** av dyp læring.



Figur 15: Teknisk kvalitet i lys av læring

I så henseende er det relevant å dokumentere læringsmappenes tekniske kvalitet, med Carneys (2004) foreslåtte nøkkelidentifikatorer som utgangspunkt, slik modellen i figur 15 illustrerer.

Innpasser man studiemiljøenes fenomenforståelse i modellen, vil en utvidet modell inkludere områdene/funksjonene de to utdanningene fremhevet ved formativ vurdering (jf. 5.2). Å dokumentere mappens tekniske kvalitet i lys av formativ vurdering vil dermed ved SU bety, at metodens evne til å ivareta 1) fagmålene, 2) nedskrevne kriterier, 3) individuell tilpasning, 4) meningsfullhet og relevans, 5) motivasjon, 6) er konkret, presis og faglig relevant, 7) er fag- og kunnskapsutvikling 8) profesjonskompetanse må dokumenteres. Ved FJU må mappens tekniske kvalitet dokumenteres ved at den ivaretar 1) bransjen, 2) profesjonskompetanse, 3) diskusjon, 4) ikke nedskrevne kriterier, 5) subjektivitet, 6) individuell tilpasning, 7) kreativitetsutvikling og 8) fagtvissyn. Den utvidede modellen vil da fremstå slik modellen i figur 16 viser.



Figur 16: Dokumentasjon av mappens tekniske kvalitet ved SU og FJU

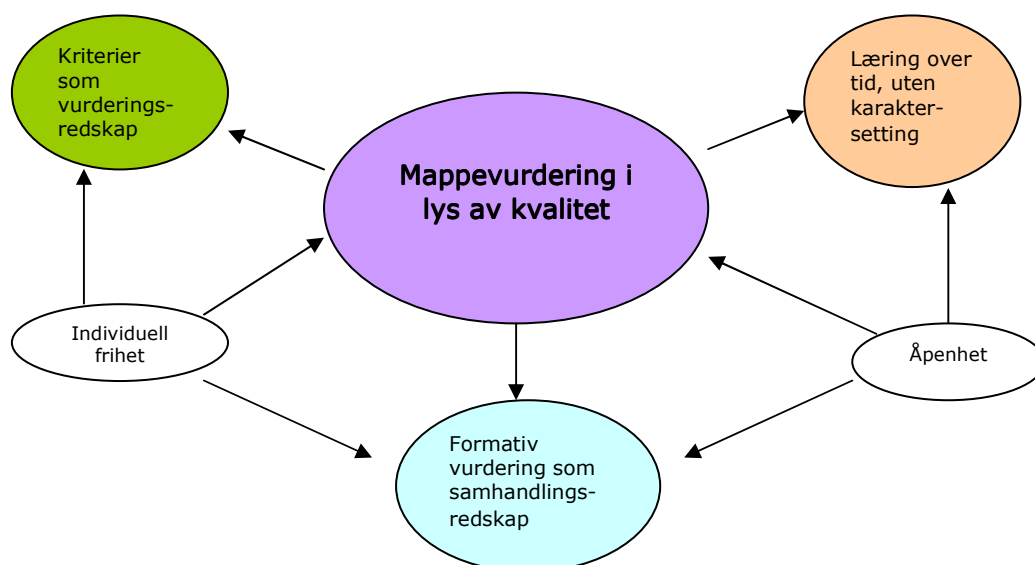
5.4 Oppsummering

Generelt var mappemetodikken hos begge utdanninger bygget på et sosiokulturelt lærings-syn, der samarbeid og organiserte aktiviteter i utvikling av mappeinnhold, sto i fokus. Det var først og fremst rene læringsmapper som ble brukt i den formative veiledningsprosess, der mappen fungerte som lærings- og samhandlingsarena. På SU hadde man i tillegg valgt å knytte formativ og summativ vurdering sammen, via de kjente vurderingskriteriene. Å kombinere formativ og summativ vurdering er generelt et omdiskutert tema. Formativ vurdering var selve kjernen i undervisningsveiledningen og var preget av syklustenking, der fremlagte bevis førte til tolking, deretter handling, og fremlegging av nye bevis. Vurderingens primære mål var å hjelpe studenten videre i læreprosessen, der kriterier var de kvalitative holdepunkter for hva som skulle vurderes og vektlegges i en oppgave. På SU brukte man 16 velkjente nedskrevne vurderingskriterier, mens de på FJU ble forhandlet frem i løpende fagdiskusjoner. Også her dokumenteres det ulik erfaring, med hensyn til hva som er å foretrekke.

Hovedhensikten med undersøkelsen var å forstå hvilken forståelse utvalgte studiemiljøer knyttet til fenomenene kvalitet, reliabilitet og validitet. Funnene kan indikere at i lærings-

miljøer der mapper brukes i forbindelse med formativ vurdering og det inngår bruk av kriterier, assosieres pålitelighet, gyldighet og kvalitet, mindre med det man normalt forbinder med disse fenomenene, og mer med arbeidsmåter og fagutvikling. Når vurdering har til hensikt å være en vurdering **for** læring, synes dette å være et godt prinsipp.

Som **kvalitet** ved arbeid med mapper fremhevet begge utdanningene spesielt: **åpenhet**, **prosessjobbing** over tid med fokus på læring og utvikling uten karaktersetting, **kontinuerlig** tilbakemelding, **frihet** til/ansvar for å gjøre individuelle valg, og bruk av **kriterier**. Visualiserer man de fremhevede kvalitetsegenskapene, i en kvalitetsmodell (figur 17) med mappevurdering i sentrum, fremstår de som satellitter i bane rundt sentrum. Kriterier som vurderingsredskap, læring over tid uten karaktersetting og formativ vurdering som samhandlingsredskap er selve kjernen i den formative vurdering. Kvalitetssatellitene åpenhet og individuell frihet forutsettes da å være usynlig (transparent) til stede.



Figur 17: Mappevurdering i lys av kvalitet

Ser man modellen i lys av teori, faller satellittene læring over tid og formativ vurdering innenfor mappekjennetegnene fremhevet av Dysthe & Engelsen (2003) og Stiggins (1994) (jf. punkt 2.1). Satellitten kriterier ivaretar imidlertid et femte punkt, nemlig bruk av et vurderingsrammeverk. Det fremheves av Wittek og Havnes (2005).

Spesielt Herman & Winters har pekt på viktigheten av å dokumentere mappekvalitet. Ser man studiemiljøenes fenomenforståelse i lys av området teknisk kvalitet (jf. 2.5.1), viser de empiriske funn, at det for det undersøkte praksisfeltet er hensiktsmessig å dokumentere mappene i lys av læring, slik Carney foreslår, og knytte dokumentasjonsinnhenting til mappers evne til påvirkning og fremvisning av dyp læring. Knyttes dette til Herman & Winters sertifiseringsrammeverk vil en mappes tekniske kvalitet ved SU, måtte dokumenteres ved at formativ vurdering er målbasert, basert på nedskrevne kriterier, individuell tilpasset, meningsfull og fagrelevant, motivasjonsfremmende, konkret, presis og faglig

relevant, fag- og kunnskapsberikende og fagkompetent. På FJU vil en mappes tekniske kvalitet dokumenteres ved at formativ vurdering er bransjetilpasset, fagkompetent, diskusjonsbasert, basert på uskrevne (implisitt) kriterier, subjektiv, individuell tilpasset, kreativitetsutviklende og fagtvisynsfremmende. Når det da kan dokumenteres at egenskapene ivaretas i utøvelse av formativ vurdering, er teknisk kvalitet akseptabel. Er den det kan formativ vurdering slik jeg ser det, også sies å besitte kvalitet.



*Den virkeligt vise
er den som formår
at forstå også den
som han ikke forstår.*

(Hein 1968:39)

Figur 18: Den virkeligt vise (Piet Hein)

6 Konklusjon

Intensjonen med undersøkelsen var å få frem **hvordan fenomenene kvalitet, reliabilitet og validitet ble forstått, utviklet og utøvd, i bruk av formativ vurdering og kriterier.**

Det var derfor helt essensielt å få kontakt med læringsmiljøer, som hadde god erfaring med mappevurdering, og jobbet i henhold til begrepet **rik mappekontekst** (jf. punkt 3).

Empirien ble samlet inn ved hjelp av datatriangulering, og hentet fra skriftlige dokumenter, intervjuer og studentoppgaver. Den skriftlige dokumentasjon ga verdifull innsikt i en arbeidsmetode, som for meg var helt ukjent, og bidro til et foreløpig forståelsesgrunnlag, av de undersøkte studiemiljøene. Telefonintervjuene og ansikt til ansikt samtalen ga uvurderlig innsikt i, hvilken personlig forståelse respondentene hadde av begrepene og fenomenene, som undersøkelsen tok sikte på å belyse. Uten disse samtalen, hadde heller ikke min tolkning og forståelse vært fyllestgjørende. Også noen få studentoppgaver ga innsikt i hvordan formativ vurdering, skriftlig (SU) eller muntlig (FJU), kunne utøves i lys av en studieoppgaves rammer. De ga konkret innblikk i, hvordan veiledning kunne uttrykkes, for å motivere en student til forbedring av en studieoppgaves innhold. Om responsgivning hadde vært et av hovedmålene i forskningsprosjektet, hadde innsyn i dette vært av stor betydning. Nå ble det en supplering til de to andre datakildene.

6.1 Undersøkelsens resultater

Undersøkelsens problemstilling ønsket å belyse 3 hovedområder:

1. Forståelse av fenomenene reliabilitet, validitet og kvalitet.
2. Utvikling og utøving av fenomenene i formativ vurdering og bruk av kriterier.
3. Innpassing i Herman & Winters (1994) sertifiseringsrammeverk.

Funnene viser at fenomenforståelse ble uttrykt i og knyttet til funksjonene den formative vurdering skulle fylle. Fenomenene **reliabilitet, validitet og kvalitet** var således en integrert del av mappemetodikken, slik den ble praktisert i studiemiljøene.

Undersøkelsen var som sagt rettet mot å forstå hvordan praksisfeltet forsto, utviklet og utøvde fenomenene. I drøftingen (jf. 5) er derfor fokusgruppens perspektiv ved arbeidsmetoden fremhevet og innpasset i en **fenomenforståelse** (jf. 5.2 og 5.3). Som det fremgår av konklusjon og oppsummering i kapittel 5, tyder de empiriske funn på at pålitelig og gyldig formativ vurdering ved de to undersøkte studiemiljøene, fant sin referanse i **samhandling** og **fagforankring** på studiet. Funnene kan derfor antyde at forståelse av pålitelig

og gyldig formativ vurdering, vil avhenge av hvordan **konteksten** utvikler og utøver arbeidsmetodikken. Pålitelig og gyldig formativ vurdering blir dermed kontekstavhengig. I drøftingskapittelet har funksjonene som ble fremhevet av studiemiljøene, derfor fremstått som styrende parametere for fenomenforståelse (jf. 5.2 og 5.3.). Funnene antyder også at ved både SU og FJU ble formativ vurdering pålitelig og gyldig, når de formative funksjonene ble ivaretatt (jf. 5.2.1.1 og 5.2.1.2).

Som medierende samhandlingsredskap benyttet man **digitale mapper**. Disse ivaretok behovet for allmenn fremvisning (publisering) og sosial samhandling/tilbakemelding, og la til rette for hensiktsmessig struktur og organisering av studieoppgaver/responser. Funnene antyder dermed, at digitale mapper som medierende redskap for formativ vurdering, kan være hensiktsmessige enheter til fremvisning, samhandling, organisering og strukturering av studieoppgaver og responser.

En grunnleggende forutsetning for arbeidsmetodikken, var **kollektiv samhandling**. Det var her forståelse, utvikling og utøving av formativ vurdering ble synlig. I læringsprosessen ved de undersøkte studiemiljøene, sto derfor språk og kommunikasjon sentralt, likeså balansen mellom det individuelle og det kollektive. Læring var påvirket av både det enkelte individ, miljøet og lærere/medstudenter, og var således et resultat av deltaking i et sosialt læringsfellesskap, der samspill mellom mennesker var en vesentlig faktor. Læring var både et resultat av det som skjedde i hodet til den enkelte og påvirkninger fra omgivelsene. Funnene kan derfor antyde, at når denne arbeidsmetodikken tas i bruk, er det fornuftig å bygge den på **sosiokulturell** læringsteori.

I utvikling og utøvelse av formativ vurdering la fokusgruppene under intervjuene vekt på å beskrive hvordan formativ vurdering som arbeidsmetode, kunne stimulere til **påvirkning** og **fremvisning** av dyp læring. Innpasser man det i Herman & Winters sertifiseringsrammeverk, og trekker konklusjoner av studiemiljøenes uttrykte forståelse, er det nærliggende å si, at når arbeid med mapper knyttes til læringsaktivitet og ikke formell vurdering, bør **teknisk kvalitet** (jf. 2.5.1) dokumenteres med utgangspunkt i nøkkelidentifikatorer fremmet av Carney (2004), og ikke nøkkelidentifikatorer foreslått av Herman og Winters (1994) (jf. 2.5.1). Uansett om man tar utgangspunkt i Carneys eller Herman & Winters nøkkelidentifikatorer, når mappens tekniske kvalitet skal dokumenteres, vil perspektivet være ivaretagelse av studenten. Men studerer man forskjellen på de to, vil mapper når de brukes som grunnlag for **formell vurdering**, reise spørsmålene: Hvor mye kan man stole på (pålitelig), og hvor gyldig (valid) er den formelle vurderingen, mens mapper brukt i **pågående læring**, mer stiller spørsmålene: Hvordan stimulerer mappen/metoden til læring og læringsaktivitet og personlig utvikling.

De empiriske funn viser altså at ved både SU og FJU, ble kvalitetsforståelse knyttet til funksjoner som skulle ivaretas, i en læringsprosess preget av sosial samhandling over tid. Når mapper brukes i formativ vurdering ved disse miljøene, er det derfor hensiktsmessig å

dokumentere mappens **tekniske kvalitet** i lys av læring, og knytte dokumentasjonsinnhenting, til belysning av funksjonene man ønsker formativ vurdering skal fylle.

6.2 Videre forskning

Å dokumentere mappekvalitet, ble allerede i 1994 fremhevet av Herman & Winters (jf. punkt 2.5). Nettopp det har og vært hensikten med denne undersøkelse, der **ett** av områdene i Herman & Winters sertifiseringsrammeverk er belyst, nemlig **teknisk kvalitet**. Forskning knyttet til å undersøke **kvalitet i arbeid med mapper**, er det stort behov for, ikke minst når mappen blir digital (jf. Barrett 2005). I lys av pågående program for digital kompetanse (2004) og innføring av nye lærerplaner i grunn- og videregående skole fra høsten 2006, mener jeg de mest presserende forskningsoppgaver på grunn- og videregående utdanningsnivå i norsk skole, er knyttet til å undersøke hvordan IKT påvirker utdanningens kvalitet, motivasjon for læring, læringsformer og læringsutbytte. Digitale mapper inngår som naturlig del i det. Det hadde derfor vært interessant å undersøke innovative læringsmiljøers bruk av mapper både som vurderings- og læringsredskap.

I lys av teoretisk rammeverk for denne undersøkelse (jf. 2), kunne det være nærliggende å gripe fatt i de tre andre dokumentasjonsområdene i Herman & Winters (1994) sertifiseringsrammeverk, **mappers egnethet, ringvirkninger ved å ta i bruk og rettferdighet**. Spesielt kvalitetsdokumentasjon knyttet til mappers egnethet (jf. 2.5.4) er et aktuelt område for forskning. Men også forskning, knyttet til hvordan mapper kan påvirke og fremvise dyp læring på disse trinnene i norsk skole, er det behov for. Uansett er Herman & Winters sertifiseringsrammeverk like aktuell som utgangspunkt, om man ønsker å forske på mappekvalitet i lærings- eller vurderingsarbeid.

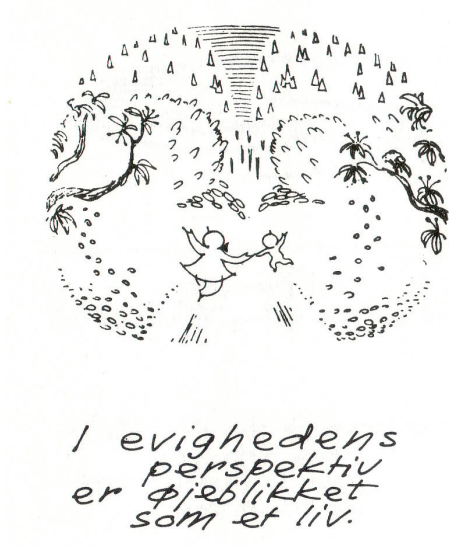
6.3 Avslutning

Som nybegynner i forskerfaget er erfaringen, at forskning er seriøst, utfordrende og spennende arbeid, som går sakte men sikkert fremover. Å holde fast i alle tråder samtidig, har ikke bestandig vært like lett, og respekten for forskning har vokst i takt med den utvidede og komplekse oppgave.

Til slutt i noen av oppgavens kapitler, finnes et lite gruk av Piet Hein, små visdomsord om livet, som kanskje kan tilføre den som leser dem et snev av livsrefleksjon og livskvalitet. Alle er vi vel til dels opptatt av at ting rundt oss skal preges av en viss form for kvalitet. Dog er det stor individuell forskjell på, hvordan begrepet oppfattes. Skal derfor kvalitet defineres og tolkes i en felles referanseramme, er det helt nødvendig med belysning og diskusjon av begrepet. Knytter man dette aspekt til denne undersøkelse, tilsier erfaringen at det tar tid å tilføre mapper kvalitet, og den blir til underveis i en prosess. Trekkes en konklusjon av dette, er kvalitet altså ikke et fastspikret og entydig begrep så snart det er definert, men

utvikles i en pågående prosess. Dette er en beskrivelse mange kan kjenne seg igjen i, når man ser postulatet i lys av den generelle samfunnsutviklingen ellers.

Program for digital kompetanse (2004) poengterer i sammendraget nødvendigheten av å vise frem suksessfaktorer, flaskehalsar og de gode eksemplene for utdanningssektoren. Skal mapper brukes som redskap i lærings- og vurderingsarbeid på ulike utdanningsnivåer, kan jeg kun gi min fulle støtte til det. Innføring av mappemetodikk medfører nye arbeidsmåter, elevroller, lærerroller og forskrifter. En fremvisning av suksessfaktorer, flaskehalsar og de gode eksemplene, vil bidra til at flere kan få en "flying start", og komme godt i gang med arbeidsmetodikken. En forutsetning er da, at det faktisk settes i gang forsøk med mappevaldering, knyttet enten til lærings- eller vurderingsarbeid, at forsøkene blir kritisk fulgt opp, og evalueres seriøst. Det er her forskeren kommer inn. Jeg ser frem til å følge utvikling og forskning innen dette feltet, både nasjonalt og World Wide!



(Hein 1968:49)

Figur 19: I evighetens perspektiv (Piet Hein)

Litteratur

Aae, T., Arntsen, R. N., Bratlie, Heidi C. M. & Hafnor, S. S. (2002). *Studenterfaringer med bruk av mappe som lærings- og vurderingsform. Innlegg på fagkritisk dag ved*

Høgskolen i Bergen. Retrieved 05/01, 2005

from: <http://www.hio.no/content/view/full/1051>

Avd. for Sykepleierutdanning (2005). *Undervisningsplan 2005-2006, deltid 1. år*. Retrieved

17.2.2006 from:

http://www.hio.no/content/download/45483/343233/file/1aar_020905.pdf

Avd. for sykepleierutdanning. (2005). *Fagplan (bachelor) for sykepleierutdanning ved*

Høgskolen i Oslo, 2005-2006. Retrieved 17.02.2006 from:

http://www.hio.no/content/download/44343/336982/file/Fagplan_trykt.pdf

Avdeling for journalistikk, bibliotek- og informasjonsfag (2005). *Studieplan for*

bachelorprogrammet i fotojournalistikk. Høgskolen i Oslo

Barrett, H. (2003). *Presentation at First International Conference on the e-Portfolio*, Poitiers,

France, October 9, 2003. Retrieved January 21, 2005 from:

<http://electronicportfolios.org/portfolios/eifel.pdf>

Barrett, H. & Carney, J. (2005). *Conflicting Paradigms and Competing Purposes in*

Electronic Portfolio Development. (Submitted to Educational Assessment, an LEA

Journal, for an issue focusing on Assessing Technology Competencies July 2005)

Barrett, H. C. (2005). *Researching Electronic Portfolios and Learner Engagement*. The

REFLECT Initiative Researching Electronic portFolios: Learning, Engagement and

Collaboration through Technology. White Paper: From:

<http://www.taskstream.com/reflect/whitepaper.pdf>

Baume, D. (2002). *A briefing on Key Concepts Formative and summative criterion norm-*

referenced assessment. Assessment Series no.7.

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning. Putting it into practice.*

Blair, L. K., & Takayoshi. (1994). *Reflections on Reading and Evaluating Electronic Portfolios.* In Black L. m. fl. ed. *New directions in Portfolio Assessment.* Boynton/Cook Publishers Heinemann. Portsmouth, NH.

Bøyesen, L., Sand, T., Øzerk, K. Z., Stålsett, U. E., & Grande, S. Ø. (1997). *Flerkulturell virkelighet i skole og samfunn.* Oslo: Cappelen akademisk forlag.

Carney, J. (2004). *Setting an Agenda for Electronic Portfolio Research: A Framework for Evaluating Portfolio Literature.* Presentation at the American Educational Research Association Conference, April 14, 2004. Retrieved January 22, 2005 from:
<http://it.wce.wvu.edu/carney/Presentations/AERA04/AERAresearchlit.pdf>

Carney, J. (2003). *Teacher Portfolio and New Technologies: Confronting the Decisions and Dilemmas.* Paper presented at the Society for Information Technology in Teacher Education annual conference, Albuquerque.

Carney, J. (2002). *The intimacies of electronic portfolio: Confronting preservice teachers' self-revelation dilemma.* Paper presented at the Society for Information Technology in Teacher Education annual conference, Nashville.

Carney, J. (2001). *Electronic and Traditional Paper Portfolios as Tools for Teacher Knowledge Representation.* Curriculum and Instruction. Seattle, University of Washington: 268

Davies, A. (2000). *Making Classroom Assessment Work.* Merville, BC: Connections Publishing

Det kongelige kirke-, utdannings- og forskningsdepartement (1996). *Læreplanverket for den 10-årige grunnskolen (L97).* Nasjonalt læremiddelsenter. From:
<http://www2.udir.no/L97/L97/index.html>

- Dunn, L., Parry, S., & Morgan, C (2002). *Seeking quality in criterion-references assessment*. Paper presented at Northumbria/EARLI Assessment Conference, 28th – 30th August.
- Dysthe, O. (1999). *Ulike teoriperspektiv på kunnskap og læring*. Bedre skole (3/99)
- Dysthe, O. & Breistein, S. (1999). *Fagskriving og rettleiing ved universitetet*. Bergen Universitet: Program for Lærerforskning, Rapport nr 3
- Dysthe, O., Engelsen, K. S., Madsen, T., & Wittek, L. (i produksjon). (2005). *Explicit criteria in portfolio assessment from student perspectives*. A theory based discussion In A. Havnes & L. McDowell (Eds.), *Assessment beyond intuition: raising standards of assessment and learning*. Book following the Assessment 2004 Northumbria/EARLI Assessment Conference
- Dysthe, O. & Engelsen, K. S (2003). *Mapper som lærings- og vurderingsform*. In O. Dysthe & K. S. Engelsen (Eds.), *Mapper som pedagogisk redskap* (pp. 13-33). Oslo: Abstrakt forlag.
- Dysthe, O. & Engelsen, K. S. (2003). *Mapper som pedagogisk redskap: Perspektiver og erfaringer*. Oslo: Abstrakt forlag.
- Dysthe, O. & Engelsen, K. S. (2003). *Digitale mapper ved to lærerutdanningsinstitusjoner*. In O. Dysthe, & K. S. Engelsen (Eds.), *Mapper som pedagogisk redskap* (pp 205-236). Oslo: Abstrakt forlag.
- Dysthe, O. (2002). *Mapper som lærings- og vurderingsreiskap*. Uniped, 2, 30-43.
- Dysthe, O. (2001). *Dialog, samspel og læring*. Oslo: Abstrakt forlag.
- Dysthe, O. & Lied, L. (2001). *Fagskriving og rettleiing ved universitetet. Intervjustudie ved Institutt for fiskeri- og marinbiologi* [Writing and supervision at the university: Interview study at the Department of Fishery and Marine Biology]. Bergen: Universitetet i Bergen.

- Dysthe, O. & Breistein, S. (1999). *Fagskriving og rettleiing ved universitet. Del 2: intervjustudie ved Institutt for administrasjon og organisasjonsvitenskap ved Universitetet i Bergen*. Universitetet i Bergen
- Ellmin, R. (2000). *Mappemetodikk*. Kommuneforlaget.
- Falchnikov, N. & Goldfinch, J. (2000). *Student peer assessment in higher education, a meta-analysis*. *Review of Educational Research*, 7
- Fog, J. (2004). *Med samtalen som utgangspunkt: Det kvalitative forskningsinterview* (2. rev. udg ed.). København: Akademisk Forlag.
- Hammersley, M. & Atkinson, P. (1987). *Feltmetodikk: grunnlaget for feltarbeid og feltforskning*. Oslo: Gyldendal.
- Hein, P. (1968). *I folkemunde. Korte gruk 1*. Herlev: Antikva Offset A/S.
- Herman, J. & Winters, L. (1994). *Portfolio research: A slim collection*. *Educational Leadership*, 52(2)
- Holter, H. & Kalleberg, R. (1996). *Kvalitative metoder i samfunnsforskning*. Oslo: Universitetsforlaget.
- Jacobsen, D. I. (2003). *Forståelse, beskrivelse og forklaring: innføring i samfunnsvitenskapelig metode for helse- og sosialfagene*. Kristiansand: Høyskoleforlag.
- Kariuki, M. T. (2001). *Creating electronic portfolios using laptops: A learning experience for preservice teachers, elementary school pupils, and elementary school teachers*. *Journal of Technology and Teacher Education*, 9(4), 567-584.
- Kvale, S (2001). *Det kvalitative forskningsintervju*. Oslo: Gyldendal Akademiske.
- Kvale, S. (2000). Eksamen som konstruksjon av kunnskap. *Uniped*, 3, 6-22.

- Launsø, L. & Rieper, O (2000). *Forskning om og med mennesker: Forskningsstyper og forskningsmetoder i samfunnsforskningen* (4. udg ed.). København: Nyt Nordisk Forlag Arnold Busck.
- Lauvås, P. *Vurdering for læring – viktigere enn eksamen (vurdering av læring)?* Retrieved 01.06.2005 from: <http://klaff.hiof.no/~pla/Formativ-vurdering-hiof.htm>
- Lindström, L., Ulriksson, L. & Elsner, C. (1999). *Portföljvärdering av elevers skapande i bild*. Utvärdering av skolan 1998 avseende läroplanernas mål (US98). Stockholm: Skolverket.
- Linnakylä, P. (2000). *Digital Portfolios in graduate studies. What we can learn by exploring and evaluating academic portfolios on the web?* Paper presented at the ECER 2000 conference in Edinburg, September 20-23, 2000 (unpublished)
- Madsen, T (2003). *Internasjonale erfaringer med mapper og mappevurdering i lærerutdanning*. In O. Dysthe, & K. S. Engelsen (Eds.): *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. (pp. 159-182). Oslo: Abstrakt forlag.
- McKinney, M. 1998). *Preservice Teachers` Electronic Portfolios: Integrating Technology, Self-Assessment, and reflection*, in Jones, A.J. ed. *Teaching Portfolios in Teacher Education*. *Teacher Education Quarterly*. Winter 1998. Caddo Gap Press. San Francisco, California, 25(1.)
- Mellin-Olsen, S. (1996). *Samtalen som forskningsmetode. Tekster om kvalitativ forskningsmetode som del av pedagogisk virksomhet*. Bergen. Caspar Forlag A/S.
- Moon, J. (1999). *Reflection in Learning and Professional Development*. London: Kogan Page.
- Moss, P. (1998). *Rethinking validity for the assessment of teaching. With portfolio in hand: Validating the new teacher professionalism*. N. Lyons. New York, Teachers College Press, 202-219.

- National Research Council. (2001). *Knowing what Students Know: The science and design of educational assessment*. Washington, D.C.: National Academy Press
- Norton, L., T. Dickens & N. McLaughlin Cook (1996): *Coursework assessment: What are tutors really looking for?* In: Gibbs, G. (ed): *Improving student learning. Using research to improve student learning*. Oxford: The Oxford Centre for Staff Development
- Office of Research. (1993). *Student Portfolios: Classroom Uses*. Retrieved 01.03.2006 from: <http://www.ed.gov/pubs/OR/ConsumerGuides/classuse.html>
- Otnes, H. (2003). *Arkivskuff eller læringsarena? Lærings- og dokumentasjonssjangrer i digitale mapper*. In O. Dysthe, & K. S. Engelsen (Eds.), *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. (pp. 85-110). Oslo: Abstrakt forlag.
- Otnes, H. (2004). *IKT og nye læreprosesser: En artikkelsamling basert på erfaringer fra et prosjekt ved avdeling for lærerutdanning*. Tønsberg: Høgskolen i Vestfold.
- Patel, R. & Davidson, B. (1995). *Forskningsmetodikkens grunnlag: Å planlegge, gjennomføre og rapportere en undersøkelse*. Oslo: Universitetsforlaget.
- Paulson, F. L. & Paulson, P. (Eds.). (1994). *Assessing Portfolios Using the Constructivist Paradigm*, in Fogarty, R. (ed.) (1996) *Student Portfolios*. Palatine: IRI Skyline Training & Publishing
- Paulson, F. L., Paulson, P. R. & Meyer, C. (1991). *What makes a portfolio a portfolio?* *Educational Leadership*, 48(5), 60-63.
- Rust, C., Price, M. & O'Donovan, B. (2003). *Improving students' learning by developing their understanding of assessment criteria and processes*. *Assessment and Evaluation in Higher Education*, 28(2)
- Sadler, R. (2001). *ASSESSMENT FOR LEARNING: A Hopeful Vision of the Future*. International Conference on Classroom Assessment

- Sadler, R. (1996). *Criteria and standards in student assessment. Different Approaches: Theory and Practice in Higher Education*. HERDSA Conference Perth, Western Australia, from: <http://www.herdsa.org.au/confs/1996/sadler.html>
- Sadler, R. (1989). *Formative assessment and the design of instructional systems*. *Instructional Science*, 18, 119-144.
- Sadler, R. (1983). *Evaluation and the improvement of academic learning*. *Journal of Higher Education* 54 (1), pp. 60-79.
- Shepard, L. A. (2001). *The role of classroom assessment in teaching and learning*. In V. Richardson (Ed.), *Handbook of research on teaching*, 4th edition. Washington, DC: American Educational Research Association.
- Sivertsen, E., Apneseth, O. & Voldgaard, J. (2004). *Evaluering av fotojournalist-utdanningen*. Høgskolen i Oslo, HiO-notat nr 8
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York: Merrill
- St.meld. nr. 27. (2000-2001). *Gjør din plikt - Krev din rett. Kvalitetsreform av høyere utdanning*. From: <http://odin.dep.no/kd/norsk/publ/stmeld/014001-040004/index-dok000-b-n-a.html>
- Taasen, I., Havnes, A. & Lauvås, P. (2004). In Havnes A., Lauvås P. (Eds.), *Mappevurdering - av og for læring: Med eksempler fra helse- og sosialfag*. Oslo: Gyldendal akademisk
- The Assessment Reform Group (2002). *Assessment for Learning: 10 principles. Research-based principles to guide classroom practice*. Retrieved 25.08.2005 from: <http://arg.educ.cam.ac.uk/CIE3.pdf>
- Topping, K. (2003). *Self and peer assessment in school and university; reliability, validity and utility*. In M. Seegers, F. Dochy & E. Cascallar (Eds.) *Optimising new modes of*

assessment: In search of qualities and standards. Dordrecht/Boston/London: Kluwer Academic Press.

Utdannings- og Forskningsdepartementet. *Program for digital kompetanse 2004-2008*.

Retrieved 10.05.2006 from:

http://odin.dep.no/filarkiv/201402/program_for_digital_kompetanse.pdf

William, D. (2000a). *Integrating formative and summative functions of assessment*. Paper presented at working group 10 of the international Congress on Mathematics Education, Makhuari, Tokyo

William, D. (2000b). *The meaning and consequences of educational assessment*. *Critical Quarterly*, 42(1), 105-127.

William, D. & Black, P. J. (1996). *Meanings and consequences: A basis for distinguishing formative and summative functions of assessment?* *British Educational Research Journal*, 22(5), 537.

Wilkerson, J. & Lang, W. S. (2003). *Portfolio, the pied piper of teacher certification assessments: Legal and psychometric issues*. Available online: <http://epaa.asu.edu/epaa/v11n45/>. *Education Policy Analysis Archives*, 11(45)

William, D. (1998). *Enculturating learners into communities of practice: raising achievement through classroom assessment*. Paper presented at European Conference on Educational Research, Ljubljana, Slovenia.

Winje, Geir (2004). *Organiske digitale mapper*. In H. Otnes (red.), *IKT og nye læreprosesser*. En artikkelsamling basert på erfaringer fra et prosjekt ved avdeling for lærerutdanning. (pp. 65 -79). Tønsberg: Høgskolen i Vestfold. Notat 2/2004

Wittek, L. & Dysthe, O. (2003). *Kvalitetskriterier og vurdering - et utviklingsområde innenfor høyere utdanning*. In O. Dysthe & K. S. Engelsen (Eds.), *Mapper som pedagogisk*

redskap. Perspektiver og erfaringer. [Portfolios as pedagogical tools] (pp. 129-156).

Oslo: Abstrakt forlag.

Wittek, L. & Havnes, A. (2005). *Vurderingsformer ved Høgskolen i Oslo*. Høgskolen i Oslo, HiO-notat nr 14

Wittek, L. (2004). *Mapper som lærings- og vurderingsform: Eksempler fra Universitetet i Oslo*. Oslo: Unipub.

Zeichner, K. & Wray, S. (2000). *The teaching portfolio in US teacher education programs: What we know and what we need to know*. 17 (2001) s. 613-621.

Zubizarreta, J. (2004). *The Learning Portfolio*. Bolton, MA: Anker Publishing.

Øhra, Mattias (2004). *Kunnskapsdeling ved bruk av digitale mapper*. In H. Otnes (red.), IKT og nye læreprosesser. En artikkelsamling basert på erfaringer fra et prosjekt ved avdeling for lærerutdanning. (pp. 15 -31). Tønsberg: Høgskolen i Vestfold. Notat 2/2004

Vedlegg

Vedlegg 1

Intervjuguide

A) Mapper sett i lys av reliabilitet, validitet og kvalitet.

1. Hvilken forståelse har dere av begrepet mapper?
2. Hvis dere skal si noe generelt om hva som er med på å gjøre mappebegrepet ved deres utdanning kvalitetssikker, hva vil dere trekke frem da?
3. Om du i en paneldebatt skulle argumentere for eller imot påliteligheten og gyldigheten av mapper som arbeidsform, hva ville du trekke frem da?
4. Pålitelighet hva favner det?
5. Gyldighet hva favner det?
6. Er dette relevante/ikke relevante begreper for dere?

Digitale mapper

1. Hvilken forståelse har du av begrepet digitale mapper?
2. Ser du noen forskjell på papirmapper og digitale mapper hva gjelder fenomenene pålitelighet, gyldighet og kvalitet?

B) Mappevurdering sett i lys av reliabilitet, validitet og kvalitet.

(Intensjonen er at mappevurdering skal være til å stole på og at den skal vurdere det som er målet for vurderingen)

1. Si litt om hva du knytter til ordet mappevurdering.
2. Ser du noen prinsipiell forskjell på begrepene mapper og mappevurdering?
3. Hvis du skal si noe generelt om kvalitet i forbindelse med mappevurdering ved hva vil du trekke frem da?
4. Er det relevant å knytte begreper som pålitelighet og gyldighet til det du trekker frem?
5. Om du i et foredrag skulle fortelle om mappevurdering som en pålitelig og målrettet vurderingsform (som vurderer det en ønsker skal være mål for vurderingen) hva ville du trekke frem da?

C) Formativ vurdering sett i lys av reliabilitet, validitet og kvalitet

1. Hvilken forståelse knytter du til begrepet formativ vurdering ved?
2. Ser du noen prinsipiell forskjell på mappevurdering og formativ vurdering?

3. Hvis du skal si noe generelt om hva som gir formativ vurdering et kvalitetsstempel, hva vil du da trekke frem?
4. Om du tenker på den formative vurderingsformen (e) slik den utøves ved din utdanning, ville det da være relevant å knytte den til frasen:
"Formativ vurdering er en vurderingsform som er pålitelig og målrettet"?
5. Hva preger en pålitelig formativ vurdering?
6. Hva preger en målrettet formativ vurdering?
7. Hvordan det arbeides det på din utdanning med å utvikle formativ vurdering som er pålitelig og som vurderer det som er målet?
8. Hvilket inntrykk har du av den reelle erfaring med formativ vurdering som pålitelig og målrettet?
9. Er det noe du spesielt vil trekke frem som fremmer/hemmer pålitelig og målrettet formativ vurdering i arbeid med mapper?

D) Kriterier sett i lys av reliabilitet, validitet og kvalitet

1. Hvilken forståelse knytter du til begrepet kriterium?
2. Mener du det er viktig at kriterier blir presentert skriftlig for studenter og lærere? Evt hvorfor?
3. Hvis du skal definere frasene "Et kriterium som er pålitelig" og "Et kriterium som er relevant" hva vil du i hovedsak trekke frem?
4. Hvem bestemmer om et kriterium er pålitelig og relevant/gyldig? Hvorfor?
5. Kan du gi meg gode eksempler på hvordan det ved din utdanning utvikles og brukes kriterier som er til å stole på og som har gyldighet/relevans! Hva er ditt inntrykk av den reelle erfaringen?
6. Hvis du skal si noe om hva som gir et kriterium kvalitet, hva vil du da trekke frem?

E) Total vurdering sett i lys av reliabilitet, validitet og kvalitet

1. Hvilke faktorer ser du som de aller viktigste for at mappevurdering, der det inngår formativ vurdering og bruk av kriterier, skal bli både pålitelig og gyldig?
2. Hvilke faktorer ser du som de aller viktigste for å sikre kvalitet i mappevurdering der det inngår formativ vurdering og bruk av kriterier?

Vedlegg 2

Intervjudokument: Notater

Sted:		
Dato:		
Tidspunkt:		
Hvem intervjues:		
Lærere:		
Studenter:		
Andre kommentarer:		
Intervjuer	Intervjuguide pkt	Intervjuperson

Oppsummeringsdokument: Egne kommentarer

Sted:	
Dato:	
Tidspunkt:	
Intervjuperson	Intervjuet: (Forståelig? Nødvendig? Eget hovedinntrykk)

Egne refleksjoner over intervjuesituasjonen:

Hvordan opptrådte respondentene?	
Hvordan fungerte jeg som intervjuer i situasjonen? (God tone, kom vi skjevt ut etc)	
Hvordan fungerte selve stedet der intervjuet foregikk? (Kunstig, ok plass etc)	
Kan selve situasjonen ha påvirket innholdet i det som kom frem? (I hvor stor grad kan vi stole på opplysningene som er kommet frem – reliabilitet.	
Hvordan fungerte videoopptaket?	
Annet	

Vedlegg 3

Test intern gyldighet ved hjelp av respondentvalidering

Tema	kjenner meg igjen	kjenner meg ikke igjen	Utdyping
Mappe			
Formativ vurdering			
Kriterier			
Formativ vurdering sett i lys av reliabilitet og validitet			
Kriterier sett i lys av reliabilitet og validitet			
Mapper sett i lys av kvalitet			

Vedlegg 4

Vurderingskriterier ved sykepleierutdanningen, Høgskolen i Oslo

Hver studieoppgave har eksplisitte vurderingskriterier, slik at studentene har innsikt i hva det forventes at de skal vektlegge i studieoppgavene. Gjennom bruk av vurderingskriteriene i studieoppgavene blir studentene også kjent med de vurderingskriterier som anvendes i den summative vurderingen. Vurderingskriteriene vil påvirke hvordan utformingen av teksten vil være.

1. Viser faglig innsikt
2. Anvender anerkjent og relevant Litteratur
3. Belyser sentrale begreper
4. Integrerer teori og praksis
5. Reflekterer over egen og/eller andres praksis
6. Viser forståelse for ulike innfallsvinkler
7. Argumenterer for egne meninger
8. Viser evne til systematisk tenkning
9. Er presis og nøyaktig
10. Vekker interesse og motivasjon hos mottaker
11. Analyserer og tolker på bakgrunn av faglig vurderinger
12. Anvender relevant forskningslitteratur
13. Anvender kunnskap fra ulike fagområder
14. Reflekterer over teori/og eller aktuell situasjon
15. Er kritisk og konstruktiv
16. Viser evne til kreativitet og selvstendighet

Vedlegg 5

Vurderingskriterier ved fotojournalistutdanningen, Høgskolen i Oslo
(For faget bildeteori)

Kriterier for vurdering av en oppgave:

God disposisjon

Klar problemstilling med avgrensning av oppgaven

Klart, presist språk med relevant begrepsbruk

Kildekritikk

Bredde i kildebruk

Selvstendig drøfting av materialet

En normal disposisjon inneholder:

Innledning (Med problemstilling, avgrensning og definisjon)

Faktadel

Drøfting

Konklusjon

Litteraturliste

I tillegg er det TO KRAV som er absolutte:

Stoff dere bruker skal refereres til, og sitater skal stå i anførselstegn.