

– De la veldig mye vekt på det derre «rette med grønt» da...

*En kvalitativ studie av hva slags opplæring i elevvurdering
fire norske allmennlærerstudenter som avsluttet
utdannelsen i 2007 fikk, sammenlignet med fire
lærerstudenter utdannet i samme periode
i Queensland, Australia*

Sverre Tveit



Masteroppgave ved Pedagogisk forskningsinstitutt,
Utdanningsvitenskapelig fakultet

UNIVERSITETET I OSLO

5. februar 2009

TITTEL:**- De la veldig vekt på det derre «rette med grønt» da...**

En kvalitativ studie av hva slags opplæring i elevvurdering fire norske allmennlærerstudenter som avsluttet utdannelsen i 2007 fikk, sammenlignet med fire lærerstudenter utdannet i samme periode i Queensland, Australia.

AV:

Sverre TVEIT

EKSAMEN:

Master i pedagogikk
Studieretning: Allmenn

SEMESTER:

Vår 2009

STIKKORD:

Elevvurdering, karaktersetting, vurdering, summativ, formativ, lærerutdanning, teoriopplæring, praksisopplæring

Forord

Da jeg for to år siden begynte å planlegge denne masteroppgaven, hadde jeg allerede arbeidet i et år med å redigere boken *Elevvurdering i skolen – grunnlag for kulturendring*. Det var derfor naturlig for meg å fortsette å arbeide med dette fagfeltet innenfor pedagogikken. Samtidig ønsket jeg å utforske mer ukjent terreng, for å utvide min egen forståelseshorisont og for å bidra til ytterligere kunnskapsproduksjon innenfor fagfeltet. Valget falt på å gjennomføre en komparativ studie av elevvurdering i allmennlærerutdanningen i Norge og en lærerutdanning i Queensland, Australia. Det har alltid tiltalt meg å reise og utforske andre kulturer, og som følge av at jeg hadde blitt kjent med vurderingssystemet i denne delstaten gjennom det nevnte bokprosjektet, var Queensland et naturlig valg.

Steinar Kvale har i sine lærebøker i forskningsmetode brukt det tyske begrepet *bildungsreise* – dannelsesreise – til å illustrere sin forskningsstrategi for kvalitative intervjuer. For meg har dette masteroppgaveprosjektet vært en slik dannelsesreise på minst to måter: Først ved å dra til Australia og utforske en svært annerledes kultur for opplæring generelt, og elevvurdering spesielt. Deretter ved at jeg i Norge og Queensland for første gang har hatt grundige samtaler med representanter fra praksisfeltet. For meg, som med bakgrunn fra Elevorganisasjonen har mest kjennskap til politiske prosesser, forskrifter og retningslinjer, forskning og annen litteratur om elevvurdering, har dette vært en svært lærerik reise. Jeg er svært takknemlig overfor de åtte lærerstudentene som gjorde denne studien mulig.

Jeg vil også rette en stor takk til min veileder Kirsti Klette, for kyndig veiledning og støtte gjennom denne nesten to år lange skriveprosessen. Gjennom flere møter i Oslo og San Francisco, og gjennom korrespondanse på telefon og e-post, har hun delt med meg fra sin faglige og metodiske ekspertise. Jeg vil også takke min mor, som har vært til stor hjelp i forbindelse med korrektur og innsamling og forsending av litteratur fra Norge til Australia og Tyskland. Takk også til mine samboere i diverse bofellesskap, først på Gold Coast og i Brisbane, og nå under mitt utvekslingsopphold fra Australia, i Tübingen i Tyskland, for deres tålmodighet og overbærenhet med mine studievener. En rekke andre personer – både i Norge og Queensland – kommer jeg til å rette en personlig takk til.

Selv opplever jeg å returnere til Norge med en langt bedre forståelse av vår utdannings- og vurderingskultur enn jeg hadde da reisen startet. Jeg håper jeg med denne masteroppgaven har lyktes i å dele disse erfaringene med andre.

Tübingen, Tyskland, 5. februar 2009

Sammendrag

Problemstilling

Hva slags opplæring i elevvurdering fikk norske allmennlærerstudenter som avsluttet utdannelsen i 2007, sammenlignet med lærerstudenter utdannet i samme periode i Queensland, Australia?

Metode

Kvalitativ casestudie med intervju av åtte lærestudenter fordelt på casene Norge og Queensland. To studenter med fordypning i morsmålsfaget, og to med fordypning i matematikk, ble intervjuet begge steder. Dokumentanalyse ble gjennomført for å kartlegge casenes kontekst.

Dokumentgrunnlag i Norge: Læreplaner, stortingsvedtak, stortingsmeldinger, lover, forskrifter, rammeplan for allmennlærerutdanningen, høyskolens program for allmennlærerutdanningen og tilhørende fagplaner og relevant pensum.

Dokumentgrunnlag i Queensland: Læreplaner; akkrediteringsprosedyrer for lærerutdanninger, lærerregistreringsprosedyrer, «profesjonelle standarder» for lærere, universitetets programplaner for utdanning av *secondary*-lærere med tilhørende fagplaner og relevant pensum.

Hovedkonklusjoner

Studien avdekket omfattende organisatoriske ulikheter mellom de to lærerutdanningene, som framholdes som en av forklaringene på de store forskjellene i lærerstudentenes kunnskaper om elevvurdering. Den norske allmennlærerutdanningen var rettet mot hele grunnskolen. Studentene hadde fått opplæring i tre fag (*Norsk, Matematikk og KRL*) i løpet av de første to årene, og deretter valgt mellom to og fem undervisningsfag de siste to årene. I Queensland var lærerutdanningen delt mellom *primary* og *secondary*. *Secondary*-studentene valgte to undervisningsfag som de fikk opplæring i gjennom alle fire årene.

Lærerne i Queensland anses i langt større grad for å være *faglærere*, mens «klasselæreren», som kan undervise i en bredde av fag og gjennom hele grunnskolen, har vært en bevisst distriktpolitisk prioritering i Norge. Studien dokumenterer hvordan de organisatoriske rammene i Queensland bidro til at man hadde mer tid til både faglig fordypning innenfor de to undervisningsfagene og i generell pedagogikk, herunder elevvurdering, sammenlignet med utdanningen i Norge.

Det er ventet at man i Norge innfører en todeling av allmennlærerutdanningen med spesialisering mot barnetrinnet og ungdomstrinnet, der ungdomsskolelærerutdanningen også skal være rettet mot mellomtrinnet. Studien anskueliggjør mulige fordeler ved en langt mer vidtrekkende reform av allmennlærerutdanningen, der utdanning av ungdomsskolelærere og lærere for videregående skole samordnes på universitetene, slik det gjøres i Queensland. Videre framholdes Queenslands modell for akkreditering av lærere og lærerutdanninger som en interessant måte å kvalitetssikre innholdet, og ikke bare omfanget, av fagene på lærerutdanningene. På denne måten kunne man blant annet kontrollert om alle lærerstudenter får tilstrekkelig opplæring i elevvurdering.

Studien viser at de norske lærerstudentene som ble intervjuet gjennomgående hadde fått lite teori- og praksisopplæring i elevvurdering. Mens temaet i liten grad var tilstede i fag og lærebøker i Norge, så hadde de i Queensland et obligatorisk fag om elevvurdering med tilhørende pensumlitteratur, i tillegg til at temaet ble tatt opp grundigere i forbindelse med læreplananalyser og planlegging av undervisningsøkter i undervisningsfagene. Alle de norske studentene etterlyste opplæring i karaktervurdering, og ga uttrykk for at de ikke var tilstrekkelig forberedt til å håndtere karaktersetting ved inngangen til læreryrket. I Queensland fikk studentene mer erfaring med karaktersetting gjennom praksis, men også her var det dette studentene i størst grad etterlyste mer opplæring i.

Funnene gir grunnlag for å reise spørsmål om det legges nok vekt på karaktersetting i teoriopplæringen i de to utdanningene. Fraværet av karaktersetting i barneskolen, og den ideologiske dragkampen som har pågått i Norge siden 1970-årenes karakterstrid, framholdes som mulige årsaker til det tilsynelatende fraværet av fokus på karaktervurdering i norsk allmennlærerutdanning. Oppgavens tittel er utdrag fra en av de norske studentenes svar på spørsmålet om hva slags pensumlitteratur de hadde hatt om elevvurdering. Tre av de norske studentene trakk fram en læreboks anvisninger om å «rette med grønt», i stedet for rødt. De forklarte at uttrykket ga anvisninger om å gi mer positive «grønne» tilbakemeldinger, i stedet for å bruke den tradisjonelle røde korrigerende pennen.

Det er godt dokumentert at generell ros og oppmuntring har liten effekt, og norske lærere har tidligere blitt kritisert for å i stor grad gi slike tilbakemeldinger. To studenter så ut til å være kjent med denne kritikken, mens de andre var opptatt av å gi «positiv feedback hele tiden». Studien gir grunnlag for å reise spørsmål om alle i norsk allmennlærerutdanning har tatt inn over seg denne kritikken. Selv om det isolert sett er bra at formativ vurdering ble sterkt vektlagt ved den norske allmennlærerutdanningen, er det betenkelig at studentene hadde lært *mest* om, var det bare halvparten som hadde oppdatert kunnskap om. En mulig grunn til at de norske lærerstudentene

festet seg ved metaforen «rette med grønt», kan være at de generelt ikke hadde lært mye om elevvurdering. Studentene i Queensland, fortalte til sammenligning mye om hva de hadde lært om utforming og bruk av vurderingskriterier.

Et paradoks avtegnet seg i teori- og analyse materialet, som gir grunnlag for å utfordre en etablert oppfatning. Mange framholder at karakterer og belønninger i liten grad bidrar til, eller kan stå i veien for, forbedrede læringsresultater. Man sier gjerne at de summative vurderingene «overkjører» de formative vurderingene. Bildet som avtegnet seg i Queensland var helt annerledes, ved at kriteriene for summativ vurdering også dannet utgangspunktet for den formative vurderingen. Innenfor en standard- og kriteriebasert tradisjon er det naturlig at også formative vurderinger forankres i vurderingskriteriene som legges til grunn ved karaktervurdering, slik at elevene lettere kan forstå hva de må gjøre for å oppnå bedre læringsresultater. Således kan det hevdes at man i det formative vurderingsarbeidet i norske skoler, særlig på barnetrinnet, lider av en svak eller fraværende summativ vurderingskultur, når man skal gjennomføre standards- eller kriteriebasert vurdering.

Implikasjoner

Studien har vist at både i Norge og Queensland fikk studentene lite eller ingen trening i karaktersetting på høyskolen eller universitetet. Hvordan man kan gi lærerstudenter bedre opplæring i karaktersetting framholdes som et viktig område å utforske.

Studien har vist at systemer for akkreditering av lærere og lærerutdanninger kan gi mer forutsigbare rammer for lærerutdanningen. Det anbefales å undersøke nærmere hvordan et slikt akkrediteringssystem kan bidra til kvalitetssikring av innholdet, og ikke bare omfanget, av fagene på lærerutdanningene.

Studien har også vist hvordan den norske allmennlærerutdanningen bar preg av at elevvurdering som fagfelt er svakt utviklet i Norge. Utvikling av formelle vurderingskriterier som kan danne grunnlag for både summativ og formativ vurdering, framholdes som et viktig område for videre forskning.

Executive summary

Research question

What training in student assessment did Norwegian teacher students who graduated in 2007 undertake, compared to teacher students trained in the same period in Queensland, Australia?

Method

Qualitative case study with interviews of eight teacher students in Norway and Queensland. Two students specialised in the mother tongue subject and two specialised in mathematics subject were interviewed both places. Document analyses were undertaken to identify the context of the two cases.

Documents from Norway: Curriculums for *Norwegian language* and *Mathematics*, parliament decisions, «white papers» to the parliament, legislations and regulations, framework for general teacher-training (*allmennlærerutdanningen*), the specific teacher education's training program, courses and readings.

Documents from Queensland: Curriculum for *English* and *Mathematics*, accreditation procedures for teacher education, the specific teacher-training programs's teaching plans and readings, teacher registration procedures, «professional standards» for teachers, the specific teacher education's training program, courses, and readings.

Conclusions

The study uncovered comprehensive organisational differences between the two teacher-training programs, which is emphasised as one of the explanations for the big differences in the students' knowledge and skills in student assessment. The Norwegian general teaching training program (*allmennlærerutdanningen*) was orientated towards all the ten compulsory years of elementary education. The teachers were trained in three compulsory subjects (*Norwegian language*, *Mathematics*, and *Christianity, Religion and Ethics Education*) the two first years, and could then choose two to five electives the last two years. In Queensland the teacher-training program was differentiated between *primary* and *secondary* education. The secondary students choose two teaching subjects, which they were trained in throughout their four year teaching degree.

The teachers in Queensland are seen more as «subject content teachers», whereas the «class teacher», who can teach in a range of subjects at any level of the ten years of elementary education, for district political considerations, have been a political priority in Norway. The study

reveals how specialisation towards subjects and age levels in Queensland provide more time for content studies and general pedagogy subjects, such as student assessment, compared to Norway.

It is expected that the Norwegian government will introduce a differentiated model to replace the current general teacher-training education. The new model will probably have specialisation towards *primary* (Year 1-7) and *lower secondary* (Year 8-10), where the latter also provides training for the *middle years* (5-7). The study illustrates possible benefits with a more far-reaching reform of the general teacher education program, where a combined teaching training program for *upper* and *lower secondary level* is placed at universities (as opposed to the university colleges), as it is done in Queensland. Furthermore, it is argued that the Queensland model for accreditation of teachers and teacher-training programs could be one way to provide quality insurance of not only the quantity, but also the quality, of the teacher-training subjects. This way it could be verified that all teacher students get sufficient training in student assessment.

The study has revealed that the interviewed teacher students from Norway had received little theory- and practical training in student assessment throughout their education. The topic was not much emphasised in the curriculum subjects and text books in Norway whereas Queensland had a compulsory course on student assessment with accompanying literature. Additionally, the Queensland students experienced that the topic was emphasised more thoroughly in the curriculum subjects when planning teaching units. All the Norwegian students called for training in grading students and expressed that they were not sufficiently prepared to handle grading when they entered the teaching profession. In Queensland the students experienced more grading training through practice. However, training in grading was what called the most for also for these students.

The findings provide a basis for questioning whether grading is emphasised enough in the two teacher-training programs. The absence of formal grading in Norwegian primary schools and the ideological rather than informed debate over student assessment, which has characterised the debate over student assessment in Norway since the 1970s, is proposed as possible explanations for what seems to be an absence of focus on grading in Norwegian general teacher education. The title of this dissertation «Well, they emphasised 'green marking' a lot...» is a quote from a Norwegian student's response to the question on what readings on student assessment they had been through. Three of the Norwegian students mentioned a textbook's advice on «marking with 'green'», instead of red. They understood the metaphor as suggesting teachers should provide more positive «green» feedback, contradictory to the traditional red marking pen.

It is well documented that general praise has low effect on students' achievements, and research has previously criticised Norwegian teachers for such practices. Two students appeared to be familiar with this criticism whereas the others emphasised providing «positive feedback all the

time». Based on this finding, one could ask whether one in all teacher education programs in Norway has acknowledged this critique. A possible reason for why the Norwegian students remembered this metaphor could be that the student assessment in general was not thoroughly discussed throughout their education. Contradictory, the students in Queensland explained extensively how they had learned to develop and use assessment criteria.

A paradox evolved from the theory- and analysis material, which provides the basis for challenging established concepts. It is often argued that grades only to a limited extent is supportive to learning, and that it in some cases can harm the learning process. It is argued that the summative assessment stands in the way for the formative assessments. What can be seen in Queensland was quite different, in that the criteria for summative assessment also formed the basis for formative assessment. Within a standard- or criterion-referenced tradition it is natural that also the formative assessment is constituted by the criteria for summative assessment, enabling the students to understand what they need to do to achieve better learning outcomes. Accordingly, it can be argued that the formative assessment practices in Norwegian schools, particularly in primary years, suffer from a weak or non-existing culture for summative assessment, when conducting standard- or criterion-referenced assessment.

Implications

The study revealed that both in Norway and Queensland, the students undertook little or no training at the university or university college in grading student work. How students can receive better training in grading is suggested as an important area for future research.

The study revealed how a system for teacher and teacher education accreditation can provide more a predictable context for teacher education. It is suggested to investigate how such an accreditation system possibly could contribute to quality insurance of the content, not merely the extent, of the subjects provided in teacher education programs.

The study also revealed how the Norwegian general teacher-training program was affected by student assessment being a weakly developed research field in Norway. The establishment of formal assessment criteria which can form the basis for both summative and formative assessment is suggested as an important area for future research.

Innholdsfortegnelse

FORORD	III
SAMMENDRAG	V
EXECUTIVE SUMMARY	VIII
INNHOLDSFORTEGNELSE	XI
LISTE OVER TABELLER OG FIGURER	XVI
1 INNLEDNING	1
1.1.1 Hvorfor utforske elevvurdering?	2
1.1.2 Hvorfor utforske allmennlærerutdanningen?	4
1.1.3 Hvorfor komparere mot Queensland?	4
1.2 Forskningsspørsmål	5
1.3 Oppgaveavgrensning	6
1.3.1 Avgrensning av elevvurderingsfeltet	6
1.3.2 Lærerutdanningsinstitusjoner og aldersgrupper i Norge og Queensland	7
1.3.3 Elevvurdering og fag	8
1.3.4 Datagrunnlag	8
1.4 Oppgavens struktur og komposisjon	9
2 INTRODUKSJON TIL OG INNRAMMING AV CASENE: ELEVVURDERING I LÆRERUTDANNINGENE I NORGE OG QUEENSLAND	11
2.1 Norge	12
2.1.1 Norges utdanningssystem- og forvaltning	12
2.1.2 Regulering av læreres arbeid med elevvurdering	13
2.1.3 Organisering og innhold i norsk allmennlærerutdanning	20
2.2 Queensland	22
2.2.1 Queenslands utdanningssystem og -forvaltning	22
2.2.2 Regulering av læreres arbeid med elevvurdering	23
2.2.3 Organisering og innhold i Queenslands lærerutdanning	29

2.3	Sammendrag av hovedforskjeller mellom Norge og Queensland	31
3	TEORETISKE PERSPEKTIVER PÅ ELEVVURDERING	35
3.1	Diskusjon av kjernebegreper i vurderingsterminologien	36
3.2	Utfordringer ved summative vurderinger	39
3.2.1	Validitet og reliabilitet	40
3.2.2	Studier av validitet og reliabilitet	42
3.2.3	Oppsummerende oversikt over krav til summativ vurdering	43
3.3	Utfordringer ved formative vurderinger	44
3.3.1	Bruk av faglige tilbakemeldinger	45
3.3.2	Formativ vurdering og kriterier	47
3.3.3	Oppsummerende oversikt over krav til formativ vurdering	48
3.4	Hva vet vi om elevvurdering i Norge?	48
3.4.1	Utviklingstrekk knyttet til elevvurdering siste århundre	49
3.4.2	Summativ vurdering i Norge	53
3.4.3	Formativ vurdering i Norge	58
3.5	Noen erfaringer med elevvurdering i Queensland	61
3.5.1	Erfaringer med modereringsprosedyrene	61
3.5.2	Problematiske aspekter ved standardbasert vurdering	62
3.6	Fagspesifikke problemstillinger knyttet til elevvurdering	63
3.6.1	Elevvurdering i morsmålsfaget	63
3.6.2	Elevvurdering i matematikkfaget	66
3.7	Elevvurdering og lærerutdanning	68
3.7.1	Betydningen av gode lærere og god lærerutdanning	68
3.7.2	Opplæring i elevvurdering i lærerutdanningen	69
3.7.3	Andre lærerutdanninger i Norden	70
3.7.4	Erfaringer med norsk allmennlærerutdanning	71
3.8	Sammendrag av teoretiske perspektiver på elevvurdering	73
4	METODE FOR GJENNOMFØRING AV STUDIEN	77
4.1	Valg av fokus for oppgaven og forberedende arbeid for studien	77
4.2	Validitet og reliabilitet	78

4.3	Valg av forskningsmetodisk tilnærming	79
4.3.1	Fenomenologi	80
4.3.2	Hermeneutisk fortolkning	81
4.4	Metodiske avveininger i forbindelse med kvalitative intervjuer	81
4.4.1	Meningsproduksjon	82
4.4.2	Intervjustrategi	83
4.4.3	«Tilsiktet naivitet»	84
4.5	Metodiske avveininger i forbindelse med case studier	84
4.5.1	Deskriptiv og eksplorerende studie	85
4.5.2	Forskningsdesign	86
4.5.3	Forskningsspørsmål	89
4.6	Utarbeiding av intervjuguide	90
4.7	Utvalg og samplingskriterier	91
4.8	Intervjuene i Norge	91
4.8.1	Utvalg og sampling	91
4.8.2	Pilotering	92
4.8.3	Intervjusituasjonen	93
4.9	Intervjuene i Queensland	93
4.9.1	Utvalg og sampling	94
4.9.2	Pilotering	95
4.9.3	Intervjusituasjonen	95
4.10	Transkribering av data	95
4.11	Analyse av data	96
4.11.1	Bakgrunnsinformasjon	96
4.11.2	Organisering av opplæring og praksis i undervisningsfagene	96
4.11.3	Organisering av innhold knyttet til elevvurdering i lærerutdanningen	96
4.11.4	Tilegnelse av kunnskaper om og erfaringer med elevvurdering	97
4.11.5	Synspunkter og erfaringer	98
4.12	Framstilling av data	99
4.13	Sammendrag av metodiske vurderinger	100
5	HVA SLAGS OPPLÆRING I ELEVVURDERING HADDE LÆRERSTUDENTENE I NORGE OG QUEENSLAND FÅTT?	103

5.1	Organisering av lærerutdanningens teori- og praksisopplæring	103
5.1.1	Organisering av teoriopplæringen	103
5.1.2	Organisering av praksis	104
5.1.3	Studentenes erfaringer med praksis i undervisningsfagene	105
5.1.4	Kvantitative forskjeller i teori- og praksisopplæringen	106
5.2	I hvilket omfang og på hvilken måte møte studentene temaet elevvurdering?	107
5.2.1	Eget fag om elevvurdering	108
5.2.2	Vektlegging av elevvurdering i undervisningsfagene	108
5.2.3	Opplæring i elevvurdering i pedagogiske fag	109
5.2.4	Pensumlitteratur om elevvurdering	110
5.3	Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer	111
5.3.1	Kunnskapstilegnelse om elevvurdering gjennom teori- eller praksisbasert opplæring?	112
5.3.2	Opplæring i bruk av verktøy og hjelpemidler for elevvurdering	113
5.3.3	Opplæring i og erfaringer med bruk av vurderingskriterier	114
5.3.4	Opplæring i og erfaringer med å diskutere vurderingskriterier	117
5.3.5	Opplæring i karaktersetting	119
5.3.6	Opplæring i å gi faglige tilbakemeldinger	121
5.3.7	Opplæring i forskrifter og retningslinjer	122
5.4	Sammenfattende sammenligning	123
5.4.1	Organisering av lærerutdanningens teori- og praksisopplæring	123
5.4.2	I hvilket omfang og på hvilken måte møte studentene temaet elevvurdering?	124
5.4.3	Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer	125
6	DRØFTING AV SENTRALE FUNN	129
6.1	Betydningen av organiseringen av lærerutdanningens teori- og praksisopplæring	130
6.2	I hvilket omfang og på hvilken måte møter studentene temaet elevvurdering i lærerutdanningen?	134
6.3	Hva slags teori- og praksisopplæring fikk studentene i utvalgte temaer knyttet til elevvurdering?	136
6.3.1	Opplæring i bruk av verktøy og hjelpemidler for elevvurdering	137
6.3.2	Opplæring i bruk av vurderingskriterier	138
6.3.3	Erfaringer med diskusjoner om bruk av kriterier	139
6.3.4	Opplæring i og erfaringer med karaktersetting	140
6.3.5	Opplæring i og erfaringer med å gi faglige tilbakemeldinger	140
6.3.6	Opplæring i forskrifter og retningslinjer	141
6.4	Forskjeller mellom morsmåls- og matematikkfaget	143

6.5	Overordnet sammenligning av opplæring i elevvurdering i Norge og Queensland	144
7	AVSLUTTENDE KOMMENTARER	147
8	REFERANSER	151
9	APPENDIKSER	161
9.1	Læreplan Norsk, Norge	161
9.2	Læreplan Matematikk, Norge	171
9.3	Læreplan English Year 1-10, Queensland: Eksempel på «level statements» for «writing and shaping – level 4»	181
9.4	Læreplan Mathematics Year 1-10, Queensland: Eksempel på «level statements» for «measurement» for level 4, 5 og 6	184
9.5	Læreplan English Year 1-10, Queensland: Eksempler på «ways to gather and record evidence from a variety of sources»	185
9.6	Læreplan Mathematics Year 1-10, Queensland: Eksempler på «ways to gather and record evidence from a variety of sources»	186
9.7	Læreplan English Senior, Queensland: Achievement levels	187
9.8	Læreplan Mathematics B, Queensland: «Minimum standards associated with exit criteria»	190
9.9	Læreplan English Senior, Queensland: Krav til «verification folio»	191
9.10	Program Consultation and Acceptance Process, Board of Teacher Education, Queensland	195
9.11	Intervjuguide Norge	196
9.12	Intervjuguide Queensland	202
9.13	Eksempel på framstilling av data (spørsmål 11)	207
9.14	Vurderingsoppgaver- og kriterier i faget Norsk 1 ved den undersøkte norske allmennlærerutdanningen 216	
9.15	Vurderingsoppgaver- og kriterier I faget English Curriculum A ved den undersøkte lærerutdanningen i Queensland	219

Liste over tabeller og figurer

TABELL 1: OBLIGATORISKE FAG I ALLMENNLERERUTDANNINGS FØRSTE TO ÅR GJELDENE FRA 2003	21
TABELL 2: EKSEMPEL PÅ STANDARDBESKRIVELSER FRA LÆREPLANEN I ENGLISH	27
TABELL 3: OBLIGATORISKE FAG I DEN FIREÅRIGE SECONDARY-LÆRERUTDANNINGEN SOM ER UTFORSKET	31
TABELL 4: REGULERING AV LÆRERES ARBEID MED ELEVVURDERING	32
TABELL 5: ORGANISERING OG INNHOLD I LÆRERUTDANNINGENE	33
FIGUR 1: ILLUSTRASJON AV ANALYSEENHETER OG REPLISERINGS- OG KONTRAKSJONSSTRATEGIER	88
TABELL 6: EKSEMPLER PÅ VALGT FAGSAMMENSETNING VED NORSK ALLMENNLERERUTDANNING	104
TABELL 7: EKSEMPLER PÅ LÆRERUTDANNINGSPROGRAM FOR ENGLISH OG SOCIAL SCIENCE/HISTORY	104

1 Innledning

Denne masteroppgaven er en kvalitativ komparativ intervjustudie om elevvurdering i lærerutdanningen i Norge og Queensland. Hovedproblemstillingen for oppgaven er:

Hva slags opplæring i elevvurdering fikk norske allmennlærerstudenter som avsluttet utdannelsen i 2007, sammenlignet med lærerstudenter utdannet i samme periode i Queensland, Australia?

Grunnlaget for valg av problemstilling er et mangeårig engasjement for temaet elevvurdering, som jeg har behandlet i flere publikasjoner (Tveit, 2003, 2007b, 2007c, 2008a, 2008b). Mine kunnskaper om de formelle retningslinjene, og den politiske konteksten disse er blitt utviklet innenfor, gjorde meg nysgjerrig på hvordan det forholder seg i lærerutdanningsfeltet så vel som yrkesprofesjonsfeltet. Jeg mente det var grunnlag for å reise spørsmål ved læreres forutsetninger, spesielt for å gjøre rettferdige karaktervurderinger, men også de formative aspektene ved elevvurdering. I Stortingsmelding nr. 16 (2006-2007) *...og ingen stod igjen* slås det fast at forskning på praksis knyttet til elevvurdering er mangelfull. Jeg la til grunn at allmennlærerutdanningen uttrykker skolens sentrale holdninger og praksiser generelt. En offentlig evaluering av samtlige norske lærerutdanningsinstitusjoner, gjennomført av *Nasjonalt organ for kvalitet i utdanningen* (NOKUT) i 2006, avdekket betydelige svakheter i organisering og innhold i allmennlærerutdanningen generelt. Dette så jeg som interessant å undersøke nærmere, i tilknytning til temaet elevvurdering.

Jeg hadde også som utgangspunkt at elevvurdering og lærerutdanningen i Norge bedre kunne belyses ved å sammenligne med et utdanningssystem som er vesentlig forskjellig fra vårt. Gjennom tidligere arbeider var jeg blitt kjent med vurderingssystemet i Queensland, Australia (Tveit, 2007d, s. 225-230). Denne delstaten har fått betydelig oppmerksomhet internasjonalt på bakgrunn av en radikal satsing på læreres profesjonalitet og kompetanse til å vurdere standarder på elevenes måloppnåelse. Ettersom man i Norge er i ferd med å utvikle felles standarder for måloppnåelse, så jeg på Queensland som et særlig interessant sammenligningscase. Jeg antok også at man i Australia generelt, og i Queensland spesielt, har mer fokus på elevvurdering i lærerutdanningen. En sammenligning med Queensland ville således både kunne bidra til å kaste et nytt lys over den norske tilnærmingen til elevvurdering i lærerutdanningen og anskueliggjøre konkrete praktiske forbedringsmuligheter. Nedenfor gis en grundigere tematisk presentasjon av

vurderingene knyttet til henholdsvis oppgavens fokus på elevvurdering, allmennlærerutdanningens opplæring i dette og komparering mot Queensland.

1.1.1 Hvorfor utforske elevvurdering?

Temaet elevvurdering har blitt viet mye oppmerksomhet i forbindelse med implementeringen av den nye reformen for grunnsopplæringen, *Kunnskapsløftet*. Rett etter Bondevik II-regjeringens tiltredelse høsten 2001 hadde vi i Norge en omfattende debatt om kvaliteten på norsk skole generelt. Denne debatten tok hovedsaklig utgangspunkt i Norges skuffende resultater i internasjonale komparative kunnskapstester som PISA (*Programme for International Student Assessment*), PIRLS (*Progress in International Reading Literacy Study*) og TIMSS (*Trends in International Mathematics and Science Study*).

PISA2000 viste at norske 15-åringer kun oppnådde 13. plass blant OECD-landene i lesing og naturfag, og 17-plass i matematikk (Lie, Kjærnsli, Roe og Turmo, 2001). Disse resultatene skapte voldsom debatt, og førte til et kritisk søkelys på norsk grunnsopplæring. Daværende utdannings- og forskningsminister, Kristin Clemet, utvidet mandatet til en offentlig utredning ledet av Astrid Søgne (ofte omtalt som Kvalitetsutvalget) til å utferdige noe som lignet på en offentlig selvransakelse av norsk utdanningspolitikk, som beredte grunnen for omfattende endringer i norsk grunnskole og videregående opplæring. Utvalget leverte først en delinnstilling NOU 2002: 10 *Første klasses fra første klasse*, med forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem for norsk grunnsopplæring. Hovedtrekkene i innstillingen ble fulgt opp, inkludert nasjonale prøver på utvalgte trinn og tilhørende kontroversiell offentliggjøring av skolers prøveresultater og øvrige resultater (eksamens- og standpunktkarakterer) gjennom den såkalte *kvalitetsportalen*. Beslutningen om offentliggjøring av resultater ble et stort politisk stridstema i årene som fulgte, og etter stortingsvalget i 2005 besluttet Stoltenberg II-regjeringen å ta ett års prøvepause og å gjennomføre en kraftig revidering av rammeverket for prøvene.

Søgne-utvalgets hovedinnstilling NOU 2003: 16 *I første rekke*, inneholdt 117 forslag til endringer i den norske grunnsopplæringen. En bred høringsrunde ga grunnlag for forslag om en omfattende reform av hele grunnsopplæringen, som ble presentert i Stortingsmelding nr. 30 (2003-2004) *Kultur for læring*. Til forskjell fra rammeverket for nasjonale prøver var det bred politisk enighet om de fleste tiltakene i reformen, hvilket innebar at Stoltenberg II-regjeringen stort sett videreførte reformens innhold. Nye læreplaner ble tatt i bruk i barneskolen, på 8. trinn og videregående trinn 1 fra høsten 2006 og var i bruk av alle elever to år senere.

Med de nye læreplanene fulgte en omfattende revidering av forskrift til opplæringsloven, gjeldende fra 1. august 2006. Det ble gjort vesentlige endringer av forskriftene for elevvurdering.

Selve vurderingsordningene for læreplanene i Kunnskapsløftet, som det er grunn til å tro ville skapt stor strid, ble imidlertid ikke presentert før senere. Regler for eksamen ble presentert trinnvis etter hvert som elevene begynte på årstrinn med eksamen i fagene. Etter implementeringen av det nye læreplanverket ble det klart at ytterligere presiseringer i regelverket for elevvurdering var nødvendig. I desember 2006 tilkjennega Stoltenberg II-regjeringen, i Stortingsmelding nr. 16 (2006-2007), en rekke utfordringer knyttet til elevvurdering. Kunnskapsdepartementet slo fast at:

- regelverket for individvurdering oppfattes ikke som klart nok
- både lærerutdanningen og skolen mangler tilstrekkelig kompetanse
- det er svak vurderingskultur og vurderingspraksis i skolen
- det er forsket lite på elevvurdering i Norge (Stortingsmelding nr 16 (2006-2007), s. 79).

Kunnskapsdepartementet presenterte 30. januar 2007 en rekke tiltak for å bøte på disse problemene. Læreplanene i Kunnskapsløftet inneholder ikke fastsatte kriterier og standard for måloppnåelse, og kan således ikke defineres som rent kriterie- eller standardbaserte, selv om det i offentlige styingsdokumenter blir framholdt at vi har et standardbasert vurderingssystem. Skoler blir imidlertid oppfordret til å utarbeide vurderingskriterier med utgangspunkt i kompetansemålene i læreplanene, og Utdanningsdirektoratet har fått i oppgave å gjennomføre et omfattende utviklingsarbeid over flere år, der utforming av et klarere regelverk og utprøving av «felles kjennetegn på måloppnåelse» er blant de viktigste tiltakene for å oppnå «en mer likeverdig og rettferdig elevvurdering» (Kunnskapsdepartementet, 2007).

Som en del av dette arbeidet ble det med virkning fra 1. august 2007 fastsatt nye og felles karakterbeskrivelser (standarder) for ungdomsskole og videregående skole. Karakterene uttrykkes nå med forskjellige kompetansebeskrivelser for hvert nivå, fra «svært lav kompetanse i faget» (karakteren 1), til «fremragende kompetanse i faget» (karakteren 6) (Forskrift til opplæringsloven § 3-8 og § 4-8). Dette var ledd i utviklingen av et «standardbasert vurderingssystem», slik det var blitt omtalt i Stortingsmelding 30 (2003-2004), og svar på en kritikk om at de gamle bestemmelsene var basert på et normalfordelingsprinsipp (Utdanningsdirektoratet, 2007a, s. 5-6). Denne endringen, og de pågående utprøvingene av felles kjennetegn på måloppnåelse som ledsaget forskriftsendringen, er ventet å få stor betydning for hvordan elevvurdering gjennomføres i hele grunnopplæringen. Det kan se ut som at vi er «på vei mot» en form for kriterie- og standardbasert vurdering i Norge. Selv om dette utviklingsarbeidet langt fra er avsluttet, ser man alt nå konturene av et vurderingssystem med sterkere nasjonale føringer for lærernes karaktersetning i alle fag.

Til sammen fører dette til betydelige endringer i rammebetingelsene for lærernes yrkesutøvelse knyttet til elevvurdering. Jeg så det derfor som relevant å gjennomføre en undersøkelse som kunne belyse våre fremtidige læreres forutsetninger for å utføre slike oppgaver.

1.1.2 Hvorfor utforske allmennlærerutdanningen?

For å få mer informasjon om hvordan det står til med vurderingskompetansen blant norske lærere, så jeg det tidlig som interessant å rette fokus mot den vanligste veien til læreryrket, *allmennlærerutdanningen*. Denne utdanningen er i Norge sterkt regulert med felles nasjonale rammeplaner. Intervjuer med siste års allmennlærerstudenter ville dermed kunne fortelle en del om nyutdannede læreres forutsetninger for å gjøre elevvurdering. Jeg anså det som mer hensiktsmessig å intervjuer nyutdannede lærere enn erfarne lærere, ettersom utsagn fra sistnevnte gruppe like gjerne vil kunne knyttes til deres arbeidserfaring som deres utdanning. Som de viktigste leverandørene av lærere til norsk skole antas allmennlærerutdanningen å uttrykke sentrale trekk og praksiser knyttet til gjeldende forskrifter og ordninger for elevvurdering.

NOKUT hadde høsten 2006 presentert en omfattende evaluering av samtlige offentlige høyere utdanningsinstitusjoner som tilbyr allmennlærerutdanning. I evalueringspanelets samstemmige vurdering ble det pekt på en rekke svakheter. Det ble blant annet konkludert med at «integrering av praksis, fagstudium, fagdidaktisk- og pedagogisk teori er en stor utfordring for allmennlærerutdanningen»; samspillet mellom teori og praksis er svakt artikulert; og det er manglende samspill mellom pedagogikk og didaktikk på den ene siden og mellom ulike fagdidaktiske områder på den andre siden (NOKUT, 2006, s. 4). Evalueringspanelet stilte også spørsmål ved den brede kompetansen til allmennlærere for arbeid i hele grunnskolen (NOKUT, 2006, s. 77). NOKUT-panelet hadde imidlertid *ikke* undersøkt temaet elevvurdering, en indikasjon på den mangelfulle forskningen på elevvurdering i Norge, slik det ble framholdt i Stortingsmelding nr. 16 (2006-2007). Til sammen tilsa dette at en studie av opplæringen i elevvurdering i allmennlærerutdanningen ville være relevant.

1.1.3 Hvorfor komparere mot Queensland?

Gjennom tidligere arbeider var jeg som nevnt blitt kjent med vurderingssystemet i den australske delstaten Queensland. Delstatens unike tilnærming til elevvurdering er blitt henviset til av flere forskere innenfor fagfeltet (Black, 2004; Harlen, 2006; Sebba, 2006; Sebba og Maxwell, 2005). I Queensland avskaffet man alle statlige eksamener i 1972. Siden midten av 1980-årene har man gjennomført en spesiell form for standardbasert vurdering der lærere modererer¹ hverandres karaktersetning basert på autentiske elevarbeider. Royce Sadler utformet det teoretiske grunnlaget for vurderingssystemet, som han har drøftet i en rekke publikasjoner. Artikkelen *Specifying and*

¹ En form for kvalitetssikring av karaktersetning der lærerne vurderer hverandres karaktersetning, og gjennom faglige diskusjoner oppnår en klarere felles forståelse av karakterstandardene. Nærmere forklart i kapittel 2.2.2.

Promulgating Achievement Standards (R. Sadler, 1987) er den viktigste klargjøringen av prinsippene bak vurderingssystemet beskrevet som «standard-referenced assessment».

Queensland framsto som et aktuelt kompareringscase ettersom prinsipper for vurderingsarbeidet i delstaten har vesentlige fellestrekk med de nye retningslinjene for elevvurdering som er under utvikling i Norge. Det må presiseres at det i norske offentlige styringsdokumenter både er brukt begrepene «standardbasert» og «standardrelatert» og at det teoretiske grunnlaget for dette så langt ikke har blitt klargjort. Queensland kan sees på som *ett eksempel* på et vurderingssystem der man benytter felles standarder, og at staten derfor er et relevant kompareringscase. Jeg antok dessuten at lærerutdanningen i Queensland er mer spesialisert enn i Norge, noe som kunne lede til interessante komparasjoner.

1.2 Forskningsspørsmål

I denne oppgaven er Norge og Queensland valgt som case innenfor et utforskende (eksplorerende) forskningsdesign. Jeg intervjuet fire studenter hvert sted, som hadde hatt fordypning i morsmåls- og matematikkfagene. Et av målene med oppgaven var å finne ut hvilken betydning organisering av teori- og praksisopplæringen i lærerutdanningen hadde for opplæring av studentene i elevvurdering. Noen problemstillinger antas anskueliggjort gjennom å komparere med Queenslands lærerutdanning, på grunnlag av statens lange erfaring med standardbasert karaktervurdering og en mer spesialisert lærerutdanningen rettet mot alderstrinn der formell elevvurdering er mer brukt. For en grundig redegjørelse for forskningstilnærming henvises det til metodekapittelet (kapittel 4). Jeg gjentar her hovedproblemstillingen, og presenterer de tre forskningsspørsmålene jeg utviklet for å utforske problemstillingen:

Hva slags opplæring i elevvurdering fikk norske allmennlærerstudenter som avsluttet utdannelsen i 2007, sammenlignet med lærerstudenter utdannet i samme periode i Queensland, Australia?

1. Hvordan påvirket organiseringen av teori- og praksisopplæring studentenes opplæring i elevvurdering?
2. I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?
3. Hva slags teori- og praksisopplæring fikk studentene i
 - bruk av verktøy og hjelpemidler for elevvurdering;
 - bruk av vurderingskriterier;
 - diskusjoner om vurderingskriterier;

- karaktersetting;
- å gi faglige tilbakemeldinger; og
- forskrifter og retningslinjer for elevvurdering?

Disse forskningsspørsmålene dannet grunnlaget for utforming av intervjuguide og drøfting av funn i intervjudata, som har ledet fram til mine konklusjoner og svar på problemstillingen.

1.3 Oppgaveavgrensning

Jeg hadde i utgangspunktet planlagt en grundig teoretisk gjennomgang av temaet elevvurdering, inkludert en omfattende gjennomgang av norsk teori. Jeg har imidlertid ikke funnet omfattende teori om elevvurdering innenfor en norsk kontekst som er relevant for denne oppgaven. Mye av det offentlige ordskiftet om elevvurdering har i Norge dreid seg om en argumentasjon for og imot karakterer, framfor å drøfte praktiske tilnærminger til *hvordan* elevvurdering bør gjøres og i hvilken grad man lykkes med dette i Norge. Dette er i seg selv et interessant funn som jeg drøfter, men det bidrar også til at oppgaven dessverre er svakere teoretisk forankret enn ønskelig. En komparativ oppgave med datainnsamling i to land vil dessuten måtte basere seg på en grundig introduksjon til og innramming av casene. Dette har jeg valgt å samle i et eget kapittel (kapittel 2), men utgjør essensielt en del av litteraturanalsen. For å ikke sprengte oppgavens rammer har jeg derfor måttet avgrense de teoretiske perspektivene i kapittel 3.

1.3.1 Avgrensning av elevvurderingsfeltet

I Norge har det blitt forsket lite på elevvurdering, men *student assessment* er et omfattende fagfelt internasjonalt innenfor utdanningsadministrasjon, pedagogikk og fagdidaktikk. Fagfeltet har også relevante koblinger til fagområder innenfor sosiologi og psykologi. Generell teori om disse spørsmål er her lagt til side, til fordel for en mer fokusert analyse av de pedagogiske aspektene ved elevvurdering og den formelle funksjon lærerens karaktersetting representerer. I den internasjonale faglitteraturen er disse aspektene ofte diskutert under overskriftene *formativ* (læringsstøttende) og *summativ* (læringsmålende) vurdering.

Mens det formative aspektet kan hevdes å være av en viss universell karakter (det ligger i lærerrollen at man skal støtte elevene i læringsarbeidet), gjennomføres summativ vurdering på forskjellige måter på tvers av land og utdanningssystemer. I denne oppgaven har jeg valgt å benytte et bredt teoritilfang om formativ vurdering, mens teori om summativ vurdering begrenses til *lærerens* sluttvurdering og karaktersetting. I Norge og Queensland eksisterer det henholdsvis fagspesifikke eksterne eksamener og tverrfaglige prøver som på forskjellige måter har til hensikt å

supplere, korrigere eller kvalitetssikre lærernes karaktersetninger, for å bidra til rettferdig seleksjon. Disse blir presentert for å forklare konteksten lærernes elevvurdering foregår innenfor, men er ikke en del av det teoretiske grunnlaget for oppgaven. Dette innebærer at teorier om psykometrisk testing, *scaling* og andre teorier som i det internasjonale forskningslitteratur hører til fagfeltet *measurement* er valgt bort.

Som nevnt ovenfor kan det hevdes at man med Kunnskapsløftet er «på vei mot» en form for kriterie- og standardbasert vurdering i Norge. Hovedvekten legges derfor på det som i engelskspråklig litteratur omtales som «classroom assessment» eller «school-based assessment», og prinsipper som er forankret i teorier om *criterion-referenced assessment* (kriteriebasert vurdering), hvor *standard-referenced assessment* (standardbasert vurdering) inngår som et mer spesifisert form. Kriterie- og standardbasert vurdering kan, som de fleste andre vurderingsprinsipper, diskuteres i relasjon til læringsteorier. Med utgangspunkt i *kognitive*, *behavioristiske* og *sosio-kulturelle* læringsteorier, kunne man dratt opp ulike perspektiver på hva som kjennetegner god praksis for elevvurdering. Dette har jeg også måttet legge til side, til fordel for en mer fokusert analyse av litteratur som spesifikt omhandler elevvurdering.

1.3.2 Lærerutdanningsinstitusjoner og aldersgrupper i Norge og Queensland

I Norge har vi tradisjonelt hatt tre veier til læreryrket:

- A. *Allmennlærerutdanningen*: en fire års utdanning som kvalifiserer til jobb i barne- og ungdomsskolen.
- B. *Praktisk-pedagogisk utdanning* (PPU): en ettårig utdanning i pedagogikk og fagdidaktikk som bygges på en grad fra universitetsnivå og kvalifiserer til utdanning i jobb i ungdomsskolen og videregående skole.
- C. *Faglærerutdanning*: en utdanning primært rettet mot lærere på yrkesfaglige studieprogram i videregående skole.

Queensland har lignende modeller, med en vesentlig forskjell fra den norske allmennlærerutdanningen. Denne veien til læreryrket er i Queensland delt i to helt separate løp; for *primary* (1-8. klasse) og *secondary* (*junior*: 9-10. klasse; *senior*: 11. og 12. klasse).

Som forklart over ønsket jeg å fokusere på allmennlærerutdanningen i Norge. Etersom karaktersetning var et av temaene jeg ønsket å belyse, og det i Norge ikke gis karakterer på barnetrinnet, var det lite relevant å intervju norske allmennlærerstudenter som hadde prioritert undervisningsfag rettet mot barneskolen. Følgelig er *primary* programmet i Queensland lite relevant. Oppgaven er derfor avgrenset til allmennlærerstudenter som fordyper seg mot

ungdomstrinnet i Norge, og *secondary*-programmet i Queensland, som kvalifiserer lærere til undervisning både på ungdomstrinnet (*junior secondary*) og i videregående skole (*senior secondary*). Betydningen av disse strukturelle forskjellene er diskutert nærmere i metodekapittelet, og står også sentralt i analyse- og drøftningsdelen av oppgaven.

1.3.3 Elevvurdering og fag

En kvalitativ komparativ studie av kvalifikasjoner knyttet til elevvurdering vil måtte være sentrert rundt og begrenset til ett eller noen få fag. Oppgavens format gir ikke rom for å utforske mange fag. Det var ønskelig å ha med to fag, da dette åpner for å utforske forskjeller mellom fagene. I denne studien er morsmålsfaget (*Norsk/English*) og matematikkfaget valgt. Morsmålsfaget valgte jeg fordi dette er undervisningsspråket i alle fag utenom fremmedspråkene, og derfor gjerne regnes som det viktigste skolefaget. Videre blir det regnet som svært krevende å gjennomføre en pålitelig vurdering av læringsresultater i skrijving og muntlige ferdigheter, hvilket er en sentral del av morsmålsfaget. Dette så jeg som interessant å belyse. Matematikkfaget valgte jeg fordi det baserer seg på flere objektive regler, og derfor av mange regnes for å skille seg vesentlig fra morsmålsfaget ved å at det er lettere å gjennomføre pålitelig vurderinger av elevenes læringsresultater. Det blir også regnet som det mest grunnleggende realfaget i skolen, og er i så måte en interessant motsats til det mest grunnleggende humanistiske faget. En grundigere begrunnelse for valg av fag er gitt i metodekapittelet (jf. kapittel 4.5.2).

1.3.4 Datagrunnlag

Jeg bestemte meg for å benytte meg av halvstrukturerte intervjuer som den primære strategi for innsamling av data, fordi dette gir gode muligheter for å undersøke hva studentene opplever å ha lært gjennom utdanningen. Innenfor rammen av en masteroppgave er det begrenset hvor mange slike intervjuer man kan gjennomføre. Jeg valgte åtte informanter fordi jeg anså dette for å være overkommelig, og fordi det tilfredsstilte minimumskrav til replisering innenfor og på tvers av casene (to studenter med fordypning i samme fag begge steder). Begrunnelser for dette er nærmere forklart i metodekapittelet (jf. kapittel 4.5.3). Jeg har til sammen mer enn seks timer intervjudata, som i transkribert utgave teller ca 50 000 ord fordelt på 180 sider. Det hadde ikke vært mulig å gjennomføre flere eller mer omfattende intervjuer innenfor oppgavens rammer. I tillegg til intervjudata utgjør offentlige styringsdokumenter for Norge, Australia og Queensland, samt de to lærerutdanningenes studie- og fagplaner, studiens datagrunnlag.

1.4 Oppgavens struktur og komposisjon

I oppgavens andre kapittel gis en introduksjon til og avgrensning av forskningsobjektet. Dette er skilt fra teoridelen for å gi en ryddig innføring i hvordan elevvurdering og lærerutdanning gjennomføres de to stedene, med utgangspunkt i policydokumenter, lover, forskrifter, læreplaner og andre retningslinjer. I kapittel 3 presenteres en analyse av teoretiske perspektiver på elevvurdering som jeg har funnet relevante for studien. Kapittel 4 er oppgavens metodekapittel, der jeg presenterer og begrunner forskningsstrategien jeg har valgt, de metodiske problemstillingene og avveiningene jeg har tatt hensyn til, og prosessen med innsamling, analyse og framstilling av intervjudata. I kapittel 5 er intervjudataene analysert og strukturert etter forskningsspørsmålene som ligger til grunn for studien. I kapittel 6 drøfter jeg funnene og peker på likheter og forskjeller, på tvers av Norge og Queensland og mellom fagene. I kapittel 7 presenteres oppgavens konklusjoner og aktuelle temaer for videre forskning. Oppgavens appendikser er å finne helt bakerst i oppgaven, etter referanselisten.

2 Introduksjon til og innramming av casene: elevvurdering i lærerutdanningene i Norge og Queensland

Dette kapittelet inneholder introduksjon til og innramming av casene i studien. Hensikten er å gi en oversikt over rammer og prinsipper for elevvurdering og lærerutdanning som er relevante for studenter som avsluttet lærerutdanningen i 2007. De norske lærerstudentene i studien begynte utdanningen i august 2003 og avsluttet i juni 2007, mens studentene fra Queensland begynte utdanningen i februar 2003 og avsluttet i november 2007. Beskrivelse av utviklingstrekk knyttet til utdanningsforvaltning, lovgivning og læreplaner er begrenset til det som er gjennomført rett i forkant av og underveis i disse studentenes utdanningsløp.

I introduksjonen til Norge gis først en oversikt over kjennetegn ved det norske utdanningssystemet generelt. Både utdanningsforvaltningen og læreplanene i Norge har gjennomgått omfattende endringer de siste fem årene. I løpet av de norske lærerstudentenes studietid har det skjedd omfattende endringer i reguleringen av lærernes oppgaver både knyttet til læreplanene generelt og knyttet til elevvurdering spesielt. Nye læreplaner ble implementert før studentenes siste studieår (høsten 2006), men det offentlige ordskiftet og høringsprosessene fram mot det nye læreplanverket har trolig bidratt til fokus på disse læreplanene så tidlig som høsten 2005. Følgelig er to sett av læreplaner relevante for denne studien, L97 og Kunnskapsløftet.

Formelle reguleringer av elevvurdering er ytterligere komplisert å presentere, ettersom det her ble gjort vesentlige endringer både i 2006 og 2007. Dette innebærer at *gjeldende regelverk* er en uklar betegnelse. Lover og forskrifter som hadde vært gjeldende i en årrekke var gjeldende da studentene begynte på lærerutdanningen i 2003; nye forskrifter ble gjeldende før deres siste studieår (1. august 2006); og atter nye forskriftsendringer ble innført straks de hadde avsluttet lærerutdanningen (1. august 2007). De norske lærerstudentene fulgte en helt ny nasjonal rammeplan (2003). Endringene som ble gjort i forhold til forrige rammeplan presenteres kort, da vurderingene som lå til grunn for disse endringene er relevante å drøfte.

I introduksjonen til Queensland forklares først Australias føderale struktur og noen kjennetegn ved og utviklingstrekk i utdanningssystemet i Queensland. Utdanningsforvaltningen, læreplanene og regelverk for elevvurdering i Queensland har ikke gjennomgått like omfattende endringer de siste årene, og presenteres derfor noe kortere. Det er imidlertid nødvendig å gi en desto grundigere kontekstuell presentasjon av hvordan elevvurdering gjennomføres i Queensland og Australia, ettersom denne regionen er ukjent for de fleste norske lesere. Organiseringen av og innholdet i lærerutdanningen i Queensland er ikke regelstyrt, som i Norge. I stedet er det etablert

en statlig institusjon som er ansvarlig for godkjenning og kvalitetssikring av lærerutdanningsinstitusjoner. Denne ordningen presenteres kort. Avslutningsvis presenterer jeg i tabellform en oppsummering av likheter og forskjeller knyttet til elevvurdering og lærerutdanningen i Norge og Queensland.

2.1 Norge

Med 4,7 millioner innbyggere fordelt på 323 802 kvadratkilometer bærer Norge preg av rimelig spredt bosetning (12/km²). Landet har 616 400 elever i den obligatoriske grunnskolen, fordelt på 3100 skoler, og har således svært mange små skoler. Nesten halvparten (47 prosent) av skolene har færre enn 300 elever (SSB, 2008a). Den offentlige skolen er gratis. I 2007 gikk 97 prosent av grunnskoleelevene i offentlige skoler, mens tallet for videregående skole var 93 prosent (SSB, 2008b).

2.1.1 Norges utdanningsystem- og forvaltning

Rettigheter og plikter knyttet til utdanning for både skoleeiere, skoler, lærere, elever og foreldre, er forankret i Opplæringsloven og den tilhørende forskriften. Opplæringslovens bestemmelser utarbeides av departementet, fremmes av regjeringen og vedtas av Stortinget. Forskriftens bestemmelser er presisering av opplæringsloven og fastsettes av Utdanningsdirektoratet, på oppdrag fra Kunnskapsdepartementet.

Læreplanverket for L97 og Reform 94 ble utformet i regi av det som da het Kirke-, utdannings- og forskningsdepartementet med Gudmund Hernes (Arbeiderpartiet) som statsråd. I etterkant av disse reformene ble den statlige utdanningsadministrasjonen samordnet. Før omorganiseringen bestod den statlige utdanningsadministrasjonen av departementet, Nasjonalt læremiddelsenter, statens utdanningskontorer i hvert fylke og Eksamenssekretariatet (underlagt Statens Utdanningskontor i Oslo og Akershus). For å samordne arbeidet med læremidler og eksamen, ble Nasjonalt læremiddelsenter og Eksamenssekretariatet slått sammen til Læringscenteret i 1999. En del oppgaver ble delegert fra departementet til Læringscenteret, blant annet konkretisering av forskningsoppdrag og utvikling av årlige resultatrapporter (Stortingsmelding nr. 28 (1998-99), kap. 7).

Læreplanverket for Kunnskapsløftet ble utarbeidet på initiativ fra det som da het Utdannings- og forskningsdepartementet, med Kristin Clemet (Høyre) som statsråd. Hun sørget i 2004 for å omorganisere Læringscenteret til et direktorat – Utdanningsdirektoratet – og delegerte ytterligere departementsoppgaver til direktoratet, blant annet regien for utformingen av læreplanverket og forvaltningen av forskriftene. Samtidig ble statens utdanningskontor

underordnet Fylkesmannen, og Utdanningsdirektoratet fikk ansvar for oppfølging av fylkesmennenes utdanningskontorer.

2.1.2 Regulering av læreres arbeid med elevvurdering

Både det gamle og det nye læreplanverket skriver seg tilbake til 1993, da *den generelle læreplanen* ble utarbeidet og enstemmig vedtatt i Stortinget. Læreplanen beskriver de overordnede målene for skolen og lå til grunn for utformingen av alle de fagspesifikke læreplanene i 1990-årenes reformer av grunnskolen og videregående skole.

L97 og R94

Det gamle læreplanverket for videregående opplæring kom med Reform 94. Dette var en omfattende strukturreform der en rekke fag og studieretninger ble samordnet. Med reformen fikk alle elever rett på videregående opplæring, noe som førte til at forskriftens prinsipp om tilpasset opplæring fikk et nytt innhold. Det ble tydeligere at elever skulle få en opplæring som tok hensyn til deres ulike forutsetninger innenfor alle fag og studieretninger (Dale og Wærness, 2003, s. 20).

Den gamle læreplanen for grunnskolen – L97 – var en svært sentralstyrt læreplan, både når det gjaldt kunnskapsinnhold og arbeidsformer. Det ble lagt stor vekt på temaorganisering av innholdet på tvers av fag, med bakgrunn i hovedmomentene i flere fagplaner. Det ble gitt retningslinjer for bruk av «temaarbeid» og «prosjektarbeid», der det førstnevnte ble definert som mer lærerstyrt, mens prosjektarbeid i større grad skulle gi elevene «ansvar for egen læring». «På småskoletrinnet skulle 60 prosent av arbeidet gjøres som tema, på mellomsteget 30 prosent som tema- og prosjektarbeid, og på ungdomssteget 20 prosent som prosjektarbeid» (Haug, 2004, s. 38). Dette ble senere endret til kun å være veiledende. I forbindelse med innføringen av L97 ble det også innført *Prinsipper for opplæringa*, ofte omtalt som «broen», fordi den anga sammenhengen mellom den generelle læreplanen og de fagspesifikke læreplanene.

I L97 ble faglige krav angitt som mål på hvert av grunnskolens hovedtrinn. I tillegg anga læreplanen i stor detalj fagstoff som hovedmomenter. Læreplanene for videregående opplæring uttrykte også krav til elevenes kompetanse angitt som mål og hovedmomenter. Disse ble ofte svært detaljert utformet med angivelse av alt fra detaljert fagstoff til arbeidsmåter (Stortingsmelding nr. 30 (2003-2004), s. 34).

Kunnskapsløftet

Med Kunnskapsløftet ble det innført noen *grunnleggende ferdigheter* som forutsetninger for videre læring. Disse er: Å kunne uttrykke seg muntlig og skriftlig, å kunne lese, å kunne regne og å kunne

bruke digitale verktøy (Kunnskapsdepartementet, 2006, s. 39). De grunnleggende ferdighetene er forsøkt integrert i de kompetansebaserte læreplanene for fagene. Hensikten har vært å gi alle lærere og skolen som organisasjon et klarere ansvar for elevenes utvikling av grunnleggende ferdigheter. De nye læreplanene har kompetansemål som uttrykker ambisjoner for elevene etter 2, 4, 7. og 10. trinn i grunnskolen. I videregående skole er målene formulert for hvert årstrinn i de fleste fagene. Læreplanene for fagene *Norsk*, *Matematikk*, *Naturfag*, *Engelsk*, *Fremmedspråk 2* og *Kroppsøving* er gjennomgående fra begynnelsen i grunnskolen til fagene avsluttes i videregående skole (Kunnskapsdepartementet, 2006, s. 9). Slik legges det til rette for at elever kan jobbe med mål for neste trinn når de har nådd målene for trinnet de går på. Kompetansemålene angir hva elevene skal kunne etter endt opplæring på ulike trinn. I læreplanverket er det presisert at «elevene vil i ulik grad nå, eller kunne nå, de fastsatte kompetansemålene. Skolen skal gi tilpasset opplæring slik at hver enkelt elev stimuleres til høyest mulig grad av måloppnåelse» (Kunnskapsdepartementet, 2006, s. 9). På bakgrunn av dårlige erfaringer med pålegget om tema- og prosjektarbeid i L97 ble metodetvungen avvirket. Ansvar for utforming av innhold, organisering og bruk av arbeidsmåter i opplæringen ble overlatt til kommunene og fylkeskommunene:

Skoleeier kan fastsette lokale læreplaner i fagene som ramme for den enkelte skoles videre arbeid med planer for opplæringen. Skoler og bedrifter må selv vurdere hvilken organisering og hvilke arbeidsmåter og metoder som er best egnet til å realisere innholdet i læreplanen for den enkelte elev, lærling og lærekandidat (Kunnskapsdepartementet, 2006, s. 10).

I det nye læreplanverket ble den generelle læreplanen videreført. «Broen» i L97 ble erstattet med *Prinsipper for opplæringen*, som gjelder for hele grunnopplæringen (Kunnskapsdepartementet, 2006, s. 31). Her inngår *Læringsplakaten*, beskrevet som et «rammeverk for kvalitet», som skal tydeliggjøre skoleeiers ansvar (Stortingsmelding nr. 30 (2003-2004), s. 36). I prinsipper for opplæringen omtales mål for sosial og kulturell kompetanse, motivasjon for læring og læringsstrategier, elevmedvirkning, tilpasset opplæring og likeverdige muligheter, lærere og instruktørers kompetanse og rolle, samarbeid med hjemmet og samarbeid med lokalsamfunnet (Kunnskapsdepartementet, 2006, s. 31-35). Læreplanene for *Norsk* og *Matematikk*, med kompetansemål for 10. trinn, er lagt ved i kompendium 1 og 2.

Retningslinjer for elevvurdering

I forbindelse med implementeringen av Kunnskapsløftet, ble det gjennomført en omfattende revidering av forskrift til opplæringsloven der det ble gjort vesentlige endringer knyttet til

elevvurdering. I den nye forskriften skilles det gjennomgående mellom underveis- og sluttvurdering. Tidligere ble det skilt mellom «formell» og «uformell» vurdering. De nye begrepene kan sees på som en presisering av lærerens roller i tilknytning til elevvurdering, og ikke endring av intensjoner i forhold til den gamle forskriften.

Underveisvurdering

Formålet med underveisvurdering er å gi en løpende veiledning som fremmer læring og bidrar til å utvikle elevenes kompetanse. Underveisvurdering er viktig for å finne elevenes læreforutsetninger som grunnlag for tilpasset opplæring (Forskrift til opplæringsloven § 3-3 og § 4-4). Prinsippet om tilpasset opplæring framgår av opplæringslovens formålsparagraf²: «Opplæringa skal tilpassast den enkelte elev/lærling/lærekandidat sine evner og forutsetningar» (Opplæringsloven § 1-2). Under prinsipper for opplæringen i Kunnskapsløftet heter det at «elevene skal kunne delta i planlegging, gjennomføring og vurdering av opplæringen» (Kunnskapsdepartementet, 2006, s. 33). Det understrekes at omfanget av medvirkning vil variere i forhold til alder og utviklingsnivå. Sistnevnte kan sees i sammenheng med opplæringslovens formålsparagraf: Læreren må vurdere elevenes evner og forutsetninger til å ta ansvar. I de gamle bestemmelsene (før 1. august 2006) var dette forskriftsfestet for elever i videregående opplæring (Forskrift til opplæringsloven av 1998, § 4-11), mens det ikke var en tilsvarende bestemmelse for grunnskolen.

Underveisvurdering kan gis med og uten karakterer (Forskrift til opplæringsloven § 3-3 og § 4-4). Vurdering uten karakter er en blanding av *individrelatert vurdering* og *målrelatert vurdering* (Dale og Wærness, 2007, s. 105). Tilbakemeldingene skal ta hensyn til elevenes individuelle forutsetninger og motivere dem til å arbeide med målene i læreplanen. Denne formen kjenner vi fra barnetrinnet, der det ikke brukes karakterer, men den utgjør også en viktig del av underveisvurderingen i ungdomsskolen og på videregående trinn. I hele grunnopplæringen skal læreren gi vurdering uten karakter i form av en beskrivende vurdering av hvordan eleven står i forhold til kompetansemålene i de forskjellige fagene og de andre målene i læreplanverket for Kunnskapsløftet, med sikte på at elevene på beste måte skal kunne nå disse målene. I de nye bestemmelsene heter det at det skal kunne dokumenteres at vurdering er gitt (Forskrift til opplæringsloven § 3-4 og § 4-5). Elevene skal delta i vurderingen av eget arbeid. I grunnskolen skal kontaktlæreren minst to ganger i året ha samtaler med foreldre eller foresatte. Etter fylte 12 år har eleven rett til å være med i samtalen (Forskrift til opplæringsloven § 3-2). I videregående skole

² Det er fremmet forslag til ny formålsparagraf der bestemmelsene om tilpasset opplæring er utelatt. De nye bestemmelser forventes innlemmet i opplæringsloven i 2009, og ligger utenfor denne oppgaven å diskutere.

skal det gjennomføres en såkalt *elevsamtale* minst én gang hvert halvår (Forskrift til opplæringsloven § 4-5).

Karaktersetting

Vurdering med karakterer inntreer først i ungdomsskolen og skal knyttes til målene i de fagspesifikke læreplanene. Da læreplanene for Reform 94 og Reform 97 ble utarbeidet, het det at «måla i læreplanane for faga og dei måla i den generelle delen av læreplanen som er nedfelte i måla for faga, medregna fellesmål» skulle inngå i vurderingen med karakter (Stortingsmelding nr. 47 (1995–96), s. 13). Slik skulle elevenes «helhetlige kompetanse» vurderes. I Stortingsmelding nr. 30 (2003–2004) *Kultur for læring*, konstaterte departementet at dette hadde vært vanskelig, og bidratt til å gjøre vurderingsgrunnlaget uklart både for lærere og elever, særlig i videregående skole (Stortingsmelding nr. 30 (2003-2004), s. 39). I de nye retningslinjene heter det at karakteren skal gi uttrykk for den kompetansen eleven har oppnådd på tidspunktet for vurderingen, og ut fra det som er forventet på dette tidspunktet. Det er presisert at verken elevens individuelle forutsetninger eller vurdering i orden og oppførsel skal inngå i karaktervurderingen i fagene (Forskrift til opplæringsloven § 3-7 og § 4-7).

For karaktervurderingen i videregående skole het det fra 1998 og inntil 1. august 2007 at måloppnåelsen skal deles inn i *over middels* (karakterene 5 og 6), *middels* (karakterene 3 og 4) og *under middels* (karakterene 1 og 2) (Forskrift til opplæringsloven av 1998, § 4-8). Denne inndelingen hadde bakgrunn i normering av karakterene (se kapittel 3.4.2). Departementet understreket likevel i Stortingsmelding nr. 30 (2003-2004) at standpunktvurderingen skulle være standardbasert, med utgangspunkt i kompetansemålene for faget. «Eventuelle avvik i gjennomsnittskarakterer skal ideelt sett synliggjøre forskjeller i oppnådd kompetanse for et helt kull». For å bidra til dette lovet departementet «å sørge for at det utvikles nødvendige redskaper for at standpunktvurderingen i størst mulig grad skal være standardbasert» (Stortingsmelding nr. 30 (2003-2004), s. 40). Dette ble fulgt opp i departementets oppdragsbrev til Utdanningsdirektoratet av januar 2007, om utprøving av «felles kjennetegn på måloppnåelse» (jf. nedenfor). Som nevnt tidligere ble det som en del av dette arbeidet med virkning fra 1. august 2007 fastsatt nye og felles karakterbeskrivelser (standarder) for ungdomsskole og videregående skole. Karakterene uttrykkes nå med forskjellige kompetansebeskrivelser for hvert nivå, fra «svært lav kompetanse i faget» (karakteren 1), til «fremragende kompetanse i faget» (karakteren 6) (Forskrift til opplæringsloven § 3-8 og § 4-8).

I ungdomsskolen skal det gis terminkarakterer i fag, orden og oppførsel midt i opplæringsperioden på hvert årstrinn og ved avslutningen av 8. og 9. trinn. I videregående skole

skal det gis terminkarakter etter 2. termin på de årstrinnene der det ikke er fastsatt nasjonale kompetansemål for faget. Terminkarakterene skal suppleres med begrunnelse og veiledning. Det samme gjelder når det blir gitt karakter på prøver og annet arbeid underveis i opplæringen (Forskrift til opplæringsloven § 3-5 og § 4-8). I ungdomsskolen gis det standpunktkarakterer ved slutten av opplæringen i fagene, og endelig karakter på 10. trinn i orden og oppførsel. I norskfaget skal elevene ha tre standpunktkarakterer: En i norsk hovedmål skriftlig, en i norsk sidemål skriftlig og en i norsk muntlig. I engelsk er det en muntlig og en skriftlig standpunktkarakter. I de andre fagene er det én standpunktkarakter (Kunnskapsdepartementet, 2006). Med forskriftsendringen av 2006 ble det presisert tydeligere at elevens innsats, orden og oppførsel ikke skal knyttes til karakterene i fag, men i karakterene for orden og oppførsel. For å gjøre det tydeligere hva disse karakterene uttrykker skal det gis separate karakterer for orden og oppførsel også for ungdomsskolen, slik det tidligere kun ble gjort i videregående skole. Utdanningsdirektoratet har sendt ut forslag til nye vurderingsbestemmelser med ytterligere presiseringer, der vurderingsforskriftene for grunnskole og videregående skole er samordnet i ett kapittel (Utdanningsdirektoratet, 2008a). Det ligger utenfor denne oppgaven å drøfte de foreslåtte endringene.

Eksamen

Over er beskrevet retningslinjer for elevvurdering foretatt av lærerne selv. I alle fag på 10. trinn på ungdomsskolen og i avsluttende fag på videregående trinn kan elevene trekkes ut til eksamen. Utdanningsdirektoratet er ansvarlig for å produksjon og sensur av *sentralt gitte* eksamener, mens skolene er ansvarlig for gjennomføring av disse eksamenene. For de *lokalt gitte* eksamenene er Utdanningsdirektoratet kun ansvarlig for å forvalte retningslinjer for gjennomføring og sensur. Eksamensformene i fagene fremgår av læreplanen. De sentralt gitte eksamenene er kun skriftlige, mens lokalt gitte eksamener kan være skriftlige, muntlige eller praktiske. Ansvar for forberedelser, gjennomføring og sensur av lokalt gitte eksamener i grunnskolen ligger hos kommunen, mens fylkeskommunen har ansvar for disse eksamenene i videregående skoler. Dette ansvaret er imidlertid oftest delegert til skolene og lærerne selv.

På 10. trinn blir alle elever trukket ut til skriftlig, sentralt gitt eksamen i enten *Norsk*, *Matematikk* eller *Engelsk*. Elevene vet at de kommer opp i en eksamen, men ikke i hvilket fag. De skriftlige eksamenene varer normalt i fem timer. Elevene kan også trekkes ut til en lokalt gitt muntlig eksamen i enten *Norsk*, *Engelsk*, *Matematikk*, *Religion Livssyn og Etikk* (RLE, tidligere KRL), *Samfunnsfag* eller *Natur- og miljøfag* (Kunnskapsdepartementet, 2006). I videregående skole varierer antallet eksamener mellom de forskjellige utdanningsprogrammene. I

studieforberedende utdanningsprogram blir 20 prosent av elevene trukket ut til en skriftlig eksamen ved slutten av det første året og alle elever ved avslutningen av andre år. I det tredje og siste året gjennomfører alle elever tre eller fire eksamener. Type eksamener varierer mellom hvilket studiespesialiserende utdanningsprogram elevene tar. Alle elevene må imidlertid gjennomføre den eksterne eksamenen i *Norsk skriftlig* (Utdanningsdirektoratet, 2008b). Hovedregelen er at elevene ikke vet om og i tilfelle hvilke eksamener de vil bli trukket ut til før to dager før eksamensdagen, hvilket fungerer som et incentiv for å stå på i alle fagene gjennom hele skoleåret.

Selv om det i utgangspunktet ikke er noen nasjonale føringer for lærernes standpunktvurdering, er det grunn til å tro at sentralt gitte eksamener har en normerende funksjon. Evalueringen av avgangsprøva i norsk (eksamen på 10. trinn) tilsier at en omfattende satsing på utvikling av felles kriterier for vurdering av elevsvarene har bidratt til å utvikle et tolkningsfellesskap som også til en viss grad gjør seg gjeldende i standpunktvurderingen (Vagle, Berge, Evensen og Hertzberg, 2007, s. 76-77). Eksamenssensorene og andre lærere som leser sensorveiledningene og bruker gamle eksamensoppgaver, blir også en del av dette tolkningsfellesskapet (Tveit, 2007a, s. 215). Eksamenskarakterene rapporteres separat på elevenes vitnemål og utgjør under 20 prosent av elevenes karakterer, mens standpunkt karakter utgjør resten av vitnemålet (Tveit, 2007a, s. 194-195, 212). Forskjellig karakternivå på eksamen og til standpunkt kan ses i sammenheng med at vurderingsgrunnlagene er vidt forskjellige. Det er derfor liten grunn for rektorer til å reagere overfor lærere som gir lavere eller høyere standpunkt karakterer enn elevene får på eksamen (Tveit, 2007a, s. 215).

Nasjonale prøver

Som en del av et nasjonalt rammeverk for kvalitetsvurdering ble det i 2004 introdusert et nytt system for prøver, omtalt som *nasjonale prøver*, ettersom de var utarbeidet i statlig regi og obligatorisk for alle elever. Sammen med prøvene fulgte et kontroversielt prinsipp om offentliggjøring av skolenes resultater på prøvene, sammen med standpunkt- og eksamenskarakterer. Prøvene og prinsippet om offentliggjøring skapte en voldsom debatt, der blant andre Elevorganisasjonen og Utdanningsforbundet markerte seg som sterke motstandere av slik offentliggjøring. På tross av skarp faglig kritikk og motstand fra skolenes organisasjoner bestemte utdanningsdepartementet at arbeidet med nye prøver skulle fortsette. Elevorganisasjonen oppnådde stor oppslutning om boikott av prøvene i matematikk, lesing og engelsk skrivning. Mellom mellom 35 og 46 prosent av elevene boikottet prøvene i videregående skole i 2005 (Hølleland, 2007, s. 37). Svein Lie m.fl. evaluerte prøvene begge årene, og konkluderte med at prøvene hadde store

svakheter på en rekke områder (Lie, 2007, s. 88-91). De faglige svakhetene ved prøvene, erkjent senere, var blant annet uklarhet om formålet og indre konsistens i oppgavesettene (Utdanningsdirektoratet, 2007b, s. 82). Det ble gitt uklare beskjeder om hvorvidt prøveresultatene skulle ha betydning for elevenes individuelle standpunkt karakterer, eller om de kun var målinger av gruppeprestasjoner (Hølleland, 2007, s. 36). En av de viktigste lærdommene fra de første årene med nasjonale prøver, var at norske lærere hadde liten erfaring med og derfor svake kunnskaper om hvordan man skal vurdere slike prøver.

Stoltenberg II-regjeringen bestemte å ta et års prøvepause mens et mer solid rammeverk for prøvene ble utviklet. I 2007 ble det introdusert nye prøver i matematikk og lesing på norsk og engelsk på 5. og 8. trinn (Utdanningsdirektoratet, 2007b, s. 82). Implementeringen av disse prøvene gikk mer i henhold til planene (Kavli, 2008, s. 4). Prøvene gjennomføres nå ved starten av mellomtrinnet og ungdomstrinnet. Slik er det gjort klarere at prøveresultatene ikke skal inngå i elevenes standpunkt karakterer. De nye prøvene skal «gi økt kunnskap og være et grunnlag for kvalitetsutvikling både på lokalt og sentralt nivå» (Utdanningsdirektoratet, 2008c, s. 7). Prøveresultatene publiseres basert på en felles skala og gjøres offentlig tilgjengelig, men myndighetene legger ikke til rette for rangering av skoler (Utdanningsdirektoratet, 2007b, s. 82). Det er planlagt å utvide prøveordningen til også å omfatte videregående skole.

Felles kjennetegn på måloppnåelse

I januar 2007 ga Kunnskapsdepartementet Utdanningsdirektoratet i oppgave å gjennomføre et omfattende utviklingsarbeid over flere år, for å oppnå «en mer likeverdig og rettferdig elevvurdering» (Stortingsmelding nr. 16 (2006-2007), s. 79). Målet med denne satsingen er ikke bare mer rettferdig karaktersetning på tvers av skoler og klasserom – de nye vurderingsordningene skal også bedre det formative aspektet ved elevvurdering og sees i sammenheng med Kunnskapsløftet overordnede mål om å øke læringsutbyttet for alle elever. Utdanningsdirektoratet har iverksatt en utprøving av ulike modeller for kjennetegn på måloppnåelse i fagene *Norsk*, *Matematikk*, *Mat og helse* og *Samfunnsfag* (Kunnskapsdepartementet, 2007). Det ligger utenfor denne oppgaven å drøfte denne utprøvingen nærmere, men det kan bemerkes at en av modellene går ut på å utforme kjennetegn på måloppnåelse som kan definere forskriftens seks karakterbeskrivelser nærmere for de respektive fagene, både for ungdomstrinnet og videregående trinn.

2.1.3 Organisering og innhold i norsk allmennlærerutdanning

Innhold og organisering i allmennlærerutdanningen i Norge er regulert av nasjonale myndigheter og fastsatt av stortinget. Allmennlærerutdanningen har siden 1992 vært fireårig. Siden har det gjentatte ganger blitt vurdert om lærerutdanningen burde spesialiseres mer mot alderstrinn og om det skulle stilles mer krav til fagkompetanse for å undervise på ungdomsskolen. I Stortingsmelding nr 48. (1996-97) *Om lærerutdanning* holdt departementet fast ved at allmennlærerutdanningen bør kvalifisere for undervisning på alle trinn i grunnskolen. Departementet ga tre begrunnelser for dette:

1. Allmennlærerens rolle som klassestyrer
2. Allmennlærerens rolle i å bidra til helhet og sammenheng mellom hovedtrinnene i grunnskolen
3. Hensynet til de mange fådelte eller små barne- og ungdomsskoler med et lite lærerkollegium, som skaper behov for lærere med bred fagkompetanse som både kan undervise mange fag, i samme klasse og på ulike trinn (Stortingsmelding nr. 48 (1996-97), s. 28).

I den samme stortingsmeldingen gikk departementet inn for å øke andelen obligatoriske fag for å oppnå en større felles basis og styrke den faglige bredden for alle allmennlærerstudenter (Stortingsmelding nr. 48 (1996-97), s. 40). Stortingsflertallet sluttet seg til departementet forslag til omfang av obligatoriske fag og prinsippet om at allmennlærerutdanning skulle gi undervisningskompetanse i hele grunnskolen, men doblet omfanget av KRL-faget i forhold til departementets forslag til 10 vekttall (30 studiepoeng) på bekostning av engelsk som obligatorisk fag (Innst. S. nr. 285 (1996-97)).

I forbindelse med Kvalitetsreformen ble strukturen på lærerutdanningen gått igjennom på ny og tilpasset den nye ordningen med studiepoeng i stedet for vekttall. Fireårig lærerutdanning ble beholdt, selv om dette avviker fra bachelor- og mastergradsinndelingen på henholdsvis tre og to år. På bakgrunn av en evaluering av erfaringene med rammeplanen av 1998 ble det bestemt å redusere antallet obligatoriske fag og å åpne for fleksibilitet i organiseringen. Det ble vurdert en modell med to atskilte studieløp der studentene måtte velge utdanning rettet mot undervisning for årstrinn 1-7 eller 5-10 ved starten av utdanningen. Departementet gikk imidlertid inn for å beholde et enhetlig løp, slik flertallet av høringsinstansene hadde argumentert for. Departementet begrunnet dette denne gangen med at den gir studentene, lærerutdanningsinstitusjonene og skoleeiere større grad av fleksibilitet, og at det er mindre ressurskrevende enn modellen med to atskilte studieløp (Stortingsmelding, 2001-2002:16).

Departementet la fram en modell der det ble kortet ned på antall obligatoriske fag fra 180 til 120 studiepoeng. Det ble foreslått å beholde 30 studiepoeng i hvert av fagene *Pedagogikk*, *Norsk* og *Matematikk*. Et eget fag i *Grunnleggende lese-, skrive og matematikkopplæring* med 10 studiepoeng ble foreslått, sammen med et fag på 10 studiepoeng kalt *Verdi- og kulturformidling*, som i hovedsak skulle knyttes opp mot KRL-faget. Det ble også foreslått et fag kalt *Profesjonskunnskap* med 10 studiepoeng. I Stortingets behandling ble det gjort omfattende endringer. Stortingsflertallet gikk inn for at KRL-faget fortsatt skulle være obligatorisk i allmennlærerutdanningen med et omfang på 20 studiepoeng, på bekostning av de foreslåtte fagene *Profesjonskunnskap* og *Verdi- og kulturformidling*. Flertallet mente at innholdet i disse fagene kunne ivaretas både i KRL-faget og i pedagogikkfaget (Innst. S. nr. 262 (2001-2002)).

Stortingets vedtak la premissene for rammeplanen av 2003 som har regulert allmennlærerutdanningen siden. Det er en fireårig utdanning (240 studiepoeng) og omfatter en obligatorisk del på 120 studiepoeng og en like stor valgbar del. Tabell 1 viser utdanningens obligatoriske del som gjennomføres de to første årene.

Tabell 1: Obligatoriske fag i allmennlærerutdannings første to år gjeldende fra 2003

Pedagogikk	30 studiepoeng
Kristendoms-, religions- og livssynskunnskap	20 studiepoeng
Matematikk	30 studiepoeng
Norsk	30 studiepoeng
Grunnleggende lese-, skrive og matematikkopplæring	10 studiepoeng
Sum obligatorisk innhold	120 studiepoeng
Praksisopplæring som en integrert del av studieenhetene	20-22 uker

Den valgbare delen på 120 studiepoeng gjennomføres de to siste årene og skal omfatte minst 60 studiepoeng i fag som tilsvarer fag i grunnskolen. Fagene kan være påbygging av fag i obligatorisk del eller nye fag med minst 30 studiepoengs omfang og skal omfatte fagdidaktikk og praksisopplæring. Inntil 60 studiepoeng kan være andre skoler relevante studier av ulikt omfang knyttet til særskilte temaer (Forskrift til rammeplan for allmennlærerutdanning, 2003).

Denne gjennomgangen viser at det i Norge har vært stor politisk strid om organisering og innhold i allmennlærerutdanningen. Stortingsflertallet sørget både i 1997/98 og 2001/2002 for at KRL-faget fikk en sterkere posisjon enn departementet la opp til. Stortingsflertallet sluttet imidlertid opp om å fjerne praktiske fag, estetiske fag og faget *Natur, Samfunn og Miljø* som

obligatorisk del av allmennlærerutdanningen. Forslag om å innføre to atskilte studieløp med ett rettet spesielt mot trinnene 5-10 ble vurdert, men avvist av flertallet av høringsinstansene, departementet og stortingsflertallet. I stedet ble det lagt til rette for at lærerutdanningsinstitusjonene selv organiserer innholdet de to siste årene, og eventuelt tilby fordypning rettet mot ungdomstrinnet. Det norske caset i denne studien har tilbudt fagene *Norsk for ungdomstrinnet* og *Matematikk for ungdomstrinnet* (begge 30 studiepoeng) og undersøkelsen har derfor vært basert på studenter som har valgt slik fordypning.

2.2 Queensland

Australia er en føderasjon bestående av seks *stater* og to *territorier*. Landet har ca 20 millioner innbyggere (2006) og majoriteten av befolkningen bor i og rundt de henholdsvis fem største hovedstedene Sydney (New South Wales), Melbourne (Victoria), Brisbane (Queensland), Perth (Western Australia) og Adelaide (South Australia). Utdanning er konstitusjonelt statenes og territorienes ansvar, et selvstyre som stadig utfordres av føderale myndigheter, men som blir sterkt beskyttet av statene. Det er imidlertid enighet mellom statene og territoriene om omfattende samordningsprosesser i utdanningssektoren av hensyn til befolkningsmobilitet. Staten Queensland har 4,2 millioner innbyggere fordelt på hele 1,8 millioner kvadratkilometer og er, som Australia sett under ett, et av verdens mest spredt befolkede områder (2,4/km²).

I 2006 hadde Queensland 58 748 elever i pre-school, 348 514 elever i *primary* (*Year 1-7*) og 236 685 elever i *secondary* (*Year 8-12*). De totalt 643 947 elevene fordelte seg på totalt 1739 skoler (Queensland Government, 2008), hvilket innebærer at hver skole i gjennomsnitt har nesten dobbelt så mange elever som i Norge. Queensland har et offentlig utdanningssystem med betydelige private innslag, særlig på *secondary* nivå. I 2006 gikk 71 prosent av elevene i *primary* på offentlige skoler, 17 prosent i katolske skoler og 12 prosent i andre ikke-offentlige skoler (andre religiøse eller «uavhengige» skoler). I *secondary* gikk i underkant av 62 prosent av elevene i offentlige skoler, mens resten av elevene fordelte seg omtrent likt på katolske og andre ikke-offentlige skoler (Queensland Government, 2008). Skoleårene går fra februar til desember. For å begynne i *Year 1* må elevene fylle seks år før 30. juni det aktuelle året.

2.2.1 Queenslands utdanningssystem og -forvaltning

Queensland Studies Authority (QSA) i hovedstaden Brisbane har ansvar for utdanningssystemet i Queensland, fra de urbane strøkene rundt Brisbane, Gold Coast og Sunshine Coast i kystområdene i sørøst, til fjerntliggende rurale samfunn 2000 kilometer nord (Cumming og Maxwell, 2004, s. 90). Før QSA ble etablert i 2002, var den statlige utdanningsadministrasjonen delt i tre enheter.

Queensland School Curriculum Council (QSCC) hadde ansvar for retningslinjer for *primary*, herunder læreplaner. *Queensland Board of Senior Secondary School Studies* (QBSSSS) hadde tilsvarende ansvar for *junior secondary* (Year 8-10.) og *senior secondary* (Year 11-12). *Tertiary Entrance Procedures Authority* (TEPA) var ansvarlig for opptak til høyere utdanning. For å bidra til bedre sammenheng i utdanningsløpet ble dette samorganisert til QSA, hvis viktigste oppgaver er: utarbeiding og revidering av læreplaner; prosedyrer for elevvurdering sertifisering av elever etter fullført skolegang; og opptak til høyere utdanning. Organisasjonen ledes av et styre satt sammen av representanter fra både offentlige og private skoler, foreldre, lærere og politikere.

2.2.2 Regulering av læreres arbeid med elevvurdering

Til tross for et sterkt press på utvikling av nasjonale læreplaner og vurderingsordninger, er dette fremdeles statenes ansvarsområde. I 1999 ble man ved *Adelaide-erklæringen* imidlertid enig om en del felles prinsipper utdanningssystemene skulle organiseres ut fra i *English* og *Mathematics* og seks andre kjernefag (*Key Learning Areas*). Tidligere forsøk på å utvikle en nasjonal læreplan har bidratt til at det allerede er en rekke fellestrekk i statenes utdanningssystemer (Cumming og Maxwell, 2004, s. 90-93). Nasjonale «Statements of Learning for English» ble utarbeidet i 2005 (NCCOSE, 2005) og tilsvarende i *Mathematics* i 2006 (NCCOSE, 2006). I Queensland pågår det i perioden 2005 til 2009 et omfattende arbeid med utarbeiding av *Queensland Curriculum, Assessment and Reporting framework* (QCAR) for *Year 1-10*, som er en oppfølging av de nasjonale føringene. I følge utdanningsdepartementet i Queensland, skal dette rammeverket «provide a common frame of reference and a shared language for communicating student achievement» som vil bidra til «greater consistency in teacher judgments» (Department of Education and the Arts, 2005, s. 2).

Det er forskjellige ordninger for eksterne eksamener i de fleste australske statene, i tillegg til læreres karaktervurderinger («school based assessment»). Retningslinjer for læreres karaktervurdering varierer mellom statene, men et fellestrekk er at de er basert på prinsipper for kriteriebasert (*criterion-referenced*) eller standardbasert (*standard-referenced*) vurdering. Cumming og Maxwell beskriver minst to grunner for at standardbasert vurdering utført av læreren har en sterk posisjon i Queensland: Ønsket om å ivareta en bred læreplan og respekt for lærerens profesjonalitet i å vurdere elevers måloppnåelse (Cumming og Maxwell, 2004, s. 90-93).

Queensland skiller seg fra de fleste australske delstatene ved å ikke ha noen eksterne eksamener hvis resultater blir rapportert på elevenes vitnemål, verken for *junior secondary* eller *senior secondary*. Ved avslutningen av *Year 12* gjennomfører imidlertid alle elever to tester kalt «Queensland Core Skills» (QCS) test. Skolenes resultater på disse prøvene brukes til å kalibrere

alle elevenes vitnemål i beregningen av *Overall Performance* (OP) som legges til grunn ved opptak til høyere utdanning. Hensikten med denne *scaling*-prosedyren er å forhindre at elever gis fordeler eller ulemper i forhold til hvilke skoler de går på (QSA, 2008b). Lærerne har imidlertid lite med QCS-testen å gjøre, og dette drøftes derfor ikke nærmere her.

Mange skoler i Queensland har elever fra 1. til 12. klasse, mens det også er vanlig å skille mellom *primary* og *secondary*. Læreplanene har imidlertid vært organisert for årene 1-10 og 11-12, hvilket innebærer at mange secondary-lærere har måttet forholde seg til to forskjellige sett av læreplaner. For *English Year 1-10* ble det introdusert en ny læreplan for åpen utprøving i 2005 (QSA, 2005a). Denne var ment for utprøving for semester 2, 2005 og semester 1, 2006, men var fortsatt ikke erstattet av noen ny læreplan i 2007. Læreplanen er sammen med en åpen utprøving av QCAR-initiativet *Essential Learnings and Standards for English for Year 9* av 2007 de viktigste styringsdokumentene for lærerne som underviser i *English* i *junior secondary*. I *senior secondary English* gjelder en læreplan som ble utarbeidet i 2002 (QBSSSS, 2002).

Læreplanen for *Mathematics Year 1-10* ble introdusert i 2004 (QSA, 2004). Også i dette faget utgjør en åpen utprøving av *Essential Learnings and Standards for Mathematics Year 9* av 2007 det viktigste styringsdokumentene ved siden av læreplanen for lærerne som underviser i *junior secondary*. For *Year 11 & 12* er det tre forskjellige matematikkfag – A, B og C – med litt ulik tilnærming og vanskelighetsgrad. Læreplanene for disse fagene ble implementert i 2001. *Mathematics B* (QBSSSS, 2001) er det klart mest vanlige faget.

Også i Queensland har det altså pågått et omfattende utviklingsarbeid knyttet til læreplanene i årene denne studien omfatter, som en oppfølging av de nasjonale føringene. Ettersom det er forskjellige læreplaner og vurderingsordninger for *Year 1-10* og *Year 11 & 12*, presenteres disse separat. *English* gis en litt fyldigere presentasjon, for å illustrere hvordan mål, kriterier og standarder er organisert i læreplanene.

Elevvurdering i Years 1-10

I læreplanene for *Year 1-10* er læreplanmålene (*learning outcomes*) organisert i seks nivåer langs en rekke *strands*. I læreplanen for *English* er det tre *strands*: *Cultural strand* (*making meaning contexts*), *Operational strand* (*using language systems*) og *Critical strand* (*evaluating and reconstructing meanings in texts*) (QSA, 2005a, s. 5). Hver *strand* har tre *substrands*: *Speaking and listening*, *Reading and viewing* og *Writing and shaping*. Hver *strand* inneholder standarder for måloppnåelse knyttet til klassenivå. Elever som demonstrerer nivå 1-oppnåelse er forventet å være i midten av *Year 2*, mens elever som demonstrerer nivå 6-oppnåelse er forventet å være i slutten av *Year 10* (QSA, 2005a, s. 7, 9). I *Mathematics* er læreplanens *strands* kalt: *Number concepts*;

Patterns and Algebra; Measurement; Chance and Data; og Space (QSA, 2004). Eksempler på *strands og levels* fra læreplanene i *English og Mathematics* ligger vedlagt som henholdsvis appendiks 3 og 4.

De formelle retningslinjene for elevvurdering for *Year 1-10* framgår av læreplanen. Vurderingskapittelet i læreplanene er omfattende (åtte sider), både for *English og Mathematics*. De beskriver formål, prinsipper og prosesser for vurdering og er tilnærmedesvis like i de to læreplanene. Formålet med elevvurdering er beskrevet slik: «providing ongoing feedback on the progress of individual students and groups of students throughout the learning and teaching process» og å informere «students, teachers, parents/carers, others in the community and/or school authorities about students' learning» (QSA, 2005a). Elevvurdering skal også hjelpe læreren til å gjøre beslutninger knyttet for eksempel til «student needs, the learning and teaching process, and resource requirements»; «plan learning and teaching programs for individuals, classes and the whole school»; «discuss future learning pathways with students and parents/carers»; «make decisions about providing learning support to particular groups of students»; og «develop learning resources and curriculum materials» (QSA, 2005a).

Prinsippene for elevvurdering har følgende overskrifter: *demonstration of learning; comprehensive range of evidence; valid and reliable evidence; og individual learners*. Det blir lagt vekt på at bedømminger av elevers læringsresultater skal være basert på en bredde av «evidence» samlet over tid. Lærere må legge til rette for at elevene får en rekke muligheter innenfor varierte kontekster til å demonstrere hva de kan (QSA, 2005a, s. 69). Læreplanene inneholder en tabell med eksempler på måter å samle og lagre *evidence* fra en bredde av kilder. Tabellene varierer i henhold til fagenes natur. Tabellene for *English og Mathematics B* ligger vedlagt som henholdsvis appendiks 5 og 6. For å sørge for valide og reliable vurderinger, blir det lagt vekt på at «teachers' judgments should be consistent within their own classes for different students, for different assessment opportunities, and at different times. They should also be consistent with the judgments of other teachers in their own school and other schools» (QSA, 2005a, s. 69). Læreplanen beskriver en rekke strategier for å sikre konsistens i lærernes vurderinger:

- Sharing understandings about the learning
- Planning collaboratively
- Using common assessment tasks
- Designing statements of anticipated evidence
- Conducting moderation processes (formal and informal)
- Collecting and using samples of typical responses (QSA, 2005a, s. 75).

Formelle modereringsprosedyrer beskrives, men i motsetning til *Year 11 & 12* er det ingen obligatoriske prosedyrer for ekstern moderering for *Year 1–10*.

Elevvurdering i Year 11 & 12

Elevvurdering for *Year 11 & 12* er regulert gjennom et omfattende system for kvalitetssikring av skolers implementering av vurderingsprosedyrer og ekstern moderering av lærereres karaktervurdering i alle de vanligste fagene³. QSA er ansvarlig for å utvikle læreplanene og for å organisere og tilrettelegge for disse prosessene.

Læreplanene

Læreplanene i alle de vanligste fagene beskriver læreplanmål organisert i kriterier og standarder for vurdering av elevenes måloppnåelse.

Læreplanene for *English* og *Mathematics B* er på henholdsvis 20 og 10 sider og beskriver fire generelle mål. Tre av disse tar også form av kriterier og standarder, mens det fjerde, *affective objectives*, ikke skal inngå i den formelle karaktervurderingen. Affektive objektiver er beskrevet med stikkordene *enjoying*, *engaging*, *relating*, *appreciating* og *playing* (QBSSSS, 2002, s. 5). I læreplanene gis det ingen forklaring på hvorfor disse objektivene er utelatt fra den summative vurderingen, men det kan antas at formell vurdering av slike faktorer blir sett på som både vanskelig og uetisk å gjennomføre, og at det vil lede til urettferdige vurderinger.

Læreplanene beskriver fem nivåer av måloppnåelse: *Very High Achievement* (VHA); *High Achievement* (HA), *Sound Achievement* (SA), *Limited Achievement* (LA) og *Very Limited Achievement* (VLA). Læreplanene inneholder rubrikker med kriterier og standarder for oppnåelse av målene. De tre vurderingskriteriene i *English* er:

1. Knowledge and control of texts in their contexts
2. Knowledge and control of textual features
3. Knowledge and application of the constructedness of text (QBSSSS, 2002, s. 11).

De tre vurderingskriteriene skal integreres i *units of work*. Læreplanen skisserer en rekke måter dette kan gjøres på. Mens det kun *foreslås* innhold, er retningslinjene for organiseringen av opplæringsenhetene mer rigide. Dette er trolig for å sikre forståelse på tvers av skoler i forbindelse

³Alle de mest vanlige fagene kalles *authority subjects*. Skoler kan også tilby *authority-registered subjects* eller utvikle egne fag, som det er knyttet færre formelle retningslinjer til. *English* og *Mathematics B* er begge *authority subjects*.

med modereringsprosedyrene. Læreplanen har retningslinjer og minstekrav til bruk av litteratur og kilder i forbindelse med opplæringsenhetene. Det understrekes imidlertid at det ikke er forventet at hver kilde vil generere en vurderingsoppgave (QBSSSS, 2002, s. 5). Tabell 2 illustrerer hvordan standardene for det første *sub criteria* for kriterie 3 er uttrykt i læreplanen:

Tabell 2: Eksempel på standardbeskrivelser fra læreplanen i English⁴

Knowledge and application of the constructedness of texts		
The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by	evaluating how, <i>and</i> exploiting the ways cultural assumptions, values, beliefs and attitudes underpin texts	VHA
	identifying and explaining <i>and</i> making use of the ways cultural assumptions, values, beliefs and attitudes underpin texts	SA
	sometimes identifying some attitudes and beliefs in texts	VLA

I disse beskrivelsene skjernes det mellom standarder både ved hjelp av ulike verb (*evaluating; identifying and exploiting; explaining*) og kvantitative begrep (*sometimes; some*). Hver standard inneholder mellom fem og åtte beskrivelser av standarder for hvert kriterium for det høyeste nivået, og mellom to og syv standardbeskrivelser for det laveste nivået. Noen av kriteriene er det altså ikke forventet at elevene skal nå for å oppnå de laveste karakternivåene. Sammenlignbarheten i lærernes karaktervurderinger er betinget av en felles forståelse av disse standardbeskrivelsene og deres fortolkning av forskjellene mellom dem. Vedlagt er de fem standardbeskrivelsene og minimumskravene for *Sound Achievement* for *English* (appendiks 7). Sistnevnte er svært viktig, ettersom mange universiteter krever *Sound Achievement* eller bedre i *English* for å ta opp studenter (QBSSSS, 2002, s. 33-38). Vedlagt er også standardbeskrivelsene for *Mathematics* (appendiks 8).

Læreplanene gir retningslinjer for hva som skal inkluderes i elevenes *folio*. For *English* er kravet 6-7 oppgaver av ulike typer. De relativt detaljerte kravene er knyttet til: *predominantly spoken* eller *written; analytic* eller *imaginative; og supervised conditions* eller *extensive access to human and documental resources* (QBSSSS, 2002, s. 30-31). Vedlagt er en kopi av kravene til *verification folio* for *English* (appendiks 9). Dette er sammen med standardbeskrivelsene rammeverket for læreres utarbeiding av oppgavespesifikke kriterier og standarder, som skal presenteres når en oppgave gis for å vurdere elevenes måloppnåelse.

⁴ Tabellen har blitt rekonstruert for å tilrettelegge for en sammenligning av *sub standards*. Bruken av *kursiv* i standardene indikerer at alle dimensjonene av en beskrivelse vil bli demonstrert på tvers av faget, og ikke nødvendigvis innenfor en enkelt oppgave.

Ekstern moderering

I dokumentet *Moderation Processes for Senior Certification* (2005) – ofte omtalt som *the moderation handbook*, forklarer QSA prosedyrene for ekstern moderering. Formålet er å sikre at «students who take the same authority subject in different schools and who achieve the same standard through the assessment programs based on a common syllabus will be awarded the same level of achievement» (QSA, 2005b, s. 11). For hvert fag er det *review panels* bestående av erfarne praktiserende lærere, som veileder skoler i hvert trinn modereringsprosedyrene (Pitman, O'Brien og McCollow, 1999, s. 9). I de vanligste fagene er det både statlige *review panels* og opp til 13 distriktpaneller, som i *English* og *Mathematics B*. Hvert distriktpanel består av mellom 10 og 20 lærere.

Læreplaner utarbeidet av QSA danner grunnlaget for skolenes utvikling av deres egne *work programs*. Alle opplæringsprogram må godkjennes av distriktpanelet før implementering. Denne godkjenningprosessen (*approval*) er det første trinnet i modereringsprosedyrene. I det andre trinnet (*monitoring*), som foregår omtrent halvveis i opplæringsløpet i faget (etter *Year 11* for de fleste fag), sender skoler inn fem eksempler på elevmapper, som inneholder både opplæringsprogrammet og elevenes arbeid. Distriktpanelene gir skolene tilbakemelding på deres implementering av læreplanen og bruk av kriterier og standarder.

Det tredje trinnet består av kvalitetssikring (*verification*) av skolenes sluttvurdering, og skjer ved avslutningen av den tredje av fire terminer i *Year 12*. Skolene sender inn ni eksemplarer av elevmapper, som inneholder elevens arbeid (besvarelser av seks eller syv oppgaver), lærernes vurdering av elevenes oppnåelse av målene og tilhørende opplæringsprogram som inneholder beskrivelser av konteksten for oppgavene (for eksempel om de ble utført med eller uten tilgang til hjelpemidler og tiden elevene hadde til rådighet). Valg av elevmapper skjer tilfeldig, stratifert etter elevenes måloppnåelse, for å kontrollere medium- og terskelnivåer for standardene lærerne bruker.

Panelmedlemmene forhåndsvurderer elevmappene før verifiseringsmøtet, der to og to panelmedlemmer bytter mapper for kryssjekking. Deretter gjennomføres diskusjoner der man søker å oppnå enighet om karakternivå, før rapportering til panelets leder. Panelets leder er ansvarlig for å veilede skolene. Dersom panelets vurderinger avviker vesentlig fra lærerens, gjennomfører panellederen forhandlinger med skolen. I ytterste konsekvens kan en skoles elevmapper sendes til det statlige panelet og QSA kan fastsette endelig karakter, men normalt oppnås det enighet gjennom forhandlingene mellom skole og distriktpanel (Maxwell, 2006, s. 7). Dersom man oppnår enighet om at en skole har gitt noen av elevenes eksempelmapper høyere eller lavere karakter enn de skulle ha hatt, så skal karakternivået også til resten av elevene bli justert

tilsvarende. Dette kan påvirke hvilke resultater det vil være nødvendig for elever å oppnå i den siste terminen, for å oppnå et særskilt karakternivå på vitnemålet (QSA, 2005b). Det siste trinnet (*confirmation*) i modereringsprosedyrene er å utferdige vitnemål etter at den siste terminen er avsluttet. Skolenes sluttvurdering må godkjennes av distriktspanelene.

Det statlige panelets oppgave er å sikre at standardene i deres respektive fag er sammenlignbare på tvers av distriktspanelene. Dette søker man å oppnå ved å kvalitetssikre eksempler på verifiseringer av skoler i hvert distrikt. Denne sammenligningsfasen avsluttes ved at leder for det statlige panelet skriver under på at enighet er oppnådd (QSA, 2005b, s. 12).

2.2.3 Organisering og innhold i Queenslands lærerutdanning

Høyere utdanning er i Australia regulert gjennom nasjonal lovgivning, men denne er ikke særlig detaljert og inneholder ingen spesifikke føringer for lærerutdanning. Dette er derfor et statlig anliggende. I motsetning til Norge, har ikke Queensland noen detaljerte statlige føringer på organiseringen av lærerutdanningen. Staten har imidlertid en lang tradisjon for registrering av lærere. *Board of Teacher Education (BTE)* ble grunnlagt i 1971 for å overvåke standarder og praksiser i lærerutdanningen, og for å utvikle og opprettholde et register over godkjente lærere. «With unions, employers and teachers recognising the benefits of teacher registration, Queensland became the first state to make registration compulsory for employment as a teacher in any school in 1975» (McMeniman, 2004, s. 19). Siden da har Queensland vært en pioner i utviklingen av lærerregistrering både i Australia og internasjonalt. Lovgivningen har i stor grad vært knyttet til «disciplinary powers and the need to introduce tighter controls in relation to child protection, rather than legislative change relating to the regulation of professional practice and professional standards» (McMeniman, 2004, s. 19).

Krav til lærerutdanningsprogrammet som studentene i denne undersøkelsen har gjennomført, er forankret i *Education (Teacher Registration) Act 1988*. Loven slo fast at BTE var ansvarlig for registrering av lærere i Queensland med rapporteringsplikt til utdanningsdepartementet i delstaten. BTE var også ansvarlig for å:

confer and collaborate with teacher employing authorities, teacher education institutions, the teaching profession, teacher organisations and the general community in relation to standards of courses of teacher education acceptable for the purpose of teacher registration and to advise the Minister accordingly (BRT, 2002, s. 1).

Programmet gikk ut på at universiteter som ville tilby lærerutdanning først måtte igjennom en *programutviklingsfase*, der man sender inn et utkast til lærerutdanningsprogram med begrunnelse

for hvordan man ønsker å organisere teori- og praksisopplæringen og hvilket faglig innhold man vil tilby. På bakgrunn av tilbakemeldinger fra BTE måtte man deretter revidere og få godkjent lærerutdanningsprogrammet for *initiell implementering*. Etter lærerutdanningsprogrammet var introdusert, måtte man *kontinuerlig*, hvert fjerde år, få godkjent lærerutdanningsprogrammet. Dokumentet *Professional Standards for Graduates and Guidelines for Preservice Teacher Education Programs*, beskriver akkrediteringsprosedyrene som er relevant for denne studien (BRT, 2002, s. 16). En kopi som illustrerer disse prosedyrene er vedlagt i appendiks 10.

Med *Education (Queensland College of Teachers) Act 2005*, gjeldende fra 1. januar 2006, kom en ny modell for kvalitetssikring av lærerprofesjonen, som erstattet BTE. Det er en institusjon som er finansiert av lærerprofesjonen selv, med ansvar blant annet for å utvikle og forvalte profesjonelle standarder for å bli tatt opp i og forbli medlem av lærerprofesjonen, og for å akkreditere lærerutdanningsprogrammer (QCT, 2007, s. 4). De nye godkjenningsmyndighetene har utarbeidet nye *Program Approval Guidelines (2007)* og *Professional Standards for Queensland Teachers (2007)*. En oppdatert utgave av førstnevnte kom i januar 2009⁵.

Lærerutdanningsprogrammet for det spesifikke caset

Universitetene i Queensland tilbyr ulike typer lærerutdanningsprogrammer. Derfor kan ikke et «universelt» program presenteres, slik det kan for allmennlærerutdanningen i Norge. I tabell 3 presenteres lærerutdanningsprogrammet som ligger til grunn for denne studien. Det presiseres at andre universiteter kan ha organisert utdanningen vesentlig annerledes.

⁵ For mer om Queenslands nye ordninger for registrering og godkjenning av lærere, og akkreditering av lærerutdanningsprogrammer, se nettsidene til *Queensland College of Teachers*: <http://www.qct.edu.au>.

Tabell 3: Obligatoriske fag i den fireårige secondary-lærerutdanningen som er utforsket⁶

Teacing area content	Teaching Area 1		Teaching Area 2	
	Teaching Area 1		Teaching Area 2	
	Teaching Area 1		Teaching Area 2	
	Teaching Area 1		Teaching Area 2	
Curriculum	Teaching Area 1 Curriculum A	Teaching Area 2 Curriculum A	Teaching Strategies	Professional practice 1
	Teaching Area 1 Curriculum B	Teaching Area 2 Curriculum B	Educational Psychology	Professional practice 2
	Assessing Student Learning	Teaching to difference	The Teacher Researcher	Professional practice 3
	Education as social construction	Philosophy and values in Education	Work, education and vocational education	Elective
(Credits) Studiepoeng	(10) 7,5	(10) 7,5	(10) 7,5	(10) 7,5

2.3 Sammendrag av hovedforskjeller mellom Norge og Queensland

Tabell 4 og 5 gir en systematisk oversikt over likheter og forskjeller mellom henholdsvis regulering av læreres arbeid med elevvurdering og organisering av lærerutdanningen i Norge og Queensland.

⁶ For å gjøre sammenligning enklere er de australske studiepoengene (credits) i tabellen og i den videre teksten omregnet til norske studiepoeng, basert på en forutsetning om at omfanget av begge utdanningene er fulltidsstudium (i Norge 30 studiepoeng, i Queensland 40 credits, per semester).

Tabell 4: Regulering av læreres arbeid med elevvurdering

ASPEKTER	NORGE	QUEENSLAND
Læreplaner	<p>Gamle: Én læreplan for hvert fag i grunnskolen (1-10), gjeldende fra 1997. Én for hvert fag i videregående gjeldende fra 1994.</p> <p>Ny av 2006: Gjennomgående læreplaner for fagene.</p>	<p>Gamle: Én læreplan for hvert fag i Years 1-10 gjeldende fra 2005 (English) og 2004 (Mathematics). Delvis forankret i nasjonale føringer.</p> <p>QCAR drafts 2007: Queensland Curriculum, Assessment and Reporting (QCAR) framework med «Essential learning and standards draft 2» av 2007. Ferdig implementert i 2009.</p>
Nasjonale prøver	<p>Kontroversielle nasjonale prøver i 2004 og 2005. Stopp i prøvene 2006. Nye prøver i lesing på norsk og engelsk og regning på 5. og 8. trinn fra 2007. Systemet varslet utvidet til videregående trinn.</p>	<p>QSA gjennomførte statlige <i>literacy</i> og <i>numeracy</i> prøver fram til 2007. Fra 2008 gjennomføres nasjonale <i>literacy</i> og <i>numeracy</i> prøver på trinn 3, 5, 7 og 9. «Queensland Core Skills» (QCS) prøve i slutten av <i>Year 12</i>, for kalibrering av elevenes vitnemål i beregningen av <i>Overall Performance</i> (OP) som legges til grunn ved opptak til høyere utdanning.</p>
Eksamener	<p>Det gjennomføres lokale og eksterne eksamener på 10. trinn og ved avslutningen av hvert fag på videregående.</p>	<p>Ingen eksterne eksamener.</p>
Retningslinjer for elevvurdering	<p>Retningslinjer i forskrift til Opplæringsloven.</p> <p>Ingen retningslinjer i det gamle læreplanverket.</p> <p>Foreløpig ingen retningslinjer i det nye læreplanverket.</p>	<p>Retningslinjer i læreplanene både for <i>Year 1-10</i> og <i>Year 11 & 12</i> som beskriver prosedyrer for å sikre bredde i evidence og konsistens i lærernes vurderinger. Prosedyrer for ekstern moderering for <i>Year 11 & 12</i>.</p>

Tabell 5: Organisering og innhold i lærerutdanningene

ASPEKTER	NORGE	QUEENSLAND
Lengde	4 år (240 studiepoeng)	4 år (240 studiepoeng)
Spesialisering mot aldersgrupper	Generelt program: Spesialisering mot barnetrinnet eller ungdomstrinnet kan gjøres ved valg av undervisningsfag i 3. og 4. studieår.	Program spesialisert mot <i>Primary</i> eller <i>Secondary</i> alle fire år.
To første år:	De to første årene består kun av obligatoriske fag: <ul style="list-style-type: none"> - 30 studiepoeng Norsk - 30 studiepoeng Matematikk - 30 studiepoeng Pedagogikk - 20 studiepoeng KRL - 10 studiepoeng GLSM 	De to første årene er bare opplæring i «content» i tilknytning til de to valgte undervisningsfagene, 60 studiepoeng i hvert av de to undervisningsfagene.
To siste år	Valgfri fordypning: 30 studiepoengs fordypningsenheter i to til fire undervisningsfag, eventuelt rettet spesifikt mot barne- eller ungdomstrinnet. Mulighet for valg av andre temafag.	Obligatorisk fordypning: 15 studiepoeng (<i>Curriculum A og B</i>) i hvert av de to undervisningsfagene (totalt 30 studiepoeng). Åtte obligatoriske pedagogiske fag (alle 7,5 studiepoeng, totalt 60 studiepoeng). Ett valgfag (7,5 studiepoeng).
Praksis – omfang	20-22 uker som inngår i studiepoengene for fagene.	15 uker fordelt på tre separate studieenheter i 5, 6, og 7. Semester (totalt 22,5 studiepoeng).
Praksis – fag	Følger veilederlærerens timeplan (som kan inneholde en rekke fag). Det tilstrebes å finne lærere med fag som er relevante for studentenes fagfordypninger.	Følger veilederlærerens timeplan (som sjelden har mer enn to fag). Det tilstrebes å finne lærere med fag som er relevante for studentenes fagfordypninger.
Elevvurdering	Ikke noe særskilt fag.	Det obligatoriske faget <i>Assessing Student learning</i> (7,5 studiepoeng).

3 Teoretiske perspektiver på elevvurdering

I dette kapitlet trekker jeg fram relevant teori om elevvurdering. Siktemålet er å belyse hvorfor gode kunnskaper og ferdigheter knyttet til elevvurdering er en nødvendig del av læreres kompetanse. Gjennomgangen vil derfor i hovedsak fokusere på elevvurdering som forskningsfelt, men også diskutere forskning om elevvurdering i lærerutdanningen. Som nevnt innledningsvis er den teoretiske gjennomgangen selektiv, og vil konsentrere seg om pedagogiske utfordringer knyttet til formativ og summativ vurdering i klasserommet.

Mens den internasjonale forskningslitteraturen knyttet til elevvurdering er omfattende, er dette fagfeltet lite utforsket i Norge (jf. kapittel 1.1.1). Mye av teorigrunnet for betydningen av elevvurdering hentes derfor fra internasjonal litteratur. Hensikten med denne gjennomgangen er å få fram sentrale trekk ved elevvurdering, for deretter å kunne analysere i hvilken grad denne kunnskapen beskrives og gjenfortelles som innsikt hos de intervjuede lærerstudenter.

I del 3.1 avklares begrepene *vurdering*, *evaluering* og *measurement*. Jeg gir en kort forklaring på hvorfor vurderingsbegrepet foretrekkes, og deler deretter begrepet inn i formativ og summativ vurdering, som beskrives som to ulike aspekter ved lærerens vurderingsarbeid. I del 3.2 gjennomgås utfordringer forskning har avdekket knyttet til de summative aspektene ved lærerrollen. Vurderingskriterier og utfordringer knyttet til validitet og reliabilitet i summative vurderinger diskuteres. I del 3.3 presenteres teori om formativ vurdering, etterfulgt av diskusjoner om virkningen av faglige tilbakemeldinger. Kjennetegn på kvalitet i formative vurderinger drøftes.

Del 3.4 er omhandler kunnskap om elevvurdering i Norge. Viktige funn fra relevante undersøkelser diskuteres. Jeg gjør en kort historisk gjennomgang av retningslinjer for elevvurdering, og framholder særskilte utviklingstrekk som mulige forklaringer på problemer i Norge knyttet til både formativ og summativ vurdering. I del 3.5 gir jeg en kort oversikt over elevvurdering i Queensland. Positive og negative erfaringer med Queenslands tilnærming til elevvurdering drøftes. I del 3.6 diskuteres kort problemstillinger knyttet til elevvurdering i morsmåls- og matematikkfagene, spesielt i Norge, men også i Queensland. Hensikten her er å få fram særtrekk og forskjeller mellom fagene. Del 3.7 omhandler elevvurdering i tilknytning til lærerutdanning. Jeg peker på betydningen av kompetente lærere generelt, og knyttet til elevvurdering spesielt. I del 3.8 presenterer jeg et sammendrag av de viktigste teoretiske aspektene som har blitt gjennomgått i kapitlet, og som sammen med intervjudata utgjør grunnlaget for oppgavens drøftningskapittel.

3.1 Diskusjon av kjernebegreper i vurderingsterminologien

I Norge har begrepene *vurdering* og *evaluering* historisk blitt brukt om hverandre med samme begrepsinnhold. Med 90-årenes reformer ble *vurdering* imidlertid etablert som det foretrukne begrep knyttet til lærerrollen (Engh, Dobson og Høihilder, 2007, s. 26). I forskriftene som regulerer lærerens rolle i karaktersetning og tilbakemelding er nå utelukkende begrepet *vurdering* brukt (Engh m.fl., 2007, s. 26), derfor brukes det som hovedbegrep i denne oppgaven.

Begrepet «vurdering» kommer fra middelnedertysk «werdêren», som i følge Trond Ålvik (1991) betyr «å verdsette, gjøre seg opp en mening om, tillegge verdi, bestemme verdien av noe» (Ålvik, 1991, s. 13). Han peker på at nærheten mellom opprinnelig og aktuell bruk av ordet er tydeligere på svensk, hvor «värdering» bygger direkte på «värde» (verdi). Tilsvarende bygger det engelske «evaluate» på kjernen «value». Ålvik mener det rent språklig ikke er noe som skiller «vurdering» og «evaluering», ettersom begge handler om å verdsette eller bedømme, og begge har sin rot i «verdi» (Ålvik, 1991, s. 13). Forskjellen er altså at vi har hentet «vurdering» fra det tyske språket, og «evaluering» fra engelsk.

Innenfor en engelskspråklig pedagogisk kontekst, er det likevel nødvendig å belyse ulikheter i terminologien knyttet til begrepene «assessment» (vurdering) og «evaluation» (evaluering). *Assessment* er i følge Lysne (1999) opprinnelig hentet fra psykometrien. Fokuset var på individet og det ble tilstrebet å oppnå objektivitet ved å begrense kommunikasjon mellom eleven og den som forestod testingen. Den klassiske betydningen kommer i følge Wohl (1967) til uttrykk i det psykometriske postulatet om at «if something exists, it does so in some measurable amount» (Lysne, 1999, s. 307). Etter hvert som elevinvolvering har blitt sentralt i teori om *assessment* (hvilket bryter fullstendig med den opprinnelige psykometriske forståelsen), har den snevrere psykometriske forståelsen blitt uttrykt gjennom nettopp begrepet *measurement*.

Evaluering har i Norge etter hvert blitt knyttet mer til kvalitetsvurdering av institusjoner og programmer. Denne forståelsen finner vi igjen i amerikansk fagterminologi der *evaluering*, i følge Bloom (1967), først og fremst brukes i forbindelse med vurdering av innholdsvaliditeten av læreprogram, undervisningsteknikker, prosesser og lærestrategier. *Assessment* ble knyttet til måling og bedømmelse av elevenes utbytte av undervisningen og læringsaktivitetene (Lysne, 1999, s. 308).

Siden har begrepet *assessment* fått et mer omfangsrikt innhold der læringsaspektet blir mer vektlagt. I den ofte siterte artikkelen *Inside the black box: Raising standards through*

classroom assessment, ga Paul Black og Dylan Wiliam (1998) følgende definisjon av assessment: «all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged» (Black og Wiliam, 1998b, s. 2).

Denne definisjonen går langt i å vektlegge tilbakemeldingsaspektet, hvilket kan sies å gjenspeile en trend i forskningsfeltet. Her skal vi forlate både *measurement* og *evaluation* og forsøke å definere elevvurdering nærmere med utgangspunkt i *assessment*-begrepet. Flere av kildene det refereres til i denne oppgaven bruker imidlertid ikke den samme begrepsavklaringen. Når jeg refererer til kilder der begrepene *evaluation* eller *measurement* er brukt, så er det innenfor den brede forståelsen av elevvurdering (assessment) jeg her har forsøkt å skape.

High stakes assessment er et uttrykk som brukes mye i internasjonal forskningslitteratur for å beskrive vurderinger som er risikofylte (*the stakes are high*). En eksamen eller en prøve som inngår i standpunkt karaktergrunnet utgjør en høy risiko for eleven, ettersom resultatet er viktig for kvalifisering til studier eller jobb. Vurderinger kan også være risikofylte for større grupper eller organisasjoner, som for eksempel en skole eller kommunes samlede resultat på nasjonale komparative tester. Generelt kan man si at validitet og reliabilitet er svært viktig ved *high stakes* vurderinger, mens det ikke er like viktig der prøveresultatene har små eller ingen konsekvenser (*the stakes are low*) (Matthiesen, 2007, s. 49; Tveit, 2008b).

Det er imidlertid viktig å skille mellom om for eksempel en prøve er *high stakes* for elevene som individer på den ene siden, eller for elevene som gruppe, skolen, kommunen på den andre siden. De nasjonale prøvene i norsk grunnskole er eksempel på vurderinger som til en viss grad er risikofylte for skoler og skoleeier (først og fremst som følge av prestisjen som følger med offentliggjøring og medienes rangering av skolenes resultater), mens de er *low stakes* for elevene ettersom de gjennomføres helt på starten av 5. og 8. trinn, og dermed ikke kan tillegges særlig vekt når standpunkt karakteren settes på slutten av 10. trinn (Tveit, 2007a, s. 203). Risikoaspektet for skolene er imidlertid lite ved de nasjonale prøvene i Norge sammenlignet med *accountability*-tester i Storbritannia og USA, der skoler i verste fall risiker mindre pengeoverføringer eller å bli stengt dersom elevene gjør det dårlig (Darling-Hammond, 2004). Ettersom fokuset i denne oppgaven er på lærerens vurderingsarbeid, skal jeg ikke drøfte risiko knyttet til eksamener og nasjonale prøver nærmere. I stedet rettes fokuset på prøver, prosjekter og oppgaver som lærere utarbeider og vurderer, som legges til

grunn ved den endelige karaktersettingen, og som derfor er *high stakes* for elevene. For disse vurderingene må man stille strenge krav til validitet og reliabilitet.

I den internasjonale forskningslitteraturen om elevvurdering har et begrepspar blitt særlig populært: *formative* og *summative assessment*. Skillet har sitt utgangspunkt i Michael Scrivens definisjon fra 1967:

Formative evaluation is evaluation designed, done, and intended to support the process of improvement, and normally commissioned or done by, and delivered to, somebody who can make improvements.

Summative evaluation is the rest of evaluation: in terms of intentions, it is evaluation done for, or by, any observers or decision makers (by contrast with developers) who need evaluative conclusions for any other reasons beside development (Scriven, 1967, s. 20).

Begrepene formativ og summativ vurdering har siden blitt brukt i skolesammenheng til å illustrere de ulike aspektene ved elevvurdering. Mange ulike fortolkninger av sammenhengen mellom de to begrepene har utviklet seg siden. Wynne Harlen (2006) drøfter hvorvidt det er egnet å diskutere forholdene mellom begrepene som en *dikotomi* eller som en *dimensjon*. Hun peker på at det viktigste for å forstå skillet mellom de to begrepene er erkjennelsen at vurderingens *formål* er forskjellig (Harlen, 2006, s. 116). Resonnementet framgår av hennes presisering av hva summativ vurdering er:

The term «summative assessment» refers to an assessment with a particular purpose – that of providing a record of a student’s overall achievement in a specific area of learning at a certain time. It is the purpose that distinguishes it from assessment described as formative, diagnostic, or evaluative, rather than any particular method of gathering information about students’ performance (Harlen, 2004, s. 1).

Mens det i teorien kan høres enkelt ut å skille mellom disse to formålene, kan det for lærere i praksis være svært vanskelig. Formativ og summativ vurdering har to forskjellige funksjoner, og det kan være problematisk at læreren er ansvarlig for begge. En parallell til jussen kan bidra til å illustrere dette:

I en rettssak er det utenkelig at forsvaren og dommeren er samme person. I skolen er imidlertid læreren den viktigste forvalteren av disse to oppgavene. I den formative rollen trer læreren fram som elevens *forsvarer*, og skal hjelpe han eller henne til best mulig resultater med utgangspunkt i elevens forutsetninger. I den summative rollen skal læreren felle en «uhildet» *dom* over elevens resultater med utgangspunkt i felles mål og retningslinjer for

karaktersetting. Relevansen av denne parallellen vil variere mellom utdanningssystemer, men med mindre all risikofylt vurdering er overlatt til eksterne eksamener, gjør denne dobbeltrollen seg gjeldende (Tveit, 2007a, s. 199).

I engelskspråklig forskningslitteratur kan man se en økende trend mot ansvarliggjøring (*accountability*) av skoler og lærere for elevenes læringsresultater, hvilket medfører økt fokus på summative vurdering. En annen trend er imidlertid at utdanningsmyndigheter, i respons til svake resultater på internasjonale komparative (summative) studier, som for eksempel PISA-prøvene, erkjenner betydningen av å også legge vekt på formativ vurdering (OECD, 2005, s. 22). Sagt annerledes: En sterk satsing på formative vurdering forventes å gi gevinst i de summative målingene.

I denne oppgaven benyttes følgende to definisjoner av lærerrollen til å klargjøre lærerens mandat knyttet til elevvurdering.

- Lærerens summative oppgaver: *Å gjøre valide og pålitelige vurderinger av elevenes oppnåelse av målene i læreplanen.*
- Lærerens formative oppgaver: *Å gi elever tilbakemeldinger på hvordan de ligger an i forhold til målene i læreplanen og hva de bør gjøre for å oppnå bedre læringsresultater.*

Som det framgår av definisjonene utgjør læreplanen den sentrale referansen både for formativ og summativ vurdering. Nedenfor drøfter jeg betydningen av disse to dimensjonene ved lærerrollen med utgangspunkt i internasjonal forskning. Summativ vurdering diskuteres først, ettersom grunnlaget for lærernes vurdering av elevenes måloppnåelse burde være retningsgivende for de formative vurderingene.

3.2 Utfordringer ved summative vurderinger

Lærerens rolle knyttet til summativ vurdering dreier seg om å gjøre valide og reliable vurderinger av elevenes læringsresultater. Enkelt uttrykt handler validitet (gyldighet) i denne sammenhengen om hvorvidt vurderingsoppgavene reflekterer de kunnskaper, ferdigheter og verdier som er ment vurdert. Reliabilitet (pålitelighet) handler om vurderingsoppgavenes konsistens og stabilitet – resultatet bør være det samme uavhengig av hvem som vurderer elevsvarene (Brady og Kennedy, 2005, s. 4). Som summativ og formativ vurdering, er validitet og reliabilitet nært knyttet sammen. Teoretisk sett kan en vurdering anses å være reliabel uten å være valid, men den kan ikke være valid uten å også være reliabel. Hensynet til validitet blir da også ofte overskygget av hensynet til reliabilitet. Et kjent sitat fra boken

Alice's Adventures in Wonderland av Charles Dodgson (1865), også kjent som «Lewis Carroll», kan brukes til å illustrere dette:

«...how can you possibly award prizes when everyone missed the target?» said Alice.

«Well» said the Queen, «some missed by more than others and we have a fine normal distribution of misses, which means we can forget about the target» (Dodgson, 1865).

Det gir liten mening, som kan forstås som det underliggende budskapet i dette sitatet, å gjøre en reliabel vurdering dersom målingen ikke har noen relevans overhodet. Følgelig bør hensynet til reliabilitet ivaretas på en måte som er forenlig med hensynet til validitet. Jeg vil her gå inn i en litt grundigere diskusjon av hva som kjennetegner valide og reliable vurderinger av elevers læringsresultater.

3.2.1 Validitet og reliabilitet

En mer presis definisjon av validitet enn den ovennevnte ble utarbeidet av Messick, som definerte validitet som en «overall evaluative judgment of the degree to which empirical evidence and theoretical rationale support the adequacy and appropriateness of interpretations and actions based on test scores and other modes of assessment» (Messick, 1996, s. 7) Det er *meningen* av resultatene, og ikke egenskaper ved prøven eller vurderingen som sådan, som vurderes. Ettersom slike meninger er en sosial *konstruksjon*, brukes «construct validity» som en mer avgrenset definisjon av validitet i psykometrisk litteratur. Et «construct» defineres som en psykologisk forestilling som gjenspeiler et teoretisk felt. I skolesammenheng kan eksempelvis et fag, eller avgrensede emner innenfor et fag, sees på som et «construct». I følge Messick (1996) er det to hovedkilder til ugyldighet i en vurdering:

- *Construct under representation*: The test is too narrow and fails to include important dimensions (...).
- *Construct-irrelevant variance*: the assessment is too broad, containing excess reliable variance that is irrelevant to the interpreted construct.

For læreren og eleven representerer læreplanen det grunnleggende konstruktet i et fag. Læreplaner og fagplaner uttrykker ofte både målsetninger for kunnskapsinnhold som skal læres, og anvisninger i prosesser for å tilegne seg disse kunnskapene. For at vurderinger av elevers læringsresultater skal ha høy validitet, må oppgavene eller prøvene gjenspeile dette konstruktet på en tilfredsstillende måte. For å lykkes med dette, må den interne strukturen på vurderingsoppgaver utformes på en måte som er konsistent med den interne strukturen i faget

(Black og Wiliam, 2006, s. 120). Læreren kan umulig vurdere om elevene har nådd *alle* målene i læreplanen. Han må derfor sørge for at de forskjellige vurderingsoppgavene samlet sett måler målene i læreplanen på en representativ måte.

Dette strukturelle aspektet ved validitet er nært knyttet til reliabilitet. Dersom oppgavene ikke er strukturert på en representativ måte, kan elever som «egentlig» er på samme nivå prestere bedre eller dårlige på grunn av tilfeldigheter knyttet til oppgavene de får. En av modellene for å kvalitetssikre prøvers «indre konsistens» er *Cronbachs alfa* (Lie, 2007, s. 88). Slike kvantitative prosedyrer kan benyttes i standardiserte prøver eller andre prøver utformet med utgangspunkt i avansert programvare, men er i praksis vanskelig gjennomførbart for lærere, og drøftes derfor ikke nærmere her. Læreren bør imidlertid gjøre en kvalitativ vurdering av hvorvidt det samlede vurderingsgrunnlaget i faget utgjør en representativ vurdering av målene i læreplanen.

En annen feilkilde knyttet til reliabilitet er at elevens prestasjoner kan variere avhengig av hvilken dag han tar prøven (Black og Wiliam, 2006, s. 120). Dette er et mye brukt argument mot ensidig vektlegging av eksamener i vurderingssystemer. I vurderingssystemer som utelukkende baseres på eksterne eksamener, er det i praksis umulig å ivareta dette hensynet til reliabilitet i utstrakt grad, ettersom det er svært kostbart å avholde slike prøver. Vurderinger foretatt av læreren, på den annen side, gir gode muligheter for å ivareta dette reliabilitetshensynet, ettersom vurderingsgrunnlaget kan spres utover en rekke prøver eller andre vurderingsordninger i løpet av skoleåret.

En tredje feilkilde, som nok er den de fleste forbinder med reliabilitet, er såkalt sensorreliabilitet: Forskjellige bedømmere kan vurdere samme besvarelse ulikt. Måling av dette kalles gjerne *inter-rater reliability*. I forbindelse med standardiserte tester er det vanlig å benytte minst tre strategier for å begrense denne feilkilden:

1. Å velge gode sensorer og skolere dem godt;
2. Utarbeiding av strenge retningslinjer for framgangsmåte og kriterier for karaktersetting;
3. Kryssretting der to eller flere sensorer vurderer besvarelsene og enes om et karakternivå (Black og Wiliam, 2006, s. 121).

Ved standardiserte prøver er ofte elementer av alle disse tre strategiene valgt. I praksis er man av økonomiske hensyn likevel tvunget til å gjøre kompromiss. Hvor godt man lykkes avhenger av: (1) kvaliteten på sensorene og skoleringen av dem; (2) kvaliteten på retningslinjer og kriterier; og (3) antall sensorer.

Et problem ved klasseromsbaserte vurderinger er at det ofte ikke eksisterer noen felles referanse på tvers av skoler, som vurderingene kan relateres til. I følge Sadler er eksplisitte kriterier nødvendig for å gjøre pålitelige vurderinger enten de som vurderer er elevenes lærere eller profesjonelle bedømmere (R. Sadler, 1986, s. 1-3). Det forutsetter trolig en viss utstrekning av standardiserte oppgaver, hvilket tradisjonelt kun har blitt brukt på en systematisk måte i forbindelse med eksterne eksamener. I artikkelen *Subjektive bedömningar och objektiva tolkningar* framholder Christina Wikström (2007) tydelige målbeskrivelser og vurderingskriterier som avgjørende for å gjennomføre en pålitelig karaktersetning. Hun peker på hvor krevende elevvurdering er uten slike hjelpemidler. «Utan dessa viktig styrdokument och hjälmedel som til exempel nationella prov står läraren inför ett nästintill omöjligt, men framförallt otacksamt, uppdrag» (Wikström, 2007, s. 33).

Under overskriften *Hur blir vi bättre på att bedöma?*, framholder Jörgen Tholin (2007) «vekollaborativt arbetssätt» som en viktig forutsetning. Han peker på at det svenske Skolverket i flere sammenhenger har lagt vekt på lærersamarbeid i vurderingsarbeidet, for eksempel ved at man bedømmer hverandres elever eller er to lærere som bedømmer samme prestasjoner (Tholin, 2007, s. 16). Tholin trekker fram flere fordeler med en slik praksis: Foruten at bedømmingene blir mer rettferdig for elevene, så får læreren hjelp i vurderingsarbeidet og man utvikler sammen med sine kolleger et profesjonelt språk knyttet til vurdering (Tholin, 2007, s. 17). Tholin framholder et «bred repertoar av bedömningsinstrument» som en annen viktig forutsetning for at bedømmingene skal bli rettferdige (Tholin, 2007, s. 19).

3.2.2 Studier av validitet og reliabilitet

Standardiserte prøver med eksterne sensorer har stor troverdighet blant folk flest, på bakgrunn av *uavhengigheten* de representerer. Men studier av reliabiliteten og validiteten til standardiserte prøver tyder på at de ikke er så pålitelige som mange tror. En studie av Rogosa (1999) viste for eksempel at 43 prosent av elevene på en matematikkprøve ble feilklassifisert med mer enn ti prosentpoeng (Black og Wiliam, 2006; Gardner og Cowan, 2005, s. 123). Gardner og Cowan presenterte i 2005 funnene fra en omfattende studie av påliteligheten ved *11-plus* testene i Nord-Irland, gjennomført skoleåret 1999-2000. *11-plus* består av to parallelle tester som dekker fagene *English* og *Mathematics*, og resultatene er direkte avgjørende for opptak til *Grammar schools*. Ved å undersøke både prøvenes indre konsistens og konsistensen på tvers av dem, ble det avdekket at en kandidat som oppnådde toppkarakter, i virkeligheten kunne ha en «true score» hvor som helst innenfor de fire øverste av de seks karakternivåene.

Gitt at 6-7000 kandidater fikk skoleplass, var det grunn til å tro at omtrent 3000 av dem var feilklassifisert enten ved at de fikk en plass de ikke skulle hatt, eller motsatt (Gardner og Cowan, 2005, s. 115).

Disse eksemplene viser at det er svært krevende å gjennomføre valide og reliable vurderinger av elevers læringsutbytte, til og med når man ved standardiserte prøver benytter omfattende prosedyrer for å ivareta pålitelighet i vurderingene. Det er enda mer krevende for lærere å gjennomføre valide og reliable vurderinger av elevers læringsresultater. Som påpekt ovenfor finnes det likevel en del strategier som kan bidra til å minske tilfeldigheter i vurderingene.

3.2.3 Oppsummerende oversikt over krav til summativ vurdering

Med utgangspunkt i diskusjonen overfor, har jeg utarbeidet noen grunnleggende problemstillinger knyttet til validitet og reliabilitet ved læreres summative vurderinger. Som vi har sett er det viktig:

- A. at læreren har tilstrekkelige strategier for å sikre at de enkelte oppgavene måler de målene i læreplanen som de gir seg ut for å måle (*construct validity*);
- B. at de forskjellige oppgavene læreren utarbeider samlet sett gjenspeiler målene i læreplanen på en representativ måte (*structure validity*/konsistens på tvers av prøver);
- C. at læreren gir elevene tilstrekkelig muligheter til å vise hva de kan (*day-to-day reliability*); og
- D. at læreren tilrettelegger for at andre lærere kan gjøre en vurdering av det samme vurderingsgrunnlaget (*inter-rater reliability*)

Punkt B og C har potensial til å bli møtt best ved summativ vurdering gjennomført av læreren, ettersom det i praksis forutsetter en rekke prøvesituasjoner, hvilket vil være for kostbart å gjennomføre som eksterne prøver. Punkt A kan potensielt bli møtt best ved summative eksterne prøver, ettersom man kan gjennomføre omfattende valideringsprosedyrer. Punkt D har tradisjonelt blitt møtt best med eksterne prøver, der man har to ulike sensorer, men er fullt mulig å praktisere i en vanlig klasseromskontekst. Lærere bør så langt det lar seg gjøre forsøke å ivareta alle hensynene, for å sikre en valid og reliabel vurdering av elevenes læringsresultater.

3.3 *Utfordringer ved formative vurderinger*

I 1998 presenterte Black og Wiliam en omfattende gjennomgang av 250 publikasjoner basert på forskning om formativ vurdering. Gjennomgangen viste at formativ vurdering er blant de mest effektive tiltakene for å heve nivået på elevers kunnskaper som noen gang er rapportert (Black og Wiliam, 1998a, s. 61). I den nevnte vesentlig kortere og mer politisk orientert publikasjonen (Black og Wiliam, 1998b) anbefalte forfatterne en rekke tiltak for å lykkes med formativ vurdering. Jeg velger her å først gå inn i analysegrunnlaget for hovedpublikasjonen. Det ligger utenfor denne oppgaven å rapportere funnene i tilknytning til alle punktene, hensikten er i stedet å belyse elementer som er sentrale for at læreren skal lykkes med formativ vurdering. Deres gjennomgang av forskningslitteraturen var basert på en tese om at formativ vurdering er en sekvens med to elementer:

- The perception by the learner of a gap between the desired goal and his or her present state (of knowledge, and/or understand, and/or skill) (*vurderingen*).
- The action taken by the learner to close the gap in order to attain the desired goal (*handlingen*) (Black og Wiliam, 1998a, s. 20).

Kunnskap om elevens nåværende nivå tilegnes primært enten gjennom en egenvurdering, eller ved at en annen person, især læreren, gjør en vurdering av elevens nivå, som deretter kommuniseres til eleven. Black og Wiliam identifiserte fire faktorer som påvirker elevens mottakelse av og reaksjon på tilbakemeldinger. Dette knytter seg til elevenes oppfatninger av læringsmålene; egen evne til å respondere; risiko knyttet til ulike måter å respondere; og hvordan et læringsarbeid bør gjennomføres (Black og Wiliam, 1998a, s. 21).

Alle disse faktorene har innvirkning på elevenes motivasjon, valg av handlinger og innsatsvilje (Black og Wiliam, 1998a, s. 21). Det grunnleggende utgangspunktet i Black og Wiliam sine perspektiver på formativ vurdering er at elevene må *forstå* målene. Eleven får så tilbakemelding, enten fra seg selv, læreren eller andre elever, om hvordan han ligger an i forhold til målene. Dette danner grunnlaget for valg av handlinger for å nå målene (Black og Wiliam, 1998a, s. 21). Forfatterne framholder *egenvurdering* som en helt nødvendig forutsetning for at formativ vurdering skal være vellykket. «For formative assessment to be productive, pupils should be trained in self-assessment so that they can understand the main purposes of their learning and thereby grasp what they need to do to achieve» (Black og Wiliam, 1998b, s. 10). De hevder videre at vurderingsoppgaver bør planlegges slik at elever

får mulighet til å *uttrykke* forståelsen av målene underveis i læringsarbeidet (Black og Wiliam, 1998b, s. 10-11).

3.3.1 Bruk av faglige tilbakemeldinger

For at elever skal kunne gjøre en god vurdering av eget arbeid, er de avhengig av gode tilbakemeldinger. Jeg ønsker å rette søkelyset på effekten av tilbakemeldinger, da dette viser seg å være en av de viktigste suksessfaktorene for å bedre elevenes læringsresultater. John Hattie presenterte i 1999 en omfattende syntese av mer enn 500 metaanalyser knyttet til effekten av ulike tiltak iverksatt for å forbedre elevprestasjoner. Analysen inkluderte mer enn 450 000 effektstørrelser fra 180 000 studier, som representerte mellom 20 og 30 millioner elever. Studien identifiserte mer enn 100 forskjellige faktorer som påvirker elevprestasjoner, deriblant egenskaper ved skoler, hjem, elever, lærere og læreplaner. Tilbakemeldingsfaktoren (*feedback*) hadde en gjennomsnittlig effekt på 0,79, nesten dobbelt så høyt som gjennomsnittseffekten (0,40) av faktorer i analysen (Hattie og Timperley, 2007, s. 82). Brutt ned viste imidlertid metaanalysen av tilbakemeldinger at det var store variasjoner i effekten av forskjellige typer tilbakemelding.

The most effective forms of feedback provide cues or reinforcement to learners; are in the form of video-, audio- or computer-assisted feedback; and/or relate to goals. Programmed instruction, praise, punishment, and extrinsic rewards where the least effective for enhancing achievement (Hattie og Timperley, 2007, s. 84).

Vi ser at tre typer tilbakemeldinger skilte seg ut som lite effektive: ros, straff og ytre belønning. At ros har liten effekt henger trolig sammen med at det inneholder lite læringsrelatert informasjon (Hattie og Timperley, 2007, s. 86). Harlen har i en annen metaanalyse funnet grunnlag for å fastslå at «using grades as rewards and punishments is harmful to students' learning by encouraging extrinsic motivation» (Harlen, 2004, s. 5).

Deci, Koestner og Ryan (1999) har påpekt at konkrete belønninger («tangible rewards»: klistremerker, priser, osv) ikke bør kalles tilbakemelding, ettersom de inneholder så lite informasjon. I en metaanalyse av virkningen av ytre belønning for indre motivasjon fant forfatterne en negativ korrelasjon (0-34) mellom ytre belønning og oppgaveprestasjoner. Konkrete belønninger underminerte indre motivasjon, spesielt knyttet til interessante oppgaver (-0.68) sammenlignet med uinteressante oppgaver (0.18). Når tilbakemeldinger ble gitt i en kontrollerende form, var virkningen enda mer negativ (-0.78). Forfatterne konkluderte med at ytre belønning er negativ fordi de «undermine people's taking responsibility for

motivating or regulating themselves» (Deci, Koestner og Ryan, 1999, s. 659). Black og Wiliam har også pekt på at: «where the classroom culture focuses on rewards, 'gold stars', grades or place-in-the-class ranking, then pupils look for the ways to obtain the best marks rather than at the needs of their learning which these marks ought to» (Black og Wiliam, 1998b, s. 8-9).

Det ville føre for langt å diskutere alle tilbakemeldingstypene som viser høy effekt. Hattie og Timperleys er mest opptatt av effekten av tilbakemeldinger knyttet til mål (*goal*). De viser også til en annen metaanalyse av Kluger og DeNisi (1996), og hevder at samlet sett ser det ut til at virkningen av tilbakemelding er påvirket av retningen av tilbakemeldinger relatert til oppgaveprestasjoner. «Feedback is more effective when it provides information on correct rather than incorrect responses and when it builds on changes from previous trails» (Hattie og Timperley, 2007, s. 85). De mener hovedhensikten med tilbakemelding bør være å redusere avviket mellom gjeldende oppnåelse og målet (Hattie og Timperley, 2007, s. 86).

I forbindelse med disse effektstudiene er det nærliggende å vise til Lev Vygotskys klassiske pedagogiske tekst *Interaksjon mellom læring og utvikling* som kom ut på engelsk i boken *Mind in Society* i 1978, men som han forfattet mer enn 40 år tidligere. Vygotskys arbeider har fått en slags renesanse i Norge de siste ti årene. Her brukes en norsk oversettelse fra 1996. I artikkelen lanserte Vygotsky teorien om *sonen for den nærmeste utvikling*, som han definerte på følgende måte:

Avstanden mellom det eksisterende utviklingsnivået slik det bestemmes gjennom selvstendig problemløsning og det potensielle utviklingsnivået slik det bestemmes gjennom problemløsning under voksen veiledning eller i samarbeid med dyktigere jevnaldrende (Vygotsky, 1996, s. 159).

Vygotsky pekte på betydningen av at elevene får tilbakemelding fra voksne eller flinkere jevnaldrende, på hva som skiller deres aktuelle måloppnåelse fra ønsket måloppnåelse, og at det er lærernes oppgave å sørge for at så skjer. Dette er blitt et vanlig pedagogisk prinsipp. Interessant er det derfor å se at tre uavhengige metaanalyser av forskning på effektivitet i opplæringen bekrefter denne teoriens gyldighet, og framholder dette prinsippet som det mest effektive virkemidlet for å øke elevens læringsutbytte.

Tilbakemelding på læringsarbeid kan deles inn i to hovedtyper: muntlige og skriftlige tilbakemeldinger. Muntlige tilbakemeldinger blir i all hovedsak gitt i klasserommet, i forbindelse med konkret og akutt oppgaveløsning, mens skriftlige tilbakemeldinger ofte formidles når elevene får tilbake skriftlige arbeider de har levert. Felles for muntlige og

skriftlige tilbakemeldinger er at de ikke bør være generelle, men knyttet til konkrete problemstillinger og forbedringspunkter. Rask og hyppig tilbakemelding er gunstig ved alle typer tilbakemelding, men ofte lettest å gjennomføre muntlig. Skriftlige tilbakemeldinger gjennom IKT-baserte læringsplattformer begynner imidlertid å bli mer vanlig, og denne formen for korrespondanse åpner for mer frekventert respons enn det som gjerne forbindes med vurdering av tradisjonelle innleveringsoppgaver.

3.3.2 Formativ vurdering og kriterier

Vurderingskriterier kan framholdes som sentrale verktøy for å anskueliggjøre forskjellen mellom de fastsatte målene og elevenes aktuelle måloppnåelse. Kriterier kan utformes på flere måter, fra helt generelle (globale kriterier) til mer eksplisitte kriterier med flere kategorier (analytiske kriterier). At kriteriene er eksplisitte er særlig viktig med tanke på likeverdighet i utdanningen. I boken *La reproduction*, som kom ut i Frankrike i 1970 og siden er blitt en klassiker innenfor pedagogisk sosiologi, presenterte Pierre Bourdieu og Jean-Claude Passeron en kritikk av seleksjonsmekanismene i skolen. De hevdet at sosial reproduksjon ikke entydig forklares med ulike økonomiske betingelser, men først og fremst har å gjøre med kulturelle faktorer. Bourdieus begrep *språklig kapital* henspeiler på hvordan noen kulturelle grupper har et kulturelt fortrinn i konkurransearenaen skolen utgjør. Bourdieu og Passeron framholdt at et tydelig og forståelig språk kan begrense favorisering av elever med kulturell bakgrunn som tilsvarer den herskende skolekulturen (barn av høyt utdannede foreldre) og som derfor har lettere for å forstå implisitt kunnskap (Bourdieu og Passeron, 2006, s. 153). I tråd med deres argumentasjon kan det hevdes at elever med lav sosioøkonomisk bakgrunn tjene på at vurderingskriteriene er mest mulig eksplisitte.

Sadler har pekt på en rekke fordeler ved å bruke eksplisitte vurderingskriterier i den formative vurderingen: De gjør det mulig for lærere og elever å tenke klarere om læringsmålene; kriteriene tar bort en del av uklarheten som kan bli forbundet med vurdering; kriterier gjør det lettere for foreldre å forstå bakgrunnen for oppgavene; kriterier gjør det mulig for elevene å vurdere sitt eget arbeid (egenvurdering); kriterier bidrar til tettere kontakt mellom elever og lærere i forbindelse med oppgaveløsning; og kriterier kompenserer for uerfarne lærere, som vil finne det mye vanskeligere å få tilgang på felles «skjult kunnskap» (R. Sadler, 1986, s. 1-3).

Torrance (2007) har rettet søkelys på mulige kritiske følger av den økende fokuseringen på vurderingskriterier. Han peker på hvordan kriteriebasert vurdering kan bli altopplukende og lede til instrumentalisme: «The more clearly requirements are stated, the

easier it would appear for them to be pursued and accomplished» (Torrance, 2007, s. 291). Han mener den angivelige formative prosessen det er å identifisere og dele vurderingskriterier, og gi tilbakemelding til elever på deres styrker og svakheter i tilknytning til disse, ofte bidrar til å svekke heller enn å styrke elevenes autonomi (Torrance, 2007, s. 291). Her ligger et stort paradoks: *For* eksplisitte kriterier kan lede til en mekanisk eller instrumentalistisk tilnærming til læring. Dette er en viktig problemstilling å peke på, men det ville være urimelig å bevisst være uklar på læringskravene. Dette er heller en diskusjon om detaljgraden av slike vurderingskriterier, og i hvilken grad de bør være fastsatt i læreplanen eller bare utformes spesifikt for konkrete oppgaver.

3.3.3 Oppsummerende oversikt over krav til formativ vurdering

Med utgangspunkt i diskusjonen overfor, har jeg utarbeidet noen grunnleggende problemstillinger knyttet til kvaliteten av læreres formative vurderinger. Det er viktig:

- A. at læreren legger til rette for at elevene kan vurdere egen måloppnåelse;
- B. at læreren ikke overdriver bruken av karakterer og andre symboler som er konkurransefremmende og slik tar oppmerksomheten bort fra konkrete forbedringsstrategier;
- C. at læreren gir elevene hyppige og tydelige tilbakemeldinger på læringsarbeidet; og
- D. at læreren bruker eksplisitte vurderingskriterier til å beskrive kunnskapsmålene og elevenes måloppnåelse.

3.4 Hva vet vi om elevvurdering i Norge?

Som nevnt innledningsvis pågår det et betydelig utviklingsarbeid knyttet til elevvurdering i Norge, som følge av myndighetenes erkjennelse av omfattende problemer knyttet til både retningslinjer og praksis. Kunnskapsdepartementet har slått fast at regelverket ikke oppfattes som klart nok, at både lærerutdanningen og skolen mangler tilstrekkelig kompetanse og at det er svak vurderingskultur og vurderingspraksis i norsk skole (Kunnskapsdepartementet, 2007). En av forklaringene på disse problemene er uklarhet i retningslinjer for elevvurdering, som skriver seg helt tilbake til introduksjonen av Normalplanen av 1939. Siden 1970-årenes karakterstrid har elevvurdering vært et særdeles politisert tema, hvilket har ført til mer ideologiske enn faglig funderte diskusjoner og reformer. Jeg presenterer her et kort historisk tilbakeblikk for å presentere bakgrunnsinformasjon som kan være med på å forklare problemene knyttet til elevvurdering som vi opplever i Norge i dag.

3.4.1 Utviklingstrekk knyttet til elevvurdering siste århundre

En statistisk undersøkelse av karaktergivningen i *Den høyere skolen* førte i 1920- og 1930-årene førte til bekymring for inflasjon i karaktergivningen. Tendensen til å gi elevene bedre karakterer enn de «fortjente» var et stort problem i de fleste land i den vestlige verden. Inspirert av Gauss-kurven, som illustrerer hvordan en rekke målte egenskaper statistisk fordeler seg symmetrisk om gjennomsnittsverdien, innførte man i Normalplanen av 1939 en veiledende vurderingsnorm langs en fempunktsskala (Lysne, 1999, s. 33, 177):

Særs godt	Meget godt	Godt	Nokså godt	Lite godt
S: 4 %	M: 24 %	G: 44 %	Ng: 24 %	Lg: 4 %

I følge Lysne hadde det såkalte «relative graderingssystemet» aldri blitt kritisk diskutert, ikke engang det faktum at en bestemt prosentandel ble dømt til å mislykkes (Lysne, 2006, s. 343). Det ble presisert at veiledende normalfordeling bare skulle ha gyldighet for store grupper på tusen individer eller flere, og at det i de fleste klasser på cirka 30 elever så å si alltid ville være avvik (Lysne, 1999, s. 133). Det kan her innskytes at ingen lærere har flere tusen elever, og at forutsetningene derfor er dårlige for å plassere elevene riktig i forhold til en større, og aller helst nasjonal, gruppe. Følgelig medfører slike retningslinjer stor sannsynlighet for at lærere følger normalfordelingen mer enn det er grunnlag for.

I 1960-årene økte frustrasjonen og misnøyen med den veiledende fordelingsnormen. Den ytterste konsekvensen av et slikt karaktersystem – at den gjorde klassekamerater til konkurrenter – ble brukt som et av argumentene mot karakterer i den svært opprivende karakterstriden i 1960- og 1970-årene (Lysne, 1999, s. 37). Dette var tiden da nye konsepter for gradering, inspirert av amerikanske teorier, kolliderte med kravene fra den politiske venstresiden i Skandinavia om avskaffing av karakterene (Egelund, 2005, s. 208). Begrunnelser for å gi elever karakterer går typisk langs tre linjer: *motivasjon*, *seleksjon* og *informasjon* (Wikström, 2006, s. 117). I den borgerlige skolen hadde karakterer og eksamen en fremtredende funksjon, ikke bare som seleksjonsverktøy, men også som et motivasjonsverktøy. Motivasjonen var ikke bare forankret i den enkelte elevs målsetninger, men i en konkurransesituasjon (Tønnessen og Telhaug, 1996, s. 22).

I Norge ledet Arbeiderpartiet an i kampen mot bruken av karakterer som motivasjonsfaktor, som man mente kunne føre til at elever som gjennomgående fikk dårlige karakterer utviklet et negativt selvbilde. Videre ble det hevdet at karakterer har en negativ

innvirkning også på gjennomsnittlig og høyt presterende elever ved å fokusere på ytre faktorer (ytre motivasjon) framfor elevenes glede og interesse for læring (indre motivasjon). Det borgerlige systemet med konkurranse som motivasjonsfaktor ble etter hvert avvirket og barnetrinnet så godt som rensket for karakterer og eksamener (Telhaug, 1997, s. 29).

I 1970-årene ble to offentlige utredninger (EVA1 og EVA2) premissleverandører i debatten om ytterligere reduksjon i karakterbruken. Flertallet i utvalget gikk inn for å avskaffe karaktervurdering i grunnskolen (NOU, 1974: 41). Flere av utvalgets medlemmer gikk også inn for å avskaffe karaktervurdering i videregående skole. Forslagene ga startskuddet til den mest kompromissløse skoledebatt vi har hatt i dette landet noen gang (Lysne, 2004, s. 113). Etter hvert som det folkelige engasjementet i debatten tiltok, ble imidlertid karaktermotstanderne marginalisert, og forslaget strandet blant annet i mangel på en troverdig løsning på seleksjonsproblemet. Arbeiderpartiet opplevde en enorm motstand, og måtte erkjenne at de ikke hadde støtte i folket. Retretten kom likevel for sent. Arbeiderpartiet mistet regjeringsmakten etter Stortingsvalget i 1981, og mange mente striden om karakterer i skolen ble deres bane (Lysne, 2004, s. 113). Telhaug har slått fast at karaktergivningen i grunnskolen i lang tid etter var «et tabubelagt tema i norsk skolepolitikk» (Telhaug, 1997, s. 61).

De amerikanske teoriene om mål- og kriteriebasert vurdering representerte et alternativ til normalfordeling av karakterene. Fra før hadde to forskere ved University of Chicago skapt det teoretiske fundamentet for et paradigmeskifte. Ralph W. Tylers klassiker *Basic principles for learning and instruction* (1949) dannet grunnlaget for utvikling av teori om målstyring i skolen, mens Benjamin S. Blooms *Taxonomy of Educational Objectives* (1956) ga anvisninger for spesifisering og kategorisering av elevers måloppnåelse. En artikkel på bare tre sider kalt *Instructional technology and the measurement of learning outcomes: Some questions*, skrevet av Robert Glaser (1963), ble starten på utviklingen av konsepter for å relatere vurderinger til absolutte i stedet for relative (normalfordelte) standarder. Dette prinsippet innebærer at brede målbeskrivelser må utledes i mer eksakte vurderingskriterier som angir grad av måloppnåelse, en framgangsmåte som etter hvert skulle komme til å dominere i vurderingssystemer verden over.

For karaktermotstanderne i Norge var det imidlertid ikke bedre om man skulle sammenlignes med en såkalt «objektiv standard» enn med andre elever. De fryktet en ensidig vektlegging av kunnskap som kunne måles og kontrolleres (Lysne, 1999, s. 37, 39). En vei ut av normalfordelingsregimet var det likevel helt nødvendig å finne. Derfor ble prinsippet om «tillempet målrelatering», som alt hadde blitt anvendt i videregående skole siden 1968, tatt i bruk i grunnskolen fra 1981. Prinsippet gikk ut på at det ikke skulle stilles absolutte krav om

oppfylging av bestemte kunnskapsmål eller dokumentasjon av bestemte ferdigheter. Det ble lagt vekt på valg av lærestoff, læremidler og arbeidsmåter med sikte på at læreprosessen hos den enkelte elev i større grad skulle innrettes mot elevenes personlige vekst mot en bredde av mål for opplæringen. Dette innebar i praksis en skjønsmessig helhetsvurdering (Lysne, 2004, s. 120).

I reformene i 1990-årene ble dette prinsippet formalisert ved at man forsøkte å forankre den generelle delen av læreplanverket i fagplanene gjennom det man kalte for «helhetlig kompetanse». I Prinsipper og retningslinjer for L97 ble det gitt anvisninger om organisering og arbeidsmåter, slik at den generelle læreplanen og fagplanenes innhold ble knyttet sammen. I Reform 94 ble sammenhengen dekket i «felles mål for faget», der læreplanens generelle del ble forsøkt integrert i fagene. I Stortingsmelding nr. 30 (2003–2004) erkjente departementet at dette hadde vært vanskelig, og slo fast at «Vurdering av elevenes «helhetlige kompetanse» har vært en medvirkende årsak til at bestemmelser om individvurdering og læreplaner kan framstå som uklare, særlig i videregående opplæring» (Stortingsmelding nr. 30 (2003-2004), s. 39).

Ved innføringen av Kunnskapsløftet ble det tatt et radikalt oppgjør med ideen om «tilpasset målrelatering» og «helhetlig kompetanse». Kunnskapsløftet representerer, i hvert fall om man legger fagplanene til grunn, et fullstendig gjennomslag for målrelateringsprinsippet som var blitt avvist i 1960- og 1970-årene, og en nølende tilslutning til en kriteriebasert tilnærming til karaktersetting. «Fluktigheten» og uklarheten i retningslinjer for elevvurdering gjør det imidlertid vanskelig, for ikke å si umulig, å forklare hva slags teorier for karaktersetting som har vært dominerende i Norge.

Den offentlige debatten om elevvurdering og karaktersetting har i Norge, også etter 1970-årenes karakterstrid, vært preget av en ideologisk steinkasting som har bidratt til å at man har hatt en politisk og filosofisk, heller enn praktisk, tilnærming til temaet. Sist debatten «for eller imot karakterer» hadde et offentlig ordskifte av et visst omfang, var under innføringen av L97. Den gangen tok Helga Hjetland, daværende leder av Norsk Lærerlag, igjen til orde for å avskaffe karakterene på ungdomstrinnet. Hun mente man burde utnytte muligheten til å legge grunnlaget for «eit permanent system med vurdering utan karakterar i heile grunnskulen frå skuleåret 2000-2001» (Hjetland, 1996, s. 185).

Nestlederen i Norsk Lærerlag på denne tiden, Ragnhild Midtbø, hadde i Dagbladet to år tidligere en kronikk med overskriften «Vekk med karakterene». Der begrunnet hun karaktermotstanden med at «uansett hvor mye opplæring lærerne får, og hvor mye tid det avsettes til samråding lærerne imellom, vil karakterene alltid være beheftet med den

vilkårighet som følger et system som i stor grad må bygge på skjønn» (gjengitt i Skagen, 1996, s. 48). Som i 1970-årenes karakterstrid manglet imidlertid lærerorganisasjonen et overbevisende alternativ til karaktersystemet, og politikken kunne derfor vanskelig få gjennomslag (Skagen, 1996, s. 48). Siden sammenslåingen med Lærerforbundet, til Utdanningsforbundet i 2001, har motstanden mot karakterer til lærernes største interesseorganisasjon vært begrenset til barnetrinnet. Utdanningsforbundet har siden vært svært skeptiske til den pågående utprøvingen av felles kjennetegn på måloppnåelse, som også omfatter modeller for å definere kjennetegn på høy og lav måloppnåelse på barnetrinnet. Leder for Utdanningsforbundets grunnskoleseksjon, Haldis Holst, begrunnet skepsisen slik i et intervju med Aftenposten:

Vi er skeptiske til vurderingskriteriene. I første rekke fordi de kan oppfattes som en konkurrerende læreplan og innsnevre den profesjonelle lærers handlingsrom. Men vi ser også at kriteriene kan brukes på en måte som nærmer seg karakterer. Alle symboluttrykk for prestasjoner er vi motstandere av (Holst, 5. mai 2006).

Utdanningsforbundet sluttet likevel nølende opp om utprøvingen. På bakgrunn av ekstern evaluering og Utdanningsdirektoratets anbefalinger i juni 2009, skal regjeringen ta stilling til hvordan prosjekt *Bedre vurderingspraksis* skal følges opp. Om en ny karakterstrid er i emning gjenstår å se.

Dette korte historiske tilbakeblikket kan bidra til å forklare de mange problemene vi har i norsk skole, både knyttet til summativ og formativ vurdering. Før vi ser nærmere på dette, skal vi slå fast at den ideologiske tilnærmingen til elevvurdering beskrevet ovenfor har relevans utover selve vurderingsfeltet. Erling Lars Dale (2008) peker på at empirisk forskning på elevenes læringsutbytte ikke ble en viktig del av den profesjonelle praksisen i det norske vurderingssystemet fra begynnelsen av 1970-årene og i en generasjon fremover. Dette kan sees i sammenheng med motstanden mot den såkalte *Tylerrasjonalen* om målstyring (Tyler, 1949). Dale slår fast at avvisningen av empirisk forskning på elevers læringsutbytte skjedde stikk i strid med intensjonene til den såkalte reformpedagogikken fra 1930-årene. «Den førende utdanningspolitiske og pedagogiske ideologien som dominerte fra begynnelsen av 1970-tallet, medførte derfor en oppløsning og ikke fornyelse av det reformpedagogiske skoleregimet i fra 1930-tallet» (Dale, 2008, s. 248). Det er grunnlag for å hevde at den politiske dragkampen elevvurdering som fagfelt har blitt offer for i Norge, har stått i veien for

utvikling av gode praktiske og pedagogiske tilnærminger både til karaktersetting og til elevvurdering generelt.

3.4.2 Summativ vurdering i Norge

Som påpekt ovenfor (jf. kapittel 3.2.2) er det svært krevende å gjennomføre valide og reliable vurderinger av elevers læringsresultater. I Norge har de mest omfattende reliabilitetsstudiene blitt gjennomført i norskfaget. Kjell Lars Berge (1996) undersøkte sensorreliabiliteten i vurderingen av norskeksamen i 1992 og 1993. Undersøkelsen viste en korrelasjon⁷ mellom sensorene på svake 0,56, som i følge Berge er på linje med hva som er oppnådd i tilsvarende studier i andre land (Berge, 1996, s. 181). Senere har Berge med flere gjort lignende reliabilitetsstudier av avgangsprøva (eksamen) i norsk på 10. trinn. Sensorreliabiliteten for vurderingene av norskeksamenene fra 1998 til 2001 viste en gjennomsnittlig korrelasjon på 0,69. Dette innebærer at bare 50 prosent av vurderingene var forutsigbare fra en sensor til en annen. Forskerne påpekte at påliteligheten dermed ikke var høy nok til å tilfredsstille internasjonale teststandarder, men at den likevel var høy sammenlignet med lignende skriveprøver internasjonalt (Berge, Evensen, Hertzberg og Vagle, 2005, s. 104; Vagle m.fl., 2007, s. 76).

Standpunktvurderingen

Det skulle gå fram av denne diskusjonen at det generelt er gode grunner for å stille spørsmål ved reliabiliteten ved sensorers vurdering av eksamener. Hvordan står det da til med påliteligheten i lærernes vurderinger av oppgaver de selv har utarbeidet? Det vet vi dessverre lite om, av flere grunner. Om vi ser bort fra muligheten for å ha to bedømmere til stede i klasserommet (slik man gjør ved muntlige eksamener), vil det være enklest å gjennomføre *inter-rater*-reliabilitetsvurderinger med utgangspunkt i skriftlige arbeider. Preasjonsbaserte vurderinger (presentasjoner, rollespill, fysiske aktiviteter osv.) vil på en eller annen måte gjøres tilgjengelig for andre (for eksempel ved hjelp av bilde, lyd eller video). Det er grunn til å tro at skriftlige prøver i de fleste fag utgjør en betydelig andel av vurderingsgrunnlaget, hvilket igjør slike reliabilitetskontroller enklere. Men med mindre det er etablert systemer og praksisformer for å ta vare på denne dokumentasjonen, så er vurderingsgrunnlaget læreren benytter ikke tilgjengelig for andre. I Norge har det ikke vært stilt slike krav til lærerne, og dermed er en av forutsetningene for å måle påliteligheten i lærernes karaktersetting

⁷ Korrelasjonene som beskrives i dette avsnittet er uttrykt i Pearsons korrelasjonmodell.

fraværende. Dette er trolig en av årsakene til at ikke eksisterer noen kjente undersøkelser av reliabilitet i standpunkt vurderingen i Norge.

Det er imidlertid gjennomført flere statistiske sammenligninger av resultater på standpunkt- og eksamens karakterer. Et problem med slike sammenligninger er at eksamensresultatene er en tvilsom referanse, ettersom de måler prestasjoner i løpet av en dag, mens standpunkt karakterene er resultat av en samlet vurdering av en rekke besvarelser over et helt skoleår. Svakheter knyttet til påliteligheten ved standardiserte prøver, gir da også grunn til å stille spørsmål ved gyldigheten av dette som referansepunkt. I en vurdering av store grupper kan det likevel, i mangel på andre referansepunkter, være en indikasjon på systematiske forskjeller. En slik statistisk sammenligning, gjennomført av Leif Kristensen på oppdrag fra Eksamenssekretariatet i 1999, viste at elever på den «snilleste» skolen samlet sett fikk hele 6,3 karakterpoeng mer på vitnemålet sammenlignet med elever ved den «strengeste» skolen (Kristensen, 1999; Tveit, 2007a, s. 210).

Statistisk Sentralbyrå har kartlagt sammenhengen mellom eksamens- og standpunkt karakterer for 2004 og fant også signifikante forskjeller. Forskerne mente forskjellene trolig var «drevet av et nokså lite antall skoler som synes å ha en 'systematisk' avvikende karakter praksis» (Hægeland, Kirkebøen og Raaum, 2005, s. 52). Den offentlige utredningen *Nasjonalt vurderingssystem for grunnskolen*, også kjent som *Moe-utvalget*, slo fast at lærerne vurderer elevenes prestasjoner med «forskjellige typer målestokk», og at det derfor heftet usikkerhet ved påliteligheten i karakter- og eksamensstatistikkene (NOU, 1997, s. 47). Læringscenteret (nå Utdanningsdirektoratet) konkluderte i 2003 med at «undersøkelser viser at det i norsk ungdomsskole og videregående skole har utviklet seg ulike kulturer for elevvurdering, mellom skoler og mellom lærere» (Læringscenteret, 2003, s. 35).

Dale og Wærness har pekt på at den norske tradisjonen for karaktervurdering er «målrelatert i den forstand at målene for vurderingen er innsikt og forståelse. Men vurderingen er også grupperelatert, siden elevene blir sammenlignet med hverandre i forhold til hvem som er over og under middels» (Dale og Wærness, 2007, s. 105). Evalueringen av lærernes vurderinger av de nasjonale prøvene i engelsk skriving i 2005 viste at sensorene hadde store problemer å avgjøre nivået i forhold til en gitt vurderingsskala (Lie, 2007, s. 89). Dette var til tross for at de benyttet seg av prøver med fastsatte vurderingskriterier. Man kan spørre seg hva man relaterer vurderingene til i standpunkt vurderingen, der man ikke har slike felles retningslinjer. Uten absolutte standardbeskrivelser er det særdeles vanskelig, for ikke å si umulig, å relatere vurderingen til faglige nivåer uavhengig av gruppens kvaliteter. Således kan det reises spørsmål om normalfordeling fortsatt gjør seg gjeldende i norske klasserom.

De få kjente pålitelighetsstudiene vi har i Norge gir altså grunn for bekymring både for eksamen- og standpunktvurdering. Man kan reise spørsmål om forutsetningene er tilstede for at norske lærere kan gjøre pålitelige karaktervurderinger. Moe-utvalget pekte i 1997 på at «elevene møter i dag ofte eksamen eller prøver uten å ha tilstrekkelig klarhet i hvilke kriterier de skal vurderes etter» (NOU, 1997, s. 47). Som vist i gjennomgangen ovenfor, er en viktig forutsetning for pålitelige karaktervurderinger at grunnlaget for karaktersettingen er klargjort for både lærere og elever. Det kan synes som at mangel på dette er en vesentlig forklaring på problemene knyttet til pålitelighet i karaktersettingen i Norge.

Uklarhet i vurderingsgrunnlaget

En gjenganger i norske klasserom er kravet til «muntlig aktivitet i timen». Uttrykket har trolig slektskap til begrepene «muntlige fag» og «skriftlige fag». Det er imidlertid ikke noe som heter muntlige og skriftlige fag. I norskfaget har vi tre standpunkt karakterer: muntlig, hovedmål skriftlig og sidemål skriftlig. I engelskfaget har vi en skriftlig og en muntlig standpunkt karakter. Det er bare i morsmålsfagene og engelsk vi har et slikt skille⁸, mens det i andre fag gis én standpunkt karakter – uten noen spesifisering av muntlig eller skriftlig. Dette henger sammen med at muntlige og skriftlige ferdigheter er et mål i seg selv i språkfagene, mens de primært utgjør grunnleggende ferdigheter og vurderingsformer i alle de andre fagene (Tveit, 2007a, s. 209-210). Årsaken til oppfatningene om muntlige og skriftlige *fag* har trolig sin opprinnelse i eksamensformen for de forskjellige fagene. Fag som ikke er språkfag og heller ikke har sentralt gitte skriftlige eksamener, har blitt hetende «muntlige fag», trolig fordi eksamen som hovedregel har vært muntlig. Og selv om det ikke har forankring i noen styringsdokumenter, er det tilsynelatende en vanlig oppfatning at det i disse fagene i større grad skal legges vekt på «muntlig aktivitet i timen» (Tveit, 2007a, s. 209-210).

Sterk vektlegging av muntlig aktivitet kan imidlertid lede til at elevene oppfatter deltakelse i seg selv som det viktigste, og ikke *hva* man bidrar med. Noen elever mestrer dette spillet bedre enn andre, og kan slike skåre poeng hos læreren ved simpelthen å være aktive. Innenfor en gruppe vil det alltid være noen som ikke ønsker å markere seg så mye som andre, men som ville kunne ha dokumentert like høy eller høyere måloppnåelse på andre måter (Tveit, 2007a, s. 209-210). En slik praksis med stor vektlegging av elevenes muntlige aktivitet kan sees på som urettferdig overfor disse elevene.

⁸ *Samisk* og *Finsk* har også to standpunkt karakterer, mens det for *Fremmedspråk 2* kun er én standpunkt karakter.

Muntlig aktivitet i timen oppfattes ofte som et uttrykk for at man viser stor *innsats*, en tilbakevendende problemstilling ved karaktervurdering. Forskrift til opplæringsloven slår fast at det kun er elevenes oppnåelse av målene, altså ikke elevenes innsats, som skal inngå i karaktervurderingen (Forskrift til opplæringsloven § 3-7 og § 4-7). Karakterer for «flid» ble avvirket så tidlig som i 1920-årene, men har blitt forsøkt gjeninnført en rekke ganger (Lysne, 2004, s. 200). Nyere undersøkelser viser imidlertid at innsats i stor grad oppfattes som et legitimt kriterium. I en undersøkelse av *Sammenhengen mellom eksamen og læring* (SEL) som Erling Lars Dale og Jarl Inge Wærness gjennomførte på oppdrag fra Utdanningsdirektoratet, rapporterte lærere og elever både i ungdomsskoler og videregående skoler at det ble lagt stor vekt på elevenes innsats når lærerne ga standpunkt-karakterer (Dale og Wærness, 2006, s. 192-193).

Ønsket om å premiere innsats er forståelig. Det virker både logisk og rettfærdig at elevene «får noe igjen» for å legge ned ekstra innsats. Det er imidlertid viktig å skjelle mellom innsats som gir dokumentert høyere måloppnåelse og innsats som en selvstendig faktor. Å «premiere» innsats uten at eleven har vist at det ekstra arbeidet har ledet til høyere måloppnåelse, er ikke forenlig med regelverket vi har hatt for elevvurdering i Norge. I den nye forskriften til opplæringsloven er det presisert tydeligere enn tidligere at det kun er oppnåelse av målene som skal vurderes (Forskrift til opplæringsloven § 3-7 og § 4-7). Bakgrunnen for dette er at «faglige prestasjoner på den ene siden, og atferd, orden og innsats på den andre – skal holdes fra hverandre, slik at det blir klarere for elevene og lærlingene hvilke kriterier som legges til grunn ved vurderingen» (*Kultur for læring*, 2003-2004: 30, s. 39).

Dette bringer oss over til regelverket for orden- og oppførselskarakterene, som det også har hersket stor forvirring rundt. I Dale og Wærness sin undersøkelse ga en lærer i videregående skole uttrykk for at «eleven skal vite at det ikke holder å sitte i kantina og gjøre det bra på prøvene. Alle kan gjøre en og annen god enkeltprestasjon. Men de må møte opp og gjøre jobben sin» (Dale og Wærness, 2006, s. 195). Denne uttalelsen er altså fra en lærer i videregående skole, der forvirringen knyttet til orden- og oppførselskarakterene har vært særlig stor, men problemstillingen er også relevant for ungdomstrinnet. Som nevnt tidligere har internasjonale undersøker vist hvordan fagkarakterer ofte benyttes som disiplineringstøyt, til å kontrollere elevoppførsel. Metaanalysene til Hattie og Timperley (2007), Deci, Koestner og Ryan (1999) og Wynne Harlen (2004) viste tydelig at dette kan føre til at elevene ikke verdsetter læring i fagene (jf. kapittel 3.3.1).

Av den nye forskriften til opplæringsloven går det klart fram at det ikke er lov å straffe elever for lavt oppmøte med å sette ned karakterer i fagene. Tidligere har orden og oppførsel vært slått sammen i én karakter i videregående skole. I den nye forskriften er regelverket samordnet slik at det er en ordenskarakter og en oppførselskarakter både i ungdomsskolen og i videregående skole (Forskrift til opplæringsloven § 3-9 og § 4-9). Av forskriften framgår det at vurdering av om elevene viser «vanlig god arbeidsinnsats» og respekt for skolens ordensreglement, hører inn under ordenskarakteren. Karakter i oppførsel skal gis på grunnlag av elevenes oppførsel i klasserommet, i skolegården og på skoleveien (Forskrift til opplæringsloven § 3-9 og § 4-9).

Trynefaktor kan i skolesammenheng defineres som *en faktor ved vurdering av elevers oppnåelse av målene for opplæringa som ikke er forankret i relevant dokumentasjon* (Tveit, 2007b, s. 285). Trynefaktor er et mye diskutert tema blant mange elever. Det påstås ofte at læreren er for snill med enkelte elever og gir et «trynetillegg», eller at noen blir uforholdsmessig strengt bedømt. Disse synspunktene hviler på den naturlige oppfatningen om at trynefaktor ikke er en legitim faktor i karaktervurderingen. Overraskende er det derfor at enkelte lærere mener det er legitimt å gi trynetillegg når de gir elevene karakterer. I Dales og Wærness' undersøkelse sa en lærer at «Positiv holdning i klasserommet er viktig. Trynefaktoren skal ikke undervurderes. Den teller også i arbeidslivet» (Dale og Wærness, 2006, s. 194). Trynefaktor er et omstridt tema som gjør seg gjeldende i de fleste samfunnsområder.⁹

Flere elever i ungdomsskolen og videregående skole rapporterte i Dale og Wærness undersøkelse om at trynefaktor hadde stor betydning når lærerne ga karakterer (Dale og Wærness, 2006, s. 195). Det er viktig å påpeke at trynetillegg ikke behøver å være tilsiktet eller bevisst for å falle inn under definisjonen ovenfor. Det er fort gjort å la seg påvirke av egne *forventninger* til elevenes nivå. Dette skapes ofte av elevenes tidligere resultater, på engelsk kjent som *halo-effekten* (glorie-effekten) (Brady og Kennedy, 2005, s. 70). Dersom man ikke benytter eksplisitte vurderingskriterier ligger det særlig godt til rette for at ubevisste faktorer spiller inn i vurderingen.

⁹ I Danmark skapte en dansk lektor overskrifter ved å oppfordre studenter til å spille på utseendet ved eksamen. Uttalelsen «Hvis man har en god røv, er der ingen grund til at skjule den til eksamen», fikk de fleste til å se rødt, men lektoren fikk støtte fra den danske arbeidsgiverforeningen *Dansk Erhverv*, som mente lektoren var «havnet i unåde for at si det højt, som alle udmærket ved», og fastslo at «køn og udseende betyder noget – både ved eksamensbordet og ved ansættelsessamtaler» (Berlingske Tidende 16.05.2007).

Uklarhetene som standpunktvurderingen har vært heftet ved knyttet til vurdering av muntlig aktivitet, innsats, orden, oppførsel, oppmøte og trynefaktor representerer til sammen en stor trussel mot pålitelighet i lærernes karaktervurderinger. Det er imidlertid ikke bare i karaktersettingen dette har vært et problem. Uklarhet omkring retningslinjer står også sentralt i den videre drøftningen av formativ vurdering i Norge.

3.4.3 Formativ vurdering i Norge

Kirsti Klettes (2003) klasseromsforskning av elever og lærere i barne- og ungdomsskoler, som utgjorde en del av evalueringen av L97, avdekket viktige utfordringer knyttet til formativ vurdering i norsk grunnskole. I studien ble det observert at lærerne svært ofte kommenterte elevenes arbeid som «fint» og «flott», uavhengig av elevenes arbeidsinnsats og kvaliteten på produktet de framviste. «Gjennomgående gir lærerne få eksplisitte produktkrav og standarder for hva som kjennetegner gode arbeidsbesvarelser og framføringer» (Klette, 2003, s. 53).

Haugstveit, Sjølie og Øygarden (2006) undersøkte hvordan lærere på mellomtrinnet i sin undervisning utviklet vurderingskompetanse og på hvilke måter dette kom til uttrykk i deres praksis (Haugstveit, Sjølie og Øygarden, 2006, s. 3). De kom fram til at det er viktig at elevene og lærerne samarbeider om å utvikle kriterier for vurdering av nivået på elevarbeidene i forhold til læreplanmål. «Når elevene skal velge ut for eksempel en fortelling, hva velger de ut fra da? Når læreren skal vurdere fortellingen, hva vurderer hun da?» (Haugstveit m.fl., 2006, s. 77). Forskerne hevdet at vurderingskriterier som elevene har et eierforhold til, vil bidra til at elevene får større tillit til lærernes vurdering. De mente dette også vil bidra til å styrke elevens forståelse og utbytte av lærerens kommentarer, og foreslo å bruke vurderingsskjemaer til å systematisere kriteriene for vurderingen (Haugstveit m.fl., 2006, s. 99).

Forskningsprosjektet PISA+ (*Prosjekt om Lærings- og Undervisnings-Strategier i Skolen*), som er en kvalitativ oppfølgingsstudie av PISA-prøvene, har vist at norske lærere er svakt skolert og mangler verktøy for å legge til rette for gode læringssituasjoner for utvikling av elevenes leseforståelse og lesestrategier (Klette m.fl., 2008). Klette (2008) peker på at forskere i perioder har vært aktivt *imot* å utvikle lærernes metodikkompetanse. Hun hevder «metodikk, oppskrifter, undervisningsteknologi og remedier har i perioder vært sett på som antitesen til reflektert profesjonskunnskap» (Klette, 2008). Klette mener verktøy og hjelpemidler i Norge har blitt sett på som uttrykk for «det instrumentalistiske mistaket» (Skjervheim, 1976), og at vi derfor mangler konkrete og fleksible instrumenter til bruk i forbindelse med opplæring. Hun mener det er mot denne bakgrunnen – mangelen på

instrumenter – man kan forstå den omfattende bruken av for eksempel arbeidsplaner i norske klasserom (Steen, 2007). Til vurderingsarbeidet er det imidlertid få kjente instrumenter og hjelpemidler lærere kan benytte seg av. Klette etterlyser flere verktøy lærere kan benytte i opplæringen. Det er først helt nylig vi har sett et systematisk arbeid med vurderingsskjemaer og kriterier, initiert fra myndighetenes side, i forbindelse med Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis* (Utdanningsdirektoratet, 2007c).

Dale og Wærness (2007) har pekt på problemer det medfører at kriteriene for middels og under middels måloppnåelse har vært beskrevet gjennom uttrykk for mangel på eller fravær av oppnåelse av kriteriene for det øverste nivået. For at en tekst skulle kommunisere på høyeste nivå, måtte den ha «et relevant innhold, en klar struktur og et funksjonelt språk» (Dale og Wærness, 2007, s. 105). En tekst på middels nivå ville ha «mindre fullstendige svar» og «noe uklar og upresis» formidling av innholdet (Dale og Wærness, 2007, s. 105). Dale og Wærness pekte på at beskrivelser av mestring var forbeholdt elever som fikk karakterene 5 og 6. De viste også hvordan elever på det laveste nivået kunne bli vurdert: «Din tekst har ingen klar struktur, den har en springende logikk, budskapet er lite relevant, innholdet er magert, og språket ditt er upresist og preges av mange feil» (Dale og Wærness, 2007, s. 106). Forfatterne slo fast at slike formuleringer ikke fungerer som kompetansemål for mestring.

Med virkning fra 1. august 2007 er det i Forskrift til Opplæringsloven tatt høyde for denne kritikken, og fastsatt nye og felles karakterbeskrivelser (standarder) for ungdomsskole og videregående skole. Karakterene uttrykkes med forskjellige kompetansebeskrivelser for hvert nivå, fra «svært lav kompetanse i faget» (karakteren 1), til «fremragende kompetanse i faget» (karakteren 6) (Forskrift til opplæringsloven § 3-8 og § 4-8).

Haugstveit m.fl. (2006) har som nevnt understreket betydningen av at elevene er aktive deltakere i utviklingen av vurderingskriterier. Deres undersøkelse viser at dette kan virke motiverende på det videre arbeidet, ved at elevene får en klarere forståelse for hva de skal lære og hva de blir vurdert ut fra (Haugstveit m.fl., 2006, s. 63). De fant også at lærerne savnet «grundige fellesdrøftinger med kolleger om målrelatert vurderingspraksis, for eksempel om kriterier og om grunnlag for vurdering» (Haugstveit m.fl., 2006, s. 227). Forskerne peker på at dette trolig henger sammen med at målrelatert vurdering ikke har noen sterk tradisjon i barneskolen, der slik vurdering i prinsippet er forbudt. Særlig lærere på 6. og 7. trinn ga uttrykk for at dette var problematisk. En lærer sa blant annet at «vi [barneskolen og ungdomsskolen] er like ved hverandre, men det er et hav mellom oss» (Haugstveit m.fl., 2006, s. 227). Forskerne konkluderer med at lærerne på de siste trinnene i barneskolen har en

stor utfordring i å utvikle en vurderingspraksis i spenningsfeltet mellom individrelatert og målrelatert vurdering (Haugstveit m.fl., 2006, s. 227).

Det kan reises spørsmål om fraværet av karakter på barnetrinnet bidrar til mindre fokus på faglig måloppnåelse på disse trinnene, hvilket igjen fører til lavere læringsutbytte enn disse elevene ellers ville hatt (jf. betydningen av faglige tilbakemeldinger, kapittel 3.3.1). Ettersom det er forbudt å bruke kriterier for å vurdere karakternivå, får elevene trolig i mindre grad tilbakemeldinger som kan fortelle dem hva som må til for å bli flinkere. Således kan man reise spørsmål om den formative vurderingen faktisk lider under fraværet av et summativt karakterregime.

I forbindelse med introduksjonen av det nye læreplanverket for Kunnskapsløftet høsten 2006, ble forskriften til opplæringsloven revidert. Den gamle forskriftens omtale av «formell» og «uformell» vurdering ble erstattet av et mer konsistent skille mellom «underveisvurdering» og «sluttvurdering». En av årsakene til denne endringen er at det i Norge har hersket stor forvirring om hva som er *formålet* med elevvurdering. Begrepene underveis- og sluttvurdering angir tydeligere elevvurderingens formål. Underveisvurderingen kjennetegnes av at elevene er *på vei* mot et (slutt)mål. Det bør derfor være rom for prøving og feiling (Tveit, 2007e, s. 198-199). Gjennom underveisvurderingen skal elevene få tilbakemeldinger som hjelper dem til best mulig måloppnåelse. Den kjennetegnes således primært av det jeg har beskrevet som *formativ* vurdering. Formålet med sluttvurdering er å gi relevant og pålitelig informasjon til elevene selv, foreldre, høyere utdanning og arbeidsliv om nivået elevene til slutt har oppnådd. Det vi har beskrevet som *summativ* vurdering.

Dette medfører at «likeverdig og rettferdig elevvurdering» får to vidt forskjellige betydninger avhengig av hvilken av de to formene det er snakk om. Likeverdig og rettferdig sluttvurdering kjennetegnes ved at elevenes oppnåelse av kompetansemålene måles mot karakterkrav som er felles for alle elevene. Likeverdig og rettferdig underveisvurdering kjennetegnes derimot av at tilbakemeldingene tilpasses den enkelte elevs forskjellige evner og forutsetninger. Mens det i sluttvurderingen skal være størst mulig likebehandling, er det et mål i seg selv å forskjellsbehandle i underveisvurderingen. De to formålene er følgelig ikke alltid forenlig i samme vurderingstiltak. Georg Matthiesen spør retorisk: «Hvilken elev vil legge seg åpen for læring og avdekke sine feil og mangler dersom en slik innrømmelse medvirker til dårlige sluttresultater?» (Matthiesen, 2007, s. 53). For å oppnå et læringsmiljø der elevene får mulighet til å prøve og feile, er det nødvendig med «risikofrie soner». Det er derfor viktig at lærere underveis opplyser elevene om *hva* som inngår i sluttvurderingen og *når* oppnåelse av målene vil bli vurdert.

3.5 Noen erfaringer med elevvurdering i Queensland

Det vil bli for omfattende å gjennomføre en grundig analyse av undersøkelser knyttet til elevvurdering i Queensland. Jeg har valgt å trekke fram noen undersøkelser, som kan fortelle oss noe om påliteligheten i lærernes vurderinger og kritikk som har blitt reist mot tilnærmingen til standardbasert vurdering. Det er gjennomført få studier av påliteligheten i læreres karaktervurdering i *Years 1-10* i Queensland. For *Year 11 & 12* er det imidlertid gjennomført flere pålitelighetsstudier i tilknytning til prosedyrene for ekstern moderering. Nedenfor gis en kort gjennomgang av disse studiene. Deretter belyses noen problematiske aspekter ved delstatens tilnærming til standardbasert vurdering.

3.5.1 Erfaringer med modereringsprosedyrene

Som beskrevet i kapittel 2, har man i Queensland et omfattende system for å sikre sammenlignbare standarder på tvers av skoler på *senior secondary* nivå. Men hvor sammenlignbare er standardene? På bakgrunn av bekymringer knyttet til dette blant elever, foreldre og lærere, tok TEPA i begynnelsen av 1990-årene initiativ til en omfattende pålitelighetsundersøkelse. Geofferey N. Masters og Bruce McBryde gjennomførte en kryssrettingsundersøkelse av 546 elevers mapper fra 1992 innenfor fagene *Chemistry*, *English*, *Mathematics* og *Modern History* på tvers av hele delstaten. Seks uavhengige vurderinger av elevmapper ble gjort av bedømmere som var kjent med modereringsprosedyrene. Alle vurderinger ble gjort på en 50-poengs skala der hvert *level of achievement* var delt inn i ti «rungs». De samlede resultatene viste en korrelasjon på eksepsjonelle 0.95. Mer detaljert var korrelasjonene slik:

Approximately 98 per cent of paired assessments differed by less than one Level of Achievement (10 rungs); 90 per cent differed by less than half a Level of Achievement (5 rungs). A similarly high level of agreement (correlation 0.92) was obtained between the assessments made in this study and the original exit Levels of Achievement allocated to these folios¹⁰ (Masters og McBryde, 1994, s. vi).

¹⁰ Det bør presiseres at undersøkelsen ikke hadde et klassisk «inter-rater reliability design», der hver vurdering gjøres uten kjennskap til andre bedømmers vurdering. «Assessments were made independently of each other, but each assessor had access to teachers' prior assessment of folios. Many folios consisted of collections of marked tests and assignments. Numerical scores were available to assessors, as were written comments made by teachers, although care was taken to remove any indication of teachers' allocated Levels of Achievement» (Masters og McBryde, 1994, s. vii).

Siden 1994 har utdanningsmyndighetene i Queensland gjennomført årlige stikkprøver, for å overvåke sammenlignbarheten i standarder på tvers av skoler og fag¹¹. Disse undersøkelsene bekrefter høy grad av sammenlignbarhet på tvers av skoler og distrikter. Rapporten fra 1994 viste at 79 prosent av elevmappene ble vurdert til å være riktig standard, et tall som siden har variert mellom 85 og 93 prosent (QSA, 2006, s. 11).

Det er nå 35 år siden introduksjonen av «school-based assessment» i Queensland. Siden implementeringen av standardbasert vurdering i midten av 1980-årene har flere forbedringstiltak blitt iverksatt, men prinsippene har vært de same. I følge Graham Maxwell er gjeldende prosedyrer «well entrenched, have substantial support and are managed effectively and efficiently» og offentlighetens tillit til systemet har blitt opprettholdt (Maxwell, 2007, s. 1).

3.5.2 Problematiske aspekter ved standardbasert vurdering

Som ved alle prinsipper for elevvurdering, kan det pekes på svakheter ved kriterie- og standardbasert vurdering. I en studie av kriteriebasert vurdering i *senior secondary English* i Queensland fra 1993, utfordret Jill Freiberg ideen om at kriteriebasert vurdering som en nøytral teknologi som sikrer at *high stakes* vurderinger alltid er valide og reliable. Hun hevdet at en slik forståelse var dominerende i teorier om vurderingspraksiser i Queensland på denne tiden. Freiberg hevdet praksis var basert på en ide om at kriteriene for vurdering av en definert standard var «universal, ‘fixed’ and value free and that a particular educated sub-culture can interpret and apply these criteria and standards in a value-neutral way» (Freiberg, 1993, s. 23).

Med utgangspunkt i kritisk diskursteori (*critical discourse theory*) påpekte hun at mens det er vanskelig nok å oppnå enighet om standarder mellom lærere, så er elevenes forutsetninger for å fortolke lærernes fortolkninger desto dårligere. Hun hevdet at «because these interpretations will be different for members of different cultures and sub-cultures, equity issues remain significant even when assessment is based on pre-specified criteria» (Freiberg, 1993, s. 24). Freiberg pekte på at beskrivelsene av de forskjellige standardene ikke var godt forankret fordi kriteriene de er knyttet til ikke er godt beskrevet og fordi man skjelner mellom standarder ved hjelp av sammenlignbare kvanitative begreper som «great» og «reasonable»; «broad» og «fine» (Freiberg, 1993, s. 128).

¹¹ Ved stikkprøvene samler det statlige panelet inn en tilfeldig gruppe av syv elevmapper og skolens vurdering av måloppnåelse og «rung placement» fra 1 til 10 innenfor hver standard, fra et tilfeldig utvalg skoler på tvers av delstaten. Det statlige panelet bestemmer hvilke elevs mapper som skal kontrolleres basert på tilfeldig utvalg, stratifisert etter medium- og terskelnivåer for standardene (QSA, 2006, s. 4-5).

Freibergs poeng er at elever vil ha ulike forutsetninger for å avkode lærernes fortolkninger. Følgelig argumenterte hun for at enighet mellom bedømmere ikke innebærer at man er i stand til å gjøre objektive vurderinger. Hun hevder at måten lærere fortolker vurderingskriterier og -standarder på, demonstrerer at bedømminger er basert på vurderinger av elevenes kulturelle kapital (Freiberg, 1993, s. 9, 131). Slik utfordrer hun blant annet Bordieu og Passersons teori om at eksplisitte kriterier kan motvirke sosial reproduksjon (jf. kapittel 3.3.2).

Man kan imidlertid reise spørsmål om noe vurderingsprinsipp er i stand til å løse dette problemet. Det er likefullt viktig å være oppmerksom på at heller ikke kriterie- eller standardbasert vurdering er objektiv, selv om man søker å oppnå størst mulig grad av pålitelighet ved å bruke eksplisitte kriterie- og standardbeskrivelser. Som nevnt tidligere bør fokus rettes mot i hvilken grad elevene er i stand til å forstå kriteriene, samtidig som man avveier hvor detaljert det er ønskelig at kriteriene er.

3.6 Fagspesifikke problemstillinger knyttet til elevvurdering

Så langt har jeg diskutert elevvurdering uavhengig av fag. Under vil jeg kort drøfte problemstillinger knyttet til elevvurdering i fagene studien omfatter, morsmåls- og matematikkfagene.

3.6.1 Elevvurdering i morsmålsfaget

I diskusjonen av morsmålsfaget legger jeg størst vekt på norskfaget, men jeg trekker også fram noen erfaringer fra faget *English* i Queensland.

Norskfaget

I norskfaget har *proessorientert skrivning* og *mappevurdering* lenge vært de mest kjente pedagogiske prinsippene som omhandler elevvurdering. Olga Dysthes innføringsbok i proessorientert skrivepedagogikk, *Ord på nye spor*, utgitt første gang i 1987, er en kjent bok for mange norsklærere. Proessorientert skrivning er en metode basert på et sosiokulturelt perspektiv på læring og legger vekt på systematisk veiledning i alle faser av skrivearbeidet (Dysthe, 1999, s. 18, 58). I bokens kapittel om vurdering peker Dysthe på at man i Norge, til forskjell fra for eksempel USA, har en sterk tradisjon for å gjøre helhetsvurdering (holistisk vurdering) av skriftlige arbeider. Hun advarer mot å følge etter USAs objektivistiske tilnærming, der man tester begrensede fenomener som språkbruk og grammatikk, blant annet ved hjelp av flervalgsoppgaver. Samtidig peker hun på at en svakhet ved helhetlig vurdering

er at den ofte blir intuitiv. Hun argumenterer for at man bør utvikle vurderingskriterier knyttet til ulike oppgavetyper og sjangere. I prosessorientert skriveopplæring kan utarbeiding og diskusjon om vurderingskriterier utgjøre et viktig element innledningsvis, som læreren kan vise tilbake til når han til slutt vurderer arbeidene (Dysthe, 1999, s. 220).

Dysthe tar til orde for å utvikle vurderingsskjemaer, delt inn i innhold, språkføring, kildebruk, rettskriving osv. Hun forklarer at man ved analytisk vurdering setter en poengsum for hvert punkt og at man får den endelige karakteren ved å summere punktpoengene. I holistisk vurdering noterer man stikkord og kommentarer for hvert punkt, men man tar hensyn til at helheten i en tekst er mer enn summen av delene. Dysthe mener analytisk vurdering er mindre problematisk i fag som matematikk, men at vi i Norge bør ta vare på tradisjonen med holistisk vurdering i de humanistiske fagene:

Delvurdering gir eit skeivt bilete av kva som totalt sett er ein god tekst, spesielt når oppgåvene er meir avanserte og krev ei rekkje ulike ferdigheiter. Ingen har hittil vore i stand til å setje opp ei fullstendig liste over enkeltelement som til saman gjer ein tekst god. Å vurdere berre nokre få trek, fremjar eit forenkla syn på kva god skrivning er (Dysthe, 1999, s. 222).

Dysthe bidro også til å introdusere mappevurdering i Norge, en vurderingsform som for språkfagene er nært knyttet til prosessorientert skriveopplæring. Hensikten er blant annet å la elevene bli deltakere i vurderingsprosessen ved å vurdere egne arbeider. Ved å være med å velge hvilke arbeider i «arbeidsmappen» som skal bli en del av «presentasjonsmappen», er elevene også med på å bestemme hvilke arbeider som skal utgjøre grunnlaget for sluttvurderingen i faget (Dysthe, 2007, s. 136). Dysthe har tatt til orde for at mappevurdering er en egnet måte å gjennomføre standpunktavurdering i norskfaget (Dysthe, 1999, s. 238).

Elevvurdering i norskfaget handler om mer enn skrivning. Frøydís Hertzberg og Astrid Roe (1999) peker i forordet på boken *Muntlig norsk* på at mens muntlige ferdigheter ble mye vektlagt i latinskolen, så var lesing og skrivning de viktigste norskfaglige ferdighetene gjennom hele 1900-tallet (Hertzberg og Roe, 1999, s. 7). I L97 fikk imidlertid muntlige ferdigheter en mer sentral plass, og med Kunnskapsløftet er muntlig en av de grunnleggende ferdighetene. Mange norsklærere har imidlertid dårlig samvittighet for ikke å legge nok vekt på muntlighet, som følge av at det er en svært tidkrevende aktivitet (Penne, 1999, s. 85).

Hertzberg (1999) framholder at mye av det norsklærere kan om elevvurdering av skriftlige arbeider, kan komme til nytte ved vurdering av muntlige prestasjoner. Til forskjell fra Dysthe, argumenterer Hertzberg for en analytisk tilnærming, og viser de pedagogiske

fordelene ved å synliggjøre og diskutere vurderingskriterier. Hun framholder også betydningen av å gjøre en så valid og reliabel vurdering som mulig også ved vurdering av muntlig, selv om dette er fjernt fra tradisjonelle testsituasjoner. Hertzberg advarer mot å nøye seg med å vurdere bare på grunnlag av såkalt «aktivitet i timen» (Hertzberg, 1999, s. 191), som tilsynelatende har vært utbredt i Norge (jf. kapittel 3.4.2). For å ivareta hensynet til reliabilitet anbefaler hun at lærere slår seg sammen for å vurdere elevenes framføringer, og på denne måten utvikler et tolkningsfellesskap (Hertzberg, 1999, s. 192). Som for vurdering av skriftlige arbeider, argumenterer Hertzberg for at elevene involveres i utviklingen av kriterier for vurdering av muntlig måloppnåelse.

English i Queensland

Claire Wyatt-Smith og Geraldine Castleton (2005) har gjennomført en interessant undersøkelse av vurderingspraksisene til *Year 5* engelsklærere i Queensland. De identifiserte blant annet hvordan lærerne etablerte latente «in-the-head» standarder, som Sadler har kalt det (Wyatt-Smith og Castleton, 2005, s. 139). I Queensland er det som nevnt ingen obligatoriske prosedyrer for å veilede eller kontrollere lærernes vurderingspraksiser for *Year 1-10*. Wyatt-Smith og Castleton hevder at lærerne baserte sine vurdering på førstehånds observasjoner av elevenes kunnskaper og skolens kulturelle kontekst, framfor læreplaner og andre offisielle retningslinjer (Wyatt-Smith og Castleton, 2005, s. 136). Ved en rekke tilfeller fant de at lærernes oppfatninger av elevenes måloppnåelser ble påvirket av lærernes gjenbruk av tidligere observasjoner av elevenes evner, motivasjon, personlighet og innstilling til skolearbeidet. Selv om disse observasjonene kun var latente, hadde de stor innvirkning på lærernes vurderinger (Wyatt-Smith og Castleton, 2005, s. 140).

Når lærerne ble spurt om å vurdere autentiske elevarbeider fra andre skoler, opplevde de det som vanskelig og ukomfortabelt å ikke ha tilgang til den sosiale konteksten elevenes arbeider var blitt gjort i. «Greater uncertainty was caused by the teachers' interest, even persistence, in trying to read the student in the writing» (Wyatt-Smith og Castleton, 2005, s. 146). I undersøkelsen beskriver en lærers frustrasjon over ikke å vite den pedagogiske konteksten for elevarbeidene, ved vurdering av elevenes måloppnåelse:

I don't know how much the teacher then expects either ... whether this has just been a short theme or thing, I'm not sure. Whereas we did it over a long time, over quite a few weeks, and our entries were ... longer, so again not knowing the context and how long they had to do this, so if this was only a day's exercises, I'd have to put it up a bit. But then, if it's supposed to be more than that, the exercise, I don't know (Wyatt-Smith og Castleton, 2005, s. 148).

Dette problemet er forsøkt ivaretatt i læreplanene for *senior secondary* i Queensland, ved at det stilles krav om at skolene dokumenterer den pedagogiske konteksten de forskjellige elevarbeidene er produsert i. Som undersøkelsen viste er det imidlertid ingen slike føringer eller kontroller for *Year 1-10*.

3.6.2 Elevvurdering i matematikkfaget

I diskusjonen av matematikkfaget legger jeg vekt på det norske faget, og prioriterer blant annet en diskusjon av noe av litteraturen som de norske lærerstudentene i undersøkelsen trakk fram. Tilsvarende har jeg svært kort pekt på et utviklingstrekk i knyttet til elevvurdering i *Mathematics* i Queensland som er relevante, uten å gjøre en grundig analyse av dette.

Matematikk i Norge

Brekke og Gjone (2001) peker i matematikkapitlet i boken *Fagdebatikk* på at vurdering ofte blir oppfattet som en lett øvelse i *Matematikk*, ettersom et regnestykk er «enten riktig eller galt» (Brekke og Gjone, 2001, s. 231). De peker på at det nødvendigvis ikke er så enkelt, ettersom elever kan gjøre det meste riktig i et komplisert regnestykke, men få galt resultat som følge av en slurvfeil. Man kan spørre seg om det er rimelig at elevene straffes for slike følgefeil, eller om de bør få noe uttelling for å ha brukt riktig regnestrategi. Et annet aspekt som gjør elevvurdering vanskelig i matematikk, kan også reises. Hva er *vanskelighetsgraden* på oppgaven? Og hvordan kan dette sammenlignes på tvers av klasserom og skoler? Spørsmålene antyder at det kan være vanskelig å utvikle et tolkningsfellesskap, også innenfor matematikkfaget.

Geir Botten (2003) peker på at den tradisjonelle vurderingspraksisen i matematikkfaget har vært å gjennomføre prøver etter at man er ferdig med et emne eller kapittel i matematikkboka. Han hevder at matematikkprøver har vært den viktigste, og noen ganger den eneste, kilden til å bestemme karakter. Botten peker imidlertid på at det de senere årene har blitt lagt mer vekt på at prøver både skal være en forlengelse og en integrert del av læreprosessen. Fokuset har flyttet seg fra prøver som kontrollapparat og rangeringsverktøy, til læringseffekten av prøveopplegg. Dette har resultert i både endret innhold i prøvene og nye prøveformer (Botten, 2003, s. 179).

Botten viser til matematikkprosjektet *Matte er gøy – vurdering som bindeledd mellom undervisning og læring*, som ble gjennomført i regi av Eksamenssekretariatet, der det ble gjort forsøk med å la elevene selv få velge og lage sine egne oppgaver. Jon Møretrø, som hadde

ansvar for dette prosjektet på Rå ungdomsskole i Bergen, viser til positive erfaringer med en rekke elevaktive vurderingsformer, blant andre vurdering av egen besvarelse, tegning av egen kompetanseprofil og utarbeidelse av og vurdering av egne oppgaver (Møretrø, 2007, s. 158-162).

Botten problematiserer den gjengse oppfatningen om at utregning av poeng på enkelt svar i skriftlige oppgaver gir en rettferdig vurdering. Han peker blant annet på at muntlige matematikkferdigheter også skal vurderes, og at noen elever har bedre forutsetninger for å presentere utregninger muntlig. Han framholder også samarbeid som en viktig egenskap i matematikk (Botten, 2003, s. 182). Under overskriften *Å «rette» med grønt*, kritiserer Botten den tradisjonelle formen for «retting» som matematikklærere har brukt. Han mener selve ordet er uttrykk for et syn på læring som står i kontrast til det rådende læringssyn, både i matematikk og i andre fag. Han framholder at man ved vurdering bør lete etter elevens tanke bak svarene, og ikke bare bedømme om svaret er rett eller galt:

Ved å bruke grønn penn markerer en at det er elevens tanke og oppfatning, og da særlig det positive og utviklingspotensialet, en er interessert i. Også denne «rettingen» inneholder indirekte en markering av hva som «ikke er rett» (Botten, 2003, s. 190)

Botten forklarer at å «rette» med grønt er brukt som en metafor for det å lete etter det positive og utviklingsmulighetene til eleven, men at det også kan være et poeng at en fysisk bytter ut rødpennen med en grønn penn for å tydeliggjøre hensikten. Han viser til at han har møtt mange personer som forteller at deres «sterkeste og mest fastspikrede minner fra skoletida er side etter side, med rødt på rødt, nederlag etter nederlag», og spør retorisk: «Kanskje grønnpennen kan være et bidrag til å redusere nederlaget?» (Botten, 2003, s. 190-191). Botten forklarer hvordan han selv som en symbolsk handling knakk alle rødblyanter og rødpennene han hadde i penalet sitt og erstattet dem med grønne pennen. Slik ble han siden minnet på at «det var det positive og utviklingsmulighetene jeg skulle lete etter, ikke feilene og det negative» (Botten, 2003, s. 191).

Botten presiserer at å «rette» med grønt ikke betyr at en er «snillere» med elevene eller lettere lar unøyaktigheter og slurv passere, og at man tvert i mot blir mer oppmerksom på helheten i det eleven gjør, også det som absolutt ikke fortjener ros. Han anbefaler også å erstatte ordet «rette» med ordene lese, kommentere eller vurdere, som mer presist uttrykker hva man gjør, og som ikke skaper inntrykk av at det alltid er læreren som har de riktige svarene, de riktige tankene og de riktige oppfatningene (Botten, 2003, s. 191).

Mathematics i Queensland

I Queensland har noe av det viktigste utviklingsarbeidet knyttet til elevvurdering i *Mathematics* de siste årene handlet om informasjons- og kommunikasjonsteknologi (IKT). *Australian Mathematical Science Institute* (AMSI) er et offentlig støttet institutt som blant annet har utviklet CD-rom med prøver som korresponderer med lærebøkene *ICE-EM Mathematics* for elever på mellom- og ungdomstrinnet (*upper primary* til *Year 10*) (AMSI, 2007). Ved å gjøre prøver tilgjengelig digitalt, er det mulig for lærere å tilpasse oppgaver til egen undervisning på en effektiv måte. Slike oppgaver, som er utarbeidet av erfarne læremiddellforfattere, gir bedre forutsetninger for å gjøre valide og reliable vurderinger av elevenes matematikkunnskaper.

3.7 Elevvurdering og lærerutdanning

Jeg ønsker her kort å løfte fram forhold ved elevvurdering i forbindelse med lærerutdanning. Lærerutdanningsfeltet generelt har i liten grad vært gjenstand for forskning. Pamela Grossman (2008) hevder det internasjonalt produseres for lite forskning som kan si noe om effektiviteten i lærerutdanningsprogrammer. Hun påstår at lærerutdanningsfeltet er i en krise, der mangel på forskningsbasert kunnskap om lærerutdanning underminerer lærerprofesjonens kredibilitet. Hun etterlyser mer forskning som undersøker organisering, innhold og utbytte av lærerutdanning både for nyutdannede lærere og deres elever (Grossman, 2008, s. 14, 16, 22). Helen Timperley og Adrienne Alton-Lee (2008) forklarer det svake empiriske innslaget i lærerutdanningen med den uklare og fragmenterte tilknytningen lærerutdanninger ofte har til forskningsmiljøene ved universitetene og at fagfolk innenfor akademiske institutter identifiserer seg med deres forskningsdisipliner framfor å videreformidle kunnskapen til profesjonsfeltet (Timperley og Alton-Lee, 2008, s. 330).

3.7.1 Betydningen av gode lærere og god lærerutdanning

Betydningen av gode lærere og en god lærerutdanning bør være åpenbar. Timperley og Alton-Lee (2008) viser til effektstudien *Tennessee Class Size Experiment* i USA, som viste at det hadde større virkning på elevers læringsresultater om lærerne hadde ett standardavvik høyere effektivitet, enn om man reduserte klassestørrelsen fra 25 til 15 elever. De hevder også at det er samfunnsøkonomisk fornuftig å investere i å heve lærernes kompetanse, da dette har stor økonomisk virkning (Timperley og Alton-Lee, 2008, s. 331-332).

Kari Smith (2005) påviser at lærere som har blitt utdannet i et omfattende lærerutdanningsprogram er mer trygg på egen undervisning enn de uten eller med mindre

grundige forberedelser (Smith, 2005, s. 191). I Finland trekkes den femårige utdanning av barneskolelærere fram som en forklaring på landets fremragende resultater i internasjonale komparative studier som PISA (Timperley og Alton-Lee, 2008, s. 333).

I en omfattende metaanalyse av effektstudier identifiserte Timperley og Alton-Lee effektive og mindre effektive strategier for etterutdanning av lærere. De fant to som hadde *liten* virkning. Den første gikk ut på å anse lærere for å være selvregulerende profesjonsutøvere som, om de ble gitt nok tid og ressurser, selv er i stand til å utarbeide hensiktsmessige læringssituasjoner for å heve egen kompetanse i samarbeid med andre lærere. Denne strategien feilet blant annet fordi lærerne ikke ble utfordret på gale eller problematiske holdninger og praksisformer. Den andre mislykkede strategien gikk ut på at eksterne ekspertter lagde «oppskrifter» for undervisning basert på forskning, og presenterte dette for lærere, sammen med et system for oppfølging for å sikre at lærerne implementerte det de hadde lært på riktig måte. Timperley og Alton-Lee mener denne strategien også er lite hensiktsmessig, fordi «experienced teachers do not approach professional learning or teaching situation as empty vessels but rather as people who have rich theories about how students learn, how best to teach them, and what constitutes desired content and outcomes» (Timperley og Alton-Lee, 2008, s. 348).

Den mest vellykkede strategien for etterutdanning av lærere var å søke å styrke læreres grunnleggende pedagogiske fag- og vurderingskompetanse innenfor et konseptuelt rammeverk som ga grunnlag for beslutninger om praktiske undervisningsstrategier. Fortolkning av informasjon innhentet gjennom tester, observasjon, innleveringer og andre former for elevvurdering, ga lærerne muligheten til å forstå i detalj hva elevene alt kunne og hva de trengte å lære. Alle strategier med stor virkning fokuserte på «assisting teachers to translate the pedagogical and assessment knowledge into teaching practice, through coaching, workshops, and participation in communities with colleagues» (Timperley og Alton-Lee, 2008, s. 350).

3.7.2 Opplæring i elevvurdering i lærerutdanningen

Timperley og Alton-Lee viser hvordan kunnskaper om elevvurdering er viktig i lærerutdanningen. De peker på at læreres evne til å finne ut av elevenes kunnskapsnivåer for å avdekke læringsbehov kan ha stor betydning for elevenes læringsresultater (Timperley og Alton-Lee, 2008, s. 343). Dette viser at elevvurdering bør utgjøre en vesentlig del av opplæringen, både i lærerutdannings- og etterutdanningsprogrammer.

Men er dette tilfellet i Norge? Jeg har ikke klart å finne forskning på norsk allmennlærerutdanning som omhandler elevvurdering spesielt. Som nevnt innledningsvis undersøkte ikke NOKUT temaet elevvurdering i undersøkelsen av allmennlærerutdanningene i 2006 (jf. kapittel 1.1.2). Deler av kritikken, blant annet manglende integrering av teori og praksis, pedagogikk og didaktikk, rammer imidlertid også fagområdet elevvurdering (NOKUT, 2006, s. 4). I Stortingsmelding nr. 16 (2006-2007), slo som nevnt Kunnskapsdepartementet fast at både lærerutdanningen og skolen mangler tilstrekkelig kompetanse på elevvering.

Den manglende oversikten over elevvurdering lærerutdanningen i Norge, gjør det nærliggende å se til hvordan det forholder seg i Sverige. Vårt nærmeste naboland hadde en tilsvarende strid om «betyg» i 1970-årene, og kriteriebasert vurdering fikk heller ikke der fotfeste i 1970 og 1980-årene, slik de gjorde i mange andre land. Men i forbindelse reformer i 1990-årene, forlot man i 1994 normalfordelingen og tok i bruk kriteriebasert vurdering som teoretisk rammeverk for vurderingssystemet. Alle elever som avsluttet gymnaset i 1997 ble vurdert basert på et slikt vurderingsprinsipp (Wikström, 2006, s. 119). Sverige var altså mer enn 10 år «tidligere ute». Likevel kan det se ut som at politiske og ideologiske perspektiver har virket begrensende på utviklingen av teori om, og opplæring i, elevvurdering ved lærerutdanningen i Sverige.

Christina Wikström (2007) peker på det paradoksale i at pedagogiske målinger (*mätningar*) i Sverige behandles «styvmoderlig» i utdanningsammenheng, til tross for at kunnskaper og prestasjoner måles i skolen hele tiden. Som eksempel trekker hun fram den svenske lærerutdanningen, der de får ytterst begrenset, hvis i det hele tatt noen, plass (Wikström, 2007, s. 21). Wikström mener lærerne må få bedre forutsetninger i form av utdanning om kunnskapsmålingers teori og praksis, uavhengig om de kommer til å arbeide med karaktersetning eller ikke. «Det är inte en bra lösning att hoppas på att lärare själva ska skaffa sig denna kompetens genom att praktisera 'learning by doing'» (Wikström, 2007, s. 33). Det hevder hun er ineffektivt, og fører til at de fleste gjør som man alltid har gjort eller selv opplevd.

3.7.3 Andre lærerutdanninger i Norden

I kapittel 2 presenterte jeg organiseringen av den norske allmennlærerutdanningen, og de politiske prioriteringene som lå til grunn for at vi har en allmennlærerutdanning som gir undervisningskompetanse i en rekke fag og for hele grunnskolen. Før jeg diskuterer nærmere

erfaringene med vår modell, ønsker jeg å se litt på hvordan lærerutdanningen er organisert i noen av våre naboland.

I Sverige gjennomførte man i slutten av 1990-årene en markant endring av lærerutdanningene. Pedagogikkutdanningen og lærerutdanningen rettet mot de yngste elevene (3-12 år) ble slått sammen til en førskolelærerutdannelse, med en varighet på 3 ½ år. Lærerutdanningen rettet mot de høyeste trinnene i grunnskolen og gymnasen (12-19 år) ble utvidet til 4 ½ år, og lærerstudentene fikk kun undervisningskompetanse i to fag (Egelund, 2008, s. 248). I Danmark la man i 1998 om lærerutdanningen til å være fireårig, og å gi undervisningskompetanse i fire fag, hvorav ett skulle være enten dansk eller matematikk (Egelund, 2008, s. 245). I 2007 ble utdanningen revidert på ny, og strammet inn ytterligere til tre fag. Samtidig ble det foretatt en trinndeling av fagene dansk og matematikk, slik at alle studentene ble utdannet til å undervise på mellomtrinnet (4.-6. klasse), men at man utover det utdannes enten til «indskoling» (1.-3. klasse) eller «udskoling» (7.-9. klasse) (Egelund, 2008, s. 247).

Finland har i nordisk sammenheng gått sin egen vei, ved at man for mer enn 20 år siden la utdanningen til universitetene. Det er en utdannelse til «klasselærer» («Master of Education») som gir undervisningskompetanse i alle fag for trinn 1 til 6, men med en spesialisering i to fag. Utdanningen tar 4-5 år. For å undervise på trinn 7 til 9 må man utdanne seg til «ämneslærer» («Master of arts/science»). Utdanningen omfatter 1-3 fag med spesialisering i ett av fagene, og tar 5-7 år. For begge gradene kreves det en universitetsgrad på masternivå (Egelund, 2008, s. 249). Ved siden av å ha en lærerutdanning organisert som masterutdannelse på universitetet, er det bemerkelsesverdig at fag, fagdidaktikk og pedagogikk inngår i et samspill gjennom utdanningene i Finland. Man har klart å unngå kampene mellom fag og pedagogikk, som norsk og svensk lærerutdanning har vært heftet med (Egelund, 2008, s. 249).

Disse tre nabolandene våre har altså litt ulike modeller for organisering av lærerutdanning. Et felles utviklingstrekk er større grad av spesialisering både mot alder og fag.

3.7.4 Erfaringer med norsk allmennlærerutdanning

Til slutt ønsker jeg å peke på noen erfaringer med norsk allmennlærerutdanning, som vil være relevante for drøftningen. Dette er ingen «uttømmende» presentasjon av feltet, ettersom jeg som nevnt har funnet lite forskning på allmennlærerutdanningen, og det dessuten strekker seg utenfor denne oppgaven å gjøre en grundig gjennomgang av allmennlærerutdanningen. Jeg

ønsker i stedet kort å peke på noen historiske utviklingstrekk, slik de fremstilles av Peder Haug i artikkelen *Skulen og allmennlærerutdanninga* (2008).

I følge Haug var hovedspørsmålet i tiden under «Arbeiderpartistaten» (1935-1980) om man skulle prioritere klasselæreren framfor faglæreren, og hva som var det mest sentrale i læreren sitt arbeid, fag eller pedagogikk. «Skulle læreren først og fremst ha som si oppgave å presentere fagstoff, eller skulle han vere rettleiar, inspirator og vere oppteken av sosiale prosessar og individuelle behov og ønske hjå elevane innanfor ein fellesskap?» (Haug, 2008, s. 260). Haug slår fast at prinsippet om klasselæreren og det som ble kalt en utvidet lærerrolle ble resultatet i plan- og lovrevisjonene, både i 1965 og i 1973/74.

Som jeg har påpekt i kapittel 2, ble dette også resultatet i stortingsbehandlingen i 1998 og 2003 (jf. kapittel 2.1.3). Dette er altså en diskusjon som har sin opprinnelse helt tilbake i mellomkrigstida. Med jevne mellomrom har man diskutert om man skal gjøre allmennlærerutdanningen om til en mer faglig orientert utdanning, men hver gang har man konkludert med at lærerens rolle som «klasselærer» er viktigst. Som påpekt tidligere har den norske allmennlærerutdanningen fått mye kritikk de senere årene. Haug er ikke videre optimistisk når han oppsummerer utfordringene til norsk allmennlærerutdanning:

Svakare faglege nivå på studentane ved inntak, få krav til fag og fagkombinasjonar i studiet, ingen krav om forkunnskapar, låge krav til studieomfang i norsk, lågt faglege nivå i norsk med mange stryk, lite realfag, svake matematikkunnskapar, få lærarar har naturfagkompetanse, mange studentar legg lite tid og arbeid ned i studia, fagmiljøa ved lærerutdanningane er små, osv. (Haug, 2008, s. 266).

Hvorfor er problemene så omfattende? Haug forsøker å sette oss på sporet av ett av problemene:

Ein har forsøkt å løyse den profesjonsfaglege kvalifiseringa gjennom administrative og organisatoriske tiltak. Difor har det vorte mange modellar og ordningar for allmennlærerutdanninga. Ein har prøvd alt, frå at alle fag var obligatoriske, og til stor valfridom. Fag har til ulike tider fått større eller mindre omfang, og vore meir eller mindre sentrale i utdanninga. Ein har hatt sentraliserte og desentraliserte styringsmodellar, opne og lukka planar, osv (Haug, 2008, s. 270).

Som påpekt i kapittel 2, har det blitt initiert endringer av rammeplanen for den norske allmennlærerutdanningen gjentatte ganger, både av departementet og stortinget (jf. kapittel 2.1.3). Mangel på *forutsigbarhet* framstår som et åpenbart problem for høyskolene som tilbyr allmennlærerutdanning. Haug peker på at endringene i lærerutdanningen oftest er legitimert

gjennom endringer i grunnskolen. Han hevder at så lenge disse endringene var veldefinerte og entydige, så klarte lærerutdanningen å leve opp til de, med et etterslep. Men da endringstakten økte fra 1990-årene av, og idealene ble mer varierte og komplekse, fikk utdanningen større problemer både med å orientere seg og å henge med (Haug, 2008, s. 271). Han konkluderer med at «dei reformpedagogiske ideane frå tidlegare har gitt ein skule som samfunnet ikkje lenger vil ha» (Haug, 2008, s. 273). Nå gjelder andre idealer, med større vekt på faglige prestasjoner enn tidligere. Haug mener både grunnskolen og lærerutdanningen har hatt store problemer med å tilpasse seg denne utviklinga, blant annet som følge av at de ikke har blitt gitt nok tid.

3.8 Sammendrag av teoretiske perspektiver på elevvurdering

I dette teorikapitlet har jeg ønsket å få fram viktigheten av at lærere har gode kunnskaper og ferdigheter i elevvurdering. Gjennom en diskusjon av kjernebegreper i vurderingsterminologien har jeg avgrenset diskusjonen til summative og formative aspekter ved elevvurdering. Jeg utviklet noen grunnleggende problemstillinger knyttet til summative vurderinger, forankret i teori om validitet (*construct* og *structure validity*) og reliabilitet (konsistens på tvers av prøver; *day-to-day* og *inter reliability*). Jeg har framholdt en bredde av vurderingsoppgaver og –kontekster som en forutsetning for å gjøre valide vurderinger, og systemer for kryssretting eller moderering som en forutsetning for å gjennomføre reliable vurderinger. Tilsvarende utviklet jeg noen grunnleggende problemstillinger knyttet kvaliteten av læreres formativ vurderinger. Jeg framholdt det som viktig at læreren legger til rette for at elevene kan vurdere egen måloppnåelse og ikke overdriver bruken av karakterer og andre symboler. Videre bør læreren gi elevene hyppige tilbakemeldinger, med utgangspunkt i vurderingskriterier som eksplisitt beskriver målene for opplæringen og elevenes måloppnåelse.

Videre gjorde jeg et historisk tilbakeblikk for å peke på mulige grunner til utfordringene vi har knyttet til elevvurdering i Norge. Jeg pekte på at 1970-årenes karakterstrid med jevne mellomrom har gjenoppstått og framholdt at vurderingsfeltet var gjenstand for ideologisk steinkasting helt inn til slutten av 1990-årene, med den største lærerorganisasjonen som en av de sterkeste motstanderne av karaktervurdering, også på ungdomstrinnet. Dette hevdet jeg har bidratt til at spesielt karaktersetning i Norge er behandlet som et politisk og ideologisk, framfor praktisk, fagfelt i norsk utdanningsteori, hvilket har ført til at det i liten grad er utviklet pedagogiske prinsipper, systemer og verktøy for karaktersetning i Norge.

Jeg har gått igjennom svakheter knyttet til både summativ og formativ vurdering i Norge. Jeg har reist spørsmål om fraværet av karakter på barnetrinnet bidrar til mindre fokus på faglig måloppnåelse på disse trinnene, som følge av at lærere i mangel på kriterier for vurdering av elevenes måloppnåelse har dårligere forutsetninger for å gjøre formative vurderinger. Jeg har dokumentert at det er gjennomført svært få undersøkelser av påliteligheten i lærernes karaktersetning, men at det er grunnlag for å tro at dette er en stor utfordring, som følge av at man verken har felles kriterier og standardbeskrivelser i læreplanen eller systemer for intern eller ekstern moderering. Jeg har pekt på forskning som har avdekket at norske grunnskolelærere i for liten grad gir tydelige kravspesifikasjoner og konkrete tilbakemeldinger. De formative tilbakemeldingene er mer kjennetegnet av generell ros, som det er dokumentert at har liten virkning.

Videre viste jeg til noen undersøkelser i Queensland som viser at man i stor grad har lyktes i å utvikle et tolkningsfellesskap mellom lærere i *senior secondary*, som følge av systemet for ekstern moderering. Samtidig har jeg trukket fram kritikk av delstatens tilnærming til standardbasert, som påpeker at man har en urealitisk forventning om at de felles kriteriene og standardene sikrer en objektiv vurdering. Videre blir det hevdet at eksplisitte kriterier ikke nødvendigvis motvirker sosial reproduksjon, ettersom noen elever har bedre forutsetninger for å tolke kriteriene enn andre.

I en diskusjon av fagspesifikke problemstillinger knyttet til elevvurdering, har jeg pekt på at man gjennom metodene prosessorientert skriveopplæring og mappevurdering har utviklet en forståelse av kompleksiteten knyttet til karaktersetning i norskfaget. I matematikkfaget ser det ut til at man tar lettere på dette, fordi det finnes en «fasit». Det pekes riktignok på problemstillingen med vurdering av følgefeil, men det ser ikke ut til at man i dette faget har vært opptatt av vanskelighetsgraden i vurderingsoppgavene og om disse er sammenlignbare på tvers av skoler. I matematikkfaget kritiserer en av lærebøkene den tradisjonelle strenge røde rettepenner, og «'retting' med grønt», brukes som en metafor for å uttrykke at man i større grad bør legge vekt på det positive. I begge fag fremheves det at vurdering av muntlige ferdigheter også er viktig. I Queensland har en undersøkelse vist hvordan engelsklærere er påvirket av den sosiale konteksten når de setter karakter, og at de prøvde å «se eleven i teksten», når de ble bedt om å vurdere andre elevers arbeider. I matematikkfaget ser det ut til at man i Australia og Queensland har utviklet gode IKT-baserte løsninger for å gi lærerne tilgang på vurderingsoppgaver sammen med lærebøker for faget, som gir bedre forutsetninger for å gjøre pålitelige vurderinger.

Avslutningsvis har jeg diskutert betydningen av gode lærere og en god lærerutdanning. Undersøkelser viser at elevers læringsresultater i stor grad kan knyttes til lærernes kompetanse. Undersøkelser har vist at den mest vellykkede strategien for etterutdanning av lærere var å søke å styrke læreres grunnleggende pædagogiske fag- og vurderingskompetanse innenfor et konseptuelt rammeverk som ga grunnlag for beslutninger om praktiske undervisningsstrategier. Jeg har vist til organiseringsformer for andre lærerutdanninger i Norden, der utdanningene i større grad er spesialisert mot fag og alderstrinn. Videre har jeg trukket fram noen problemstillinger som har blitt reist som mulige årsaker til svakheter i den norske allmennlærerutdanningen. De stadige endringene i de nasjonale rammeplanene, og uforutsigbarheten dette medfører for høyskolene, framholdes som en mulig forklaring på en del av allmennlærerutdanningens problemer.

4 Metode for gjennomføring av studien

I dette kapittelet beskriver jeg de metodiske hensynene jeg har tatt og utdyper framgangsmåten som er brukt for intervjuene i studien. Metodekapittelet skal gi tilstrekkelig informasjon til å vurdere grunnlaget for analyser, drøftninger og konklusjoner jeg har gjort i de neste kapitlene. I del 4.1 beskriver jeg prosessen med valg av fokus for oppgaven og det forberedende arbeidet for studien. I del 4.2 peker jeg på grunnleggende problemstillinger knyttet til validitet og reliabilitet, som jeg har forsøkt å ta hensyn til i utarbeidelsen av forskningsmetoden. I del 4.3 redegjør jeg for forankringen studien har i metodiske disipliner innenfor kvalitativ forskning: fenomenologi og hermeneutisk fortolkning. I del 4.4 begrunner jeg valg av kvalitative intervjuer som primær datainnsamlingsmetode, og prinsipper for å tilstrebe en gyldig meningsproduksjon gjennom slike intervjuer. Del 4.5 inneholder en presentasjon av metodiske avveininger for utforskningen av de to casene Norge og Queensland. I del 4.6 til 4.12 drøfter jeg de ulike fasene i det metodiske arbeidet: utarbeiding av intervjuguide; utvalg og samplingskriterier; gjennomføring av intervjuer i Norge og Queensland; transkribering, analyse og framstilling av data. I del 4.13 sammenfatter jeg metodekapittelet, og oppsummerer strategiene jeg har brukt for å gjennomføre en så valid og reliabel studie som mulig.

4.1 Valg av fokus for oppgaven og forberedende arbeid for studien

Beslutningen om å gjennomføre en komparativ kvalitativ studie av elevvurdering i lærerutdanningen i Norge og Queensland, ble bestemt i møte med veileder i januar 2007. I møtet ble prosjektbeskrivelse utarbeidet og søknad til *Norsk Samfunnsvitenskapelig Datatjeneste* (NSD) forberedt. På dette tidspunktet var et utvekslingsopphold ved Griffith University i Queensland allerede planlagt. Et nytt møte ble avholdt i slutten av februar, der søknaden til NSD ble ferdigstilt og sendt. Samme dag ble også de nødvendige papirer i forbindelse med veiledningssamarbeidet undertegnet og levert til Pedagogisk forskningsinstitutt, før avreise til Australia. De første to ukene på Gold Coast, en times kjøring sør for Queensland's hovedstad Brisbane, gikk med til å finne seg til rette med studentbolig og på universitetet. I begynnelsen av mars tok jeg kontakt med Joy Cumming ved Griffith University, som jeg ønsket å knytte meg til i forbindelse med studiene der.

Selv om det ikke er knyttet til masterstudiet mitt ved Universitetet i Oslo, vil jeg nevne at jeg alt på dette tidspunktet bestemte meg for å gjennomføre en 40 credits «dissertation» om vurderingssystemet i *senior secondary* i Queensland, ved siden av den norske

masteroppgaven. Som følge av at jeg alt hadde avsluttet mange fag i masterutdanningen min, hadde jeg kun anledning til å gjennomføre ett semesters utveksling fra Universitet i Oslo. Ettersom de fleste lærerstudenter i Queensland avsluttet lærerutdanningen i desember, var det imidlertid ikke mulig å gjennomføre datainnsamlingen der før ved avslutningen av et eventuelt andre semester. Som følge av strenge visumregler var det derfor hensiktsmessig å gjennomføre en separat ettårig mastergrad i Queensland, for å kunne bli lenger i landet (alternativet ville vært å plukke bananer!). Til sammen ble jeg derfor tre semester i Queensland, og jobbet med to masteroppgaver parallelt, med veiledning fra Cumming på Griffith-oppgaven. Følgelig tok dette norske masteroppgaveprosjektet lenger tid enn normalt, men jeg fikk tilsvarende bedre forutsetninger for å bli kjent med rammer og systemer for elevvurdering og lærerutdanning i Queensland enn jeg ellers ville hatt.

I forbindelse med oppgaven om vurderingssystemet i videregående skole i Queensland fikk jeg mye hjelp av Graham Maxwell, som tidligere hadde ansvar for systemet for ekstern moderering i QSA, og som jeg hadde kommet i kontakt med ett år i forveien. Han organiserte møter med de ansvarlige for to særskilte områder innenfor vurderingsfeltet i Queensland: satsingen QCAR-satsingen for Year 1-10 og systemet for ekstern moderering av læreres karaktersetning i Year 11 & 12, som jeg har omtalt i kapittel 2. Disse møtene bidro til å gi meg innsikt i hvordan skolesystemet generelt, og elevvurdering spesielt, er organisert, og var derfor svært verdifulle også for forberedelsene av denne studien.

4.2 Validitet og reliabilitet

Ved ethvert forskningsprosjekt er en av forskernes viktigste oppgaver å sørge for valide og reliable funn. Som ved elevvurdering handler validitet om funnernes gyldighet, om studien måler det den gir seg ut for å måle, mens reliabilitet handler om undersøkelsens pålitelighet. For å sikre valid datainnsamling er *begrepsbruk* svært viktig. For å kunne sammenligne informantenes utsagn, er det viktig at de har en så lik som mulig forståelse av spørsmålene man stiller. På samme måte er det viktig at forskeren forstår informantenes svar, slik de er ment. I forbindelse med en komparativ studie på to forskjellige språk er dette en særlig utfordring. Et annet viktig krav som må møtes for at en studie skal kunne regnes for å være valid, er forskerens uavhengighet. Strategier må benyttes for å unngå at forskerens egne synspunkter påvirker informantene og slik legger føringer for konklusjonene som trekkes.

Det viktigste hensynet til reliabilitet i forskningsstudier er studiens etterprøvbarehet. Full etterprøvbarehet er ikke mulig i kvalitative studier, da det ikke vil være mulig å gjennomføre de samme intervjuene, ettersom disse er situasjonsbestemte. Intervjuer man de

samme subjektene ved en senere anledning, vil deres situasjon være annerledes enn da de sist ble intervjuet. Og intervjuer man andre subjekter, så vil konteksten være annerledes enn for subjektene jeg intervjuet. Likevel er det viktig å gjøre forskningsstrategien så transparent som mulig, slik at det er mulig for andre å gjøre en mest mulig tilsvarende undersøkelse. For å gjøre en replikasjon av studien mulig har jeg lagt vekt på:

- å gjøre alle spørsmålene tilgjengelig, ved å legge ved intervjuguiden som appendiks (appendiks 11 og 12);
- å redegjøre grundig for samplingskriterier og prosedyrer, slik at det lett kan gjentas; og
- å være grundig i kildegjengivelsen av analyserte dokumenter, slik at de er lett for andre å undersøke.

I den følgende gjennomgangen av forskningsmetoden forsøker jeg gjennomgående å vise strategier jeg har brukt for å ivareta studiens validitet og reliabilitet.

4.3 Valg av forskningsmetodisk tilnærming

Et grunnleggende spørsmål i forbindelse med ethvert kvalitativt forskningsprosjekt er dette: Kan og bør sosial virkelighet sees på som objektive realiteter som eksisterer uavhengig av de sosiale aktørene, eller kan og bør de sees på som sosiale konstruksjoner skapt av sosiale aktørers oppfatninger og handlinger? Peter L. Berger og Thomas Luckmanns bok *Den samfunnsskapte virkelighet* (1967) er blitt stående som et av de mest betydningsfulle innleggene i denne debatten, som var på sitt sterkeste under den såkalte positivismestriden i 1960 og 1970-årene. Berger og Luckmanns standpunkt framgår av bokens tittel. Hovedpoenget deres kan sammenfattes i følgende korte utdrag: «Spesifikke ansamlinger av 'virkelighet' og 'kunnskap' kan knyttes til spesifikke sosiale kontekster» (Berger og Luckmann, 2004, s. 25-26). Det eksisterer en rekke forskjellige sosiale kontekster, derav forskjellige konstruksjoner av kunnskap. Uten å undervurdere viktige bidrag fra den «objektivistiske» forståelseshorisonten, kan man slå fast at oppfatningen av kunnskap som en sosial konstruksjon har fått stadig større utbredelse.

For samfunnsforskeren som gjennomfører et kvalitativt intervju er et overordnet spørsmål: Hvordan skal man tolke informantenes utsagn? Mer konkret: Innenfor hva slags sosiale kontekst er utsagnene gitt? I de aller fleste studier har man et ønske om at kunnskapen man produserer ikke bare skal kunne knyttes til denne personen som har blitt intervjuet – at de i det minste til en viss grad kan generaliseres. Derfor brukes gjerne flere informanter. Men for hver nye informant man intervjuer må man forholde seg til en helt ny ansamling av

virkelighet og kunnskap skapt innenfor andre sosiale kontekster. Utviklingen av en forskningsmetode handler blant annet om å utvikle strategier som bidrar til at kunnskapen man konstruerer er så valid (gyldig) som mulig. I denne studien var valg av lærerutdanningen som forskningsarena et sentralt virkemiddel for å oppnå størst mulig grad av likhet i de intervjuedes sosiale kontekster, for å ha et bedre grunnlag for å trekke slutninger på tvers av casene. Hadde jeg for eksempel intervjuet lærere som hadde ti års undervisningserfaring, ville deres sosiale kontekster være mer forskjellig, som følge av ulik undervisningserfaring, enn tilfellet var for disse nyutdannede lærerne.

Det finnes en rekke forskningstilnæringer, som med utgangspunkt i fokus og teoretiske perspektiver argumenterer for ulike metodiske løsninger for å finne fram til gyldige funn (blant andre *fenomenologi*, *etnometodelogi*, *hermenutikk*, *narrative analyser*) (Patton, 2002, s. 104-120). Oppgavens fokus, komparative ambisjon og omfang har imidlertid ikke gjort det mulig å følge én av disse metodene eksklusivt. Jeg har dratt veksler på teorigrunnlaget for både fenomenologisk og hermeneutisk metode, og vil her drøfte noen aspekter ved disse to forskningsmetodiske retningene som jeg har hatt nytte av i utviklingen av forskningsmetode for denne oppgaven.

4.3.1 Fenomenologi

Fenomenologi er en retning innenfor kvalitativ forskningsmetode der man søker å finne ut av et fenomens *natur*, utelukkende med utgangspunkt i menneskers subjektive oppfatninger. Denne forskningsmetoden har blitt svært populær og er blitt anvendt i en rekke forskjellige varianter, hvilket også har gitt metoden et noe utvannet innhold. Det er en svært empirinær tilnærming som, i sin reneste form, kan kritiseres for å ikke ta i betraktning strukturene som ligger til grunn for aktørenes intensjonelle oppførsel (Scott, 1996, s. 66).

Når jeg i denne oppgaven drar veksler på fenomenologiske prinsipper, er det ved primært å støtte meg til intervjuer som datakilder i undersøkelsen av «fenomenet» elevvurdering i lærerutdanningen. Videre har jeg i intervjuanalysen i stor grad gitt plass til informantenes egne stemmer. Jeg har også unnlatt å teste hypoteser overfor informantene direkte (jf. avsnittet om tilsiktet naivitet nedenfor). Hele studien er imidlertid designet med utgangspunkt i mine egne forhåndsantagelser og litteraturanalyser, og er derfor delvis teoridrevet, hvilket bryter med et av fenomenologiens grunnprinsipper (Kvale, 1996, s. 54).

4.3.2 Hermeneutisk fortolkning

Hermeneutikk dreier seg essensielt om fortolkning av kunnskap, som står sentralt særlig i analyse- og drøftningsdelen av denne studien. Et fundamentalt spørsmål ved hermeneutisk fortolkning er dette: «What are the conditions under which a human act took place or a product was produced that make it possible to interpret its meanings?» (Kvale, 1996, s. 46). Selv om hermeneutikk tradisjonelt har vært knyttet til fortolkning av tekster fra litteratur, religion og juss, så har begrepet «tekst» etter hvert blitt utvidet til å inkludere diskurs (samtale) og handlinger. Ved intervjustudier som denne er ikke en slik forståelse av tekst kontroversiell, ettersom det med utgangspunkt i samtalen vitterlig produseres en tekst – transkripsjonen. Sett med et hermeneutisk perspektiv har vi da to tekster som må fortolkes: først dialogen som skaper intervjueteksten; og deretter den skapte teksten (Kvale, 1996, s. 49). Denne «doble hermeneutikken», slik det ble uttrykt av Giddens (1982), gjør samfunnsvitenskapelige data vanskelig å fortolke (Giddens, 1987, s. 20).

Fortolkning av tekst vil alltid pregs av fortolkerens antagelser. «The interpreter cannot 'jump outside' the tradition of understanding he or she lives in» (Kvale, 1996, s. 48). En vanlig strategi for å begrense dette problemet er å gjøre egne antagelser eksplisitte. Slik kan man oppnå større bevissthet om hvordan visse formuleringer og begreper kan legge føringer for fortolkningen, og ta hensyn til dette når man tolker data. Dette har jeg blant annet gjort i intrduksjons- og teorikapitlet, ved å henvise til egne publikasjoner.

4.4 Metodiske avveininger i forbindelse med kvalitative intervjuer

En viktig grunn til å velge intervju som den primære datainnsamlingskilde er at intervjuer, til forskjell fra for eksempel observasjoner, kan gi forskeren tilgang til informantens synspunkter, også knyttet til tidligere hendelser. Intervjuer gir tilgang til situasjoner som forskeren selv ikke har tilgang til, eller blir nektet adgang til (Scott, 1996, s. 67). Intervjuer kan imidlertid skape mye misforståelser. Scott beskriver to hovedkilder til misforståelser som kan lede til at intervjuer gir en fundamental feilrepresentasjon av subjektets erfaringer:

- misforståelse mellom intervjueren og subjektet; og
- misforståelse mellom subjektets redegjørelse og hendelsene og perspektivene subjektet beskriver (Scott, 1996, s. 66).

Representasjoner vil aldri kunne bli helt presise (jf. diskusjonen om virkeligheten som en sosial konstruksjon), men en rekke strategier kan benyttes for å i størst mulig grad unngå eller begrense disse kildene til misforståelser. Den første typen misforståelser kan begrenses ved hjelp av en rekke «intervjutekniske» strategier og virkemidler. Den andre typen misforståelser

knytter seg til mer grunnleggende faktorer: subjektene hukommelse og representasjon av hendelser. I den klassiske boken *Mind, Self, and Society*, første gang utgitt i 1934, pekte George Hebert Mead på at selvet alltid er i utvikling og konstituerer seg selv ved forskjellige øyeblikk og rekonstruerer virkeligheten både i fortid og nåtid. Aktører forklarer ikke bare hva de gjorde, men også hva de føler de burde ha gjort. Sosiale aktører vil alltid gi mening til fortiden i lys av nåtiden (Mead, 2005, s. 144). Intervjuutsagn må derfor alltid fortolkes i lys av både en samtidig og fortidig kontekst. Her presenteres derfor noen metodiske avveininger jeg har gjort for å minske muligheten for at slike misforståelser skal forringe gyldigheten (validiteten) av analysene og konklusjonene som trekkes på grunnlag av intervjuene.

4.4.1 Meningsproduksjon

Kvale definerer kvalitative forskningsintervjuer som «attempts to understand the world from the subjects' point of view, to unfold the meaning of peoples' experiences, to uncover their lived world prior to scientific explanations» (Kvale, 1996, s. 6). Han kaller intervjustrategien han har utarbeidet for «a semistructured life world interview». Dette beskriver han som «an interview whose purpose is to obtain descriptions of the life world of the interviewee with respect to interpreting the meaning of the described phenomena» (Kvale, 1996, s. 4).

Kvale benytter to metaforer til å illustrere hvordan forståelse bør skapes med utgangspunkt i subjektene selv. *Gruvearbeideren (the miner)* ser på kunnskap som skjult metall som skal avdekkes. Enten man søker objektive fakta som kan kvantifiseres, eller biter av essensiell mening, så er hensikten at kunnskap som hentes ut av subjektene ikke skal «forurenses» av gruvearbeideren. Den andre metaforen forstår intervjueren som *en reisende (the traveler)*, hvis reise leder til en historie han kan fortelle om når han kommer hjem igjen. Den reisende intervjueren reiser i ukjent terreng eller med kart, og blir kjent med forskjellige temaer ved å stille spørsmål som leder menneskene til å fortelle deres egne historier. Kvale viser til den forvandlende virkning av reising, slik det uttrykkes i det tyske begrepet *bildungsreise* – dannelsesreise – i beskrivelsen av denne reisefilosofien. Reisen behøver ikke bare lede til ny forståelse av mennesker og steder man oppsøker, den reisende kan også forandre seg. Refleksjoner under og etter reisen kan lede intervjueren til nye forståelser av seg selv, sine egne oppfatning og hans egen kultur (Kvale, 1996, s. 5-10).

Mens gruvearbeidermetaforen illustrerer en forståelse av kunnskap som gitt, bygger reisemetaforen på en konstruktivistisk forståelse som impliserer en samtalende tilnærming til forskning. Kvale hevder at det i samfunnsforskningen er en dreining fra eksterne observasjoner og manipulering av menneskelige subjekter til en forståelse av mennesker ved

hjelp av samtaler. Han støtter seg til reisemetaforen og bygger sine metodiske anvisninger på en forståelse av intervju som en samtale. (Kvale, 1996, s. 13).

4.4.2 Intervjustrategi

Kvale peker på at det til forskjell fra kvantitative undersøkelser finnes få standardiserte prosedyrer for hvordan man skal gjennomføre kvalitative intervjuer. Mange metodiske valg må derfor gjøres på stedet, under intervjuet (Kvale, 1996, s. 32). Jeg forsøkte derfor å i størst mulig grad gjøre meg kjent med både temaet for intervjuet og de metodiske utfordringene på forhånd. Formålet med kvalitative forskningsintervjuer er å beskrive og forstå sentrale temaer i subjektene erfaringer. Det er nødvendig å lytte til både eksplisitte meninger og det som blir sagt «mellom linjene». En måte å forsikre seg om at man oppfatter det implisitte korrekt, som jeg har brukt mye i mine intervjuer, er å «sende tilbake» fortolkningen man selv gjør, slik at subjektet kan bekrefte, avkrefte eller presisere (Kvale, 1996, s. 32).

Et kvalitativt forskningsintervju bør være *fokusert og spesifikt*. Den primære hensikten er å beskrive spesifikke situasjoner fra informantens verden. Intervjuren bør lede informantene mot særskilte temaer, men ikke særskilte meninger om disse temaene (Kvale, 1996, s. 33). Kvale peker på at en viktig målsetning med kvalitative intervjuer er å erverve *ufortolkede* beskrivelser. *Hvorfor* de erfarer og oppfører seg som de gjør er det primært opp til forskeren å finne ut av. Kvale illustrerer dette poenget med analogi til en leges diagnostisering:

The doctor does not start by asking the patient why he is sick, but rather asks the patient what is wrong, what he is feeling, and what the symptoms are. On the basis of the information obtained, the doctor may then formulate a hypothesis of which illness may be likely. Further questioning proceeds from this hypothesis, and on the basis of the patient's answers and results from other methods of investigation, the doctor then makes the diagnosis (Kvale, 1996, s. 33).

Kvale peker på at det både for doktoren og forskeren vil være tilfeller der det er viktig å vite subjektene egen forklaring av tilstanden. Den primære oppgaven er imidlertid å erverve beskrivelser slik at forskeren har et relevant og persist grunnlag for å gjøre egne fortolkninger (Kvale, 1996, s. 33). Dette var et sentralt prinsipp jeg fulgte for å unngå å infiltrere i informantens meningskonstruksjon. Jeg la ned mye arbeid i å utarbeide en omfattende intervjuguide med spørsmål strukturert på en måte som kunne gi best mulig grunnlag for «diagnosearbeidet». Samtidig la jeg vekt på å strukturere intervjuguiden på en måte som bidro til at jeg unngikk ledende spørsmål. Gjennomgående benyttet jeg åpne spørsmål. Samtidig

benyttet jeg rigide spørsmålskategorier (blant annet adskilte spørsmål for hva man lærte gjennom teori- og praksisopplæring), for å balansere disse to hensynene.

4.4.3 «Tilsiktet naivitet»

Kvale peker på betydningen av at intervjueren er godt kjent med temaet man samtaler om, for å kunne oppfatte de mange nyansene i subjektens beskrivelser. Dette står imidlertid i motsetningsforhold til et annet viktig prinsipp – å ikke la egne antagelser og oppfatninger bli styrende for samtalen. En strategi for å balansere disse to hensynene er hva Kvale kaller «tilsiktet naivitet»: «The interviewer should be curious, sensitive to what is said—as well as to what is not said—and critical of his or her own presuppositions and hypotheses during the interview» (Kvale, 1996, s. 33-34). Dette forutsetter en kritisk bevissthet om egne antagelser og oppfatninger. Da jeg planla denne studen, var det på bakgrunn av en oppfatning om at studenter ved norsk allmennlærerutdanning lærer lite om elevvurdering, og at studentene i Queensland lærer mer enn studentene i Norge. Spesielt under intervjuene i Norge, der jeg fra før av hadde relativt god kjennskap til en rekke problemstillinger og svakheter knyttet til elevvurdering, var det svært viktig å unngå at disse oppfatningene la føringer på intervjuet.

En slik «tilsiktet naivitet» ville imidlertid ikke la seg gjennomføre dersom informantene på forhånd var kjent med mine oppfatninger. Jeg var derfor påpasselig med å ikke fortelle informantene om mine egne synspunkter, slik de blant annet har kommet til uttrykk i boken *Elevvurdering i skolen – grunnlag for kulturendring* (Tveit, 2007c), i forkant av intervjuene. Lanseringen av boken i februar 2007 medførte omtale i det største fagbladet for lærere i Norge – Utdanning – både i bladet og på nettsidene. Intervjuene i Norge ble imidlertid først gjennomført i juni, og om informantene i det hele tatt hadde lest disse artiklene, så var det liten grunn til å tro de ville huske det fire måneder senere. Jeg var likevel forberedt på at det kunne skje. I tilfelle ville det ikke være forenlig med forskningsetiske prinsipper å legge skjul på min bakgrunn. Selv om det går på akkord med prinsippet om å ikke tilkjenne egne oppfatninger, var planen å bekrefte at dette er et spørsmål jeg har engasjert meg mye i tidligere, men at dette ikke burde være styrende for intervjuet, og noe vi derfor i tilfelle kunne snakke om etterpå. Dette ble som forventet ikke nødvendig, hvilket gjorde det mulig å benytte strategien for «tilsiktet naivitet».

4.5 Metodiske avveininger i forbindelse med case studier

Case studier er en mye brukt forskningsmetode, ikke minst i forbindelse med kvalitative komparative studier. Robert E. Stake (1997) mener case studier er særlig egnet for å utforske

«bounded systems which have a character and a totality which cannot be represented by scores, and are conducted to understand the complexity in dynamic systems» (Stake, 1997, s. 406). Det finnes svært mange metodebøker for case studier. Strategien for designet av dette case studiet har jeg i stor grad utviklet med utgangspunkt i Robert K. Yin sine anbefalinger, slik han presenterer dem i tredje utgave av boken *Case study research: design and methods* fra 2003.

Yin mener case studier er en egnet strategi for empirisk undersøkelser av samtidige fenomener, særlig der skillelinjene mellom fenomener og konteksten er uklar (Yin, 2003, s. 13). Han beskriver case studier som undersøkelser som har å gjøre med en «technically distinctive situation in which there will be many more variables of interest than data points» og der man derfor er avhengig av ulike kilder som trianguleres (Yin, 2003, s. 13-14). Han presenterer en omfattende forskningsstrategi som inkluderer design av studien og teknikker for både innsamling og analyse av data. Omfanget av denne studien har ikke gitt rom for bruk av mange kilder som kan trianguleres, men jeg har likevel designet studien på en slik måte at prinsipper for triangulering kan benyttes.

Yin lister opp fem nøkkelspørsmål som er vanlig å benytte i samfunnsforskning: *hvem, hva, hvor, hvordan og hvorfor*. Han mener at de to siste spørsmålsformene er eksplorerende og at de derfor leder til bruk av case studier, historiske undersøkelser eller eksperimentelle studier som den foretrukne metode. Det underliggende rasjonale er at spørsmålene *hvordan* og *hvorfor* har å gjøre med operasjonelle koblinger som må spores over tid, til forskjell fra størrelser og frekvenser (der *surveys* ville vært å foretrekke) (Yin, 2003, s. 6). Videre peker han på at case studier har en klar fordel når et *hvordan* eller *hvorfor* spørsmål blir stilt om samtidige hendelser (hvilket ekskluderer historiske tilnærminger) som forskeren har liten eller ingen kontroll (hvilket ekskluderer eksperimenter) (Yin, 2003, s. 9). Yin er altså nøye med å forklare hvorfor case studier er en egnet forskningsstrategi for eksplorerende studier.

4.5.1 Deskriptiv og eksplorerende studie

Denne studien er en kombinert *deskriptiv* og *eksplorerende* studie. I deskriptive studier blir situasjoner beskrevet, uten at man forsøker å finne forklaringer på resultatene. Eksplorerende studier gjennomføres med den hensikt å *generere* forskningsspørsmål, heller enn å svare på dem. Som følge av den manglende forskningen på feltet i Norge, ville en ren deskriptiv studie av elevvurdering i allmennlærerutdanningen vært et viktig bidrag. Her anvendes imidlertid også et eksplorerende fokus med den hensikt å generere teorier og forskningsspørsmål som kan utforskes i senere studier. For å skille mine deskriptive og eksplorerende tilnærminger til

datamaterialet har jeg valgt å dele analyse- og drøftningsdelene i denne oppgaven i to kapitler (kapittel 5 og 6).

Studien baserer seg utelukkende på kvalitative data. Den har imidlertid ikke et rent induktivt fokus, som i rent fenomenologiske studier. I følge Yin bør man også i case studier på forhånd utarbeide teoretiske forestillinger som kan veilede datainnsamling og analyser. Teori om hva som undersøkes bør utvikles, enten formålet med studien er å utvikle eller å teste teorier (Yin, 2003, s. 13-15, 28). Yin hevder at mens noviser gjerne tror at hensikten med en litteraturanalyse er å finne *sva*r på hva som er kjent innenfor et forskningsfelt, så går erfarne forskere gjennom litteraturen for å utvikle skarpere og mer innsiktsfulle *spørsmål* (Yin, 2003, s. 378). Gode forskningsspørsmål er målsetningen for teorigjennomgangen man gjennomfører før datainnsamlingen starter. Hensikten er å ha et teoretisk grunnlag som kan veilede forskeren i valg av hva slags data som skal innsamles og hvordan disse bør analyseres (Yin, 2003, s. 9). Selv om jeg kun benytter kvalitative data i min studie, vil det følgelig være galt å beskrive studien som rent induktiv. Litteraturanalysen (kapittel 3) og forskningsstrategien jeg utviklet, formet spørsmålene jeg stilte. Studien kan derfor sies å ha en kombinert induktiv-deduktiv tilnærming (Patton, 2002, s. 67).

4.5.2 Forskningsdesign

Jeg har innledningsvis forklart og begrunnet valg av Norge og Queensland som hovedenheter i case studiet (jf. kapittel 1.1.3), men jeg har ikke grundig forklart hvorfor jeg har valgt å intervju lærerstudenter med fordypning i morsmåls- og matematikkfaget. Her følger en begrunnelse for valg av disse fagene.

Valg av fag

Hovedgrunnen for å velge fagene *Norsk* og *English* er at morsmålet er grunnleggende i alle andre fag i skolen og i samfunnet som sådan. Videre er det en vanlig oppfatning at elevvurdering i språkfag er svært vanskelig, som følge av de mange subjektive aspektene ved slike vurderinger. Hensikten med denne studien var imidlertid ikke å utforske elevvurdering i morsmålsfaget alene, men å bruke disse til å si noe om opplæring i elevvurdering ved lærerutdanningene. Lærerutdanningen inneholder åpenbart en hel rekke fag, og man må anta at det er betydelige forskjeller mellom fagene. For å kunne utforske lærerutdanningen mer generelt ville det være ønskelig å utforske en rekke fag, men det ville kreve mer ressurser enn jeg hadde tilgjengelig for denne studien. For å kunne utvikle en forståelse av noen likheter og forskjeller mellom fag var det imidlertid nødvendig å utforske minst ett annet fag. Yin

presenterer flere argumenter for å velge «multiple-cases» framfor «single-case» design, selv om bare en to-case studie er mulig. Single-case studier er sårbare fordi man «put all your eggs in one basket» (Yin, 2003, s. 53), mens multiple-case design muliggjør *repliseringsstrategier* for å validere funnene. Med replisering mener Yin at man gjennomfører parallelle undersøkelser der man for eksempel spør to representanter fra det samme utvalget de samme spørsmålene. Dette framholder han at gir funnene større troverdighet, og at det gir mer interessante analyser, ettersom det trolig vil være noen forskjeller mellom casene (Yin, 2003, s. 53).

En av fordelene med multiple-case studier som Yin peker på, er muligheten for å «predict contrasting results for predictable reasons» (Yin, 2003, s. 53). Ved å velge et annet fag som skiller seg betydelig fra *Norsk* og *English*, kan man regne med å finne andre særtrekk knyttet til elevvurdering. For meg var det naturlig å kontrastere morsmålsfaget, som hører til språk- og humanistiske fag, med et fag fra den naturvitenskapelige fagkretsen. Blant disse er det gode grunner for å velge matematikkfaget. For det første blir matematikk sett på som et svært viktig fag fordi det gjerne anses for å være grunnleggende for alle de naturvitenskapelige fagene. En amerikansk studie har bekreftet dette empirisk. Elever som gjennomførte avanserte matematikkfag i amerikansk *high school*, oppnådde bedre resultater i biologi, kjemi og fysikk når de begynte på *college*, mens disse fagene ikke ga tilsvarende tverrfaglig utbytte (P. M. Sadler og Tai, 2007, s. 458). Videre ser det ut til å være en vanlig oppfatning at karakternivå i matematikk, i motsetning til *Norsk* og *English*, er relativt lett å bli enige om. Slik sett gjør de to fagene en kontrasterende komparativ analyse mulig.

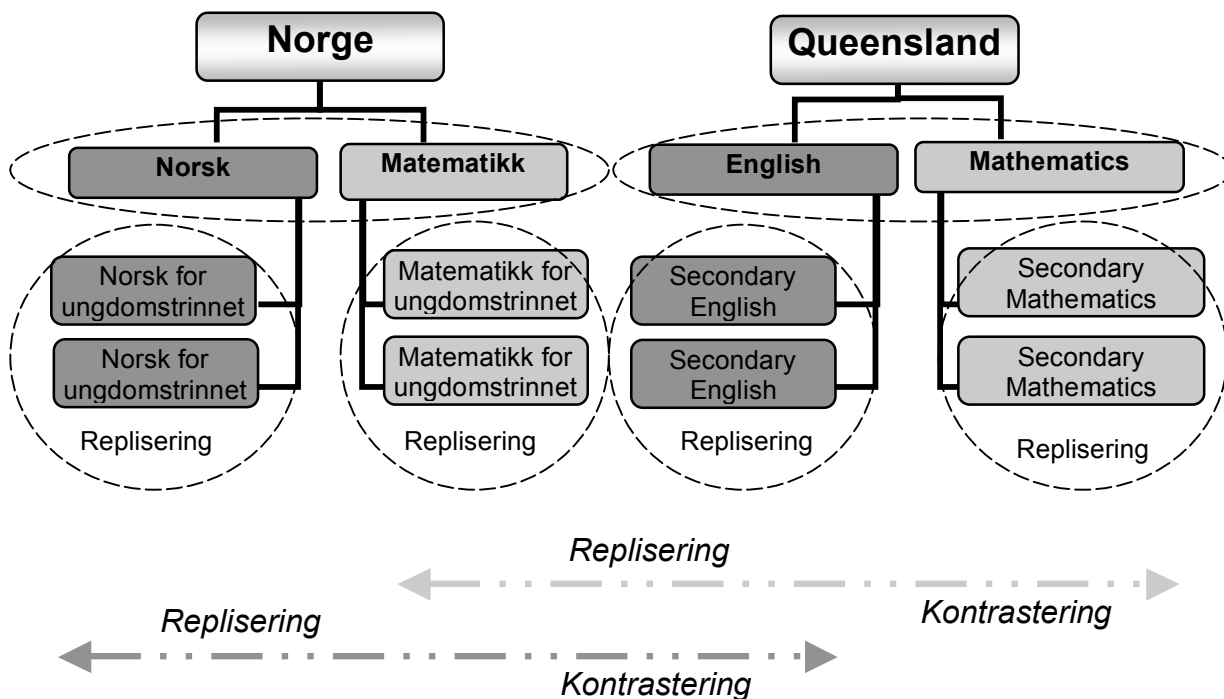
Design

Med utgangspunkt i Yins forståelse av analysenheter kan Norge og Queensland kalles *hovedenheter*. Hver hovedenhet inneholder flere case. Hver students erfaringer kan sees på som single case, som utgjør to underenheter for hver hovedenhet (Yin, 2003, s. 26). I Norge: *Norsk* og *Matematikk*. I Queensland: *English* og *Mathematics*. Et minstekrav for å kunne gjennomføre valideringsprosedyrene Yin foreslår, er å ha to informanter med samme bakgrunn. Jeg valgte å intervju til sammen åtte informanter, slik at jeg kunne sammenligne fire studenters utsagn fra hvert land, og to fra hvert fag i hvert land der det er aktuelt å trekke fram likheter og forskjeller knyttet til fagtyper. En kontrasterende forestilling for hovedcasene Norge og Queensland, er antagelsen om at nyutdannede lærere i Norge opplever å ha fått mindre opplæring i elevvurdering enn nyutdannede lærere i Queensland. Grunnlaget for

denne forhåndsantagelsen har jeg beskrevet med utgangspunkt i offentlige dokumenter i kapittel 2.

Figur 1 illustrerer de ulike casene som er med i studien, og mellom hvilke case jeg har benyttet repliserings- og kontrasteringsstrategier.

Figur 1: Illustrasjon av forskningsdesign med repliserings- og kontrasteringsstrategier



En tentativ definisjon av analyseenheter (casene) bør utledes fra de initielle forskningsspørsmålene (Yin, 2003, s. 26). Ved utledning av analyseenheter bør spesifikke tidsbindinger definere starten og slutten på casene (Yin, 2003, s. 40), hvilket legger føringer for analysen av offentlige dokumenter og annen litteratur relevant for casestudien. Studiens omfang tillater ikke en uttømmende dokumentanalyse. Jeg valgt å primært konsentrere meg om offentlige dokumenter om elevvurdering og lærerutdanningene i Norge og Queensland som har hatt betydning for utdanningsløpet mine informanter gjennomførte fra 2003 til 2007. For elevvurderingsfeltet har jeg undersøkt norske stortingsmeldinger og læreplaner i begge land. Dokumentanalysen knyttet til lærerutdanning har utgangspunkt i offisielle styringsdokumenter og flertallsbeslutninger.

- I Norge: Stortingsvedtak, stortingsmeldinger, lover, forskrifter, rammeplan for allmennlærerutdanningen, høyskolens program for allmennlærerutdanningen og tilhørende fagplaner og relevant pensum.

- I Queensland: Akkrediteringsretningslinjer for lærerutdanninger, lærerregistreringsprosedyrer, «profesjonelle standarder» for lærere, universitetets programplaner for utdanning av *secondary*-lærere med tilhørende fagplaner og relevant pensum.

4.5.3 Forskningsspørsmål

Forskningsdesignet danner rammeverket for utviklingen av forskningsspørsmål. Når man utforsker forskningsspørsmålene er det viktig å forstå at «the *verbal* line of inquiry is different from the *mental* line of inquiry» (Yin, 2003, s. 75). Yin lister opp fem *nivåer* av spørsmål:

Level 1: Questions asked of specific interviewees

Level 2: Questions asked of the individual case

Level 3: Questions asked of the pattern of findings across multiple cases

Level 4: Questions asked of an entire study

Level 5: Normative questions about policy recommendations and conclusions (Yin, 2003, s. 74).

Nivå 1-spørsmål hører til den verbale linjen av spørsmål som blir stilt til informantene i studien. For å sikre pålitelighet i studien er det imidlertid viktig å ikke bare undersøke den verbale linjen av forskningsspørsmål eksplisitt. Nivå 2-spørsmål er spørsmål stilt til det individuelle caset som ikke er eksplisitt stilt til informantene. Disse spørsmålene bør, som spørsmål på kryss av casene (nivå 3) og overordnede spørsmål i studien (nivå 4), også drøftes, slik at potensielle replikasjoner av studien kan bli gjennomført innenfor det samme rammeverket for å undersøke påliteligheten i funnene (Yin, 2003, s. 74).

Forskingsspørsmålene for studien ble utarbeidet, diskutert og justert i samarbeid med veileder flere ganger. I problemstillingen og spørsmålene jeg endte opp med å bruke, presentert i kapittel 1, har jeg lagt vekt på å få fram at studien omfatter studenter som avsluttet utdanningen i 2007 (jf. kapittel 1.2). Problemstillingen kan sees på som det Yin kaller et nivå 4-spørsmål, som ligger til grunn for hele studien. Ved å bruke preteritumsform i spørsmålene, har jeg forsøkt å tydeliggjøre at det ikke er meningen at man skal generalisere til generell praksis utover de aktuelle institusjonene og studentene. Forskingsspørsmålene kan sees på det Yin kaller nivå 2 og 3-spørsmål, som er rettet mot casene, men er ikke stilt på samme måten til informantene.

Hele denne studien hadde sitt utgangspunkt i et nivå 5-spørsmål, nemlig målet om å utforske en lærerutdanningsinstitusjon i et annet land der jeg hadde en formening om at studentene gis mer og bedre opplæring i elevvurdering enn det norske allmennlærerstudenter

får. Jeg har dessuten tidligere selv kommet med omfattende kritikk av retningslinjer og praksis for elevvurdering i Norge. Ettersom jeg hadde slike normative oppfatninger ved starten av studien, var det svært viktig at jeg maktet å legge disse oppfatningene til side da jeg samlet inn data, for å begrense muligheten for at undersøkelsen i stor grad ble påvirket av mine subjektive oppfatninger.

4.6 Utarbeiding av intervjuguide

I midten av mai 2007 reiste jeg hjem til Norge for å ha møter med veileder og planlegge de norske intervjuene til studien. Basert på tidligere utformet skisse ble intervjuguiden ferdigstilt i samarbeid med veileder. Intervjuguidene som ble brukt inneholdt forholdsvis strengt strukturerte spørsmål. Dette var nødvendig for at svarene skulle være mulig å komparere. Intervjuguiden hadde først noen innledende bakgrunnsspørsmål (kjønn og alder). Selve intervjuet var delt i fire deler:

Del 1 omhandlet spørsmål om hvordan opplæring og praksis i undervisningsfagene var organisert i studentenes lærerutdanning. Hensikten med disse spørsmålene var å finne ut hvor *mye* teori- og praksisopplæring studentene hadde fått i undervisningsfaget, og hvordan denne opplæringen var organisert. Delen inneholdt også noen andre spørsmål om studentenes utdannelsesbakgrunn.

Del 2 omhandlet spørsmål om hvordan innholdet knyttet til elevvurdering var organisert i lærerutdanningen, og i hvilken grad elevvurdering ble vektlagt i undervisningsfaget. Hensikten med disse spørsmålene var å finne ut *hvordan* og i hvilket *omfang* studentene møtte temaet elevvurdering.

Del 3 omhandlet spørsmål om i hvilken grad opplæringen på lærerutdanningen, praksis som en del av lærerutdanningen og annen arbeidserfaring hadde bidratt til at studentene fikk kunnskaper og erfaringer med elevvurdering. Her ble det stilt syv spørsmål knyttet til elevvurdering som studentene ble bedt om å svare på innenfor kategorier som skilte mellom opplæring på lærerutdanningen, gjennom praksis og annen arbeidserfaring for undervisningsfaget og for andre fag. De ble også spurt om hva de hadde lært «på andre måter». Hensikten med denne oppdelingen var å skape bevissthet hos studentene om hvor de hadde lært hva, slik at studien så presist som mulig kunne avdekke hva som hadde vært de viktigste faktorene for disse studentenes kunnskaper og erfaringer om elevvurdering og/eller mangel på dette.

Del 4 omhandlet studentenes synspunkter på og erfaringer med elevvurdering ved avslutningen lærerutdanningen. Hensikten med disse spørsmålene var å få et innblikk i

studentenes holdninger og synspunkter knyttet til elevvurdering, slik de kom til uttrykk etter avsluttet lærerutdanning. Kopi av intervjuguidene for Norge og Queensland ligger vedlagt (henholdsvis appendiks 11 og 12).

4.7 Utvalg og samplingskriterier

Følgene utvalgs-kriterier ble utarbeidet for studien:

- 2 nyutdannede lærere i Norge som har hatt faget *Norsk for ungdomstrinnet*
- 2 nyutdannede lærere i Queensland som har valgt *secondary English*
- 2 nyutdannede lærere i Norge som har hatt faget *Matematikk for ungdomstrinnet*
- 2 nyutdannede lærere i Queensland som har valgt *secondary Mathematics*

Det ble også tilstrebet å finne studenter av begge kjønn i begge fag, og at studentene var under 25 år og således ikke hadde noen vesentlig annen utdanning. Intervjuene ble gjort på slutten av studentens siste studieår, når de hadde fått all undervisning, tatt alle eksamener og tilnærmevis var ferdig utdannede lærere. Dette innebar at intervjuene i Norge måtte gjennomføres i juni måned, mens de i Queensland måtte gjennomføres i november/desember.

4.8 Intervjuene i Norge

Sampling av studenter for den norske delen av studien ble forberedt over telefon og e-post fra Australia. En norsk høyskole på det sentrale østlandsområdet som utdannet allmennlærere, ble kontaktet med spørsmål om hjelp til å sample studenter. Det var ikke nødvendig å få noen særskilt tillatelse fra høyskolen ut over den som allerede var gitt fra NSD. En av høyskolelektorene satte meg i kontakt med en tidligere studentrepresentant, som angivelig ville være den beste veien til å finne egnede intervjukandidater. Denne studenten, som selv passet til samplingskriteriene, sa seg villig til å spørre om noen av hans studentkolleger kunne tenke seg å stille opp til en times intervju. Dette ga imidlertid ikke resultater, jeg fikk derfor hjelp av en IT-ansvarlig ved høyskolen til å sample aktuelle intervjukandidater.

4.8.1 Utvalg og sampling

Datamaterialet jeg ble forelagt inneholdt en tabell med fire navn og en annen tabell med seks navn på studenter som avsluttet lærerutdanningen våren 2007 og som hadde valgt fordypning i henholdsvis *Matematikk for ungdomstrinnet* og *Norsk for ungdomstrinnet* i det siste semesteret. Samplingsmaterialet for førstnevnte inneholdt to kvinner og to menn, mens det for sistnevnte inneholdt fem kvinner og en mann.

Prosedyren som ble benyttet impliserte at studentene som ble presentert som aktuelle for min studie hadde det aktuelle undervisningsfaget i det *siste* semesteret, selv om dette ikke var et kriterium for samplingsprosedyren. Det viste seg dessverre at det ikke fantes noen oversikt på institusjonen som gjorde det mulig å koble de to kriteriene a) avslutter våren 2007 med b) har hatt [norsk/matematikk] som undervisningsfag, uavhengig av når denne fordypningen ble gjennomført. Hadde dette vært mulig er det grunn til å tro at utvalgsgrunlaget ville vært vesentlig større. På den annen side kan det sees på som gunstig at de aktuelle kandidatene hadde deltatt i den samme undervisningen.

I begynnelsen av juni ble fire potensielle intervjukandidater kontaktet på telefon. Alle sa seg villig til å bli intervjuet, og jeg avtalte intervjuer med dem enkeltvis fordelt på dagene 20. og 21. juni i 2007. De hadde på dette tidspunktet gjort seg ferdig med alle eksamener og var – med forbehold om at de bestod disse eksamenene – offisielt ferdigutdannede lærere i tråd med samplingskriteriene. Det har ikke vært aktuelt å undersøke hvorvidt noen mot formodning strøk disse eksamenene, da det ikke ville være aktuelt å gjennomføre nye intervjuer av den grunn. De fire intervjukandidatene fra Norge har blitt tildelt følgende pseudonymer:

- Camilla (C): *Norsk for ungdomstrinnet*, Norge (kvinne f. 1981)
- Trygve (T): *Norsk for ungdomstrinnet*, Norge (mann f. 1982)
- Olav (O): *Matematikk for ungdomstrinnet*, Norge (mann f. 1976)
- Eli (E): *Matematikk for ungdomstrinnet*, Norge (kvinne f. 1977)

Kjønnsfordelingen møtte kriteriene som ble tilstrebet. De to med fordypning i *Norsk for ungdomstrinnet* hadde alderen som ble tilstrebet (ca 25 år), mens de to med fordypning i *Matematikk for ungdomstrinnet* var fem år eldre enn tilstrebet. Dette var uunngåelig som følge av det svake utvalgsgrunlaget.

4.8.2 Pilotering

For å foreberede meg på intervjusituasjonen forsøkte jeg meg fram med båndopptaker og spørsmålsstillinger i den tidlige fasen av arbeidet med intervjuguiden. Da denne var klar, tok jeg kontakt med studentrepresentanten jeg hadde hatt kontakt med tidligere, som også møtte samplingskriteriene. Han sa seg villig til å gjennomføre en pilotering av intervjuguiden, som vi gjennomførte over telefon. Resultatet av piloteringen var noen begrepstilpasninger og strukturendringer (spørsmålsrekkefølge). I hovedsak fungerte intervjuguiden etter hensikten.

4.8.3 Intervjusituasjonen

Intervjuene ble gjennomført 20. og 21. juni som planlagt og forløp stort sett uten problemer. Ved en feiltagelse ble Trygve ikke stilt spørsmål 16, om kunnskaper og erfaringer med å sette karakter på elevers arbeid. Kun i forbindelse med ett spørsmål, spørsmål 13 om verktøy og hjelpemidler for elevvurdering, opplevde jeg behov for å presisere spørsmålet overfor flere av informantene. Intervjuet med Eli trekkes her fram for å illustrere hvordan dette utspant seg i praksis:

ST: Da vil jeg spørre mer konkret om hva slags verktøy og hjelpemidler til vurderingsarbeidet har du lært å bruke? Da tar vi først *Matematikk* i den samme oversikten. Er det noen spesielle verktøy eller hjelpemidler som du har lært å bruke?

E: Tenker du da på type tester og sånn, eller?

ST: Ja, hva som helst av ting du kan bruke, som du har lært deg å bruke til vurderingsarbeidet rett og slett.

E: Ja, vi har hatt litt om diagnostiske prøver, for eksempel.

ST: På lærerutdanninga?

E: Ja. Ja, det er egentlig det jeg kommer på sånn først.

ST: Ja.

Samtalen fortsatte deretter om *Engelsk*, der Eli fortalte at hun hadde lært å bruke såkalte «grids» (Eli-intervju, s. 8). Ettersom verktøy og hjelpemidler for elevvurdering i Norge er lite utviklet (jf. kapittel 3.4.3) er det ikke så rart at studentene hadde problemer med å forstå spørsmålsstillingen. I ettertid ser jeg at jeg burde ha forberedt dette spørsmålet bedre, og vurdert å benytte flere begreper for å forklare spørsmålsstillingen. På bakgrunn av disse metodiske problemene har jeg tatt forbehold ved konklusjonene jeg har trukket på bakgrunn av studentenes svar på dette spørsmålet.

4.9 Intervjuene i Queensland

I september og oktober ble mye tid brukt på å innhente de nødvendige tillatelser til å gjennomføre intervjuer i Queensland. Som en del av forberedelsene til den australske

masteroppgaven hadde jeg lagt ned mye arbeid med tillatelser, som bidro til å gjøre arbeidet med søknaden til den norske oppgaven vesentlig enklere. Det viste seg dessuten at jeg ved å skrive en oppgave ved en annen institusjon og i et annet land enn der jeg skulle gjennomføre intervjuer, faktisk hadde en enklere prosedyre for å få de nødvendige tillatelsene. Det var kun nødvendig med tillatelse fra det aktuelle universitetet.

4.9.1 Utvalg og sampling

Sampling av intervjukandidater ble gjennomført ved hjelp av en universitetslektor ved den aktuelle lærerutdanningsinstitusjonen. Jeg fikk en liste med 23 studenter med fordypning i *English*, hvorav 22 var kvinner og bare én mann (!). 20 av disse var under 25 år. Listen over fordypning i *Mathematics* inneholdt 11 studenter, hvorav tre var kvinner og åtte menn. Seks av disse var under 25 år, hvorav bare én kvinne. Utvalgsgrunnlaget var likevel bedre enn for den norske studien. Det viste seg å være umulig å sample studenter som hadde fått opplæring i undervisningsfagene i siste semesteret. Dette skyldes at lærerutdanningen ved institusjonen ikke la opp til opplæring i undervisningsfag, men i stedet en rekke andre temafag samt praksis de to siste semestrene.

Universitetets regler tilsa at universitetslektoren måtte opprette kontakt med studentene først, før han etter samtykke kunne gi meg studentenes telefonnummer. To studenter sa ja med en gang, mens to ikke hadde mulighet fordi de var henholdsvis bortreist og opptatt med jobb. De to neste som ble kontaktet sa imidlertid ja, og dermed hadde jeg de fire kandidatene som trengtes. De fire intervjukandidatene fra Queensland har blitt tildelt pseudonymene:

- Amy (A): *English*, Queensland (kvinne f. 1983)
- Royce (R): *English*, Queensland (mann f. 1986)
- Philip (P): *Mathematics*, Queensland (mann f. 1980)
- Damion (D): *Mathematics*, Queensland (mann f. 1984)

Kjønnsfordelingen møtte kriteriene som ble tilstrebet for *English*, men i *Mathematics* var det dessverre ikke mulig å finne noen kvinne til studien til tross for at kjønnsbalansen ble gitt prioritet foran alder. Årsaken var kombinasjonen av få kvinner med denne fordypningen (bare tre) og at disse enten ikke var mulig å få tak i eller ikke kunne stille opp. Alderssammensetningen er imidlertid i motsetning til den norske delen av studien helt i henhold til kriteriene.

4.9.2 Pilotering

I november ble intervjuguiden oversatt til engelsk og sendt til veileder for tilbakemelding. På bakgrunn av tilbakemeldingene ble guiden språklig forbedret. Den ble deretter forbedret ytterligere, først ved hjelp av en profesjonell engelsklærer med bakgrunn fra undervisning i Queenslands skoler. Deretter ble begreper forbedret ved hjelp av en lærerstudent med engelsk som morsmål (Canadisk), som studerte ved samme universitet som intervjukandidatene. Intervjuguiden ble deretter sendt på ny til veileder for kontroll og noen få nye justeringer ble gjort. Denne omfattende oversettelsesprosedyren ble brukt for å kvalitetssikre begrepsforståelse på tvers av de to casene, for å minimalisere faren for feilkilder som følge av ulike forståelse av spørsmålene mellom alle intervjukandidater, og spesielt på tvers av Norge og Queensland. Det ble altså lagt ned mye arbeid for å sikre valid begrepsbruk på tvers av casene.

4.9.3 Intervjusituasjonen

De fire intervjuene ble gjennomført over en drøy uke, fra 30. november til 7. desember, et par uker etter siste semester var avsluttet. Tidspunktene for intervjuene var altså tilnærmedesvis tilsvarende den norske delen av studien. Også i Queensland hadde studentene problemer med å forstå spørsmål 13. Her var det trolig i større grad begrepsmessige problemer som gjorde seg gjeldende. På grunnlag av pilotguiden og litterære kilder ble begrepene «hjelpemidler og verktøy» oversatt med «instruments and other resources». Begrepet «instruments» var imidlertid uklart for flere av studentene, og i ettertid er det lett å se at jeg også burde ha inkludert «tools» i spørsmålsstillingen. Spørsmålet ble, som i Norge, forsøkt presisert med stikkord som «rubrics», «specific tests». Likevel knyttet flere av studentene dette spørsmålet nærmere opp mot IKT (trolig på grunn av begrepet «instruments»). Som for de norske studentene har jeg i analysene og drøftningene tatt forbehold om disse metodiske problemene.

4.10 Transkribering av data

Jeg valgte å transkribere intervjuene med fokus på innholdet i intervjuene, og med et forholdsvis lavt lingvistisk presisjonsnivå. Der kandidaten lo eller tok lange tenkepauser noterte jeg dette i transkripsjonene, men jeg lagde ingen oversikt over hvor mange sekunder pauser varte. Jeg noterte heller ikke nøling, hosting eller annen ikke-verbal kommunikasjon, da jeg ikke så på dette som relevant for denne typen studie.

4.11 Analyse av data

Som nevnt tidligere hadde intervjuguiden fire hoveddeler i tillegg til en del om studentenes bakgrunn. Transkriberingene viste at noen justeringer av spørsmålsgrupperingene kunne være hensiktsmessig. Spørsmål 6 (Hvor mye lærererfaring har du hatt utenom lærerutdanningen?) og 7 (Har du noen annen høyere utdanning?) tilhører essensielt bakgrunnskategoriene. De ble derfor flyttet sammen med spørsmål 1 (fødselsår) og 2 (kjønn). Her ble også spørsmål 25 (Kan jeg få lov til å ta kontakt med deg senere ved en eventuell oppfølgingsundersøkelse?) plassert. Til sammen kan dette kalles *bakgrunnsinformasjon*.

4.11.1 Bakgrunnsinformasjon

Bakgrunnsinformasjon om kjønn og alder var med for å kartlegge i hvilken grad man nådde tilstrebningskriteriene for utvalget. Ettersom dette er en kvalitativ undersøkelse med bare åtte informanter, vil man ikke kunne peke på sammenhenger mellom informantenes bakgrunn og responser. Bakgrunns spørsmålene om annen arbeidserfaring og annen høyere utdanning ble inkludert fordi de, dersom et tydelig mønster avtegnet seg, kunne gi grunnlag for generering av noen interessante hypoteser. Det gjorde det imidlertid ikke.

4.11.2 Organisering av opplæring og praksis i undervisningsfagene

Intervjuguiden inneholdt følgende spørsmål om organisering av opplæring og praksis i undervisningsfagene:

- 3) Hvilke undervisningsfag har du hatt på lærerutdanningen og hvilket omfang hadde opplæringen i disse fagene?
- 4) Hvor mye praksis har du hatt som del av lærerutdanningen?
- 5) Hvor mye praksis har du hatt i de ulike undervisningsfagene?

Hensikten med disse spørsmålene var å finne ut hvor *mye* teori- og praksisopplæring studentene hadde fått i undervisningsfagene.

4.11.3 Organisering av innhold knyttet til elevvurdering i lærerutdanningen

Intervjuguiden inneholdt følgende spørsmål om organisering av innhold knyttet til elevvurdering i lærerutdanningen:

- 8) Har du hatt et eget fag på lærerutdanningen som dreide seg spesielt om elevvurdering?

-
- 9) Ble emnet elevvurdering vektlagt i matematikk?
 - 10) Har du hatt emner under andre fag som dreide seg spesielt om elevvurdering?
 - 11) Hva slags pensumlitteratur eller tilleggslitteratur har du hatt som dreide seg om elevvurdering? Nevn navn på bøker eller forfattere, om mulig.

Hensikten med disse spørsmålene var å finne ut *hvordan* og i hvilket *omfang* studentene hadde møtt temaet elevvurdering.

4.11.4 Tilegnelse av kunnskaper om og erfaringer med elevvurdering

Intervjuguiden inneholdt følgende spørsmål om tilegnelse av kunnskaper om og erfaringer med elevvurdering:

- 12) På hvilke arenaer mener du at du har lært mye om elevvurdering og hva var det viktigste du lærte?
- 13) Hva slags verktøy og hjelpemidler til vurderingsarbeidet har du lært å bruke?
- 14) Hvilke kunnskaper har du fått om bruk av kriterier for elevenes oppnåelse av målene i læreplanen? (stikkord om nødvendig: kriterier for vurdering av nasjonale prøver, prøver utviklet for skolen eller distriktet, sentralt gitte eksamener eller lignende)
- 15) På hvilke måter har du med andre lærere eller studenter diskutert bruk av kriterier for vurdering av elevenes oppnåelse av målene i læreplanen?
- 16) Fortell om dine kunnskaper om og erfaringer med å sette karakter på elevers arbeid.
- 17) Fortell om dine kunnskaper om og erfaringer med å gi elever tilbakemelding på læringsarbeidet. Skill mellom muntlige og skriftlige tilbakemeldinger.
- 18) Hva slags opplæring har du fått i forskrifter og retningslinjer for elevvurdering gjennom lærerutdanningen?

Ved å stille disse spørsmålene forsøkte jeg å finne ut i hvilken grad opplæringen på lærerutdanningen, praksis som en del av lærerutdanningen og annen arbeidserfaring hadde bidratt til at studentene fikk kunnskaper og erfaringer med elevvurdering. Å skape bevissthet hos informantene og gjøre deres svar mer distinkte var som nevnt en av grunnene til at å be studentene svare innenfor kategorier for spørsmålene 12 til 18. Følgende kategorioversikt ble

utdelt på ark til informantene, og de ble bedt om å svare punktvis innenfor spørsmålskategoriene som var relevante for dem:

- A) Gjennom opplæring i matematikk på lærerutdanningen?
- B) Gjennom praksis i matematikk som en del av lærerutdanningen?
- C) Gjennom arbeidserfaring i matematikk utenom lærerutdanningen?

- D) Gjennom opplæring i andre fag på lærerutdanningen?
- E) Gjennom praksis i andre fag som en del av lærerutdanningen?
- F) Gjennom arbeidserfaring i andre fag utenom lærerutdanningen?

- G) På andre måter?

På denne måten ønsket jeg å forsikre meg om at studentene skilte klart mellom hva de har lært gjennom opplæring på lærerutdanningen, gjennom praksis på lærerutdanningen og gjennom annen arbeidserfaring i faget de i utgangspunktet ble intervjuet om, og tilsvarende for andre fag, samt på andre måter. Hensikten var å sørge for at informantene ga uttømmende forklaringer, i riktig kategori.

Analysene viste at studentene hadde gitt mye informasjon om hva de hadde lært på lærerutdanningen og gjennom praksis. Det var imidlertid bare to av studentene som hadde annen arbeidserfaring. Det var også svært få som svarte på kategorien «på andre måter». Følgelig er svarene innenfor spørsmålskategoriene knyttet til annen arbeidserfaring utelatt i analysen, med mindre svarene essensielt hørte til en annen spørsmålskategori. Selv om datamaterialet ikke har kommet til nytte i analysen og drøftningen, var spørsmålene viktige for å skape bevissthet hos informantene om hva slags bakgrunn de har og hvor de har lært hva, slik at svarene ble så distinkte som mulig. For studentene som hadde svart at de ikke hadde noen annen arbeidserfaring på spørsmål seks ble punkt C og F strøket over alt før skjemaet ble presentert for studenten. I analysedelen rapporteres det altså bare fra spørsmålene A, B, D og E, med mindre svarene i de andre kategoriene skapte fortolkningsproblemer.

4.11.5 Synspunkter og erfaringer

Avslutningsvis ble kandidatene stilt følgende spørsmål om synspunkter og erfaringer på noen utvalgte problemstillinger:

-
- 19) Hva mener du er viktig når man gir tilbakemeldinger for at disse tilbakemeldingene skal bidra til mer læring?
 - 20) Hvilke tanker gjør du deg om å sette en samlet fagkarakter for hver elev?
 - 21) Hva tror du vil være dine bekymringer når du setter karakterer?
 - 22) Hvordan kan du sikre at din karaktersetning blir så rettferdig som mulig?
 - 23) Mange lærere gir uttrykk for følgende: Det er vanskelig å forholde seg til krav om felles kriterier for karaktervurdering, samtidig som opplæringen skal tilpasses den enkelte elev. Hvilke tanker gjør du deg om dette?
 - 24) Hva skulle du ønske at du hadde lært mer om i lærerutdanningen knyttet til elevvurdering?

Hensikten med disse spørsmålene var å få et innblikk i studentenes holdninger og synspunkter knyttet til elevvurdering, slik de kom til uttrykk etter avsluttet lærerutdanning. Hensikten var blant annet å undersøke hva som kjennetegner vurderingskulturen i Norge og Queensland generelt, og ikke nødvendigvis bare forankret i hva de spesifikt hadde lært gjennom lærerutdanningen.

I den endelige analysen er alle seks spørsmål i denne siste spørsmålskategorien utelatt. Det er to grunner til at jeg har nedprioritert disse dataene. For det første var for mye sprik i studentenes oppfatninger til at noen klare slutninger kunne trekkes. Dette henger trolig sammen med at spørsmålene utfordret informantene på personlige oppfatninger og mindre kontekstuelle problemstillinger knyttet til elevvurdering. Det er ikke overraskende at det var større sprik i svarene på slike spørsmål, enn på konkrete spørsmål om hva slags opplæring de hadde fått. For det andre viste det seg at omfanget av intervjudata var for stort til at jeg kunne gjøre grundige analyser av alle spørsmålene. Jeg kom derfor fram til at det var mest hensiktsmessig å prioritere en grundig analyse av dataene som inneholder klare funn, og utelate de siste spørsmålene. I drøftningsdelen har jeg likevel trukket fram noen studentutsagn som jeg mener egner seg til å illustrere viktige aspekter jeg diskuterer. Studentenes svar på det siste spørsmålet om hva de skulle ønske de hadde lært mer om i lærerutdanningen knyttet til elevvurdering, drøfter jeg i tilknytning til det aktuelle temaet – karaktervurdering.

4.12 Framstilling av data

Framstillingen av data var en treleddet prosess. Først grupperte jeg intervjudataene for hvert spørsmål i de to landene. Deretter gjorde jeg noen innledende analyser av svarene der jeg tok med alt jeg mente hadde relevans, og knyttet noen egne kommentarer til dataene. Disse

plasserte jeg i en egen kolonne i analysedokumentet. Når jeg hadde gjort dette for alle 25 spørsmålene, lagde jeg en tredje kolonne, der jeg samlet alt jeg ønsket å bruke. Grunnlaget for dette utvalget var da i hovedsak den andre kolonnen, men ved å ha alt samlet i forskjellige kolonner i det samme dokumentet, kunne jeg enkelt gå til rådataene i den første kolonnen og se om det var noe mer jeg ville ha med. Jeg kunne også lettere se om jeg hadde misforstått noen utsagn. Et eksempel på oppsettet jeg brukte for framstilling av data er vedlagt (appendiks 13). Det er altså den tredje kolonnen i disse analysedokumentene, som har vært grunnlaget når jeg har «skrevet ut» analysekapitlet. Etter justeringene av spørsmålsgruppene, omtalt ovenfor, består den endelige dataframstillingen i kapittel 5 av tre deler:

- Organisering av lærerutdanningens teori- og praksisopplæring
- I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?
- Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer

4.13 Sammen drag av metodiske vurderinger

I dette metodekapittelet har jeg først beskrevet bakgrunn for valg av fokus for oppgaven. Det har vært forsket lite både på elevvurdering og allmennlærerutdanningen i Norge, og en sammenligning med Queensland antok jeg kunne bidra til å belyse problemstillinger knyttet til både elevvurdering og organisering av lærerutdanningen bedre. Videre har jeg pekt på grunnleggende problemstillinger knyttet til validitet og reliabilitet og forankret studien i forskningsmetodene fenomenologi og hermeneutisk fortolkning. Jeg har begrunnet valg av kvalitative intervjuer som primær datainnsamlingsmetode, og drøftet prinsipper for å tilstrebe en gyldig meningsproduksjon gjennom slike intervjuer. Videre har jeg designet studien som en case studie med Norge og Queensland som hovedenheter og morsmåls- og matematikkfaget som underenheter begge steder. I designet har jeg lagt opp til to repliseringer for hvert fag begge steder.

Videre har jeg forklart hvordan jeg gjennom hvert ledd i det metodiske arbeidet har forsøkt å ta hensyn til problemstillinger knyttet til validitet og reliabilitet. Forskingsspørsmålene for studien ble utarbeidet, diskutert og justert i samarbeid med veileder flere ganger. Jeg la vekt på å få fram at studien omfatter studenter som avsluttet utdanningen i 2007, og ikke kan generaliseres utover de aktuelle institusjonene og studentene. Jeg pilorte intervjuguiden i begge land for å forsikre meg om at den fungerte etter hensikten og at informantene ville forstå begrepene. Jeg tilstrebet et bredt utvalgsgrunnlag, men av praktiske grunner hadde jeg ikke mange studenter å velge mellom noen av stedene. Dette holdes ikke for å være viktig, i og med at det er en begrenset kvalitativ undersøkelse. De overordnede

kriteriene til fagbakgrunn er ivaretatt, og tilstrebningskriteriene for alder og kjønn er også i stor møtt. Ved gjennomføringen av intervjuene benyttet jeg en strategi omtalt som «tilsiktet naivitet» for å unngå at mine oppfatninger infiltrerte informantenes meningsproduksjon. Jeg benyttet en transkriberingsstil med lav lingvistisk presisjon, ettersom dette ikke ble sett på som viktig for denne typen studier.

I analysen grupperte jeg spørsmålene litt annerledes for å skape bedre oversikt. Jeg prioriterte bort spørsmålgruppen om studentenes synspunkter og erfaringer, av hensyn til oppgavens omfang og fordi dette datamaterialet hadde få interessante funn. I framstillingen av data samlet jeg alle studentenes svar i ett dokument for hvert spørsmål. Jeg lot rådataene stå i en egen kolonne, mens jeg trakk ut interessante utsagn i en andre. I en tredje kolonne satt jeg dette sammen og analyserte likheter og forskjeller mellom Norge og Queensland. Slik kunne jeg enkelt gå til rådataene i den første kolonnen og se om det var noe mer jeg ville hatt med. Ved å legge ved eksempel på denne analyseteknikken, ved siden av intervjuguiden for begge stedene, har jeg gjort studien så transparent som mulig. Alle transkripsjoner er anonymisert og lagret hos NSD. Det viktigste hensynet til reliabilitet er følgelig ivaretatt. Dersom noen ønsker å kontrollere analysene eller gjennomføre en replikasjon av undersøkelsen, kan de søke om tilgang på alt rådata fra NSD. Etter justeringen av spørsmålsgrupper består den endelige dataframstillingen i neste kapittel av tre deler: *Organisering av lærerutdanningens teori- og praksisopplæring*; *I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?*; og *Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer*.

5 Hva slags opplæring i elevvurdering hadde lærerstudentene i Norge og Queensland fått?

I dette kapitlet har jeg analysert de åtte intervjuene som ble gjennomført med nyutdannede lærere i Norge og Queensland. Analysen er delt i tre hoveddeler, i henhold til forskningsspørsmålene: betydningen av teori- og praksisopplæringens organisering for opplæring i elevvurdering; i hvilket omfang og på hvilken måte studentene møtte temaet elevvurdering; og hva slags teori- og praksisopplæring studentene hadde fått i bruk av verktøy og hjelpemidler for elevvurdering, bruk av vurderingskriterier, diskusjoner om vurderingskriterier, karaktersetting, å gi faglige tilbakemeldinger og forskrifter og retningslinjer for elevvurdering.

5.1 Organisering av lærerutdanningens teori- og praksisopplæring

Det var betydelige forskjeller mellom de to lærerutdanningenes organisering av teori- og praksisopplæringen generelt. I det følgende analyseres først organiseringen av teoriopplæringen, etterfulgt av organiseringen av praksisopplæringen, før jeg drøfter noen erfaringer med praksis og de kvantitative forskjellene i opplæringen som ble gitt.

5.1.1 Organisering av teoriopplæringen

Som forklart i introduksjonskapitlet er alle fag i den norske allmennlærerutdanningen obligatoriske de to første årene. Man får opplæring i *Pedagogikk*, *KRL*, *Grunnleggende Lesing*, *Skriving* og *Matematikkopplæring (GLSM)*, *Matematikk* og *Norsk*. I de to siste årene har studentene frihet til å velge fag innenfor retningslinjene i rammeplanen (jf. kapittel 2.1.3). Dette gir studentene frihet til å velge i hvilken grad de vil spesialisere seg. I tabell 6 har jeg trukket fram to eksempler på studenter som hadde valgt en henholdsvis bred og spesialisert fagsammensetning. Trygve hadde valgt *Pedagogikk 2*, *Kroppsoving* og *Samfunnsfag* i tillegg til *Norsk for ungdomstrinnet*, mens Eli hadde valgt maksimal fordypning i *Matematikk* og *Engelsk* (for både barne- og ungdomstrinnet).

Secondary-programmet ved den aktuelle lærerutdanningen i Queensland består av en rekke forskjellige kombinasjoner av *to* undervisningsfag, som man velger ved starten av utdanningen. Med unntak av et 7,5 studiepoengs valgfag i siste semester, består hele graden av obligatoriske fag. De første to årene består utelukkende av *Content studies*, mens de to siste årene består av *Curriculum studies* som inneholder fagdidaktikk (*Curriculum A* og *B*) i de to fagene, og en rekke andre temaer i tilknytning til undervisningsfagene eller lærerrollen

generelt. Som valgfritt emne kan studentene velge mellom en liste på 14 emner i blant annet språk, musikk, IKT og rådgivning. Det går også an å velge 20 dagers *internship* – ekstra ulønnet praksis – som valgfritt emne. Tabell 7 viser hvordan opplæringen til Amy var organisert.

Tabell 6: Eksempler på valgt fagsammensetning ved norsk allmennlærerutdanning

Trygve – Norsk				Eli – Matematikk			
Avdeling 1 obligatorisk	Pedagogikk			Avdeling 1 obligatorisk	Pedagogikk		
	KRL	GLSM			KRL	GLSM	
	Matematikk				Matematikk		
	Norsk				Norsk		
Avdeling 2 fordypning	Pedagogikk 2			Avdeling 2 fordypning	Engelsk 1		
	Kroppsoving				Engelsk 2		
	Samfunnsfag				Matematikk for barnetrinnet		
	Norsk for ungdomstrinnet				Matematikk for ungdomstrinnet		
Poeng	10	10	10	Poeng	10	10	10

Tabell 7: Eksempler på lærerutdanningsprogram for English og Social Science/History

Amy – English og Social Science/History				
Teaching area content	English		Social Science/History	
	English		Social Science/History	
	English		Social Science/History	
	English		Social Science/History	
Curriculum studies	English Curriculum A	Social Science Curriculum A	Teaching Strategies	Professional practice 1
	English Curriculum B	Social Science Curriculum B	Educational Psychology	Professional practice 2
	Assessing Student Learning	Teaching to difference	The Teacher Researcher	Professional practice 3
	Education as social construction	Philosophy and values in Education	Work, education and vocational education	Elective
Poeng	7,5	7,5	7,5	7,5

5.1.2 Organisering av praksis

I lærerutdanningen i Norge er det lagt opp til at det til sammen skal være 20 uker med praksis. Disse inngår som en del av opplæringen gjennom alle fire undervisningsårene, og ikke som selvstendige studiepoengenheter. Olav ga denne oversikten over hvordan hans praksis hadde blitt fordelt over de fire årene:

- Første klasse: seks uker (2+4)
- Andre klasse: seks uker (4+2)
- Tredje klasse: seks uker (2+4)
- Fjerde klasse: to uker (Olav-intervju, s. 3)

Som det framgår av tabell 7 er praksisperiodene i Queensland fordelt på tre studiepoengenheter, hver på 7,5 studiepoeng, i 5, 6, og 7. semester. Hver av disse periodene er på fem uker. I praksisperiodene i både Norge og Queensland legges det i hovedsak opp til at lærerstudentene hjelper til eller overtar en øvingslærers undervisningsansvar i praksisperioden.

Tre av de fire norske studentene fortalte om en ordning med *skoleovertakelse*, som de blant annet var igjennom det siste året. Denne ordningen gikk ut på at en gruppe lærerstudenter dro til en skole og hjalp til i en uke, før de i den påfølgende uken overtok hele skolen, mens lærerne og skoleadministrasjonen var på kurs. Studentene i undersøkelsen hadde ulike erfaringer med denne ordningen. Camilla ga uttrykk for at hun ikke samarbeidet med øvingslærer noe særlig i praksisperiodene de to siste årene, fordi de hadde hatt skoleovertakelse. Dette hadde ført til at hun fikk prøve seg mye alene, noe hun opplevde som verdifullt. Olav, derimot, ble plassert i skoleadministrasjonen under en av skoleovertagelsene og hadde derfor ikke fått undervisningserfaring den uken. Studentene i Queensland hadde ikke hatt noen ordning med skoleovertakelse. Ingen av studentene hadde benyttet muligheten til å velge et *internship* som valgfag siste semesteret, og de ga inntrykk av at det var få som gjorde dette.

5.1.3 Studentenes erfaringer med praksis i undervisningsfagene

I hvilket omfang fikk lærerstudentene praksis i undervisningsfagene de fordypet seg i? De norske studentene ga uttrykk for at de hadde praksis i en rekke fag. Hvor mye praksis de fikk i de to siste årenes undervisningsfag var relativt tilfeldig, avhengig av hvilken øvingslærer de hadde. Studentene hadde derfor vanskeligheter med å huske akkurat hvor mye praksis de har hatt i de forskjellige fagene. Dette kan illustreres med Elis svar på spørsmålet om hvor mye praksis hun hadde i *Engelsk* og *Matematikk*: «Hvis jeg tenker på alle fire årene, så er det jo naturlig at det har vært mer *Matte*, fordi de har mer *Matte* på skolen. Men vi har jo fulgt de timene som elevene har, så vi har hatt alle fag» (Eli-intervju, s. 3).

Camilla ga uttrykk for at hun hadde vært heldig og fått en øvingslærer med fag som passet henne, derfor hadde hun fått undervist mye både i *Norsk* og *Matematikk*. Trygve fortalte at han hadde hatt praksis i de fleste fag, men mest på barnetrinnet. Han fortalte at «de har prøvd å legge det til rette, men det er ikke alltid det går opp» (Trygve-intervju, s. 3). Olav fortalte at han ikke hadde hatt noe særlig praksis det siste året, som følge av at han satt i administrasjonsgruppen under skoleovertakelsen.

Lærerstudentene i Queensland ga også uttrykk for at utbyttet av praksis var litt tilfeldig, avhengig av hvilken øvingslærer de fikk. Amy fortalte at hun hadde *English* i første

praksisperiode, *Social Science* i den andre og i begge i den tredje. Hun ga uttrykk for at hun fikk undervist mye i *English* i den første perioden, men fordi hun ikke fikk plass i en *senior secondary* skole i den andre perioden, fikk hun da ikke undervist i *History*. Dette fikk hun imidlertid tatt igjen i den tredje praksisperioden: Hun forklarte hva slags oppfølging hun fikk i den siste praksisperioden: «I had one lady who I had to see, she supervised my *English* teaching. Then I had another lady who supervised my *History* teaching. It's the big one – the final one is the big one, where you take on both» (Amy-intervju, s. 7). Royce hadde undervist i *Health and Physical Education* (HPE) i første praksisperiode, *English* i andre og begge i den tredje. Han ga uttrykk for at det var slik det var ment å fungere, men at det ofte ikke gjorde det. Han oppfattet seg selv som heldig, som fikk fullt utbytte av praksisperiodene.

Philip fikk praksis i henholdsvis *Mathematics*, *Physics* og begge, som planlagt. Damion hadde undervist litt i begge fag i alle praksisperiodene, men mest i *Mathematics* i første periode, og *Physics/Science* i andre periode, som planlagt. Han fortalte at når han kom på praksisskolen «they didn't really know that it was supposed to be either Maths or Physics, and whatever teaching you got, and whatever subjects they had, is what you got, so...» (Damion-intervju, s. 3).

5.1.4 Kvantitative forskjeller i teori- og praksisopplæringen

Vi ser at det er svært store forskjeller mellom hvordan de to utdanningsløpene var organisert. I Norge fikk studentene undervisningskompetanse i minst tre fag, og de kunne, dersom de valgte bredest mulig utdanning, få formell undervisningskompetanse i hele syv fag. De kunne også valgt kun formell undervisningskompetanse i tre fag, og seksti poeng i «andre skoler relevante studier» (jf. kapittel 2.1.3). Studenter i Norge som valgte maksimal fordypning i *Norsk* eller *Matematikk* kunne få 90 studiepoeng fagfordypning i disse fagene, mens det for andre fag bare var mulig å oppnå 60 studiepoeng fagfordypning. Det bør minnes om at det kun var mulig å få 30 studiepoeng fagfordypning rettet mot ungdomstrinnet, mens de resterende fagene var rettet enten mot hele grunnskolen eller spesifikt mot barnetrinnet. I Queensland derimot, var all fagopplæring rettet mot *secondary* og begrenset til to undervisningsfag.

For å gjøre det sammenlignbart med den norske allmennlærerutdanningen, er det nødvendig å gjøre noen utregninger av det totale omfanget av opplæringen i fagene i Queensland. Studentene hadde fire emner på 15 studiepoeng i *content* (totalt 60 studiepoeng) og to emner på 7,5 studiepoeng i *curriculum* (totalt 15 studiepoeng). For å kunne sammenligne med den norske utdanningen der praksis er integrert i fagene, må

praksisenhetene i Queensland regnes med. De to første praksisperiodene var inndelt i henholdsvis *teaching area 1* og *2*, mens den siste var delt mellom de to undervisningsfagene. Studentene fikk således halvannen praksisenhet av 7,5 studiepoeng i hvert fag (totalt 11,25 studiepoeng). Til sammen fikk studentene en kompetanse på 86,25 studiepoeng i hvert av de to undervisningsfagene i løpet av hele utdanningen.

Alle studenter i Queensland fikk altså tilnærmet like mye fagkompetanse i sine to fag, som de norske studentene som hadde valgt full fagfordypning i *Norsk* og *Matematikk*. Studentene i Queensland fikk imidlertid tre ganger mer fagkompetanse rettet spesifikt mot *secondary*, enn de norske studentene maksimalt kunne få for ungdomstrinnet.

Når det gjelder andre fag enn undervisningsfagene, så utgjør pedagogikkfaget og GLSM i Norge til sammen 40 studiepoeng, mens de obligatoriske fagene i faggruppen *Professional studies* i Queensland til sammenligning tilsvarer 60 studiepoeng. 7,5 studiepoeng valgfag kommer i tillegg til dette i Queensland. Samlet sett fikk lærerstudentene i Queensland følgelig over 50 prosent mer generell pedagogisk kompetanse enn de norske lærerstudentene.

Den norske lærerutdanningen har fem uker mer praksis enn lærerutdanningen i Queensland. I Queensland ble praksis sentrert rundt øvingslærere som hadde samme undervisningsfag som dem selv, og fordelt på *junior-* og *senior secondary*. I Norge ble praksistiden fordelt mellom en rekke fag. Alle studentene fortalte om en del tilfeldigheter i tilknytning til utplassering i praksis, men det ser ut til at de norske studentene hadde vært mer utsatt for slike tilfeldigheter. Dette skyldes trolig at kabalen som skulle legges opp bestod av tre fag de to første årene, og 2-4 fag i løpet av de siste to årene, mens det i Queensland kun var praksis i to fag som skulle organiseres alle de fire årene.

I Queensland var det tilfeldigheter knyttet til antall timer og fordelingen av praksis mellom de to undervisningsfagene. Ingen fortalte om problemer med at de måtte undervise i andre fag enn undervisningsfagene deres. En sentral problemstilling var imidlertid om man fikk utplassering på ungdomstrinnet eller videregående trinn. Fikk man ikke plass på en *senior secondary* skole i en av praksisperiodene, så fikk man mindre erfaring med undervisning i spesialiseringen innenfor faget.

5.2 I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?

Hensikten med disse spørsmålene var å finne ut *hvordan* og i hvilket *omfang* studentene møtte temaet elevvurdering i lærerutdanningen. Målet var å finne ut om de hadde fått opplæring

innenfor et eget fag om elevvurdering, hvordan de møtte temaet innenfor undervisningsfagene og pedagogiske fag, og hva slags pensumlitteratur de hadde hatt om elevvurdering.

5.2.1 Eget fag om elevvurdering

De norske studentene fortalte at de ikke hadde hatt et eget fag om elevvurdering. To av dem (Camilla og Trygve) fortalte at de hadde hatt litt om det i de forskjellige fagene, som didaktikkdel i undervisningsfagene, eller i pedagogikkfaget. Olav fortalte også at temaet til en viss grad hadde vært bakt inn i de andre fagene, mens Eli kort svarte at det hadde hun ikke hatt.

Alle studentene fra Queensland hadde hatt faget *Assessing Student Learning* (7,5 studiepoeng). Hovedtema i faget var hvordan man lager gode vurderingsoppgaver. Royce fortalte at det ble lagt vekt på at man må benytte et bredt utvalg av vurderingsoppgaver, samt forskjellige teknikker og strategier for formativ og summativ vurdering. I følge Amy var hovedtemaet i faget å lage vurderingsoppgaver: «Creating pieces of assessment (...) the actual criteria sheet, how to mark them, so not only coming up with the task but also coming up with a fair and equitable way of assessing» (Amy-intervju, s. 9).

Amy og Damion ga uttrykk for at de likte faget veldig godt. Damion mente at «without it, you'd really be struggling to do assessment in schools (...) I found it relevant to what I had to do in the classroom» (Damion-intervju, s. 5). Royce ga også uttrykk for at faget var rimelig bra. Philip ga uttrykk for at deler av faget var interessant men at det ikke var så relevant for *Mathematics* og *Physics*:

The Maths and Physics side is a lot more exam based, where it's just answering questions, where a lot applied to more performance based subjects like HPE and Drama and more assignment, yes, more subjects that have assignments in them rather than straight exams (Philip-intervju, s. 5).

5.2.2 Vektlegging av elevvurdering i undervisningsfagene

Camilla og Trygve ga uttrykk for at elevvurdering var en del av fagdidaktikken i *Norsk*, først og fremst gjennom innføring i prosessorientert skriving og mappevurdering. Trygve fortalte at underveivurderingen hadde vært sentralt i dette. Det ble lagt vekt på at man bør gi konkrete og konstruktive tilbakemeldinger.

I følge Olav hadde de i *Matematikk* hatt noen få timer på slutten om elevvurdering. Da hadde de i hovedsak lært om diagnostisk testing. I følge Eli ble elevvurdering vektlagt i veldig liten grad i *Matematikk*: «Det var kanskje en og en halv time på nest siste forelesning, hvor vi snakket om elevvurdering». Temaet i forelesningen var «hvilke typer man vurderer, altså

hvorfor man vurderer og hvordan, også så vi også på noen nasjonale prøver i *Matte* og snakka om hvorfor man gjør det så dårlig» (Eli-intervju, s. 4).

Amy fortalte at elevvurdering inngikk som en viktig del av *undervisningsplanene* (*units of work*) som de må utarbeide i alle de fagdidaktiske fagene (*Curriculum A* og *B* i hvert fag). Hun hevdet at «your unit will definitely show how much you have learned about the cohesion among, between the assessment task and what you're doing, what you're teaching» (Amy-intervju, s. 11). Hun ga imidlertid uttrykk for at hun fremdeles ikke følte seg trygg på vurdering i faget, selv om de hadde lært en del om oppgaveutforming. Royce mente elevvurdering var en viktig del av fagdidaktikken i *English*, fordi mye av undervisningen var knyttet opp mot læreplanen og utarbeiding av forskjellige typer vurderingsoppgaver – i hovedsak med et summativt fokus.

Philip fortalte at de ikke lærte noe særlig om elevvurdering i *Mathematics*: «It was more looking at the syllabus and what we had to teach» (Philip-intervju, s. 4). Damion mente heller ikke han lærte så mye om elevvurdering i *Mathematics*. Han nevnte blant at de lærte å lage undervisningsplaner. På direkte spørsmål om det var en vurderingsdel i disse undervisningsplanene svarte han: «There was, but it wasn't a huge part of it. It was only a small... like we didn't have to go into any detail, it was just basically, a little small bit in the end of all your planning of units» (Damion-intervju, s. 5).

Det ser ut til at elevvurdering ble mer vektlagt i morsmålsfagene enn i matematikkfaget, både i Norge og Queensland.

5.2.3 Opplæring i elevvurdering i pedagogiske fag

Camilla fortalte at hun lærte litt om elevvurdering i valgfaget *Tilpasset opplæring*, der det blant annet ble satt et kritisk søkelys på utfordringer knyttet til vurdering mot felles standarder. Trygve fortalte ikke om noen opplæring i elevvurdering i pedagogiske fag. Olav fortalte at de hadde litt om elevvurdering i *Pedagogikk* i det første eller andre studieåret. Han mente at pedagogikkfaget «har nok tatt vare på litt av den delen, fordi vi også har spurt om det» (Olav-intervju, s. 6). Han ga uttrykk for at de ikke hadde hatt nok om dette. Eli pekte på at didaktikken skal være bakt inn i de forskjellige fagene og mente at «det er den til en viss grad, men det har vært lite fokus på det» (Eli-intervju, s. 4). Hun fortalte at de hadde jobbet litt mer med det i *Engelsk* enn i *Matematikk*.

Amy fortalte at hun i faget *Teaching and Reading* også måtte lage en undervisningsplan hvor elevvurdering inngikk, men at elevvurdering var mindre sentralt i dette faget enn i *English* og *Social Science*. Royce fortalte at det i *HPE* ble lagt stor vekt på

formative tilbakemeldinger, og han mente det kom av at det er lettere å gi tilbakemeldinger på fysiske oppgaver «because you actually see them and you can talk to them» (Royce-intervju, s. 8). Royces mente *HPE* slik skilte seg fra *English*, der han opplevde et sterkt summativ fokus.

Philip fortalte at de i *Physics* studerte elevvurdering i forbindelse med revidering av læreplanen: «Because it changed, a few of the necessary assessment tasks, that they had to do, because they were changing that, so we had a look at that» (Philip-intervju, s. 5). Damion mente vektleggingen av elevvurdering i *Science/Physics* var svært likt *Mathematics*: «We did like, how to plan a unit, and we did that, and maybe what assessment you would use, but we didn't actually go into how to assess them» (Damion-intervju, s. 6).

5.2.4 Pensumlitteratur om elevvurdering

På spørsmål om hva slags pensumlitteratur om elevvurdering hun hadde hatt, nevnte Camilla en spesifikk bok om prosessorientert skriving der det stod litt om mappevurdering. Denne boken var en del av det obligatoriske norsk- eller pedagogikkfaget et av de to første årene. Trygve kom ikke på noen spesielle bøker umiddelbart. Han nevnte en bok kalt *Fagdebatikk*, men la til at den nok dreide seg mer om fagets innhold og ikke så mye om vurdering. På spørsmål om han kunne huske noe der elevvurdering ble vektlagt spesielt, svarte Trygve: «Nei, egentlig ikke, annet enn hva lærerne har fortalt selv ut fra sine egne erfaringer, kanskje» (Trygve-intervju, s. 6).

Olav fortalte om en bok i *Matematikk* som tok opp den diagnostiske testingen. Eli fortalte at i *Matematikk* var det eneste de hadde fått et foredragshefte som foreleseren hadde laget om vurdering. Etter å ha tenkt seg litt om kom hun også på en bok av Geir Botten kalt *Meningsfylt matematikk* som hun hadde i *Matematikk* på første avdeling. «Vi hadde det mye moro med hans 'retting med grønt' (...) Han var veldig i mot at man skulle bruke *rødt* da. Så det var retting med grønt, det gjorde det mye hyggeligere», fortalte Eli og lo (Eli-intervju, s. 5). Dette uttrykket ble også trukket fram av to andre norske studenter. Trygve forklarte uttrykket på denne måten: «du skulle tvinge deg selv til 'å rette med grønt', fordi da skulle du se etter muligheter og ikke feilene (...) ut i fra totalinntrykket, ikke fra poengskalaen, så skulle man sette karakter» (Trygve-intervju, s. 5). Camilla, som hadde hatt fordypning i *Matematikk* i tillegg til *Norsk*, trakk også fram dette uttrykket fra matematikkopplæringen:

Ja, da var det veldig mye fokus på dette her med at Matte er et fag hvor det er veldig rett og galt, og at det er røde streker og grønn, altså... De la veldig mye vekt på det derre «rette med grønt» da... For at man ikke skal miste motet og... (Camilla-intervju, s. 6).

Eli nevnte også boken *Kan læring planlegges*, av Britt Ulstrup Engelsen (2006). De norske studentene fortalte at de hadde hatt noe litteratur om elevvurdering i fagene *Norsk*, *Matematikk*, *KRL* og *Pedagogikk*, men det var ikke et sentralt tema og det var ingen egne bøker om det.

Alle studentene i Queensland, med unntak av Philip, fortalte om en spesifikk bok i faget *Assessing Student Learning*. Boken, som het *Celebrating Student Achievement*, ble av disse beskrevet som nyttig. Med forbehold om at hans pensum var annerledes, er det grunn til å tro at Philip ikke var oppmerksom på, eller hadde glemt, at denne boken var på pensum. Han fortalte imidlertid om et kompendium i *Assessing Student Learning*-faget, som læreren hadde satt sammen. Damion fortalte også om et slikt kompendium, mens dette ikke ble nevnt av Royce. Amy hadde også hatt et slikt kompendium i faget. På bakgrunn av undersøkelse av fagplanen er det grunn til å tro at alle studentene hadde et slikt kompendium, selv om det også her var en av dem som ikke nevnte det.

Amy fortalte i tillegg om et kompendium i *English*, der læreren hadde satt sammen en rekke artikler, hvor noe av innholdet handlet om elevvurdering. Hun mente det også må ha vært noe om elevvurdering i *History (Social Science)*, men kom ikke på noen konkret litteratur. Royce trakk fram læreplanen som sentral litteratur i fagene og nevnte også en bok i *HPE* som han ikke husker hva het og som han aldri kjøpte. Verken Philip eller Damion fortalte om litteratur på elevvurdering i *Mathematics* eller i andre fag utover *Assessing Student Learning*.

De norske studenter så ut til å ha hatt en del pensum i norskfaget. Interessant nok trakk tre av dem fram prinsippet «retting med grønt» fra en av pensumbøkene i *Matematikk*. I Queensland så de ut til å ha hatt betydelig mer pensumlitteratur om elevvurdering, først og fremst som en del av faget *Assessing Student Learning*.

5.3 Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer

I denne delen ønsket jeg å belyse hva slags teori- og praksisopplæring studentene hadde fått i særskilte temaer knyttet til elevvurdering. Jeg analyserer dette først med utgangspunkt i et globalt spørsmål som ble stilt om betydningen av teori- og praksisopplæring. Deretter

analyseres spesifikke problemstillinger knyttet til elevvurdering, og hva slags opplæring studentene hadde fått gjennom teori- og praksisopplæring.

5.3.1 Kunnskapstilegnelse om elevvurdering gjennom teori- eller praksisbasert opplæring?

Hovedinntrykket man fikk gjennom Camilla var at det lille hun hadde lært om elevvurdering var gjennom teoriopplæring i undervisningsfagene, og ikke gjennom praksis. Trygve fortalte derimot at han hadde lært mest om elevvurdering gjennom praksis. Han fortalte at de «gjennomførte en del prøver og sånn, kom fram til en skala på en måte» (Trygve-intervju, s. 8). De to studentene med fordypning i *Norsk for ungdomstrinnet* hadde altså svært ulike erfaringer fra teoriopplæringen på høyskolen.

Olav ga uttrykk for at han lærte mest om elevvurdering gjennom praksis i *Matematikk*, samt vikarjobbingen utenom studiene. Det samme gjaldt for andre fag. Eli understrekte at hun generelt ikke hadde hatt mye om elevvurdering, men at hun hadde lært noe gjennom opplæringa på lærerutdanningen og annen arbeidserfaring. Hun kunne ikke huske å ha lært noe særlig gjennom praksis i *Matematikk*. Hun understreket at hun hadde hatt lite veiledet praksis, blant annet fordi hun henholdsvis hadde vært på studieutveksling til utlandet og på skoleovertakelse de to siste praksisperiodene. De to studentene med fordypning i *Matematikk for ungdomstrinnet* (Olav og Eli) hadde altså også svært ulike historier.

I Queensland fortalte Amy at hun hadde lært mye om elevvurdering gjennom faget *Assessing Student Learning*, og at hun ellers hadde lært mye gjennom praksis i både *English* og *History (Social Science)*. Royce fortalte at han hadde lært mye om elevvurdering i *English Curriculum B*, og særlig gjennom hans to siste praksisperioder. Han mente han lærte mer om elevvurdering gjennom praksis enn på universitetet. «It sort of actualised what we had been talking about in uni». Han mente han lærte mer om elevvurdering i *HPE* enn i *English*, men at praksisen også var viktig: «You learn probably more in university, but what you learn in prac is how to manage it I suppose» (Royce-intervju, s. 13).

Philip mente han hadde lært en del om elevvurdering på universitetet, både som en del av faget *Assessing Student Learning* men også i andre fag. Han hadde dessuten lært en del gjennom praksis i *Mathematics*: «I actually made an exam paper with my student supervisor, in the first one, so, he basically showed me how he went through making an exam, which are good questions and why we choose that questions» (Philip-intervju, s. 7). Damion fortalte at han definitivt lærte mest om elevvurdering i faget *Assessing Student Learning*. I *Mathematics* lærte han ikke så mye på universitetet, utover at de studerte vurderingskapitlet i læreplanen.

Gjennom praksis i *Mathematics* lærte han en del om karaktersetting. Han pekte på at man på universitetet ikke får erfaring med konkrete eksempler, men det gjør man gjennom praksis: «you could say, ok, this is right or this is wrong, or this is an A standard, this is a B standard. And so that's where I really sort of started to see how different students' work was at different levels» (Damion-intervju, s. 9).

5.3.2 Opplæring i bruk av verktøy og hjelpemidler for elevvurdering

Som belyst i metodekapittelet (jf. kapittel 4.8.3 og 4.9.3) viste spørsmål 13 seg å være vanskelig å forstå for mer enn halvparten av studentene, og jeg måtte presisere hva jeg mente med spørsmålet. Med dette metodiske forbeholdet ønsker jeg likevel å trekke fram noen studentutsagn, fordi de er med på å anskueliggjøre hvor lite sentralt opplæring i vurderingsverktøy- og hjelpemidler tilsynelatende var i lærerutdanningene.

Analysen viser at det ikke var noen systematiske forskjeller, usikkerheten var knyttet til morsmåls- og matematikkfagene i Norge og Queensland. Både Camilla og Trygve ga uttrykk for at de hadde lært lite om bruk av hjelpemidler til vurderingsarbeidet i *Norsk*. Camilla fortalte at de har lært litt om bruk av vurderingsskjemaer, og Trygve fortalte at de har lært å bruke *ClassFronter* til levering av og å gi tilbakemelding på oppgaver.

Både Olav og Eli foralte at de hadde lært litt om diagnostisk testing i *Matematikk*. Olav fortalte at det i praksisen ble lagt stor vekt på dokumentasjon: «at du tar vare på alt du har gjort (...) Men det er ikke noen sånne konkrete verktøy i den forstand» (Olav-intervju, s. 8).

Camilla ble intervjuet på bakgrunn av hennes fordypning i *Norsk for ungdomstrinnet*, men hadde også fordypet seg i *Matematikk for ungdomstrinnet*. Hun var derfor i posisjon til å sammenligne de to fagene, og særlig knyttet til opplæring i vurderingsverktøy er hennes utsagn interessante:

I Norsken (...) praksisen på en skole i fjor (...) da viste hun oss et skjema etter, ok rettskriving, helhet, innhold, sånne ting (...). Så da fikk vi vel fire-fem sånne kategorier og vi satt enkeltkarakterer på en hel stil, og så var det å legge de sammen som ble sluttresultatet. Det var ganske nyttig, for akkurat det å sette karakter på et norsk skriftlig arbeid, det er ikke så veldig lett. I Matten derimot er det jo rett og slett bare å telle feil, som vi i hvert fall er blitt lært opp til, men gjerne skrive en skriftlig kommentar som oppmuntrer litt på slutten (Camilla-intervju, s. 8).

Utsagnet til Camilla er et eksempel på en tilsynelatende vanlige oppfatning om at karaktersetting er vanskeligere i språkfag enn i matematikkfaget.

Amy fortalte at hun gjennom praksis lærte mye om bruk av hjelpemidler, fordi «you have to face units that you've never done before, and you certainly haven't created them, so you don't know them very well, so you just have to learn it as you go» (Amy-intervju, s. 21). Royce fortalte at han i praksisperiodene i *English* likte å «get them to do just a bit of writing every week» (Royce-intervju, s. 14). Han nevnte ingen konkrete verktøy, men så ikke ut til å ha forstått spørsmålsstillingen etter intensjonen.

Philip foralte at de på universitetet fikk trening i å lage vurderingsoppgaver i *Mathematics*, og at de utvekslet oppgavene de lagde på CD slik at de kunne lære av hverandre. Gjennom praksis opplevde han både på *junior* og *senior secondary* at de gikk igjennom gamle eksamensoppgaver «so they can see what was done in the past and sometimes take questions and change them» (Philip-intervju, s. 8). Philip fortalte også at lærebøker i *Physics*, som de fikk bruke på universitetet, var rettet mot bare læreren, og kom med en CD med mange oppgaver og spørsmål man kan stille elevene for å vurdere måloppnåelsen. Damion fortalte at det i *Mathematics Curriculum A* og *B* ble lagt vekt på å bygge teknologi inn i vurderingsoppgavene. Men når det gjaldt å vurdere elevene, så handlet teknologibruk ikke om så mye annet enn «throwing all the data into the computer and sending it out» (Damion-intervju, s. 9).

5.3.3 Opplæring i og erfaringer med bruk av vurderingskriterier

Camilla ga uttrykk for at det var veldig bra da hun i praksisperioden i *Norsk* fikk utdelt et skjema med forskjellige kriterier de skulle vurdere etter. Dette var den eneste erfaringen hun fikk med bruk av kriterier i norskfaget (Camilla:XX). Trygve fortalte at de ikke hadde lært noe om bruk av kriterier, «annet en hvordan vi har fått karakterer selv da, altså erfart det på kroppen selv, (...) hvordan de har definert karakterene her». Han fortalte videre at «de har ikke satt noen bestemte kriterier» og mente dette var veldig individuelt, i forhold til hva hver enkelt lærer vektlegger. På spørsmål om han hadde sett noen kriterier brukt, for eksempel ved nasjonale prøver eller eksamensoppgaver, svarte Trygve:

Nei, eneste med nasjonale prøver er at elevene må lese innenfor en viss tid. Vi setter jo på en måte vurderingen ut fra hvor fort eller sakte de leser og hvor mye de forstår i løpet av å lese en tekst over en begrenset tidsperiode, så det er det eneste jeg har vært borti (Trygve-intervju, s. 11).

Olav trakk igjen fram diagnostisk testing, da han ble spurt om hvilke kunnskaper han hadde fått om bruk av vurderingskriterier i *Matematikk*. Han hadde lært mest om dette på høyskolen,

men også noe gjennom praksis. På spørsmål om han hadde fått noen ferdige kriterier som beskrev forskjellig grad av måloppnåelse i tilknytning til nasjonale prøver eller eksamensoppgaver, svarte Olav: «Vi har vel også fått utlevert de nasjonale prøvene, men der har ikke jeg sett noen sånne kriterier for hvor mange poeng du skulle ha og litt sånne ting» (Olav-intervju, s. 9). På det samme spørsmålet måtte Eli tenke seg om i noen sekunder, før hun spurte:

E: Kriterier... Hva tenker du på da?

ST: Hva som kreves for å få forskjellige karakterer da, om du har lært, for eksempel brukt noen eksempler fra nasjonale prøver eller noe sånt?

E: Ja, altså vi har jo diskutert sånne ting som hvordan man skal vurdere et svar som riktig eller feil, eller hvor mye man skal få i uttelling på svar, men jeg kan ikke huske at vi har noen helt sånn bestemte... nei (Eli-intervju, s. 9).

Igjen kan Camillas erfaringer med *Matematikk* trekkes fram: Hun fortalte at hun hadde rettet en del matematikkprøver i praksisperiodene. Det ser likevel ikke ut som at kriterier stod i sentrum hos Camilla i dette arbeidet: «Det blir jo bare at du teller riktige og gale, og en del halvpoeng». På spørsmål om ikke vanskelighetsgraden på oppgaven har betydning, svarte hun at «de fleste mattebøker nå har jo tre ulike linjer, som er tre ulike nivåer, og det har jeg erfart veldig mye i praksis.» Hun fortalte imidlertid at hun har lite erfaring med å lage og rette slike prøver, fordi hun «har kun rettet matteprøver på barneskolen» og der var ikke prøvene linjedelt (Camilla-intervju, s. 10).

Amy mente at det ikke ble lagt stor vekt på kriterier i opplæringen i *English* på universitetet: «I just got a model and I changed a few words in it, and that's pretty much it» (Amy-intervju, s. 22). Denne modellen fikk de fra læreplanen. Hun uttrykte frustrasjon over at det ble lagt lite vekt på dette på universitetet, både i *English* og *History (Social Science)*:

It is not a big emphasis on the criteria, on making judgments. The only time I learned how to make a judgment was in the *Assessing Student Learning*. It should be embedded in the curriculum studies, but I don't feel it is (...) the curriculum studies, both in *English* and *History*, I felt, it was mainly on unit production, unit production and creating tasks. Not on the actual making judgments and marking. So... That's just assumed that you're gonna know that (Amy-intervju, s. 23).

Amy fortalte at de i faget *Assessing Student Learning* hadde lært å ta utgangspunkt i læreplanen når man utformer vurderingskriterier: «I learned that it's adapted from, where, what the sources, because you're analysing the syllabuses, and how that's implemented to an actual task sheet, how it's changed in a way, not changed, but adapted to a task, like the wording» (Amy-intervju, s. 21). Hun mente at hun lærte mye om kriterier i *English* gjennom praksis:

I was there when my mentors were marking, so we were marking together, so she would first ask me «What do you think we should grade this paper?», and then she would check me and say «Well, actually I would have put it here on this criteria» (Amy-intervju, s. 22).

Selv om hun fikk en del trening i bruk av kriterier mente Amy at det var vanskelig: «It's hard, it's still hard to know, criteria are really tuff, particularly in English because you just, it's not straight-forward. Like I said, it's just your interpretation of a word» (Amy-intervju, s. 22). På spørsmål om de hadde lært å bruke kriteriene som står i læreplanen svarte Royce: «Well uni teaches us to use them specifically. But what prac sort of teaches you, is how to communicate to kids what it meant and what you are looking for» (Royce-intervju, s. 16). Som Amy, ga han uttrykk for at praksis ga den viktigste opplæringen i å forstå hvordan man skal bruke kriterier.

Philip fortalte at læreplanen i *Mathematics* har tre forskjellige kriterer som de bruker ved karaktersettingen, og at dette ble gjennomgått på universitetet. Han mente han fikk enda mer erfaring med vurderingskriterier gjennom praksis: «Because I actually got to mark them I had to think about it a lot more than normal, (...) and also talk with the teacher about what they consider an A was, what a B was, in the different areas» (Philip-intervju, s. 10). Philip lærte også en del om vurderingskriterier i faget *Assessing Student Learning*:

In the *Assessing Student Learning* we did go through using criteria sheets, and what makes a good criteria sheet. Also, yes, because a couple of the readings was on using criteria sheets. And one of the assessment pieces, we had to have criteria sheets with it as well (Philip-intervju, s. 10).

Damion fortalte at han gjennom *Mathematics Curriculum A* og *B* ikke hadde fått noen opplæring i hvordan man bør bruke kriterieskjemaer. «When I got to the *Assessing Student Learning* subject, in fourth year, then a lot of that fit into those... and that's when I really sort of learned how to do them properly and make them worthwhile» (Damion-intervju, s. 11). Han fikk ikke noe særlig erfaring med utarbeiding av kriterieskjemaer i praksis, fordi dette alt var utarbeidet av skolene. Han fikk imidlertid trening i å fortolke dem. Damion foralte

imidlertid at han i *Science* fikk trening i bruk av kriterier på universitetet, og særlig i en av praksisperiodene der elevene måtte gjennomføre og rapportere et eksperiment: «It was good. It was heaps easier to give them feedback, because you could show why they were A or why they were B» (Damion-intervju, s. 11).

Alle studentene i Queensland var opptatt av at vurderingskriterier må være eksplisitte. På spørsmål (19) om hva som er viktig når man gir tilbakemeldinger for at disse tilbakemeldingene skal bidra til mer læring, trakk Amy fram betydningen av å «undervise eksplisitt». Dette beskrev hun som å være tydelig på hva man forventer av elevene. Hun ga uttrykk for at det er et problem at mange lærere «don't really state what they want. They don't put it in dot points, they don't put it in sequences, steps for them to do. And it's just guess-work, like it's guess what the teacher wants» (Amy-intervju, s. 33). Royce understreket også betydningen av å bruke konkrete kriterier, som svar på dette spørsmålet: «If it's not specific, if it's general, it's generally not worthwhile» (Royce-intervju, s. 24).

5.3.4 Opplæring i og erfaringer med å diskutere vurderingskriterier

Camilla mente de ikke hadde diskutert vurderingskriterier noe særlig i *Norsk*, verken på høyskolen eller i praksis. Trygve fortalte at han ikke hadde diskutert kriterier direkte på høyskolen. Om praksis i *Norsk* fortalte han at de hadde diskutert det litt med øvingslærer. Han forklarte at dette går mest på etterarbeid:

Da må man vurdere elevene og i hvilken grad de har oppnådd det målet som man har satt for den økta eller timen eller den perioden som man underviser i. Altså, det er viktig å sette mål først og så må man analysere i etterkant om elevene har oppnådd de målene. Det må man også da teste, gjennom prøver eller andre måter å fastslå at de har det. Men (...) jeg tror ikke vi har bestemt noen egne kriterier, det har vi ikke gjort (Trygve-intervju, s. 11).

Olav fortalte at de hadde diskutert bruk av kriterier med andre matematikklærere når de lagde prøver i praksisperiodene:

Så diskuterte vi litte granne med tanke på poeng, spesielt, at du fikk et poeng for hver riktig og sånne ting. Slik at der ble det da diskutert om hvor mange poeng du måtte ha for å få de forskjellige karakterene, og litte granne sånt noe. At du går ut ifra cirka halvparten for å få en treer og litt sånn ut i fra det (Olav-intervju, s. 9).

Eli fortalte at de i opplæringen i *Matematikk* blant annet hadde gjennomgått nasjonale prøver, og diskutert i grupper og i plenum hvilken skår man skal gi de forskjellige elevsvarene. Hun

kunne ikke huske å ha diskutert kriterier med andre lærere i praksisperiodene, og forklarte dette med at det nå var tre år siden sist hun hadde veiledet praksis i *Matematikk*.

Amy fortalte at hun ikke hadde diskutert kriterier med andre studenter på universitetet. Dette forklarte hun med at «you don't really have the students' work to look at» (Amy-intervju, s. 24). Hun hadde imidlertid diskutert kriterier mye med kolleger gjennom praksis:

Yes, definitely, because you have this whole cross-marking thing, so I was, being a prac student you're not finally responsible for those marks that the students get, but you get to practice on it, and then you're mentor gets to check you (Amy-intervju, s. 23).

Amy fortalte at hun også hadde diskutert kriterier med andre lærerstudenter i praksisperiodene. Til forskjell fra Amy, fortalte Royce at de hadde diskutert kriterier på universitetet en hel del:

We had to work in groups of two or three to create quite a lengthy, one we had to do an English unit, and one we had to do a two year English work plan, and, so that requires you to speak about assessment, and also write up your criteria sheets, so your marking rubric I suppose (Royce-intervju, s. 17).

Han hadde ikke fått noen erfaring med karaktersetting på universitetet, men at satt karakter på mange elevarbeider gjennom praksis, og diskuterte da karaktersettingen uformelt med andre lærere også.

Philip fortalte at de ikke hadde diskutert bruken av vurderingskriterier noe særlig med andre studenter på universitetet. I praksisperiodene, derimot, drøftet han kriterier med lærerne, spesielt det ene av de tre kriteriene som ble regnet for å bli vurdert veldig subjektivt. Philip fortalte også at de på *junior secondary* hadde et system for intern moderering, der lærerne i samme fag og på samme trinn rettet hverandres prøver: «All the teachers would get the same test paper and then there was the moderation on, ok, what marks did you get that for, and stuff like that» (Philip-intervju, s. 12). På *senior secondary* var det som regel bare én lærer, så der var ikke dette mulig, men der var det til gjengjeld ekstern moderering: «After giving them what they think, they put it all to a folder and send it to another board outside the school, that does all the schools, and they decide whether the mark that was given was at the same level as all the others» (Philip-intervju, s. 12).

Damion fortalte at de ikke diskuterte kriterier i *Mathematics* på universitetet. Han ga heller ikke uttrykk for å ha diskutert det med lærere under praksis, men fortalte i stedet om

erfaringene med vurdering av eksperimentrapporter i *Science*. Han mener grunnen til at han fikk mer erfaring med dette i *Science* hadde å gjøre med «fagets natur»:

I think because the assessment techniques in Maths were more exam based, where yes, it is less subjective, like you just have to do, for example add fractions together, whereas in *Science* you might have to explain a process or show your understanding rather than actually just do something, like to show your understanding, so that's where the criteria kicked in more (Damion-intervju, s. 14).

5.3.5 Opplæring i karaktersetting

Camilla hadde ikke fått noen erfaring med karaktersetting på høyskolen. Hun fortalte imidlertid om én gang hun i praksis i *Norsk* fikk utdelt skjemaer som karaktersettingen skulle basere på. Da Olav ble bedt om å fortelle om hans kunnskaper og erfaringer med å sette karakter på elevers arbeid, lo han og sa «Ja, jeg kan prøve i hvert fall!» (Olav-intervju, s. 10). Han fortalte at han hadde fått noen få oppgaver som de skulle lage en skala på for så å diskutere i grupper hvordan det bør gjøres, men ellers lite. Han forklarte nærmere hvorfor han lo av spørsmålsstillingen: «Vi har vel ikke noen sånne enorme erfaringer når vi går ut, med tanke på det, det er vel noe alle ønsker seg mer av, føler jeg. I hvert fall jeg ønsker meg mer av det, og det er klart, det vi ofte sier er at vi lærer mest av praksisen vår» (Olav-intervju, s. 10).

Olav hadde et interessant svar på spørsmålet (20) om hva han tenkte om å sette en samlet karakter i faget: «Det er klart det er, sånn som skolesystemet er bygd opp så er det nødt til å være sånn, føler jeg. Det er nå en gang sånn at skolesystemet er bygd opp på den måten at du får en karakter» (Olav-intervju, s. 13). Det ser ut til at Olav oppfattet spørsmålet politisk, heller en praktisk. På spørsmål om hvordan han kan sikre at karaktersettingen blir så rettferdig som mulig (spørsmål 22), svarte han:

Hvis du er stille og gjør det bra på prøver, for å si det sånn. For du må ha et helhetlig inntrykk. Det er jo det man skal ha nå, eller ha i karaktersettingen. Og hvis du da ikke ytrer noen som helst meninger i timene så vil du ikke få en sekser, for å si det sånn (Olav-intervju, s. 15).

På spørsmål om hvordan han kan holde oversikt over dette, svarte han:

Det blir sånn sett, med tanke på dokumentasjon, at du hele tiden har prøver, nesten å ha en oversikt over hvem som svarer i timen omtrentt (...) når du har hatt en time, så tenk igjennom hvem som faktisk har sagt noe i timen, hvem du har jobbet litt granne sammen med, hvem som har kommet med fornuftige/ufornuftige ting da (Olav-intervju, s. 15).

På spørsmål om kunnskaper og erfaringer med å sette karakter brøt også Eli ut i latter og sa: «Ja, det er ikke så mye!» (Eli-intervju, s. 11). Hun viste til den ene timen der de hadde diskutert nasjonale prøver, men kunne ikke helt huske om de skulle diskutere seg fram til en karakter. «Jo, vi må jo ha gjort det», konkluderte hun, og lo mer. Hun hadde ikke fått noen erfaring med karaktersetting fra praksis i *Matematikk* (Eli-intervju, s. 11).

Ingen av studentene fra Queensland hadde fått vesentlig erfaring med karaktersetting på universitet. De ga uttrykk for at de fikk en del erfaring med karaktersetting fra praksis i fagene, og at de fikk hjelp av øvingslæreren i dette arbeidet. Amy forklarte hvordan dette hadde foregått:

In prac I would take not all of them, but I would take, say, a third of them home, throughout the week and mark a couple of nights and take them back to my mentor and she would assess how I'm marking. And tell me if I'm off the mark, or ask me to justify why I've chosen that (Amy-intervju, s. 26).

Amy fortalte at tilbakemeldingene fra støttelæreren oftest var at hun var for streng i karaktersettingen. Både Amy og Royce ga uttrykk for bekymring for om de ga riktige karakter. Amy ga uttrykk for at hun var redd at det store prøvepresset kunne gå ut over rettferdigheten i karaktersettingen, mens Royce var bekymret for hvordan man skal ta hensyn til spesielle tilfeller. Begge ga uttrykk for at kommunikasjon med andre lærere er det viktigste virkemiddelet for å sikre en rettferdig karaktersetting. De mente modereringsprosedyrene bidrar til å begrense deres bekymringer knyttet til rettferdighet i karaktersettingen.

Philip fortalte at han ikke hadde fått noen erfaring med karaktersetting på universitetet i *Mathematics*. I praksis derimot har han erfaring med karaktersette eksamensbesvarelser for en hel klasse. Han forklarte hvordan øvingslæreren ga tilbakemeldinger når han trengte hjelp: «Like when I wasn't sure whether to give some marks for something that a student had done, I'd go see him and say 'how much would you give for doing that?'». (Philip-intervju, s. 14). Damion hadde ikke fått noen trening i karaktersetting på universitetet. Han fortalte at han hadde fått litt erfaring med karaktersetting i alle de tre praksisperiodene, men ikke veldig mye.

På spørsmålet (24) om hva de skulle ønske de hadde lært mer om i lærerutdanningen knyttet til elevvurdering, etterlyste *alle* studentene, både i Norge og Queensland, mer opplæring i karaktersetting.

5.3.6 Opplæring i å gi faglige tilbakemeldinger

Camilla fortalte at hun hadde fått «veldig lite» erfaring med å gi elevene tilbakemeldinger gjennom lærerutdanningen. Hun hadde lært at det er viktig «å hele tiden skryte, eller om de svarer noe riktig, så er det det å gi dem feedback da, positiv feedback hele tiden» (Camilla-intervju, s. 11). I praksis hadde hun gitt karakter og skriftlig kommentar på matematikkprøver, stiler og andre typer skriftlige produkter, men hun mente det var like viktig å gi muntlig feedback i timene.

Trygve var ikke helt på linje med Camilla. Han fortalte at han gjennom opplæringen i *Norsk* på høyskolen lærte at «muntlige tilbakemeldinger er ofte de spontane, som man må ta der og da, men man også da passe på at det er en konstruktiv tilbakemelding, ikke bare ros. Eller elevene skal få ros når de fortjener å få ros, men ikke rose bare for å rose» (Trygve-intervju, s. 12). Han fortalte at han på høyskolen hadde lært litt om hvordan man skal gi tilbakemeldinger, mens han gjennom praksis hadde fått erfaring med hvilken virkning tilbakemeldingene har på elevene. Trygve fortalte videre at han på høyskolen hadde lært at det ved skriftlige oppgaver i *Norsk* er viktig å ta en ting av gangen: «Først skal man kanskje få elevene til å i det hele tatt skrive en hel tekst og så kan man se på språk, så kan man se på innhold, og så kan man se på grammatikk og så videre». Han fortalte at han i praksis hadde fått erfare hvordan elevene reagerer på tilbakemeldingene man gir: «Om de da er motiverende, om de er demotiverende, eller hvis de stiller seg likegyldig til det» (Trygve-intervju, s. 12).

Olav fortalte at de ikke fikk noen trening i å gi tilbakemelding, verken muntlig eller skriftlig, i *Matematikk* på høyskolen. Eli fortalte mer om den ene timen de diskuterte vurdering i *Matematikk*. Der ble det lagt vekt på at man i Norge gir for mye generelle tilbakemeldinger. «Sånn 'flott' og 'fint' og alt er bra på en måte, og at elevene derfor (...), de vet ikke hvor de står, de tror de er mye flinkere enn det de egentlig er». Hun fortalte at hun gjennom praksis rettet bøker til elever, og da var hun opptatt av dette og skrev «Fint, men se mer på det og det» (Eli-intervju, s. 12).

Amy mente hun ikke hadde lært noe om hvordan man skal gi tilbakemeldinger på læringsarbeidet i *English* på universitetet. I praksis fikk hun i hovedsak erfaring med å gi skriftlig tilbakemelding på utkast til skriftlige arbeider. Muntlige tilbakemeldinger gjorde seg mest gjeldende ved at man skulle «make yourself available, in lunch time, in class when doing independent work, to be approached by the students or go and seek if they need any help» (Amy-intervju, s. 28). Royce fortalte at han hadde lært «absolutely nothing» om hvordan man

skal gi muntlige og skriftlige tilbakemeldinger i *English* på universitetet. Gjennom praksis hadde han lært at «oral feedback is something that's encouraging them really, you know, not necessarily curriculum-wise but behaviour-wise with students and everything» (Royce-intervju, s. 21). Han fortalte at skriftlig tilbakemeldinger ble gitt i forbindelse med kommentarer på elevs utkast til oppgavebesvarelser.

Philip hadde heller ikke lært noe om å gi tilbakemeldinger i *Mathematics* på universitetet, og mente dessuten at *Assessing Student Learning* handlet mest om å utarbeide vurderingsoppgaver. I praksis hadde han, særlig ved retting av eksamensoppgaver, vært opptatt av å «exactly showing them where they've done a step wrong, where they've gone wrong and why they, always, whenever they've lost a mark, always to show why they've lost marks» (Philip-intervju, s. 15). Damion fortalte også at han ikke hadde lært noe om hvordan man bør å gi muntlige og skriftlige tilbakemeldinger i *Mathematics* på universitetet. Han trakk imidlertid fram at man måtte gi tilbakemeldinger på hverandres presentasjoner i faget *Assessing Student Learning*, noe han opplevde som nyttig. Han hadde fått mye erfaring med å gi muntlige formative tilbakemeldinger gjennom praksis. Skriftlige tilbakemeldinger var kun i tilknytning til karaktersetting, og det hadde han ikke fått så mye erfaring med.

5.3.7 Opplæring i forskrifter og retningslinjer

På spørsmål om hva slags opplæring hun hadde fått i forskrifter og retningslinjer for elevvurdering svarte Camilla at hun i faget *Tilpasset opplæring* hadde lært at elever som ikke har utbytte av vanlig undervisning hadde rett på individuell opplæringsplan. Trygve svarte: «Ingen, egentlig. Det har ikke vært så mye snakk om, annet enn at vi er pålagt å gi en vurdering. Men de har ikke sagt at det står der eller der, i den paragrafen» (Trygve-intervju, s. 13).

Olav fortalte at de i *Matematikk* på høyskolen gikk igjennom læreplanene, både L97 og Kunnskapsløftet. Han fortalte videre at de hadde sett litt på opplæringsloven:

At du har krav på både en sluttevaluering eller –vurdering og gjennom, eller underveisvurderinger hele tiden. (...) vi har jo sett på de lovene som står der, det har vi jo. At det er lovbestemt at du skal ha vurderinger gjennom hele tiden, men ikke noe sånn om kriterier rundt det, det er det ikke, for de forskjellige karakterene i hvert fall (Olav-intervju, s. 12).

Eli fortalte at de i *Matematikk* på høyskolen hadde snakket om at det ikke står noe om elevvurdering i Kunnskapsløftet, men at det er en egen forskrift til Opplæringsloven. «Ja, den

har vi gått igjennom. Jeg husker ikke alt, men blant annet det der at elevene skal være med å vurdere, at det skal legges til rette for egenvurdering» (Eli-intervju, s. 13).

Alle studentene i Queensland fortalte at de i fagdidaktikken på universitetet studerte læreplanene, og at det er der de viktigste retningslinjene står. Amy og Royce fortalte at de også studerte gamle læreplaner og dermed fikk sett utviklingen fram mot den gjeldende læreplanen. Royce foralte at særlig retningslinjene i *senior syllabus* er viktige, fordi disse gir føringer på hvor mange og hva slags vurderingsoppgaver man skal gi elevene.

Philip fortalte at de i *Mathematics* på universitetet studerte læreplanen, blant annet karakterkravene knyttet til de tre kriteriene for å oppnå de forskjellige karakterene: «Yes, to get a VHA they have to get I think it's two VHA's in two different criteria and HA in an other criterion» (Philip-intervju, s. 17). I praksisperiodene var også læreplanen sentral. Han hadde også lært litt om hvordan man sender elevarbeider til ekstern moderering og hvordan man rapporterer elevenes måloppnåelse på «report cards». Damion fortalte at han hadde lært litt om forskjellige policy-dokumenter i *Assessing Student Learning*, blant annet QCAR (Queensland Curriculum Assessment and Reporting framework). Selv om de ikke hadde snakket om dette i praksisperiodene, opplevde han at det de hadde lært om policy-dokumenter på universitetet kom til nytte i praksis.

5.4 Sammenfattende sammenligning

Nedenfor følger en sammenfattende gjennomgang av de mest interessante funnene.

5.4.1 Organisering av lærerutdanningens teori- og praksisopplæring

Lærerutdanningene i Norge og Queensland har svært ulik organisering. I Norge har studentene de to første årene et felles obligatorisk løp med opplæring *Norsk*, *Matematikk*, *KRL*, *GLSM* de to første årene, før de fikk anledning til å velge fag og eventuelt fordype seg mot ungdomstrinnet de to siste årene. *Secondary*-programmet i Queensland er et gjennomgående obligatorisk løp der studentene de første to årene får opplæring i to undervisningsfag, og de to siste årene i fagdidaktikk og pedagogikk.

Valgfriheten i den norske modellen gjør det krevende å sammenligne de to casene. En norsk student som valgte full fagfordypning i *Norsk* og *Matematikk*, ville i prinsippet kunne få omtrent like mye fagkompetanse i disse to fagene, som alle lærerstudentene i Queensland fikk i sine respektive to fag. Studentene i Queensland fikk imidlertid *tre ganger så* mye fagkompetanse rettet spesifikt mot ungdomstrinnet, som det de norske studentene maksimalt kunne få. Blant disse studentenevar det imidlertid ingen som hadde valgt full fordypning i

Norsk og Matematikk. Dette er dessuten de fagene det er mulig å oppnå mest fordypning i som følge av de er gitt høy prioritert blant de obligatoriske fagene de to første studieårene. Alle norske allmennlærere som fullfører utdanningen med fordypning i andre fag, får mindre formell kompetanse i disse fagene enn lærere i Queensland får. Lærere som begynner å undervise på ungdomstrinnet i Norge, står faglig langt tilbake for sine kolleger i Queensland, som følge av at de har fått en hel del lite relevant opplæring i hvordan de bør undervise barneskoleelever. Spesialisering mot *secondary*, og begrensningen til to fag, frigir dessuten mer tid til generelle pedagogiske fag i Queensland. Samlet sett får disse lærerstudentene i overkant av 50 prosent mer opplæring i generelle pedagogisk fag som lærerstudentene i Norge.

Den norske lærerutdanningen har fem uker mer praksis enn Queenslands lærerutdanning. I sistenevnte er imidlertid praksis sentrert rundt øvingslærere som har samme undervisningsfag som dem selv og fordelt på *junior* og *senior secondary*, mens den i Norge må fordeles mellom en rekke fag. Følgelig er det grunn til å anta at studentene i Queensland fikk mer undervisningspraksis i hvert undervisningsfag enn de norske studentene. Færre undervisningsfag bidro trolig også til mindre tilfeldigheter knyttet til fordeling av praksisplasser i Queensland, og hvor relevante disse var for studentene. Mens ingen studenter der fortalte om at de måtte undervise i andre fag enn undervisningsfagene, så dette ut til å være et betydelig problem i Norge. Flere av de norske studentene fortalte dessuten om en skoleovertakelsesordning, som hadde blandede innvirkning på muligheten til å få praksis i undervisningsfaget. På den ene siden ga det studentene det fulle undervisningsansvaret, som kunne være nyttig dersom man fikk dette i undervisningsfagene. På den annen side fortalte studenter om at de var blitt plassert i skoleadministrasjonen og derfor ikke fikk noen praksis den uken.

5.4.2 I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?

En av de mest markante forskjellene mellom Norge og Queensland, knytter seg til det faget om elevvurdering (7,5 studiepoeng) som studentene i Queensland hadde hatt, og som det ikke fantes noen norsk parallell til. Det er også markante forskjeller på vektleggingen av elevvurdering i undervisningsfagene. Det ser ut til at elevvurdering hadde en sterkere posisjon i morsmålsfagene enn i matematikkfaget, både i Norge og Queensland. Både Amy og Royce viste til undervisningsplanene de måtte utarbeide i *English*, og det kan det se ut som temaet var mer strukturert som en vesentlig del av morsmålsfaget i Queensland enn det var i Norge.

Det er også interessant å se at både Camilla og Trygve kunne berette om at fokuset var på formativ vurdering (proessorientert skrijving, mappevurdering, tilbakemeldinger) i norskfaget, mens det i Queensland ble lagt mest vekt på de formelle og summative sidene i *English*. Alle studentene i Queensland fortalte at temaet elevvurdering var tatt opp i andre fag på lærerutdanningen. Som for undervisningsfagene ser det ut til at disse studentene hadde møtt temaet på en mer strukturert måte, i forbindelse med behandling av konkrete undervisningsplaner (units) og læreplananalyser.

Det trer også fram betydelig forskjeller mellom landene og fagene når det gjelder pensumlitteratur om elevvurdering. Ingen av de norske studentene hadde hatt noen bøker der elevvurdering var hovedtema. Fra norskfaget ble det foralt om en bok om *proessorientert skriveopplæring*, der mappevurdering var sentralt. I *Matematikk* nevnte en student boken *Meningsfylt matematikk* av Geir Botten, der det blant annet ble lagt vekt på å «rette med grønt», et uttrykk to andre studenter også var kjent med. I tillegg ble nevnte en student en uspesifisert bok der diagnostisk testing ble tatt opp. En annen student nevnte også boken *Fagdebatikk*, uten å si noe konkret om hva han hadde lært om elevvurdering der. Ellers fortalte studentene at de trodde elevvurdering ble tatt opp i bøker i *KRL* og *Pedagogikk*. Den eneste boken som her ble nevnt var *Kan læring planlegges* av Britt Ulstrup Engelsen.

I Queensland hadde de omfangsrik pensumlitteratur på elevvurdering i tilknytning til det obligatoriske faget *Assessing Student Learning*, blant annet en egen lærebok om elevvurdering kalt *Celebrating Student Achievement* og et kompendium med artikler satt sammen av læreren. Disse studentene viste også til læreplanene som sentral litteratur for vurderingsfeltet, mens dette ikke ble nevnt av noen av de norske studentene. Som i Norge, trakk studentene fra Queensland til sammen fram flere eksempler på pensumlitteratur om elevvurdering innenfor andre fag. Samlet sett rapporterte studentene fra Queensland om vesentlig større omfang av pensumlitteratur som tok opp temaet elevvurdering enn de norske studentene gjorde.

5.4.3 Teori- og praksisbasert opplæring i elevvurdering knyttet til særskilte temaer

De syv spørsmålene som ble stilt med den hensikt å skille mellom hva man lærer gjennom teori- og praksisbasert opplæring har avdekket en rekke interessante aspekter. Jeg har valgt å sammenfatte disse syv spørsmålene punktvis.

Kunnskapstilegnelse om elevvurdering gjennom opplæring eller praksis?

Det trer fram betydelige forskjeller mellom landene knyttet til kunnskapstilegnelse gjennom teori- eller praksisopplæring generelt: To av de fire norske studentene ga uttrykk for at de lærte mest gjennom praksis, mens de to andre ga uttrykk for at det lille de hadde lært, var på høyskolen. Det er ingen systematisk forskjeller mellom fagene. Det er grunn til å tro at forskjellene mellom studentenes oppfatning av hvor mye de lærte på høyskolen var påvirket av hvor mye de opplevde å ha lært gjennom praksis, hvilket det knytter seg mye tilfældigheter til som følge av vanskeligheter med overensstemmelse mellom studentene og øvingslærernes undervisningsfag. Studentene fra Queensland ga et mer entydig bilde. Alle ga uttrykk for å ha lært en del både på universitet og gjennom praksis. Praksis ser imidlertid ut til å ha vært klart viktigst, mens faget *Assessing Student Learning* også ble framholdt som viktig. De opplevde også at de hadde lært litt i undervisningsfagene på universitetet. Det så ut til å hefte litt mindre tilfældigheter knyttet til relevansen av praksis i Queensland, sammenlignet med Norge.

Opplæring i bruk av verktøy og hjelpemidler for elevvurdering

Jeg har tatt et metodisk forbehold om studentenes svar på dette spørsmålet, som følge av at flere av dem hadde problemer med å forstå spørsmålsstillinge, trolig som følge av forvirring om betydningen av begrepene «verktøy» og «instruments». Dette kan imidlertid være en indikasjon på at dette er lite utviklet knyttet til elevvurderingsfeltet både i Norge og Queensland. De norske studentene ga uttrykk for at de hadde fått lite opplæring i verktøy og hjelpemidler for elevvurdering. Eksempler som ble trukket fram var *ClassFronter*, vurderingsskjemaer i *Norsk* og diagnostiske tester i *Matematikk*. Studentene fra Queensland hadde også fått lite opplæring i vurderingsverktøy og hjelpemidler. Philip fortalte eksempelvis om et opplegg i *Mathematics*, der de fikk trening i å lage vurderingsoppgaver og utvekslet disse på en CD. Selv om det gjaldt et annet fag, er det også interessant å trekke fram hans erfaringer med lærebøker i *Physics* som var rettet mot bare læreren, og kom med en CD med mange spørsmål som de kunne stille elevene, en hjelp i å vurdere elevenes kompetansenivå.

Opplæring i og erfaringer med bruk av vurderingskriterier

De norske studentene tegnet et bilde av en allmennlærerutdanning der kriterier ikke var en sentral del av opplæringen, og de ga heller ikke uttrykk for å ha lært noe særlig om dette i praksisperiodene. I Norge var dette, sagt med Trygves ord, «veldig individuelt da i forhold til hva hver enkelt lærer vil vektlegge» (Trygve-intervju, s. 10). Både Amy og Royce ga uttrykk

for at kriterier var lite sentralt i *English* på universitetet. Amy fortalte at det eneste stedet hun der hadde lært om vurderingskriterier, var i faget *Assessing Student learning*. Både Amy og Royce ga imidlertid uttrykk for at de hadde lært mye om bruk av kriterier gjennom praksis. Vurderingskriterier framstod her som viktig, først og fremst fordi det er en viktig del av læreplanen og følgelig alt vurderingsarbeidet som foregår i skolen. Amy fortalte at hun hadde lært mye gjennom å rette prøver sammen med øvingslæreren, mens Royce mente han i praksisperioden hadde lært mye om hvordan man skal kommunisere kriteriene til elevene slik at de kan forstå dem. Philip fortalte at læreplanen i *Mathematics* har tre forskjellige kriterier som de bruker ved karaktersettingen, og at dette ble gjennomgått på universitetet, men at han fikk enda mer erfaring med vurderingskriterier gjennom praksis. Damion mente han ikke hadde hatt noen idé om hvordan man bør bruke kriterieskjemaer før han tok faget *Assessing Student Learning*. Han fikk ikke noe særlig erfaring med utarbeiding av kriterieskjemaer i praksis, fordi dette alt var utarbeidet av skolene., men han fikk trening i å bruke dem.

Opplæring i og erfaringer med å diskutere vurderingskriterier

De norske studentene fortalte at de i liten grad hadde diskutert kriterier på høyskolen. Trygve og Olav fortalte at de hadde diskutert det litt med øvingslærer gjennom praksis, mens Camilla ikke sa noe om dette. Eli skilte seg ut ved ikke å ha noen erfaring med dette fra praksis, men fortalte på den annen side om retting av nasjonale prøver i plenum på høyskolen. Mens Amy, Philip og Damion fortalte at de ikke diskuterte noe om bruk av kriterier på universitetet, så fortalte Royce at de gjorde dette en del» ved at de ble satt sammen i grupper som utformet undervisningsplaner der kriterieskjemaer inngikk. Amy, Royce og Philip fortalte at de diskuterte kriterier med øvingslærere og andre lærere gjennom praksis. De ga uttrykk for at det ble lagt stor vekt på dette som følge av systemene for intern og ekstern moderering. Damion ga uttrykk for at de verken har diskutert kriterier på universitetet eller i praksis i *Mathematics*, og mente kriteriene er mye mer sentralt i *Science*. Samlet sett framstår ikke vurderingsskriterier som et sentralt diskusjonstema verken på høyskolen eller i praksis i Norge. I Queensland derimot, framstod dette som langt mer sentralt i studentenes praksis, blant annet på grunn av skolenes modereringsprosedyrer, mens bare en fortalte om erfaringer med dette fra universitetet.

Opplæring i karaktersetting

Verken i Norge eller Queensland hadde studentene fått noen erfaring med karaktersetting gjennom høyskolen eller universitetet. I Norge kan det se ut som at studentene fikk noe

trening med karaktersetting dersom de hadde praksis i ungdomsskolen, men det gjaldt ikke alle. I barneskolen vil retting av prøver måtte være begrenset til å gi tilbakemelding på hva som var rett og galt, ettersom man ikke har lov til å gi karakterer. I Queensland svarte alle studentene at de hadde fått en del erfaring med karaktersetting gjennom praksis. Det så ut til å være etablert rutiner for at studentene skal få trening i dette.

Opplæring i og erfaringer med å gi faglige tilbakemeldinger

De norske studentene hadde forskjellige oppfatninger om hva som kjennetegnet gode faglige tilbakemeldinger. En ga klart uttrykk for at det var viktig å gi «positiv feedback hele tiden», mens to andre også pekte på at man ikke må «rose bare for å rose». De ga i liten grad uttrykk for å ha lært hvordan man bør gi tilbakemeldinger på høyskolen, mens så ut til å ha fått noen erfaringer med dette gjennom praksis. *Ingen* av studentene i Queensland fortalte at de hadde lært noe på universitetet om hvordan man skal gi faglige tilbakemeldinger, med unntak av en student som trakk fram at de lærte av å gi hverandre tilbakemeldinger på presentasjoner i faget *Assessing Student Learning*. Studentene ga uttrykk for at de hadde fått en erfaring med å gi tilbakemelding gjennom praksis, både skriftlig i forbindelse med innleveringer og prøver, og muntlig, gjennom vanlig interaksjon med elevene i klasserommet.

Opplæring i forskrifter og retningslinjer

Det var vesentlige forskjeller mellom fagene i Norge, knyttet til opplæring i forskrifter og retningslinjer for elevvurdering. Både Camilla og Trygve fortalte at de ikke hadde lært noe om retningslinjer for elevvurdering i *Norsk*, verken på høyskolen eller gjennom praksis. Både Olav og Eli fortalte på den annen side at dette temaet hadde blitt tatt opp i matematikkopplæringen på høyskolen. Det er ikke grunnlag for å knytte dette til vesensforskjeller mellom *fagene*. Forskjellene henger trolig sammen med at en matematikklærer på høyskolen har lagt spesielt vekt på opplæring i dette. I Queensland hadde alle studentene blitt kjent med retningslinjer for elevvurdering gjennom studier av læreplanene i undervisningsfagene. Noen av dem fikk også opplæring i forskjellene mellom gamle og gjeldende læreplaner. En av studentene hadde gjennom praksis fått kjennskap til retningslinjene for moderering, og en annen var blitt kjent med de nye rammeverket for læreplaner og vurdering (QCAR) gjennom faget *Assessing Student Learning*.

6 Drøfting av sentrale funn

I dette kapittelet drøfter jeg de viktigste funnene i analysen. Jeg trekker fram forskjeller mellom Norge og Queensland, og mellom fagene, og diskuterer mulige forklaringer med utgangspunkt i teoretiske perspektiver fra kapittel 2 og 3. I en studie som denne må man ta forbehold om at studentenes utsagn verken behøver å være representativ for studentene som studerer der, for hva slags opplæring lærerutdanningen har forsøkt å gi dem. Heller enn å felle en dom over norsk allmennlærerutdannings tilnærming til elevvurdering, bør funnene brukes til å undersøke problemstillingene som reises nærmere. Utgangspunktet for drøftingen er problemstillingen jeg presenterte innledningsvis og tilhørende forskningsspørsmål.

Hva slags opplæring i elevvurdering fikk norske allmennlærerstudenter som avsluttet utdannelsen i 2007, sammenlignet med lærerstudenter utdannet i samme periode i Queensland, Australia?

1. Hvordan påvirket organiseringen av teori- og praksisopplæring studentenes opplæring i elevvurdering?
2. I hvilket omfang og på hvilken måte møtte studentene temaet elevvurdering?
3. Hva slags teori- og praksisopplæring fikk studentene i
 - verktøy og hjelpemidler for elevvurdering;
 - bruk av vurderingskriterier;
 - diskusjoner om vurderingskriterier;
 - karaktersetting;
 - å gi faglige tilbakemeldinger; og
 - forskrifter og retningslinjer for elevvurdering

Det er naturlig å innlede drøftningskapittelet med en diskusjon om betydningen av de svært ulike modellene for organisering av lærerutdanningen generelt, blant annet mulige fordeler og ulemper ved de ulike forvaltningsmessige tilnærmingene til lærerutdanning, før jeg diskuterer konsekvensene dette har for opplæring i elevvurdering. Videre drøfter jeg hvilket omfang og på hvilken måte studentene møtte temaet elevvurdering. Deretter drøfter jeg hva slags teori- og praksisopplæring studentene hadde fått i temaene i det tredje forskningsspørsmålet. Avslutningsvis ser jeg på overordnede likheter og forskjeller på tvers av casene, både mellom norsk- og matematikkfaget i Norge, og mellom de to landene generelt.

6.1 Betydningen av organiseringen av lærerutdanningens teori- og praksisopplæring

Analysen av organiseringen av de to lærerutdanningene viste svært store forskjeller på nyutdannede læreres samlede opplæring i undervisningsfagene, som følge av ulik grad av spesialisering i Norge og Queensland. I Queensland valgte studentene *to* fag de ønsket å undervise i på *secondary* nivå, alt ved starten av opplæringen. I Norge gjennomgikk studentene først en bred og generell opplæring i *Norsk, Matematikk, KRL* og *GLSM* rettet mot hele grunnskolen de to første årene, før de fikk anledning til å velge fag og eventuelt fordype seg mot ungdomstrinnet.

En norsk student som valgte å fordype seg mot barne- og ungdomstrinnet både i *Norsk og Matematikk*, kunne til sammen få 90 studiepoeng fordypning i hvert av disse fagene. Som vist i analysedelen fikk studentene i Queensland tilnærmedesvis like mye (86,5 studiepoeng) faglig fordypning i hvert av undervisningsfagene. Studentene i Queensland fikk imidlertid nesten tre ganger så mye fagkompetanse rettet spesifikt mot ungdomstrinnet, som det de norske studentene maksimalt kunne få (30 studiepoeng). Norske allmennlærere som fullførte utdanningen med fordypning i andre fag enn *Norsk og Matematikk* ved denne høyskolen, fikk mindre formell kompetanse i sine respektive undervisningsfag enn studentene ved lærerutdanningen i Queensland. Spesielt norske lærere som begynte å undervise på ungdomstrinnet hadde mindre teoretisk og praktisk opplæring enn deres kolleger i Queensland. Spesialisering mot *secondary* og begrensningen til to fag, frigjorde dessuten mer tid til generelle pedagogiske fag i Queensland. Samlet sett fikk disse lærerstudentene omtrent 50 prosent mer opplæring i generelle pedagogiske fag enn allmennlærerstudentene i Norge.

Den norske lærerutdanningen hadde fem uker mer praksis enn Queenslands lærerutdanning. Ettersom praksis i Norge må deles mellom mange undervisningsfag, er det imidlertid grunn til å tro at studentene i Queensland også fikk mer undervisningspraksis i hvert av sine to undervisningsfag enn de norske studentene fikk i hver av sine. De mange fagene førte også til større grad av tilfeldigheter knyttet til hvilke fag de norske studentene fikk undervise i. Det er derfor grunnlag for å hevde at lærerstudentene ved denne norske allmennlærerutdanningen generelt fikk mindre praksiserfaring i fagene de skulle undervise i, enn lærerstudentene ved den aktuelle lærerutdanningen i Queensland. Et utsagn fra den norske studenten Eli, som hadde fordypet seg i *Matematikk for ungdomstrinnet* kan illustrere trekk ved vilkårligheten knyttet til organiseringen av praksisopplæringen ved den norske høyskolen:

Første året så var vi seks uker på samme skole. Ja, det var to uker først og så var det fire uker. Og da var det mer veiledet praksis hele tiden, da var jeg i 5. klasse. Og så året etter så var det samme opplegget, men da var jeg i en 7. klasse. To uker først, og så fire uker. Og så tredje året så var det åpent for at man kunne ta praksis i utlandet. Så hadde jeg først to uker skoleovertakelse, og så hadde jeg praksis i utlandet. Og så har det vært skoleovertakelse i år. Så derfor er det lenge siden det har vært noe... (Eli-intervju, s. 14).

Eli fikk altså knapt noe praksisopplæring de to siste årene, som var da hun tok fordypning i *Matematikk for ungdomstrinnet*. Dersom hun begynte å jobbe på en ungdomsskole den påfølgende høsten, var det altså uten noen praksiserfaring fra ungdomstrinnet i løpet av lærerutdanningen. I forbindelse med fordypning i matematikk og andre undervisningsfag de to siste årene hadde hun ikke hatt noen praksiserfaring i det hele tatt, verken på barnetrinnet eller ungdomstrinnet. Den manglende praksisen kan knyttes til flere aspekter ved organiseringen av den norske allmennlærerutdanningen. Dersom praksisperiodene hadde vært organisert i egne studieenheter, slik de var i Queensland, så hadde det kan hende vært enklere å stille krav som gjorde det umulig å «unnslippe» en praksisperiode, slik Eli gjorde tredje studieår. Man kunne også med fordel stilt strengere krav om utplassering på barne- eller ungdomsskoler, i tråd med studentenes fordypning.

Som påpekt i teorikapitlet (jf. kapittel 3.7.1) er det godt dokumentert at læreres formelle opplæring har stor betydning for elevers læringsresultater. Det er derfor bekymringsfullt at norske allmennlærere står betydelig svakere rustet for undervisning på ungdomstrinnet enn de nyutdannede lærerne i Queensland. At fagkompetanse blir høyt verdsatt i Queensland kommer også til uttrykk gjennom de strenge registrerings- og kvalitetssikringsprosedyrene for lærere og lærerutdanninger (jf. kapittel 2.2.3). I Queensland stilles det et absolutt krav om utdanning på *secondary*-nivå, for å undervise i *junior*- eller *senior secondary* skoler. Allmennlærerutdanning er det eneste kravet til utdanning for å bli fast ansatt som ungdomsskolelærer i Norge. Det er opp til skoleledere, i ansettelsesprosessen, å avgjøre om det skal stilles krav til fordypning i «ungdomsskoledidaktikk». Dersom skolens ledere ikke stiller slike krav, kan allmennlærere få undervise i ungdomsskolen uten noen opplæring rettet spesifikt mot dette alderstrinnet. Det ligger utenfor denne oppgaven å utforske hvor mange slike lærere som underviser i norske ungdomsskoler. Jeg ønsker å peke på at *muligheten* for det i seg selv er et uttrykk for at lærernes fagkompetanse blir lite vektlagt i Norge.

Tilfeldighetene knyttet til utplassering i praksis, slik dette fremstår i det norske caset, innebærer at det er mulig å få formell kompetanse til å undervise i norsk grunnskole, inkludert

opplæring på ungdomstrinnet, *uten* å ha gjennomført praksis i fag man skal undervise i. At krav til både teori- og praksisopplæring i undervisningsfagene ikke er regelstyrt mener jeg er problematisk, ettersom det kan føre til at slike lærere blir ansatt dersom skolen har mangel på bedre kvalifiserte søkere, i stedet for at stillingene utlyses på ny.

Det samlede inntrykket man får når man sammenligner forskjellene i krav til spesialisering mot både fag og alderstrinn, og krav til fagkompetanse for å undervise, er at lærerne i Queensland i langt større grad anses for å være *faglærere*. Som vist i teorikapittelet (jf. kapittel 3.7.1) har «klasselæreren», som kan undervise i en bredde av fag gjennom hele grunnskolen, vært en bevisst politisk prioritering, på bekostning av faglig fordypning. Siden mellomkrigstiden har det gjentatte ganger blitt diskutert å spesialisere allmennlærerutdanningen. Som påpekt i introduksjonskapittelet, anførte man i Stortingsmelding nr. 48 (1996-97) tre argumenter for å beholde en bred allmennlærerutdanning: 1) Allmennlærerens rolle som klassestyrer; 2) allmennlærerens rolle i å bidra til helhet og sammenheng mellom hovedtrinnene i grunnskolen; og 3) hensynet til de mange fådelte eller små barne- og ungdomsskoler med et lite lærerkollegium (jf. kapittel 2.1.3).

Den første begrunnelsen er et eksempel på hvordan lærerens rolle som fagperson i Norge tilsynelatende er underordnet lærerens sosiale og praktiske ansvarsoppgaver. Departementet presenterte i stortingsmeldingen ingen argumentasjon for behovet for «helhet og sammenheng mellom hovedtrinnene», hvilket gjør denne begrunnelsen vanskelig å diskutere. Den tredje begrunnelsen, om at man må ta hensyn til behovet for lærere med bred fagkompetanse på skoler med et lite lærerkollegium, innebærer at lærerens fagkompetanse i landet som sådan må vike for distriktpolitiske hensyn. Jeg stiller spørsmål om det er rimelig å la hensynet til distriktene være styrende for utformingen av allmennlærerutdanninger for hele landet. Sammenligningen med lærerutdanningen i Queensland anskueliggjør flere fordeler ved å dele den norske allmennlærerutdanningen i to separate løp for barne- og ungdomstrinnet, og at lærerne i sistnevnte program kun får undervisningskompetanse i to fag.

Som påpekt av Haug (2008) har det i norsk allmennlærerutdanning pågått en strid mellom skolefagene og pedagogikkfaget (jf. kapittel 3.7.4). Allmennlærerutdanning framholdes som en motsetning til faglærerutdanning, og implisitt hevdes det gjerne at en spesialisering mot fag vil innebære en nedprioritering av pedagogikkfaget og de mer «allmenne» sidene ved læreryrket. Sammenligningen med Queensland viser tvert imot at en spesialisering mot fag og alderstrinn gir bedre rammer for denne typen fag, rett og slett fordi man har mer tid. Ved å bruke Queenslands modell med sterk spesialisering mot fag, kunne

man sørget for at alle lærerstudentene fikk 50 prosent mer opplæring i pedagogiske fag, enn den obligatoriske delen av dagens rammeplan legger opp til.

Som vist i teorikapitlet har alle våre nordiske naboland gått mer i retning av spesialisering mot fag og alderstrinn (jf. kapittel 3.7.3). Kunnskapsminister Bård Vegard Solhjell og forsknings- og høyere utdanningsminister, Tora Aasland, har varslet at dette vil bli foreslått i den kommende stortingsmeldingen om allmennlærerutdanningen (Aasland og Solhjell, 6. desember 2008). Det er grunn til å forvente stor enighet om at man bør gå i denne retningen, og at debatten i stedet vil dreie seg om hvor mange undervisningsfag ungdomsskolelærerne skal få undervisningskompetanse i, hvilken grad det skal legges opp til overlapping mellom ulike trinn og hvilke opplæringstrinn som skal dekkes av lærerutdanninger på høyskolene og universitetene.

I Sverige har man, som i Queensland, en utdanning rettet både mot ungdomstrinnet og gymnaset, mens utdanningen i Finland er delt mellom 1. - 6. trinn og 7. - 9. trinn. I Danmark har man en mellomløsning der alle lærere blir utdannet til mellomtrinnet (4. - 6. trinn) for deretter å velge spesialisering mot barne- eller ungdomstrinnet. Det er grunn til å forvente at regjeringen vil foreslå en modell med to separate løp, et rettet mot barnetrinnet (1.-7.) og et mot ungdomstrinnet, men at sistnevnte også vil gi undervisningskompetanse både på mellom- og ungdomstrinnet (5. - 10.).

Med en slik løsning vil det trolig ikke være mulig å begrense utdanningen for ungdomstrinnet til to fag. Selv om man et stykke på vei bøter på problemene denne studien har avdekket, så vil det ikke gi like godt utdannede faglærere som de har i Queensland. Studien anskueliggjør mulige fordeler ved en langt mer vidtrekkende reform av allmennlærerutdanningen, der utdanning av ungdomsskolelærere og lærere for videregående skole samordnes. Med en slik modell kunne man lagt utdanning av ungdomsskolelærere til universitetene, og slik knyttet profesjonsopplæringen mer opp mot fagkretsene de tilhører i universitetene. Det ligger utenfor denne oppgaven å drøfte en slik løsning nærmere, men på bakgrunn av erfaringene med modellen i Queensland er det grunnlag for å anbefale å vurdere en slik ordning.

Med utgangspunkt i Queensland kan man imidlertid reise spørsmål om man bør ha noen nasjonal modell overhodet. Som påpekt i introduksjonskapitlet er ikke lærerutdanningen i Queensland regelstyrt, det er opp til universitetene å utforme sine egne lærerutdanningsprogrammer (jf. kapittel 2.2.3). I stedet har man omfattende prosedyrer for akkreditering og kvalitetssikring av lærerutdanningene. Queensland har i en årrekke vært verdensledende i arbeidet med akkreditering av lærere, og særlig de helt siste

utviklingstrekkene knyttet til dette er interessante. Siden 2006 har lærerprofesjonen selv, gjennom en uavhengig profesjonsorganisasjon, finansiert og håndhevet akkrediteringsprosedyrene både for lærere og lærerutdanning. Slik har lærerne som yrkesgruppe oppnådd en høyere profesjonsstatus.

Nå er det ikke mulig å overføre erfaringene i Queensland direkte til Norge, og å bryte med tradisjonen for at rammeplanen for lærerutdanningen fastsettes av nasjonale myndigheter. Men det går an å innføre nasjonale rammer som gir rom for et visst mangfold i organiseringen av lærerutdanningen, og det er uansett gode grunner for at innholdet i lærerutdanningenes fag, ikke bare omfanget av dem, kvalitetssikres av en uavhengig institusjon. Slik kan man blant annet sikre at alle lærerstudenter får tilstrekkelig opplæring i elevvurdering.

6.2 I hvilket omfang og på hvilken måte møter studentene temaet elevvurdering i lærerutdanningen?

Intervjuanalysen viste at lærerstudentene i Queensland fikk betydelig mer teoribasert opplæring i elevvurdering enn de norske allmennlærerstudentene. Den mest markante forskjellen knyttet seg til faget *Assessing Student Learning* (7,5 studiepoeng) som de hadde i Queensland, og som det ikke var noen norsk parallell til. Å ha elevvurdering som et eget fag bidrar til å løfte dette fram som en sentral del av lærerutdanningen i Queensland, slik det til sammenligning er gjort med faget *Tilpasset opplæring* på den norske lærerutdanningen. Til forskjell fra faget *Tilpasset opplæring*, er *Assessing Student Learning* obligatorisk del av lærerutdanningen i Queensland. Studentene ga uttrykk for at de hadde lært mye om elevvurdering i dette faget, og det kan framholdes som en av de viktigste faktorene for bildet som tegner seg av at lærerutdanningen i Queensland vektla elevvurdering betydelig mer enn den norske allmennlærerutdanningen gjorde.

Som vist ble elevvurdering også mer vektlagt i undervisningsfagene i Queensland enn i Norge. Dette kan også sees i sammenheng med at utdanningen i Queensland er rettet mot *secondary*. I Queensland får elevene karakterer fra første klasse, men det er først mot slutten av *secondary* lærernes vurderinger er «high stakes». I Norge er utdanningen som helhet rettet mot hele grunnskolen, og i barneskolen har vi som kjent ingen karaktervurdering. Nå har riktignok den norske delen av studien dreid seg om studenter som hadde valgt fordypning mot ungdomstrinnet, men det ser ut til at de summative aspektene ved elevvurdering er lite tilstede også i disse fagene. I Queensland var elevvurdering en mer strukturert del av opplæringen i

undervisningsfagene, ved at studentene måtte utarbeide undervisningsplaner, der elevvurdering utgjorde en viktig del.

Pensumlitteratur om elevvurdering er også blant områdene forskjellene mellom Norge og Queensland er svært store. Mens ingen av studentene i Norge hadde hatt noen bøker der elevvurdering var hovedtema, hadde de i Queensland hatt en lærebok om elevvurdering kalt *Celebrating Student Achievement* (Brady og Kennedy, 2005) og et kompendium med artikler satt sammen av læreren. Begge var pensum i faget *Assessing Student Learning*. Læreboken var på 140 sider, mens kompendiet var på 250 sider. Studentene ga uttrykk for at både læreboken og kompendiet var nyttig. De nevnte også andre bøker som de ikke husket navnene på. Disse har det derfor ikke vært mulig å undersøke nærmere.

En kort gjennomgang av bøkene de norske studentene trakk fram gir følgende resultat med hensyn til omfang:

- *Ord på nye spor* (Dysthe, 1999, s. 18, 58): 30 sider om elevvurdering
- *Meningsfylt matematikk* (Geir Botten): 20 sider om elevvurdering
- *Kan læring planlegges* (Engelsen, 2006): 30 sider om elevvurdering
- *Fagdebattikk* (Sjøberg, red, 2001): 2 sider om elevvurdering i kapittelet om *Matematikk* ((Brekke og Gjone, 2001, s. 231).

Det må tas forbehold om det svake datagrunnlaget for en analyse av denne litteraturen, ettersom studentene i begge land nevnte noen bøker de ikke husket navnet på. Men på bakgrunn av litteraturen som er identifisert kan man klart konkludere med at lærerstudentene i Queensland har hatt flere ganger mer pensumlitteratur om elevvurdering enn de norske lærerstudentene. Mens vi i de norske bøkene som er rapportert finner i underkant av 100 sider, hadde de i Queensland hatt i underkant av 400 sider litteratur. Disse forskjellene må man kunne si er som forventet, på bakgrunn av faget de hadde hatt om elevvurdering i Queensland.

Foruten et mye større omfang av pensumlitteratur om elevvurdering tilknyttet faget *Assessing Student Learning*, kan det større omfanget av temaet i lærerutdanningen i Queensland enn i Norge forklares med at elevvurdering er sentralt i læreplanene. Som påpekt i introduksjonskapitlet inneholder læreplanen for *Year 1-10* åtte sider om elevvurdering, mens vurderingskapitlet i læreplanene for *Year 11 & 12* for *English* og *Mathematics B* er på henholdsvis 20 og 10 sider (jf. kapittel 2.2.2). Til sammenligning inneholder læreplanverket for L97 og Kunnskapsløftet *ingen* sider om elevvurdering (KUF, 1997; Kunnskapsdepartementet, 2006). Dette innebærer trolig at elevvurdering utgjorde en langt mer sentral del av læreplananalysen som gjennomføres i teoriopplæringen i Queensland.

Tilsvarende har lærerne i Queensland trolig mye mer fokus på elevvurdering, noe lærerstudentene hadde utbytte av gjennom praksisopplæringen.

I Norge må som nevnt lærere og lærerstudenter gjøre seg kjent med forskrift til opplæringsloven, for å få kjennskap til de nasjonale retningslinjene for elevvurdering. Det er grunn for å tro at terskelen er mye større for å gjøre seg kjent med retningslinjene når de er plassert i forskriften, enn om de var en del av læreplanene. En mulighet for å gjøre retningslinjene mer tilgjengelig, kunne være å inkludere de viktigste av de til enhver tid gjeldende forskriftsbestemmelsene i læreplanene for Kunnskapsløftet, som er gjort tilgjengelig på nettet.

Som presentert i introduksjonskapittelet er en betydelig forskjell mellom de to utdanningssystemene, at vi i Norge har eksamener både på 10. trinn og videregående trinn, mens de i Queensland ikke har hatt noen statlige eksamener siden 1972 (jf. kapittel 2.3). De norske lærerstudentene så imidlertid ikke ut til å ha fått særlig kjennskap til eksamenssystemet. Introduksjonen av nasjonale prøver i 2004 og 2005 førte til mye støy i media, hvilket kan være en av grunnene til at flere av studentene nevnte at de hadde hatt litt om dette. En av studentene fortalte at nasjonale prøver var tema i en av timene i *Matematikk*, og at de der både hadde snakket om hvorfor «hvorfor man gjør det så dårlig» og gått igjennom og diskutert i grupper hvilke skår man skulle gi de forskjellige elevsvarene. En annen hadde blitt kjent med kriterier for lesing, mens en tredje mente han hadde fått utlevert prøvene, men at han ikke hadde sett noen kriterier for hvordan man vurderer dem.

Det er bemerkelsesverdig at de nasjonale prøvene, som var helt nye da disse studentene tok utdannelsen, fikk relativt mye oppmerksomhet sammenlignet med avgangsprøvene og eksamene, som de selv har hatt på skolen og som tradisjonelt har vært noe lærerne må forholde seg mer til. I Queensland hadde til sammenligning studentene fått lære om det nye initiativet QCAR, som også blant annet inneholder statlige prøver. Der hadde de blant annet hatt policy-dokumenter om dette som pensum i faget *Assessing Student Learning*. Det er nærliggende å tro at nasjonale prøver og eksamenssystemet ville fått en mer grundig behandling om man hadde hatt et eget fag om elevvurdering i Norge.

6.3 Hva slags teori- og praksisopplæring fikk studentene i utvalgte temaer knyttet til elevvurdering?

I denne delen drøfter jeg måten studentene i Norge og Queensland møtte temaet elevvurdering gjennom lærerutdanningen, og peker på forskjeller knyttet til hvorvidt opplæring på høyskolen eller universitetet på den ene siden, og praksis på den andre siden, hadde gitt

studentene kunnskaper om og erfaringer med elevvurdering knyttet til særskilte temaer. Det globale spørsmålet om hvor de hadde lært mest om elevvurdering (gjennom teori- eller praksisopplæring) drøftes som en del av den overordnede sammenligningen mellom Norge og Queensland og mellom fagene til slutt.

6.3.1 Opplæring i bruk av verktøy og hjelpemidler for elevvurdering

Som nevnt både i metode- og analysekapittelet (jf. kapittel 4.8.3, 4.9.3 og 5.3.2) er datagrunnlaget svakere for spørsmålet om vurderingsverktøy og hjelpemidler, som følge av at flere av studentene hadde problemer med å forstå hva jeg mente med spørsmålsstillingen. Dette kan være et uttrykk for svakheter i begrepsbruk og spørsmålsstillingen i intervjuguidene. På den annen side kan det være en indikasjon på at begge lærestedene i liten grad vektla opplæring i bruk av verktøy og hjelpemidler for elevvurdering, og at studentene derfor hadde problemer med å svare på spørsmålet. I Queensland skilte *Mathematics* seg ut ved at studentene fikk trening i å lage vurderingsoppgaver og siden utvekslet disse på en CD. En av studentene med fordypning i matematikk hadde også fått en CD med vurderingsoppgaver sammen med læreboka i faget. Samlet sett må man likevel konkludere med at verktøy og hjelpemidler for elevvurdering ikke er sentralt i noen av lærerutdanningene.

Man skulle tro at det var sentralt i lærerutdanningens mandat å gi lærerstudenter trening i bruk av vurderingsverktøy og hjelpemidler, både for summative og formative formål. Det være seg oppgavebanker, vurderingsskjemaer, testproduksjonssoftware eller lignende. Mangelen på slike verktøy og hjelpemidler for elevvurdering i Norge, kan sees i sammenheng med hvordan metodiske instrumenter generelt er blitt angrepet i norsk skole. Som anført tidligere har Klette (2008) pekt på hvordan metodikkompetanse og opplæring i konkrete verktøy har blitt sett på som «antitesen» til reflektert profesjonskunnskap og et uttrykk for det Skjervheim kalte et «instrumentalistisk mistak». Klette mener denne koblingen er misforstått. Dale (2008) peker også på hvordan man i Norge har «styrt unna» teknisk kunnskap, med henvisning til at det er instrumentalistisk. Det er grunn for å reise spørsmål om dette tankegodset er noe av forklaringen på det tilsynelatende fraværet av opplæring i bruk av verktøy og hjelpemidler for elevvurderingen i den norske allmennlærerutdanningen som ble undersøkt.

Ettersom heller ikke studentene i Queensland hadde fått noe særlig opplæring i verktøy og hjelpemidler for elevvurdering, kan det imidlertid se ut som at dette er et mer generelt problem ved utdanning av lærere. Det kan virke som at holdningen fra begge lærerutdanningsinstitusjonene var at dette er noe man må lære seg når man er ute i praksis

eller når man begynner å jobbe. Det er overraskende at praktiske hjelpemidler blir gitt så liten plass, ettersom dette burde være svært konkret lærestoff som studentene lett ser nytteverdien av. Lærerutdanningene så ut til å i liten grad dra nytte av utviklingen knyttet til IKT for å styrke lærerstudentenes vurderingskompetanse.

Det er grunn for å reise spørsmål om hvorfor slik teknologi ikke ble benyttet i særlig grad ved noen av lærestedene. Ser vi litt utover våre case, er nettopp dette området elevvurderingsfeltet i stor utvikling. Som eksempel kan her nevnes New Zeland's program *Assessment and Tools for Teaching and Learning* (AsTTle), der lærere har tilgang på en oppgavedatabank med prøver og andre oppgaver med tilhørende vurderingskriterier- og standarder, for vurdering av elevers læringsresultater i forskjellige fag og på forskjellige alderstrinn (TKI, 2009). Skottlands program *Assessment is for Learning* (AifL) er utviklet for å bidra til å utvikle et tolkningsfellesskap på tvers av skoler og klasserom. Det er blant annet utviklet flere verktøy for å gjennomføre intern moderering (LTS, 2008). Som en del av utviklingsprosjektet QCAR (jf. kapittel 2.2.2) pågår det nå i Queensland et arbeid med å utvikle en oppgavebank for de mest sentrale fagene (QSA, 2008a).

I Norge kan Utdanningssetaten i Oslo sitt prosjekt *Skriflig vurdering på barnetrinnet*, som inngår i satsingen *Vurdering for læring* (VFL), trekkes fram som eksempel på utviklingsarbeid som pågår i flere kommuner med bruk av vurderingsskjemaer til å gi elever tilbakemelding på læringsarbeidet (Utdanningssetaten, 2009). Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis* (jf. kapittel 2.1.2) omfatter 77 skoler fordelt over hele landet (Utdanningsdirektoratet, 2009). Det pågår altså en del utviklingsarbeid, og det er grunn til å tro at det vil komme flere slike utviklingsprosjekter, som gjennom IKT-baserte løsninger kan gjøres tilgjengelig for alle. Det er viktig at lærerutdanningene aktivt følger opp dette utviklingsarbeidet, slik at de kan gi lærerstudentene oppdatert kunnskap om tilgjengelige verktøy og hjelpemidler for elevvurdering.

6.3.2 Opplæring i bruk av vurderingskriterier

Analysen viste at lærerstudentene i Norge hadde fått liten trening i bruk av vurderingskriterier, både på høyskolen og i praksis. I Queensland hadde studentene lært litt om bruk av vurderingskriterier i faget *Assessing Student Learning*. Vurderingskriterier var også en viktig del av praksis i Queensland, først og fremst fordi det er en viktig del av læreplanen og følgelig alt vurderingsarbeidet som foregår i skolen. Damion pekte på et interessant poeng i forhold til erfaring med vurderingskriterier som jeg ønsker å drøfte nærmere:

I think just from being in uni, and having criteria myself, so like for any particular subject and on assignment you get a criteria sheet, and seeing different subjects, some were really good, some were really bad (Damion-intervju, s. 12).

Han hadde en bestemt oppfatning om hva slags kriterier som fungerte bra og dårlig, slik de hadde blitt presentert for han i fagene på universitetet. Han forklarte nærmere hva han mener kjennetegner dårlige kriterier:

Like for one criteria sheet an A said... (...) The column A was «Excellent», B was «Good», C was «Adequate», so like that didn't explain anything (laughing). And then from my experience that was really useless, then I could sort of see what criteria sheets I found to be useful and effective and use that as well (Damion-intervju, s. 12).

Damion illustrerer her hvor intetsigende vurderingskriterier kan være. Han minner oss på hvordan man med utgangspunkt i egne erfaringer kan få innsikt i hva som er god vurderingspraksis. Følgelig er lærestedenes vurderingspraksis av stor betydning. Trygve fortalte at han ikke hadde lært noe om bruk av kriterier «annet en hvordan vi har fått karakterer selv da» (Trygve-intervju, s. 10). Jeg ønsket derfor å se litt nærmere på hva de norske lærerstudentene kunne lære gjennom vurderingspraksisen som gjorde seg gjeldende på høyskolen.

Sammenligner man vurderingskriteriene som benyttes ved den norske høyskolen og universitetet i Queensland (henholdsvis appendiks 5 og 6), så trer det fram store forskjeller. I begge planene angis vurderingsoppgaver, men i Queensland er man langt mer spesifikke i kravspesifikasjonene. Det kan synes som at det er store kulturforskjeller, ikke bare knyttet til opplæring i elevvurdering, men også hvordan lærerestedene praktiserer det selv. Vektleggingen av vurderingskriterier er mye større i Queensland enn i Norge, både i læreplanene for alle alderstrinn og i undervisningsplanene for lærerutdanningen. Jeg vil hevde at lærere i Queensland som følge av dette får betydelig mer utbytte av egne erfaringer med å bli vurdert, enn de norske lærerstudentene.

6.3.3 Erfaringer med diskusjoner om bruk av kriterier

Som påpekt i teorikapitlet (kapittel 3.2.1) er diskusjon og erfaringsdeling om bruk av kriterier en nødvendig forutsetning for å utvikle et tolkningsfellesskap. Rettferdig karaktergivning, hviler derfor på at bruken av vurderingskriterier er forankret i et kollegialt samarbeid innenfor

og på tvers av skoler. Analysen viste at studentene i Norge ikke fikk noen erfaring med dette, verken på høyskolen eller gjennom praksis. I Queensland bidro heller ikke universitet vesentlig til diskusjoner om vurderingskriterier. Gjennom praksis derimot, fikk studentene erfaring både med interne og eksterne modereringsprosedyrer. Det er betenkelig at de norske studentene verken hadde fått kjennskap til eller erfaring med noen prosedyrer for kvalitetssikring av lærernes karaktersetting, all den tid standpunkt karakterene er like «high stakes» som karakterene lærerne i Queensland gir. Slike prosedyrer måtte naturligvis i all hovedsak studentene møtt gjennom praksis. Fraværet av diskusjoner om bruk av kriterier på allmennlærerutdanning bør derfor i stor grad tilskrives en svak vurderingskultur i Norge generelt (jf. avnitt 3.4), og ikke bare allmennlærerutdanningene.

6.3.4 Opplæring i og erfaringer med karaktersetting

Verken i Norge eller Queensland hadde studentene fått noen erfaring med karaktersetting gjennom høyskolen eller universitetet. I Queensland fikk studentene en del trening i karaktersetting som en del av praksis. Det så ut til å være etablert rutiner for at studentene skal få trening i dette. På universitetet var temaet imidlertid helt fraværende. I Norge kan det se ut som at studentene som hadde praksis i ungdomsskolen fikk noe trening med karaktersetting. I barneskolen vil imidlertid retting av prøver være begrenset til å gi tilbakemelding på hva som var rett og galt, ettersom man ikke har lov til å gi karakterer.

Man kan spørre seg hvorfor lærerutdanningene, verken i Norge eller Queensland, tilsynelatende ikke ser på det som deres oppgave å gi studentene trening i karaktersetting. Summativ vurdering utgjør en svært viktig del av lærernes ansvarsoppgaver knyttet til elevvurdering (jf. kapittel 3.1). På linje med bruk av autentiske eksempler i forbindelse med undervisningsopplegg (undervisningsplaner, lærebøker etc), er det mulig å benytte autentiske eksempler på elevarbeider, som studentene får prøve seg på å sette karakter på. Her vil jeg igjen trekke fram arbeidet som pågår i blant annet i New Zealand, Skottland og Queensland (jf. kapittel 6.3.2), med blant annet utvikling av oppgavebanker med tilhørende kriterier og ordninger for karaktersetting. Dersom man utvikler lignende i Norge, vil det være svært enkelt å gjøre slike oppgaver tilgjengelig for høyskolene slik at de kan gi studentene trening i karaktersetting.

6.3.5 Opplæring i og erfaringer med å gi faglige tilbakemeldinger

Mens studentene i Norge ga uttrykk for at de hadde fått en del teoriopplæring om hvordan man bør gi tilbakemeldinger på lærerutdanningen, så fortalte studentene i Queensland at de

knappt hadde lært noe om dette på universitetet. Det kan se ut som at man i Queensland hadde et mer summativt fokus enn i Norge. Dette kan henge sammen med at vurderingskulturen i Queensland er mer summativ orientert enn i Norge, som følge av karaktervurdering gjennomføres fra første klasse. Tilsvarende kan det synes som at vi i Norge har en sterkere tradisjon for formativ vurdering enn man har i Queensland, nettopp som følge av at summative vurderinger har blitt sterkt kritisert og er forbudt på barnetrinnet (jf. kapittel 3.4.1).

De norske studentene hadde imidlertid sprikende oppfatninger om hva som kjennetegner gode formative vurderinger. Tre av studentene framholdt at de hadde lært «å rette med grønt». De forklarte dette uttrykket, som de hadde blitt kjent med gjennom en lærebok, med at man i stedet for å bruke den tradisjonelle røde korrigerende pennen bør gi mer positive «grønne» tilbakemeldinger. Dette uttrykket har studentene fra en av lærebøkene i *Matematikk*, skrevet av Geir Botten, som jeg har gjengitt i teorikapitlet (jf. kapittel 3.6.2). Botten er påpasselig med å peke på at «'retting' med grønt» ikke betyr at en er «snillere» med elevene eller lettere lar unøyaktigheter og slurv passere. Han peker også på at at man må kommentere det som absolutt ikke fortjener ros. Hovedperspektivet, forsterket av en virkningsfull metafor, er likevel en kritikk av den tradisjonelle «rettingen» og at man i stedet bør gi mer positive tilbakemeldinger. Det er derfor ikke rart om flere av studentene satt igjen med en forståelse av at budskapet var at man bør gi «positiv feedback hele tiden» (jf. kapittel 5.3.6).

Det er godt dokumentert at generell ros og oppmuntring har liten effekt. Som påpekt i teorikapitlet (jf. kapittel 3.4.3) har Klette (2003) tidligere kritisert norske lærere for å i for stor grad gi slike tilbakemeldinger. To av studentene trakk fram nettopp dette, og var tydeligvis blitt gjort oppmerksom på denne kritikken, mens dette ikke var tilfellet for de to andre. Ettersom tre av fire studenter hadde merket seg uttrykket «'retting' med grønt» og ga uttrykk for at noe av det viktigste budskapet i pensumlitteraturen om elevvurdering var at man skal gi positive tilbakemeldinger, mener jeg det er grunnlag for å hevde at vurderingspraksisen Klette kritiserte fremdeles har et visst fotfeste. Dette er overraskende, ettersom dette er en svært kjent kritikk fra evalueringen av den forrige grunnskolereformen. En mulig forklaring kan være at man i Norge, i mangel på felles kriterier og standarder for måloppnåelse, har dårlige forutsetninger for å gi «eksplisitte produktkrav og standarder», slik Klette har etterlyst.

6.3.6 Opplæring i forskrifter og retningslinjer

I Norge trer det fram en overraskende forskjell knyttet til opplæring i forskrifter og retningslinjer for elevvurdering. Olav og Eli fortalte at dette temaet hadde blitt tatt opp i

matematikkopplæringen på høyskolen, mens Camilla og Trygve ga uttrykk for at de ikke hadde lært noe om dette i norskfaget, verken på høyskolen eller gjennom praksis. Dette kan like gjerne henge sammen med enkeltlærere på høyskolens prioriteringer. Jeg har ikke funnet holdepunkter i teori eller det øvrige intervjudata til å hevde at dette kan knyttes til systematiske forskjeller mellom disse fagene i Norge.

I Queensland hadde alle studentene blitt kjent med retningslinjer for elevvurdering gjennom studier av læreplanene i opplæringen i fagene. Noen av dem hadde også satt seg inn i forskjellene mellom gamle og gjeldende læreplaner. Studentene hadde dessuten, gjennom faget *Assessing Student Learning*, i tillegg fått kjennskap til de nye retningslinjene som vil komme i forbindelse med QCAR (jf. kapittel 2.2.2). Den sentrale posisjonen elevvurdering har i læreplanene i Queensland, er trolig den viktigste årsaken til at studentene der ble kjent med retningslinjer for elevvurdering både gjennom opplæring på universitet og gjennom praksis.

Som drøftet i teorikapitlet (jf. kapittel 3.4.1) er det er grunn til å problematisere uklarhetene i retningslinjene for elevvurdering, som enkelte av studentene hadde fått gjennomgått. Urealistiske politiske ambisjoner om å inkludere «alt» i karaktervurderingen, har ført til en situasjon der vurderingsgrunnlaget har framstått som uklart, hvilket har banet vei for illegitim vektlegging av innsats og trynefaktor, og for mye vektlegging av muntlig aktivitet i timen (jf. kapittel 3.4.2). Olav så ut til å være del av en slik vurderingskultur. For å sikre en så rettferdig karaktersetning som mulig, var han opptatt av at er «å ha en oversikt over hvem som svarer i timen omtrent (...) når du har hatt en time, så tenk igjennom hvem som faktisk har sagt noe i timen» (Olav-intervju, s. 15).

Manglende opplæring i forskrifter og retningslinjer for elevvurdering i allmennlærerutdanningen kan imidlertid i like stor grad knyttes til svakheter i regelverket i seg selv. De siste to årene har det blitt gjennomført omfattende endringer i dette regelverket, og flere endringer er ventet som ledd i Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis*¹². For allmennlærerutdanningene blir det viktig å følge denne utviklingen, og gi lærerstudentene en oppdatert innføring i forskrifter og retningslinjer for elevvurdering. Som nevnt bør man kan hende vurdere å inkludere de viktigste bestemmelsene i de enkelte læreplanene for fagene i Kunnskapsløftet (jf. kapittel 6.2).

¹² Utdanningsdirektoratet har sendt nye forskriftsbestemmelser om elevvurdering ut på høring, som trolig vil bli gjeldende fra høsten 2009 (Utdanningsdirektoratet, 2008a). Det ligger utenfor denne oppgaven å drøfte disse forslagene nærmere.

6.4 Forskjeller mellom morsmåls- og matematikkfaget

Som nevnt innledningsvis var en av grunnene for at jeg valgte å sammenligne morsmålsfaget med matematikkfaget, at sistnevnte baserer seg på flere objektive regler, og derfor av mange regnes som et fag der det er enklere å gjøre pålitelige vurderinger av elevenes læringsresultater. Nå er ikke dette en hypotesetestende studie, og med bare åtte informanter er det ikke mulig å trekke generelle slutninger. Hensikten var i stedet å undersøke denne oppfatningen nærmere, og generere teorier for senere undersøkelser. I så måte kan studien bidra til å anskueliggjøre hva slags oppfatninger enkelte lærerstudenter har om temaet, og jeg mener disse oppfatningene gir grunnlag for å utforske denne problemstillingen nærmere.

I Norge ga lærerstudentene inntrykk av at man i morsmålsfaget viet mer oppmerksomhet til hvor vanskelig karaktersetning er, sammenlignet med matematikkfaget. Jeg vil trekke fram et sitat fra Camilla, den eneste studenten i utvalget som hadde hatt fordypning både i *Norsk for ungdomstrinnet* og *Matematikk for ungdomstrinnet*, for å vise ulike tilnærminger til summativ vurdering i de to fagene. Hun fortalte at hun gjennom praksis i norskfaget hadde fått et skjema som beskrev forskjellige kriterier:

Da viste hun oss et skjema etter rettskriving, helhet, innhold, sånne ting (...) Så da fikk vi vel fire-fem sånne kategorier og vi satt enkeltkarakterer på en hel stil, og så var det å legge de sammen som ble sluttresultatet. Det var ganske nyttig, for akkurat det å sette karakter på et norsk skriftlig arbeid, det er ikke så veldig lett (Camilla-intervju, s. 8).

Camilla ga uttrykk for at det var nyttig å få trening i bruk av vurderingskriterier, ettersom hun syntes det var vanskelig å sette karakter på norsk skriftlig arbeid. Hun fortsatte med å trekke en interessant sammenligning til matematikkfaget: «I Matten derimot er det jo rett og slett bare å telle feil, som vi i hvert fall er blitt lært opp til, men gjerne skrive en skriftlig kommentar som oppmuntrer litt på slutten» (Camilla-intervju, s. 8).

I Queensland kunne man se lignende perspektiver. Damion mente han hadde fått mer erfaring med bruk av vurderingskriterier i forbindelse med vurdering av eksperimentrapporter i *Science*, og forklarte dette med «fagets natur»:

I think because the assessment techniques in Maths were more exam based, where yes, it is less subjective, like you just have to do, for example add fractions together, whereas in *Science* you might have to explain a process or show your understanding rather than actually just do something, like to show your understanding, so that's where the criteria kicked in more (Damion-intervju, s. 14).

Det kan virke som at verken Camilla og Damon hadde blitt gjort oppmerksom på at selv om det er enklere å «telle feil» i matematikkfaget, så er det ikke nødvendigvis så lett å fastsette karakternivå, ettersom elevenes prøveresultater må sees i sammenheng med prøvenes vanskelighetsgrad. Dessuten må man, som Botten (2003) påpeker, ta stilling til hvordan man skal vurdere følgefeil (jf. kapittel 3.6).

Trygve ga også inntrykk av at karaktersetting ikke ble presentert som en særlig krevende øvelse i opplæringen i *Matematikk* på høyskolen. Han fortalte at hans matematikkforeleser påsto at «en god matematikklærer trenger egentlig ikke ha noen prøver for å kunne bestemme nivået på elevene, det kan gå med daglig interaksjon og så videre» (Tryge-intervju, s. 13). Har man en slik innstilling undervurderer man fullstendig kompleksiteten knyttet til å gjøre pålitelige vurderinger i matematikkfaget. Det må karakteriseres som oppsiktsvekkende at enkelte lærere på denne allmennlærerutdanningen bidrar til å overføre slike holdninger til kommende lærere.

Til sammen tegnet studentene et bilde av karaktervurdering i matematikkfaget som en mekanisk og enkel øvelse, mens de i morsmålsfaget ble eksponert for karaktersetting som en mye mer problematisk og krevende oppgave. Jeg mener det er betenkelig om man kan gå ut av lærerutdanning og tro at det er lett å gjennomføre en pålitelig karaktervurdering i *Matematikk*. Samtidig ønsker jeg å framheve at man i morsmålsfaget ser ut til å ha fått klart fram hvor vanskelig karaktersetting er. I Norge kan problematiseringen av elevvurdering i norskfaget sees i sammenheng med det omfattende arbeidet som er gjort knyttet til pålitelighetsvurdering av sentralt gitte eksamener (jf. kapittel 3.4.2).

Både i Norge og Queensland så det ut til at det ble gitt mer opplæring i elevvurdering i morsmålsfaget enn i matematikkfaget. Elevvurderingens relativt sterke posisjon i *Norsk*, sammenlignet med *Matematikk*, kan sees i sammenheng med den sterke vektleggingen av prosessorientert skriveopplæring og mappevurdering (jf. kapittel 3.6.1).

6.5 Overordnet sammenligning av opplæring i elevvurdering i Norge og Queensland

Det var store forskjeller mellom de organisatoriske rammene for de to lærerutdanningene. Det er liten tvil om at forholdene ligger mye bedre til rette for en grundigere opplæring i både fag, pedagogikk og elevvurdering i Queensland, som følge av at utdanningen er spesialisert mot *secondary* og begrenset til to undervisningsfag. Det er på denne bakgrunn helt urimelig å ha de samme forventningene til norsk allmennlærerutdanning om opplæring i elevvurdering. Jeg mener funnene anskueliggjør flere negative følger å politikken på å utdanne *allmennlærere*.

Det er viktig å peke på at også andre pedagogiske «nisjefelt» kan tape på denne brede tilnærmingen, rett og slett i mangel på tid.

Analysen viste at de norske studentene hadde ulik oppfatning av hvorvidt opplæringen på høyskolen eller praksis var viktigst for kunnskapstilegnelsen om elevvurdering. Enkelte av de norske studentene opplevde høyskolen som den viktigste arenaen, men forklarte dette med at de hadde fått lite utbytte av praksisopplæring i elevvurdering i denne perioden (blant annet som følge av skoleovertakelse, ekskursjon etc). Andre av de norske studentene følte de hadde lært litt om elevvurdering gjennom praksis. I Queensland ble både opplæring på universitetet og gjennom praksis pekt på som viktig. Praksis framstod der, sammen med faget *Assessing Student Learning*, som de viktigste arenaene for tilegnelse av kunnskaper om elevvurdering, mens noen også opplevde at de hadde lært en del i undervisningsfagene på universitetet.

I Queensland fikk de generelt svært mye mer opplæring i elevvurdering på universitetet, blant annet som følge av at de hadde et eget fag om elevvurdering med tilhørende pensumlitteratur og fordi elevvurdering utgjør en mye mer sentral del av læreplanene for fagene. Av samme grunn lærte de også betydelig mer om elevvurdering gjennom praksis. I tillegg bidro systemer for intern og ekstern moderering som studentene møtte gjennom praksis, til at de i Queensland fikk en del opplæring i *bruk av* og *diskusjon* om vurderingskriterier, og mer opplæring i karaktersetting.

Det eneste området det så ut til at de norske lærerstudentene hadde fått *mer* opplæring enn studentene i Queensland, var knyttet til faglige tilbakemeldinger. Til sammen tegnet studentene et bilde av Norge som svært formativt orientert, mens man i Queensland i større grad hadde et summativt fokus. Disse forskjellene var særlig tydelige i morsmålsfagene. I *Norsk for ungdomstrinnet* fortalte at de hadde lært en del om prosessorientert skriving, mappevurdering og tilbakemeldinger, mens *English*-studentene i Queensland hadde lært mer om utarbeiding av oppgaver, vurderingskriterier og karaktersetting.

På bakgrunn av hva vi vet om de to utdanningssystemene (jf. kapittel 2.1.2 og 2.2.2), er ikke denne forskjellen overraskende. Men selv om det isolert sett er gledelig at formativ vurdering blir sterkt vektlagt i norsk allmennlærerutdanning, mener jeg kunnskapene to av disse fire studentene presenterte, om at det er viktig å gi «positiv feedback hele tiden», er urovekkende. Det de norske allmennlærerstudentene hadde lært *mest* om, var det bare halvparten som hadde oppdatert kunnskap om.

Når jeg har trukket fram «retting med grønt» i tittelen på oppgaven, så er det fordi dette studentutsagnet illustrerer de store forskjellene mellom casene. Denne studentens nølende svar på spørsmålet om hva slags pensumlitteratur de hadde hatt om elevvurdering,

var betegnende for en rekke av forskjellene som kom til uttrykk gjennom intervjuene. Studentene i Queensland hadde gjennomgående mer å fortelle, som svar på mine spørsmål. Oppsummert var det to viktige grunner for dette:

1. Opplæringen gjennom faget *Assessing Student Learning* og tilhørende pensumlitteratur.
2. Vektleggingen av elevvurdering i læreplanene for fagene, herunder kriterier og standarder for karaktersetting.

Der de norske studentene svarte noe undrende på spørsmålene om vurderingskriterier, kunne studentene i Queensland i detalj berette om hvordan kriteriene og standardene for elevvurdering var utformet i læreplanene og hvordan man lager gode vurderingsoppgaver. Disse store forskjellene var forutsibare, som følge av Queensland lange tradisjon med standardbasert vurdering og Norges helt ferske tilnærming til kriteriebasert vurdering. Sammenligningen er således ikke til forkleinelse av de norske studentene som ble intervjuet, men en anskueliggjøring av hvilke store utfordringer man står overfor i Norge i arbeidet med å utvikle et standardbasert vurderingssystem der vurderingskriterier utgjør grunnlaget for både gode formative tilbakemeldinger, og valide og reliable summative karaktervurderinger.

7 Avsluttende kommentarer

Jeg har gjennom denne undersøkelsen identifisert en rekke svakheter og problemstillinger knyttet til teori- og praksisbasert opplæring av lærerstudenter i elevvurdering. Særlig i Norge framstår utfordringene som svært store. Det må naturligvis understrekes at en kvalitativ dybdeintervjustudie som denne overhodet ikke er egnet til å felle en dom, verken over denne høyskolens opplæringstilbud, eller norsk allmennlærerutdanning generelt. Jeg mener likevel at jeg har kunnet presentere interessante funn, som kan danne grunnlag for videre forskning.

Selv om allmennlærerutdanningenes obligatoriske fag har blitt stadig mer innskrenket med årene, så viser studien med all tydelighet konsekvenser av at man i 1997 slo fra seg ideen om et todelt løp hvorav det ene skulle være spesialisert mot ungdomsskolen og begrenset til noen få undervisningsfag. I forbindelse med den forestående revideringen av allmennlærerutdanningen mener jeg det bør legges vekt på forskningen som viser betydningen av faglig kompetente lærere og erfaringene med mer spesialiserte utdanninger, som i Queensland. Man trenger ikke reise lenger enn til Finland og Sverige for å finne interessante modeller. I lys av denne studien vil jeg framholde det som særlig interessant å undersøke den svenske modellen med en utdannelse rettet mot ungdomsskolen og gymnaset, samt at man begrenser lærernes undervisningskompetanse til to fag.

Samtidig mener jeg det er viktig at man tar Peder Haugs advarsel på alvor, når han sier at «dei reformpedagogiske ideane frå tidlegare har gitt ein skule som samfunnet ikkje lenger vil ha» (Haug, 2008, s. 273). Kan hende er det ikke nødvendigvis de reformpedagogiske ideene, men *reformene*, både av allmennlærerutdanningen og hele utdanningssystemet, som Haug også peker på, som må bære ansvaret for de store utfordringene i dagens allmennlærerutdanning. Systemet for akkreditering av lærerutdanninger i Queensland framstår som en mye mer forutsigbar ordning, som også bidrar til å sikre kvaliteten, og ikke bare kvantiteten av fagene i lærerutdanningene. Ved selv å ta ansvaret for akkrediteringsprosedyrene for både lærere og lærerutdanninger, har lærerne som yrkesgruppe oppnådd en høyere profesjonsstatus i Queensland. I Norge har profesjonsmedlemmene i blant andre lege- og advokatyrkene, gjennom sine organisasjoner, langt større mulighet til å ta ansvar for å opprettholde autoriteten og respekten for egen profesjon enn det lærerne har i dag. Kanskje kan Queensland være et eksempel til etterfølgelse, i tilstrebelserne for å «gjenreise respekten for læreryrket», som det stadig snakkes om?

Studien identifisert en rekke problemstillinger knyttet til den norske «vurderingskulturen». Den har som nevnt Kunnskapsdepartementet alt slått fast at er «svak»,

og mine funn gir på ingen måte grunnlag for å betvile den analysen. Studien er ikke egnet til å si noe om hvor omfattende problemene er, men jeg mener den kan bidra til å anskueliggjøre en rekke problemstillinger. Et paradoks avtegnet seg i teori- og analyse materialet, som gir grunnlag for å utfordre en etablert oppfatning:

Mange forskere jeg har vist til i denne oppgaven (Black og Wiliam, 1998; Deci, Koestner og Ryan, 1999; Hattie og Timperley, 2007) framholder at karakterer og belønninger i liten grad bidrar til, eller kan stå i veien for, forbedrede læringsresultater. Man sier gjerne at de summative vurderingene «overkjører» de formative vurderingene. Bildet som avtegnet seg i Queensland var helt annerledes, ved at kriteriene for summativ vurdering også dannet utgangspunktet for den formative vurderingen.

Haugstveit med flere (2006) pekte på fraværet av karakterer på barnetrinnet som en mulig årsak til at mange lærere ikke har god kjennskap til hvordan man gjennomfører målrelatert og kriteriebasert vurdering (jf. kapittel 3.4.3). Jeg mener det er holdepunkter for å hevde at en av grunnene til at elevvurdering generelt er lite tilstedeværende i norsk allmennlærerutdanning, skyldes fraværet av karakterer på barnetrinnet. Dette forbudet innebærer at det heller ikke er tillatt å bruke vurderingskriterier som angir nivåer av måloppnåelse. Innenfor en standard- og kriteriebasert tradisjon er det naturlig at også formative vurderinger forankres i vurderingskriteriene som legges til grunn ved karaktervurdering, slik at elevene kan forstå hva de må gjøre for å oppnå bedre læringsresultater. Således kan det hevdes at man i det formative vurderingsarbeidet i norske barneskoler, lider under fraværet av en summativ vurderingskultur, når man skal gjennomføre kriteriebasert vurdering.

Følgelig tror jeg de som har vært redde for en «snikinnføring» av karakterer på barnetrinnet til en viss grad har grunnlag for sin bekymring. Det kan etter hvert vise seg at i hvert fall noe som ligner på karakterer tvinger seg fram. Jeg vil reise spørsmålet: Hvordan skal man identifisere, uttrykke og forklare barn hva som kjennetegner kvalitet i fagene, *uten* bruk av nivåer? Erling Lars Dale argumenterer for at vi i Norge trenger «å utvikle en *kriterieorientert testtradisjon* som gjør det mulig å måle forbedringer i læringen» (Dale, 2008, s. 137). Det er nærliggende å spørre om vi ikke også bør måle forbedringer i barneskoleelevers læring?

Jeg tar ikke med dette til orde for bruk av karakterer som en form for «pisk og gullerot» mot våre barn. Som nevnt har begrunnelser for bruk av karakterer gått langs tre linjer: *Seleksjon*, *motivasjon* og *informasjon*. Seleksjon er den viktigste grunnen for at vi har karakterer i ungdomsskolen, videregående skoler og i høyere utdanning, men ikke en grunn

for å bruke karakterer på barnetrinnet. Motivasjon ble brukt som begrunnelse tidligere, og selv om det for noen har en motiverende funksjon er det godt dokumentert at det for flertallet av elever ikke utgjør en motivasjonsfaktor som bidrar til bedre læring. Med standard- eller kriteriebasert vurdering får imidlertid informasjonsargumentet gyldighet. Nettopp *informasjonsverdien* nivåbeskrivelser av elevarbeider kan ha for elever, lærere og foresatte, gjør at karaktervurdering også kan ha en formativ funksjon. Jeg mener det kan være hensiktsmessig å utvide utprøvingen av kjennetegn på måloppnåelse på barnetrinnet fra høy og lav måloppnåelse, til å inkludere flere standarder, for å vinne erfaring med dette.

Et slikt forslag kan man regne med at fører til sterke reaksjoner. Det offentlige ordskiftet om karakterer har i Norge stort sett vært begrenset til debatter for eller mot karakterer. Dette gjaldt ikke bare 1960- og 1970-årene. Norsk Lærerlag var, helt fram til de slo seg sammen med Lærerforbundet til Utdanningsforbundet i 1999, motstandere av karaktervurdering på ungdomstrinnet. Som gjengitt i teorikapittelet (jf. kapittel 3.4.1) var et av hovedargumentene grunnskolelærernes organisasjon brukte for motstanden mot karakterer på ungdomstrinnet at «(...) karakterene alltid [vil] være beheftet med den vilkårlighet som følger et system som i stor grad må bygge på skjønn». Få vil si seg uenig i dette. Men i stedet for å drøfte hvordan man kunne utøve dette skjønnet best mulig, hadde grunnskolelærernes organisasjon en enkel løsning på dette problemet: «vekk med karakterene».

Det er ingen logisk grunn for at de politiske og ideologiske formene debatten om karaktervurdering har hatt, skulle medføre at karaktersetting defineres helt bort fra praksisfeltet. Likevel kan det se ut som at det er nettopp det som har skjedd, og at skolen og lærerutdanningene etter dette «instrumentalismehysteriet» står igjen helt ribbet for teorier, konsepter, instrumenter og verktøy for å gjennomføre god elevvurdering. Det er ikke slik at karakterer, standarder og kriterier nødvendigvis leder til standardisering og instrumentalisme. Queensland er et godt eksempel på et utdanningssystem der et stort mangfold er ivaretatt innenfor et system for standardbasert vurdering med rigide ordninger for sikre valide og pålitelige vurderinger av elevenes måloppnåelse.

Uavhengig av hvordan man ender opp med å utforme de nye vurderingsordningene i Norge, er det viktig at det skjer i en forutsigbar form slik at både lærere og lærerstudenter kan få en grundig opplæring i hvordan de skal bruke det nye regelverket og de nye vurderingsordningene. Det mest oppsiktsvekkende funnet i denne studien er at verken i Norge eller Queensland hadde studentene fått noen trening i karaktersettingen på høyskolen eller universitetet. Det ble etterlyst av samtlige studenter. Forståelig nok, det er tross alt den «mektigste» delen av lærergjernen, og får enorme konsekvenser for elevenes framtid.

8 Referanser

- AMSI (2007). ICE-EM update. *Australian Mathematical Sciences Institute, Summer 2007*.
Hentet fra: http://www.ice-em.org.au/images/stories/FileDownloads/Newsletters/ICE-EM_Newsletter_SUMMER_2007.pdf.
- Berge, K. L. (1996). *Norsksensorenes tekstnormer og doxa. En kultursemiotisk og sosiotekstologisk analyse*. Trondheim: Norges Teknisk-Naturvitenskapelige Universitet.
- Berge, K. L., Evensen, L. S., Hertzberg, F. og Vagle, W. (2005). *Ungdommers skrivekompetanse. Norsksensuren som kvalitetsvurdering*. Oslo: Universitetsforlaget.
- Berger, P. L. og Luckmann, T. (2004). *Den samfunnsskapte virkelighet* (3. utgave). Bergen: Fagbokforlaget.
- Black, P. J. (2004). *The nature and value of formative assessment for learning*. London: King's College.
- Black, P. J. og Wiliam, D. (1998a). *Assessment and classroom learning. Assessment in Education: Principles, Policy and Practice, 5, 7-74*.
- Black, P. J. og Wiliam, D. (1998b). *Inside the black box. Raising standards through classroom assessment*. London: King's College.
- Black, P. J. og Wiliam, D. (2006). *The Reliability of Assessments*. I: J. Gardner (red.), *Assessment and learning*. London: Sage.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals; Handbook I, Cognitive Domain*. New York: David McKay.
- Botten, G. (2003). *Meningsfylt matematikk – nærhet og engasjement i læringen*. Bergen: Caspar forlag.
- Bourdieu, P. og Passeron, J.-C. (2006). *Reproduksjonen. Bidrag til en teori om undervisningssystemet*. København: Han Reitzels Forlag.
- Brady, L. og Kennedy, K. (2005). *Celebrating Student Achievement: assessment and reporting* (2 ed.). French Forest: Pearson Education Australia.
- Brekke, G. og Gjone, G. (2001). *Matematikk. I: S. Sjøberg (red.), Fagdebatikk. Fagdidaktisk innføring i sentrale skolefag*. Oslo: Gyldendal Akademisk.
- BRT (2002). *Professional Standards for Graduates and Guidelines for Preservice Teacher Education Programs*.
- Cumming, J. J. og Maxwell, G. (2004). *Assessment in Australian schools: current practice and trends. Assessment in Education: Principles, Policy and Practice, 11*.

- Dale, E. L. (2008). *Felleskolen - reproduksjon av sosial ulikhet*. Oslo: Cappelen akademisk forlag.
- Dale, E. L. og Wærness, J. I. (2003). *Differensiering og tilpansing i grunnopplæringen*. Oslo: Cappelen Akademisk Forlag.
- Dale, E. L. og Wærness, J. I. (2006). *Vurdering og læring i en elevaktiv skole*. Oslo: Universitetsforlaget.
- Dale, E. L. og Wærness, J. I. (2007). Tilpasset opplæring og elevvurdering. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Darling-Hammond, L. (2004). From «Separate but Equal» to «No Child Left Behind»: The Collision of New Standards and Old Inequalities. In D. Meier og G. Wood (red.), *Many Children Left Behind*. Boston: Beacon Press.
- Deci, E. L., Koestner, R. og Ryan, R. M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125 (6), 627-668.
- Dodgson, C. L. C. (1865). *Alice's Adventures in Wonderland*. England: Macmillan.
- Dysthe, O. (1999). *Ord på nye spor. Innføring i prosessorientert skriveopplæring* (2. utgave). Oslo: Det Norske Samlaget.
- Dysthe, O. (2007). Læring og læringsformer i Kunnskapsløftet. I: H. Hølleland (red.), *På vei mot Kunnskapsløftet*. Oslo: Cappelen akademiske forlag.
- Egelund, N. (2005). Educational assessment in Danish schools. *Assessment in Education: Principles, Policy and Practice*, 12 (2), 203-212.
- Egelund, N. (2008). Dansk læreruddannelse i et reformperspektiv. I: K.-A. Madssen (red.), *Pedagogikken i reformene – reformene i pedagogikken*. Oslo: Abstrakt forlag.
- Engelsen, B.U. (2006). *Kan læring planlegges?* (5. utgave). Oslo: Gyldendal Akademisk.
- Engh, R., Dobson, S. og Høihilder, E. K. (2007). *Vurdering for læring*. Kristiansand: Høyskoleforlaget.
- Forskrift til opplæringsloven av 17. juli 1998. Hentet 20. januar 2009 fra:
<http://www.lovdatab.no/cgi-wift/ldles?doc=/sf/sf/sf-20060623-0724.html>.
- Forskrift til opplæringsloven. Oppdatert: 30. juni 2006. Hentet 20. januar 2009 fra:
<http://www.lovdatab.no/cgi-wift/ldles?doc=/sf/sf/sf-20060623-0724.html>.
- Forskrift til rammeplan for allmennlærerutdanning, 2003. Hentet 2. januar 2009 fra:
http://lovdatab.no/cgi-wift/wiftldles?doc=/usr/www/lovdatab/ltavd1/filer/sf-20030403-0504.html&emne=rammeplan*%20'allmennl%E6rerutdanning*!&.

- Freiberg, J. (1993). *The Operation Of Cultural Capital in Summative Criteria-Based assessment in Senior Secondary English*. Deakin University.
- Gardner, J. og Cowan, P. (2005). The fallibility of high stakes '11-plus' testing in Northern Ireland. *Assessment in Education: Principles, Policy and Practice*, 12 (2), 145-165.
- Giddens, A. (1987). *Social Theory and Modern Sociology*. Cambridge: Polity Press.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18, 519-521.
- Grossman, P. (2008). Responding to Our Critics: From Crisis to Opportunity in Research on Teacher Education. *Journal of Teacher Education*, 59 (1), 10-23.
- Hægeland, T., Kirkebøen, L. J. og Raaum, O. (2005). *Skoleresultater 2004. En kartlegging av karakterer fra grunn- og videregående skoler i Norge*. Oslo: Statistisk Sentralbyrå.
- Harlen, W. (2004). *A systematic review of the evidence of the impact on students, teachers and the curriculum of the process of using assessment by teachers for summative purposes*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In *Assessment and Learning*. London: Sage.
- Hattie, J. og Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77 (1).
- Haug, P. (2004) *Resultat frå evalueringa av Reform 97*. Oslo: Noregs forskingsråd.
- Haug, P. (2008). Skulen og allmennlærerutdanninga. I: K.-A. Madssen (red.), *Pedagogikken i reformene – reformene i pedagogikken*. Oslo: Abstrakt forlag.
- Haugstveit, T. B., Sjølie, G. og Øygarden, B. (2006). *Vurdering som profesjonsfaglig kompetanse*. Elverum: Høgskolen i Hedmark.
- Hertzberg, F. (1999). Vurdering av muntlig – det går an. I: F. Hertzberg og A. Roe (red.), *Muntlig norsk*. Oslo: Tano Aschehoug.
- Hertzberg, F. og Roe, A. (1999). Forord. I: F. Hertzberg og A. Roe (red.), *Muntlig norsk*. Oslo: Tano Aschehoug.
- Hjetland, H. (1996). Elevvurdering for å fremje læring og utvikling. I: K. Skagen (red.), *Karakterboka. Om karakterer og vurdering i ny skole*. Oslo: Universitetsforlaget.
- Holst, H. (5. mai 2006). *Foreslår å bedømme elevene*. Oslo: Intervju med Aftenposten. Hentet 2. februar 2009 fra: <http://www.aftenposten.no/nyheter/iriks/article1304852.ece>.
- Hølleland, H. (2007). Nasjonale prøver og kvalitetsutvikling i skolen. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.

- Innst. S. nr. 262 (2001-2002). Oslo: Stortinget. Hentet 2. februar 2009 fra:
<http://www.stortinget.no/inns/2001/inns-200102-262.html>.
- Innst. S. nr. 285 (1996-97). Oslo: Stortinget. Hentet 2. februar 2009 fra:
<http://www.stortinget.no/inns/1996/inns-199697-285.html>.
- Kavli, H. (2008). *Nasjonale prøver 2007 - Brukernes evaluering av gjennomføringen*. Oslo: Synovate.
- Klette, K. (2003). Lærernes klasseromsarbeid; Interkasjons- og arbeidsformer i norske klasserom etter Reform 97. I: K. Klette (red.), *Klasserommets praksisformer etter Reform 97*. Oslo: Pedagogisk Forskningsinstitutt, Universitetet i Oslo.
- Klette, K. (2008). *Recipes for Busy Kitchens: Arbeidsplaner som læringsverktøy*. Oslo: Overlevert personlig 19. januar 2009 på Pedagogisk forskningsinstitutt, Universitetet i Oslo.
- Klette, K., Lie, S., Ødegaard, M., Anmarksrud, Ø., Arnesen, N., Bergem, O. K. og Roe, A. (2008). *Rapport om forskningsprosjektet PISA+*. Oslo: Universitetet i Oslo.
- Kristensen, L. (1999). *Analyse*. Upublisert: Statens utdanningskontor, eksamenssekretariatet.
- KUF (1997). *Læreplanverket for den 10-årige grunnskolen*. Oslo: Kirke- utdannings- og forskningsdepartementet. Hentet 2. februar 2009 fra:
<http://www.udir.no/L97/L97/index.html>
- Kunnskapsdepartementet (2006). *Læreplanverket for Kunnskapsløftet. Midlertidig utgave juni 2006*. Oslo: Kunnskapsdepartementet. Hentet fra:
http://www.utdanningsdirektoratet.no/upload/larerplaner/Fastsatte_lareplaner_for_Kunnskapsloeftet/Kunnskapsloftet_midlertidig_utgave_2006_tekstdel.pdf
- Kunnskapsdepartementet (2007). *Oppdragsbrev nr. 06 om tiltak knyttet til individvurdering i skole og fag- og yrkesopplæring*. Oslo: Kunnskapsdepartementet.
- Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks: Sage.
- Læringscenteret (2003). *Karakter - mer enn karakterer*. Oslo: Læringscenteret.
- Lie, S. (2007). *Evalueringen av de nasjonale prøvene og hva vi kan lære av dem*. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Lie, S., Kjærnsli, M., Roe, A. og Turmo, A. (2001). *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv*. Oslo: Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.

-
- LTS (2008). AifL – Assessment is for Learning. Glasgow: Learning and Teaching Scotland.
Sist oppdatert: 18. november 2008. Hentet 2. februar 2009 fra:
<http://www.ltscotland.org.uk/assess/of/index.asp>
- Lysne, A. (1999). *Karakterer og kompetanse. Stridstema i norsk skolehistorie. Bind 1*.
Haslum: AVA Forlag.
- Lysne, A. (2004). *Kampen om skolen. 1970-2000. Karakterer og kompetanse. Stridstema i norsk skolehistorie. Bind 2*. Haslum: AVA Forlag.
- Lysne, A. (2006). Assessment Theory and Practice of Students' Outcomes in the Nordic Countries. *Scandinavian Journal of Educational Research*, 50 (3), 327-359.
- Masters, G. N. og McBryde, B. (1994). *An Investigation of the Comparability of Teachers' Assessments of Student Folios*. Brisbane: Tertiary Entrance Procedures Authority.
- Matthiesen, G. (2007). Vurdering i skole og opplæring - problemstillinger og erfaringer. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Maxwell, G. (2006). *Quality management of school-based assessments: Moderation of teacher judgments*. Paper presentert ved den 32. årlige IAEA konferansen i Singapore.
- Maxwell, G. (2007). *Implications for moderation of proposed changes to senior secondary school syllabuses*. Brisbane: Queensland Studies Authority.
- McMeniman, M. (2004). *Review of the powers and functions of the Board of Teacher Registration*. Brisbane: Griffith University.
- Mead, G. H. (2005). *Sindet, selvet og samfundet*. Århus: Akademisk forlag.
- Messick, S. (1996). *Validity and Washback in Language Testing*. Princeton, New Jersey: Educational Testing Service.
- Møretrø, J. (2007). Tilpasset opplæring og elevvurdering. I: S. Tveit (red.), *Elevdeltakelse i vurderingsarbeidet i matematikkfaget*. Oslo: Universitetsforlaget.
- NCCOSE (2005). *Statements of Learning for English*. Curriculum Cooperation. Hentet 2. februar 2009 fra: http://www.mceetya.edu.au/verve/_resources/SOL_english_06.pdf.
- NCCOSE (2006). *Statements of Learning for Mathematics*. Curriculum Cooperation. Hentet 2. februar 2009 fra: http://hrd.apecwiki.org/images/2/2d/SOL_Mathematics_2006.pdf.
- NOKUT. (2006). *Evaluering av allmennlærerutdanningen i Norge 2006. Del 1: Hovedrapport*. Oslo: NOKUT.
- NOU (1974: 41). *Karakterer, eksamener, kompetanse m.v. i skoleverket*. Oslo: Kirke- og undervisningsdepartementet.

- NOU (1997). *Nasjonalt vurderingssystem for grunnskolen (Moe-utvalget)*. Oslo: Kirke, utdannings- og forskningsdepartementet
- NOU (2002: 10). *Første klasser fra første klasse*. Oslo: Utdannings- og forskningsdepartementet. Hentet 2. februar 2009 fra: <http://www.regjeringen.no/nb/dep/kd/dok/NOUer/2002/NOU-2002-10.html?id=145378&showdetailedtableofcontents=true>.
- NOU (2003: 16). *I første rekke*. Oslo: Utdannings- og forskningsdepartementet. Hentet 2. februar 2009 fra: <http://www.regjeringen.no/nb/dep/kd/dok/NOUer/2003/NOU-2003-16.html?id=147077>.
- OECD (2005). *Formative Assessment. Improving learning in secondary classrooms*. Paris: OECD Publishing.
- Opplæringsloven. Oppdatert: 19. desember 2008. Hentet 2. februar 2009 fra: <http://lovdata.no/all/hl-19980717-061.html>.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3. utgave). Thousand Oaks: Sage
- Penne, S. (1999). Nadia og retorikken. I: F. Hertzberg og A. Roe (red.), *Muntlig norsk*. Oslo: Tano Aschehoug.
- Pitman, J., O'Brien, J. og McCollow, J. (1999). *High-Quality Assessment: We are what we believe and do*. Paper presentert ved den 24. årlige IAEA konferansen i Bled, Slovenia. Hentet 2. februar 2009 fra: http://pan.search.qld.gov.au/search/click.cgi?url=http%3A%2F%2Fwww.qsa.qld.edu.au%2Fdownloads%2Fpublications%2Fresearch_qbssss_assessment_quality_99.pdf&rank=2&collection=qld-gov.
- QBSSSS (2001). *Mathematics B Senior Syllabus*. Brisbane: Queensland Board of Senior Secondary School Studies.
- QBSSSS (2002). *English Senior Syllabus*. Brisbane: Queensland Board of Senior Secondary School Studies.
- QCT (2007). *Program approval guidelines for Preservice Teacher Education*. Brisbane: Queensland College of Teachers.
- QSA (2004). *Mathematics Year 1 to 10 Syllabus*. Brisbane: Queensland Studies Authority. Hentet 2. februar 2009 fra: http://www.qsa.qld.edu.au/downloads/syllabus/kla_maths_syll.pdf.

-
- QSA (2005a). *English Year 1 to 10 Syllabus Open Trial Semester 2, 2005 - Semester 1, 2006*. Brisbane: Queensland Studies Authority. Hentet 2. februar 2009 fra:
http://www.beenleigss.eq.edu.au/requested_sites/english/english.htm.
- QSA (2005b). *Moderation Processes for Senior Certification (including prevailing information statements, memoranda and policy statements)*. Brisbane: Queensland Studies Authority.
- QSA (2006). *Random Sampling of Assessments in Authority Subjects 2006 Report*. Brisbane: Queensland Studies Authority.
- QSA (2008a). Assessment Bank. Brisbane: Queensland Studies Authority. Oppdatert: 1. oktober 2008. Hentet 30. januar 2009 fra:
<http://www.qsa.qld.edu.au/assessment/3162.html>.
- QSA (2008b). QCS Test general information. Brisbane: Queensland Studies Authority. Oppdatert 10. Desember 2007. Hentet 21. august 2008 fra:
<http://www.qsa.qld.edu.au/assessment/2318.html>.
- Queensland Government. (2008). *Census 2006 Bulletin 5: Education in Queensland*. Brisbane: Queensland Government. Hentet 2. februar 2009 fra:
<http://www.oesr.qld.gov.au/queensland-by-theme/society/education-training/bulletins/census-2006/education-qld-c06/education-qld-c06.pdf>.
- Sadler, P. M. og Tai, R. H. (2007). The Two High-School Pillars Supporting College Science. *Education Forum*, 317, 2.
- Sadler, R. (1986). *The place of numerical marks in criteria-based assessment*. Brisbane: Queensland Board of Secondary School Studies.
- Sadler, R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13 (2).
- Scott, D. (1996). Methods and data in educational research. I: D. Scott og R. Usher (red.), *Understanding educational research*. London: Routledge.
- Scriven, M. (1967). The methodology of evaluation. I: R. W. Tyler (red.), *Perspectives of Curriculum Evaluation*. Chicago: Rand McNally.
- Sebba, J. (2006). Policy and Practice in Assessment for Learning: the Experience of Selected OECD Countries. I: *Assessment and Learning*. London: Sage.
- Sebba, J. og Maxwell, G. (2005). Queensland, Australia: An Outcomes-based Curriculum. In *Formative assessment. Improving learning in secondary classrooms*. Paris: OECD.
- Skagen, K. (1996). Vurdering og karakterer i 1990-årene. I: K. Skagen (red.), *Karakterboka. Om karakterer og vurdering i ny skole*. Oslo: Universitetsforlaget.

- Skjervheim, H. (1976). *Det instrumentalistiske mistaket*. Oslo: Filosofi og Samfunn.
- Smith, K. (2005). Teacher educators' expertise: what do novice teachers and teacher educators say? . *Teaching and Teacher Education*, 21, 177-192.
- Solhjell, B.V. og Aasland, T. (6. desember 2008). *Vil splitte opp lærerutdanningen*. Oslo: Intervju med Norsk Telegrambyrå. Hentet 2. februar 2009 fra:
<http://www.vg.no/nyheter/innenriks/elevavisen/artikkel.php?artid=555131>
- SSB (2008). Utdanningsstatistikk. Elevar i grunnskolen. Endelege tal, 1. oktober 2007. Oslo; Statistisk Sentralbyrå. Oppdatert 28. april 2008. Hentet 22. oktober 2008 fra:
<http://www.ssb.no/vis/emner/04/02/20/utgrs/main.html>
- SSB (2008b). Utvalgte nøkkeltall, fylkeskommuner - nivå 1. Gjennomsnitt alle fylkeskommuner 2007. Oslo: Statistisk Sentralbyrå. Hentet 23. mai 2008 fra:
<http://www.ssb.no/kostra/stt/index.cgi?spraak=norsk&nivaa=1®ionstype=fylkeskommune&faktaark=101040642883067®ioner=default@default&kolonne=0&event=ny&mal=region&cookie=0>
- Stake, R. E. (1997). Case Study Methods in Educational Research: Seeking Sweet Water. I: R. M. Jaeger (red.), *Complementary Methods for Research in Education*. Washington: American Educational Research Association.
- Steen, N. O. (2007). *From doing to learning. Læringsorientering i elevers arbeidsplaner og i bruk av arbeidsplaner i klasserommet*. Masteroppgave. Oslo: Universitet i Oslo.
- Stortingsmelding nr. 47 (1995–96). *Om elevvurdering, skolebasert vurdering og nasjonalt vurderingssystem*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 48 (1996-97). *Om lærerutdanning*. Oslo: Kyrkje-, utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 28 (1998-99). *Mot rikare mål*. Oslo: Kyrkje-, utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 16 (2001-2002). *Kvalitetsreformen Om ny lærerutdanning Mangfoldig - krevende - relevant*. Oslo: Utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 20 (2003-2004). *Kultur for læring*. Oslo: Utdannings- og forskningsdepartementet.
- Stortingsmelding nr. 16 (2006-2007) *...og ingen stod igjen*. Oslo: Kunnskapsdepartementet.
- Telhaug, A. O. (1997). *Utdanningsreformene. Oversikt og analyse*. Oslo: Didakta Norsk Forlag.
- Tholin, J. (2007). Vilken kunnskap räknas? I: A. Petterson (red.), *Sporre eller otyg – om bedömning och betyg*. Mamö: Lärarforbundets förlag.

-
- Timperley, H. og Alton-Lee, A. (2008). Reframing teacher professional learning: An alternative policy approach to strengthening valued outcomes for diverse learners. In G. Kelly, A. Luke og J. Green (red.), *Disciplines, knowledge and pedagogy. Review of Research in Education* (Vol. 32). Washington D.C.: Sage.
- TKI (2009). Assessment Tools for Teaching and Learning. Hentet 2. februar 2009 fra: <http://www.tki.org.nz/r/asttle/>
- Tønnessen, L. K. B. og Telhaug, A. O. (1996). Elevvurderingen i norsk skole i etterkrigstida. In K. Skagen (red.), *Karakterboka. Om karakterer og vurdering i ny skole*. Oslo: Universitetsforlaget.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14 (3), 281-294.
- Tveit, S. (2003). Jeg vil se mappa mi! I: Læringscenteret (red.), *Karakter - mer enn karakterer*. Oslo: Læringscenteret.
- Tveit, S. (2007a). Analyse av retningslinjer og praksis for elevvurdering. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Tveit, S. (2007b). Elevvurdering i Kunnskapsløftet. I: H. Hølleland (red.), *På vei mot Kunnskapsløftet*. Oslo: Cappelen akademiske forlag.
- Tveit, S. (2007c). *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Tveit, S. (2007d). En mer læringsstøttende og rettferdig elevvurdering. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Tveit, S. (2007e). Formål og retningslinjer for elevvurdering i Kunnskapsløftet. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Tveit, S. (2008a). Mot en standardbasert vurdering i norsk skole. *Bedre skole*, 2, 66-72.
- Tveit, S. (2008b). *The national reform for enhanced comparability of teacher judgments in Norwegian secondary education*. Presentert ved Association for Educational Assessment Europe sin årlige konferanse i 2008, i Hisar, Bulgaria.
- Tyler, R. W. (1949). *Basic principles of learning and instruction*. Chicago: University of Chicago Press.
- Utdanningsdirektoratet (2007a). *Høringsbrev om forslag til endringer i forskrift til opplæringsloven 23.06.2006 nr. 724 og endring i Læreplanverket for Kunnskapsløftet*. Retrieved from.

- Utdanningsdirektoratet (2007b). *The Education Mirror 2006. Analysis of primary and lower and upper secondary education in Norway*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2007c). *Vurdering. Et felles løft for bedre vurderingspraksis – en veiledning. Revidert utgave*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2008a). *Høringsbrev endring i forskrift til opplæringsloven kapittel 3 og 4 og endring i forskrift til privatskoleloven kapittel 3 og 4*. Oslo: Utdanningsdirektoratet. Hentet 2. februar 2009 fra: http://udir.no/templates/udir/TM_Artikkel.aspx?id=4147.
- Utdanningsdirektoratet (2008b). *Trekkordning ved eksamen i Kunnskapsløftet*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2008c). *Utdanningsspeilet 2007*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2009). *Vurdering*. Oslo: Utdanningsdirektoratet. Oppdatert 26. januar 2009. Hentet 2. februar fra: <http://www.skolenettet.no/vurdering>.
- Utdanningsetaten. (2009). *Vurdering for læring*. Oslo: Utdanningsetaten i Oslo. Oppdatert 14. januar 2009. Hentet 2. februar 2009 fra: http://www.utdanningsetaten.oslo.kommune.no/vurdering_for_laring/.
- Vagle, W., Berge, K. L., Evensen, L. S. og Hertzberg, F. (2007). Det umulige er nesten mulig. Om norskeksamen på ungdomsskolen og utvikling av vurderingsnormer. I: S. Tveit (red.), *Elevvurdering i skolen - grunnlag for kulturendring*. Oslo: Universitetsforlaget.
- Vygotsky, L. S. (1996). Interaksjon mellom læring og utvikling. I: E. L. Dale (red.), *Om utdanning. Klassiske tekster*. Oslo: Gyldendal norsk forlag.
- Wikström, C. (2006). Education and assessment in Sweden. *Assessment in Education: Principles, Policy and Practice*, 13 (1), 113-128.
- Wikström, C. (2007). Subjektive bedömningar och objektiva tolkningar. I: A. Petterson (red.), *Sporre eller otyg – om bedömning och betyg*. Mamö: Lärarforbundets förlag.
- Wyatt-Smith, C. M. og Castleton, G. (2005). Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies*, 37(2), 24.
- Yin, R. K. (2003). *Case study research : design and methods* (3. utgave): Thousand Oakes: Sage.
- Ålvik, T. (1991). Noen sentrale spørsmål vedrørende skolebasert vurdering. I: T. Ålvik (red.), *Skolebasert vurdering – en innføring*. Oslo: Ad Notam.

9 Appendikser

9.1 Læreplan Norsk, Norge¹³

Formål med faget

Norskfaget er et sentralt fag for kulturforståelse, kommunikasjon, dannelse og identitetsutvikling. Gjennom aktiv bruk av det norske språket i arbeid med egne tekster og i møte med andres innlemmes barn og unge i kultur og samfunnsliv. Norskfaget åpner en arena der de får anledning til å finne sine egne stemmer, ytre seg, bli hørt og få svar. Slik representerer faget en demokratisk offentlighet som rustet til deltakelse i samfunnsliv og arbeidsliv. Mer enn noen gang krever samfunnet mennesker som mestrer språk og et stort mangfold av tekster. Norskfaget skal legge til rette for utvikling av den enkelte elevs språk- og tekstkompetanse ut fra elevens evner og forutsetninger. Lese- og skrivekompetanse er både et mål i seg selv og et nødvendig grunnlag for læring og forståelse i alle fag på alle trinn.

Et hovedmål for opplæringen i norsk gjennom hele grunnopplæringen er språklig selvtillit og trygghet i egen kultur som grunnlag for utvikling av identitet, respekt for andre kulturer, aktiv samfunnsdeltakelse og livslang læring. Norskfaget etablerer seg i spenningsfeltet mellom det historiske og det samtidige, det nasjonale og det globale. Å se norsk språk og kultur i et historisk og nasjonalt perspektiv kan gi elevene innsikt i og forståelse for det samfunnet de er en del av. Internasjonale perspektiver i norskfaget kan bidra til å utvikle kulturforståelse, toleranse og respekt for mennesker fra andre kulturer. Dagens situasjon er preget av kulturutveksling og kommunikasjon på tvers av tidligere grenser — språklig, kulturelt, sosialt og geografisk. I denne sammenheng byr norsk kulturarv på et forråd av tekster som kan få ny og uventet betydning nettopp når kommunikasjonen får nye former og perspektivene utvides. Kulturarven er slik sett en levende tradisjon som forandres og skapes på nytt, og norskfaget skal oppmuntre elevene til å bli aktive bidragsytere i denne prosessen.

I Norge er både bokmål, nynorsk og samisk offisielle skriftspråk, og det tales mange ulike

¹³ Læreplanen har kompetansemål for 2, 4, 7, 10. og alle videregående trinn. Her er kun målene for 10. trinn tatt med.

dialekter og sosiolekter, men også andre språk enn norsk. Norsk språk og kultur utvikles i en situasjon preget av kulturelt mangfold og internasjonalisering, i samspill med nordiske nabospråk og minoritetsspråk i Norge og med impulser fra engelsk. I dette språklige og kulturelle mangfoldet utvikler barn og unge sin språkkompetanse. Med utgangspunkt i denne språksituasjonen skal det legges til rette for at barn og unge får et bevisst forhold til språklig mangfold og lærer å lese og skrive både bokmål og nynorsk.

Norskfaget forholder seg til et bredt spekter av tekster, både muntlige, skriftlige og sammensatte tekster, der skrift, lyd og bilder spiller sammen. Et norskfag for vår tid bygger på et utvidet tekstbegrep som inkluderer alle disse tekstformene. Faget skal hjelpe elevene til å orientere seg i mangfoldet av tekster og gi rom for opplevelse, læring, refleksjon og vurdering. I tillegg skal faget stimulere til utvikling av gode læringsstrategier og evne til kritisk tenkning og motivere til utvikling av lese- og skrivevane og gode lese- og skrivevaner. Gjennom lesing og skriving kan barn tidlig utvikle tanker, utforske nye verdener og stå fram med egne meninger og vurderinger. Etter hvert vil de unge få orientere seg i og tolke skjønnlitteratur og sakprosa fra fortid og samtid og fordype seg i selvvalgte emner. Slik får de mulighet til å utvikle perspektiv på teksthistoriens lange linjer, brudd og konflikter.

Hovedområder i faget

Faget er strukturert i hovedområder som det er formulert kompetansemål innenfor. Hovedområdene utfyller hverandre og må ses i sammenheng.

Norsk har kompetansemål etter 2., 4., 7. og 10. årstrinn i grunnskolen og etter Vg1, Vg2 og Vg3 i studieforberedende utdanningsprogram i videregående opplæring. I yrkesfaglige utdanningsprogram er det kompetansemål etter Vg2 og etter påbygging til generell studiekompetanse.

Kompetansemål som omhandler skriftlig sidemål, gjelder ikke for yrkesfaglige utdanningsprogram etter Vg2.

For elever som har samisk som første - eller andrespråk eller finsk som andrespråk, gjelder

rett til fritak fra opplæring og vurdering i norsk sidemål.

Oversikt over hovedområder:

Årstrinn	Hovedområder			
1.–10.	Muntlige	Skriftlige	Sammensatte	Språk
Vg1-Vg3	tekster	tekster	tekster	og kultur

Muntlige tekster

Hovedområdet *muntlige tekster* dreier seg om muntlig kommunikasjon, det vil si å lytte og tale og utforskning av muntlig tekst. Sentralt er utvikling av ulike språklige roller og sjangere og forståelse for hvordan man tilpasser språk og form til mottaker og til formål med teksten. Å lytte og tale er en del av hverdagens omgangsformer og er sentralt for sosial og kulturell kompetanse og for estetisk utfoldelse og opplevelse.

Skriftlige tekster

Hovedområdet *skriftlige tekster* dreier seg om skriftlig kommunikasjon, det vil si å lese og skrive norsk. Lesing og skriving er parallelle prosesser i den enkeltes læringsforløp. Eleven utvikler skrivekompetanse gjennom å skrive og lese og lesekompetanse gjennom å lese og skrive. Dette skjer gjennom arbeid med ulike sjangere på bokmål og nynorsk med økende krav til forståelse for sammenhenger mellom tekstenes form og funksjon. Det arbeides med å stimulere til lese- og skrivelyst og –evne og til å utvikle elevenes lese- og skrivestrategier i en sammenhengende progresjon gjennom hele det 13-årige løpet. Det er også lagt vekt på elevens evne til å lese ulike tekster på ulike måter, både for å lære og for å oppleve, samt på elevens forståelse av sin egen utvikling som leser og skriver.

Sammensatte tekster

Hovedområdet *sammensatte tekster* viser til et utvidet tekstbegrep der tekst kan være satt sammen av skrift, lyd og bilder i et samlet uttrykk. Det innebærer arbeid med tekster som bildebøker, tegneserier, aviser, reklame, nettsider, sangtekster, film og teater. Hovedområdet omfatter både elevens egen tekstproduksjon og opplevelse, kritisk vurdering og analyse av sammensatte tekster. Å kunne lese en sammensatt tekst dreier seg om å finne mening i helheten av ulike uttrykksformer i teksten.

Språk og kultur

Hovedområdet *språk og kultur* dreier seg om norsk og nordisk språk- og tekstkultur, men med internasjonale perspektiver. Det er lagt vekt på at elever skal utvikle en selvstendig forståelse av norsk språk og litteratur og innsikt i hvordan språk og tekst har endret seg over tid og fortsatt er i endring. Elevene skal få kunnskap om språket som system og språket i bruk i en rekke eldre og nyere tekstformer. De får mulighet til å utforske og oppleve både gode norske forfattere og forfattere fra verdenslitteraturen. I tillegg forholder de seg til tradisjoner i norsk teksthistorie i et sammenlignende perspektiv mellom nåtid og fortid og i forhold til impulser utenfra.

Timetall i faget

Timetall oppgitt i 60-minutters enheter:

BARNETRINNET

1.-4. årstrinn: 931 timer - gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

5.-7. årstrinn: 441 timer – gjelder elever som ikke har samisk som første- eller andrespråk

1.-4. årstrinn: 642 timer – gjelder for elever som har samisk eller finsk som andrespråk

5.-7. årstrinn: 312 timer - gjelder for elever som har samisk eller finsk som andrespråk

UNGDOMSTRINNET

8.-10. årstrinn: 398 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

8.-10. årstrinn:

Alternativ 1: 335 timer – gjelder for elever med samisk som andrespråk, men som ikke velger fremmedspråk/språklig fordypning

Alternativ 2: 278 timer – gjelder for elever med samisk som andrespråk som velger fremmedspråk/språklig fordypning

8.–10. årstrinn: 258 timer – gjelder for elever som har finsk som andrespråk

STUDIEFORBEREDENDE UTDANNINGSPROGRAM

Vg1: 113 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg1: 103 timer – gjelder for elever som har samisk eller finsk som andrespråk

Vg2: 112 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg2: 103 timer – gjelder for elever som har samisk eller finsk som andrespråk

Vg3: 168 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg3: 103 timer – gjelder for elever som har samisk eller finsk som andrespråk

YRKESFAGLIGE UTDANNINGSPROGRAM

Vg1: 56 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg1: 45 timer – gjelder for elever som har samisk eller finsk som andrespråk

Vg2: 56 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg2: 45 timer – gjelder for elever som har samisk eller finsk som andrespråk

PÅBYGGING TIL GENERELL STUDIEKOMPETANSE FOR YRKESFAGLIGE UTDANNINGSPROGRAM

Vg3: 281 timer – gjelder elever som ikke har samisk som første- eller andrespråk eller finsk som andrespråk

Vg3: 219 timer – gjelder for elever som har samisk eller finsk som andrespråk

Grunnleggende ferdigheter i faget

Grunnleggende ferdigheter er integrert i kompetansemålene der de bidrar til utvikling av og er en del av fagkompetansen. I norsk forstås grunnleggende ferdigheter slik:

Å kunne uttrykke seg muntlig i norsk er å ha evnen til å lytte og tale og kunne vurdere elementene i en sammensatt talesituasjon som er en forutsetning for kommunikasjon med andre både når det gjelder sosialt samvær, arbeidsliv og deltakelse i offentlig liv. Samtaler om tekster har en avgjørende betydning for elevenes læring og utvikling. Å tale og lytte er grunnleggende menneskelige aktiviteter som i norskfaget blir videreutviklet gjennom systematisk opplæring i ulike muntlige sjangere og aktiviteter.

Å kunne uttrykke seg skriftlig i norsk er også et ansvarsområde i norskfaget, fra den første skriveopplæringen til den videre skriveopplæringen gjennom 13 år. Skriftligheten i samfunnet er økende, ikke minst gjennom utviklingen av digitale kommunikasjonsformer, og kravet til å mestre skriftlig framstilling i ulike sjangere har blitt større. Skrivning er en måte å utvikle og strukturere ideer og tanker på, men det er også en kommunikasjonsform og en metode for å lære.

Å kunne lese i norsk er en grunnleggende ferdighet som norskfaget tar et særskilt ansvar for gjennom den første leseopplæringen og den videre leseopplæringen som foregår gjennom hele det 13-årige løpet. Lesing er både en ferdighet og en kulturell kompetanse. Å kunne lese omfatter både å kunne finne informasjon i ulike tekster, å lære fag og å oppleve og forstå resonnementer og framstillinger i et bredt spekter av tekstformer. Lesing er avhengig av kulturforståelse, og samtidig utvikler lesing kulturforståelse. Gjennom lesing får elevene del i tekstkulturen, og kan utvikle evnen til å tolke og forstå ulike tekster. Dermed får de erfaringer som gir mulighet for læring og opplevelse og for å forstå seg selv og kunne være deltaker i samfunnet.

Å kunne regne i norsk er en ferdighet som forutsetter et annet språk enn verbalspråket. Men disse språkene har et felles kunnskapsområde når det gjelder begrepsutvikling, logisk resonnement og problemløsning. Det gjelder også forståelse for form, system og komposisjon.

Ved lesing av sammensatte tekster og sakprosa blir arbeidet med grafiske framstillinger, tabeller og statistikk viktig for forståelse.

Å kunne bruke digitale verktøy i norsk er nødvendig for å mestre nye tekstformer og uttrykk. Dette åpner for nye læringsarenaer og gir nye muligheter i lese- og skriveopplæringen, i produksjon, komponering og redigering av tekster. I denne sammenheng er det viktig å utvikle evne til kritisk vurdering og bruk av kilder. Bruk av digitale verktøy kan støtte og utvikle elevenes kommunikasjonsferdigheter og presentasjoner.

Kompetansemål i faget

Kompetansemål etter 10. årstrinn

Muntlige tekster

Mål for opplæringen er at eleven skal kunne

- uttrykke egne meninger i diskusjoner og vurdere hva som er saklig argumentasjon
- drøfte hvordan språkbruk kan virke diskriminerende og trakasserende
- delta i utforskende samtaler om litteratur, teater og film
- forstå og gjengi informasjon fra svensk og dansk dagligtale
- lede og referere møter og diskusjoner
- vurdere egne og andres muntlige framføringer
- gjennomføre enkle foredrag, presentasjoner, tolkende opplesing, rollespill og dramatisering, tilpasset ulike mottakere

Skriftlige tekster

Mål for opplæringen er at eleven skal kunne

- lese og skrive tekster i ulike sjangere, både skjønnlitterære og sakspregede på bokmål og nynorsk: artikkel, diskusjonsinnlegg, formelt brev, novelle, fortelling, dikt, dramatekst og kåseri
- bruke lesestrategier variert og fleksibelt i lesing av skjønnlitteratur og sakprosa

- orientere seg i store tekstmengder for å finne relevant informasjon
- lese kritisk og vurdere teksters troverdighet
- gjenkjenne ulike måter å argumentere på i tekster
- begrunne egne tekstvalg med utgangspunkt i egne lesepreferanser og formålet med lesingen
- lese og gjengi innholdet i et utvalg tekster på svensk og dansk
- formidle muntlig og skriftlig egne leseerfaringer og leseopplevelser basert på tolkning og refleksjon
- gjenkjenne de språklige virkemidlene humor, ironi, kontraster og sammenligninger, symboler og språklige bilder og bruke dem i egne tekster
- uttrykke seg presist og med et variert og nyansert ordforråd i ulike typer tekster på bokmål og nynorsk
- vise hvordan tekster i ulike sjangere kan bygges opp på ulike måter
- vurdere egne tekster og egen skriveutvikling ved hjelp av kunnskap om språk og tekst
- bruke tekstbehandlingsverktøy til arkivering og systematisering av eget arbeid
- bruke tekster hentet fra bibliotek, Internett og massemedier på en kritisk måte, drøfte tekstene og referere til benyttede kilder

Sammensatte tekster

- tolke og vurdere ulike former for sammensatte tekster
- bruke ulike medier, kilder og estetiske uttrykk i egne norskfaglige og tverrfaglige tekster
- vurdere estetiske virkemidler i sammensatte tekster hentet fra informasjons- og underholdningsmedier, reklame og kunst og reflektere over hvordan vi påvirkes av lyd, språk og bilder
- gjøre rede for grunnleggende prinsipper for personvern og opphavsrett knyttet til publisering og bruk av andres tekster

Språk og kultur

Mål for opplæringen er at eleven skal kunne

- presentere viktige temaer og uttrykksmåter i sentrale samtidstekster og sammenligne

dem med framstillinger i klassiske verk fra norsk litteraturarv: kjærlighet og kjønnsroller, helt og antihelt, virkelighet og fantasi, makt og motmakt, løgn og sannhet, oppbrudd og ansvar

- gjøre rede for hvordan samfunnsforhold, verdier og tenkemåter framstilles i oversatte tekster fra samisk og andre språk
- presentere resultatet av fordypning i tre selvvalgte emner: et forfatterskap, et litterært tema og et språklig tema
- gjøre rede for noen kjennetegn ved hovedgrupper av norske talemål
- forklare bakgrunnen for at det er to likestilte norske skriftspråk og gjøre rede for språkdebatt og språklig variasjon i Norge i dag
- gjøre rede for rettigheter i forbindelse med samisk språk og om utbredelsen av de samiske språkene i Norge, Sverige, Finland og Russland
- forklare hvordan mening og uttrykk videreføres og endres når enkle fortellinger, tegneserier og poplyrikk oversettes til norsk

Vurdering i faget

Bestemmelser for sluttvurdering:

Standpunktvurdering

Årstrinn	Ordning
10. årstrinn	Elevene skal ha tre standpunktkarakterer, én i norsk hovedmål skriftlig, én i norsk sidemål skriftlig og én i norsk muntlig.
Vg1 studieforbereende utdanningsprogram	Elevene skal ha tre standpunktkarakterer, én i norsk hovedmål skriftlig, én i norsk sidemål skriftlig og én i norsk muntlig.
Vg2 yrkesfaglige utdanningsprogram	Elevene skal ha to standpunktkarakterer, én i norsk hovedmål skriftlig og én i norsk muntlig.
Vg2 studieforbereende utdanningsprogram Vg3 studieforbereende utdanningsprogram Påbygging til generell studiekompetanse	Elevene skal ha tre standpunktkarakterer, én i norsk hovedmål skriftlig, én i norsk sidemål skriftlig og én i norsk muntlig.

Der faget går over flere år, er det bare standpunktvurdering fra det øverste nivået som eleven har i faget som framkommer på kompetansebevis eller vitnemål.

For elever som har samisk som første- eller andrespråk eller finsk som andrespråk, gjelder rett til fritak fra opplæring og vurdering i norsk sidemål.

Eksamen for elever

Årstrinn	Ordning
10. årstrinn	Elevene kan trekkes ut til én skriftlig eksamen som omfatter norsk hovedmål og sidemål. Skriftlig eksamen blir utarbeidet og sensurert sentralt. Elevene kan også trekkes ut til muntlig eksamen i norsk. Muntlig eksamen blir utarbeidet og sensurert lokalt.
Vg2 yrkesfaglige utdanningsprogram	Elevene kan trekkes ut til skriftlig eksamen i norsk hovedmål. Skriftlig eksamen blir utarbeidet og sensurert lokalt. Elevene kan også trekkes ut til muntlig eksamen i norsk. Muntlig eksamen blir utarbeidet og sensurert lokalt. Eksamen omfatter hele faget (112 timer).
Vg3 studieforbredende utdanningsprogram Påbygging til generell studiekompetanse	Elevene skal opp til skriftlig eksamen i norsk hovedmål. Elevene kan også trekkes ut til skriftlig eksamen i norsk sidemål. Skriftlig eksamen blir utarbeidet og sensurert sentralt. Elevene kan i tillegg trekkes ut til muntlig eksamen i norsk. Muntlig eksamen blir utarbeidet og sensurert lokalt. Eksamen omfatter hele faget (393 timer).

Eksamen for privatister

Årstrinn	Ordning
10. årstrinn	Se gjeldende ordning for grunnskole-opplæring for voksne.
Vg2 yrkesfaglige utdanningsprogram	Privatistene skal opp til skriftlig eksamen i norsk hovedmål. Skriftlig eksamen blir utarbeidet og sensurert lokalt. I tillegg skal privatistene opp til muntlig eksamen i norsk. Muntlig eksamen blir utarbeidet og sensurert lokalt. Eksamen omfatter hele faget (112 timer).
Vg3 studieforbredende utdanningsprogram Påbygging til generell studiekompetanse	Privatistene skal opp til skriftlig eksamen i norsk hovedmål og i norsk sidemål. Skriftlig eksamen blir utarbeidet og sensurert sentralt. I tillegg skal privatistene opp til muntlig eksamen i norsk. Muntlig eksamen blir utarbeidet og sensurert lokalt. Eksamen omfatter hele faget (393 timer).

De generelle bestemmelsene om vurdering er fastsatt i forskriften til opplæringsloven.

9.2 Læreplan Matematikk, Norge

Føremål med faget

Matematikk er ein del av den globale kulturarven vår. Mennesket har til alle tider brukt og utvikla matematikk for å utforske universet, for å systematisere erfaringar og for å beskrive og forstå samanhengar i naturen og i samfunnet. Ei anna inspirasjonskjelde til utviklinga av faget har vore glede hos menneske over arbeid med matematikk i seg sjølv. Faget grip inn i mange vitale samfunnsområde, som medisin, økonomi, teknologi, kommunikasjon, energiforvaltning og byggjeverksemd. Solid kompetanse i matematikk er dermed ein føresetnad for utvikling av samfunnet. Eit aktivt demokrati treng borgarar som kan setje seg inn i, forstå og kritisk vurdere kvantitativ informasjon, statistiske analysar og økonomiske prognosar. På den måten er matematisk kompetanse nødvendig for å forstå og kunne påverke prosessar i samfunnet.

Problemløysing høyrer med til den matematiske kompetansen. Det er å analysere og omforme eit problem til matematisk form, løyse det og vurdere kor gyldig det er. Dette har òg språklege aspekt, som det å resonnerer og kommunisere idear. I det meste av matematisk aktivitet nyttar ein hjelpemiddel og teknologi. Både det å kunne bruke og vurdere hjelpemiddel og teknologi og det å kjenne til avgrensinga deira er viktige delar av faget. Kompetanse i matematikk er ein viktig reiskap for den einskilde, og faget kan leggje grunnlag for å ta vidare utdanning og for deltaking i yrkesliv og fritidsaktivitetar. Matematikk ligg til grunn for viktige delar av kulturhistoria vår og for utviklinga av logisk tenking. På den måten spelar faget ei sentral rolle i den allmenne danninga ved å påverke identitet, tenkjemåte og sjølvforståing.

Matematikkfaget i skolen medverkar til å utvikle den matematiske kompetansen som samfunnet og den einskilde treng. For å oppnå dette må elevane få høve til å arbeide både praktisk og teoretisk. Opplæringa vekslar mellom utforskande, leikande, kreative og problemløysande aktivitetar og ferdigheitstrening. I arbeid med teknologi og design og i praktisk bruk viser matematikk sin nytte som reiskapsfag. I skolearbeidet utnyttar ein sentrale idear, former, strukturar og samanhengar i faget. Det må leggjast til rette for at både jenter og gutar får rike erfaringar som skaper positive haldningar og ein solid fagkompetanse. Slik blir det lagt eit grunnlag for livslang læring.

Hovudområde i faget

Faget er strukturert i hovudområde som det er formulert kompetansemål for. Hovudområda utfyller kvarandre og må sjåast i samanheng.

Matematikk har kompetansemål etter 2., 4., 7. og 10. årssteget i grunnskolen og etter Vg1 og Vg2 i studieførebuande og yrkesfaglege utdanningsprogram i vidaregåande opplæring.

Det er to læreplanar i faget for Vg1 og to læreplanar for Vg2. Læreplan T er meir teoretisk orientert, medan læreplan P er meir praktisk orientert. Begge variantane gjev generell studiekompetanse.

Yrkesfagelevar skal ha tre femdelar av læreplan Vg1P eller Vg1T:

Vg1P: hovudområda tal og algebra
 geometri
 økonomi

Vg1T: hovudområda tal og algebra (kompetansemåla 1, 2, 3 og 5)
 geometri (heile hovudområdet)
 sannsyn (kompetansemåla 1, 2 og 3)

Elevar i yrkesfagleg utdanningsprogram og dei som har fagbrev, sveinebrev eller annan yrkeskompetanse, og som ynskjer generell studiekompetanse, følgjer resten av læreplanen frå Vg1P eller Vg1T og læreplanen som høyrer til Vg2.

Oversikt over hovudområde:

Årssteg	Hovudområde				
	1.-4.	Tal	Geometri	Måling	Statistikk
5.-7.	Tal og algebra	Geometri	Måling	Statistikk og sannsyn (bm.: sannsynlighet)	
8.-10.	Tal og algebra	Geometri	Måling	Statistikk, sannsyn og kombinatorikk	Funksjonar

Vg1T	Tal og algebra	Geometri		Sannsyn	Funksjonar
Vg1P	Tal og algebra	Geometri	Økonomi	Sannsyn	Funksjonar
Vg2T	Geometri	Kombinatorikk sannsyn	Kultur og modellering		
Vg2P	Tal og algebra i praksis	Statistikk	Modellering		

Tal og algebra

Hovudområdet *tal og algebra* handlar om å utvikle talforståing og innsikt i korleis tal og talbehandling inngår i system og mønster. Med tal kan ein kvantifisere mengder og storleikar. Tal omfattar både heile tal, brøk, desimaltal og prosent. Algebra i skolen generaliserer talrekning ved at bokstavar eller andre symbol representerer tal. Det gjev høve til å beskrive og analysere mønster og samanhengar. Algebra blir òg nytta i samband med hovudområda *geometri* og *funksjonar*.

Geometri

Geometri i skolen handlar mellom anna om å analysere eigenskapar ved to- og tredimensjonale figurar og gjere konstruksjonar og berekningar. Ein studerer dynamiske prosessar, som spegling, rotasjon og forskyving. Hovudområdet omfattar òg det å utføre og beskrive lokalisering og flytting.

Måling

Måling vil seie å samanlikne og oftast knyte ein talstorleik til eit objekt eller ei mengd. Denne prosessen krev at ein bruker måleiningar og høvelege teknikkar, målereiskapar og formlar. Vurdering av resultatet og drøfting av måleusikkerheit er viktige delar av måleprosessen.

Statistikk, sannsyn og kombinatorikk

Statistikk omfattar å planleggje, samle inn, organisere, analysere og presentere data. I analysen av data høyrer det med å beskrive generelle trekk ved datamaterialet. Å vurdere og sjå kritisk på konklusjonar og framstilling av data er sentralt i statistikk. I sannsynsrekning talfester ein kor stor sjans det er for at ei hending skal skje. I kombinatorikk arbeider ein med systematiske måtar å finne tal på, og det er ofte nødvendig for å kunne berekne sannsyn.

Funksjonar

Ein funksjon beskriv endring eller utvikling av ein storleik som er avhengig av ein annan, på ein eintydig måte. Funksjonar kan uttrykkjast på fleire måtar, til dømes med formlar, tabellar og grafar. Analyse av funksjonar går ut på å leite etter spesielle eigenskapar, som kor raskt ei utvikling går, og når utviklinga får spesielle verdiar.

Økonomi

Hovudområdet *økonomi* handlar om berekningar og vurderingar som gjeld økonomiske forhold.

Kultur og modellering

Hovudområdet *kultur og modellering* gjev eit overordna perspektiv på faget matematikk. Hovudområdet beskriv den logiske strukturen i faget og viser historia og den kulturelle rolla til faget. Modellering er ein fundamental prosess i faget, der utgangspunktet er noko som verkeleg finst. Dette blir beskrive matematisk med ein modell som blir tilarbeidd, og resultatata av det blir tolka i lys av den opphavlege situasjonen.

Timetal i faget

Timetala er oppgjevne i einingar på 60 minutt.

BARNESTEGET

1.–7. årssteget: 812 timar

UNGDOMSSTEGET

8.–10. årssteget: 313 timar

STUDIEFØREBUANDE UTDANNINGSPROGRAM

Vg1: 140 timar

Vg2: 84 timar

YRKESFAGLEGE UTDANNINGSPROGRAM

Vg1: 84 timar

PÅBYGGING TIL GENERELL STUDIEKOMPETANSE FOR YRKESFAGLEGE
UTDANNINGSPROGRAM

Vg3: 140 timar

Grunnleggjande ferdigheiter i faget

Grunnleggjande ferdigheiter er integrerte i kompetansemåla, der dei medverkar til å utvikle fagkompetansen og er ein del av han. I matematikk forstår ein grunnleggjande ferdigheiter slik:

Å kunne uttrykkje seg munnleg i matematikk inneber å gjere seg opp ei meining, stille spørsmål, argumentere og forklare ein tankegang ved hjelp av matematikk. Det inneber òg å vere med i samtalar, kommunisere idear og drøfte problem og løysingsstrategiar med andre.

Å kunne uttrykkje seg skrifleg i matematikk inneber å løyse problem ved hjelp av matematikk, beskrive og forklare ein tankegang og setje ord på oppdagingar og idear. Ein lagar teikningar, skisser, figurar, tabellar og diagram. I tillegg nyttar ein matematiske symbol og det formelle språket i faget.

Å kunne lese i matematikk inneber å tolke og dra nytte av tekstar med matematisk innhald og med innhald frå daglegliv og yrkesliv. Slike tekstar kan innehalde matematiske uttrykk, diagram, tabellar, symbol, formlar og logiske resonnement.

Å kunne rekne i matematikk utgjer ei grunnstamme i matematikkfaget. Det handlar om problemløysing og utforsking som tek utgangspunkt i praktiske, daglegdagse situasjonar og matematiske problem. For å greie det må ein kjenne godt til og meistre rekneoperasjonane, ha evne til å bruke varierte strategiar, gjere overslag og vurdere kor rimelege svara er.

Å kunne bruke digitale verktøy i matematikk handlar om å bruke slike verktøy til spel, utforsking, visualisering og publisering. Det handlar òg om å kjenne til, bruke og vurdere

digitale hjelpemiddel til problemløsning, simulering og modellering. I tillegg er det viktig å finne informasjon, analysere, behandle og presentere data med høvelege hjelpemiddel, og vere kritisk til kjelder, analysar og resultat.

Kompetansemål i faget

Kompetansemål etter 10. årssteget

Tal og algebra

Mål for opplæringa er at eleven skal kunne

- samanlikne og rekne om heile tal, desimaltal, brøkar, prosent, promille og tal på standardform, og uttrykkje slike tal på varierte måtar
- rekne med brøk, utføre divisjon av brøkar og forenkle brøkuttrykk
- bruke faktorar, potensar, kvadratrøter og primtal i berekningar
- utvikle, bruke og gjere greie for metodar i hovudrekning, overslagsrekning og skriftleg rekning med dei fire rekneartane
- behandle og faktorisere enkle algebrauttrykk, og rekne med formlar, parentesar og brøkuttrykk med eitt ledd i nemnaren
- løyse likningar og ulikskapar av første grad og enkle likningssystem med to ukjende
- setje opp enkle budsjett og gjere berekningar omkring privatøkonomi
- bruke, med og utan digitale hjelpemiddel, tal og variablar i utforsking, eksperimentering, praktisk og teoretisk problemløsning og i prosjekt med teknologi og design

Geometri

Mål for opplæringa er at eleven skal kunne

- analysere, også digitalt, eigenskapar ved to- og tredimensjonale figurar og bruke dei i samband med konstruksjonar og berekningar
- utføre og grunngje geometriske konstruksjonar og avbildingar med passar og linjal og andre hjelpemiddel
- bruke formlikskap og Pytagoras' setning i berekning av ukjende storleikar

- tolke og lage arbeidsteikningar og perspektivteikningar med fleire forsvinningspunkt ved å bruke ulike hjelpemiddel
- bruke koordinatar til å avbilde figurar og finne eigenskapar ved geometriske former
- utforske, eksperimentere med og formulere logiske resonnement ved hjelp av geometriske idear, og gjere greie for geometriske forhold som har særleg mykje å seie i teknologi, kunst og arkitektur

Måling

Mål for opplæringa er at eleven skal kunne

- gjere overslag over og berekne lengd, omkrins, vinkel, areal, overflate, volum og tid, og bruke og endre målestokk
- velje høvelege måleiningar, forklare samanhengar og rekne om mellom ulike måleiningar, bruke og vurdere måleinstrument og målemetodar i praktisk måling, og drøfte presisjon og måleusikkerheit
- gjere greie for talet π og bruke det i berekningar av omkrins, areal og volum

Statistikk, sannsyn og kombinatorikk

Mål for opplæringa er at eleven skal kunne

- gjennomføre undersøkingar og bruke databasar til å søkje etter og analysere statistiske data og vise kjeldekritikk
- ordne og gruppere data, finne og drøfte median, typetal, gjennomsnitt og variasjonsbreidd, og presentere data med og utan digitale verktøy
- finne sannsyn gjennom eksperimentering, simulering og berekning i daglegdagse samanhengar og spell
- beskrive utfallsrom og uttrykkje sannsyn som brøk, prosent og desimaltal
- vise med døme og finne dei moglege løysingane på enkle kombinatoriske problem

Funksjonar

Mål for opplæringa er at eleven skal kunne

- lage, på papiret og digitalt, funksjonar som beskriv numeriske samanhengar og praktiske situasjonar, tolke dei og omsetje mellom ulike representasjonar av funksjonar,

som grafar, tabellar, formlar og tekst

- identifisere og utnytte eigenskapane til proporsjonale, omvendt proporsjonale, lineære og enkle kvadratiske funksjonar, og gje døme på praktiske situasjonar som kan beskrivast med desse funksjonane

Vurdering i faget

Retningsliner for sluttvurdering:

Standpunktvurdering

Årssteg	Ordning
10. årssteget	Elevane skal ha ein standpunktkarakter.
Vg1 yrkesfaglege utdanningsprogram Vg1 studieførebuande utdanningsprogram Vg2 studieførebuande utdanningsprogram Påbygging til generell studiekompetanse	Elevane skal ha ein standpunktkarakter.

Der faget går over fleire år, er det berre standpunktvurderinga på det øvste nivået som eleven har i faget, som kjem fram på kompetansebeviset eller vitnemålet. Unnateke frå dette er dokumentasjon for påbygging til generell studiekompetanse, der òg standpunktkarakteren frå Vg1 yrkesfaglege utdanningsprogram skal først.

Eksamen for elevar

Årssteg	Ordning
10. årssteget	Elevane kan trekkjast ut til skriftleg eksamen. Skriftleg eksamen blir utarbeidd og sensurert sentralt. Elevane kan òg trekkjast ut til munnleg eksamen. Munnleg eksamen blir utarbeidd og sensurert lokalt.
Vg1 yrkesfaglege utdanningsprogram	Elevane kan trekkjast ut til skriftleg eller munnleg eksamen. Skriftleg eksamen blir utarbeidd og sensurert lokalt. Munnleg eksamen blir utarbeidd og sensurert lokalt.
Vg2 studieførebuande utdanningsprogram	Elevane kan trekkjast ut til skriftleg eller munnleg eksamen. Skriftleg eksamen blir utarbeidd og sensurert sentralt. Munnleg eksamen blir utarbeidd og sensurert lokalt. Eksamen omfattar heile faget (224 timar).
Påbygging til generell studiekompetanse	Elevane kan trekkjast ut til skriftleg eller munnleg eksamen. Skriftleg eksamen blir utarbeidd og sensurert sentralt. Munnleg eksamen blir utarbeidd og sensurert lokalt. Eksamen omfattar berre faget i påbygging til generell studiekompetanse (140 timar).

Eksamen for privatistar

Årssteg	Ordning
10. årssteget	Sjå ordninga som gjeld for grunnskole-opplæring for vaksne.

Vg1 yrkesfaglege utdanningsprogram	Privatistane skal opp til skriftleg eksamen. Eksamen blir utarbeidd og sensurert lokalt.
Vg2 studieførebuande utdanningsprogram	Privatistane skal opp til skriftleg eksamen. Eksamen blir utarbeidd og sensurert sentralt. Eksamen omfattar heile faget (224 timar).
Påbygging til generell studiekompetanse	Privatistane skal opp til skriftleg eksamen. Eksamen blir utarbeidd og sensurert sentralt. Eksamen omfattar berre faget i påbygging til generell studiekompetanse (140 timar).

Dei generelle retningslinene om vurdering er fastsette i forskrifta til opplæringslova.

9.3 Læreplan English Year 1-10, Queensland: Eksempel på «level statements» for «writing and shaping – level 4»

Level statement			
Students interpret and construct texts, from a range of generic categories that explore familiar and unfamiliar subject matter, by considering text type, purposes, implied meanings, and known and unknown audiences. They identify and use a range of textual resources that structure texts, extend and elaborate ideas and information, and express opinions. They explore how particular representations appeal to certain groups.			
CORE LEARNING OUTCOMES	Cultural strand Making meanings in contexts	Operational strand Using language systems	Critical strand Evaluating and reconstructing meanings in texts
	Cu 4.3 When writing and shaping, students: <ul style="list-style-type: none"> • select subject matter according to purpose, text type, audience and medium • organise subject matter that develops a topic or storyline, supports a point of view or offers an explanation • develop characterisation that is relevant to the storyline using descriptions, actions and dialogue. 	Op 4.3 When writing and shaping, students: <ul style="list-style-type: none"> • organise and link ideas using generic structure, paragraphs, topic sentences and theme position • use evaluative verbs and attributes and simple figurative language when developing plot and characterisation and description • indicate certainty using modals • elaborate ideas and support a position through dependent and independent clauses, extended noun groups, phrases, words, and visual resources • use sound, visual and meaning patterns and knowledge of word function to spell multisyllabic words with regular patterns • indicate direct speech with punctuation. 	Cr 4.3 When writing and shaping, students: <ul style="list-style-type: none"> • choose aspects of subject matter, attributes, processes and visual resources to construct representations of people, places, events and things in ways that appeal to certain groups.
	Knowledge of texts in contexts	Knowledge of textual resources	Understanding of knowledge, values and practices of groups (discourse)
	To demonstrate these writing and shaping outcomes students should know: Text types <ul style="list-style-type: none"> • text types can be grouped into genre categories according to cultural purposes • names of the 	To demonstrate these writing and shaping outcomes students should know: Linguistic resources Text types <ul style="list-style-type: none"> • text types can be grouped into genre categories according to generic structure and patterns of textual resources • the names and functions of 	To demonstrate these writing and shaping outcomes students should know: Discourse <ul style="list-style-type: none"> • the knowledge, values and practices of different groups

<p style="text-align: center;">LEVEL-SPECIFIC CORE CONTENT</p>	<p>genre categories (narratives, transactions, procedures, reports and expositions) and a range of text types within these categories</p> <p>Subject matter</p> <ul style="list-style-type: none"> subject matter is organised differently according to text type subject matter is selected to support and provide evidence for a main idea or point of view how to draw on familiar and unfamiliar subject matter to construct texts <p>Roles and relationships</p> <ul style="list-style-type: none"> the role taken on by the writer will affect the degree of formality of the language assumptions about the characteristics of audiences can be used to engage their interest and attention through subject matter selection writers select language to show that some people have more authority than others (power), know each other better than others (distance) or like each other more than others (affect) <p>Mode and medium</p> <ul style="list-style-type: none"> likely associations 	<p>the common stages of the generic structure of a range of text types</p> <ul style="list-style-type: none"> the patterns of use of textual resources commonly associated with the generic categories of narratives, transactions, procedures, reports and expositions <p>Subject matter</p> <ul style="list-style-type: none"> subject matter is developed through a range of general, abstract and technical nouns and adjectives and a wide range of different types of verbs (e.g. thinking, feeling, saying, action, relating) characters are constructed using dialogue, extended noun groups and phrases that describe appearances and actions settings and plot are developed through extended noun groups and phrases and a range of verbs there are dependent and independent relationships among clauses and these work together to extend and elaborate ideas and information figurative language, such as simile, metaphor and personification, can be used to develop imagery <p>Roles and relationships</p> <ul style="list-style-type: none"> choice of modals indicates a degree of obligation, certainty or probability (e.g. should, may, must) written resources (processes, attributes) are chosen to appeal to certain groups nouns, verbs, adjective and adverbs can express opinions or give an evaluation of the behaviours of people and of the places, events or things being described (evaluative and opinion words) <p>Mode and medium</p> <ul style="list-style-type: none"> there is subject–verb agreement in simple, compound and complex sentences particular conjunctions are 	<ul style="list-style-type: none"> that assumptions about the characteristics of these groups influence the construction of texts <p>Subject matter</p> <ul style="list-style-type: none"> aspects of subject matter are selected to appeal to, and influence, different audiences textual resources are chosen to appeal to different groups.
---	---	---	---

	<p>between mode, medium and text type</p> <ul style="list-style-type: none"> choice of mode and medium determines the textual resources that will be available to develop, organise the flow of and link information in ways that guide the listener through the text. 	<p>selected to indicate the type of relationship between clauses</p> <ul style="list-style-type: none"> topic sentences are generally at the beginning of a paragraph and, together with the functional component in theme position in a sentence or clause, can be used to focus the reader's attention <p>Spelling (word structure) and punctuation</p> <ul style="list-style-type: none"> sound, visual, meaning patterns and word functions can be used in a multistrategy approach to spelling multisyllabic words that are regular in construction punctuation such as quotation marks adds to the meaning and status of published texts <p>Visual, gestural, spatial and audio resources</p> <ul style="list-style-type: none"> visual resources (shot type, layout, colour, print style) and audio resources (sound effects and music) can be used to emphasise meaning or to appeal to certain groups 	
--	---	---	--

9.4 Læreplan Mathematics Year 1-10, Queensland: Eksempel på «level statements» for «measurement» for level 4, 5 og 6

Core content		
Measurement — Length, mass, area and volume		
Level 4	Level 5	Level 6
<p>Measurement terms and attributes</p> <ul style="list-style-type: none"> • perimeter • circumference • square and cubic units <p>Units of measure</p> <ul style="list-style-type: none"> • millimetres (mm), centimetres (cm), metres (m) and kilometres (km) • tonnes (t) and kilograms (kg) • square metres (m²) • square centimetres (cm²) • cubic metres (m³) • cubic centimetres (cm³) • measuring instruments • related historical units of measure <p>Relationships</p> <ul style="list-style-type: none"> • the larger the unit the fewer required to measure and vice versa <ul style="list-style-type: none"> – metres, centimetres and millimetres – kilograms and tonnes – square centimetres and square metres • relationships between: <ul style="list-style-type: none"> – length, width and area of rectangle – length, width and height, and volume of a prism – length of side and perimeter 	<p>Measurement terms and attributes</p> <ul style="list-style-type: none"> • perimeter • circumference • diameter • radius • pi (π) <p>Units of measure</p> <ul style="list-style-type: none"> • square metres (m²) • hectares (ha) and square kilometres (km²) • measuring instruments • historical units of measure <p>Relationships</p> <ul style="list-style-type: none"> • the larger the unit the fewer required to measure and vice versa <ul style="list-style-type: none"> – millilitres and cubic centimetres – hectares and square metres • relationships between: <ul style="list-style-type: none"> – diameter and circumference of circle (π) – length and width (height), and areas of triangles and parallelograms – areas of triangles and areas of rectangles – areas of rectangles and areas of parallelograms (same length, same width or height) – areas of circles and irregular shapes • formulae <ul style="list-style-type: none"> – area of rectangle – volume of prism – perimeter of rectangles including squares 	<p>Measurement terms and attributes</p> <ul style="list-style-type: none"> • tangent • opposite and adjacent sides • hypotenuse <p>Units of measure</p> <p>Relationships</p> <ul style="list-style-type: none"> • within right-angled triangles <ul style="list-style-type: none"> – Pythagoras' Theorem – tangent ratio • formulae <ul style="list-style-type: none"> – circumference of circle – area of circle – area of triangle – volume of cylinder – volumes of pyramids and cones • compound shapes • objects

9.5 Læreplan English Year 1-10, Queensland: Eksempler på «ways to gather and record evidence from a variety of sources»

Sources of evidence	Assessment techniques	Recording instruments
<p>Students can provide evidence about what they know and can do with what they know, in a variety of forms. These include:</p> <ul style="list-style-type: none"> • computer-generated presentations • discussions with students • group discussions • conversations • interviews • role plays • persuasive speeches • concept maps • individual and group presentations • prepared speaking notes • work in progress • learning logs • reflective journals • commentaries on processes or justifications of choices • guided reading activities • note-taking • miscue analyses • reflections on processes • storyboards • spoken, written, visual and/or multimodal student-constructed texts. 	<p>Observation involves teachers observing students in focused and systematic ways during learning experiences. Teacher observation occurs continually as part of the learning and teaching process and can be used to gather a broad range of information about students' demonstrations of learning. Teacher observations can also be structured to gather particular kinds of information.</p> <p>Consultation involves teachers discussing student work with students, colleagues, parents/carers or other paraprofessionals. The varying perspectives of the participants in consultations can help enrich the evidence gathered about students' demonstrations of learning. Consultation can be used to verify the evidence gathered using other techniques. Some consultations may reveal a need for more detailed assessment.</p> <p>Focused analysis requires teachers to examine in detail students' responses to tasks or learning experiences. This technique provides detailed evidence about students' demonstrations of learning.</p> <p>Self- and peer-assessment involve students in using the above techniques to assess their own work and the work of their peers. Self- and peer-assessment allow teachers to take account of students' perceptions when gathering evidence.</p>	<p>Teachers can record their judgments about students' learning using a variety of recording instruments, including:</p> <ul style="list-style-type: none"> • anecdotal records • annotations on work samples • observation notes • checklists • statements of anticipated evidence (criteria sheets) • reflection sheets or diaries, learning logs • student-teacher conference or interview records • teachers' record books or file cards • student profiles • audio and visual (including photographic and video) recordings • self- and peer assessment sheets • student folios.

9.6 Læreplan Mathematics Year 1-10, Queensland: Eksempler på «ways to gather and record evidence from a variety of sources»

Sources of evidence	Assessment techniques	Recording instruments
<ul style="list-style-type: none"> • computer-generated presentations • concept maps • debates • discussions with students • games • journals • observations of written work in progress • plans of approach to investigations • projects/assignments • questioning led by teacher or student • reports (e.g. on investigations) • research projects • sketches and drawings • structured whole- or small-group discussions • student explanations of work in progress • student folios • working notes and jottings • written tests (e.g. Years 3, 5 and 7 test reports) 	<p>Observation involves teachers observing students as they participate in planned activities. Teacher observation occurs continually as a natural part of the learning and teaching process and can be used to gather a broad range of evidence about students' demonstrations of learning outcomes. Teacher observations can also be structured to gather particular kinds of evidence in relation to learning outcomes.</p> <p>Consultation involves teachers discussing student work with students, colleagues, parents/ carers or other paraprofessionals. The varying perspectives of the participants in consultations can help enrich the evidence gathered about students' demonstrations of learning outcomes. Consultation can be used to verify the evidence gathered using other techniques. Some consultations may reveal a need for more detailed assessment.</p> <p>Focused analysis involves teachers examining in detail student responses to tasks or activities (e.g. computer-generated presentations, group discussions, tests, debates or research projects). This technique provides detailed evidence about students' demonstrations of learning outcomes.</p> <p>Self- and peer-assessment involves students using the above techniques to assess their own work and the work of their peers. Self- and peer-assessment allow teachers to take account of students' perceptions when gathering evidence.</p>	<ul style="list-style-type: none"> • anecdotal records • annotated work samples • audio and visual (including photographic and video) recordings • checklists • feedback sheets • folios • learning logs • observation notes • reflection sheets, diaries or scrapbooks • reports of test results • self- and peer-assessment sheets • self- and peer-reflective journals • statements of anticipated evidence or criteria sheets • student/teacher journals • worksheets

9.7 Læreplan English Senior, Queensland: Achievement levels

Very High Achievement (VHA)

Table 2: Very High Achievement

At Very High Achievement , the student has consistently shown a discriminating knowledge and understanding of how texts are constructed across a range of texts in a range of social and cultural contexts. The student has demonstrated a comprehensive knowledge, understanding and control of the interrelatedness of purpose, genre, register, textual features and discourses in texts. The student has demonstrated a thorough understanding of the implications of tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> exploiting the patterns and conventions of genres to achieve cultural purposes selecting and synthesising substantial, relevant subject matter interpreting, inferring from, analysing and evaluating information, ideas, argument and images in great depth substantiating opinions with well balanced and relevant argument and evidence exploiting the ways writers, speakers/signers, shapers' roles and their relationships with readers, listeners, viewers are affected by power, distance and affect exploiting modes and media, integrating them to effect if appropriate. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> exploiting the sequencing and organisation of subject matter in stages making discerning use of cohesive ties (and hyperlinks) to emphasise ideas and connect parts of texts exploiting an extensive range of apt vocabulary, including figurative uses combining a wide range of clause and sentence structures for specific effects, while sustaining grammatical accuracy sustaining control of paragraphing and a wide range of punctuation controlling conventional spelling integrating visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features to enhance meaning in layout or presentation of texts sustaining use of a wide range of spoken/signed and nonverbal features that contribute to meaning: <ul style="list-style-type: none"> pronunciation, phrasing and pausing for emphasis, audibility and clarity, volume, pace facial expressions, gestures, proximity, stance, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> thoroughly examining how, <i>and</i>³ exploiting the ways discourses in texts shape and are shaped by language choices evaluating how, <i>and</i> exploiting the ways cultural assumptions, values, beliefs and attitudes underpin texts making subtle and complex distinctions when evaluating <i>and</i> when shaping representations of concepts and of the relationships and identities of individuals, groups, times and places thoroughly analysing how readers are invited to take up positions in relation to texts <i>and</i> demonstrating with subtlety and complexity the position(s) s/he adopts as a reader making purposeful and discerning choices that very effectively invite readers to take up positions.

³ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to examine the ways ... or exploit the ways ..., and so on for each descriptor.

High Achievement (HA)

Table 3: High Achievement

At High Achievement , the student has consistently shown knowledge and understanding of how texts are constructed across a range of texts in a range of social and cultural contexts. The student has demonstrated considerable knowledge, understanding and control of the interrelatedness of purpose, genre, register, textual features and discourses in texts. The student has demonstrated understanding of the implications of tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> employing the patterns and conventions of genres to achieve cultural purposes selecting and usually synthesising considerable relevant subject matter interpreting, inferring from, analysing and evaluating information, ideas, argument and images in depth substantiating opinions with relevant argument and evidence establishing writers, speaker/signers, shapers' roles and controlling the ways their relationships with readers, listeners, viewers are influenced by power, distance and affect exploiting modes and media, integrating them if appropriate. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> sequencing and organising subject matter logically in stages controlling the use of cohesive ties (and hyperlinks) to connect ideas and parts of texts selecting, with occasional lapses, a wide range of suitable vocabulary, including figurative uses controlling a wide range of clause and sentence structures, while generally maintaining grammatical accuracy sustaining control of paragraphing and a wide range of punctuation controlling conventional spelling, with occasional lapses integrating visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features to make meaning in layout or presentation of texts using a wide range of spoken/signed and nonverbal features that contribute to meaning: <ul style="list-style-type: none"> pronunciation, phrasing and pausing for emphasis, audibility and clarity, volume, pace facial expressions, gestures, proximity, stance, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> examining how, <i>and</i>⁴ making effective use of the ways discourses in texts shape and are shaped by language choices examining how, <i>and</i> making effective use of the ways cultural assumptions, values, beliefs and attitudes underpin texts making fine distinctions both when evaluating <i>and</i> when shaping representations of concepts and of the relationships and identities of individuals, groups, times and places analysing how readers are invited to take up positions in relation to texts <i>and</i> clearly demonstrating the position(s) s/he adopts as a reader making purposeful choices that effectively invite readers to take up positions.

⁴ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to examine how ... or make effective use of ..., and so on for each descriptor.

Sound Achievement (SA)

Table 4: Sound Achievement

At Sound Achievement, the student has shown reasonable knowledge and understanding of how texts are constructed across a range of texts in a range of social and cultural contexts. The student has demonstrated a general knowledge, understanding and control of the interrelatedness of purpose, genre, register, textual features and discourses in texts. For the most part, the student has demonstrated understanding of the implications of tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> • in the main, employing the patterns and conventions of genres to achieve particular cultural purposes • selecting sufficient relevant subject matter • interpreting and explaining information, ideas, argument and images, with some analysis and evaluation • supporting opinions with relevant argument and evidence • establishing writers, speakers/signers, shapers' roles and maintaining the ways their relationships with readers, listeners, viewers are influenced by power, distance and affect • usually making effective use of modes and media, in combination if appropriate. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> • in the main, sequencing and organising subject matter in stages • usually linking ideas with cohesive ties (and connecting parts of texts using hyperlinks) • using suitable vocabulary • using a range of clause and sentence structures with occasional lapses in grammatical accuracy • controlling paragraphing and punctuation, such as commas, apostrophes, capitals and full stops • using conventional spelling, in the main • using, with some success, visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features to make meaning in layout or presentation of texts • using a range of spoken/signed and nonverbal features that contribute to meaning: <ul style="list-style-type: none"> – pronunciation, phrasing and pausing for emphasis, audibility and clarity, volume, pace – facial expressions, gestures, proximity, stance, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> • explaining how, <i>and</i>⁵ employing ways discourses in texts shape and are shaped by language choices • identifying and explaining <i>and</i> making use of the ways cultural assumptions, values, beliefs and attitudes underpin texts • making broad distinctions when identifying and explaining <i>and</i> when shaping representations of concepts and of the relationships and identities of individuals, groups, times and places • identifying and explaining ways readers have been invited to take up positions in relation to texts <i>and</i> broadly demonstrating the position(s) s/he adopts as a reader • making purposeful choices that, in the main, invite readers to take up positions.

⁵ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to explain how ... or make use of ..., and so on for each descriptor.

Limited Achievement (LA)

Table 5: Limited Achievement

At Limited Achievement, the student has shown some knowledge and understanding of how texts are constructed across a range of texts in a range of social and cultural contexts. The student has demonstrated a very general knowledge, understanding and control of choices of genre, register and textual features in texts. The student has demonstrated an understanding of the implications of some of the tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> • unevenly using the patterns and conventions of genres to achieve cultural purposes • selecting some relevant subject matter • interpreting and explaining some information, ideas, images • supporting opinions with a little argument and evidence • generally establishing writers, speakers/signers, shaper's roles and sometimes maintaining the ways their relationships with readers, listeners, viewers are influenced by power or distance or affect • using some modes and media, with occasional effectiveness. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> • occasionally sequencing and organising subject matter in stages • making lapses in linking ideas with cohesive ties (or when using hyperlinks) • using basic vocabulary • using clause and sentence structures accurately in places, but with frequent grammatical lapses in subject-verb agreement, continuity of tenses and pronoun references • using paragraphing and punctuation accurately in places, but with frequent lapses • using conventional spelling, with frequent lapses • using some visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features in layout or presentation of texts • using some spoken/signed and nonverbal features that contribute to meaning: <ul style="list-style-type: none"> – pronunciation with some lapses, audibility and clarity, volume – facial expressions, gestures, proximity, stance, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> • identifying some ways language choices are shaped by discourses • identifying some of the ways cultural assumptions, values, beliefs and attitudes underpin texts <i>and</i>⁶ sometimes making use of them • making general distinctions when identifying representations of concepts and of the relationships and identities of individuals, groups, times and places <i>and</i> unevenly shaping the representations in own texts • recognising and describing some ways readers have been invited to take up positions in relation to texts • making choices that sometimes invite readers to take up positions.

⁶ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to identify some cultural assumptions ... or make use of them ..., and so on for each descriptor.

Very Limited Achievement (VLA)

Table 6: Very Limited Achievement

At Very Limited Achievement, the student has shown a little knowledge and understanding of how texts are constructed in some texts in social and cultural contexts. The student has demonstrated a little knowledge, understanding and control of genre, register and textual features. The student has demonstrated some understanding of some of the implications of tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> occasionally using some conventions of genres to achieve some purposes selecting some subject matter that relates to tasks stating opinions identifying roles of writers, speakers/signers, shapers and making some use of their relationships with readers, listeners, viewers in some texts using some modes and media. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> linking some ideas with conjunctions using a narrow range of basic vocabulary using a narrow range of clause and sentence structures with frequent grammatical lapses that impede understanding using some punctuation though not paragraphing using some conventional spelling, but lapses impede understanding using some visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features using some spoken/signed and nonverbal features: <ul style="list-style-type: none"> pronunciation with some lapses, audibility and clarity facial expressions, gestures, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> sometimes identifying some attitudes and beliefs in texts making very general distinctions when identifying representations of concepts and of the relationships and identities of individuals, groups, times and places <i>and</i> occasionally shaping some representations in own texts.

⁷ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to identify ... or to shape some ..., and so on for each descriptor.

Minimum requirements for Sound Achievement

6.7.4 Minimum requirements for Sound Achievement

For the awarding of a minimum Sound Achievement, the student's folio must meet the descriptors in both modes, i.e. in writing and in speaking/signing. To determine this, the responses to the modes are to be examined separately. The predominantly written responses, when taken together, must for the most part meet the minimum standard in table 7; the predominantly spoken/signed responses, when taken together, must for the most part meet the minimum standard in table 7. Higher achievement in spoken/signed responses cannot compensate for weaknesses in written responses or vice versa. Once the standard has been determined for each mode, the judgment for the folio follows.

Table 7: Minimum requirements for Sound Achievement

At minimum Sound Achievement, the student has shown reasonable control of a range of texts and their textual features across a range of social and cultural contexts demonstrating a general understanding of the interrelatedness of purpose, genre, register, textual features and discourses in the construction of texts. S/he demonstrates that s/he has understood, for the most part, the implications of tasks.		
Knowledge and control of texts in their contexts	Knowledge and control of textual features	Knowledge and application of the constructedness of texts
<p>The student has demonstrated knowledge that meanings in texts are shaped by purpose, cultural context and social situation by:</p> <ul style="list-style-type: none"> employing some of the patterns and conventions of genres used to achieve cultural purposes selecting subject matter that is mostly relevant interpreting and explaining information, ideas, argument and images, with some analysis providing argument and evidence to support most opinions establishing writers, speakers/signers, shapers' roles and generally maintaining the ways their relationships with readers, listeners, viewers are affected by power, distance and affect making some effective use of modes and media, combining where appropriate. 	<p>The student has demonstrated knowledge of appropriateness of textual features for purpose, genre, and register by:</p> <ul style="list-style-type: none"> In the main, sequencing and organising subject matter in stages, although there may be some significant lapses in a few texts usually linking ideas with cohesive ties (and using hyperlinks) using basic but usually suitable vocabulary using a range of clause and sentence structures, with some lapses in grammatical accuracy, including subject-verb agreement, continuity of tense and pronoun reference using paragraphing, with some lapses, and usually controlling commas, apostrophes, capitals and full stops, despite some run-on sentences using conventional spelling, in the main using, with some success, visual (graphic, still and moving images), auditory (music, silence and sound effects) and/or digital (graphic design elements) features in layout or presentation of texts using a range of spoken/signed and nonverbal features that usually contribute to meaning: <ul style="list-style-type: none"> pronunciation, phrasing and pausing, audibility and clarity, volume, pace facial expressions, gestures, proximity, stance, movement. 	<p>The student has demonstrated knowledge of the ways in which texts are selectively constructed and read by:</p> <ul style="list-style-type: none"> describing how <i>and</i>⁸ employing some of the ways discourses shape and are shaped by language choices identifying and describing, <i>and</i> making use of, the ways some cultural assumptions, attitudes, beliefs and values underpin texts making broad distinctions, both when identifying and describing, <i>and</i> when shaping representations of concepts and of the relationships and identities of individuals, groups, times and places identifying and describing ways readers have been invited to take up positions in relation to texts <i>and</i> broadly demonstrating reading position(s) s/he adopts sometimes making purposeful choices that invite readers to take up positions.

⁸ The use of the italic indicates that all dimensions of the descriptor will be demonstrated across the course, and not necessarily within an individual task. In this descriptor, a task might ask students either to identify ... or make some effort to ..., and so on for each descriptor.

9.8 Læreplan Mathematics B, Queensland: «Minimum standards associated with exit criteria»

	Standard A	Standard B	Standard C	Standard D	Standard E
Criterion: Knowledge and procedures	<p>The overall quality of a student's achievement across the full range within the contexts of application, technology and complexity, and across topics, consistently demonstrates:</p> <ul style="list-style-type: none"> accurate recall, selection and use of definitions, results and rules appropriate use of technology appropriate selection, and accurate and proficient use of procedures effective transfer and application of mathematical procedures 	<p>The overall quality of a student's achievement across a range within the contexts of application, technology and complexity, and across topics, generally demonstrates:</p> <ul style="list-style-type: none"> accurate recall, selection and use of definitions, results and rules appropriate use of technology appropriate selection and accurate use of procedures. 	<p>The overall quality of a student's achievement in the contexts of application, technology and complexity generally demonstrates:</p> <ul style="list-style-type: none"> accurate recall and use of basic definitions, results and rules appropriate use of some technology accurate use of basic procedures. 	<p>The overall quality of a student's achievement in the contexts of application, technology and complexity sometimes demonstrates:</p> <ul style="list-style-type: none"> accurate recall and use of some definitions, results and rules appropriate use of some technology. 	<p>The overall quality of a student's achievement rarely demonstrates knowledge and use of procedures.</p>

	Standard A	Standard B	Standard C	Standard D	Standard E
Criterion: Modelling and problem solving	<p>The overall quality of a student's achievement across the full range within each context, and across topics generally demonstrates mathematical thinking which includes:</p> <ul style="list-style-type: none"> interpreting, clarifying and analysing a range of situations, identifying assumptions and variables selecting and using effective strategies selecting appropriate procedures required to solve a wide range of problems appropriate synthesis of procedures and strategies; <p>... and in some contexts and topics demonstrates mathematical thinking which includes:</p> <ul style="list-style-type: none"> synthesis of procedures and strategies to solve problems initiative and insight in exploring the problem exploring strengths and limitations of models extending and generalising from solutions. 	<p>The overall quality of a student's achievement across a range within each context, and across topics, generally demonstrates mathematical thinking which includes:</p> <ul style="list-style-type: none"> interpreting, clarifying and analysing a range of situations, identifying assumptions and variables selecting and using effective strategies selecting appropriate procedures required to solve a range of problems; <p>... and in some contexts and topics demonstrates mathematical thinking which includes appropriate synthesis of procedures and strategies.</p>	<p>The overall quality of a student's achievement in all contexts generally demonstrates mathematical thinking which includes:</p> <ul style="list-style-type: none"> interpreting and clarifying a range of situations selecting strategies and/or procedures appropriate to problems. 	<p>The overall quality of a student's achievement sometimes demonstrates mathematical thinking which includes following basic procedures and/or using basic strategies.</p>	<p>The overall quality of a student's achievement rarely demonstrates mathematical thinking which includes following basic procedures and/or using basic strategies.</p>

	Standard A	Standard B	Standard C	Standard D	Standard E
Criterion: Communication and justification	<p>The overall quality of a student's achievement across the full range within each context consistently demonstrates:</p> <ul style="list-style-type: none"> accurate and appropriate use of mathematical terms and symbols accurate and appropriate use of language collection and organisation of information into various forms of presentation suitable for a given use or audience use of mathematical reasoning to develop logical arguments in support of conclusions, results and/or propositions justification of procedures and conclusions recognition of the effects of assumptions used evaluation of the validity of arguments. 	<p>The overall quality of a student's achievement across a range within each context generally demonstrates:</p> <ul style="list-style-type: none"> accurate and appropriate use of mathematical terms and symbols accurate and appropriate use of language collection and organisation of information into various forms of presentation suitable for a given use or audience use of mathematical reasoning to develop simple logical arguments in support of conclusions, results and/or propositions justification of procedures 	<p>The overall quality of a student's achievement in all contexts generally demonstrates:</p> <ul style="list-style-type: none"> accurate and appropriate use of basic mathematical terms and symbols accurate and appropriate use of basic language collection and organisation of information into various forms of presentation; use of some mathematical reasoning to develop simple logical arguments in support of conclusions, results and/or propositions. 	<p>The overall quality of a student's achievement sometimes demonstrates evidence of the use of the basic conventions of language and mathematics.</p>	<p>The overall quality of a student's achievement rarely demonstrates use of the basic conventions of language or mathematics.</p>

9.9 Læreplan English Senior, Queensland: Krav til «verification folio»

REQUIREMENTS FOR VERIFICATION FOLIOS

The verification folio requirements for English stipulate the minimum and maximum number of assessment instruments. Schools must ensure that the verification folios presented in October contain all summative assessment instruments and corresponding student responses upon which judgments about interim levels of achievement have been made up to that time.

Teachers must also ensure that sufficient material is included to support judgments and that the minimum requirements are completed and included in sample folios by the due dates for meetings.

A verification folio is compiled over time and comprises a representative sample of a student's work. The folios provide information on which level of achievement proposals are to be based. The verification folio should be genuinely representative of a student's standard of achievement across a variety of tasks and conditions as specified in the work program. These tasks should assess across the range of general objectives, deal with the principles of school-based assessment and permit students to demonstrate achievement across the exit criteria and standards.

Verification requirements

An individual student folio for October verification *must* contain:

- 6–7 tasks with the characteristics outlined in section 6.3
 - 4–5 of these are to be predominantly written (see minimum requirements below)
 - 2–3 of these are to be predominantly spoken/signed (see minimum requirements below)
- a student profile detailing the pattern of achievement in summative tasks
- documentation of spoken/signed tasks.

Minimum requirements — mandatory written categories

Among the tasks in the verification folio there must be at least:

- one analytical exposition in response to literature, e.g. an analysis of how Australian

identities are constructed through selections of poetry; a critical comparison of the ways the viewer has been positioned in two film versions of a play or novel that has been studied

- one imaginative text, e.g. short story, drama script, interior monologue, epistolary narrative
- one persuasive/reflective text suitable for a public audience, e.g. feature article, profile, or column; obituary; interview; biography; review.

All tasks should be carefully contextualised in terms of purpose and audience.

Two written tasks are to be *under supervised conditions*. *One* of these is *in response to an unseen question*. An unseen question is defined as one that students have not previously sighted. The “unseenness” relates to the question and not the resources that may support the response. The task is to be completed in one uninterrupted session.

At least one written task, but no more than two, is to be completed by students after the opportunity for extensive access to human, library, and electronic resources that should be referenced. Tasks completed under assignment conditions must include student documentation of the text production process as outlined in 6.3.5.

Once the mandated categories are met in a folio, the remaining written tasks do not necessarily have to be drawn from these three categories.

Schools planning to use letters to the editor, individually or in sets, personal letters, diary or journal entries, résumés, job applications and similar as the basis of a summative assessment task should consider how this material and the task demands are congruent with the principle of increasing complexity of challenge as applied in the school work program.

Minimum requirements — mandatory spoken/signed categories

Among the tasks in the verification folio there *must* be at least two of the following:

- an expository text, e.g. media analysis, seminar, panel discussion on a significant issue
- an imaginative text, e.g. storytelling, dramatic re-creation (such as a transformation of a text), monologue (student-constructed text)
- a persuasive/reflective text suitable for a public audience (live, audiotaped or videotaped, multimodal), e.g. eulogy, Australia Day address, testimonial, taking a position at a public meeting, forum.

All tasks should be carefully contextualised in terms of purpose and audience.

Of the predominantly spoken/signed texts, *at least one* is to be a solo presentation. If a group task is undertaken, assessment criteria are not to assess group outcomes but are to assess the achievement of each individual student. For spoken/signed tasks, the assessment is to be made of the demonstrated achievement and not the group or individual planning processes. The inclusion of student notes and other supporting evidence of student response, as well as an annotated criteria sheet with teacher comments to inform the judgment of the level of achievement, must be provided. This contributes to documentation of the text production process as outlined in 6.3.5.

6.6.1 Confirmation of spoken/signed tasks

It is necessary to supplement the school submission of folios for monitoring and verification with confirmation of two students' spoken/signed assessments. This evidence is to be in one of two forms: videotape or audiotape.

This means that each *school's submission* for Year 11 monitoring and Year 12 verification *must* include a videotape or an audiotape that has two presentations of the one spoken/signed task — *one* representing a typical *A* standard and *one* representing a typical *C* standard. If there is no *A* standard available then the next highest presentation in the cohort should be supplied.

The tape should be cued and clearly labelled with school name and school code. It should be accompanied by the task sheet including task-specific criteria and standards. The achievement of each of the two students in the presentations should also be labelled on the tape with a statement about their placement within the band, such as “mid Sound”. Scripts or student notes, such as palm cards, must also accompany the sample tapes.

Videotapes or audiotapes do not have to illustrate presentations of the individual students whose folios have been included in the submission.

The student response submitted as an A standard for the spoken/signed task does not have to be that of a student whose overall level of achievement may be VHA. Similarly, the student

response submitted as a C standard for the spoken/signed task does not have to be that of a student whose overall level of achievement may be SA. At verification, the tape is meant to confirm teacher judgments, and so sophisticated recording techniques and extensive editing are not required. The tape provides an authentic record of the two students' presentations as they occurred in their selected contexts.

6.6.2 Authentication of 'prepared' tasks

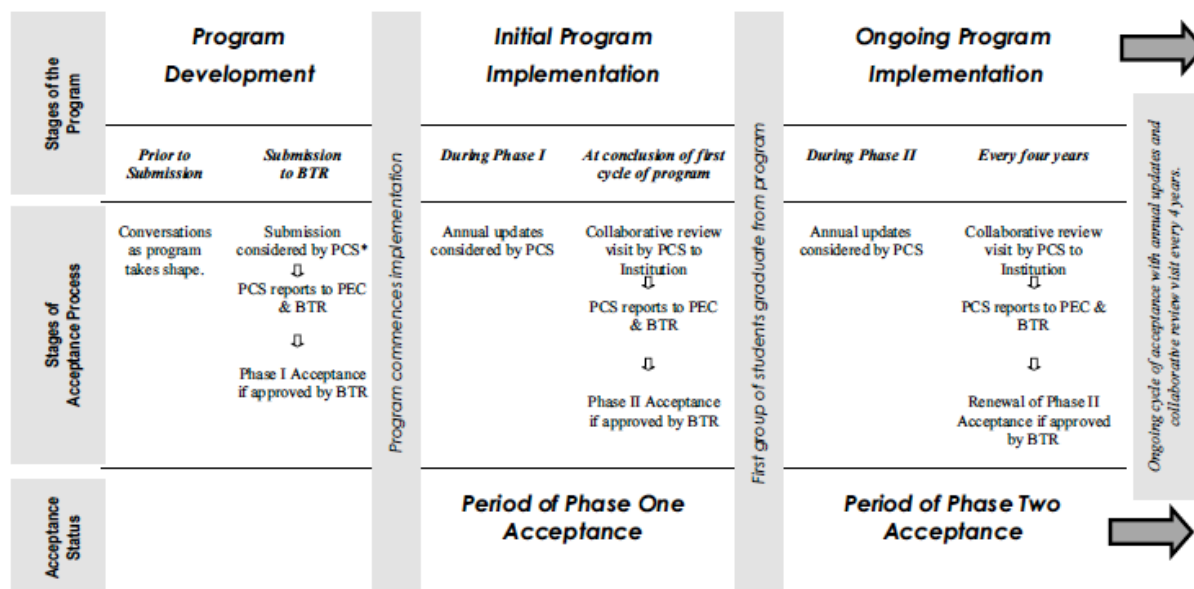
The following procedures are recommended as a way of affirming that a response to a task is genuinely that of the student:

- the teacher must monitor the development of the task by seeing plans and a draft of the student's work
- the student must produce and maintain documentation (including drafts) of the development of the response
- the student must acknowledge all resources used. This will include text and source material and the type of assistance received.

Schools will need to develop guidelines and proformas for students in relation to print and electronic source materials and resources, and to other types of assistance (including human) that have been accessed.

9.10 Program Consultation and Acceptance Process, Board of Teacher Education, Queensland

The Program Consultation and Acceptance Process



9.11 Intervjuguide Norge

Innledning

Mitt navn er Sverre Tveit. Jeg er mastergradsstudent i Pedagogikk ved Universitet i Oslo og gjør en sammenliknende studie av norsk og australsk lærerutdanning i forhold til temaet elevvurdering. Jeg har Kirsti Klette, professor ved Pedagogisk Forskningsinstitutt, som veileder.

- Opplysninger innhentes ved hjelp av dokumentanalyse av kursplanene for lærerutdanningene, analyse av lærebøker og intervju med lærerstudenter.
- Opplysningene vil bli brukt til å analysere lærerutdanningenes fokus på elevvurdering og lærerstudentenes kunnskaper om elevvurdering.
- Det vil bli benyttet pseudonymer for å aidentifisere respondentene.
- Prosjektet varer fra mai 2007 til mai 2008. Opplysningene vil være aidentifisert innen 31.12.2007 og bli lagret med tanke på videre bruk i forbindelse med senere prosjekter.
- Deltakelse er frivillig og samtykke kan trekkes tilbake på hvilket som helst tidspunkt uten at dette grunngis.
- Jeg er underlagt taushetsplikt og dataene behandles konfidensielt.
- Prosjektet er meldt til og godkjent av Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste AS.

I den norske delen av studien har jeg spesielt fokus på fagene matematikk og norsk. Du er valgt ut som respondent på bakgrunn av at du har avsluttet din lærerutdanning i vår og at du har hatt fordypning i Matematikk for ungdomstrinnet. Du vil få spørsmål både knyttet til matematikkfaget spesielt, og utdanningen din generelt. Jeg ber deg svare så ærlig som mulig, helt uavhengig av hvilke svar du tror jeg ønsker. Slik kan du bidra til å gi meg dypere kjennskap til innholdet i din utdanning, mer enn det jeg kan få ved å studere pensumlitteratur og fagplaner. Og jeg vil få et bilde av hvilke tanker du gjør deg om temaet elevvurdering – nå ved avslutningen av din lærerutdanning.

Jeg trenger et muntlig samtykke fra deg for å kunne gjennomføre denne studien.

Jeg vil gjennomføre et forholdsvis strengt strukturert intervju. Det vil si at jeg har forberedt en del spørsmål og for noen av spørsmålene vil jeg be deg svare innenfor kategoriene jeg presenterer for deg. Om det er noe du ikke oppfatter, eller noe du vil be meg gjenta, står vi begge fritt til oppklaringsspørsmål for avklaringer underveis. Jeg kan også komme til å stille noen oppfølgingsspørsmål til det du sier. Vi kan komme til å oppleve utover i intervjuet at du alt har svart på noen av spørsmålene. Jeg vil da likevel at du skal svare på nytt med mindre jeg sier noe annet.

Personalia

Dataene vil bli anonymisert og du vil i undersøkelsen bli benyttet pseudonymer, det vil si at ditt og alle andres navn i datamaterialet vil bli endret. Jeg trenger ellers å vite:

1. Hvilket år ble du født:
2. Kjønn:

Bakgrunn

Jeg vil nå stille deg noen spørsmål om omfanget av din utdanning og erfaring som lærer.

3. Hvilke undervisningsfag har du hatt på lærerutdanningen og hvilket omfang hadde opplæringen i disse fagene? (fag og studiepoeng)

(Informanten blir bedt om å fylle ut en liste med undervisningsfag 1-4 evt. flere)
4. Hvor mye praksis har du hatt som del av lærerutdanningen? (Antall hele uker: + eventuelt antall dager)?
5. Hvor mye praksis har du hatt i de ulike undervisningsfagene? (Uker/timer)
(Samme rekkefølge som i listen for spørsmål 3 benyttes)
6. Hvor mye lærererfaring har du hatt utenom lærerutdanningen? (år eventuelt måneder?)
7. Har du noen annen høyere utdanning?
 - a. Hvis ja, hva slags?

Organisering av innhold knyttet til elevvurdering i lærerutdanningen

Jeg vil nå stille deg om noen spørsmål om organiseringen av innhold knyttet til elevvurdering i lærerutdanningen. Med elevvurdering menes både underveisvurdering og sluttvurdering. Det vil si at både tilbakemeldinger som skal bidra til læring og karaktervurdering er tema her.

8. Har du hatt et eget fag på lærerutdanningen som dreide seg spesielt om elevvurdering?
 - a. Hvis ja: Hvilket fag?
 - b. Hvis ja: Hvilket omfang (antall poeng/timer)
 - c. Hvis ja: Hva var hovedtemaer i dette faget?

9. Ble emnet elevvurdering vektlagt i matematikk?
 - a. Hvis ja: Hvilket omfang hadde faget/fagene? (antall poeng/timer/oppgaver)
 - b. Hvis ja: Hvor stor andel handlet om elevvurdering? (antall poeng/timer/oppgaver)
 - c. Hvis ja: Hva var hovedtemaer i dette emnet?

10. Har du hatt emner under andre fag som dreide seg spesielt om elevvurdering?
 - a. Hvis ja: Hvilke fag?
 - b. Hvis ja: Hvilket omfang hadde faget (Antall poeng/timer/oppgaver)
 - c. Hvis ja: Hvor stor andel av faget handlet om elevvurdering? (Antall poeng/timer/oppgaver)
 - d. Hvis ja: Hva var hovedtemaene i dette emnet?

11. Hva slags pensumlitteratur eller tilleggslitteratur har du hatt som dreide seg om elevvurdering? Nevn navn på bøker eller forfattere, om mulig.

Tilegnelse av kunnskaper om og erfaringer med elevvurdering

Nå ønsker jeg å stille deg syv spørsmål (spørsmål 12 – 18) om hvordan du har tilegnet deg kunnskaper om og erfaringer med elevvurdering. Jeg kommer til å be deg svare innenfor følgende kategorier og vil be deg svare punktvis for de kategoriene som er relevante for deg.

(Denne oversikten utdeles på ark til informantene):

- A) Gjennom opplæring i matematikk på lærerutdanningen?

-
- B) Gjennom praksis i matematikk som en del av lærerutdanningen?
 - C) Gjennom arbeidserfaring i matematikk utenom lærerutdanningen?

 - D) Gjennom opplæring i andre fag på lærerutdanningen?
 - E) Gjennom praksis i andre fag som en del av lærerutdanningen?
 - F) Gjennom arbeidserfaring i andre fag utenom lærerutdanningen?

 - G) På andre måter?

12. På hvilke arenaer mener du at du har lært mye om elevvurdering og hva var det viktigste du lærte?

13. Hva slags verktøy og hjelpemidler til vurderingsarbeidet har du lært å bruke?

14. Hvilke kunnskaper har du fått om bruk av kriterier for elevenes oppnåelse av målene i læreplanen? (stikkord om nødvendig: kriterier for vurdering av nasjonale prøver, prøver utviklet for skolen eller distriktet, sentralt gitte eksamener eller lignende)

15. På hvilke måter har du med andre lærere eller studenter diskutert bruk av kriterier for vurdering av elevenes oppnåelse av målene i læreplanen?

16. Fortel om dine kunnskaper om og erfaringer med å sette karakter på elevs arbeid.

17. Fortel om dine kunnskaper om og erfaringer med å gi elever tilbakemelding på læringsarbeidet. Skill mellom muntlige og skriftlige tilbakemeldinger.

18. Hva slags opplæring har du fått i forskrifter og retningslinjer for elevvurdering?

Synspunkter og erfaringer

Jeg går nå over til å spørre deg om dine synspunkter og erfaringer på noen utvalgte problemstillinger.

19. Hva mener du er viktig når man gir tilbakemeldinger for at disse tilbakemeldingene skal bidra til mer læring?

20. Hvilke tanker gjør du deg om å sette en samlet fagkarakter for hver elev?
21. Hva tror du vil være dine bekymringer når du setter karakterer?
22. Hvordan kan du sikre at din karaktersetting blir så rettferdig som mulig?
23. Mange lærere gir uttrykk for følgende: Det er vanskelig å forholde seg til krav om felles kriterier for karaktervurdering, samtidig som opplæringen skal tilpasses den enkelte elev. Hvilke tanker gjør du deg om dette?
24. Hva skulle du ønske at du hadde lært mer om i lærerutdanningen knyttet til elevvurdering?
25. Kan jeg få lov til å ta kontakt med deg senere ved en eventuell oppfølgingsundersøkelse?

Fordypningsfag

- a. [Fordypningsfag1]
- b. [Fordypningsfag 2]
- c. [Fordypningsfag 3]
- d. [Fordypningsfag 4]

Spørsmål 12 – 17:

- A) Gjennom opplæring i matematikk på lærerutdanningen?
- B) Gjennom praksis i matematikk som en del av lærerutdanningen?
- C) Gjennom arbeidserfaring i matematikk utenom lærerutdanningen?

- D) Gjennom opplæring i andre fag på lærerutdanningen?
- E) Gjennom praksis i andre fag som en del av lærerutdanningen?
- F) Gjennom arbeidserfaring i andre fag utenom lærerutdanningen?

- G) På andre måter?

9.12 Intervjuguide Queensland

Personal information

The data will be made anonymous and pseudonyms will be used, which means that your's and others' names will be changed. Of personal information I will need to know:

1. In what year where you born?
2. Gender:

Background

I will now ask you some questions about the duration of education and experience as a teacher.

3. What curriculum specific courses have you undertaken in your educational studies (teacher training studies) and what were the duration (what time-length) of these courses? (credits)
 - The interviewee will be asked to fill a list with in-depth courses
4. How much in-service work experience did you have as part of the educational studies (practicum and internship)?
 - e. Number of full weeks + possibly days
5. How much in-service work experience did you have in your curriculum specific subjects as part of the educational studies (practicum and internship)?
 - f. Number of full weeks + possibly days.
 - g. Same order as in question 3
6. How much prior experience with teaching have you had outside the educational studies?
 - h. Years, possibly months
7. Have you had any other courses / classes in higher education (from Universities, Colleges)?

- i. If yes, what?

Organising of content related to student assessment in the educational studies

I would now like to ask you questions about the organisation of the content related to student assessment in the educational studies. When using the term «student assessment» both formative and summative aspects will be included. That is, both feedback and marking will be discussed.

8. Did you have a specific course during your educational studies (teacher training) with a special emphasise on student assessment?
 - a. If yes: Type of course?
 - b. If yes: Duration? (credits/hours)
 - c. If yes: Main topics of this course?

9. Was the topic student assessment emphasised in your Mathematics/English courses/classes? [Course code to be added]?
 - a. If yes: Duration? (credits/hours/assignments or exams)
 - b. If yes: The amount of time used on assessment specifically (roughly estimated)? (credits/hours/assignments or exams)
 - c. If yes: Main topics (about assessment?) in this course?

10. Have you had sections of other courses which emphasised/ payed attention to student assessment?
 - a. If yes: Duration of the course/courses? (credits/hours/assignments or exams)
 - b. If yes: The amount of time used on assessment specifically (roughly estimated?) (credits/hours/assignments or exams)
 - c. If yes: Main topics (about assessment?) in this course?

11. What required texts, books (and other reading materials) within the courses focusing on student assessment did you have?
 - d. Please list names on books, texts and authors if possible.

Acquire of knowledge and experience on student assessment

Now I would like to ask you seven questions (question 12-18) about how you have acquired knowledge and experience with and about student assessment. I will ask you to respond throughout the following categories, and I will ask you to respond point by point, referring to the categories which seem relevant to you.

(This list will be provided to the interviewees while answering question 12-18):

- A) During courses/classes in English in the university?
- B) Through in-service work experience (practicum and internship) in English as part of the educational studies?
- C) Through prior experiences in English outside of the educational studies (professional career within or outside the education system)?

- D) During other courses/classes in the university?
- E) Through in-service work experience (practicum and internship) in other courses/classes as part of the educational studies?
- F) Through prior experiences in other courses outside of the educational studies (professional career within or outside the education system)?

- G) Through other sources?

12. Where do you think you learned a lot about student assessment and what was the most important thing you learned?

13. What kind of instruments and other resources in students' assessment have you learned to use?

14. What did you learn about the use of criteria for identifying students' achievements (regarding for example learning objectives)?

- e. If necessary mention criteria for state tests, school or district developed tests, examinations etc

15. In what ways have you discussed the use of criteria for assessing students' achievements of the learning objectives with other teachers or university students?

16. Can you tell me about your knowledge and experiences with marking students' work?
17. Can you tell me about your knowledge and experiences on giving feedback on students' work? Please elaborate on oral versus written feedback.
18. What training have you had in the use of legislations and formal guidelines for student assessment?

Perspectives and experiences

I will now be asking you questions about your perspectives and experiences on some selected issues.

19. What do you think is important, when giving feedback to students, and for this feedback to contribute to improved learning?
20. What concerns or worries do you have about giving a final subject grade for the students?
21. What do you think will be your worries while marking students?
22. How can you make sure that your marking is as fair as possible?
23. Many teachers express the following statement: «It is difficult to cope with the shared requirements for marking, at the same time as the teaching should be adapted to the individual student». What are your comments to this?
24. What would you have liked to learn more about related to student assessment?
25. Will I be allowed to contact you for a possible follow-up study?

Handout

- A) During courses/classes in English in the university?

- B) Through in-service work experience (practicum and internship) in English as part of the educational studies?
- C) Through prior experiences in English outside of the educational studies (professional career within or outside the education system)?
- D) During other courses/classes in the university?
- E) Through in-service work experience (practicum and internship) in other courses/classes as part of the educational studies?
- F) Through prior experiences in other courses outside of the educational studies (professional career within or outside the education system)?
- G) Through other sources?

9.13 Eksempel på framstilling av data (spørsmål 11)

Spørsmål 11	Hva slags pensumlitteratur eller tilleggslitteratur har du hatt som dreide seg om elevvurdering? Nevn navn på bøker eller forfattere, om mulig.	Analyse av enkeltintervjuer	Analyse på tvers
Camilla, <i>Norsk U</i>	<p>Ja, vi har en bok som heter. Som er om prosessorientert skriving, og der stod det jo litt om mappevurdering da. Men det er først og fremst det vi har hatt i forbindelse med... jeg husker vi hadde litt om det i Matten, men vi hadde ikke noen egen bok på det der. Men den boken lurer jeg på om var en ped-bok i andre klasse, den jeg sa til deg nå. Så det har ikke vært noe sånn...</p> <p>ST: Den prosessorientert skriving...?</p> <p>C: Ja. Eller norskbok var det kanskje, men det var sånn obligatorisk i løpet av de to første årene.</p>	<p>Camilla forteller om en bok om prosessorientert skriving der det stod litt om mappevurdering. Hun er usikker på om den var i Norsk eller i Pedagogikk, men den var en del av de obligatoriske fagene de to første årene.</p> <p>Camilla forteller også at hun hadde litt pensum om elevvurdering i Matematikk, men ikke en egen bok på det.</p>	<p>Camilla nevner spesifikt en bok om prosessorientert skriving der det stod litt om mappevurdering som var en del av det obligatoriske Norsk- eller Pedagogikkfaget et av de to første årene.</p> <p>Trygve kom ikke på noen spesielle bøker umiddelbart. Han nevnte en bok kalt</p>
Trygve, <i>Norsk U</i>	<p>T: (5 sekunder stille)</p> <p>ST: Om det er noen spesielle bøker eller...</p> <p>T: Nei, jeg kan ikke komme på i farta noen spesielle bøker... Men det er vel... Vi hadde noen fag... En som het fagdebatikk i hvert fall, men den dreier seg kanskje mer om fagets innhold og ikke så mye om vurdering. Eller så har det vært en del, ja noen didaktikkbøker, som KRL-didaktikk og Norsk-didaktikk og sånne ting. Men det har ikke vært noen egen som heter</p>	<p>Trygve kom ikke på noen spesielle bøker umiddelbart. Han nevnte en bok kalt «Fagdebatikk» men la til at den nok dreide seg mer om fagets innhold og ikke så mye om vurdering. Han nevner også didaktikkbøker i</p>	<p>«Fagdebatikk» men la til at den nok dreide seg mer om fagets innhold og ikke så mye om vurdering.</p> <p>Olav forteller om en bok i Matematikk som tok opp den</p>

	<p>på en måte elevvurdering, det har vi aldri hatt nei.</p> <p>ST: Men har du, har det vært vektlagt... Du husker ikke noe som er spesielt, noe du har lært spesielt bra om det?</p> <p>T: Nei, egentlig ikke, annet enn hva lærerne har fortalt selv ut fra sine egne erfaringer, kanskje.</p>	<p>KRL og Norsk, men sier at det ikke har vært noen egen på elevvurdering.</p> <p>På spørsmål om han kan huske noe der det har blitt vektlagt spesielt, der han har lært spesielt bra om det, så svarer Trygve « Nei, egentlig ikke, annet enn hva lærerne har fortalt selv ut fra sine egne erfaringer, kanskje».</p>	<p>diagnostiske testingen.</p> <p>Eli forteller at i matematikk er «Det eneste vi har fått var et sånt foredragshefte eller notater som hun hadde laget til oss, om vurdering». Etter å ha tenkt seg litt om kommer hun også på en bok av Geir Botten kalt <i>Meningsfylt matematikk</i> som hun hadde i matematikken på første avdeling. «vi hadde det mye moro med hans «retting med grønt» (...) Han var veldig i mot at man skulle bruke <i>rødt</i> da. Så det var retting med grønt, det gjorde det mye hyggeligere» (ler).</p> <p>Eli nevner også boka <i>Kan læring planlegges</i>, av Engelsen.</p>
<p>Amy, <i>English</i></p>	<p>A: How many, like how many text books or?</p> <p>ST: No, yes, what type of text books that were focusing on student assessment.</p> <p>A: Ok. In the subject Assessing Student Learning there was actually a text book. I don't know who is was done by but it was called Assessing Student Learning, so it was pretty much tailored from the text book. Yes, so I don't know whether it was a QSA text book, or... Did you want some? (pouring water). I couldn't remember. But with the curriculum subjects, there were just different sections of my convenor, of my the curriculum subjects, actually did up his own course notes. So they were photo copies from a number of different resources, and he put together it in a book, so... And in that it had his personal notes as well, so it wasn't really a text book, with English curriculum it was more a combination...</p> <p>ST: So that was the English subject?</p>	<p>Amy forteller om en bok i faget Assessing Student Learning. Hun mente den hadde samme navn som faget, men fant senere ut at den heter «Celebrate Students' Achievements». Hun synes denne boken var veldig nyttig fordi den var delt opp tematisk i «making judgments, creating tasks, you know...»</p> <p>For faget English forteller Amy om forskjellige «course notes» som foreleseren hadde satt sammen til et kompendium som også inneholdt</p>	<p>av Geir Botten kalt <i>Meningsfylt matematikk</i> som hun hadde i matematikken på første avdeling. «vi hadde det mye moro med hans «retting med grønt» (...) Han var veldig i mot at man skulle bruke <i>rødt</i> da. Så det var retting med grønt, det gjorde det mye hyggeligere» (ler).</p> <p>Eli nevner også boka <i>Kan læring planlegges</i>, av Engelsen.</p>

	<p>A: Yes.</p> <p>ST: So in Assessing Student Learning, you had a book called Assessing Student Learning? (A: Yes). Which was the only book in that subject? (A: Yep). So that was like the whole book was part of your course readings?</p> <p>A: Yes, it was pretty much divided, it was a really handy text book because it divided, I mean when they find a text book, and they tailor the subject to the text book, you know how they do it... And so, the chapters were all divided into, your making judgments, creating tasks, you know... So it was quite handy.</p> <p>ST: And then in English, some of the course readings were on student assessment?</p> <p>A: Yes, that's right. So he do, say, a reading from somewhere, and then he pull out of the book, an article or something...</p> <p>ST: So when did you do this Assessing Student Learning course?</p> <p>A: Um, fourth year, first semester.</p> <p>ST: So that was in January? No...</p> <p>A: Yes, earlier this year, yes.</p> <p>ST: And when did you do English?</p> <p>A: Last year. English A? So, February/March last year. And curriculum B was about July/August.</p> <p>ST: Do you still have these books?</p>	<p>kommentarer fra ham.</p> <p>Amy mener også det må ha vært noe litteratur på assessment i History (Social Science) men kan ikke huske noen spesielle.</p>	<p>Ellers så nevner de norske studentene fagene Norsk, Matematikk, KRL og Pedagogikk der de mener det var noe om elevvurdering, men det var ikke sentralt og det var ingen egne bøker om temaet.</p> <p>På spørsmål om han kan huske noe der det har blitt vektlagt spesielt, der han har lært spesielt bra om det, så svarer Trygve « Nei, egentlig ikke, annet enn hva lærerne har fortalt selv ut fra sine egne erfaringer, kanskje».</p> <p>Begge English-studentene i Queensland forteller om en spesifikk bok i faget Assessing Student</p>
--	---	---	---

	<p>A: Yes.</p> <p>ST: Would it be possible for me to borrow them?</p> <p>A: Yes.</p> <p>ST: Because I will need to analyse them and... I could... I'll probably buy the one that is a book, but the one that is made together by your convenor is probably difficult for me to find in other ways than through you actually...</p> <p>A: Sure. Yep, that'll be fine.</p> <p>ST: Yes, excellent, I'll get back to you on that some later.</p> <p>A: Ok, no problem.</p> <p>ST: Yes, so that's about it. These course reading materials in the English subject and the Assessing Student Learning book that was about student assessment. Is there any other...? In History...?</p> <p>A: You see it's hard to say that these books were mainly about, with curriculum studies were mainly about assessment, because really it was only sections of these books that were about assessment, because a lot of it was about, you know, the actual structure of the 1 to 10 syllabus, so... What was the question again?</p> <p>ST: In History, for example, or other subjects, did you have any about student assessment there?</p> <p>A: Um, yeah. I got another, didn't I...? History is a very, more straight forward, and everything</p>	<p>Learning. Boken, som heter Celebrating Student Achievement, blir av begge beskrevet som nyttig.</p> <p>Philip nevner ikke denne boken. Med forbehold om at hans pensum var annerledes er det grunn til å tro at Philip ikke har vært oppmerksom på at det var en bok kalt «Celebrating Student Achievement» på pensum. Han forteller imidlertid om et kompendium i Assessing Student Learning-faget, som læreren hadde satt sammen.</p> <p>Amy forteller i tillegg om et kompendium i English der læreren hadde satt sammen en</p>
--	--	---

	<p>is on the net, so you can have unit outlines and assessment tasks on the net, that you can download. And so, they tell you how much time you should spend on each area before you start getting into the assessment task and things like that, so... Whether any other books, or, is that the question?</p> <p>ST: Yes.</p> <p>A: Yes, there must have been, but I can't remember (laughing)</p>		<p>rekke artikler, der hun mener noe av innholdet gikk på elevvurdering.</p> <p>Amy mener det også må ha vært noe i History (Social Science), men husker ingen spesielle.</p>
Royce, <i>English</i>	<p>R: We, oh, we had the syllabus that we had to use, they were very important things. There was another text book that I didn't actually buy or use, but for that Assessing Student Learning course there was a small text book that was pretty good I suppose.</p> <p>ST: But you didn't buy it?</p> <p>R: I did, I did buy that one?</p> <p>ST: Which course was that in, the one you didn't buy?</p> <p>R: The HP curriculum subjects.</p> <p>ST: Ok.</p>	<p>Royce forteller at læreplanene var sentrale deler i fagene og at vurdering er en sentral del av dem.</p> <p>Royce forteller også om en lærebok i Assessing Student Learning (Celebrating student achievement) som var «pretty good I suppose».</p> <p>Han nevner også en bok i HPE som han ikke husker hva heter og som han aldri kjøpte.</p>	<p>Royce trekker også fram læreplanen som sentral litteratur i fagene.</p> <p>Royce nevner også en bok i HPE som han ikke husker hva heter og som han aldri kjøpte.</p> <p>Damion forteller at de hadde en lærebok i Assessing Student Learning og en «book of readings» med «references to,</p>
Olav, <i>Matematikk</i>	<p>ST: Ja, da går jeg til spørsmål 11: Hva slags pensumlitteratur eller tilleggslitteratur har du hatt som dreide seg om elevvurdering? Nevn navn på bøker eller forfattere, om mulig.</p> <p>O: Ja. Vi hadde, i Matematikken så hadde vi</p>	<p>Olav forteller om en bok som tok opp den diagnostiske testingen. Han mener også han fikk noen stensiler i</p>	<p>like Queensland Education site, like QCAR, and different sort of Government stuff.» Han</p>

	<p>som gikk på den, i hvert fall den diagnostiske testingen, jeg mener den står på pensumlista vår, hvis ikke så kunne jeg sikkert finne det, og sende en tekstmelding på det eventuelt. Det kunne jeg gjort, for jeg har ikke det i hodet. Og så fikk vi vel noen ark, mener jeg, eller stensiler, som gikk så vidt det var innom vurderingen. Men ikke noe sånn rene bøker som har tatt for seg vurdering, det har det ikke vært. Det kan, det er klart nå husker ikke jeg helt hva som var i Ped'n i første og andre klasse, men der lurte jeg på om det var noe litteratur som gikk på den vurderingsbiten.</p> <p>ST: Ja, det skal jeg undersøke nærmere også.</p> <p>O: Mh.</p>	<p>Matematikken. Han lurte også på om det var noe litteratur på elevvurdering i Ped'n, men har ikke hatt noen rene bøker som gikk på elevvurdering.</p>	<p>forteller at det var nyttig. Han har ikke hatt noen annen litteratur om elevvurdering i andre fag.</p> <p>Noen hovedforskjeller mellom landene og fagene kommer fram her:</p> <p>I Queensland har</p>
<p>Eli, <i>Matematikk</i></p>	<p>ST: Hva slags pensumlitteratur eller tillegglitteratur har du hatt som dreide seg om elevvurdering? Husker du noe av det?</p> <p>E: I Matten i år så har vi ikke... Det eneste vi har fått var et sånt foredragshefte eller notater som hun hadde laget til oss, om vurdering. Vi har ikke hatt om det i noen bøker.</p> <p>ST: Så det var et kompendium, eller?</p> <p>E: Ja, vi kan godt kalle det det. Altså, det var en powerpoint-presentasjon som hun har kjørt ut bare, på print.</p> <p>ST: Å, så du føler ikke det har vært tema i noe av det som har vært pensum, altså pensumbøker?</p> <p>E: Nei.</p> <p>ST: Ikke i Matematikk, eller noen andre fag heller?</p>	<p>Eli forteller at i matematikk er «Det eneste vi har fått var et sånt foredragshefte eller notater som hun hadde laget til oss, om vurdering». Etter å ha tenkt seg litt om kommer hun også på en bok av Geir Botten kalt <i>Meningsfylt matematikk</i> som hun hadde i matematikken på første avdeling. «vi hadde det mye moro med hans «retting med grønt» (...) Han var veldig i mot at man skulle bruke <i>rødt</i> da. Så det var retting med grønt, det gjorde</p>	<p>de betydelig mer litteratur på elevvurdering i tilknytning til det obligatoriske faget Assessing Student Learning, blant annet en egen lærebok om elevvurdering «Celebrating Student Achievement» og et kompendium med artikler satt sammen av læreren.</p> <p>I Queensland henviser også studentene til læreplanene som</p>

	<p>E: Eh... Nå er det lenge siden vi har hatt det andre faget. Vi har jo hatt Ped som eget fag, og der har det vært egne kapitler om det.</p> <p>ST: Husker du hvilke bøker det var?</p> <p>E: Vi har blant annet hatt Engelsen, ja.</p> <p>ST: Engelsen. Er det <i>Kan læring planlegges</i> boka?</p> <p>E: Ja. Og så har vi, jeg kom på han Geir Botten (sjekk navn!), vi hadde det mye moro med hans «retting med grønt». Det var en egen Matte-bok, <i>Meningsfylt Matematikk</i>, der stod også... Men det var ikke fordypningsfag, men det var...</p> <p>ST: Nei, men hva het han sa du?</p> <p>E: Geir Botten.</p> <p>ST: Geir Botten. Og den het <i>Retting med grønt</i>?</p> <p>E: Neida (ler). Han var veldig i mot at man skulle bruke <i>rødt</i> da. Så det var retting med grønt, det gjorde det mye hyggeligere. <i>Meningsfylt Matematikk</i> het den boka.</p> <p>ST: Ok (ler). Men det var ikke fordypning, det var tidligere...?</p> <p>E: Ja.</p>	<p>det mye hyggeligere» (ler).</p> <p>Eli nevner også boka <i>Kan læring planlegges</i>, av Engelsen.</p>	<p>sentral litteratur for vurderingsfeltet. Læreplanen i Queensland inneholder flere siders omtale av elevvurdering i tillegg til «learning outcomes», mens det norske læreplanverket kun består av kompetansemålene.</p> <p>I Norge har ingen studenter hatt noe spesifikt fag på elevvurdering og de har ikke hatt noen bøker som går bare på dette. En Norsk-student forteller om en bok kalt «Proessorientert skriveopplæring» der mappevurdering var sentralt. I</p>
Philip, <i>Mathematics</i>	<p>ST: What required texts, books and other reading materials within the courses focusing on student assessment did you have?</p> <p>P: We had the readings for Assessing Student Learning, which was a whole book on looking at assessment, and all different topics about</p>	<p>Philip forteller om et kompendium i <i>Assessing Student Learning</i> som læreren hadde satt sammen. På direkte spørsmål om det var noen</p>	<p>Matematikk nevnes det at diagnostisk testing ble tatt opp i en bok. Ellers forteller studentene at de</p>

	<p>assessment.</p> <p>ST: So that was a book? And was there also some course readings?</p> <p>P: Yes, there were just course readings. There wasn't an actual text book. He just put together a whole heap of readings from...</p> <p>ST: When did you take this course?</p> <p>P: Um, in the beginning, first semester this year, so first semester of fourth year.</p> <p>ST: Did you like it?</p> <p>P: Um, yes there were some parts that were interesting, but I think a lot of it, because of the Maths and Physics side is a lot more exam based, where it's just answering questions, where a lot applied to more performance based subjects like HPE and Drama and more assignment, yes, more subjects that have assignments in them rather than straight exams.</p> <p>ST: So you didn't feel that it was that relevant to your subjects?</p> <p>P: Yes, not as relevant as the other subjects.</p> <p>ST: Sure.</p>	<p>bøker svarer han nei.</p> <p>Med forbehold om at hans pensum var annerledes er det grunn til å tro at Philip ikke har vært oppmerksom på at det var en bok kalt «Celebrating Student Achievement» på pensum .</p> <p>På spørsmål om han likte faget gir Philip uttrykk for at deler av faget var interessant men at det ikke var så relevant for han som fordypte seg Mathematics og Physics: «the Maths and Physics side is a lot more exam based, where it's just answering questions, where a lot applied to more performance based subjects like HPE and Drama and more assignment, yes, more subjects that have assignments in them rather than straight exams».</p> <p>[Dette er flyttet til spørsmål 8).</p>	<p>tror elevvurdering ble tatt opp i noen bøker i KRL og Pedagogikk.</p> <p>Studentenes fra Queenslands rapporter om litteratur gir grunn til å stille spørsmål ved om elevvurdering i større grad er representert på pensum i English enn i Mathematics. En senere sub-analyse av Curriculum-fagenes pensum vil kunne avdekke dette. Tilsvarende undersøkelse kan være interessant å gjøre av de norske pensumlistene.</p>
Damion, <i>Mathematics</i>	ST: Yes. Ok. What required texts, books and other reading materials within the courses	Damion forteller at de hadde en lærebok i	

	<p>focusing on student assessment did you have?</p> <p>D: Um, there was, in that subject Assessing Student Learning there was a text book and a book of readings, and then there was also a lot of references to, like Queensland Education site, like QCAR, and different sort of Government stuff.</p> <p>ST: Ok. So that was the Assessing Student Learning, was there any texts about assessment in the other courses, or in the curriculum subjects?</p> <p>D: No.</p> <p>ST: So did you find this text book and the course readings for Assessing Student Learning useful?</p> <p>D: Yes, in that subject, well because that subject was after the curriculum ones, and I didn't really need to have very much assessment during the curriculum, I didn't really need it, and then when I did the Assessing Student Learning subject, I used that text book a lot. Especially for when doing assignments, getting quotes and backing up ideas (unclear).</p> <p>ST: Yes, so you found it useful?</p> <p>D: Yes.</p> <p>ST: And the course readings also?</p> <p>D: Yes.</p> <p>ST: Great.</p>	<p>Assessing Student Learning og en «book of readings» med «references to, like Queensland Education site, like QCAR, and different sort of Government stuff.» Han forteller at det var nyttig. Han har ikke hatt noen annen litteratur om elevvurdering i andre fag.</p>	
--	---	---	--

9.14 Vurderingsoppgaver- og kriterier i faget Norsk 1 ved den undersøkte norske allmennlærerutdanningen

Arbeidskrav

Retten til å gå opp til eksamen føreset godkjende arbeidskrav. Arbeidskrav som ikkje er levert/utført innan fastsett frist, blir ikkje godkjende, og studenten misser retten til å gå opp til eksamen. Studentane er sjølv ansvarlege for å halde seg informert om dei gjeldande fristane. Ved sjukdom eller liknande må studenten, innan fristen for innlevering av arbeidskravet, gjere avtale med faglærer om utsett frist. Arbeidskrav som er levert/utført innan fristen, men som ikkje er blitt godkjent, må leverast på nytt til ny frist. Studentar som har mista retten til å gå opp til eksamen på grunn av ikkje godkjent arbeidskrav, må sjølv kontakte faglærer for avtale om ny innlevering/utføring av arbeidskravet i neste semester/studieår.

Generelt

Vurderingane gjeld dei arbeida som studentane er pålagde å gjere. Vurderinga tar sikte på å øve opp studentane si evne til å reflektere over praktiske og teoretiske tilnæringsmåtar til faget. Studentane skal lære å analysere, vurdere og formulere faglege og metodiske problemstillingar. I studiet lærer studentane å evaluere arbeida til kvarandre og eigne arbeid og prestasjonar. Dei vert oppmuntra til å drøfte vurderingskriterium og premisser for resposarbeid og vert i dette arbeidet oppmoda til å tenke sjølvstendig og nyansert.

Arbeidskrav

Munnleg forteljing (eventyr)

Arbeidet med språk og læring

I tillegg til tekstane i skrivemappa skal studentane lage ei biletbok som gruppearbeid

Eksamen

Kasusoppgåve i norsk i samband med det fleirfaglege temaet skolestart, 3 studiepoeng.

Vurdering: Bestått/ikke bestått

Halvårsoppgåve i norsk og pedagogikk i veke 50. Oppgåva gir 5 studiepoeng i norsk.

Vurdering: Gradert karakter

Skrivemappa: Første studieår skriv studenten minst 6 tekstar knytt til fagemna i norsk, halvparten på kvar målform. Studenten vel sjølv ut 3 av desse tekstane for skrivemappa og grunnjev dette valet i ein refleksjonstekst. Av desse 4 tekstane skal 2 vere på bokmål og 2 på

nynorsk. Studenten får tilbakemelding på tekstane frå lærar og/eller medstudentar. Tekstane er i ulike sjangrar. Skrivemappa er eit elektronisk produkt. Faglærar skal ha ein papirkopi av skrivemappa.

Vurdering: Gradert karakter, 7 studiepoeng.

Andre studieår

Eksamen

Fordjupingsoppgåva vert bygd på tidlegare arbeid som i alt utgjer minst 11 studiepoeng.

Vurdering: Bestått/ikke bestått. Norsk gir 1 studiepoeng til denne oppgåva som er fleirfagleg.

Miniforelesing med fagleg notat. Forelesinga varer i 15-20 minutt og vert halde over eit oppgitt emne innafor haustens undervisningstema. Studenten lagar eit notat til forelesinga.

Vurdering: Gradert karakter, 7 studiepoeng.

Avsluttande skriftleg eksamen over 6 timar med oppgåver henta frå dei emna som ikkje er gjennomgått i haustsemesteret. Vurdering: Gradert karakter, 7 studiepoeng. Ekstern sensor.

Endeleg vurdering i faget

Samla karakter i faget vert sett ut i frå dei graderte karakterane på halvårsoppgåva og skrivemappa første studieår og miniforelesing med fagleg notat og avsluttande eksamen i andre studieår. Karakterane blir vekta etter studiepoeng.

Karakterskala

Ved gradert karakter blir det gitt bokstavkarakterar med A som beste og E som dårlegaste karakter på bestått eksamen. Karakteren F blir brukt ved ikkje bestått eksamen.

Ved ikkje gradert karakter blir det gitt bestått - ikke bestått.

Vurderingskriterium

Graderte karakterar med forklaringar

Symbol	Betegning	Kvalitativ forklaring for eksamen
A	Framifrå	Arbeidet er ein framifrå prestasjon som klart skil seg ut. Det vitnar om sær s god fagkunnskap, stor grad av sjølvstendig tenking og har sær s godt språk.

B	Mykje godt	Arbeidet er ein mykje god prestasjon som ligg over gjennomsnittet. Det vitnar om mykje god fagkunnskap, sjølvstendig tenking og har mykje godt språk.
C	Godt	Arbeidet er ein gjennomsnittleg prestasjon som har innslag av sjølvstendig tenking og er tilfredsstillande når det gjeld fagkunnskap og språk.
D	Nokså godt	Arbeidet er ein prestasjon under gjennomsnittet med manglar i fagkunnskap og språk.
E	Godt nok	Arbeidet har store manglar i fagkunnskap og språk, men tilfredsstillar minimumskrava.
F	Ikkje godkjent	Arbeidet tilfredsstillar ikkje minimumskrava til fagkunnskap eller språk.

Vurdering bestått/ikkje bestått

Bestått	Kandidaten har vist fagkunnskap og evne til formidling som tilfredsstillar det som vert kravd for å kunne praktisere som lærar i norsk.
Ikkje bestått	Kandidaten har ikkje vist fagkunnskap eller evne til formidling som tilfredsstillar det som vert kravd for å kunne praktisere som lærar i norsk.

9.15 Vurderingsoppgaver- og kriterier I faget English Curriculum A ved den undersøkte lærerutdanningen i Queensland

Item	Assessment Task	Length	Weighting	Total Marks	Relevant Learning Outcomes	Due Day and Time
1.	Assignment	2500 words	50 %	25	1, 2, 3, 4, 5,	
2.	Examination: multiple choice; short- and extended-answer questions	3 hours	50 %	50	2, 3, 5	

NB. A Pass grade must be achieved in each summative assessment item to achieve a Pass grade in this course.

Assessment in this core English Curriculum course reflects the orientation of the course to the development of knowledge, skills and processes that will enable students to make theoretically and empirically informed evaluations and recommendations pertinent to subject English and literacy curriculum: learning outcomes, texts and content, pedagogies, assessment practices and interactional formats evident in existing lower secondary and middle years English units and lessons.

During the course of the short semester, feedback is given on Formative Assessment tasks only. These formative assessment tasks, completed for tutorials and during workshops, reflect the nature and content of the summative assessment tasks.

Tertiary Standards of Literacy and the Grading of Assessment Items

Each assessment item in this course is assessed using **TWO** sets of criteria. The first set of criteria pertains to Achievement of Purpose. The second set of Criteria pertains to **Proficiency of Language Use**. In the first instance submissions are assessed and graded according to the level of achievement with respect the completion of all of the requirements of the task i.e. **Achievement of Purpose**. In addition, both written and spoken texts will be assessed according to the number of technical errors in language use, i.e., **Proficiency of**

Language Use. The latter will have no effect on the **Achievement of Purpose** grading if the technical accuracy of the submission is at an appropriate tertiary standard. However, where there are in excess of 3 technical errors each error will incur a penalty of 1 mark up to a total of 20. Penalty marks will be deducted from the initial grade.

Assessment Items: Tasks, Rationale & Criteria

Assessment Item 1 Assignment 50% (WRITTEN SUBMISSION 2500 words)

Evaluate/Justify the English curriculum apparent in the English unit your group has developed this semester

Using clear headings for each section, you are to provide theoretically grounded evaluations/justifications of the specified aspects of the Unit

Aspects of the Unit to be Evaluated/Justified	Length	Weight
1: Evaluate/justify the validity and appropriateness of the specified learning outcomes, specific content (factual knowledge, conceptual knowledge, procedural knowledge, and skills) to be taught.	500 words	10%
2: Evaluate/justify the appropriateness and coherence of the pedagogy evident in specified teaching strategies, learning activities, learning activity sequences and the selected context for learning.	500 words	10%
3: Evaluate/justify the suitability and value of specified texts and resources to be used for teaching purposes. This section must include 2 examples of text segments that will be used for teaching purposes plus an analysis of the textual elements in each text segment that will be used to model specific facts and concepts to be learned	1000 words	20%
4: Evaluate/justify the validity and appropriateness of specified formative and summative assessment tasks	500 words	10%

To complete this task you are required to demonstrate apply factual and conceptual knowledge and understanding of:

- the current requirements specified for Queensland Middle Years English curriculum,
- principles of planning and assessment suitable for Middle Years English curriculum,

- factual, conceptual and procedural knowledge of the discipline – literary and non-literary
- texts, language elements, and literacy practices – necessary for teaching English in the Middle Years
- contemporary theories of text, meaning and meaning-making, specifically: Literacy as Social Practice, Multiliteracies, and the Four Resources Model,
- contemporary theories of learning, curriculum design and pedagogy, specifically: constructivism, problem-based learning, integrated curriculum, Productive Pedagogies and the Curriculum Area Multiliteracies And Learning framework (CAMAL)

ASSESSMENT ITEM 2 Written Examination – Closed Book 50%

Multiple Choice, Short- and Extended-Answer Questions

Students are assessed on their ability to demonstrate:

Aspects of the Unit to be Evaluated/Justified Length Weight	Weight
1: Factual knowledge of text structures and textual elements of written, spoken and visual texts, English grammar, punctuation and spelling, comprehending and composing strategies and procedures, (i.e., essential knowledge for teaching and assessing student learning in subject English and in other school subjects),	20%
2: Ability to apply conceptual knowledge of the ways that language elements are used in written, spoken and visual texts to construct meaning and to position audiences.	10%
3: Ability to apply principles of unit and lesson planning appropriate for Middle Years English curriculum educational reasoning practices to select and justify specific teaching strategies, lesson sequences, learning activities and assessment tasks for specific learning outcomes. Factual and conceptual knowledge of theories of learning, curriculum design in English, language and literacy	20%

Assessment Rationale and Criteria

Assessment in this initial English Curriculum course reflects the four key aspects of effective curriculum:

1. the multi-literate teacher who has excellent knowledge of the discipline, that is, the individual teacher's knowledge about and ability to operate and evaluate the language systems and associated textual and cultural resources used to make meaning in and from texts;
2. responsive and skilful program design and evaluation,
3. effective theoretically grounded, inclusive pedagogy, and
4. appropriate, valid and reliable assessment programs and practices.

The design of assessment tasks reflects the requirements of the Queensland College of Teachers (QCT), the professional body responsible for the registration of teachers in Queensland.

A professionally oriented degree course needs to provide students with an opportunity to meet the standards required of the official bodies that oversee and register workers in the profession. The QCT requires that in seeking registration as a teacher a candidate provides evidence of:

1. Personal Literacy Levels: knowledge of subject matter to be taught and informed and skilful application of text analytic processes;
2. Knowledge of and skills in the application of current national and state curriculum policy initiatives and innovations, relevant, contemporary theories of language, literacy/multiliteracies, literacy learning and pedagogy to the tasks of evaluating planned appropriate and effective English curriculum;
3. Ability to evaluate curriculum design: choices of content, pedagogical design and assessment practices, that takes account of program knowledge components, aspects of planning and implementing appropriate and effective, inclusive English curriculum and that reflect knowledge of contemporary literacy education and pedagogical theory, and the aims and intent of English curricular specified by the Queensland Studies Authority in the Queensland Curriculum, Assessment and Reporting Framework (QCAR). The assessment items in this course provide for these requirements.

Rationale for Assessment Program

◆ Explicit knowledge and control of the elements and operation of English grammar, spelling, punctuation and other textual elements used to make meaning in and from texts is essential for the English teacher.

-
- ◆ Effective and responsible professional practice as a teacher of English in Queensland schools requires deep knowledge about and ability to critically evaluate contemporary theories of literacy and literacy curriculum and pedagogy and their expressions in Syllabus documents, State and National initiatives and policies relevant to literacy education and equity, classroom programs and lessons; such evaluations forge a necessary link between theory and practice, therefore, classroom teachers need to be skilled evaluators of observed/recorded/transcribed classroom interactions.
 - ◆ Knowledge about and ability to evaluate the language systems and associated textual resources used in written, visual, digital and multi-modal texts is essential for teachers of English in outcomes-oriented curriculum aimed at the development of multi-literate students.
 - ◆ Expert knowledge of the ways that meaning is constructed textually is essential knowledge for teachers of English.
 - ◆ Deep understanding of the links between contemporary theories of literacy and literacy education, pedagogies, and classroom practices is essential for teachers of English.
 - ◆ Advanced skills in comprehension, text analysis and de-construction of texts, are essential for teachers of English.

Rationale for Assessment Method – Item 1

Assessment Item 1 derives from collaborative work in several developmental stages that scaffold students' application of knowledge about contemporary literacy theory, pedagogy, Queensland English Essential Learnings (QSA, 2007), language systems and associated textual resources used in written, visual, digital and multi-modal texts. By working in small groups, students are involved in decision making, problem solving, critical evaluation and strategic planning processes. During group work individuals are required to make rational, well substantiated decisions regarding inclusions in unit plans. It is during negotiations and reflections on these decisions that the ability to critically evaluate and justify curriculum and pedagogical choices is developed. The tasks included in Assessment Item 1 thus derive from and build on the meanings collaboratively developed and negotiated during interactive group sessions.

- ◆ Students to develop a sense of being a part of a sharing academic community;
- ◆ Problem-based learning workshops provide a context for active construction and negotiation of meaning and mutual understandings assists learners involved in developing knowledge and understanding of difficult concepts;

- ◆ Real-life tasks as a context for the application of theoretical concepts assists learners to clarify and consolidate understandings;
- ◆ Developing deep understanding is a process that requires concerted and repetitive engagements with ideas and their applications within a range of contexts and tasks;
- ◆ The requirement that the substantive content of group discussions are written up and justified with reference to set course readings provides a process that enables students to consolidate new knowledge and understandings;
- ◆ Ability to interpret, organise, synthesise and justify ideas is essential to practising professional teachers.

