



Rapport
2018:9

Sluttrapport fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk

Idunn Seland, Elisabeth Hovdhaugen, Berit Lødding, Tine S. Prøitz og
Ester Rønsen

Rapport
2018:9

Sluttrapport fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk

Idunn Seland, Elisabeth Hovdhaugen, Berit Lødding, Tine S. Prøitz og
Ester Rønsen

Rapport	2018:9
Utgitt av Adresse	Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU) Postboks 2815 Tøyen, 0608 Oslo. Besøksadresse: Økernveien 9, 0653 Oslo.
Prosjektnr.	20504
Oppdragsgiver Adresse	Utdanningsdirektoratet Postboks 9359 Grønland, 0135 Oslo
Fotomontasje	Bøk Oslo AS
ISBN	978-82-327-0344-9
ISSN	1892-2597 (online)



Copyright NIFU: CC BY-NC 4.0

www.nifu.no

Forord

NIFU har på oppdrag fra Utdanningsdirektoratet evaluert forsøk med redusert antall karakterer ved halvårsvurdering i norsk. Evalueringsperioden har strukket seg fra 2014 til 2018, og omfatter to delrapporter samt denne sluttrapporten.

Første delrapport (2015) tok for seg litteraturgrunnlaget for evalueringen. Andre delrapport (2016) presenterte kvalitative data om planlegging og iverksetting av forsøket, samt hva som skjer med undervisning og vurdering når lærerne skal sette færre halvårskarakterer i norsk.

Sluttrapportens sentrale bidrag er en analyse av standpunktkarakterer i norsk for elever som har avsluttet studieforberedende utdanningsprogram i videregående skole, og som har vært del av forsøket. I tillegg utdypes tidligere funn med en ny spørreundersøkelse til skoleledere og analyse av nye forskningsintervjuer med lærere og elever ved ungdomsskoler. Sluttrapporten sammenfatter de to delrapportene og gir en avsluttende drøfting av evalueringens hovedfunn.

Leder for evalueringen har vært Idunn Seland. Berit Lødding og Tine S. Prøitz har som forfattere av delrapport 1 og 2 bidratt til sluttrapportens kapittel 1-3. Elisabeth Hovdhaugen har sammen med Ester Rønsen skrevet kapittel 4 med tilhørende metodeseksjon i kapittel 1. Rønsen og Seland har skrevet kapittel 5, Seland har skrevet kapittel 6 og Lødding har skrevet kapittel 7. Sluttkapittelet er ført i pennen av prosjektleder på bakgrunn av en kollektiv diskusjon i prosjektgruppa.

Roger Andre Federici har stått for kvalitetssikring av helheten i rapporten. Det vurderingsfaglige aspektet i evalueringen er kvalitetssikret av Tine S. Prøitz, førsteamanuensis ved Høgskolen i Sørøst-Norge. Karakteranalysen er kvalitetssikret av Rune B. Reiling og Astrid J. Sandsør, begge forskere ved NIFU.

Prosjektleder ønsker å rette en særlig takk til skolelederne, lærerne og elevene som lot seg intervju i forbindelse med denne evalueringen.

Oslo, juni 2018

Sveinung Skule
Direktør

Roger Andre Federici
Forskningsleder

Innhold

Sammendrag	7
Executive summary.....	11
1 Innledning.....	15
1.1 Hvorfor endre på rutineene for halvårsvurdering i norskfaget?	15
1.2 Rapportens hovedproblemstilling og evalueringens siktemål.....	16
1.3 Norskfagets formål og oppgaver i opplæringen	18
1.4 Begrepsavklaring: undervisning	20
1.5 Begrepsavklaring: vurdering	21
1.6 Tidligere forsøk med redusert antall karakterer i norsk.....	25
1.7 Rapportens særlige bidrag til forskningsfeltet	27
1.8 Metode	34
1.9 Rapportens oppbygning.....	42
2 Sentrale dokumenter som styrer undervisning og vurdering i norskfaget.....	44
2.1 Kunnskapsløftet som styringsreform og læreplanreform.....	45
2.2 Læreplan for norskfaget og veiledning til læreplanen.....	46
2.3 Forskrift om individuell vurdering og rundskriv om individuell vurdering	51
2.4 Et internasjonalt perspektiv på vurdering i norsk skole	55
2.5 Nynorskens stilling i norskfaget.....	57
2.6 Oppsummering	58
3 Undervisning og vurdering i norskfaget.....	60
3.1 Tilbakeblikk på litteratursøk i evalueringen.....	60
3.2 Endringer i norskfagets innhold, lærer- og elevrollene	62
3.3 Undervisningspraksiser i norskfaget	63
3.4 Vurdering i norskfaget.....	73
3.5 Oppsummering	82

4	Analyse av karakterdata.....	85
4.1	Avgrensninger av datamaterialet.....	85
4.2	Forventninger til forsøket.....	88
4.3	Forskjeller på tiltaksskoler og sammenligningsskoler.....	89
4.4	Grafiske analyser.....	90
4.5	Regresjonsanalyse med forskjell-i-forskjeller metode.....	92
4.6	Oppsummering og konklusjon.....	94
5	Skoleledernivåets erfaringer fra forsøket	95
5.1	Utvalg og svarprosent.....	95
5.2	Implementering av én eller to karakterer ved halvårsvurdering.....	96
5.3	Forventninger og resultater av forsøket.....	97
5.4	Oppsummering.....	107
6	Norsklærernes erfaringer fra forsøket	108
6.1	Vurdering for læring og vurdering av læring i norskfaget.....	109
6.2	Skolekonteksten.....	120
6.3	Oppsummering.....	122
7	Elevers erfaringer med norskfaget og med forsøket.....	125
7.1	Tolkning og koding av elevmaterialet fra ungdomsskolen.....	126
7.2	Hovedinntrykk fra videregående materialet.....	126
7.3	Norskfaget.....	128
7.4	Overgangen til flere karakterer i norsk.....	130
7.5	Motivasjon og læring gjennom forsøket.....	132
7.6	Oppsummering fra de to nivåene.....	134
8	Videre arbeid med å utvikle vurdering i norskfaget.....	139
8.1	Ingen systematiske forskjeller i elevenes standpunktkarakterer.....	139
8.2	For hvem har forsøket vært hensiktsmessig?.....	141
8.3	Hva styrer undervisning og vurdering i norskfaget?.....	145
8.4	Anbefalinger for videre utvikling av vurdering i norskfaget.....	147
8.5	Avslutning.....	149
	Referanser.....	151
	Vedlegg.....	160
	Vedlegg 1: Oversikt over forskningsspørsmål i evalueringen.....	161
	Vedlegg 2: Regresjonstabeller for analyse av karakterer i kapittel 4.....	165
	Tabelloversikt.....	172
	Figuroversikt.....	173

Sammendrag

Hva er sammenhengen mellom vurdering, undervisning og elevers læring? Dette er spørsmålet vi nærmer oss gjennom denne sluttrapporten fra evaluering av et forsøk hvor lærerne skal sette bare én eller to istedenfor normalt tre halvårskarakterer i norsk.

Halvårsvurdering er karakteren som elevens faglærer setter hvert halvår i ungdomsskolen og i videregående skole helt frem til standpunktkarakteren skal settes ved avslutningen av opplæringsløpet. Normalt skal elevene ha en halvårskarakter i skriftlig hovedmål, en karakter i sidemål og en karakter i norsk muntlig. I forsøket er dette slått sammen til enten én felles karakter eller to karakterer: norsk skriftlig (hovedmål + sidemål) og norsk muntlig.

Forsøket gjelder 8. og 9. trinn i ungdomsskolen og i Vg1 og Vg2 på studieforberevende utdanningsprogram i videregående skole.

Hovedfunn: ingen endring i standpunktkarakterer

I denne rapporten har vi analysert 197 000 elevers standpunktkarakter i norsk ved avslutningen av Vg3 for å se om elevenes læring påvirkes av hvorvidt elevene har vært del av forsøket mens de gikk i Vg1 og Vg2.

- Analysen viser ingen forskjeller i karakternivå mellom elever som har deltatt i forsøket sammenlignet med elever som ikke har deltatt i forsøket.
- Vi finner heller ingen forskjeller i elevenes standpunktkarakterer med utgangspunkt i hvorvidt skolene deres har gjennomført forsøket med henholdsvis én eller to karakterer ved halvårsvurdering i norsk.
- Dette funnet endrer seg ikke selv om vi skiller mellom elever som har gått ut fra ungdomsskolen med et høyere eller et lavere antall grunnskolepoeng (høyt presterende kontra lavere presterende elever), eller elevenes kjønn. Heller ikke når vi kontrollerer for skolestørrelse får vi endringer i dette resultatet.

Endringer i undervisning og vurdering under forsøket

Gjennom evalueringen har vi fått vite at forsøket har gitt lærerne mulighet til å redusere det antallet karaktergivende prøver som elevene normalt vil få gjennom semesteret. Lærerne sier at dette gir mer tid til å drive skriveopplæring, samtidig som læreren kan gi elevene mange flere tilbakemeldinger *uten* karakter. Dette gir en indikasjon på hvilke endringer som kan skje med vurdering og undervisning når rutinene for halvårsvurdering i norskfaget endres.

Vi kan ikke vite hvor mange lærere som har lagt om undervisningen som følge av forsøket med redusert antall karakterer ved halvårsvurdering. Det vi med sikkerhet kan si, er at de endringer i norskundervisningen som eventuelt har skjedd ved skolene under forsøket, ikke har påvirket elevenes læring på *negativ* måte. Dette betyr at rutinene for halvårsvurdering i norsk kan endres uten at det svekker elevenes læring i faget, når vi ser alle elevene under ett.

Hvorfor blir det færre karaktergivende prøver under forsøket?

Selv om det nasjonale regelverket for vurdering sier at halvårskarakteren skal være et uttrykk for hvor eleven befinner seg *underveis* i læringsprosessen, kan bidrag fra forskningslitteraturen og evalueringer av vurderingsfeltet i Norge likevel tyde på at også underveisvurderinger får preg av å måle *resultatet* av elevenes læring på et gitt tidspunkt. Dette skaper et behov hos lærerne og hos den lokale utdanningsforvaltningen for å ha nok dokumentasjon i form av karaktergivende prøver underveis i semesteret for at halvårskarakteren kan sies å være gyldig.

Antallet karaktergivende prøver før halvårsvurdering er ikke regulert i det nasjonale lovverket, men bygger på lokale regler og sedvane. De lokale reglene eller sedvane tilsier at hver halvårskarakter (uansett fag) skal bygge på to eller tre karaktergivende prøver gjennom semesteret. I norskfaget betyr dette at elevene må gjennom mellom seks til ni vurderingssituasjoner med karakter gjennom semesteret, mot to eller tre vurderingssituasjoner med karakter i for eksempel matematikk.

Når antallet halvårskarakterer i norsk går ned fra tre til to eller én, har skolene derfor automatisk senket kravet til det totale omfanget av karaktergivende prøver i norsk underveis i semesteret tilsvarende. Dette er hva skoleledere og lærere forteller oss skjer under forsøket, og dette frigjør tid som lærerne ellers ville ha brukt til å forberede, rette og repetere lærestoff sammen med elevene foran mange slike karaktergivende prøver. Tiden som blir frigjort, kan de bruke til å gi elevene hyppigere tilbakemeldinger uten karakter underveis i opplæringen.

Hva kan vi lære av evalueringen?

Man kan spørre seg hvorvidt vi kunne forventet endringer i elevenes læring som følge av forsøket. Vi som forskere tror ikke dette ville være et sannsynlig utfall. Undervisning og vurdering er tett integrert i skolen, og disse faktorenes innvirkning på elevenes læring følger sannsynligvis et langt mer komplekst forløp enn at man kan «slå av» en karakter ved halvårsvurdering og se endringer kommer til syne. Evalueringen av forsøket kaster samtidig lys over to mindre utforskede felt som vedrører forskning og policyutvikling for vurdering generelt, og som aktualiseres av forsøket med endrede rutiner for halvårsvurdering i norskfaget. Vi viser kort til disse feltene nedenfor.

1. Regler og rutiner for vurdering modereres av mange faktorer

Utdanningsdirektoratet la i forsøket til grunn at norsklærere opplever at vurdering påvirker undervisning. Denne evalueringen har vist at forståelsen av en slik sammenheng mellom vurdering og undervisning må inkorporere læreplanen, kunnskapssynet og kunnskapspolitikken som ligger til grunn for læreplanen, utdanningsforvaltningens og lærerprofesjonens forståelse og fortolkning av læreplanen, dokumentasjonssystemet for elevenes læring og den faktiske implementering av læreplanen i undervisningen og vurderingen.

Det er etter vår forståelse sammenhengene mellom disse faktorene, og ikke halvårskarakteren i seg selv, som sperrer for/muliggjør det å ha færre karaktergivende prøver gjennom semesteret.

2. Vurdering uten karakter krever nytt språk for vurdering

Vi vet som sagt ikke hvor utbredt det har vært for norsklærere som har deltatt i forsøket å redusere antall karaktergivende prøver til fordel for å gi elevene tilbakemelding uten karakter. Det vi vet, er at på skoler som vi har fått besøke i forbindelse med evaluering og hvor dette har vært praksis, gir elevintervjuene indikasjoner på at vurdering uten karakter ikke nødvendigvis gir dem bedre informasjon om hva de kan gjøre for å oppnå fremgang. Dette kan bidra i en forklaring på hvorfor forsøket heller ikke synes å ha påvirket elevens læring i positiv retning.

Ved noen av skolene vi har besøkt i forbindelse med evalueringen, faller forsøket sammen med en utvikling av vurderingsarbeidet i norskfaget som dreier seg om å kutte alle vurderingssituasjoner med karakter gjennom semesteret før halvårsvurdering. Lignende utviklingstendenser ved andre skoler har vakt oppmerksomhet i media. Skolene argumenterer for at karaktergivende prøver flytter elevenes oppmerksomhet fra selve læringen og over til resultatet av læringen på et for tidlig tidspunkt, og at det å få karakterer skaper unødig stress for elevene. Disse skolene ønsker å komme bort fra denne situasjonen.

Vår evaluering av forsøket med redusert antall karakterer ved halvårsvurdering i norsk bekrefter at elever har et flertydig forhold til karakterer. På den ene siden opplever de karakterer som stress, forårsaket av et press som oppstår hos elevene selv forsterket gjennom foreldrene deres, skolen og samfunnet for øvrig. På den annen side er de innstilt på at karakterer skal gi dem informasjon om hvor gode (eller hvor dårlige) de er i et fag. Får de karakteren 3/4 på en enkeltprøve, vil de at lærerens tilbakemelding skal fortelle dem hvordan de kan komme seg opp til karakteren 4 på neste prøve.

Denne evalueringen viser at når denne skalaen mangler, kan elevene lete i lærerens tilbakemeldinger etter budskap som kan gi dem tilsvarende informasjon uten å finne dette eller å forstå tilbakemeldingene. Elevene synes ikke å være like fortrolige med et språk for vurdering som utelukker karakter som hva lærerne deres tror. Vår tolkning er at reduksjon av karaktergivende prøver under forsøket ikke har blitt erstattet av kvalitativt bedre tilbakemeldinger som forstås av eleven og hjelper dem videre i læringsarbeidet. Dette er et naturlig felt for videre forskning om vurdering generelt så vel som i norskfaget spesielt.

Evalueringsens omfang, data og metodebruk

Totalt har 184 videregående skoler og 25 ungdomsskoler deltatt i forsøket i perioden 2013-2018. I videregående skoler har forsøket vært avgrenset til elever i studieforberedende utdanningsprogram.

NIFU har utført evalueringen på oppdrag fra Utdanningsdirektoratet i perioden 2014-2018. Evalueringen består av den foreliggende sluttrapporten og to delrapporter, utgitt i henholdsvis 2015 og 2016.

I evalueringen har vi ved hjelp av en litteraturgjennomgang plassert forsøket i en faglig ramme. Vi har analysert registerdata over standpunktkarakterer og eksamenskarakterer i norsk for elever før og under forsøket. Vi har også brukt spørreundersøkelser til skoleeiere og skoleledere, samt forskningsintervjuer med skoleledere, lærere og elever.

Executive summary

This is the final report of an evaluation of a trial of having a reduced number of grades given in the subject Norwegian (mother tongue) for students in Grades 8 and 9 in lower secondary school and Grades 11(Vg1) and 12 (Vg2) in upper secondary school. The report analyzes formative assessment practices and students' learning results through a combination of qualitative and quantitative methods: A quantitative analysis of data on grading at schools before and after the intervention is supplied by data from interviews with teachers, school leaders and students.

A grading system from 1-6 (6 being the best grade) constitutes the formal assessment system in Norwegian lower and upper secondary education. This grading system is utilized for formative and summative purposes alike, but the trial was concerned with grades used for formative assessment only, and in the form of bi-annual grade reports. Local administrative rules and teachers' perception of good professional practice demand that for such a bi-annual grade report to be valid, every subject grade must be based on two or three formally graded tests.

Norwegian is the only subject in this system which is assessed in three separate grades, in the form of marks, mirroring the thematic structure of the curriculum. Problems arise when teachers must be grading between six and nine tests each term. Testing takes time away from instruction, and thus reduces the amount of teaching students receive. The trial evaluated in this report was set up to find out what might happen to instruction and student learning if formative assessment in the subject Norwegian is represented by only one or two grades instead of the normal three grades on students' bi-annual report card.

In all, 184 upper secondary schools and 25 lower secondary schools took part in the experiment between 2013 and 2018. Participation was based on the school's own interest and initiative. Only a few regulations were established for how the trial should be set up and implemented locally, among these that all students should receive instructions in all three parts of the curriculum in the subject Norwegian as normal.

Main results

We analyzed 197 000 students' overall and summative results in Norwegian based on whether the teachers used one, two or three grades for formative purposes on the bi-annual report card.

- Our analysis shows no difference in overall student achievement in Norwegian based on whether students participated in the experiment or not.
- Our analysis shows no difference in overall student achievement in Norwegian based on whether students who took part in the experiment got one or two grades on their formative assessment report card during the school year.
- This main finding does not change when we control for school size, gender or the grade point average acquired from compulsory education when entering upper secondary school, and which can serve as an indicator of low versus high performing students.

How to interpret the absence of effects in this trial

Teachers who were part of the trial have told us in interviews that when they reduce the number of graded tests during the term, they can simultaneously increase the amount of feedback they give to students without giving them a formal grade. The teachers greatly appreciate the opportunity for supportive learning activities created by the reduced number of tests and indicated that this would enable more creativity and pedagogic engagement on the part of teachers.

As our information on this is based on a small sample of teachers, this finding cannot be generalized to the entire teacher population that took part in the trial. Still, we can conclude that if changes in instruction and assessment practices like these have been widespread during the trial, they do not harm students' learning results. Local rules and traditions prescribing the number of formal tests behind each grade on the report card are therefore not supported by pedagogy.

However, the students taking part in the trial reported less enthusiasm about not receiving as many formal assessments with grades as before. These students value grades as "the truth" about their achievement. This study tells us that while teachers have found new opportunities for giving more informal and formative assessment without grades, the students can feel puzzled about what to do with feedback that does not give a clear statement on their performance.

Some students seem to believe that every result of learning, be it a written text or an oral performance, can be assessed and marked according to the formal grades between 1 and 6. When grades do not appear, the students feel that this crucial piece of information is being withheld from them.

Conclusion

As this trial has shown that teachers can free up their instruction and classroom activities from a prerequisite number of formal tests each term, we have also seen that formative assessment without formal grades must present the teachers and students with an effective and qualitative “meta-language” for learning improvement

1 Innledning

I forskrift til opplæringsloven står det at elever i ungdomsskolen og videregående skole skal få *halvårsvurdering* med karakter. I læreplanen for norskfaget er det fastsatt at standpunktvurderingen i norsk består av tre karakterer: én karakter i skriftlig hovedmål, én karakter i muntlig norsk og én karakter i sidemål. Derfor skal også halvårsvurderingene i norsk gis i form av tre karakterer. Standpunktkarakter er den vurderingen som elevenes faglærer skal sette på elevenes vitnemål når de går ut av henholdsvis 10. trinn i grunnskolen og Vg3 i studieforberedende utdanningsprogram i videregående skole. Karakterene som elevene får i halvårsvurdering, blir ikke stående på elevenes vitnemål, men skal gi dem informasjon om hvordan det går med dem i faget underveis i opplæringen.

Påvirkes undervisning, vurdering og læringsresultater i norskfaget hvis skolen får lov til å organisere elevenes halvårsvurdering slik at de bare får én eller to karakterer mens de er underveis i opplæringen? Dette har vi undersøkt i evalueringen av et slikt forsøk i norske skoler i perioden 2014-2018. Skoler som deltar i forsøket skal enten slå sammen karakteren for norsk hovedmål og sidemål, slik at elevene får én karakter i norsk skriftlig og én karakter i norsk muntlig, eller slå sammen alle disse delene av norskfaget til én felles karakter ved halvårsvurderingen. Hvordan gjør lærerne dette når læreplanen og ordningene med eksamen og standpunktkarakterer skal være som før? Påvirker det arbeidet deres, og har ordningen noen fordeler for elevene deres?

1.1 Hvorfor endre på rutinene for halvårsvurdering i norskfaget?

Forsøket med redusert antall karakterer ved halvårsvurdering i norsk sprang ut av den siste læreplanrevisjonen i norsk, som ble ferdigstilt i 2013. I høringsrunden for læreplanrevisjonen ble det foreslått å endre vurderingsordningen ved å redusere antallet karakterer. Forslaget var begrunnet i høringsdokumentenes diskusjon av hvilke emner faget skulle omfatte, og hvor stor arbeidsbyrde som den gjeldende vurderingsordningen skaper for norsklærerne. Høringssvarene viste at et slikt forslag måtte bygge på mange hensyn og interesser som veide mot hverandre. Ikke

minst ble det påpekt at endring i vurderingsordningene i faget også måtte medføre endringer i eksamensbestemmelsene, noe som har fagpolitiske, språkpolitiske og pedagogiske implikasjoner. Som eksempel kan vi nevne den lovfestede sidestillingen av norsk hovedmål og sidemål i den nasjonale språkpolitikken, hvor det å eventuelt fjerne karakteren i sidemål utfordrer denne sidestillingen.

Læreplanrevisjonen ble ferdigstilt uten at diskusjonen om antall karakterer i faget ble avsluttet. Spørsmålet om hva som eventuelt vil skje hvis antall karakterer endres, ble deretter gjort til et forsøk som kun gjelder *halvårsvurderingen*, og hvor alle landets videregående skoler med studieforberedende utdanningsprogram ble invitert til å delta. Forsøket ble deretter utvidet til ungdomsskolene. Totalt har 184 videregående skoler og 25 ungdomsskoler deltatt i forsøket. Det vanligste var å gjennomføre forsøket med to karakterer: 50,5 prosent av de deltagende skolene gjorde dette, mens 41,5 prosent av skolene hadde forsøk med en karakter. De øvrige skolene har gjennomført forsøket med en blanding av én og to karakterer på ulike årstrinn.

1.2 Rapportens hovedproblemstilling og evalueringens siktemål

Denne rapporten gir ingen evaluering av vurderingsordningene i norsk som helhet. Forsøket med redusert antall karakterer ved halvårsvurdering gjelder bare 8. og 9. trinn i grunnskolen og Vg1 og Vg2 for elever som går studieforberedende utdanningsprogram i videregående skole. På skoler som deltar i forsøket, skal elever på 10. trinn i grunnskolen og på Vg3 i studieforberedende utdanningsprogram fremdeles ha tre karakterer ved halvårsvurdering, standpunkt og eksamen. Læreplanen og eksamensordningene er akkurat den samme for elever på skoler som deltar i forsøket og skoler som ikke deltar i forsøket. På denne måten omfatter forsøket mange skoler og mange elever, men kan likevel sies å være lite. Samtidig kan forsøket gi et visst inntrykk av hva som ville skje hvis vurderingsordningene i norskfaget ble gjennomgående endret ved å redusere antallet karakterer.

Hovedproblemstillingen for denne evalueringen er formulert på følgende måte:

Fører forsøket med færre karakterer ved halvårsvurdering i norsk til at skolene kan utvikle mer hensiktsmessig organisering av undervisning og vurdering i norskfaget?

Vår evaluering av forsøket har resultert i to delrapporter i tillegg til denne slutt-rapporten. I den første delrapporten (Seland, Lødding og Prøitz 2015) tok vi for oss hva som kjennetegner undervisnings- og vurderingspraksis i norskfaget, og hva som kan anses som særlige utfordringer i faget. Vi gjorde også en vurdering av i hvilken grad forsøket kunne sies å være et svar på disse utfordringene.

I den andre delrapporten fra evalueringen (Seland, Gjerustad og Lødding 2016) undersøkte vi hvordan forsøket er organisert i skolene og hvilke erfaringer lærerne, skolelederne og elevene hadde gjort med forsøket så langt. Forskningsspørsmålene skulle i særlig grad belyse hva som skjer med undervisning, vurdering og motivasjon for faget under forsøket.

Denne sluttrapporten inkorporerer funnene fra delrapport 1 og 2 fra evalueringen i en overordnet drøfting av hvorvidt forsøket legger til rette for at skolene kan utvikle mer hensiktsmessig organisering av undervisning og vurdering i faget. Dette besvarer vi ved å undersøke forsøkets betydning for elevenes læring i faget gjennom følgende forskningsspørsmål, også disse formulert av oppdragsgiver:

1. Er det mulig å identifisere systematiske forskjeller i elevenes læring som følge av deltakelse i forsøket?
2. Er det strategier for organisering av undervisning i forsøket med færre karakterer som i særlig grad ser ut til å være hensiktsmessig?
3. Kan disse strategiene bli gjenstand for en mer spesifisert utprøving, og hvordan bør dette i så fall gjennomføres?

Hovedbidraget i denne sluttrapporten er en analyse av standpunktkarakterer i norsk hovedmål, norsk sidemål og norsk muntlig for elever som har deltatt i forsøket og som et år senere har avsluttet Vg3 på studieforberedende utdanningsprogram, sammenlignet med elever som ikke har deltatt i forsøket. Disse karakterene er analysert som en gjennomsnittskarakter i norsk for eleven, men vi har også gjort separate analyser av de tre standpunktkarakterene. I denne evalueringen er analysene av standpunktkarakterer brukt for å gi et svar på om forsøket har betydning for elevenes læringsresultater. Analysen inneholder også data om elevenes eksamenskarakter i norsk hovedmål. Eksamen i skriftlig sidemål og norsk muntlig følger en trekkordning, hvilket betyr at et stort antall elever ikke har disse karakterene. Eksamenskarakterene i sidemål og muntlig er derfor utelatt fra vår analyse. I analysen skiller vi mellom skoler som har gjennomført forsøket med én kontra to karakterer ved halvårsvurdering i norsk. Dette utgjør grunnlaget for å svare på hvorvidt det avtegner seg en strategi for ordningen som i særlig grad synes å være hensiktsmessig i en mer spesifisert utprøving. Vår drøfting av dette spørsmålet utfylles av kvalitative forskningsfunn og bidrag fra forskningslitteraturen.

I dette kapitlet bruker vi materiale fra delrapport 1 i evalueringen (Seland m.fl. 2015) for å gi en oversikt over norskfagets formål, avklaringer av begrepene "undervisning" og "vurdering", tidligere forsøk med redusert antall karakterer i norsk og vårt eget valg av perspektiv for evalueringen. Til slutt i dette kapitlet gjør vi rede for valg og bruk av metoder i evalueringen som helhet.

Evalueringen har strukket seg over en periode på fire år, og delrapporter og sluttrapport dekker til sammen et bredt spekter av forskningsspørsmål. En rekke av disse spørsmålene er utviklet av forskergruppen i prosjektet, samtidig som alle

forskningsspørsmål tar sitt utgangspunkt i oppdragsgivers kunnskapsinteresse i evalueringen. Istedenfor å gjengi hele listen av forskningsspørsmål i dette kapitlet, har vi laget et vedlegg til rapporten som viser samtlige spørsmål som har ligget til grunn for arbeidet med evalueringen. For hvert kapittel i denne rapporten som presenterer empiri, presenterer vi de aktuelle forskningsspørsmålene i kapitlets innledning.

1.3 Norskfagets formål og oppgaver i opplæringen

Norskfagets formål slås i læreplanen fast med følgende formulering: «Formålet er å styrke elevenes språklige trygghet og identitet, utvikle deres språkforståelse og gi et godt grunnlag for mestring av begge målformene i samfunns- og yrkesliv.» Formålsbeskrivelsen i planen bygger på en innledende tekst som viser til at:

[n]orsk er et sentralt fag for kulturforståelse, kommunikasjon, dannelse og identitetsutvikling. Gjennom aktiv bruk av det norske språket innlemmes barn og unge i kultur- og samfunnsliv og rustes til deltakelse i arbeidsliv og demokratiske prosesser.

Kunnskapsdepartementet (2013: 1)

Litteraturgjennomgangen i den første delrapporten fra denne evalueringen (Seland m.fl. 2015) ga forskjellige eksempler på hvordan nøkkelbegrepene danning/dannelse, identitetsbygging og demokratisk deltakelse behandles i beskrivelser av norskfagets formål. Når bidragsyttere skal definere disse begrepene i litteraturen, ser vi at de gis ulikt innhold og mening. Berge (2005) omtaler for eksempel dannelse i forbindelse med skriftkyndighet eller *literacy*, noe som gir dannelsesbegrepet et innhold knyttet til fagets nytteverdi. Andre ser dannelsesbegrepet som suverent overordnet en tanke om nytte, i form av å utvikle "demokratiske samfunnsmedlemmer som kan bidra som aktive deltakere i kulturen" (Aase 2005: 17). Ifølge andre litteraturbidrag om norskfaget utgjør språkhistorie, nynorsk og det å ha erfaring som nynorsktalende et særlig grunnlag for danning (Måseide 2013; Nicolaysen 2005). Slik kan *danning* både ses som et vidt begrep og et prinsipp med stor legitimitet innen norskfaget.

Identitet og personlighetsutvikling knyttes også til dannelsesbegrepet (Måseide 2013), enten man vektlegger tilhørighet, det unike ved individet eller erfaring med forskjellighet som vesentlig for identitetsutviklingen (Nicolaysen 2005). Haaland (2009) stiller opp kontrasten mellom "innfødte" og "utstøtte" i elevenes identitetsutvikling, mens Gaare (2006) omtaler det "kosmopolitiske" som kontrast til nasjonsbygging i spørsmålet om hvilken selvforståelse norskfaget skal fremme.

Utvikling av kompetanse for deltakelse i et demokratisk samfunn er også gjennomgående i beskrivelser av hensikt med faget. Aase (2005) knytter forbindelse

mellom dannelse og demokratisk deltakelse. Et demokratiseringsarbeid er ifølge Nicolaysen (2005) uløselig knyttet til fremveksten av de to målformene nynorsk og bokmål. At skolen tar ansvar for formidling av en skjønnlitterær kanon eller kulturarv, kan ses som et spørsmål om sosial og kulturell utjevning og dermed forenlig med demokratisering. Oppøvelse av skriftkyndighet eller *literacy* har demokrati-kompetanse som overordnet målsetting. Disse to posisjonene (vektlegging av kulturarv versus skriftkyndighet) representerer to hovedargumenter eller diskurser om hva norskfaget må ivareta, og stilles i litteraturen og debatten ofte opp mot hverandre. Samtidig ser vi at disse to posisjonene benytter samme overordnede referanse som begrunnelse og legitimering: dyktiggjøring for deltakelse i et demokratisk samfunn.

Vi kan oppsummere dette som at vi i norskfaget har å gjøre med komplekse og mangetydige begreper om fagets hensikt og oppgave i skolen. Begrepene har videre hatt forskjellig meningsinnhold til forskjellige tider og for ulike formål. I mange av litteraturbidragene som vi har gjennomgått (se Seland m.fl. 2015), fremheves betydningen av bestemte elementer i faget, det være seg nynorsk, språkhistorie, kulturarv eller demokratikompetanse. Ingen av de refererte bidragene om fagets formål og funksjon i opplæringen kan sies å være enerådende eller tegne det fullstendige bildet av faget. Dermed representerer disse bidragene et mangfold av stemmer om hvordan faget skal forstås og hva faget skal gjøre for elevene og samfunnet. Med så ulike oppfatninger av hva norskfagets kjerneoppgaver er, hadde vi også forventet at litteraturen ville peke på utfordringer ved vurdering av elevers prestasjoner i faget. Dette er et område som overraskende få av litteraturbidragene ser ut til å være opptatt av.

Læreplanen i norsk under Kunnskapsløftet vektlegger skriftkyndighet (*literacy*) som del av et tverrfaglig arbeid med grunnleggende ferdigheter. Dette representerer et brudd med tidligere læreplaner, slik også en dreining bort fra det tradisjonelle nasjonsbyggingsprosjektet i faget kan ses som et slikt brudd. Av de politiske grunnlagsdokumentene for Kunnskapsløftet og litteraturen for øvrig går det frem at arbeid med grunnleggende ferdigheter er et område hvor samarbeid mellom lærere i flere fag er en forutsetning. Dette reiser spørsmålet om grensdragning mellom hvilke emner og læringsmål som skal legges til norskfaget og hva som også kan sies å høre naturlig hjemme i skolens andre fag. Her er avgrensningen mot samfunnsfagene kanskje mest aktuell, men ifølge Kunnskapsløftet skal elevene også øves i grunnleggende ferdigheter i lesing, skriving og muntlig i skolens øvrige fag.

Gjennom litteraturbidrag som er referert delrapport 1 fra evalueringen (Seland m.fl. 2015), fremstår norskfaget som sammensatt av en rekke elementer som skal besvare mange og kanskje dels også konkurrerende formål. Norskfaget omtales som et kulturfag, et dannelsesfag og et ferdighetsfag, og kan på denne måten be-

traktes som å skulle tjene mange ulike samfunnsinteresser. Faget kan således karakteriseres som et mer kontekstuelet forankret skolefag enn som et konseptuelet forankret (Muller 2009). Dette teoretiske skillet refererer til hva som danner helhet og sammenheng i et skolefag sett i lys av disipliner i høyere utdanning. Ideen er at i fag som norsk er det den samfunnsmessige politiske, kulturelle, sosiale og økonomiske konteksten som skaper sammenheng i faget slik dette fremtrer i læreplaner, mens det i fag som matematikk i større grad er bestemte begreper og logikker innenfor den akademiske disiplinen matematikk som reflekteres i skolefaget matematikk.

En direkte konsekvens av dette er for eksempel ulikhet i tilnærminger til vurdering (se også senere i dette kapitlet). Skolefag med en mer entydig disiplinforankring (eksempelvis matematikkfaget) kjennetegnes ofte av en tydeligere vertikal struktur, forstått som et sekvensielt organisert lærestoff og et strengt hierarkisk oppbygd begrepsapparat hvor det ene kunnskapselementet legger grunnlaget for det neste i opplæringen. Mer kontekstuelet oppbygde fag (eksempelvis norskfaget) kjennetegnes ved at kunnskapselementer riktignok står i en innbyrdes sammenheng, men denne sammenhengen er ikke utpreget vertikal og kan bestå av en rekke sideordnede emner. Det fremtrer ingen åpenbar sekvensiell organisering av lærestoffet som avgjør når i opplæringsløpet elevene skal tilegne seg en bestemt kompetanse i faget. Som konsekvens for vurdering innebærer dette at mens konseptuelet ordnede fag har retningsgivere og avgrensninger som tydelig angir hva som skal og bør bli underlagt vurdering, blir dette mer utydelig og er i mindre grad gitt i mer kontekstuelle fag, som norskfaget (Prøitz 2013).

1.4 Begrepsavklaring: undervisning

Undervisning som begrep viser til kompleks kommunikasjon og interaksjon i situasjoner hvor det foregår aktivitet med formål om læring. Undervisning kan på denne måten forstås som en sosial praksis, som innebærer handlinger, forståelser og tolkninger som utspiller seg mellom lærere og elever og mellom elever når de arbeider for å skape mening i undervisningssituasjonen (Bransford et al. 2000; Nordenbo et al. 2008). Det finnes en rekke tilnærminger til studier av undervisning og undervisningspraksis. I delrapport 1 i denne evalueringen så vi nærmere på to norske studier som tar for seg undervisningspraksis forstått som hva lærerne uttrykker, behersker, formidler og prioriterer i vanlige, men like fullt komplekse undervisningssituasjoner i klasserommet (Mjåtveit og Jakobsen 2013; Mausestagen og Kostøl 2009).

I denne evalueringens delrapporter har vi undersøkt henholdsvis litteratur om undervisningspraksis i norskfaget og utsagn fra skoleledere, lærere og elever om

slik undervisningspraksis. Hensikten er å forstå hvordan faget og vurderingsordningene kan sies å legge føringer på lærernes undervisningspraksis. I analysen av lærernes undervisningspraksis vektlegger vi aktiviteter knyttet til faglig tilrettelegging, emosjonell støtte, åpenhet for elevenes valgmuligheter og innspill til undervisning og stimulering av elevenes mestring og prestasjon (Mjåtveit og Jakobsen 2013). Kommunikasjonen som underbygger læreres undervisningspraksis, vil påvirke elevers motivasjon og opplevelse av mestring i faget (Mausethagen og Kostøl 2009).

1.5 Begrepsavklaring: vurdering

Formålet med vurdering i fag er «å fremje læring undervegs og uttrykke kompetansen til eleven, lærlingen og lære kandidat undervegs og ved avslutninga av opplæringa i faget. Vurderinga skal gi god tilbakemelding og rettleiing til elevane, lærlingane og lære kandidatane» (Forskrift til Opplæringsloven, kapittel 3).¹ Vurdering kan gis med og uten karakterer. Halvårsvurdering, som vi undersøker i denne evalueringen, skal gis med karakterer. Halvårsvurderingen skal samtidig følges av skriftlig eller muntlig tilbakemelding fra læreren hvor elevens læring i faget vurderes mot læringsmålene i læreplanen.

Læreren kan også gi eleven slike tilbakemeldinger eller kommentarer gjennom skoleåret som en naturlig del av kontakten mellom lærer og elev i og utenfor undervisningen. I norsk skole er det praksis at lærerne gir elevene prøver og hjemmeoppgaver som skal leveres inn og vurderes av lærer gjennom skoleåret. Elevene kan også ha forberedte presentasjoner av fagstoff for hverandre og for læreren. Alle disse eksemplene representerer vurderingssituasjoner, og i ungdomsskolen og videregående skole vil læreren ofte sette karakter på elevenes arbeid. Samtidig viser denne evalueringen at lærere også vurderer elevene *gjennom aktiviteten* som skjer i undervisningen og i klasserommet fra dag til dag. Her blir vurderingen et uttrykk for hvordan og i hvilken grad budskapet i kommunikasjonen mellom lærer og elev er mottatt og forstått. Hvis eleven ved svar på lærerens spørsmål eller i klasseromsamtalen viser at hun har forstått noe som læreren ønsker at elevene skal lære, registrerer læreren dette for seg selv, ofte bare ved å huske det. Læreren kan imidlertid også gjøre notater om slik vurdering. Lærers umiddelbare respons i undervisningssituasjonen kan bekrefte for eleven at hun har lært, forstått og gjør fremgang. Våre intervjuer med lærere i denne evalueringen viser at lærerne verdsetter denne typen vurdering i undervisningssituasjonen fordi de mener det gir bedre læring for elevene enn de forberedte prøvene og elevoppgavene som skal rettes. I denne rapporten kaller vi dette *uformelle vurderingssituasjoner*.

¹https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_4

1.5.1 Formativ og summativ vurdering - underveisvurdering og sluttvurdering

I det vurderingsfaglige språket benyttes flere sentrale begreper som har til hensikt å karakterisere ulike tilnærminger og formål for vurdering. Særlig sentralt i den internasjonale faglitteraturen står begrepene *formativ* og *summativ* vurdering (Black & Wiliam 2003; 1998; Harlen & James 1997; Sadler 1989; Taras 2005). Litteraturen bærer imidlertid preg av diskusjon om ulikheter i forståelse av hva som skiller formativ og summativ vurdering. Debatten reflekterer i stor grad også hvordan disse kategoriene ikke er gjensidig utelukkende. Diskusjonen omhandler spesielt viktige spørsmål omkring bruk av vurdering for ulike formål. Sentralt i dette står behovet for klarhet om bruk. For eksempel er det viktig at den som blir vurdert, vet og forstår hva vurderingen skal brukes til. Formålsdiskusjonen handler også om bevissthet omkring hvorfor vurderingen gjøres blant dem som skal gjennomføre og bruke resultater av vurdering. I dette er det også et element som omhandler respekt for den som skal vurderes og elevers rett til å vite når man blir vurdert for hvilke formål.

Nasjonalt kompetansemiljø i vurdering (2012: 8-9) slår fast at det er lang tradisjon for summativ vurdering i Norge, forstått som måling av kunnskap, ferdigheter og kompetanse ved hjelp av tester. Samtidig må det bemerkes at norsk skole har hatt få tradisjoner for standardisert testing av elevers læringsutbytte (Hertzberg 2008; Skedsmo 2011). Formativ vurdering fokuserer derimot den prosess i tankegang, læring og selvregulering som ligger forut for og dermed skaper læringsresultatene. Her er vurderingen rettet mot målene for læringen, og den skal gjennom å synliggjøre elevens tankegang og læringsstrategier avdekke behovene for faglig veiledning og tilpasning. Læreren skal ideelt sett også forbedre sin undervisningspraksis med utgangspunkt i informasjonen om elevenes arbeid.

I norsk forskning er det pekt på at lærere strever med å skille mellom formativ og summativ vurdering. Sandvik (2014: 45) skriver at til tross for at det synes enkelt å skille mellom formativ og summativ vurdering, viser det seg likevel at for lærere kan dette i praksis være vanskelig. Dette er ikke unikt for norske lærere, og internasjonal forskning har vist gjentatte ganger at skillet kan oppfattes som kunstig og være vanskelig å etterleve i praksis. Stobart (2008: 146) peker på at formativ vurdering tradisjonelt har blitt anvendt som «mini-summativ», der læreren har samlet dokumentasjon ment for karaktersetting ved semesterslutt (se også Nusche, Earl, Maxwell & Shewbridge 2011: 56).

Merknadene til regelverket om individuell vurdering i forskrift til opplæringsloven kapittel 3² understreker alle elevers rett til både underveisvurdering og sluttvurdering. Underveisvurdering omtales i forskriftens § 3-11 som et redskap i

² <https://www.udir.no/regelverk-og-tilsyn/finn-regelverk/etter-tema/Vurdering/individuell-vurdering-udir-5-2016/>

læreprosessen, som grunnlag for tilpasset opplæring og et bidrag til elevens økte kompetanse i faget. Elevens kompetanse på vurderingstidspunktet skal vurderes mot kompetansemålene i læreplanen. Slik underveisvurdering kan være både muntlig og skriftlig gjennom opplæringsløpet. Elevers egenvurdering og læreres halvårsvurdering av elevene inngår i underveisvurderingen. Sluttvurdering er i forskriftens § 3-17 slått fast å være informasjon om kompetansen som eleven har ved avslutningen av opplæringen i faget. Sluttvurdering gis i form av standpunkt-karakterer og eksamenskarakterer.

Arbeidet med å skape sammenheng mellom vurderingsformer kan beskrives som en trinnvis prosess der læreren samler dokumentasjon eller bevis på elevens læring, fortolker denne dokumentasjonen og deretter utvikler undervisning som kan bidra til ytterligere å fremme elevens læring (Sandvik 2014: 49, NOU 2014: 7, s. 101).

1.5.2 Nærmere om halvårsvurdering

Sentralt i arbeid med å skape sammenheng mellom underveis- og sluttvurdering står myndighetenes krav og elevenes rett til halvårsvurdering. Halvårsvurdering defineres av Utdanningsdirektoratet i Merknader til regelverket om individuell vurdering (se også tidligere i dette kapittelet) på denne måten: «*Halvårsvurderingen er en del av underveisvurderingen. Den skal altså gi informasjon om hvordan eleven mestrer etter kompetansemålene i læreplanen og gi veiledning om hvordan eleven kan øke kompetansen*» (merknader til § 3-13, s. 11). Halvårsvurdering skal gjennomføres på alle trinn i grunnopplæringen. Fra og med 8. årstrinn skal halvårsvurdering gis med karakter. Halvårsvurdering med karakter skal gis skriftlig. Det er skoleleder som har ansvar for at faglærerne gjennomfører halvårsvurdering. Evalueringen hvor denne delrapporten inngår, er i sin helhet innrettet mot forsøk med redusert antall karakterer ved halvårsvurdering i norsk.

1.5.3 Vurderingskompetanse

Flere steder i litteraturen vises det til læreres og elevers vurderingskompetanse. Utdanningsdirektoratet (2011: 91) definerer vurderingskompetanse på følgende måte:

Vurderingskompetanse forstås som bruk av metoder, vurderingsformer og vurderingsresultater, utvikling av et fagspråk om vurdering og som sammenheng mellom mål og vurdering. Å utvikle vurderingskompetanse innebærer også å etablere en vurderingskultur med fokus på elevenes og lærlingenes læring, med rom for å prøve og feile og for å gi hverandre konstruktive tilbakemeldinger.

Nasjonalt kompetansemiljø i vurdering (2012: 8) skriver om vurderingskompetanse at den skal være anti-instrumentell, noe som innebærer at slik kompetanse ikke er knyttet til definert eller avgrenset bruk av særlige metoder for vurdering. Videre skal slik kompetanse være en egenskap ved den enkelte lærer, samtidig som den er forankret i lærerkollegiet, mens skoleeier og skoleleder har ansvar for at slik kompetanse finnes og vedlikeholdes. Vurderingskompetanse har både individuelle og kollektive kjennetegn, som samlet sett kan beskrives som skolens vurderingskultur. Det vises også i styringsdokumentene til at elever skal få oppøve sin vurderingskompetanse, noe de bare kan gjøre i refleksjon sammen med andre elever eller læreren.

1.5.4 Forskning om vurdering i ulike fag

Norsk forskning har vist eksempler på at lærere i ulike fag har forskjellig praksis for vurdering og karaktersetting (Fjørtoft 2013; Prøitz og Borgen 2010; Prøitz 2013; Hovdhaugen, Seland, Lødding, Prøitz og Vibe 2014). Studiene viser også at disse variasjonene i stor grad synes å ha sitt utgangspunkt i bestemte kjennetegn ved fagene og fagenes egenart. Dette er funn som i stor grad harmonerer med internasjonal forskning (Resh 2009; Duncan & Noonan 2007; Lekholm & Cliffordson 2009; McMillan 2003). Funnene som viser at læreres vurderingspraksiser og karaktersetting er forskjellig avhengig av fag, dreier seg dels om at lærere sosialiseres og utvikler ulike typer av profesjonalitet avhengig av fagtilknytning og fagets disiplinære tilknytning. Dels handler dette om variasjoner i fagstruktur og fagstoff som påvirker beslutninger om innhold, undervisning og vurdering (Resh 2009). Fagenes rolle i vurdering og karaktersetting er studert og diskutert i en rekke studier (Eggen 2004; Lekholm & Cliffordson 2009; 2008; Duncan & Noonan 2007; Resh 2009).

Det er hevdet at en forklaring på ulikheter i praksis er å finne i ulike epistemologiske og ideologiske posisjoner lærere inntar i arbeidet med vurdering, og at dette påvirker lærernes syn på og holdninger til vurdering (Eggen 2004: 480). For eksempel er det ut fra sammenligninger av læreres praksiser i engelsk, naturfag og matematikk argumentert for at matematikk- og naturfaglærere anser sine fag å ha objektivt definerte mål, mens engelsklærere forholder seg til flere mål tilpasset den enkelte elev og undervisningssituasjon (Black, Harrison, Lee, Marshall et al. 2003: 68). Forskjeller mellom lærernes vurderingspraksiser i engelsk og matematikk basert på lærernes egne beskrivelser av egen praksis er også delt inn i ulike kategorier i forskningen. Praksiser i matematikk beskrives som basert på mer rasjonelle, analytiske, a-historiske, verdifrie og stabile standarder, mens tilsvarende i engelsk er beskrevet som basert på mer holistisk intuitive, ikke-numeriske og dialog- og

observasjonsbaserte standarder (Wyatt-Smith & Klenowski 2013). Lignende betraktninger finnes også i en studie av norske læreres praksiser for standpunkt-karaktersetting i fem fag der norsklærere oppgir at standarder i faget er mer åpent for diskusjon og forhandling enn for eksempel matematikklærerne oppgir (Prøitz og Borgen 2010).

Når det gjelder norskfaget mer spesielt er det vist til at prosedyrer og dokumenter knyttet til standpunkt-vurdering, eksamenssensur og karaktersetting synes å ha mye å si, i tillegg til lærernes syn på vurdering (Sandvik, Engvik, Fjørtoft, Langseth m.fl. 2012; Prøitz og Borgen 2010). Sistnevnte studie av norsklæreres standpunkt-karaktersetting viste at norsklærere først og fremst baserte sin vurderingspraksis på skriftlige innleverte arbeider som særømne, heldags- og halvdagsprøver og såkalt «tentamen» basert på eksempeloppgaver fra eksamen, mens muntlig fremføring ble brukt som supplement og grunnlag for vurdering til muntlig karakter. Sandvik m.fl. (2012) viser på sin side til bruk av mappevurdering i norsk som et eksempel på styrking av formativ vurdering, samtidig som elevpresentasjoner i kombinasjon med elevvurdering synes å være en viktig del både av det formative og det summative vurderingsarbeidet. Samtidig understrekes det i begge studier at norsklærerne generelt oppgir at det å ha nok tid til vurderingsarbeidet er en utfordring. Det er også vist til at det eksisterer ulike forståelser og praksiser i individuell vurdering og at det er en vei å gå for utvikling av et praksisfellesskap når det gjelder vurdering i norskfaget (Sandvik m.fl. 2012). Lærere i norskfaget viser til at dobbeltheten i faget som ferdighetsfag og dannelsesfag oppleves som en utfordring for organisering av undervisning og vurdering (Sandvik m.fl. 2012).

1.6 Tidligere forsøk med redusert antall karakterer i norsk

Skoleåret 2009-2010 startet Fana gymnas i Bergen et forsøk med én karakter i norsk på Vg1 og Vg2 (studieforberedende program). Påfølgende skoleår startet Kongsbakken videregående skole i Tromsø et tilsvarende forsøk. Erfaringene fra elever og lærere ved disse skolene er samlet og gjengitt av Rønning (2012) i rapporten *Fra karakterjag til læring – fra dommer til trener*. Rapporten bygger på kvalitative intervjuer supplert av en survey til elevene og elevresultater fra ett kull ved skolene.

De to skolene skiller seg fra hverandre i beskrivelsen av samarbeidsrutinene for norsklærerne forut for og under forsøket. På Kongsbakken videregående skole hadde man et systematisk og forpliktende samarbeid mellom lærerne som gjennomførte forsøket. I åtte år forut for forsøket hadde lærerne på skolen hatt forpliktende, timeplanlagt møtetid til bruk i såkalte faglag i 90 minutter annenhver uke. Tema for møtene i faglaget var felles planarbeid, vurdering av kollegers arbeid og utvikling av felles undervisningsopplegg. Kongsbakken videregående skole hadde

også deltatt i det nasjonale prosjektet "Bedre vurderingspraksis" og var involvert i satsingen "Vurdering for læring" da forsøket med én karakter i norsk startet opp (se beskrivelse av disse nasjonale satsingene senere i dette kapittelet). På Fana gymnas var det å delta i lærersamarbeid en avgjørelse som var opp til den enkelte lærer, og det var ikke utviklet felles lokale planer for norskfaget ved skolen. Antall undervisningstimer i norskfaget varierte fra uke til uke ved begge skoler. Ved Fana gymnas tok man ut en såkalt «blokkdag» av oppsparte timer og satte disse av til en hel dag med norskundervisning. Ved Kongsbakken videregående skole ga ordningen en «fagdag» hver femte uke.

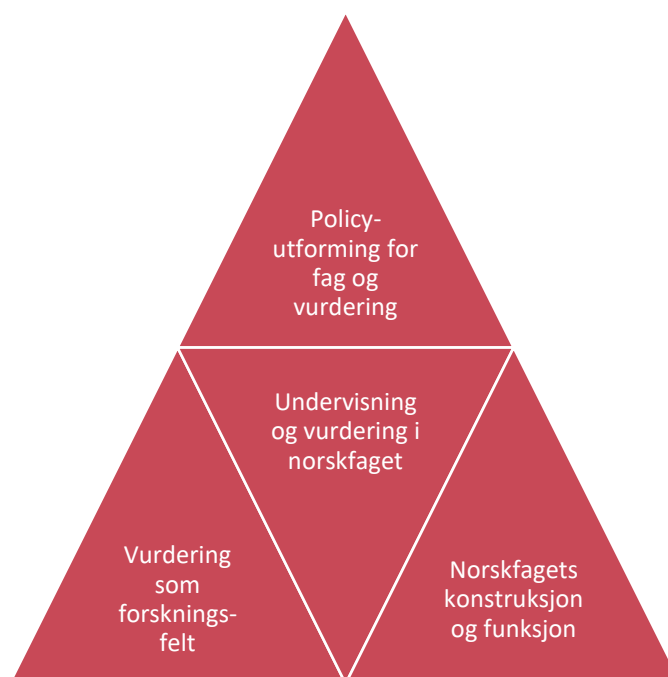
I intervju hevdet lærerne ved begge skoler at forsøket hadde bidratt til endret undervisning i norsk. Når lærerne kunne konsentrere seg om å skaffe vurderingsgrunnlag for én istedenfor tre karakterer, reduserte dette ifølge de samme lærerne krav til mengde elevarbeider og -innleveringer. Slik kunne elevene arbeide mer med hver enkelt tekst enn hva som var tilfelle før forsøket startet. Lærer og elev ble dermed mer involvert i vurderingsarbeidet av hver enkelt tekst, slik at lærerne mente at undervisningen fikk sterkere preg av læringsprosess enn hva som var tilfelle da eleven skulle ha tre karakterer i faget. Lærerne mente videre at dette ga elevene mer dybdelæring i faget. Noen av de intervjuede lærerne spurte seg imidlertid om muntlig aktivitet og lesing kunne ha fått noe mindre oppmerksomhet i forsøksperioden sammenlignet med elevenes skriftlige arbeid. Enkelte lærere førte også en form for «skyggeregnskap» med tre karakterer i norsk i forsøksperioden, for å kunne gi elever som ønsket det, en tilbakemelding om hvordan de lå an i henholdsvis skriftlig hovedmål, skriftlig sidemål og muntlig norsk. Dette gjorde disse lærerne for å sikre seg mot at eleven nedprioriterte et av disse områdene, slik at de risikerte en negativ overraskelse ved overgang til det ordinære vurderingssystemet med tre karakterer på Vg3.

Elevene som hadde deltatt i forsøket, fortalte at de oppfattet norsk som et omfattende fag med mange ulike fagområder. De ga i hovedsak støtte til at de hadde fått gode underveisvurderinger under forsøket, men uttrykte noe usikkerhet med hensyn til hvorvidt muntlig norsk hadde fått den nødvendige oppmerksomheten i undervisningen. Elevene var overveiende positive til forsøket og mente det hadde hatt en gunstig innvirkning på deres motivasjon og læring i norskfaget. Ved overgang til Vg3 sa nær halvparten av elevene at de syntes skriftlig sidemål var vanskelig. En relativt stor andel svarte også at de syntes det var vanskelig å vite hva som legges til grunn for de tre karakterene i det ordinære vurderingssystemet i norskfaget. Vel halvparten av elevene trodde ikke de hadde faglige ulemper på Vg3 som følge av å være omfattet av forsøket med én karakter på Vg1 og Vg2. Rapportens omtale av elevenes resultater etter forsøket viser kun til ett kull på ett måletidspunkt. Dette gir ingen holdepunkter for å vurdere om det har skjedd endringer i karakternivå som følge av forsøket (Rønning 2012).

Bøyum og Norendal (2011) skrev en artikkel i tidsskriftet *Norsklæreren* om forsøket med én karakter i norsk ved Fana gymnas. I artikkelen viste de to til begrunnelser for forsøket og deres inntrykk av lærernes og elevenes erfaringer. Bøyum og Norendal konkluderer at en reduksjon av karakterer i norskfaget vil føre til at læreren står friere til å styre aktiviteten basert på klassens behov, noe de mener kan gi norskfaget et kvalitetsmessig løft.

1.7 Rapportens særlige bidrag til forskningsfeltet

Forskergruppens formål med denne evalueringen har vært å samle, presentere og analysere ulike aktørers perspektiv på undervisning og vurdering i norskfaget i lys av forsøket om redusert antall karakterer ved halvårsvurdering. Dette gjør vi på bakgrunn av en forutgående gjennomgang av litteraturbidrag om norskfagets konstruksjon og funksjon i opplæringen samt vurderingsfaglig forskning, og vi ser våre data i sammenheng med de policydokumenter og -satsninger som former norskfaget og vurdering i norsk skole. Dette kan illustreres som figur 1.1.



Figur 1.1: Evalueringens oppbygning illustrert ved emnenes innbyrdes forbindelser og grenseflater

Innholdet i hver av de ulike trekantene i modellen i figur 1.1 er beskrevet utfyllende i delrapport 1 fra evalueringen (Seland m.fl. 2015) og er dessuten gjengitt i sammendrag i kapittel 1, 2 og 3 i denne sluttrapporten.

I oppsummeringen og vurderingen av høringsuttalelser om reduksjon av antall karakterer ved halvårsvurdering i norskfaget (Utdanningsdirektoratet 2013a), skriver direktoratet at en konsekvens av dette forslaget vil være at norsklærere får en bedre arbeidssituasjon, at det blir mer rom for å prioritere opplæring i grunnleggende ferdigheter, større mulighet til konsentrasjon om emner samt mindre fokus på skriftlige innleveringer og dokumentasjon for å sikre et godt nok grunnlag for standpunktvurdering. Det er dette som empirisk belyses i denne evalueringen gjennom data fra skoler som deltar i forsøket. Vi ser imidlertid at forståelser og diskusjoner av læreres arbeidssituasjon og prioritering av undervisningsemner kan kreve et bredere bakteppe enn det utsnittet som kunnskap om norskfaget i seg selv kan gi. Vi vil anskueliggjøre dette med et kort blikk på noen enkeltbidrag om læreres arbeidssituasjon generelt, norsklæreres arbeidssituasjon spesielt og betraktninger med opphav i forskningslitteraturen om hvilke forhold som kan sies å styre vurdering i fag.

1.7.1 Læreres arbeid med vurdering i kontekst av Kunnskapsløftet

Utvikling av vurderingsordninger i norskfaget og i grunnopplæringen for øvrig må forstås i lys av utdanningsreformen Kunnskapsløftet, som ble iverksatt i 2006. Aasen, Møller, Rye, Ottesen m.fl. (2012: 59, 67) beskriver Kunnskapsløftet som en reform innrettet mot strukturen, innholdet og styringen av grunnopplæringen. Nye læreplaner styrt av overordnede nasjonale mål er under Kunnskapsløftet utledet i kompetansemål for fag og årstrinn, som kommer til uttrykk gjennom sterkere vektlegging av læringsresultater og elevenes grunnleggende ferdigheter. Kvaliteten i opplæringen skal gjennom Kunnskapsløftet styrkes ved at nasjonale myndigheter stiller tydeligere krav til skolene, mens lærerne selv skal utforme lokale læreplaner og vurdere elevens resultater i tråd med de fastsatte målene for opplæringen. I dette arbeidet skal lærerprofesjonen styrkes og understøttes av aktive og tilretteleggende skoleledere og skoleeiere på kommunenivå, som til gjengjeld skal stå til ansvar for skolens resultater.

Kompetansemålene i læreplanene etter Kunnskapsløftet er bygget opp med beskrivelser av hva elevene skal mestre på ulike årstrinn fremfor hvilke elementer undervisningen skal inneholde. Aasen m.fl. (2012: 253) skriver at læreplanene under Kunnskapsløftet gir få anvisninger om individuell vurdering, samtidig som innføringen av kompetansemål innebar at læreres gjeldende vurderingspraksis måtte endres. Når lærerne skulle arbeide med å konkretisere innholdet, aktivitetene og metodene i undervisningen med utgangspunkt i kompetansemålene, måtte dette også få innvirkning på grunnlaget for vurdering av elevenes læring.

NOU 2014: 7 (s. 98) omtaler evalueringene av Kunnskapsløftet, som viste utfordringer i det lokale læreplanarbeidet som følge av at beskrivelsene av måloppnåelse i læreplanene varierte med hensyn til tydelighet og presisjon. I flere av fagene var det et stort antall mål, og målene fremsto også i mange tilfeller som svært omfattende. Det gikk også frem at læreplanene ikke skilte mellom mål som var enkle å oppfylle og mål som krever avansert bruk av kunnskaper og ferdigheter. Rapportene fra evaluering av Kunnskapsløftet viser samtidig at mange lærere vier økende oppmerksomhet til vurdering som redskap for å støtte opp om elevenes læring etter hvert som reformen har fått virke noen år i skolen (Aasen m.fl. 2012; Hodgson, Rønning, Skogvold og Tomlinson 2010). Nasjonale myndigheter har siden innføringen av Kunnskapsløftet arbeidet parallelt med å utvikle og tydeliggjøre regelverket for vurdering, styrke lærernes vurderingskompetanse og utvikle vurderingskultur- og praksis i skolen (Aasen m.fl. 2012: 253). Forsøket med redusert antall karakterer i norsk kan ses i forlengelse av dette arbeidet.

I utlysningen av dette evalueringsoppdraget skriver Utdanningsdirektoratet at norsklærere opplever at vurdering påvirker undervisning i faget. Dette kom tydelig frem i høringssvar for læreplanrevisjonen, som ble avsluttet i 2013. Bakgrunnen for evalueringsoppdraget er dermed å undersøke disse sammenhengene nærmere. Forsøket ble igangsatt for å skaffe mer informasjon om et fag er tjent med å ha tre karakterer når de fleste fag har én karakter. Et operasjonalisert forskningsspørsmål er derfor om vurdering av samme kompetansemål med tre ulike karakteruttrykk, slik det er i norskfaget, styrer undervisningen mer enn det ville ha gjort hvis faget hadde ett eller to karakteruttrykk. Dette kan videre forstås som at implementering av læreplanene delvis avhenger av hvordan vurderingssystemene er integrert i de samme planene (NOU 2014: 7, s. 109). Slik implementering, det vil si omsetting av planene til undervisning og vurdering i norskfaget, vil imidlertid også påvirkes av en rekke andre forhold, som læreplanens form og innhold, og hvordan planen blir forstått, fortolket og kommunisert av aktørene som er ansvarlig for opplæringen (ibid.).

1.7.2 Nye krav skaper endret arbeidssituasjon for lærere

Ballet & Kelchtermans (2009) viser til den såkalte intensifiseringshypotesen lansert av M. V. Apple (1986) om hvordan økende krav til opplæringens effektivitet har satt lærerens yteevne under press. Apple knytter dette til samfunnsendringer fremmet gjennom krav som politikerne og det statlige utdanningsbyråkratiet stiller til skolene, som han mener springer ut av logikk hentet fra økonomisk teori. Ballet & Kelchtermans (2009: 1150-1151) utdyper Apples opprinnelige hypotese gjennom intervjuer med 26 flamske (belgiske) lærere ansatt ved fire skoler om de-

res opplevelse av intensifisering av læreryrket. De legger til grunn at lærerprofesjonen er i endring, der lærere opplever større grad av krav om dokumentasjon og ansvarliggjøring i læringsprosesser så vel som resultater. Kravene kommer til uttrykk gjennom utdanningspolitiske reformer og tiltak i skolen, gjennom læreplanutvikling og i møte med foreldre. Lærerne kan reagere på slike krav og påfølgende økte arbeidsbyrde med å se seg nødt til å arbeide mindre kreativt med undervisningen, ved å redusere kollegasamarbeidet og renonsere på egen fritid. Apple (1986) hevdet at dette i sum førte til en de-profesjonalisering av lærerstanden.

På bakgrunn av intervjuene viser Ballet & Kelchtermans (2009: 1152-1153) til at læreres lojalitet er bundet til en rekke aktører eller interessenter i opplæringen, som rektorer, kolleger, foreldre og fremfor alt elevene selv. Her kan det oppstå interessekonflikter og krysspress, hvor lærerne har en tendens til å prioritere de krav og pålegg som de mener gagnar eleven. Samtidig hadde flere av de intervjuede lærerne problemer med å identifisere opphavet til kravene som de selv fortalte om. De generaliserte dette ved å si at kravene kom «ovenfra». Disse forskernes hovedfunn (s. 1153) er imidlertid at trekk ved samhandlings- og samarbeidskulturen ved selve skolen modifierer og former lærernes opplevelse av krav og arbeidsbyrde. Ballet & Kelchtermans (2009) beskriver slik skolens samhandlingskultur som et sett av eksplisitte og innforståtte regler og normer som styrer interaksjonen i lærerkollegiet i arbeidshverdagen. Det er dette som ifølge disse forskerne vil ha sterkest innvirkning på undervisningen som lærerne utfører og deres opplevelse av arbeidspress, utdypet gjennom lærernes opplevelse av egen profesjonell identitet (ibid. 1150).

Apples (1986) intensifiseringshypotese bekreftes langt på vei under norske forhold i en rapport som inngikk i Tidsbrukutvalgets undersøkelser med henblikk på tiltak som kunne gi bedre utnyttelse av tidsressursene i grunnskolen for å styrke elevenes læringsvilkår og -resultater. Rapporten er skrevet av Nyen, Jordfald og Seip (2009) på bakgrunn av en spørreskjemaundersøkelse blant 1177 grunnskolelærere og 1273 lærere i videregående skole. Av sammenligninger i rapporten fremgår det at klart flere lærere opplevde tidspress og å ha for mye å gjøre i 2009 sammenlignet med i 2006. Nyen m.fl. (2009) setter dette i forbindelse med innføringen av reformen Kunnskapsløftet, som de mener endret lærerrollen og gjorde denne videre og mindre avgrenset i ansvar for elevenes læringsresultater.

Lærerne i Nyens m.fl. (2009) undersøkelse mener at «alt» krever mer tid i arbeidshverdagen, men peker i svarene særlig på dokumentasjon av vurdering og etterarbeid med tester. De mener også det går mer tid til foreldrekontakt, arbeid med lokale læreplaner og planlegging av samarbeid med andre lærere enn før. Interessant for denne evalueringens vedkommende er at lærerne ønsker å bruke *mindre* tid på skriftlig dokumentasjon av vurdering, som de sier de er pålagt gjennom nye

forskrifter og lover, og *mer* tid på å gi tilbakemeldinger til elever. Det å gi tilbakemeldinger til elever er også en del av vurderingsarbeidet. Slike utsagn kan tyde på at lærerne kategoriserer og omtaler ulike aspekter ved vurderingsarbeidet på ulik måte.

1.7.3 Norsk lærernes arbeidssituasjon

Selv om det er forsket en del på læreres arbeidssituasjon de siste årene, er norsk læreres arbeidssituasjon og disse lærernes meninger og holdninger til eget fag et underutforsket felt (Breivik 2011: 42). Vi har funnet to masteroppgaver som behandler slike problemstillinger fra ulike faglige ståsted og vinkler.

Moe (2013) har utført en diskursanalyse i faget norskdidaktikk på basis av tre intervjuer med like mange norsklærere i videregående skole. Han undersøker hvordan disse norsklærerne konseptualiserer norskfaget, og hvordan dette kan sies å samsvare med læreplanen i norsk. Hans konklusjon er at de tre norsklærerne gir ulike svar på hva norskfaget er og bør være, og at de likeledes kan stå som representanter for ulike syn på norsklærerrollen. Disse to momentene henger sammen i Moes analyse, slik at norsklæreren som beskriver faget som et litteraturfag, legger vekt på formidlingsglede og «oppdragelse» til litteraturforståelse i norsklærerrollen. Norsklæreren som legger vekt på nyttesiden i faget, som opplæring i lesing og skriving, viser til dette når hun beskriver sin egen rolle.

Moe (2013: 80) skriver videre at fremfor å knytte norskfaget til læreplanen, viser norsklærerne i studien heller til undervisningssituasjoner i klasserommet og i en viss grad til egne følelser når de omtaler faget og egen undervisning. Av dette kan vi få forståelsen av at det ikke nødvendigvis er læreplanen som sterkest påvirker hvordan norsklærerne opplever det å undervise i norsk. Ettersom vurderingsordningene er tett knyttet til læreplanen, er det på denne bakgrunnen også nødvendig å sette spørsmålstegn ved oppdragets premiss om at vurderingsordningen vil ha avgjørende innflytelse på planleggingen og gjennomføringen av undervisningen. Moes (2013) arbeid bygger imidlertid på kun tre intervjuer, og kan i denne sammenheng derfor bare brukes til å utlede dette spørsmålet til videre diskusjon.

Et blikk på Skjærseths (2009: 44) masteroppgave i yrkesrettet spesialpedagogikk kan utdype sammenheng mellom læreplan, undervisning og vurdering når hun skriver at blant 80 norsklærere som svarte på en spørreskjemaundersøkelse om lærergjerningen, svarte to tredjedeler av dem at de brukte læreboka som det primære utgangspunktet for undervisningen. Skjærseth (2009: 62) skriver: «Faren ved å følge læreboka fra perm til perm, er at undervisningen lett kan gi læreren en opplevelse av å ikke rekke pensum og være i tidsnød.» Ingen av norsklærerne i de to masteroppgavene vi refererer her, avviste læreplanen som redskap, selv om 12 av Skjærseths 80 respondenter mente at undervisningen deres ikke var påvirket

av læreplanen (Skjærseth 2009: 43). Lærerne oppga videre at deres viktigste kilde til faglig egenutvikling var det kollegiale samarbeidet, som rangeres høyere enn egen utdanning og faglitteratur. Lærerne foretrakk også å reflektere over egen praksis, gjerne i samtale med kolleger. De vanligste samtaleemnene i disse situasjonene er ifølge undersøkelsen «fag og faglige opplegg» (62 av 80 respondenter svarte dette), «problematferd» (60 av 80 respondenter) og «elevvurdering (59 av 80 respondenter)» (Skjærseth 2009: 42).

Breivik (2011) skriver i en artikkel i *Norsklæreren* om Språkrådets (Fagrådet for skole og offentlig forvaltning) undersøkelse om norsklæreres meninger og holdninger til eget fag blant 1594 lærere ved henholdsvis 280 videregående og 463 ungdomsskoler. Det Breivik betegner som et stort flertall (ikke tallfestet) av norsklærerne mener at arbeidsmengden i faget er for stor. De yngste lærerne med kortest utdanning, minst erfaring og som arbeidet i ungdomsskolen, mente imidlertid at arbeidsmengden var overkommelig (s. 42). Høy utdanning og lang erfaring blant respondentene henger i undersøkelsen sammen med hvordan de vurderer de ulike emnene i norskfaget. Jo høyere utdanning og jo lengre erfaring disse lærerne har, jo mer vektlegger de betydningen av språk- og litteraturhistorie, tekstforståelse og undervisning i sidemålet. Jo lenger lærerne har arbeidet i skolen, jo mer enige er de i at norskfaget har gjennomgått store endringer i løpet av deres karriere (ibid.). Ungdomsskolelærerne i undersøkelsen har i gjennomsnitt kortere ansiennitet i skolen enn respondentene fra videregående skole, men ungdomsskolelærerne er gjennomgående mest positive til læreplanen. Norsklærerne i Nord-Norge synes arbeidsmengden i norskfaget er mindre enn hva gjennomsnittet av respondentene i undersøkelsen mener, og norsklærerne i Oslo har oftere faglige diskusjoner og er mer fornøyd med fagmiljøet enn gjennomsnittet av respondentene (Breivik 2011: 43)

1.7.4 Redusert antall karakterer i norsk: løsning søker problem?

Etter at gjeldende læreplan i norsk ble innført i august 2013, utledet utvalget bak NOU 2014: 7 *Elevenes læring i fremtidens skole* et generelt behov for fagfornyelse i skolen relatert til læreplaner og vurderingssystemer. Utvalget mente at et omfattende innhold i fag, herunder også norskfaget etter læreplanen av 2013, kommer i konflikt med det utvalget betegner som dybdelæring og progresjon. Norskfaget fremstår dermed fremdeles som omfattende etter læreplanrevisjonen i 2013. Samtidig finnes det få forslag til hvilke konsekvenser et gitt syn på norskfagets kjerneoppgaver vil ha for vurdering og vurderingspraksis. Vår litteraturgjennomgang i delrapport 1 fra evalueringen (Seland m.fl. 2015) gir grunnlag for å peke på mangelen på forskningsmessig oppdaterte og empirisk baserte studier av vurdering i norskfaget på et generelt grunnlag. Med innføringen av Kunnskapsløftet, utvikling

av tester og nye ordninger for vurdering generelt med konsekvenser for undervisning og vurdering i norskfaget, tydeliggjøres et behov for å innhente ny kunnskap om vurdering i norskfaget som helhet.

Forsberg og Lindberg (2010: 30) skriver i en kartlegging av svensk forskning i vurderingsfeltet, at det er å oppfatte som «att styra med svansen (...) når betygsystemet tar över på bekostnad av läroplaner och kursplaner och den syn på kunskap och lärande och mål som finns uttrycta i dessa.» Her uttrykker Forsberg og Lindberg (2010) et vurderingsperspektiv forankret i didaktisk tradisjon, mens Kunnskapsløftet kan sies å være sterkere forankret i en *curriculum*-tenkning, der mål-oppnåelse og resultater har en fremskutt posisjon. Vår fortolkning av Forsberg og Lindberg i deres didaktiske perspektiv er at vurderingssystemet skal ses som en funksjon av læreplanene, som igjen forankres i mål for opplæringen. I norsk kontekst er disse målene og planenes utforming forankret i utdanningsreformen Kunnskapsløftet, som gjennom rasjonale hentet fra målstyring i offentlig forvaltning gir et konsistent uttrykk for hvordan lærere skal arbeide med undervisning og følgelig vurdering.

Etter vårt syn er det derfor nødvendig å gå tilbake ikke bare til læreplanen hvis man skal diskutere vurderingsordningen i norskfaget, men også til grunnlaget og rasjonale for utdanningsreformen Kunnskapsløftet. Dette må ikke oppfattes som et partsinnlegg i en debatt for eller imot målstyring av grunnopplæringen, men en påpekning av det mest sentrale premiss for å forstå hva slags utfordringer og problemer lærere mener preger undervisning og vurdering i norskfaget, og som resulterer i ønsker og krav om endret vurderingsordning. Utdanningsdirektoratets (2013a) sammenfatning av høringsrunden som viser at mange lærere oppfatter at vurdering styrer undervisning, skal likevel ikke betraktes som en misforståelse av slike sammenhenger. Dette underbygges, men utdypes også i større kompleksitet av utvalget bak NOU 2014: 7, som skriver at implementering av læreplanene delvis avhenger av hvordan vurderingssystemene er integrert i de samme planene. Slik implementering, det vil si omsetting av planene til undervisning og vurdering i norskfaget, vil imidlertid også være avhengig av læreplanens form og innhold, og hvordan planen blir forstått og fortolket av aktørene som er ansvarlig for opplæringen (NOU 2014: 7, s. 109).

Hvis ovenstående sammenfattes, kan vi på følgende måte demonstrere hvor ulikt forbindelsene mellom læreplan, undervisning og vurdering kan fortolkes: Først har vi sammenhengen mellom læreplan og vurdering presentert på et ideelt og planmessig stringent vis der et visst kunnskapssyn og læringssyn kommer til uttrykk gjennom nasjonale mål for opplæringen, i vårt tilfelle Kunnskapsløftet. Disse forståelsene, nedfelt i dokumenter og arbeidsprosesser hvor faglige og politiske interesser brynes, resulterer i læreplaner for fag, som legger grunnlag for undervisning og vurdering. Denne ideelle sammenhengen samsvarer med Forsberg

og Lindbergs (2010) fremlegging av sammenhengen mellom læreplaner og vurdering.

Ballet og Kelchtermans (2009) videreutvikling av Apples intensifiseringshypotese ser imidlertid denne sammenhengen fra lærernes synspunkt. Da blir Kunnskapsløftets krav til vurdering en sentral faktor som påvirker undervisning og vurdering via lærernes arbeidssituasjon. Forenklet kan denne sammenhengen fremstilles som at utdanningsreformer (inkludert læreplanreformer) som i den senere tid har funnet sted sammen med en rekke tiltak for en mer effektiv skole, skaper en travel og presset arbeidssituasjon for lærerne. Lærernes arbeidssituasjon vil ifølge denne beskrivelsen påvirke deres arbeid med vurdering.

En delvis empirisk forankret sammenheng mellom læreplan og vurdering er uttrykt i NOU 2014: 7, forstått som at læreplanens form og innhold inkludert vurderingssystem vil sammen med aktørenes forståelse og fortolkning av læreplanen påvirke implementeringen av de samme læreplanene. Her gjøres i praksis vurderingssystemet til både noe som *bidrar til å forme* implementering av læreplanen og som *formes av* slik implementering, hvor begge deler må ses i sammenheng med læreplanens form og innhold.

I utlysning av dette oppdraget viser Utdanningsdirektoratet til lærernes høringssvar som setter faktorene "undervisning" og "vurdering" i en forenklet og dessuten motsatt rekkefølge, da undervisning omtales som en funksjon av vurdering. Vi kommer tilbake til samtlige av disse beskrivelsene av mulige sammenhenger mellom undervisning og vurdering i den avsluttende drøftingen i denne rapporten.

1.8 Metode

I det følgende gjør vi rede for hvordan vi har arbeidet med å samle inn og analysere evalueringsdata om henholdsvis karakterer, spørreundersøkelse til skoleledere og forskningsintervjuer blant skoleledere, lærere og elever som har deltatt i forsøket.

Forsøket har vært innrettet mot både ungdomsskolen og videregående skole. Vår analyse av karakterdata omfatter likevel kun de videregående skolene, da det ble stor oppslutning om forsøket på dette skolenivået, mens det derimot ikke ble så stor oppslutning om forsøket blant skolene på ungdomsskolenivået. Kun 25 ungdomsskoler har deltatt i forsøket, mot 184 videregående skoler. Vi har imidlertid gjennomført forskningsintervjuer på både ungdomsskoler og videregående skoler som har deltatt i forsøket.

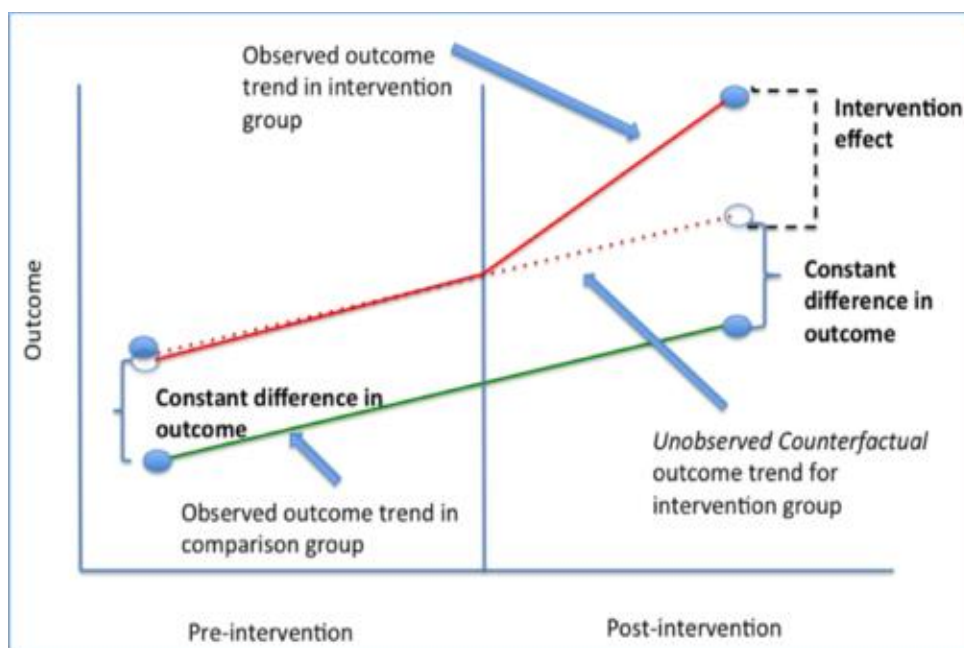
1.8.1 Analyse av karakterdata

For å undersøke eventuelle effekter av forsøket med redusert antall karakterer i norsk i Vg1 og Vg2 har vi brukt forskjell-i-forskjeller³ analyse. Forskjeller-i-forskjeller analyse er en form for kvasi-eksperimentelt forskningsdesign, det vil si en situasjon der vi har en gruppe som opplever tiltaket, vanligvis omtalt som *tiltaksgruppe*, og en gruppe som ikke opplever tiltaket, *sammenligningsgruppe*. Gruppene vi sammenligner er ikke tilfeldig fordelt, siden skolene selv fikk melde seg på forsøket, og evalueringen omtales dermed som et kvasi-eksperiment i stedet for et ekte eksperiment, siden det ikke oppfyller kravet om randomisert deltakelse. Ved å bruke longitudinelle data, der vi har informasjon om skolene også før forsøket med færre karakterer i norsk ble iverksatt, kan vi likevel undersøke om den alternative måten å vurdere norsk (med færre karakterer) i Vg1 og Vg2 har betydning for norsk-karakterene elevene oppnår i Vg3.

Kort forklart går forskjell-i-forskjeller metoden ut på å sammenligne *endringene* i utfall over tid mellom de som deltar i et forsøk og de som ikke deltar, for å se om de to gruppene utvikler seg ulikt med hensyn til utfallet før og etter forsøket. Utfallet i denne sammenheng er gjennomsnittskarakterer i norsk ved utgangen av vg3. Dersom karaktersnittet for tiltaks- og sammenligningsgruppen utvikler seg signifikant forskjellig før og etter forsøket, har vi funnet støtte for at forsøket har hatt en effekt.

For at denne metoden skal kunne si noe om effekten av forsøket, må tiltaksskolene og sammenligningsskolene være sammenlignbare over tid i perioden før forsøket. Dette kalles antakelsen om felles trend (common trend assumption) (Angrist & Pischke 2009, Lechner 2010, Stock & Watson 2015). Dersom tiltaksskoler og sammenligningsskoler er sammenlignbare kan vi være sikrere på at en eventuell effekt vi finner skyldes forsøket og ikke at de to gruppene uansett ville ha utviklet seg forskjellig i fravær av forsøket. En måte å undersøke denne antagelsen er ved å se på trenden i utfallsmålet i perioden før forsøket startet grafisk.

³ Ofte omtalt som difference-in-differences (DiD).



Figur 1.2: Forskjell-i-forskjeller, grafisk forklaring (Columbia University, 2018).

Metoden kan illustreres ved hjelp av figur 1.2. Den røde og grønne heltrukne linjen viser den observerte utfallsvariabelen for tiltaksgruppen (rød) og sammenligningsgruppen (grønn) i perioden før og etter forsøket. Vi antar at gruppene er sammenlignbare over tid i fravær av forsøket, som vil si at vi antar at den prikkete røde streken ville ha illustrert utviklingen i utfallsvariabelen i tiltaksgruppen dersom forsøket ikke hadde funnet sted. Dette blir ofte omtalt som den uobserverte kontrafaktiske situasjonen (unobserved counterfactual). For at vi skal tro på dette kontrafaktiske utfallet, er det viktig at vi observerer en konstant forskjell i utfallsmålet i tiden før forsøket som støtter antakelsen at gruppene er sammenlignbare over tid (common trend assumption). Merk at vi ikke trenger å anta at gruppene har like utfallsmål, vi trenger bare å anta at de utvikler seg likt over tid, hvilket er en mindre streng antakelse. Effekten av forsøket måles som forskjellen mellom det vi observerer og det vi tror vi ville ha observert i fravær av intervensjonen (intervention effect). Denne effekten finner vi ved å ta den observerte forskjellen mellom tiltaksgruppen og sammenligningsgruppen etter forsøket og trekke fra forskjellen mellom tiltaksgruppen og sammenligningsgruppen før forsøket.

En annen fordel med denne metoden, er at vi kan kontrollere for skolefaste effekter og årsfaste effekter, fordi vi har informasjon om hvilken skole elevene har gått på og vi har data fra perioden før og etter forsøket startet. Alle skolekjennetegn som ikke endrer seg over tid tas hensyn til med skolefaste effekter og endringer som påvirker alle i begge gruppene hvert år tas hensyn til med årsfaste effekter. I tillegg har vi informasjon om elevenes grunnskolepoeng, kjønn og utdanningspro-

gram som vi inkluderer som bakgrunnsvariabler i analysene. Gjennom å kontrollere for disse variablene undersøker vi at tiltaksskolene og sammenligningsskolene ikke er forskjellige fra hverandre på måter som kan påvirke deres trend over tid.

Avgrensninger og utvalg av dette datamaterialet sammenlignet til den totale datamengden om elevers standpunkt- og eksamenskarakterer er plassert innledningsvis i kapittel 4 i denne evalueringen, hvor også analysene presenteres.

1.8.2 Spørreundersøkelse blant skoleledere

For å hente inn data om skoleledernes erfaringer med forsøket med redusert antall karakterer ved halvårsvurdering i norsk har vi brukt Utdanningsdirektoratets faste spørreundersøkelse «Spørsmål til Skole-Norge». Denne undersøkelsen gjennomføres to ganger pr. år, og er organisert som en utvalgsundersøkelse. Det innebærer at det er laget tre faste og sammenliknbare utvalg av landets videregående skoler, og at deltakelse i spørreundersøkelsen rullerer mellom disse utvalgene hvert halvår. Omtrent en tredjedel av de videregående skolene fra hvert fylke inngår i hvert av de tre utvalgene. De videregående skolene som får tilsendt spørreundersøkelsen er derfor ikke trukket tilfeldig, men fordelt på utvalg etter kriteriene størrelse, geografi og skoletype. Resultatet er tre relativt like utvalg som hver for seg er representative for den totale populasjonen av videregående skoler i Norge. For at vi i størst mulig grad skulle få svar fra de samme skolelederne tidlig og sent i forsøksperioden, er spørsmålene som gjelder forsøket med redusert antall karakterer ved halvårsvurdering i norsk stilt til skolelederne i videregående skole høsten 2015 og våren 2017. Utvalgsmetoden tar likevel ikke høyde for at skoler kan ha byttet rektor i perioden.

«Spørsmål til Skole-Norge» blir også rutinemessig sendt til et utvalg grunnskoler, organisert etter samme prinsipp som beskrevet over. Vår opprinnelige plan i evalueringen var også å bruke denne undersøkelsen for å hente inn erfaringer fra skoleledere i ungdomsskoler i forsøket. Etter at forsøket var startet ved et stort flertall av de videregående skolene i landet, viste det seg imidlertid at bare et lite mindretall av ungdomsskolene var interessert i å delta i forsøket (se tidligere i dette kapittelet). Med et fast, roterende utvalg tilsvarende en tredjedel av landets grunnskoler i hver runde av spørreundersøkelsen, ville muligheten for å treffe disse ungdomsskolene som deltar i forsøket, være svært liten. I samråd med oppdragsgiver avgjorde vi derfor å ikke legge inn spørsmål om redusert antall karakterer ved halvårsvurdering i norsk i «Spørsmål til Skole-Norge» til grunnskolene.

«Spørsmål til Skole-Norge» høsten 2015 ble sendt til 130 videregående skoler, og deltakelsen var på 74 prosent. Representativiteten for videregående skoler ble vurdert som god (Gjerustad og Waagene 2015). Våren 2017 var svarprosenten blant videregående skoler 65 prosent, og representativiteten ble også da vurdert

som god (Federici, Gjerustad, Vaagene, Larsen m.fl. 2017). Det var 82 skoleledere i videregående som svarte på spørsmålet om skolen deltok i forsøket med utprøving av én eller to karakterer i halvårsvurdering i norskfaget, og 38 av disse (46 prosent) svarte ja. Det var også fire skoler som svarte at de ikke deltok nå, men hadde deltatt tidligere.

1.8.3 Forskningsintervjuer

Som ledd i datainnsamlingen i denne delen av evalueringen har vi besøkt fire videregående skoler og to ungdomsskoler som deltar i forsøket. De videregående skolene startet forsøket i 2013 eller tidligere, mens ungdomsskolene startet opp i 2014. Utvalgsstrategien skulle sikre at minst ett elevkull ved skolen hadde fulgt forsøket i Vg1 og Vg2/ på 8. og 9. trinn før vi besøkte skolen mens disse elevene var i Vg3 eller på 10. årstrinn og forberedte seg til standpunktvrdering med tre karakterer. Vi gjorde dette for at elevene skulle ha erfaringer fra hele forsøket og kunne sammenligne skoleår og vurderingsordninger.

Utvalget av skoler representerer variasjon med hensyn til valg av én kontra to karakterer ved halvårsvurdering. Oppdragsgiver ønsket også at utvalget skulle inneholde en skole hvor majoriteten av elevene har nynorsk som hovedmål. I tillegg ønsket vi en geografisk spredning i utvalget.

På bakgrunn av disse kriteriene og Utdanningsdirektoratets liste over deltaker-skoler/antall karakterer ved halvårsvurdering valgte vi ut i alt tolv skoler under til sammen seks skoleeiere. Vi tok kontakt med skoleeierne og informerte om at vi ønsket å ta kontakt med navngitt skole, og hvilken skole vi deretter ville ta kontakt med hvis den prioriterte skolen takket nei til besøk og forskningsintervjuer. Med hensyn til å finne en videregående skole hvor majoriteten av elevene har bokmål som sidemål, ba vi skoleeier om å foreslå en skole som vi kunne ta kontakt med. Alle de seks skolene som vi i tråd med denne prosedyren henvendte oss til først, takket ja til å få besøk og bidra i forskningsintervjuer om forsøket. Fordelingen av skoler fremgår av tabell 1.1. Her har vi utelatt opplysninger om geografi og skoleeier.

Tabell 1.1: Utvalg av skoler til forskningsintervjuer, fordeling med antall karakterer i forsøket og hovedmål for majoriteten av elevene

Skolenr.	Antall karakterer i forsøket	Hovedmål
1 (videregående skole)	Én karakter i Vg1, to karakterer i Vg2	Bokmål
2 (videregående skole)	Én karakter	Bokmål
3 (videregående skole)	To karakterer	Bokmål
4 (videregående skole)	Én karakter i Vg1, to karakterer i Vg2	Nynorsk
5 (ungdomsskole)	To karakterer	Bokmål
6 (ungdomsskole)	To karakterer	Bokmål

Slik det går frem av tabell 1.1, viser utvalget av skoler variasjon mellom skoler som prøver ut redusert antall karakterer ved halvårsvurdering i norsk med én kontra

to karakterer. Med hensyn til ungdomsskolene prøver flertallet av skolene ut ordningen med to karakterer, og vi valgte derfor to slike skoler for å styrke informan-tenes anonymitet i datainnsamlingen og -analysen. Skole 1 og skole 4 representerer begge løsninger, og kan om mulig brukes til å kontrastere valg mellom én eller to karakterer. Det må nevnes at det å kombinere karakterer og årstrinn på denne måten ikke er den vanligste løsningen blant forsøksskolene, og utvalget kan på denne måten sies å ha en skjevhet. Årsaken til at utvalget ble slik, var at skole 4 over tid har vekslet fra kombinasjonsløsning og det å ha likt antall karakterer i norsk på Vg1 og Vg2. Våre opplysninger om denne skolen skrev seg fra perioden der skolen hadde likt antall karakterer på begge årstrinn. Vi anser ikke dette som noen alvorlig skjevhet ved utvalget. De seks skolene har forskjellige skoleeiere, og utvalget favner en skole hvor majoriteten av elevene har nynorsk som hovedmål.

Ved henvendelse til rektor på hver av de seks skolene spurte vi om tillatelse og tilrettelegging for å gjennomføre i alt tre intervjuer ved hver skole. Disse intervjuene var 1) individuelt intervju med representant for skoleledelsen som hadde ansvar for forsøket, 2) gruppeintervju med mellom tre og fem norsklærere som hadde erfaring med forsøket ved skolen, 3) gruppeintervju med mellom tre og fem elever i Vg3 studieforberedende som hadde fulgt forsøket i Vg1 og Vg2 ved skolen (videregående skoler) eller tilsvarende en gruppe med tre og fem elever på 10. trinn (ungdomsskoler). Med hensyn til sammensetning av elevgruppen ba vi om å få snakke med elever med forskjellig prestasjonsnivå i norsk. Vi oppfatter at dette ble tatt til følge, bedømt ut fra hvordan elevene responderte på spørsmål som for eksempel hvilke utdanningsplaner de hadde for det påfølgende året.

Med denne fremgangsmåten har vi møtt til sammen 20 norsklærere i gruppeintervju, fordelt mellom grupper bestående fra to til seks personer på de ulike skolene. Av disse 20 norsklærerne er det 16 kvinner. Norsklærernes ansiennitet som lærer varierte fra helt nyutdannet (et halvt års ansettelse på gjeldende skole) til 38 år i læreryrket. Flertallet av de 20 norsklærerne har arbeidet som lærer i mer enn ti år. Vi har intervjuet til sammen 19 elever. Blant disse var det omtrent like mange gutter som jenter.

Under intervjuene fulgte vi en intervjuguide. Denne er trykt som vedlegg til delrapport 2 (Seland m.fl. 2016).

Individuelle forskningsintervjuer og gruppeintervjuer/fokusgruppeintervjuer skiller seg fra hverandre som metode for datainnhenting, og i vårt design hadde vi i tillegg lagt to ulike perspektiver til grunn for henholdsvis individuelt intervju med leder og fokusgruppeintervjuene med norsklærere og elever.

I det individuelle intervjuet med representant for skoleledelsen ønsket vi å kartlegge skolens forberedelse og tilrettelegging av forsøket. I dette tilfellet brukte vi intervjuet som grunnlag for faktainnhenting med en naturalistisk tilnærming

(Gubrium & Holstein 1997). Det innebærer at vi ville høre hvordan skolen har organisert forsøket, og ikke primært har vært ute etter å undersøke lederens opplevelser, holdninger eller oppfatninger forbundet med dette, selv om slike spørsmål og utsagn også finnes i disse intervjuene. I analysen av disse intervjuene har vi også sett etter og gjengitt ledernes utsagn slik de fremkommer i materialet, uten å foreta noen videre fortolkning av utsagn eller opplysninger. Skolelederintervjuene gir slik en bakgrunn med historikk og lokale løsninger for praktisk tilrettelegging forbundet med forsøket ved skolene.

Med basis i litteraturstudiets funn om at lærere organiserer undervisningen i norskfaget forskjellig, ønsket vi å belyse denne ulikheten og samtidig få fram mulige likheter mellom lærernes praksis og erfaringer ved å la lærerne selv få diskutere dette i grupper. Fokusgruppeintervjuer skiller seg fra individuelle intervjuer ved at man i gruppen kan undersøke *fokusert sosialitet* (Halkier 2012: 133), det vil si hvordan diskusjonen om et utvalgt emne utvikles gjennom det sosiale samspillet i gruppen. Hoveddelen av spørsmålene i intervjuguiden setter lærernes meninger, oppfatninger, erfaringer og holdninger i sentrum. Dette er normative problemstillinger, som med hell kan belyses gjennom gruppediskusjoner hvor deltakerne bygger på hverandres utsagn. Det samme er tilfelle i elevintervjuene, hvor vi i stor grad har spurt etter elevenes opplevelser, erfaringer, synspunkter og holdninger.

I lærerintervjuene stilte vi få og åpne spørsmål og lot diskusjonen utvikle seg. I elevintervjuene var det nødvendig å bruke en litt strammere tilnærming, det vil si at moderator (intervjuer) stilte flere spørsmål og elevene gjerne svarte én og én (selv om de satt i gruppe). Spesielt i startfasen kan det være formålstjenlig at alle deltakere får komme til orde. Det varierte hvorvidt elevgruppene kom godt i gang med diskusjoner eller ikke, noe som kan bero på at de ikke alltid var godt forberedt på intervjusituasjonen. Dette kunne ha årsaker som at lærer kanskje ikke hadde gitt eleven informasjonen fra forskerne på forhånd, eleven hadde glemt dette eller eleven ble hentet rett inn til intervjuet fra undervisningen uten at dette var avtalt med eleven på forhånd. I motsetning til lærerne kjente heller ikke alltid elevene hverandre så godt, særlig ikke på store skoler. Ved en av de store videregående skolene kom de fire elevene som deltok i intervjuet fra hver sin klasse, mens ved den minste videregående skolen var alle de fem elevene som ble intervjuet fra samme klasse. Ved en tredje skole måtte elevintervjuene gjennomføres fordelt på to grupper, hver med to elever fra forskjellige klasser. Mens mange lærere «fløt» naturlig på kollegaforholdet og sammenlignet erfaringer basert på mange tidligere faglige diskusjoner, hadde ikke elevene nødvendigvis den samme opplevelsen av at de hadde deltatt i et forsøk med klar innretning og hensikt. Moderator spurte etter elevenes generelle erfaringer med vurdering, karaktersetting og etter hvert om forberedelser til den forestående avslutningen av skoleåret med heldagsprøver

og eksamen. Elevene ble stadig oppfordret til å sammenligne erfaringene fra norsktimene i Vg3/10. årstrinn med erfaringene de hadde med faget fra de to foregående årene. Ved en av de to ungdomsskolene viste det seg imidlertid at skolen hadde kommet sent i gang med forsøket, og elevene på 10. trinn hadde derfor *ikke* erfaringer fra å få halvårsvurdering med redusert antall karakterer i norsk. Disse elevenes utsagn om norskfaget og det å bli vurdert med karakterer er like fullt beholdt i vårt materiale.

I gruppeintervjuer kan situasjonen ligge godt til rette for at forskerne kan bruke ulike hjelpemidler som kan bidra til å fokusere emnet for diskusjonen. Ifølge Halkier (2012) kan hjelpemidler være en oppgave som lar deltakerne forklare for forskeren og hverandre, gi uttrykk for sine vurderinger og respondere på andres utsagn, det vil si sette fart i diskusjonen. Til lærerintervjuene hadde vi derfor kopiert opp fem kompetansemål fra læreplanen i norsk for Vg2 (studieforberedende utdanningsprogram) og 10. årstrinn i ungdomsskolen. Kompetansemålene var hentet fra emnet «skriftlig kommunikasjon». Årsaken til at vi hadde valgt kompetansemål fra Vg2, var at vi regnet med at norsklærerne på videregående skole organiserer arbeidet i henhold til fag fremfor trinn (Sandvik og Fjørtoft 2014: 117), og at de alle derfor ville være kjent med Vg2-målene. Dette viste seg ikke å holde stikk i praksis, da flere skoler hadde lærere som bare underviste på ett trinn, for eksempel Vg1. Disse lærerne hadde fagsamarbeid med kolleger, men ikke knyttet til de spesifikke kompetansemålene på hvert trinn. Sandvik og Fjørtoft (2014: 117) skriver for så vidt også at lærere i videregående skole synes å være mer individuelt orienterte enn lærere i barne- og ungdomsskolen. Når vi på en skole møtte lærere som underviste på enten Vg1 og Vg3, var derfor kompetansemål på Vg2 uten stor relevans for diskusjonen.

Vår opprinnelige hensikt med å bruke kompetansemålene fra Vg2 var å få lærerne til å fortelle om hvordan de faktisk arbeider med undervisning og vurdering i henhold til kompetansemål før og under forsøket med redusert antall karakterer i norsk. Dette er et hovedspørsmål i evalueringen, og vi håpet på denne måten å få frem data som kunne la seg sammenligne på tvers av lærergrupper og skoler. Mens vi med oppgaven klarte å få lærerne inn i en bredere anlagt diskusjon av undervisning og kompetansemål, ble ikke dataene sammenlignbare i den grad vi hadde sett for oss med denne metoden. Norsklærernes fortellinger om undervisning og vurdering under forsøket er derfor preget av den sosiale samhandlingen som utviklet seg i gruppene, hvor data i høy grad er kontekstuellt avhengig, relasjonelt påvirket av intervjusituasjonen og – etter vår oppfatning – potensielt foranderlige (Halkier 2012). Dette gjør det vanskelig, for ikke å si uforsvarlig, å trekke konklusjoner om all undervisning og vurdering som foregår i regi av lærere og skoler i forsøket, men gir på den andre siden rike eksempler fra skolene som vi har besøkt i forbindelse

med denne evalueringen. Slik det er nevnt med bakgrunn i vår litteraturgjennomgang tidligere i dette kapitlet synes det også å være stor variasjon i norsklærernes arbeidsmåter generelt.

Lydopptakene fra skoleleder- og lærerintervjuene er lyttet gjennom og sammenfattet i tråd med spørsmålene i intervjuguiden. Dette resulterte i en mellomanalyse med lengde mellom to og fire tekstsider, som deretter har blitt brukt som utgangspunkt for datagjennomgang og analyser i kapitlene som tar for seg data fra disse nivåene. Lydopptakene av gruppeintervjuene med elever ble transkribert. Disse tekstene ble kodet i NVivo og gjengitt tema for tema, som: hva elevene synes om norskfaget; om det er bedre å ha en, to eller tre karakterer; hvordan elevene gjøres kjent med vurderingskriterier; samarbeid på tvers av norskfaget og andre fag. Temaene springer delvis ut fra konkrete spørsmål som ble stilt, men vi har også kunnet finne en uttalelse som relevant for et tema uavhengig av spørsmål.

Med unntak av transkriberingen har forfatteren av kapitlet vært involvert i alle deler av prosessen. I dette forløpet ligger det mange valg og avveininger, som er beskrevet innledningsvis i kapitlet, for å illustrere hvordan prosessen frem til formidlingen er preget av mange tolkninger og vurderinger underveis.

Beskrivelsene av elevenes erfaringer og synspunkter er gitt på grunnlag av gjennomlesning av kodet tekst, etter hvert supplert med nye beskrivelser når vi fant grunn til å utdype eller nyansere fremstillingen. På denne måten har kapitlet preg av at gjengivelsen ligger tett på elevers uttalelser før vi tolker disse og trekker linjer på tvers av materialet for å belyse de relevante problemstillingene.

1.9 Rapportens oppbygning

I denne rapporten skal vi videre gå nærmere inn i sentrale dokumenter som styrer undervisning og vurdering i norskfaget (kapittel 2). Dette, sammen med kapittel 3 om undervisning og vurdering i norskfaget, utfyller vår modell av grunnlaget for evalueringen, slik denne er gjengitt i figur 1.1 i dette innledningskapittelet.

Etter dette følger presentasjon av data. Dette skjer først i form av analyse av karakterdata (kapittel 4), som er denne sluttrapportens sentrale bidrag til evalueringen av forsøket. Deretter følger en oppdatert fremstilling av datainnhentingene fra henholdsvis skoleledernivået (kapittel 5), lærernivået (kapittel 6) og elevnivået (kapittel 7) i evalueringen. Av disse tre kapitlene bygger skolelederdataene på spørreundersøkelser og forskningsintervjuer, mens lærer- og elevkapitlet kun bygger på forskningsintervjuer. Alle disse tre kapitlene har tidligere vært publisert i en foreløpig form i delrapport 2 fra evalueringen (Seland m.fl. 2016).

Avslutningskapittelet (kapittel 8) tar for seg sluttrapportens hovedproblemstillinger, slik disse også er presentert tidlig i dette kapittelet. Etter at vi først har opp-

summert resultatene fra kapittel 4 om hvorvidt forsøket har hatt betydning for elevenes læring i norskfaget, foretar vi en diskusjon av for hvilke målgrupper forsøket kan sies å ha vært hensiktsmessig. Helt til slutt i rapporten gir vi våre anbefalinger for videre utvikling av vurderingsordninger i norskfaget.

2 Sentrale dokumenter som styrer undervisning og vurdering i norskfaget

I dette kapitlet presenterer vi politiske dokumenter og initiativ som har hatt til hensikt å forme og utvikle norskfaget inkludert læreres vurderingspraksis. Vi har avgrenset dette bakover i tid til utdanningsreformen Kunnskapsløftet, som trådte i kraft i 2006. Reformen omtales som en styrings- og læreplanreform (Møller m.fl. 2013; Aasen m.fl. 2012), og disse to aspektene virker sammen i hvordan alle lærere forventes å arbeide med undervisning og planlegging etter reformen.

Grunnlagsdokumentene for Kunnskapsløftet, læreplan for norskfaget, forskrift til opplæringsloven kapittel 3 om individuell vurdering og rundskriv om individuell vurdering er sentrale dokumenter som angir nasjonale utdanningsmyndigheters bestemmelser og føringer for undervisning og vurdering i norsk.

Vi begynner kapitlet med en kort presentasjon av utdanningsreformen Kunnskapsløftet før vi fortsetter med gjeldende læreplan for norskfaget supplert av veiledning til læreplanen. Deretter ser vi nærmere på prosessen som førte frem mot denne læreplanen i perioden 2010-2013, fordi denne revisjonen resulterte i vedtak om forsøket med én eller to karakterer i norsk. Deretter tar vi for oss vurderingsforskriften med tilhørende dokumenter og et kort blikk på satsingene Bedre vurderingspraksis og Vurdering for læring. Vi avrunder dette temaet med OECDs rapport om blant annet elevvurdering i norsk skole fra 2011, for å få et internasjonalt perspektiv på myndighetenes arbeid med utvikling av vurderingsarbeidet i Norge. Helt til sist ser vi kort på nynorsk slik dette behandles i myndighetenes språkpolitiske dokumenter.

En tidlig versjon av dette kapitlet ble trykt i delrapport 1 fra evalueringen av forsøket med én eller to karakterer ved halvårsvurdering i norsk (Seland m.fl. 2015). Kapitlet er nå oppdatert med dokumenter som er nyere enn 2015.

2.1 Kunnskapsløftet som styringsreform og læreplanreform

Utdanningsreformen Kunnskapsløftet endret styringsstrukturen i grunnsopplæringen ved å styrke skoleeierens beslutningsmyndighet og handlefrihet. Ansvaret for beslutninger og gjennomføring av disse i forvaltningen av skolen skal foretas så lokalt som mulig, for å legge til rette for en kultur for læring og utvikling av skolen som lærende organisasjon (Møller m.fl. 2013: 24, 26). Stortingsmelding nr. 30 (2003-04) *Kultur for læring* avgitt av Utdannings- og forskningsdepartementet begrunnet hovedlinjene i Kunnskapsløftet på følgende måte:

Forestillingen om at staten kan skape et likeverdig skoletilbud gjennom detaljregulering og –styring, erstattes med tillit til at den enkelte lærer, skoleleder og skoleeier selv har de beste forutsetninger for å vite hvordan god læring kan skapes og gjennomføres, innenfor rammen av nasjonale mål.

Utdannings- og forskningsdepartementet (2004: 25)

Sentrale virkemidler i reformen var blant annet å innføre kompetansemål i læreplanene, vektlegging av resultat kvalitet og dokumentasjon av oppnådde resultater (Møller m.fl. 2013: 27). Reformen etablerte på denne måten læreplanene som en funksjon av målstyring, uttrykt av Møller m.fl. (2013: 26) som at «[p]olitikerne skulle stoppe ved skoleporten, men de skulle ha stor oppmerksomhet mot elevene som gikk ut av porten, og hvilke resultater som ble oppnådd.» Disse forskerne setter på denne måten Kunnskapsløftet i kontrast til læreplanreformene på 1990-tallet, som var mer opptatt av å skape et felles kunnskapsgrunnlag under sterk statlig styring i vern av en offentlig fellesskole (ibid.: 25). I den norske læreplanhistorien representerte dette et tydelig skifte, da læreplanene gikk fra å ha en hovedsakelig innholdsmessig orientering til å vektlegge kompetansemål fremfor innhold i undervisningen (NOU 2014: 7, s. 20; 71; Tveit 2013: 222). Endringen var forberedt ved læreplanene under Reform 94, som ble utformet under en begynnende utvikling av målstyring i forvaltningen av grunnsopplæringen. Læreplanene skulle ikke være detaljerte, men kunnskapsmålene skulle suppleres med hovedmomenter som skulle bidra til presisering og oppfyllelse av målene. Ifølge NOU 2014: 7 var det å forenkle læreplanene en av intensjonene med Kunnskapsløftet (vår utheving). Dette fikk konsekvens gjennom vektlegging av metodefrihet for skoler og for lærere, underbygget av forventinger om lokalt læreplanarbeid (s. 70-71).

Aasen m.fl. (2012) understreker i sluttrapporten fra evalueringen av Kunnskapsløftet som styringsreform at skolers og læreres frihet har blitt strammet inn igjen etter innføringen av reformen. Dette har blant annet skjedd fordi internasjonale sammenligninger av elevenes læringsresultater har pekt på mulige feil og mangler ved opplæringen i norsk skole, og i tråd med hva Aasen m.fl. (2012) skriver, er det nasjonale myndigheter som må svare for svakheter ved opplæringssys-

temet. En forenklet beskrivelse av dette handlingsforløpet er at nasjonale myndigheter først tok to skritt tilbake fra lokalnivået i opplæringen ved innføringen av Kunnskapsløftet, og deretter et skritt frem for i sterkere grad å bygge opp under skolenes arbeid med nasjonale målsettinger i opplæringen. Følgen er at lærerne kan oppleve at den hierarkiske styringen av skolen har blitt sterkere og den lokale friheten mindre etter innføringen av Kunnskapsløftet. Dette inntrykket formidles også gjennom rapporten fra Nyen m.fl. (2009) om læreres tidsbruk før og etter innføringen av Kunnskapsløftet.

Som læreplanreform innførte Kunnskapsløftet nye læreplaner i alle fag. Disse læreplanene er bygget opp på en måte som skiller seg fra den tidligere norske læreplantradisjonen, med kompetansemål for hva elevene skal mestre på bestemte årstrinn uten nærmere beskrivelse av innholdet i opplæringen. Kunnskapsløftet ga derfor skolen og lærerne økt frihet til å bestemme lokale læringsmål og utforme undervisningen i tråd med dette (Karseth og Engelsen 2013).

Læreplanen i norskfaget fra 2013 oppgir dermed kompetansemål for bestemte (ikke alle) årstrinn i grunnopplæringen. For elever i ungdomsskolen oppgis det for eksempel bare kompetansemål for 10. årstrinn. Dette betyr at lærerne må omsette disse kompetansemålene til læringsmål for opplæringen også på 8. og 9. årstrinn gjennom lokalt arbeid med læreplanene. I videregående opplæring, studieforberevende utdanningsprogram, oppgir læreplanen kompetansemål for Vg1, Vg2 og Vg3. Veiledning til læreplanen gir eksempler på hvordan kompetansemål kan omformes til læringsmål på lokalt plan.

2.2 Læreplan for norskfaget og veiledning til læreplanen

Læreplanen i norskfaget har status som forskrift fastsatt av Kunnskapsdepartementet, og er gjeldende fra 1. august 2013. Fagets formål slås fast med følgende formulering: «Formålet med opplæringen er å styrke elevenes språklige trygghet og identitet, utvikle deres språkforståelse og gi et godt grunnlag for mestring av begge målformene i samfunns- og yrkesliv.»

Norskfaget skal ifølge formålsbeskrivelsen utvikle elevenes språkkompetanse med utgangspunkt i den enkeltes evner og forutsetninger. Det vises til at det å utvikle muntlige ferdigheter, lese- og skrivekompetanse både vil være et mål i seg selv og gi det nødvendige grunnlaget for elevens læring og forståelse i alle fag på alle trinn. Elevene skal lære å orientere seg i et bredt spekter av tekster, og skriftlig og muntlig være i stand til å sette ord på egne tanker og meninger. Formålsbestemmelsen omtaler også kulturarven, og fastslår at elevene gjennom norskfaget skal oppmuntres til å bli aktive bidragsytere i å utvikle kulturarven som en levende tradisjon i et spenningsfelt mellom norsk språk, kultur og litteratur i et historisk og et

internasjonalt perspektiv. Det språklige mangfoldet i Norge omtales som en ressurs for utviklingen av barn og unges språkkompetanse, og med utgangspunkt i denne språksituasjonen skal barn og unge få kjennskap til språklig mangfold, og lære å lese og skrive både på bokmål og nynorsk.

Planen har tre hovedområder: muntlig kommunikasjon, skriftlig kommunikasjon og området språk, litteratur og kultur. Muntlig kommunikasjon er definert som både å lytte og tale, gjennom å oppfatte, fortolke og vurdere andres utsagn og selv å kunne uttrykke kunnskap, tanker og ideer. Skriftlig kommunikasjon er å kunne lese og skrive norsk gjennom systematisk opplæring som dekker hele opplæringsløpet. Språk, litteratur og kultur dekker både norsk og nordisk språk- og tekstkultur, og skal også innlemme internasjonale perspektiver. Elevene skal utvikle en selvstendig forståelse av norsk språk og litteratur og få innsikt over hvordan disse størrelsene har endret seg over tid.

Som nevnt angir læreplanen i norsk kompetansemål for hvert av disse tre hovedområdene etter 2., 4. 7. og 10. trinn i ungdomsskolen og etter Vg1, Vg2 og Vg3 i studieforbereende utdanningsprogram i videregående skole.⁴ For ungdomsskoler som deltar i forsøk med én eller to karakterer i norsk på 8. og 9. trinn er det derfor undervisning og vurdering frem mot kompetansemålene for 10. trinn som vil aktualiseres, mens for videregående skole vil det være kompetansemålene for Vg1 og Vg2 i det studieforbereende utdanningsprogrammet som skal dokumenteres med redusert antall karakterer ved halvårsvurdering under forsøket. Vi gjengir ikke disse kompetansemålene her, men henviser til den punktvis fremstillingen under hvert av de tre hovedområdene for faget på det relevante årstrinnet i selve læreplanen (Kunnskapsdepartementet 2013: 8-9; 10)

Læreplanen angir at elevene på 10. årstrinn i ungdomsskolen skal ha tre standpunkt-karakterer, én i norsk hovedmål skriftlig, én i norsk sidemål skriftlig og en i norsk muntlig. Det samme er tilfelle for Vg3 studieforbereende utdanningsprogram. På 10. årstrinn kan elevene trekkes ut til to dagers skriftlig eksamen som omfatter norsk hovedmål og norsk sidemål. Elevene kan også trekkes ut til muntlig eksamen i norsk. På Vg3 studieforbereende utdanningsprogram skal elevene ha skriftlig eksamen i norsk hovedmål, og elevene kan også trekkes ut til eksamen i norsk sidemål. Elevene kan i tillegg trekkes ut til muntlig eksamen i norsk. Ordinær ordning for halvårsvurdering i faget gjenspeiler med andre ord hovedinndelingen i fagets eksamensordning.

Utdanningsdirektoratet har gitt ut *Veiledning til læreplan i norsk*, gjeldende fra 20. august 2013. Veiledningen (Utdanningsdirektoratet 2013b) viser hvordan de

⁴ For yrkesfaglige utdanningsprogrammer angir læreplanen kompetansemål etter Vg2 og etter påbygging til studiekompetanse. Forsøk med én eller to karakterer i norsk gjennomføres ikke i yrkesfaglige utdanningsprogrammer i videregående skole fordi faget der bare har én standpunkt-karakter. Vi utelater derfor slike utdanningsprogrammer i videre omtale av læreplanen i norsk.

grunnleggende ferdighetene lesing, skriving og muntlige ferdigheter kan integreres i opplæringen, og hvordan kompetansemål kan utformes på lokalt plan. Med henvisning til formålet for faget slås det i veiledningens innledning fast at norskfaget både er et dannelsesfag som skal styrke språklig identitet og aktiv deltakelse i samfunnslivet, og et ferdighetsfag som gir grunnlag for læring i alle fag. Veiledningen er bygget opp av praktiske eksempler som viser hvordan læreren kan arbeide med kompetansemål på gitte årstrinn, fra refleksjon over hva eleven skal lære (kompetansemålene operasjonalisert til læringsmål), via elevens motivasjon og interesse for læringsarbeidet, det innhold, den metode og organisering som undervisningen krever, underveisvurdering, det å skape variasjon og legge til rette for progresjon.

2.2.1 Arbeidet frem mot gjeldende læreplan i norsk

Den første læreplanen i norsk under Kunnskapsløftet, som ble avløst av gjeldende plan fra 1. august 2013, var den første eksponenten for vektleggingen av mål for elevenes kompetanse fremfor detaljerte pålegg om innholdet i dette faget. Dette er sentralt for å forstå debatten som ledet frem mot læreplanrevisjonen i 2013 og bakgrunnen for diskusjonen om antall karakterer i norsk. Vi refererer i det følgende noen hovedtrekk i debatten som gikk forut for denne læreplanrevisjonen.

Arbeidet med revisjon av den første læreplanen i norsk under Kunnskapsløftet⁵ startet formelt opp i 2010 ved oppdragsbrev fra Kunnskapsdirektoratet til Utdanningsdepartementet. Brevet inneholdt flere føringer for helhetlig gjennomgang av norskfaget, og departementet la til grunn at gjennomgangen skulle legge vekt på å utvikle elevenes grunnleggende språkferdigheter – å lese og å uttrykke seg muntlig og skriftlig. Departementet ba også om at gjennomgangen av faget skulle inkludere sluttvurderingsordningene for hovedmål og sidemål, og ba etter ny korrespondanse med direktoratet om at forslag til nye vurderingsordninger i faget ble sendt på høring sammen med forslag til ny læreplan. Direktoratet la deretter til rette for «Forum for norskfaget», som skulle komme med innspill til ny læreplan i norsk. Dette arbeidet resulterte i en rapport i 2012, og direktoratet kunngjorde forslag til hvordan faget kunne revideres (Søyland 2012: 6; Steinfeld 2012b: 84). Parallelt med den formelle revisjonsprosessen i en læreplangruppe foregikk en debatt i fagmiljøet, blant annet i (men ikke begrenset til) tidsskriftene *Norsklæreren*, *Syn og*

⁵ Den første læreplanen i norsk under Kunnskapsløftet ble fastsatt i 2005. Etter dette ble planen justert i 2008 og i 2010, første gang for å styrke leseopplæringen og andre gang for å gi planen økt relevans for Vg1 studieforberedende utdanningsprogram og Vg2 yrkesfaglige utdanningsprogram (kilde: Høringsbrev om endringer i læreplanen i norsk, sendt fra Utdanningsdirektoratet til høringsinstansene 5. desember 2012).

Segn og Ordet: kvartalsskrift for norsk språk og kultur. Debatten illustrerer en forbindelse mellom innholdet i læreplanen sett i lys av planens kompetansemål og vurderingsordningene i faget.

Hovedargumentet som de fleste av deltakerne bruker som utgangspunkt i debatten, er en påstand om at arbeidsbyrden for norsklærerne er for stor og at faget er for omfattende. Et utdrag fra en av debattartiklene i *Norsklæreren* gir denne bakgrunnen for argumentet om norsklærernes arbeidsbyrde:

Norsk har alltid hatt en stor rettebyrde, og selv om leseplikten er og har vært lavere enn for andre fag, oppveier den lave leseplikten på ingen måte merarbeidet. Med den nye læreplanen i Kunnskapsløftet ble likevel arbeidsbyrden enda tyngre enn den var før. Jeg tror at økningen i arbeidsbyrde i stor grad skyldes to forhold: tallet på karakterer og innholdet i læreplanen. Tidligere var det på allmennfag to karakterer i 1. klasse og tre i 2. og 3. klasse, hvor faget i det minste hadde fem timer. Med Kunnskapsløftet ble det innført tre karakterer på alle trinn. Forskyvningen av timetallet fra 4-5-5 til 4-4-6 samtidig som vi skulle sette tre karakterer i faget på alle trinn, har ført til et betydelig merarbeid.

(Sannrud 2010: 74)

Flere debattanter bidrar i disse tidsskriftene med innlegg som støtter argumentet om at lærernes arbeidsbyrde er for stor og at norskfaget følgelig bør nedskaleres, dog med ulike tilnærminger. Stensby (2011) og Skirbekk (2013a) viser til at faget mangler tilstrekkelig avgrensning mot andre fag i skolen, så som historie og samfunnsfag. Dette fører ifølge disse innleggene til at norskfaget får ansvar for emner som kan sies å høre inn under andre fag i opplæringen. Stensby (2011) viser også til vektleggingen av muntlige vs. sammensatte tekster i faget, et eget hovedområde i den første læreplanen i norsk under Kunnskapsløftet. Hennes forslag er at disse hovedområdene kan slås sammen.

Flere innlegg i debatten om norskfagets omfang og innhold viser også til vektingen mellom hovedmål og sidemål. De kontrasterende argumentene består av at hovedmålet må få større plass i undervisningen på bekostning av sidemålet (Nyheim 2011; 2012), motsatt at omfanget av undervisning i sidemålet må opprettholdes på dagens nivå (Stensby 2011; Eide 2012; Nygård og Sture 2013). Dette munnar i flere av innleggene ut i argumenter for eller imot redusert antall karakterer i norsk. To slike innlegg fremmer et syn om at redusert antall karakterer i norsk vil svekke sidemålets stilling (Søyland 2012; Skirbekk 2013b), mens to innlegg for et redusert antall karakterer i norsk hevder at dette vil styrke sidemålets stilling (Eide 2012; Titlestad 2010).

Høringsutkastet til læreplan i norsk ble sendt fra Utdanningsdirektoratet til høringsinstansene 5. desember 2012 med frist 5. mars 2013 (Skirbekk 2013a: 53).

Direktoratet ba høringsinstansene om å ta stilling til påstanden «Gjeldende vurderingsordninger for standpunktkarakterer bør videreføres.» Blant høringsinstansene var det om lag to tredjedeler som ønsket en revisjon av vurderingsordningene med reduksjon fra tre standpunktkarakterer i norskfaget. Mindretallet som i høringsrunden argumenterte for å opprettholde tre standpunktkarakterer i norsk, bygget i likhet med innlegg referert over på et behov for å støtte og markere den sidestilte status som er tildelt hovedmål og sidemål i opplæringen, og en bekymring for at elevers muntlige kompetanse ville bli nedprioritert til fordel for skriftlig kompetanse. Argumenter som støttet redusert antall karakterer i norsk, viste til at det å skaffe tilstrekkelig vurderingsgrunnlag skaper et prøvepress som går ut over undervisningskvaliteten i faget, og forringer lærernes arbeidssituasjon (Utdanningsdirektoratet 2013a).

Av flertallet som i høringsrunden ønsket en reduksjon av antall standpunktkarakterer i norsk, var det en mindre andel som støttet én karakter i norskfaget. Dette mindretallet viste i sine høringssvar blant annet til at skillet mellom muntlige og skriftlige arbeids- og presentasjonsformer i faget kan fremstå som kunstig, og at én karakter gir bedre tid til læring. Flere av de språkpolitiske organisasjonene som deltok i høringsrunden, var på den andre siden sterkt kritisk til én karakter fordi de fryktet for sidemålets stilling. Argumenter for to standpunktkarakterer i norsk fremkom i høringsrunden ved at norsk er det mest omfattende faget i opplæringen, noe som bør synliggjøres ved to standpunktkarakterer. Flere av høringsinstansene fryktet dessuten at én karakter i norsk vil gi et mindre nyansert bilde av elevenes kompetanse, og at resultatet kan bli mange «middels» karakterer (Utdanningsdirektoratet 2013a). Vi ser her, som i debatten i fagtidsskriftene referert over, at endring av vurderingsordningen i norskfaget har språkpolitiske så vel som fagdidaktiske og fagpolitiske dimensjoner.

2.2.2 Fagfornyelsen: arbeidet med læreplanen etter 2013

I NOU 2014: 7 *Elevenes læring i fremtidens skole* tar utvalget blant annet for seg det de mener er et generelt behov for fagfornyelse i skolen, relatert til læreplaner og vurderingssystemer. Utredningen legger til grunn at mange av skolefagene har et omfattende innhold, noe som forutsetter at undervisningen må legge vekt på breddeforståelse. Dette kan komme i konflikt med elevenes behov for hva utvalget betegner som dybdelæring (NOU 2014: 7: 11).

Utvalget tar for seg gjeldende læreplan i skolefagene, herunder også norsk. Utvalget konkluderer på bakgrunn av en historisk sammenligning at denne planen er tydeligere innrettet mot forberedelse til aktiv deltakelse i samfunnet enn det bredere mandat for kultur-, dannelses- og identitetsutvikling som preget tilsvarende

plan mot slutten av 1900-tallet. Med hensyn til bredden i norskfaget skriver utvalget:

Norskfaget har et omfattende formål og har mange timer i skolen og favner av den grunn bredt. Prinsippet om at begge målformer skal brukes i arbeidet for å nå kompetansemålene, bidrar også til at bredde ser ut til å prioriteres fremfor dybde. Riktignok ble bredden i norskplanen redusert da læreplanen ble revidert i 2013, særlig når det gjaldt mål som grenser opp mot andre fagdisipliner, for eksempel mediekunnskap, kunsthistorie eller samfunnskunnskap, eller mål som er særlig tid- eller ressurskrevende. Likevel er norsk fortsatt et omfattende fellesfag, og spenningen mellom bredden i emner og ønsket om fordypning i emnene er et tilbakevendende debatttema.

(NOU 2014: 7, s. 77)

I en sammenligning med opplæringen i morsmålsfag (nasjonalspråket) i Sverige, Danmark og Finland slår utvalget fast at når morsmålsfag skal kombinere det å være et redskapsfag og et kultur-, litteratur- og dannelsesfag, blir dette faget omfattende og får et bredt ansvar i skolen. Norskfaget, tilføyer utvalget, preges i tillegg av to sidestilte målformer (NOU 2014: 7, s. 78). Samlet sett kan dette ha som konsekvens, fremholdes det i utredningen, at det blir utfordrende å ivareta dybdelæring og god progresjon i det elevene skal lære i dybden (s. 82).

Utdanningsdirektoratet har arbeidet med å kartlegge og vurdere nytteverdien av læreplanenes kjennetegn på måloppnåelse. Vi omtaler dette arbeidet senere i kapittelet, i forlengelse av OECDs råd til norske myndigheter i utvikling av vurderingsarbeidet.

2.3 Forskrift om individuell vurdering og rundskriv om individuell vurdering

I forskrift til opplæringsloven kapittel 3 «Individuell vurdering i grunnskolen og i vidaregåande opplæring»⁶ slås det fast at elever i offentlig grunnskoleopplæring og elever, lærlinger og lære kandidater i offentlig vidaregåande opplæring har rett til vurdering. Formålet med vurdering i fag er ifølge forskriften «å fremje læring undervegs og uttrykke kompetansen til eleven, lærlingen og lære kandidaten undervegs og ved avslutninga av opplæringa i faget. Vurderinga skal gi god tilbakemelding og rettleiing til elevane, lærlingane og lære kandidatane.» I forskriften gjøres det videre klart at eleven skal være kjent med målene for opplæringen og hva som vil bli vektlagt i vurderingen av elevens kompetanse. Grunnlaget for vurdering er kompetansemålene for faget slik disse er fastsatt i læreplanverket, og det er skoleeier som

⁶ Forskrift til opplæringslova kapittel 3. Individuell vurdering i grunnskolen og i vidaregåande opplæring: https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_4#KAPITTEL_4

har ansvaret for at elevens rett til vurdering blir oppfylt. Med hensyn til vurdering med karakterer skal det i grunnskolen til og med 7. årstrinn bare gis vurdering uten karakterer. Fra 8. årstrinn og i videregående opplæring skal vurdering også gis med tallkarakter i en skala fra 1-6.

I kapittel 1 viste vi til forskriftens beskrivelse av hensikten med vurdering og forskjellen mellom underveisvurdering og sluttvurdering. Læreres halvårsvurdering av elevene inngår i underveisvurderingen. Fra og med 8. årstrinn skal elevene ha halvårsvurdering i form av karakter.

Vurderingsforskriften i sin helhet ble sist endret 1. juli 2009 og trådte i kraft fra 1. august samme år. Krav til at underveisvurdering skal fremme elevens læring i faget, ble utdypet ved denne endringen (NOU 2014: 7, s. 74).

Utdanningsdirektoratet utga opprinnelig også Rundskriv om individuell vurdering (Udir-1-2010), et fyldig dokument med tolkninger av paragrafene i Vurderingsforskriften. Rundskrivet ble gjort ugyldig 11. august 2015, og vi omtaler derfor ikke rundskrivets innhold i denne rapporten. Et nytt rundskriv, «Individuell vurdering Udir-5-2016» (Utdanningsdirektoratet 2016), er gyldig når denne slutt-rapporten skrives.⁷ Dette rundskrivet gir likeledes utfyllende tolkning av Vurderingsforskriften. I tillegg til at rundskrivet forklarer forskjellen mellom underveisvurdering og sluttvurdering og mellom halvårsvurdering og standpunkt-karakter (se kapittel 1 i denne rapporten), viser vi her til hvordan rundskrivet også presiserer at elevene skal få vurdering *uten karakter* gjennom hele opplæringsløpet. I ungdomsskolen og i videregående opplæring kommer dette i tillegg til halvårsvurderinger og sluttvurderinger med karakter. Læreren må altså gi eleven karakter ved halvårsvurdering og sluttvurdering med karakter, men utover dette er det ikke krav om at læreren setter karakter på de vurderinger som hun gjør av elevens arbeid underveis i opplæringsløpet. Denne presiseringen har relevans for de forskningsdataene og erfaringene som vi presenterer fra forsøksskolene senere i denne rapporten.

Uavhengig av om læreren velger å bruke tallkarakter for å gi eleven underveisvurdering utenom vurderingen som skal skje hvert halvår, er formålet med all underveisvurdering altså å bidra til elevens læring, gi et grunnlag for tilpasset opplæring og bidra til at eleven øker kompetansen i faget. I rundskrivet står det at underveisvurderingen skal være løpende og systematisk, som en kontinuerlig og integrert del av opplæringen (Utdanningsdirektoratet 2016). Rundskrivet omtaler også elevers egenvurdering som del av underveisvurderingen. Eleven skal delta

⁷ I en overgang mellom de to rundskrivene eksisterte et dokument kalt «Endringer i regelverket om vurdering» (Utdanningsdirektoratet 2015). Dette dokumentet ga en oversikt over de siste endringene i vurderingsforskriften gjeldende fra 18. juni 2015. Vi brukte dette dokumentet i delrapport 1 i denne evalueringen for å vise til relevante tolkninger av spørsmål i vurderingsarbeidet. Da rundskriv «Individuell vurdering Udir-5-2016» ble offentliggjort, ble dokumentet «Endringer i regelverket om vurdering» fjernet fra Utdanningsdirektoratets nettsider, og er ikke lenger tilgjengelig.

aktivt i vurderingsarbeidet i skolen gjennom å få trening i å vurdere hva som forventes av et bestemt arbeid, vurdere egen kompetanse i arbeidet og vurdere egen faglig utvikling – herunder også hvordan de lærer best

Til sist i dette avsnittet om forskriften skal vi tilbake til veiledningen til læreplanen i norsk (Utdanningsdirektoratet 2013b) og se kort på hvordan lærerne oppfordres til å arbeide med vurdering knyttet til kompetansemål i læreplanen. Dette må ses som en operasjonalisering av læreplanen i lys av regelverket om vurdering, som vi nettopp har referert.

Veiledning til læreplan i norsk gir til sammen 11 eksempler på hvordan elevene kan arbeide med grunnleggende ferdigheter i faget, på måter som kombinerer kompetansemål fra hovedområdene i læreplanen. I veiledningen oppfordres læreren til å utforme læringsmål i samarbeid med elevene og selv utvikle kjennetegn på måloppnåelse der man mener det er nødvendig. Vurdering utledes i eksemplene via lærerens refleksjon over hva elevene skal lære og hvordan kompetansemål og hovedområder kan gjenspeiles i undervisningen, hvordan dette skal formidles til elevene, hvordan elevene skal motiveres for læring og hvilke metoder som skal brukes. Her skal læreren også i planleggingen reflektere over hva slags sluttprodukt opplegget skal gi som resultat. På dette grunnlaget må læreren bestemme hvordan man best kan vurdere hvorvidt elevene har lært noe, og hva de har lært. Denne fremgangsmåten illustrerer forarbeidet til selve vurderingen, som skal bygges over svar på følgende spørsmål som læreren stiller seg selv:

- *Hvordan og av hvem får eleven læringsfremmende tilbakemeldinger underveis i læringsarbeidet?*
- *Hvordan formulerer vi eventuelle kjennetegn på måloppnåelse for dette læringsarbeidet?*
- *Hvordan viser sluttproduktet at målene er nådd?*
- *Hvordan oppsummerer og evaluerer lærer og elever opplegget og elevenes læringsutbytte?*

Utdanningsdirektoratet (2013b: 3-4)

Tveit (2013) setter vurderingsforskriften og det følgende arbeidet med å utvikle regelverket i kontekst når han hevder at arbeidet med å styrke læreres vurderingskompetanse var kommet kort ved innføringen av Kunnskapsløftet og kompetansemål i læreplanene. Dette setter han i forbindelse med forutgående tiårs holdnings- og verdidebatt i skolen, hvor han mener at forventninger til elevers læringsutbytte ble nedprioritert til fordel for det han betegner som et holistisk perspektiv på læring og utdanning (Tveit 2013: 230). Da Kunnskapsløftet ga lærerne økt metodefrihet regulert av overordnede kompetansemål i bytte mot dokumenterte læringsresultater, skjedde dette i en situasjon som ifølge Tveit var preget av manglende forståelse av regelverket for vurdering, svak vurderingskompetanse blant lærerne og lærerutdannerne, og lite forskning på feltet (Tveit 2013: 223, se også St.meld.

nr. 16 (2006-07)). Tiltak fra utdanningsmyndighetene i dette feltet har, i tillegg til å utvikle regelverket om vurdering, vært satsingen Bedre vurderingspraksis fulgt av satsingen Vurdering for læring samt økt satsing på forskning om vurdering i norsk skole. Vi skal se på innsatsen for å endre vurderingskulturen i skolen nedenfor, mens norsk forskning om vurdering etter Kunnskapsløftet ble behandlet i kapittel 1.

2.3.1 Bedre vurderingspraksis og Vurdering for læring

Satsingen Bedre vurderingspraksis ble administrert av Utdanningsdirektoratet i perioden 2007-2009, og omfattet 77 skoler. Fra Utdanningsdirektoratet har vi fått opplyst at hensikten med Bedre vurderingspraksis var å prøve ut nasjonale kjennetegn på måloppnåelse i fire fag over en periode på to år. Det ble prøvd ut flere modeller, hvorav to gikk ut på å prøve ut allerede utviklede kjennetegn, mens en modell gikk ut på at skolene selv skulle utvikle slike kjennetegn. Basert på disse erfaringene initierte Utdanningsdirektoratet satsingen Vurdering for læring med oppstart i 2010 (Hopfenbeck, Tolo, Florez og El Masri 2013: 28). Satsingen tar utgangspunkt i fire prinsipper for læring som formuleres på følgende måte av disse forskerne:

- Elevene må forstå hva de skal lære og hva som er forventet av dem
- Elevene må få tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonen
- Elevene må få råd om hvordan de skal forbedre seg
- Elevene må involveres i eget læringsarbeid ved blant annet å vurdere eget arbeid og utvikling

Satsingen Vurdering for læring bygger på en stor kunnskapsbase i det internasjonale forskningsfeltet *assessment for learning*. Begrepet understreker det formative aspektet ved vurdering. Vi skal ikke referere disse enkeltarbeidene her, men viser til «Teoretisk bakgrunnsdokument for arbeid med vurdering for læring på ungdomstrinnet», utgitt av Nasjonalt kompetansemiljø i vurdering (2012). Her skriver forfatterne at *assessment for learning*-feltet ikke utgjør noen avgrenset teori, men bygger på en rekke elementer som danner basis for formativ elevvurdering i pedagogikken. De understreker også at formativ praksis ikke må oppfattes som utførelse av på forhånd beskrevne metoder, men som verktøy som må tilpasses lærerpersonlighet, elevenes aldersgrupper, det formelle vurderingssystemet og læringens kontekst, slik at denne praksisen er i kontinuerlig utvikling. Denne utviklingen kan ikke drives av enkelte ildsjeler, men krever «en utviklingsorientert skoleledelse og et endringsvillig kollegium» (Nasjonalt kompetansemiljø i vurdering 2012: 6-7).

I Norge har skoler og skoleeiere selv fått bestemme hvordan deltakelse i satsingen Vurdering for læring skal utformes og implementeres i skolene, med felles målsetting om å bruke satsingen til å videreutvikle skolenes vurderingskultur og derigjennom forbedre elevenes læringsresultater. Lokal implementering involverer alle ledd i skoleforvaltningen og skoleorganisasjonen, og med nettverk mellom skoler organisert av skoleeier. Nettverkene skal legge til rette for faglig utvikling, erfaringsdeling og refleksjon basert på praksiseksempler som knyttes til teori og foreliggende forskning. Utdanningsdirektoratet støtter lokal implementering gjennom nasjonale konferanser, nettbaserte ressurser for lærerne og skolelederne og økonomiske tilskudd til skoleeierne (Hopfenbeck m.fl. 2013: 31).

Tveit (2013: 231) mener Vurdering for læring representerer et skifte i norsk vurderingstradisjon, med fremheving og vektlegging av formativ vurdering. Sandvik (2014: 46) skriver at Vurdering for læring kan forstås som en dynamisk, interaktiv og læringsfokusert tilnærming til formativ vurdering, der elevene skal engasjeres i egen læringsprosess gjennom forståelse av læringsmålene, gjennomføre egenvurdering og også vurdere andres ferdigheter. Lærerne på sin side skal hjelpe elevene til å forstå læringsmålene og planlegge undervisningen på basis av dette, gi elevene tilbakemeldinger som er læringsfremmende og oppmuntre elevene til å delta i de nevnte aspektene ved vurderingsarbeidet. Buland (2014: 85) forstår Vurdering for læring som et paradigmeskifte for vurdering, som inneholder endret begrepsbruk og begrepsforståelse basert på et nytt og felles sett av skrevne og uskrevne teorier og metoder med henblikk på å samle profesjonsutøverne i feltet om en bestemt praksis.

2.4 Et internasjonalt perspektiv på vurdering i norsk skole

Til slutt i dette kapittelet skal vi se kort på OECDs Norges-rapport i serien *Reviews of evaluation and Assessment for Improving School Outcomes* (Nusche et al. 2011). Denne rapporten ble skrevet mens den første læreplanen i norsk under Kunnskapsløftet ble revidert, og læreplanen fra 2013, veiledning til læreplanen og rundskriv om individuell vurdering (Udir-5-2016, se over) inkorporerer mange av OECDs anbefalinger. Det er likevel trekk ved det norske vurderingssystemet som nevnes i rapporten, som det kan være relevant å vise til i denne gjennomgangen av systemer som former vurdering i norskfaget. OECDs målsetting med rapporten var å gi en oversikt over gjeldende systemer, vurdere styrker og svakheter ved systemene og gi råd om forbedring.

Slik gir OECD ros til norske myndigheter for en gjennomgripende utvikling av det samlede vurderingssystemet i opplæringssektoren over en kort tidsperiode (oppstart ved St. meld. nr. 30 (2003-2004)). For å bedre elevvurderingen, som er

et av tre elementer ved det totale nasjonale vurderingssystemet som OECD undersøker, anbefaler imidlertid forfatterne at norske myndigheter fortsetter arbeidet med å spesifisere kompetansemålene som skal være grunnlaget for vurdering. Videre råder OECD norske myndigheter til å utvikle læringsmål og vurderingskriterier i enkeltfag, samt støtte og veilede kommuner og skoler i arbeidet med å bearbeide kompetansemålene i det lokale læreplanarbeidet:

Many teachers find it difficult to translate these competence aims into concrete lesson plans, objectives and assessment activities. The broad competence goals have the advantage of giving teachers ownership in establishing their teaching programme, but there seems to be a need for more structure for a substantial number of teachers.

Nusche et al. (2011: 52)

OECD tilføyer at læringsmålene bør spesifisere sammenhengen mellom læreplan, undervisning og vurdering. Veiledning til læreplan i norsk (Utdanningsdirektoratet 2013b, se tidligere i dette kapittelet) gir til sammen 11 eksempler på hvordan kompetansemål på ulike stadier i opplæringsløpet kan bearbeides via læringsmål til metodebruk og vurderingskriterier. OECD etterlyser like fullt nasjonale vurderingskriterier, noe veiledning til læreplanen ikke gir. OECD fremholder at det å arbeide med å gi elevene vurdering underveis i opplæringen kan synes enkelt for norske lærere, fordi dette har mange fellestrekk med hva de «alltid» har vært vant til å gjøre og hva de forbinder med lærerprofesjonen. Formativ vurdering forstått som vurdering for læring er likevel noe annet enn hva mange norske lærere hittil har praktisert når de har gitt elevene tilbakemeldinger, skriver forfatterne, og derfor kan det likevel vise seg vanskelig for disse lærerne å tilegne seg slik tenkemåte og praksis. OECD mener nemlig at underveisvurdering i norske klasserom har hatt elementer av sluttvurdering, idet resultater fra mindre prøver har blitt samlet opp og påvirket sluttkarakteren i faget:

Teachers' classroom assessments were frequently used to track students' progress and provide practice for a final summative assessment. Similarly, self-assessment was often understood in a framework for self-marking, not reflection on learning. This represents a risk that formative assessment is being accepted by teachers as just another name for what they already do. It may reinforce the use of routine assessment in schools as preparation for more summative assessment.

Nusche et al. (2011: 56)

Kontinuerlig elevvurdering, fremholdes det i rapporten, kan inkludere både formative og summative aspekter, men disse kategoriene representerer fundamental ulike hensikter (ibid.). Mens norske myndigheter kan endre regelverk og læreplan og engasjere skoler i satsinger som Vurdering for læring, har myndighetene reelt

sett liten mulighet til å påvirke lærernes praksis hvis lærerne selv ikke forstår eller ikke ønsker å endre sin vurderingsmåte. Slik kan ovenstående empiriske beskrivelse fra OECD vise seg å ha gyldighet i norske klasserom også i dag.

Mens delrapport 1 i evalueringen av forsøk med halvårsvurdering med én eller to karakterer i norsk ble skrevet, arbeidet Utdanningsdirektoratet med å kartlegge og vurdere nytteverdien av læreplanenes kjennetegn på måloppnåelse. Dette arbeidet ble utført på bakgrunn av oppdragsbrev fra Kunnskapsdepartementet, som ba direktoratet vurdere om det bør innføres nasjonale standarder/kriterier for vurdering. Direktoratet innhentet synspunkter fra fylkesmannsembetene, skoleeiere, skoleledere og en referansegruppe bestående av ungdomsskolelærere, skoleledere, representanter for skoleeiere, representanter for organisasjonene og forskere.

I svarbrevet til Kunnskapsdepartementet⁸ skriver Utdanningsdirektoratet at kjennetegnene på måloppnåelse er godt kjent, men brukes ulikt. Flertallet av fylkesmennene, skolelederne, skoleeierne og medlemmene av referansegruppen ønsker at kjennetegnene skal forbli veiledende fremfor obligatoriske, men ønsker samtidig at bruk av kjennetegnene får ytterligere støtte gjennom veiledningsmaterieell. Utdanningsdirektoratets drøfting konkluderer likeledes med at kjennetegn på måloppnåelse bør forbli veiledende fremfor obligatoriske, men at det er behov for å utvikle eksempler på hvordan kjennetegnene kan brukes. Avslutningsvis slår direktoratet fast at selv om kjennetegn på måloppnåelse er godt kjent, er disse i seg selv utilstrekkelige for å sikre likeverdig og rettferdig standpunktvurdering av elever. Direktoratet foreslår derfor en egen satsing på standpunktvurdering i norsk skole, som inkluderer læreplanforståelse og lokalt arbeid med læreplaner.

2.5 Nynorskens stilling i norskfaget

Det finnes et stort tilfang av litteratur om bakgrunnen for sidestilling mellom nynorsk og bokmål i norsk språkpolitikk. Siden vi i denne evalueringen avgrenser litteraturutvalget til tiden etter 2005 og vi i dette kapittelet tar for oss statlige dokumenter som styrer norskfaget, skal vi her kun kort vise til den siste stortingsmeldingen om norsk språkpolitikk. Dette gjør vi med bakgrunn i læreplanen for norskfaget, som fastslår at elevene skal få kjennskap til språklig mangfold, og lære å lese og skrive både på bokmål og nynorsk.

Stortingsmelding nr. 35 (2007-2008) *Mål og mening – ein heilskapeleg norsk språkpolitikk* ble avgitt fra Kultur- og kirkedepartementet under regjeringen Stoltenberg II. I regjeringserklæringen ble det slått fast at siden nynorsk har vanskeligere rammevilkår enn bokmål, er det særlig viktig å sikre en god utvikling også for

⁸ Svar på oppdragsbrev 18-14 punkt 3, kartlegge og vurdere nytteverdien av kjennetegn på måloppnåelse. Brev fra Utdanningsdirektoratet til Kunnskapsdepartementet 30. september 2015.

nynorsk. I meldingen omtaler departementet sidemålsundervisningen i grunnskolen som et av to sentrale virkemidler for sidestillingen av nynorsk og bokmål. Det andre virkemiddelet er reglene for bruk av begge målformer i offentlig forvaltning. De to virkemidlene har form av plikter i språkopplæringen og forvaltningen av det norske språket, og støtter hverandre i det departementet i meldingen betegner som norskspråklig grunnkompetanse. Departementet skriver:

Som verkemiddel i norsk språkpolitikk har dei lang tradisjon, og som grunnleggjande prinsipp er det viktig at dei blir liggjande fast også for framtida. Nødvendige unntak eller tilpassingar må utformast og praktiseras slik at dei ikkje rokkar ved sjølve det prinsipielle utgangspunktet.

St.meld. nr. 35 (2007-08), s. 197

I meldingen rammes dette inn av tiltak for et mer systematisk arbeid med tanke på å styrke nynorsk språk og den nynorske skriftkulturen i kulturlivet, i medier, i universitets- og høyskolesystemet og i forvaltningen.

2.6 Oppsummering

I dette kapittelet har vi tatt for oss policydokumenter og -satsinger som bestemmer og påvirker vurdering i norskfaget. Innledningsvis har vi gjort rede for noen hovedtrekk ved utdanningsreformen Kunnskapsløftet, som innførte nye læreplaner i alle fag. Disse læreplanene er bygget opp på en måte som skiller seg fra den tidligere norske læreplantradisjonen, med kompetansemål for hva elevene skal mestre på bestemte årstrinn i opplæringen og uten nærmere angivelse av innholdet i opplæringen. Kunnskapsløftet ga i utgangspunktet skolen og lærerne økt frihet til å bestemme lokale læringsmål og utforme undervisningen i tråd med dette. I ettertid kan det hevdes at friheten har blitt strammet inn, mens krav til skolens innsats og lærernes ansvar for elevenes læringsresultater er opprettholdt (Aasen m.fl. 2012).

Gjeldende læreplan i norskfaget oppgir dermed kompetansemål for bestemte (ikke alle) årstrinn i grunnopplæringen. For elever i ungdomsskolen oppgis det for eksempel bare kompetansemål for 10. årstrinn. Dette betyr at lærerne må omsette disse kompetansemålene til læringsmål for opplæringen også på 8. og 9. årstrinn gjennom lokalt arbeid med læreplanene. Veiledning til læreplanen gir eksempler på hvordan kompetansemål kan omformes til læringsmål på lokalt plan.

Vi har i dette kapittelet også referert fra debatten som fulgte arbeidet med revisjon av den første læreplanen i norsk under Kunnskapsløftet. Denne debatten tok utgangspunkt i at mengden lærestoff i norskfaget måtte reduseres. I debatten ble det pekt på at faget mangler tilstrekkelig avgrensning mot andre skolefag, som historie og samfunnsfag. Vekting og prioritering av sidemål kontra hovedmål ble også diskutert i lys av såkalt «slanking» av norskfaget. Etter at gjeldende læreplan i

norsk ble innført i august 2013, behandlet utvalget bak NOU 2014: 7 et overordnet behov for fagfornyelse i skolen relatert til læreplaner og vurderingssystemer. Utvalget mente at et omfattende innhold i fag, herunder også norskfaget etter læreplanen av 2013, kommer i konflikt med det utvalget betegner som dybdelæring. Norskfaget fremstår dermed fremdeles som omfattende etter læreplanrevisjonen i 2013.

I dag er det to dokument som angir regelverket for individuell vurdering i grunnopplæringen, nemlig Forskrift til opplæringslova kapittel 3 Individuell vurdering i grunnskolen og vidaregåande opplæring og rundskrivet Udir-5-2016. Dette rundskrivet hviler på et tidligere rundskriv fra 2010 og et dokument kalt «Endringer i regelverket om vurdering». Denne utviklingen av regelverket om vurdering må ses i sammenheng med de nasjonale satsingene Bedre vurderingspraksis (2007-2009) og Vurdering for læring (2010-). Disse satsingene bygger på et internasjonalt teoretisk felt om betydningen av formativ vurderingspraksis, som settes i kontrast til at det å vurdere kan oppfattes som en bestemt metode. Lærernes vurderingspraksis må ifølge dette teoretiske grunnlaget være i kontinuerlig utvikling, og den må forankres i lærerkollegiet.

OECD knyttet seg i 2011 tett til Vurdering for læring i en rapport om vurderingssystemer i norsk skole. Her spurte forfatterne om norske lærere virkelig forstår og anvender formativ vurdering, ettersom de hadde erfart at mye formativ vurdering i norsk skole i realiteten var lærerens kontroll av elevers kompetanseoppnåelse underveis i opplæringen, det vil si et element av summativ vurdering. OECD anbefalte videre at norske myndigheter fortsatte arbeidet med å spesifisere kompetansemål, utarbeide nasjonale læringsmål og vurderingskriterier. Etter at OECD-rapporten ble utgitt, er Vurderingsforskriften så vel som veiledning til læreplanen videreutviklet, men i læreplanen gis det kun et knippe eksempler på hvordan kompetansemål kan omformes til læringsmål og undervisningsopplegg. Ingen av disse dokumentene angir standardiserte vurderingskriterier. Med bakgrunn i en kartlegging og drøfting av ønsker og behov for standardiserte vurderingskriterier gjennomført i 2015, anbefaler Utdanningsdirektoratet at kjennetegn på måloppnåelse i læreplanen fortsatt blir veiledende fremfor obligatoriske.

3 Undervisning og vurdering i norskfaget

Litteraturgjennomgangen som presenteres i dette kapittelet ble opprinnelig trykt som kapittel 4 og 5 i den første delrapporten i denne evalueringen (Seland m.fl. 2015). Her er de to opprinnelige kapitlene slått sammen, redigert og noe forkortet for å gi en oversikt over litteratur som forteller om undervisning og vurdering i faget. Fremgangsmåten og resultatene for våre litteratursøk finnes i sin helhet i delrapport 1 (Seland m.fl. 2015), og er også kort referert nedenfor.

Etter dette tilbakeblikket på metode for litteratursøkene innleder vi kapittelet med en kort oversikt over hvordan lærer- og elevrollen i norskfaget kan sies å ha utviklet seg fra Normalplanen i 1939. Denne teksten ble opprinnelig trykt i kapittel 3 i Seland m.fl. (2015), og finnes også i en forkortet utgave i innledningskapittelet i delrapport 2 fra denne evalueringen (Seland m.fl. 2016).

3.1 Tilbakeblikk på litteratursøk i evalueringen

I evalueringens første delrapport (Seland m.fl. 2015) gjennomførte vi en kartlegging av litteratur på feltet. Et hovedfunn fra dette arbeidet var at litteraturfeltet for norskfaget og vurdering i norskfaget fremsto som hovedsakelig fagdidaktisk basert og med utspring i en læreboktradisjon, fremfor artikler i fagfellevurderte tidsskrifter. I tillegg fant vi flere forskningsrapporter initiert av nasjonale utdanningsmyndigheter, enkelte mastergradsoppgaver og ph.d.-avhandlinger.

I delrapport 1 gjorde vi rede for at litteraturstudien ikke utgjorde en *systematisk gjennomgang* av feltet. En systematisk litteraturgjennomgang er en metode med regler for søk, utvelgelse og sammenfatning av forskningslitteratur med hensyn til funn av hovedsakelig kvantitativ karakter og med vekt på statistisk etterprøvable effekt av innføring av et gitt tiltak. Vår litteraturgjennomgang har heller ikke siktet mot et slik ideal. Vi viste også til at dersom man i tråd med retningslinjer for *systematic reviews* gjorde litteratursøk i databaser, ville resultatene være så få at det knapt ville være mulig å gi noen faglig basert bakgrunn for å forstå og diskutere vurdering i norskfaget.

Vår tilnærming ble derfor å bruke forskningsbibliotekarens og forskergruppens samlede kunnskap for å favne så bredt som mulig i feltet, og finne frem til arbeider som på best mulig måte kunne dekke de problemstillingene vi var ute etter å besvare. Vi var i dette inspirert av metoder for *konfigurativ review*, som kan beskrives som en iterativ tilnærming som søker metning av perspektiv fremfor å være uttømmende. Dette innebar å samle inn og legge frem ulike perspektiver og funn presentert som en «mosaikk» fremfor å syntetisere dem aggregert (Gough m.fl. 2012). I dette har vi brukt internettbaserte søk, men vel så viktig og fruktbart har det vært å gjennomgå litteraturlister i artikler og bøker, delvis som resultat av slike innledende søk.

Søk på stikkordet «norskfag» ga først og fremst treff i fagdidaktisk litteratur og til en viss grad i publikasjoner som gir leserne rom for debatt, som tidsskriftet *Norsklæreren* (medlemsblad for Landslaget for norskundervisning). Vi fikk også treff på noen få mastergradsarbeider i pedagogikk og fagdidaktikk. Vi har brukt et utvalg av disse bidragene for å beskrive innholdet i faget og norsklærernes erfaring med undervisningen, men vi mener å ha avdekket et betydelig «hull» i det forskningsfaglige feltet på dette området. Arbeidene vi har hentet inn fra dette feltet er hovedsakelig bøker og særlig antologier, supplert av debattartikler fra fagblader. Disse debattartiklene er uten unntak normative, sjelden forskningsbaserte, og enkelte av dem også i en polemisk stil. Vi har likevel valgt å inkludere et utvalg slike artikler i vår gjennomgang. Artiklene er valgt ut etter relevans for stikkordene «læreplandebatt» eller «læreplanrevisjon», «arbeidsbyrde», «dannelsesfag», «redskapsfag», «grunnleggende ferdigheter» og «vurdering». Vi har også inkludert artikler fra *Norsklæreren* og lignende tidsskrifter (ikke fagfellelvurderte) som diskuterer balansen mellom hovedmål og sidemål i faget. Særlig i spørsmålet som gjelder sidemål har det vært vanskelig å få treff i litteratursøk, og treffene vi fikk, var primært av politisk og ideologisk karakter som innlegg i læreplandebatt.

Når det gjelder våre søk etter forskning om undervisning og undervisningspraksis i norskfaget spesielt gjorde vi et søk i basen Oria med stikkord «Undervisning i norskfaget» avgrenset til perioden 2005-2015. Resultatet ga 55 treff, hvorav sju av disse viste seg å være empiriske arbeider i form av seks mastergradsoppgaver, en ph.d.-avhandling og en bok – valgte vi ut til nærmere gjennomlesing. I tillegg gjorde vi søk i databasen SkrivBIB, og valgte ut to fagartikler i bladet *Norsklæreren* samt en mastergradsoppgave som kan belyse skriveundervisning. Dette er resultat av en bevisst utvalgsstrategi for å belyse evalueringens overordnede problemstilling om hvordan undervisning kan organiseres i et regime med tre kontra redusert antall karakterer i norsk. Arbeidene er ikke valgt ut eller presentert med henblikk på å gi en representativ eller uttømmende beskrivelse av undervisningspraksiser i norsk.

Vår litteraturgjennomgang kan derfor karakteriseres som utforskende og assosiativ i et felt som synes svært oppdelt i enkeltarbeider og –forskningsspørsmål, og

som ikke er preget av et avgrenset sett med teoretiske tradisjoner. Vi understreker at dette gjelder søk på «norskfag», «norsklærer», «dannelsesfag» og «ferdighetsfag» og ikke på stikkordene «undervisning» og «vurdering». Disse stikkordene leder til et felt som absolutt er definert av teoretiske tradisjoner, som til en viss grad kan sies å strukturere de ulike arbeidene.

3.2 Endringer i norskfagets innhold, lærer- og elevrollene

I delrapport 1 i denne evalueringen (Seland m.fl. 2015) refererte vi til hvordan Moslet (2009) karakteriserer og kategoriserer norskundervisningen, lærerrollen og elevrollen i ulike epoker i etterkrigstiden. Moslets betegnelser på disse epokene er «vakkernorsken» (1950-tallet), «praktisknorsken» (1960-tallet), «kritisknorsken» (1980-tallet), «kreativnorsken» (1980-tallet) og det Moslet kaller for «her-og-nå-norsken» (1990-tallet). Moslet (2009) beskriver hvordan det har foregått en tydelig dreining fra 50-tallets autoritative og kulturarv-orienterte norsklærer som var forventet å forme og øve opp elevene, til læreren som samtalepartner, regissør og iscenesetter for de prosessene av samhandling, samarbeid og samskaping som han mener kjennetegner faget i nyere tid. Elevrollen har også endret seg betydelig. I de første tiårene ble eleven betraktet som mottaker av kunnskap og et formbart objekt, mens de siste tiårene blir gradvis mer preget av et elevsyn der eleven er en meningsberettiget bidragsyter og skapende deltaker. Moslet (2009) beskriver også hvordan undervisningen i faget har utviklet seg over tid fra å være preget av hierarki og lærerens monolog til hva kan kaller for «et dialogisk klasserom.» Som bakgrunn for presentasjonen av studier av undervisningspraksis i norskfaget etter Kunnskapsløftet i 2006 refererer vi her Moslets (2009) beskrivelser av norskundervisningen på 1980-tallet og 1990-tallet.

Kreativnorsken, som Moslet (2009) bruker som betegnelse på 1980-tallets norskfag, gjorde seg gjeldende med et elevsyn hvor barnets egenverd ble fremhevet, et språksyn som betonte sammenhenger mellom språk og erfaring og språk og handling samt et læringssyn hvor egenaktivitet og skapende virksomhet hos elevene ble verdsatt. To metoder står sentralt i denne tiden: lesing på talens grunn (LTG) og «kreativ skriving». Moslet (2009) mener det særlig er på den skrivepedagogiske siden at «kreativnorsken» gjør seg gjeldene i 80-årene. Vi skal ikke gjengi beskrivelsene av disse to metodene, men det er interessant at i denne tiden ble det gjort banebrytende og samtidig omstridte forsøk med skriving på dialekt, i tråd med fremhevingen av sammenhengen mellom språk og erfaring og elevaktivitet.

Sammenlignet med 1970-tallet, da det ble lagt en hovedvekt på sakprosa, ble sjangerrepertoaret utvidet i 1980-årene. Kreativ skriving innebar personlige, skjønnlitterære skrivemønstre med rom for subjektiv innlevelse, følelser, fabulering og fantasi, med utløp for barnas eget «autentiske» barnespråk, skriver Moslet

(2009). Om læreren hevder han at lærere omdefinerte rollen slik at de ble mottakere av innhold eller budskap. Han beskriver også lærerrollen som «språklig-kulturell støttekontakt, med ansvar for stimulering, mild rettleiing og terapi» (ibid.: 24). Kreativ skriving hadde mye til felles med det en senere har kalt kommunikativ skriving, mener Moslet, og fortsetter: «I sin ytterste konsekvens betyr det at norskfaget innvikles i en 'klasseromsoffentlighet' og at norsktimene rammer inn et 'kulturelt forum' der elevene kan bidra med sin 'stemme'» (ibid.: 32). Dette kommer enda tydeligere frem i 90-årene.

Moslets (2009) uttrykk *her-og-nå-norsken* for 1990-tallet tegnes som kontrast til norsk som det han kaller for et «utholdenhetsfag» i de første tiårene etter annen verdenskrig, hvor elevenes først ville få «belønning» for strevet med lesing og skriving i en ikke nærmere bestemt fremtid. 1990-årenes vektlegging av prosessorientert skriving slo ifølge Moslet (2009) godt an hos elever som var innstilt på behovstilfredsstillelse og selvbekreftelse "her og nå". Læringen ble omfattet av hendelsesrike sosiale prosesser, hvor også samtalen ble kraftig oppgradert som faglig aktivitet. Om lærerrollen heter det nå: «samarbeidsorientert planlegger, organisator av prosesser og prosjekter» (ibid.:24), og: «Læreren er blitt en nærere person, som gruppedeltaker, samtalepartner, konsulent, veileder og ikke minst som regissør av den livs levende, berikende, frustrerende, solidariske, konfliktfylte samhandlinga» (ibid.: 33). Samtidig understreker Moslet (2009) at andre signaler pekte i andre retninger. Med vektlegging av «fellesstoff» og «felles mål» i læreplanen fra 1997 (L97) ser Moslet (2009) en norsklærer som ble gjenreist som tradisjonell kulturbærer og kulturformidler.

Etter denne korte gjennomgangen av utviklingstrekk ved norskfagets innhold og endringer i lærer- og elevrollen i nyere tid kan vi legge til denne korte observasjonen om norskfaget etter Kunnskapsløftet: Steinfeldt (2012a) skrev i en artikkel i *Norsklæreren* at når Kunnskapsløftet peker mot en levende kulturarv som forandres og skapes på nytt, og med oppmuntring til elevene om å være aktive bidragsytere i denne prosessen, fremstår dialogen og samhandlingen mellom lærer og elev i klasserommet som ytterligere aksentuert. Her møtes på et vis 1950-tallets vektlegging av kulturarv og 1980- og 90-tallets arbeidsmetoder i faget underlagt et endret syn på lærer- og elevrollen.

3.3 Undervisningspraksiser i norskfaget

Vi har delt den følgende litteraturgjennomgangen inn i arbeider som forteller om henholdsvis litteraturdelen, skriveopplæringen og sidemålsopplæringen i norskfaget. Muntlige opplæringsmål behandles særlig i studiene av litteraturundervisningen.

3.3.1 Undervisningspraksiser i litteraturdelen av norskfaget

I en ph.d.-avhandling intervjuer Kjelen (2013) i alt 18 norsklærere i ungdomsskolen om deres undervisningspraksis for å analysere deres forståelse av begrepene *dannelse*, *kanon* og *kompetanse* i litteraturundervisningen. Kjelen legger til grunn at Kunnskapsløftet har endret rammene for norsklærernes arbeid med litteraturundervisningen, og gitt lærerne større frihet i for eksempel utvalg av tekster enn hva som var tilfelle i den forutgående læreplanen L97. Kjelen hevder i avhandlingen at norskfaget under Kunnskapsløftet ikke lenger kan forstås som et litteraturfag, men et kultur- og tekstfag. Der litteraturundervisningen tidligere hadde klarere og avgrensede formål i læreplanen, forstått som å utvikle elevenes forståelse for litteraturen og gi dem kjennskap til nasjonallitteraturen, har litteraturundervisningen i dag et langt bredere mandat. Enkeltlærere i Kjelens undersøkelse hevder at litteraturen nå ikke lenger skal leses for sin egen del, men lede frem mot andre kompetansemål. Kjelen finner også en rekke ulike forståelser av norskfaget og av litteraturundervisningens mandat i lærerintervjuene.

Kjelen (2013) kategoriserer tidlig i avhandlingen en av skolene som «læreplanvelvillig» og en annen skole som «læreplanskeptisk». Han diskuterer begrepet skolekultur, og foreslår at dette begrepet kan brukes til å beskrive og forklare den kultur og diskurs (kontinuerlige fagsamtale og felles forståelse) som styrer (norsk)lærernes prioritering av fagstoff og foretrukne arbeidsmåter i undervisningen. Kjelen utvikler ikke skolekulturbegrepet i avhandlingen, men viser til at lærerne har oppfatninger om at de gjør det «på sin måte», mens naboskolen underviser i litteratur på en annen måte. Dette kan gjøre lærerne frustrert. Likevel er de trygge på at deres tolkning av læreplanen er den beste.

Undervisningspraksis blant lærerne på den «læreplanvelvillige» skolen var derfor preget av føringene i Kunnskapsløftet, med sterk tekst- og leserorientering organisert etter tekstenes tematikk, lesestrategier og sammenlignende lesing fremfor kronologisk gjennomgang av litteraturhistorien. På den «læreplanskeptiske» skolen underviste norsklærerne i litteratur på en mer «gammeldags» måte, med vekt på formidling av kulturarven, kronologisk presentasjon av epoker og verk, vektlegging av verkets og forfatterens samtidige kontekst, inkludert forfatterbiografi. På denne skolen arbeidet ikke lærerne med sammenligning av tekster eller tematisk organisering av tekstutvalg. På de to skolene valgte lærerne ut mange av de samme tekstene (romaner og noveller), men deres undervisning i tekstene var preget av stor variasjon.

Det er bare et fåtall av Kjelens 18 lærere som forteller i detalj hvordan de underviser i utvalgte tekster. En av disse lærerne, på den «læreplanvelvillige» skolen, forteller om et undervisningsopplegg med såkalt førlesing av tekst (teksten presenteres for elevene som type eller sjanger, viktige ord og sentrale begreper blir forklart) fulgt av felles gjennomlesing med såkalt «lesestopp» for å stille elevene spørsmål

om hva som skjer i teksten (observasjonsspørsmål). Når elevene på denne måten hadde blitt kjent med teksten, organiserte læreren tablåer eller «frys-situasjoner» i klasserommet der elevene skulle illustrere teksten ved hjelp av positurer. Elevene arbeidet også med teksten i grupper basert på oppgaver og spørsmål fra læreren. Undervisningen i denne særskilte teksten ble avsluttet med «etterlesing» hvor elevene oppsummerte og evaluerte hva de hadde lært. Kjelen (2013) kommenterer at i dette metodemangfoldet kan oppmerksomheten lett gli fra teksten som er gjenstand for undervisningen og over til elevenes egne følelser og reaksjoner.

Ved den «læreplanskeptiske» skolen var det langt vanligere at læreren gjorde en kort «førlesing» som nevnt over, mens elevene brukte tiden på å lese selve tekstene. Lærerne ville deretter gjennom samtale legge til rette for elevenes økte forståelse og diskusjon ved å aktualisere tekstene i elevenes samtid og erfaringshorisont, for eksempel ved å få dem til å reflektere over skikken med arrangerte ekte-skap eller mangel på moderne velferdsordninger i lesingen av 1800-tallets forfattere. Slik blir tekstene gjenstand for en sammenligning, skriver Kjelen, men sammenligning dreier seg om sosialhistoriske spørsmål fremfor tematiske sammenligninger av selve tekstene. Lærernes metodepreferanser var likevel ikke fullt ut konsistente innenfor verken den «læreplanvelvillige» eller den «læreplanskeptiske» skolen. På den førstnevnte skolen fantes det også lærere som helst ville la elevene lese gjennom hele verk og kanskje deretter utfylle lesingen med å vise elevene filmatiseringer av de samme verkene. På den sistnevnte skolen var det også lærere som valgte langt mer moderne litteratur enn sine kolleger, og som søkte å styrke elevenes personlige utvikling gjennom litterære samtaler fremfor å sammenligne sosialhistoriske problemstillinger.

Et gjennomgående trekk hos Kjelen (2013) utvalg av lærere er omtalen av såkalt emnetrengsel i norskfaget. Lærernes metodevalg og prioriteringer i henhold til læreplanen kan oppfattes som en strategi for å navigere i mangfoldet av emner og kompetansemål. Lærerne ved den «læreplanvelvillige» skolen var mest plaget av slik emnetrengsel. Konsekvensen var at de lot elevene lese utdrag av romaner, og de tok seg sjelden tid til å vise hele spillefilmer eller la elevene dramatisere større deler av et skuespill. På den «læreplanskeptiske» skolen fungerte den fag- og fagdidaktiske diskursen ifølge Kjelen (2013) som en slags beskyttelse mot lærernes opplevelse av emnetrengsel. Ettersom lærerne i fellesskap ga høyest prioritert til tekster fra sent 1800- og tidlig 1900-tall, gjorde dette det mulig for dem å gi mindre oppmerksomhet til andre emner i litteraturundervisningen. Disse lærerne kunne dermed bruke mer langsomme og tidkrevende metoder i undervisningen, som visning av spillefilm og lesing gjennom dramatisering. Lærerne på den «læreplanvelvillige» skolen var derimot mer frustrerte over læreplanens manglende prioritering av fagstoff i norsk, og lette foreløpig forgjeves etter noe de kunne kalle

for fagets «kjerne». Denne «kjernen» hadde norsklærerne på den andre skolen funnet frem til gjennom sin diskurs, som på mange måter holdt fast ved forrige læreplan.

Kjelen (2013) viser med avhandlingen at det er et mangfold av undervisningspraksiser i litteraturundervisningen i norskfaget. Inndelingen av skolene i en «læreplanvelvillig» og en «læreplanskeptisk» skole er tentativ, fordi enkeltlæreres praksis kan divergere innenfor samme skole. Samtidig gir denne inndelingen et bilde av en kultur eller felles forståelse som kan prege et lærerkollegium, og som gir retning til lærernes valg. Denne diskursen peker mot idealer for undervisningspraksis i norskfaget, og viser samtidig hvor ulike disse idealene kan være.

En masteroppgave i norskdidaktikk (Lervaag 2010) gir en detaljert fremstilling av et undervisningsopplegg i retorikk brukt i forbindelse med litteraturdelen av norskfaget. Vi refererer dette arbeidet her, fordi det gir mange eksempler på hvordan lærerne integrerer undervisningsvurdering med henblikk både på gi umiddelbar tilbakemelding til elever og karaktersetning i norsk muntlig som halvårsvurdering.

Lervaag (2010) beskriver hvordan hun som masterstudent arbeidet sammen med faglærer i norsk om å forberede og gjennomføre et undervisningsopplegg i retorikk anvendt på muntlige presentasjoner av litteraturtolkning og -sammenligning i en Vg3 klasse på studieforbereende utdanningsprogram i videregående skole. Selve undervisningsopplegget bygde på metoder hentet fra boken *Muntlige tekster i klasserommet* (Penne og Hertzberg 2008), og besto i hovedsak av lærerens forelesninger/innledninger, gruppearbeid, fremføringer av gruppearbeid i plenum, lærerens umiddelbare tilbakemeldinger til elevene og lærerens og masterstudentens etterarbeid med å vurdere hver enkelt elev individuelt med tanke på karakter. I dette arbeidet brukte Lervaag et spørreskjema til elevene (27 elever), et gruppeintervju med representanter fra de fem elevgruppene som ble etablert i forbindelse med undervisningsopplegget, et intervju med faglærer og observasjoner dokumentert med notater og videoopptak.

Lærerens og masterstudentens innledende forelesninger for elevene hadde repetisjon av litteraturanalyse og retorikkens grunnregler som tema. Elevene arbeidet deretter i grupper over flere dager med å analysere og sammenligne tekster fra litteraturdelen av norskfaget. Det går ikke frem av masteroppgaven hvordan tekstene var valgt ut. Av elevfremføringene som refereres av Lervaag (2010), kan vi likevel se at to av gruppene arbeidet med samiske dikt, sagn og eventyr, to av gruppene arbeidet med nynorske forfatterskap, og den siste gruppen arbeidet mer helhetlig med et enkeltstående norsk forfatterskap fra 1990- og 2000-tallet. Elevene brukte powerpoint, musikkinnslag, filmsnutter og delvis kostymer for å illustrere presentasjonene. Presentasjonene ble gjennomført på en fagdag i slutten av undervisningsperioden.

Gjennom masterstudentens (Lervaag) forelesninger om retorikk ble elevene presentert for mulige vurderingskriterier av presentasjonene de selv skulle holde som avslutning av undervisningsopplegget. Elevene diskuterte forslagene til vurderingskriterier, og kom selv med tilføyelser og presiseringer. Elevenes forslag til kriterier ble deretter bearbeidet av Lervaag og faglæreren til et vurderingsskjema som elevene skulle benytte både under gruppearbeidet og som publikum ved elevpresentasjonene. Elevene førte også logg før og etter hver økt med gruppearbeid. Loggen ble lest av lærer, og ga pekepinn om delmål og fremdrift i elevenes bearbeiding av stoffet. Som ledd i forberedelse til å vurdere seg selv og hverandre ble elevene presentert for ulike aspekter ved talerollen og lytterrollen, det vil si å lytte etter informasjon, å lytte kritisk og å lytte empatisk (Penne og Hertzberg 2008). Flere elever ga uttrykk for at nettopp dette var til stor hjelp for dem når de selv skulle innta den saklige vurdererrollen overfor andre.

Under presentasjonene av gruppearbeidene fikk hver gruppe først vurdere sin egen innsats, hvor loggføringen åpenbart støttet elevene i å kunne beskrive og grunngi hvordan arbeidet med stoffet og gruppa hadde utviklet seg over tid. Medelevene som var publikum, brukte vurderingsskjemaene under fremføringene og deltok deretter i tilbakemeldingene. Også faglærer ga elevene umiddelbare tilbakemeldinger basert på skjema med vurderingskriterier. Her ser vi mange uttalte eksempler på Moslets (2009) beskrivelse av læreren som «samarbeidsorientert planlegger, organisator av prosjekter og prosesser» i «læring omfattet av hendelsesrik, sosial samhandling».

Lervaag (2010) har en sluttsats i empirifremstillingen i masteroppgaven som er særlig interessant med tanke på vurdering. Når faglærer og masterstudent setter seg ned med egne vurderingsskjema for hver elev, er de i utgangspunktet ikke enige om hva slags karakterer elevene bør få. Lervaag beskriver denne situasjonen som at faglærer var mest opptatt av hvordan elevene hadde besvart de faktiske litteraturanalyseoppgavene, mens Lervaag selv, som ansvarlig for undervisningsopplegget i retorikk, var opptatt av hvordan elevene hadde brukt denne typen virkemidler under fremføringene. Her ser vi at de to var uenige i hva som er det sentrale innholdet i faget. Leervaag (2010) tilføyer at de to likevel ble enige om karakterene. Vi synes dette gir et interessant innblikk i vurdering av muntlige ferdigheter i norsk, og at to lærere kan bedømme prestasjoner ulikt til tross for at de bruker samme metode for å komme frem til sin vurdering (i dette tilfellet vurderingsskjema med faste kriterier).

Leervag (2010) skriver i begynnelsen av mastergradsoppgaven at vurdering av muntlige ferdigheter blir oppfattet som vanskelig, fordi situasjonen gjør elevene sårbare. Lervaag inkluderte spørsmål i spørreskjema og gruppeintervju for å undersøke elevenes opplevelse av å bli vurdert. I starten av undervisningsopplegget

svarte en del av elevene at de ofte grudde seg til muntlige prestasjoner og fremføringer, og enkelte av dem sa også at det var vanskelig å vite hva prestasjonene burde inneholde. Ved ny spørreskjemaundersøkelse og gruppeintervju med elevene ved avslutningen av undervisningsopplegget mente flere elever at kunnskap om retorikk hadde gitt dem gode redskaper til å strukturere og forme presentasjonene. Flere fortalte også at de hadde kjent seg komfortable under presentasjonene. Lervaaqs observasjonsnotater (gjengitt i masteroppgaven) viser imidlertid at på hver enkelt gruppe var det elever som viste uttrykk for sjenanse og ubehag ved å stå foran medelevene under presentasjonen, og en eller to elever uteble konsekvent fra gruppearbeid og fremføring. Dette kan tolkes som at arbeids- og/eller presentasjons- og vurderingsformen ble for utfordrende for noen av elevene.

3.3.2 Undervisningspraksiser i skriveopplæringen

To bidrag fra *Norsklæreren* gir synspunkter på hvordan undervisningspraksis i skriveopplæringen bør være i henhold til læreplanen i norsk. Kringstad og Kvithyld (2013) forholder seg til 2013-revisjonen av norskplanen når de på et mer overordnet nivå hevder at undervisning i skriving som grunnleggende ferdighet i norskfaget utfordres av manglende begreper om hva faglighet i norskfaget innebærer. De sammenligner med læreplanen i naturfag og denne planens omtale av skriving som grunnleggende ferdighet. Her sier planen at skriveopplæringen skal støtte elevens faglige utvikling, begrepsbruk og forståelse av naturfag. En tilsvarende presisering finnes ikke i norskfaget.

Læreplanen i norsk fra 2013 formulerer tekstenes formål («argumenterende», «fortellende») fremfor tidligere læreplaners omtale av tekster etter sjanger («kåseri», «debattinnlegg»). Kringstad og Kvithyld (2013) fremholder at formålene gir tekstene en funksjon, og at arbeidet for å oppnå dette formålet bør ses som skrivehandlinger. Det er også slik tekstene må vurderes, skriver disse, ikke som mer eller mindre gode uttrykk for bestemte sjangre. Lærerne må følgelig gi elevene formål for skrivehandlingene, og lærerne må skape meningsfylte skrivesituasjoner for elevene. Kringstad og Kvithyld (2013) avslutter med å understreke at undervisningen må bygge på en forståelse av hva som begrunner relevant skriving i norskfaget, og må legge til rette for elevenes bruk av fagbegreper, gjerne gjennom eksempeltekster.

Iversen og Otnes (2011) skriver med bakgrunn i 2010-planen i norsk når de kommenterer forekomsten av betegnelsen «grammatikk» i planen. Dette begrepet er nesten borte fra læreplanen i norsk, skriver disse to, mens det er en rekke henvisninger til «språk» og «tekst». De viser også at læreplanen er sentrert om språkets funksjon, altså å lage sammenheng, å nyansere, distansere, fremheve, fortelle, argumentere og beskrive.

Iversen og Otnes (2011) gir deretter skissemessige innspill til undervisningsopplegg eller -logikker som kan underbygge grammatikkens plass i arbeid med språkets funksjon. Valg av sjanger som fortelling, dramatisering, appeller og instruksjoner preges av ulike verbtider, idet eventyr foregår i fortiden, dramatiseringer bruker presens som virkemiddel og instruksjoner har verb i imperativform (for eksempel «skru av korken»). Iversen og Otnes viser deretter til hvordan nye ytringsformer som SMS, blogg og chat kan bygge opp under forståelsen av grammatikk og setningsbygning. Disse kanalene har en aksentuert funksjon i å kommunisere et kort og presist budskap, og det er ifølge disse forfatterne underordnet at man for eksempel i SMS ikke uttrykker seg i hele setninger.

Med hensyn til vurderingssituasjoner, som for eksempel samtale mellom lærer og elev om et skriftlig elevarbeid, viser Iversen og Otnes (2011) til betydningen av grammatikk for å gi innhold og retning til elevens forståelse av lærerens kommentarer. Ved bruk av grammatikalske grunnbegreper kan lærer og elev diskutere hvordan setninger er bygget opp, og hvordan ordvalg og derigjennom grammatikk underbygger eller tilslører et budskap i setningen.

Et mastergradsarbeid (Såheim 2012) i norskdidaktikk gir en spennende synsvinkel på undervisningspraksis gjennom å undersøke 10. trinnselevers forståelse av lærerens tilbakemelding i såkalt prosesskriving. Prosessorientert skriving som metode fikk innpass i norsk skole i løpet av 1980- og 1990-årene, og består av skriving i bestemte faser, hvor lærer leser teksten og kommenterer arbeidet mellom hver fase. Moslet (2009, se tidligere i dette kapittelet) omtaler prosessorientert skriving som et særlig medium for læreren som gruppedeltaker, samtalepartner og konsulent, med eleven som meningsberettiget bidragsyter og skapende deltaker i 90-tallets «her og nå»-norskfag. Prosesskriving er i litteraturen beskrevet som en helhetlig metode, men i dag vil svært mange norsklærere benytte seg av en variant av prosesskriving uten å ta med alle de foreskrevne fasene, skriver Såheim (2012). Til denne rapportens formål kan vi betrakte prosesskriving som et prisme på sammenhenger mellom lærerens undervisningspraksis og vurderingspraksis i norsk, noe som også behandles i lærerintervjuene som presenteres i et eget kapittel senere i denne rapporten. Gjennom bearbeidingen av lærerens «fremovermeldinger» forventes det at eleven skal utvikle sine skriveferdigheter. På denne måten glir vurdering og læring over i hverandre og aksentuerer undervisning som en gjensidig kommunikasjonshandling mellom lærer og elev.

Forut for selve skriveoppgaven hadde 10.-klassen i Såheims (2012) masteroppgave arbeidet med sjangeren «fortelling» i to uker. Elevene kunne deretter velge mellom tre oppgavetitler, som alle anga en oppdiktet fortelling om en tur, gjerne holdt i en lett og humoristisk tone. Elevene leverte første utkast i den elektroniske læringsplattformen Fronter, og læreren ga skriftlige kommentarer på elevarbeidene ved å bruke den vanlige merknadsfunksjonen i Word-programmet. Elevene

fikk også et ferdig utfylt vurderingsskjema sammen med besvarelsen, basert på kriterier som klassen hadde blitt enige om i fellesskap. (Vurderingsskjemaet var ikke gjenstand for Såheims analyse.)

Såheim (2012) kategoriserer først lærerkommentarene. Disse kan være kommentarer til tekstsegmenter, gjerne plassert i margen, eller de kan være sluttkommentarer som står til slutt i elevens utkast. Kommentarer med tilbakeskuende vurderinger viser til noe som er bra i teksten, og som læreren vil at eleven skal beholde eller videreutvikle i neste utkast. Endringsforslag viser til mangler eller problemer ved teksten. Denne inndelingen henter Såheim fra Igland (2008) *Mens teksten blir til*. Norsk læreren som hadde ansvar for elevene som Såheim intervjuer i oppgaven, var kjennetegnet ved at hun aldri påpekte skrivefeil eller manglende tegnsetting. Hun pekte heller ikke direkte på hva som skulle endres eller hvordan elevene skulle bearbeide teksten i neste innlevering. Hvordan hun arbeidet, går frem i det følgende.

Såheim (2012) intervjuer elevene i perioden etter at lærerne har gitt dem sin første tilbakemelding, og elevene skal gå gjennom besvarelsene på nytt. Såheim mener at de fleste kommentarene blir adekvat oppfattet og forstått av elevene. Særlig gjelder dette rosende kommentarer. Der elevene stusser eller mistolker lærerens kommentarer, kan dette ifølge Såheim bero på ulike forhold. En årsak er at de kan være ukjente med ord som læreren bruker i kommentaren, og derfor ikke forstår innholdet i tilbakemeldingen. Eksempler på slike ord som elevene ikke forsto, var «omstendelig» og «blasert». Elevene ønsket seg også mer konkrete kommentarer når lærer for eksempel skrev «gjentakelse». Elevene skjønte ikke alltid hva det var læreren mente de hadde gjentatt.

Såheim (2012) viser også at lærerens kommentar ofte kan være ordknapp og utformet som et spørsmål, hvor eleven overlates til å tolke hvordan kommentaren kan brukes til å forbedre teksten. Eksempler på slike lærerkommentarer er «Her mangler det kanskje noe?» og «Se på setningen en gang til.» Et par lærermerknader i teksten var også ekspressive men ufullstendige, som når læreren hadde skrevet «???!» etter elevens beskrivelse av hovedpersonens møte med en bjørn under en joggetur. I det siste tilfellet skjønte ikke eleven hva læreren mente med merknaden. Denne læreren utnyttet på denne måten ikke det vurderingsgrepet Iversen og Ottesen (2011, se over) hevder ligger i samtale basert på elevers skriftlige arbeider. Elevenes løsning når de ikke forsto merknaden, var enten å avvise lærerens kommentar eller å si i intervjuet med Såheim (2012) at de ville spørre læreren neste gang de skulle arbeide med teksten på skolen. Såheim (2012) kommenterer at i enkelte tilfeller trodde elevenes at de forsto lærerens kommentarer, mens Såheim tolket kommentaren annerledes. I praksis kan dette bety at noen av de kritiske kommentarene ble misforstått og tolket som lærers godkjenning eller ros. Såheim

avslutter drøftingen ved å slå fast at elevene gjennomgående uttrykte tillit til læreren og hennes vurderinger, og mange av elevenes ytringer i forskningsintervjuet om tilbakemeldingene viste at elevene syntes kommentarene var både rimelige og nyttige.

3.3.3 Undervisningspraksiser i sidemålsopplæringen

Vi viser her til to masteroppgaver som undersøker elev- og lærerholdninger til nynorsk med bakgrunn i blant annet lærernes undervisningspraksis. I tillegg har vi gjennomgått en undersøkelse rapportert av TNS Norsk Gallup (2006) på oppdrag fra Utdanningsdirektoratet. Vårt utvalg av litteratur om undervisningspraksis i sidemålsopplæring er lite, og viser i hovedsak til opplæring i nynorsk som sidemål (unntaket er Gallupundersøkelsen, som også undersøker bokmål som sidemål). Mastergradsarbeidene vi her har vist til, bygger også på svært små kvalitative utvalg, mens Gallupundersøkelsen har samlet data fra 587 lærere til statistisk analyse.

Gallupundersøkelsen fulgte opp situasjonsbeskrivelsen av sidemålsundervisningen i Stortingsmelding nr. 30 (2003-04) *Kultur for læring*, som påpekte et behov for systematisk forsknings- og utviklingsarbeid innenfor nynorskopplæringen. Gallupundersøkelsen (2006: 16) viser at skoler med nynorsk som sidemål brukte om lag en time per uke på slik opplæring, mens skoler med nynorsk som hovedmål brukte om lag en halvtime per uke på sidemålsopplæring i bokmål.

De to masteroppgavene undersøker nynorskundervisningen i henholdsvis to klasser i videregående skoler i Oslo (Rise 2007) og tre klasser i videregående skole (Røed 2010) fordelt på et såkalt bokmålsområde, et «grenseområde» med tradisjon for blandet tekstkultur (både bokmål og nynorsk) og et nynorskområde. Siden begge masteroppgaver er skrevet med henblikk på å undersøke holdninger til nynorsk, er det denne termen og ikke betegnelsen «sidemål» som brukes i disse arbeidene. Begge arbeider er basert på spørreskjemaundersøkelse blant elever og intervjuer med henholdsvis lærere og elever. Her er det intervjumaterialet som har interesse, idet det kan fortelle om læreres undervisningspraksis i nynorsk.

Rise (2007) har intervjuet to norsklærere i Oslo om deres undervisningspraksis i nynorsk. Disse lærerne ble valgt ut til dybdeintervju av Rise fordi elevene ved i alt seks klasser svarte på spørsmål om egen motivasjon og oppfatninger av lærernes motivasjon for nynorsk gjennom en forutgående spørreskjemaundersøkelse. Rises utvalg på to norsklærere representerer derfor den læreren som ifølge elevene har sterkest motivasjon for nynorsk og den læreren som utviser minst motivasjon for å undervise i sidemålet. Begge lærerne forteller imidlertid selv i intervju med Rise at de setter pris på nynorsk og er stolte av målformen.

Rise (2007: 63) skriver at begge disse lærerne synes å legge vekt på tradisjonelle metoder når de underviser i sidemål. Dette utdypes som formlære, grammatikk-innlæring og lesing av tekster. Læreren som hadde de mest motiverte elevene, sa han syntes det var vanskelig å skape situasjoner der sidemålet ble en naturlig del av norskfaget. Han oppfattet formverket for nynorsk som det styrende elementet for opplæringen, og sa at det var krevende å rette elevarbeider fordi nynorsk har så mange valgfrie former. I intervjuet sier denne læreren at han prøver å trekke inn kulturhistoriske perspektiver, legge vekt på det genuint norske og vise til at de fremste norske forfatterne skriver på nynorsk. «Men jeg når liksom ikke fram,» sa denne læreren (Rise 2007: 79).

Norsklæreren som ifølge Rises analyse hadde de minst motiverte elevene i sidemål, sa at hun organiserte undervisningen som forelesinger og lot elevene arbeide med oppgaver. Hun sa videre at hun forsøkte å gi elevene konstruktive tilbakemeldinger i faget, men på grunn av elevenes mange formfeil, returnerte hun elevarbeider med «rødt overalt» (Rise 2007: 81). Rise hevder på bakgrunn av analysen at det ikke er lærernes undervisningspraksis som kan forklare at elevene deres har ulik innstilling til sidemålsundervisning, men snarere lærernes egen innstilling til faget. Vi for vår del vil nyansere dette til at *lærernes tolkning av elevenes innstilling til faget* synes å være styrende for deres holdning til undervisningen. Tolkningen understreker refleksiviteten i kommunikasjonen som foregår i undervisningen.

Røed (2010) intervjuet til sammen 13 elever for å finne ut hva som kjennetegnet nynorskundervisningen i ulike deler av landet med ulik tradisjon for tekstkultur (bokmål og nynorsk). Elevene fra skoler i nynorskområdet og i «grenseområdet» var eksponert for nynorsk gjennom den ordinære norskundervisningen, og skilte i praksis ikke mellom hovedmål og sidemål. Elevene i bokmålsområdet så derimot på nynorsk som noe som kom «i tillegg» til den vanlige norskopplæringen. Ved denne videregående skolen fortalte elevene at undervisningen i nynorsk var preget av repetisjon av læringsstoff fra ungdomsskolen, mye tavleundervisning og oppgaveark med grammatikkoppgaver. Disse elevene ønsket seg bruk av alternative undervisningsopplegg som kunne gjøre nynorsk mer spennende, for eksempel ønsket de at de kunne se film eller oppleve målformen i et mer naturlig miljø (klassetur eller ekskursjon).

Begge masteroppgaver bygger på funn om at undervisningspraksis i nynorsk bygger på tradisjonelle undervisningsmetoder. Kontrasten til eksempelet om gruppearbeid om bruk av retoriske virkemidler i litteraturundervisningen (Lervaag 2010, se over) er stor. På den annen side forteller ikke litteraturen om hvor vanlig slike omfattende, kreative og potensielt spennende og utfordrende prosjekter er i norskundervisningen som helhet. Gallupundersøkelsen (2006) fant ellers at det var små forskjeller i hvordan skoler la opp undervisning i henholdsvis hovedmål

og sidemål, og at stilskriving og oppgaveløsning var de mest brukte arbeidsformene i begge målføre. Samtidig oppga lærerne i Gallupundersøkelsen at de varierte hovedmålsopplæringen med noe flere metoder enn hva de gjorde i sidemålsopplæringen. I hovedmålsopplæringen kunne lærerne bruke mellom fem og sju ulike arbeidsmetoder, mens i sidemålsopplæringen brukte lærerne mellom én og fire arbeidsmetoder.

3.3.4 Sluttsats om vår undersøkelse av litteratur om undervisningspraksis i norskfaget

En av problemstillingene i denne evalueringen ber oss svare på hva som kjenner tegner undervisningspraksis i norskfaget. Vårt utvalg av arbeidere som undersøker undervisningspraksis empirisk, gir ikke noe enhetlig inntrykk av slik praksis. Dette kan naturligvis bero på våre utvalgskriterier eller andre mer skjulte mangler ved selve utvalget, men vi tror ikke dette er årsaken. Snarere trekker vi den slutning at vårt litteraturutvalg gir et bilde av undervisningspraksis i norskfaget som svært mangfoldig, i noen grad samkjørt innenfor den enkelte skole eller mellom kolleger som samarbeider mye, men også utformet og tolket av enkeltlærere i deres oppfatning av hva norskfaget er og skal være og hvordan det best kan gjøres levende for elevene i klasserommet. Det må legges til at det har vært vanskelig å finne relevant litteratur om dette emnet, som har preg av forskning og ikke er lærebokorientert. På denne måten er det betegnende at det er en overvekt av masteroppgaver i vårt litteraturutvalg om dette emnet.

Professor Emeritus Jon Smidt ved Høgskolen i Sør-Trøndelag er sitert i en av disse masteroppgavene (Moe 2013) for beskrivelsen av norskfaget «ikke som ett fag, men som mange fag, hvert klasserom har sitt norskfag.» Slik tror også vi at undervisning i norskfaget ikke lar seg favne av én beskrivelse, men først og fremst må forstås som resultat av ulike vektlegginger, tolkninger, preferanser og idealer holdt oppe og utviklet av norsklærerne selv.

3.4 Vurdering i norskfaget

I det følgende presenterer vi et utvalg norske studier som tar for seg vurdering i norskfaget. Litteraturen som presenteres er et utvalg av akademisk litteratur publisert i fagfellelvurderte tidsskrifter, artikler som ikke er fagfellelvurdert, arbeidsrapporter og mer populærvitenskapelige publikasjoner. Kapitlet tar ikke sikte på å diskutere emnet uttømmende, men er et forsøk på å gi et bredt bilde av litteratur på feltet.

Det finnes bidrag som berører vurdering i norskfaget mer generelt der det gjøres oppsummerende betraktninger (se for eksempel Hertzberg 2008; Hvistendahl

og Roe 2014). Andre omhandler i første rekke vurdering av skriving eller er spesielt knyttet til skriftlig eksamen i hovedmål (Berge m.fl. 2007; Hertzberg 2007; Hertzberg 2008; Fjørtoft 2010). Noen bidrag handler om opplæring i muntlige ferdigheter (Svenkerud m.fl. 2012; Svenkerud 2013). I tillegg finnes det bidrag i *Norsk Pedagogisk Tidsskrift* og flere populærbidrag i tidsskriftet *Norsklæreren* som berører deltemaene skriving, lesing og muntlig kommunikasjon knyttet til vurdering, karaktersetting og særlig eksamen. Boken *Vurdering for læring i fag* (Dobson m.fl. 2010) henviser til hvordan norsk skriftlig kan oppsummeres i to hovedretninger – vurdering preget av 1) fokus på retting og språkstil og 2) prosessorientert skrivepedagogikk. Den første hovedretningen vektlegger produkt og har likhetstrekk med sluttvurdering, mens den andre er karakterisert som en form for vurdering for læring.

I kapittelet om norsk skriftlig i Dobson m.fl. (2010) diskuterer Fjørtoft (2010) potensialet for et nytt blikk på prosessorientert skriving knyttet til et forsterket fokus på kriteriebasert vurdering i læreplanen for Kunnskapsløftet (LK06). Bakke og Kverndokken (2010) skriver i samme bok om vurdering i muntlighet og hvordan muntlighet ikke bare er et tema for norskfaget, men at læreres kompetanse i vurdering av muntlighet må utvikles i alle fag. Når vi ser på andre typer av bidrag, for eksempel publisert i rapporter, finner vi noen eksempler på studier med en mer generell tilnærming til vurdering og karaktersetting orientert mot flere fag. I disse studiene er ikke norskfaget det hovedsakelige forskningsobjektet, men behandles som ett fag blant flere andre. Ofte er dette studier som har til hensikt å gi et oversiktsbilde av vurdering og karaktersetting, og der norsk inngår (Prøitz og Borgen 2010; Hovdhaugen m.fl. 2014; Sandvik m.fl. 2012). Senere i kapittelet trekker vi ut hovedmomenter fra noen av disse studiene for et mer konsentrert blikk på norskfaget spesielt.

Når det gjelder norsk muntlig viser studiene at til tross for at muntlighet er en viktig ferdighet i dagens samfunn, finner forskerne at det i liten grad undervises i muntlighet i 9. klasserom (Svenkerud 2012; 2013). Sammenliknet med undervisning i skriving blir dette funnet spesielt tydelig. Forskerne forklarer dette med at det er lange tradisjoner for undervisning i skriving og grammatikk, rettskriving, tekststruktur og sjanger i norskfaget. Dette er også undervisning som er preget av omfattende bruk av tilbakemeldinger til elever. Forskerne ser dette i sammenheng med utvikling av en vurderingskultur forankret i det faktum at skriftlig har vært et eksamensfag i generasjoner i Norge, så vel som i Norden. Denne vurderingskulturen er i tillegg styrket av prosesskrivingens fokus på veiledning underveis. Dette er tradisjoner man ikke har i muntlig. Forskerne fremholder samtidig at eksamen i muntlig med KL06 har endret seg fra å være en prøve i pensum til å bli en forberedt framføring – der hjelpemidler for fremføringen har vært tatt i bruk, som for eksempel PowerPoint-presentasjon der retoriske grep kommer i forgrunnen og får større

betydning. Til tross for endringer i læreplan og vektlegging av muntlig kommunikasjon som grunnleggende ferdighet, viser forskerne til at deres foreløpige resultater tyder på at det er eksamensform fremfor læreplan som gir status i norskfaget. De regner med at nye studier om noen år vil kunne avdekke mer bevisst arbeid med muntlighet (Svenkerud 2012). Svenkerud (2013) viser i sin avhandling om muntlige ferdigheter på 9. trinn at det er en generell mangel på fagspråk om muntlige ferdigheter. Hun peker også på at forskning viser at det gis gjennomgående lite veiledning, og at de tilbakemeldingene som gis til elevers arbeid med muntlig, er svake.

I den norske litteraturen vi har identifisert, tydeliggjøres det gjennomgående hvordan vurdering i norskfaget har vært og fortsatt «er» vurdering i skriving. Dette ses som er tett sammenbundet med sentralt gitt skriftlig eksamen, eksamensveiledninger og det fokus som legges til grunn her. Tradisjonen i norskfaget har vært preget av at man vurderte med karakterer basert på prøver i stilskriving på tentamen og eksamensdager (Berge 2014). Litteraturen tyder også på at det har skjedd en utvikling som er dels knyttet til utvikling og endring i vurderingsfeltet internasjonalt, beskrevet som en bevegelse fra det summative og mer testorienterte til det formative og en mer sammensatt bruk av ulike vurderingsformer. I Norge, som i mange andre land, har denne utviklingen bragt med seg et sterkere fokus på det formative og vurdering for læring (Baird m.fl. 2015), men også en introduksjon av tester og testing av elevferdigheter, som for eksempel de nye læringsstøttende prøvene i skriving på 5. og 8. trinn (Evensen 2014). Dels handler dette også om en utvikling og endring i Norge og i norske læreplaner med innføring av skriving som grunnleggende ferdighet og som læringsstøttende element i LK06 (Berge 2014).

3.4.1 Norsk læreres vurderingspraksis undersøkt gjennom FIVIS-prosjektet

I første delrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS) behandles vurdering i norskfaget i grunnskolen (Sandvik m.fl. 2012). Det er fremhevet i rapporten at det å være kompetent skriveleærer og leseleærer også betyr å være en kompetent vurderer. For at vurderingen skal være gyldig må læreren ha forståelse for det som skal vurderes. Det vises til at fordi dagens virkelighet er preget av multimodalitet, skapes det utfordringer i å bestemme hva som skal ligge til grunn for lesekyndighet og skriftkyndighet. I forlengelse av dette fremmes problemstillinger om at dårlig forståelse for fagets kompleksitet kan resultere i vurderinger som ikke er valide, sett i lys av læreplan og norskfaget forstått i sin historiske og kulturelle kontekst.

Med utgangspunkt i materiale fra fokusgruppeintervjuer med lærere finner forskerne i studien tre sentrale aspekter ved lærernes vurderingspraksiser i norsk:

1) bruk av mappe som vurderingsform, 2) prosessorientert skriving for fremme av skrivekompetansen hos elever og 3) utstrakt bruk av muntlige presentasjoner. Den observerte bruken av mappevurdering og muntlig presentasjoner i denne undersøkelsen fremheves som en grunn til undersøkelsens høye skåre på medelevvurdering i norskfaget. Det vises til at informanter peker på at de hadde positive erfaringer med denne typen vurderingsformer fra lærerutdanningen da de studerte. I rapporten pekes det på at formative vurderingspraksiser fra lærerutdanningen synes å ha fungert som modellering for informanternes yrkespraksis.

Lærer på ungdomstrinnet fremhever i Sandvik m.fl. (2012) forskjellige kriteriebaserte vurderingsverktøy for måloppnåelse og med kriterier. Informanter beskriver hvordan slike verktøy brukes i samtale med eleven for å snakke om gråsoner og nyanser som vurderingsverktøy alene ikke viser. I rapporten vises det også til at informanter beskriver at de har endret praksis. Mens de tidligere ikke ga elever tips til tentamen, anser de dette i dag ut fra et mer prosessuelt ståsted, der de gjerne gir hjelp og veiledning. Undersøkelsen viser også til at mappevurdering fremheves av informanter som kanskje den mest valide vurderingsformen for vurdering av elevenes samlede kompetanse mot slutten av året – men at dette er meget arbeidskrevende med tanke på retting. Informanten oppgir at dette er stressende. Lærerne i studien viser i intervju at det hersker uklarhet omkring hva slags kompetanse elever kan vise gjennom ulike oppgavetyper i norskfaget. Norsk lærerne er opptatt av hva som må til for å få beste karakter og sikre reliabilitet. Det blir vist til at Utdanningsdirektoratets eksamensveiledninger er en kilde til informasjon som hjelper med inndeling i under og middels, og over middels prestasjon.

I FIVIS delrapport 1 (Sandvik m.fl. 2012) pekes det også på at vurderingsprosedurene, dokumenter knyttet til standpunktvurdering, eksamenssensur og karaktersetting former videregående læreres syn på vurdering i norskfaget, for eksempel vises det til at eksamensveiledningene anses som konkrete beskrivelser av måloppnåelse. I studien pekes det på at materialet viser at vurderingens formative effekt er i utvikling i skolens norskfag. Materialet får også frem kompleksiteten i forhandlingsfeltet som lærere forholder seg til: mellom elevers ønsker, ytre stimuli og formativ vurdering, spenningen mellom tradisjoner og vaner knyttet til sluttvurdering og endring i vurderingsformer og funksjoner. Forskerne peker også på at de har observert ulike forståelser og praksiser av individuell vurdering i norskfaget og at de på dette grunnlag mener at det er en vei å gå mot praksisfellesskap i vurdering i norsk. I delrapporten reises det også et spørsmål om lesing, ettersom ingen av lærerne diskuterer lesekompetanse når det snakkes om vurdering. Dialogen de har hatt med informantene fra norskfaget handler i første rekke om inndelingen mellom muntlig og skriftlig, noe som de anser reflekterer eksamensveiledninger og ikke en kompetanseorientert forståelse av norsk som fag.

I FIVIS delrapport 2 (Sandvik og Buland 2013) behandles vurderingspraksiser i norskfaget i et eget kapittel. Sentralt i kapitlet står spørsmål om hvordan nasjonal læreplan henger sammen med lokale operasjonaliseringer og vurderingsformer. Det fokuseres også på hvilke begreper lærere bruker i beskrivelser av egen og andres vurderingskompetanse og praksis og om læringsprosesser i undervisning og læring (Fjørtoft 2013). I studien ser forskerne spesielt på skriving og lesing på småskoletrinnet og ut videregående. I studien er det også lagt til grunn et perspektiv knyttet til begrepet "backwash"-effekt for å beskrive hvordan vurdering virker tilbake på undervisning og læring (Fjørtoft 2013). I rapporten (Sandvik og Buland 2013) peker forskerne på forskjeller mellom norsklærernes vurderingspraksiser på barnetrinnet og på trinn med karakterer. Blant annet viser rapporten at det synes å være en balanse mellom standardiserte vurderingsformer og lærerdesignede vurderingspraksiser som sikrer reliabilitet og som bidrar til sterkere praksis- og fortolkningsfellesskap i vurdering av lesekompetanse på barnetrinnet. I ungdomsskolen og i videregående mangler lærerne denne typen kartleggingsverktøy i lesing. Studiens funn peker på at det mangler kunnskap og bevisstgjøring om lesedidaktikk i både ungdomstrinn og i videregående. Forskerne i prosjektet ser også en tendens til at det skjer en refortolkning av prosessorientert skriving som formativ praksis i faget. Samtidig tyder studien på at dette er praksiser som det er vanskelig å forene med summativ vurdering.

I studien (Sandvik og Buland 2013) vises det også til at lærere på ungdomstrinn og i videregående i mindre grad har en forståelse av lesing som er i tråd med læreplanens intensjoner og nyere forskning på leseforståelse. I rapporten pekes det spesielt på noen hovedbekymringer knyttet til mangelen på sterke fortolkningsfellesskap blant lærerne og klare føringer for hvordan læreplanmål kan kobles til ulike karakterer noe som kan ha konsekvenser for både validitet og reliabilitet i vurderinger med karakter. Med utgangspunkt i prosjektets funn understrekes behovet for norskdidaktisk utvikling i vurdering og da først og fremst på de øverste trinnene i opplæringssystemet.

3.4.2 Norsklæreres vurderingspraksis undersøkt som standpunkt-karaktersetting

NIFU har i tidligere studier undersøkt sider ved læreres arbeid med standpunkt-karaktersetting i flere fag, deriblant norsk. Intervjuer med lærere gir i disse studiene et innblikk i arbeidet med vurdering i fag generelt. Her gjengir og sammenstiller vi de mest sentrale funnene knyttet til norskfaget fra to av disse rapportene (Prøitz og Borgen 2010; Hovdhaugen m.fl. 2014). Hver for seg omfatter studiene data fra intervjuer med henholdsvis 16 norsklærere i gruppeintervjuer (Hovdhaugen m.fl. 2014) og ni norsklærere i individuelle dybdeintervjuer (Prøitz og Borgen 2010).

I Prøitz og Borgens studie fra 2010 viser de intervjuede norsklærerne til at de i arbeid med standpunktvrdering må skille mellom tre karakterer; skriftlig, muntlig og sidemål. I hovedsak er det skriftlige prøver og særlig eksempel-eksamensoppgaver gitt som heldagsprøve som utgjør grunnlaget for standpunkt-karaktersettingen i norsk skriftlig. I muntlig peker informanter på at «hver time er en vurderingssituasjon». I tillegg kommer ulike former for forberedte fremføringer. Et poeng som blir fremhevet når det gjelder sidemål, er at det er forskjell mellom hvordan man vurderer i de to målformene. I sidemålet ser lærerne etter ortografiske og generelle språklige problemer elevene måtte ha, mens i hovedmål er det innholdet som kommer i forgrunnen.

I rapporten presenteres en oversikt over hvilke kilder lærerne i studien oppgir at de vektlegger når de vurderer for standpunkt-karaktersetting i ungdomsskole og i videregående. I tabell 3.2 har vi trukket ut resultatene for norskfaget. Selv om dette gjelder standpunktvrdering, viser lærernes kildebruk (se tabellen) hvordan øvrige vurderingssituasjoner gjennom skoleåret får betydning for sluttvurderingen. I lærerintervjuene som ligger til grunn for evaluering av forsøket med redusert antall karakterer ved halvårsvurdering i norsk (se eget kapittel senere i denne rapporten) kommer vi tilbake til hvordan vurderingssituasjoner underveis i skoleåret får relevans for sluttvurderingen, og hvordan sluttvurderingen dermed også legger premisser for de forutgående vurderingssituasjonene. Dette for at lærerne skal ha tilstrekkelig vurderingsgrunnlag når eleven skal få karakter.

Tabell 3.1: Kilder norsklærere vektlegger i arbeid med standpunkt-karaktersetting (bearbeidet utdrag fra NIFU rapport 16-2010)

Norsk	Viktigste kilde	Supplerende kilder
Ungdomsskole	Skriftlig: tentamen basert på eksamensoppgaver, skriftlige prøver Muntlig: fremføringer og aktivitet i timen Sidemål: skriftlige prøver	Skriftlige prøver Lærerens inntrykk av eleven
Videregående skole	Hovedmål skriftlig; skriftlige prøver, eksamensoppgaver, delprøver Sidemål; skriftlige prøver, muntlig aktivitet Muntlig; forberedt muntlig presentasjon	Hver time er en vurderingssituasjon for muntlig karakter Kjennskap til elevene over tre år (gjelder på studieforberedende utdanningsprogram)

Norsklærerne i Prøitz og Borgens (2010) undersøkelse viser til at skjønn er viktig når de setter karakter, men i dette inngår både eget og kollegaers skjønn. En lærer viser i studien til at også hun som lærer kan være i tvil om en prestasjon skal vurderes til for eksempel 5 eller 6. Mange av lærerne vektlegger derfor tolkningsfellesskapet med kollegene.

I Hovdhaugen m.fl. (2014) undersøkelse forteller norsklærerne at de ikke anser det som gjennomførbart å la en heldagsprøve måle alle kompetansemål ved avslut-

ningen av Vg3. De må derfor "lukke" kompetansemål i norsk underveis i opplæringen. I rapporten vises det til at norsklærerne har utviklet praksiser som gjør at de foretar vurdering av elevenes sluttkompetanse over et lengre tidsrom og med en bredere innfallsvinkel enn hva matematikklærerne i samme studie gjør. Det vises til at på et overordnet nivå anser norsklærere seg som avhengig av flere vurderingssituasjoner over tid for å kunne sette en riktig standpunktkarakter. Dette er i kontrast til matematikklærerne, som mener at årets siste vurderingssituasjon vil være tilstrekkelig for å sette standpunktkarakter. Norsklærerne fremhever en gjennomgående og fortolkende bruk av vurderingskriterier i arbeidet. De tilstreber også at vurderingskriterier skal kunne brukes på et ensartet vis av kolleger som vurderer den samme elevbesvarelsen. På denne bakgrunnen fremtrer behovet for kollegasamarbeid i vurdering.

3.4.3 KAL-prosjektet

Når forskningsarbeid med vurdering i norskfaget skal studeres, kommer man ikke utenom prosjektet "Kvalitetsvurdering av læringsutbyttet i norsk skriftlig" (KAL-prosjektet). Formålet med prosjektet var å vurdere læringsutbyttet i norsk skriftlig i grunnskolen, slik det kom til uttrykk gjennom avsluttende eksamen på tiende trinn. KAL-prosjektets hovedfokus - eksamen - ligger utenfor denne evalueringens siktemål der vi ser på problemstillinger knyttet til halvårsvurdering. KAL-prosjektets fokus var også avgrenset til norsk skriftlig, en avgrensning som blir for smal for denne rapportens formål. For det tredje omhandlet KAL-prosjektet en annen reform, prosjektet samlet inn eksamensdata fra 1998,1999 knyttet til Mønsterplan av 1987 og eksamensdata fra 2000 og 2001 innenfor den nye Læreplanen av 1997. Likevel gjengir vi her noen hovedkonklusjoner fra studien fordi vi mener det er elementer i studien som fortsatt har aktualitet. Prosjektet har hatt stor betydning for arbeid med og utvikling av skriftlig eksamen i norsk og diskusjoner omkring sensur og betydningen av tolkningsfellesskap mellom lærere i vurdering.

I rapporten pekes det på at grunnskolen har en egen, sammensatt tekstkultur, med klare delkulturer. En spesifikk litterær form for fortellende skriving står særlig sterkt i denne kulturen. Videre vises det til at avgangsprøven hadde en bestemt skriveideologi karakterisert som en ekspressivistisk orientert modell for personlig vekst, og at modellen står nært den prosessorienterte tradisjonen for skriveopplæring. Videre beskrives hvordan «[d]en prosessorienterte modellen har fått avgjørende konsekvenser for forholdet mellom skriver og leser, for synet på kandidatens rettigheter og posisjoneringsarbeid i eksamenssituasjonen og for posisjonering av sensorenes vurderingsarbeid, for forholdet mellom situasjons- og kulturkontekst og for forståelsen av kulturelle kategorier som sjanger, tekst og språk» (Evensen 2003:16). Forskerne konkluderte med at prioriteringen av en fortellende skriving

bunner i en underliggende skriveideologi og kommunikasjonskontrakt som tilstreber symmetri mellom skriver og leser, det er en ekspressiv ideologi som vektlegger privat erfaring fra intimsfæren og som fører til en eksamen hvor bredden av kulturelle skrivemåter blir liten. Prosjektet viste at reliabiliteten i læreres vurdering og karaktersetning til eksamen etter 10. trinn var høyere enn funn i tilsvarende studier i videregående (Evensen 2014). En tolkning av dette funnet presentert senere er at sensorsamtaler om tekster og mer klart uttalte kriterier hadde bidratt til utvikling av tolkningsfellesskap (Evensen 2014).

3.4.4 Classroom assessment in writing

Med utgangspunkt i betraktninger i den norske litteraturen har vi søkt etter internasjonale bidrag som diskuterer vurdering i majoritetens morsmålsfag. På samme måte som for norske bidrag er det vanskelig å identifisere litteratur som fokuserer på vurdering i morsmåls-språkfag som helhet. Det vi imidlertid identifiserte med våre søk, er at det på samme måte som i den norske litteraturen finnes bidrag innenfor spesielt delelementene lesing/literacy og skriving. Avslutningsvis i dette kapittelet konsentrerer vi oss derfor om et sentralt bidrag med en bred orientering mot vurdering og skriving.

Parr (2013) konkluderer sitt kapittel "Classroom assessment in Writing" i *SAGE Handbook of research on classroom assessment* med at det gjenstår mye arbeid på dette feltet fordi forskningslitteraturen om vurdering i skriving er begrenset. Hun skriver at dagens klasseromsvurdering gir mange muligheter for å vurdere ulike former for skriftlig kommunikasjon under varierte forhold, men på grunn av mangel på empirisk forskning er det vanskelig å peke ut bestemte modeller som er bedre enn andre for elevers utvikling av skriveferdigheter. Generell uenighet omkring hva kvalitet er i skriving, gjør at lærere har behov for faglig støtte i valg av læringsaktiviteter og arbeid med elevers skrevne tekster. Sett fra et teoretisk ståsted og overordnet perspektiv har det skjedd en bevegelse fra det å se på skriving som et produkt og som en fagovergripende og overførbar ferdighet som skulle undervises og vurderes, til å se på skriving som en sosial handling påvirket av formål, innhold og kontekst (Parr 2013). En sentral betraktning er at koblingene mellom dagens skriveteorier, pedagogikk og vurdering er utydelige. Parr hevder altså at vurdering av skriving er svakt koblet til gjeldende skriveteorier.

Internasjonal litteratur viser at selv om litteraturen om vurdering generelt har beveget seg i en mer formativ retning, domineres feltet fremdeles av vurderingsformer som summativt tester elevers skriveprestasjoner i henhold til standarder. Parr advarer mot at disse vurderingsformene påvirker valg av undervisningsform og skriveprosess i klasserom, og at tester har en tendens til å fremme uintendert læring, for eksempel et for smalt skrivebegrep. Et svar på disse utfordringene har

vært å ta i bruk såkalt autentisk vurdering, som kan representere elevers skriveprestasjoner fra varierte kontekster ved bruk av portfolio (ofte kalt mappevurdering i Norge). Portfolio komprimerer informasjon fra flere skrivesituasjoner samlet inn over tid og gir således rikere og mer sammensatt informasjon om elevers innsats, progresjon og skriveprestasjoner (Parr 2013).

Bruk av portfolio i skriving er imidlertid omdiskutert, og det er stilt spørsmål om hvorvidt portfolio faktisk kan anses å være en vurderingsform. Det er for eksempel pekt på at portfolio kan betraktes som en pedagogisk hensiktsmessig prosedyre for samling tekster som har *likhetstrekk* med vurderingsprosedyrer, uten nødvendigvis å være det. Den såkalte portfolio-bevegelsen viste uansett at det var behov for flere vurderingssituasjoner og innsamling av elevarbeider over tid for å kunne gjøre mer solide vurderinger. Til tross for at portfolio-tilnærmingen så ut til å kunne skape en kobling mellom vurdering og aktiviteter i klasserommet, har forskning i liten grad klart å påvise en direkte kobling mellom undervisning og læring ved hjelp av portfolio. Samtidig blir det fremhevet som viktig at «literacy»-lærere driver systematisk og gjennomgående innsamling og observasjon av elevers lesing og skriving som en integrert del av undervisningen (Parr 2013).

Det finnes flere former for vurdering i skriving. En av disse er selvsagt knyttet til lærerens daglige observasjon, veiledning og diskusjon med elever om deres skriving. Dette er situasjoner der elevens skriveutvikling gis en umiddelbar reaksjon fra læreren. I mer formelle vurderingssituasjoner samles observasjoner (ofte nedskrevet) og beskrivelser av skrivesituasjoner, det være seg elevens autonome skriving, utholdenhet, ressursbruk og egne betraktninger om egen skriving som grunnlag for vurdering. I det hele tatt fremheves gjennomgående viktigheten av at elever får mulighet til å utvikle seg som skrivere, der de selv tar ansvar for egen skriving og refleksjon omkring skrivingen. I denne sammenheng fremheves såkalt «conferencing» som en metode som gir muligheter for situasjonsbestemt vurdering og påfølgende beslutningstaking fra lærere om det videre arbeidet som kan bidra til å bringe elevers utvikling fremover. Dette krever imidlertid at læreren kjenner eleven godt, har fulgt elevens utvikling over tid, at eleven tar initiativ til interaksjon og gir informasjon om egen evne og mulighet til å følge sin egen utvikling, at eleven kan uttrykke hva han/hun har oppnådd og hvordan læreren bruker denne informasjonen og annen kunnskap til å identifisere og tas i bruk situasjoner som oppstår. Dette er imidlertid ansett å være en kompleks og krevende undervisnings-/vurderingsform selv for dyktige lærere (Parr 2013).

Parr (2013) beskriver det hun kaller planlagt formativ skrivevurdering (“planned formative writing assessment”) med likhetstrekk til norsklæreres arbeid med vurdering knyttet til elevenes skriving. Dette er vurdering som i stor grad dreier seg om å studere utvalg og/eller samlinger av skrivearbeider som utkast eller i en-

delig form. Parr (2013) understreker at det å lage oppgaver for slik vurdering krever at lærere utvikler evnen til å lage «rike» oppgaver for elever og kunne studere svarene på oppgavene på en nyansert måte slik at de bestemte skrivefunksjonene eller dimensjonene som skal vurderes, faktisk blir vurdert. Parr (2013) viser til at mange oppgaver ikke gir informasjon som er detaljert nok, enten det dreier seg om setningsoppbygging, syntaks eller elevens evne til å organisere og strukturere informasjon og kunnskap.

3.4.5 Slutsats om vår undersøkelse av litteratur om vurdering i norskfaget

Våre litteratursøk har identifisert en begrenset mengde litteratur om vurdering i norskfaget som helhet. Vi har imidlertid identifisert bidrag knyttet til særlig skrive- men også muntlig i norsk. Det ser ut til å være få empirisk forankrede studier av nyere dato, dette bekreftes blant annet når det gjelder muntlig av Svenkerud (2013), og gjennom søk i databaser som Oria, NorART og SkrivBiB. Litteraturgjennomgangen tyder på et fravær av lesing som isolert tema i litteraturen knyttet til vurdering mot karakter. På den annen side er lesing en omfattende tematikk ofte knyttet til testing av lesing i bred forstand forstått som literacy. Ofte er dette vurdering med andre formål enn læreres vurdering for å sette karakter, som ved kartleggings- og diagnostiske formål eller for mer overordnede systemformål som i internasjonale sammenlikningsstudier.

Vår litteraturgjennomgang i forbindelse med denne evalueringen gir et grunnlag for å peke på mangelen på forskningsmessig oppdaterte og empirisk baserte studier av vurdering i norskfaget på et generelt grunnlag. Med innføringen av Kunnskapsløftet, utvikling av tester og nye ordninger for vurdering generelt med konsekvenser for undervisning og vurdering i norskfaget som helhet tydeliggjør et behov for å innhente ny kunnskap om vurdering i norskfaget.

3.5 Oppsummering

I dette kapittelet har vi sett læreres undervisningspraksis i norskfaget eksemplifisert ved forelesninger/innledninger, tilrettelegging for gruppeoppgaver og gjennomføring av slike oppgaver som kan strekke seg over lengre perioder, kortere illustrerende aktiviteter (tablåer), elevers lesing av tekster i utdrag eller hele originalverk, filmvisning, dramatisering, lærers "førlesing", læreres tilrettelegging av elevers lesing med såkalte observasjonsspørsmål underveis, læreres tilrettelegging av en samtale for ulike typer sammenligning av tekst, tema eller problemstilling, eller med henblikk på elevers personlige identitetsutvikling. Vi har også sett hvordan gruppearbeid og elevfremføringer benyttes i vurderingssituasjoner, vi har

sett elevers loggføring i praksis og vi har fått et elevperspektiv på prosessorientert skriving. I sidemålsundervisningen har vi særlig sett oppgaveløsning knyttet til grammatikk. Her er det et mangfold av metoder og kommunikasjonsformer mellom lærer og elev og mellom elever, og læreren blir en tilrettelegger, en formidler, en faglig autoritet (på sjangertyper, på retorikkregler, på litteraturanalyse), en regissør, en samtalepartner, planlegger og organisator i «læring som hendelsesrike sosiale prosesser,» slik Moslet (2009) beskriver dette. I tillegg har vi sett flere eksempler på lærer som støtter elevers utvikling og mestring gjennom umiddelbare tilbakemeldinger (gruppearbeid, elevfremføringer) og gjennom prosessorientert skriving (særlig ved bruk av rosende kommentarer).

Gjennom eksemplene gjengitt i dette kapittelet fremstår undervisningen som peker frem mot den muntlige karakteren i norskfaget som den mest hendelsesrike, og undervisningen i skriving som den mest vurderingsorienterte med et tilretteleggende og utviklende siktemål. Eksemplene på undervisningspraksis i nynorsk som sidemål skiller seg ut ved å være mindre «hendelsesrike», forstått som mer orientert mot innlæring av formverk og grammatikk og mindre orientert mot opplevelsesrik samhandling i klasserommet. Nynorskundervisningen representerer kanskje også de mest formaliserte tilbakemeldingsrutinene.

Hvis vi ser på de to rapportene fra FIVIS (Sandvik m.fl. 2012; Sandvik og Buland 2013) og de to NIFU-rapportene (Prøitz og Borgen 2010; Hovdhaugen m.fl. 2014), gir de samlet et bilde av norsklæreres vurderingsarbeid som kjennetegnet av noen overordnede fellestrekk. Norsklærere beskriver norskfaget som preget av arbeidet med tre karakterer, skriftlig, muntlig og sidemål. Vurderingsgrunnlaget består av ulike typer av elevarbeider, men der skriftlige kilder ser ut til å dominere. Heldagsprøve er den vurderingsformen som gir det tyngste grunnlaget for å sette standpunkt-karakter. Ofte er dette prøver som er inspirert av Utdanningsdirektoratets eksempeloppgaver for eksamen og tilhørende vurderingskriterier i eksamensveiledninger. Muntlig karakter settes på grunnlag av aktivitet i timer og særlig med utgangspunkt i forberedt muntlig presentasjon. Samtidig er det tydelig at skillene mellom skriftlig og muntlig ikke nødvendigvis alltid opprettholdes når det er tvil om elevers karakterer. Norsklærere har behov for flere vurderingssituasjoner gjennom året for å sette karakterer som oppleves som valide. Kriterier og kollegasamarbeid fremheves som viktige for å sikre reliabel og rettferdig vurdering. Det er behov for norskdidaktisk utvikling i vurdering i norskfaget, først og fremst orientert mot de øverste trinnene i opplæringssystemet.

Våre litteratursøk har identifisert en begrenset mengde litteratur om vurdering i norskfaget som helhet. Vi har imidlertid identifisert bidrag knyttet til særlig skriving men også muntlig i norsk. Ut fra disse observasjonene er det vanskelig å si noe veldig bestemt om vurdering i norskfaget generelt utover å peke på noen hovedtrekk og hovedtemaer som går igjen i faglitteraturen:

- Skrivning ser ut til å være hoved-vurderingsformen i norskfaget. Skrivning studeres og diskuteres sett i sammenheng med eksamen og/eller eksamensveiledninger. Eksamen og eksamensveiledninger i norsk skriftlig brukes også ofte som begrunnelse for ulike vurderingspraksiser i norsk.
- Muntlig beskrives ofte som et undervurdert delaspekt av norskfaget. Muntlighet i norskfaget og som element i alle fag gjennom Kunnskapsløftets grunnleggende ferdigheter er spådd en mer fremtredende rolle i både undervisning og i vurdering.
- Mappe som verktøy for vurdering i norskfaget fremtrer i flere av bidragene vi har sett på i dette kapittelet. Dels som arbeids- og undervisningsmåte/metode, dels som vurderingsform. Samlet synes disse bidragene å peke på at det viktigste med denne måten å arbeide med vurdering på, er betydningen av flere og varierte vurderingssituasjoner og vurderingsformer i norskfaget. Dette er til nytte både for lærerens vurdering og for elevenes læring.
- Litteraturen fremhever også betydningen av tid og rom for muligheter for å utvikle og drive tolkningsfellesskap for å fremme reliabel og valid vurdering i norskfaget.

Lesere som er interessert i metoder for våre søk, kilder og avgrensning av litteratursøk henvises til delrapport 1 (Seland m.fl. 2015) for helhetlig gjennomgang av dette.

4 Analyse av karakterdata

I dette kapittelet analyserer vi karakterdata for å undersøke om det er mulig å identifisere systematiske forskjeller i elevenes læring som følge av deltakelse i forsøket. Dette utgjør det overordnede forskningsspørsmålet som belyses i denne sluttrapporten.

Som hovedutfallsmål i analysene bruker vi gjennomsnitt av standpunktkarakterene i norsk hovedmål, norsk sidemål og norsk muntlig, men vi tester også med utfallsmål basert på enkeltkarakterer (nærmere beskrevet nedenfor). Avveininger ved bruk av standpunktkarakter som mål på læring er nærmere drøftet i kapittel 8, hvor vi kommer tilbake til funnene som presenteres i dette kapittelet.

Basert på analysen av karakterdata skal vi også i dette kapittelet gi indikasjoner på hvorvidt det peker seg ut modeller for videre utprøving av forsøket. Vi oppfatter skolenes eller skoleeiernes valg av én versus to karakterer ved halvårsvurdering som forsøkets to modeller for gjennomføring. Slike indikasjoner vil følgelig komme til syne hvis det er signifikante forskjeller mellom elevers gjennomsnittlige standpunktkarakter i norsk avhengig av hvorvidt de har deltatt i gjennomføringen av forsøket med én karakter i motsetning til to karakterer ved halvårsvurdering.

Metoden vi har brukt omtales som forskjell-i-forskjeller analyse ("difference-in-differences"). Dette er en form for kvasi-eksperimentelt forskningsdesign hvor vi sammenligner *endringene* i utfall over tid mellom dem som deltar i et forsøk og dem som ikke deltar, for å se om de to gruppene utvikler seg ulikt med hensyn til utfallet før og etter forsøket. Dersom utfallsmålet for tiltaks- og sammenligningsgruppen utvikler seg signifikant forskjellig før og etter forsøket, har vi funnet støtte for at forsøket har hatt en effekt. Valg av metode for analyse av karakterdata er nærmere presentert og diskutert i kapittel 1 i denne rapporten.

4.1 Avgrensninger av datamaterialet

For å kunne gjennomføre karakteranalysene har vi fått data fra Utdanningsdirektoratet. Disse dataene inneholder informasjon om standpunktkarakterer i norsk (hovedmål, sidemål og muntlig) og informasjon om eksamenskarakterer for i alt åtte kull som startet på studieforberedende utdanningsprogram i skoleårene

2007/08 til 2014/15. Vi har i avgrensningen av hva som inngår i vår avhengige variabel valgt å bruke gjennomsnitt av standpunktkarakterer i norsk muntlig, skriftlig hovedmål og skriftlig sidemål som hovedutfallsmål.

For å undersøke om forsøket kan ha påvirkning på eksamenskarakterer og på sidemålskarakter har vi også gjennomført kontrollanalyser der vi tar utgangspunkt i eksamenskarakteren for skriftlig hovedmål, eksamenskarakter i skriftlig sidemål samt standpunktkarakter i sidemål. Alle som går Vg3 kommer opp i eksamen i skriftlig hovedmål, slik at vi har data for alle elever for dette utfallsmålet. Eksamen i skriftlig sidemål følger en trekkordning, men siden det likevel er relativt mange som kommer opp i skriftlig sidemål til eksamen har vi også tatt med dette som utfallsmål, sammen med standpunkt i sidemål som alle elever uten fritak har. Derimot har vi ikke kunnet gjøre analyser av eksamen i muntlig da dette er en eksamen med trekkordning, og det er et stort antall elever som ikke har denne karakteren.

Fra tidligere forskning på karakterer vet vi at det er systematiske variasjoner i eksamenskarakterer sammenlignet med standpunktkarakterer: elever pleier å få noe lavere eksamenskarakter enn standpunktkarakter på skriftlig eksamen, mens karakteren på muntlig eksamen ofte er den samme eller litt høyere enn tilsvarende standpunktkarakter (Hovdhaugen m.fl. 2014). Dette har imidlertid lite påvirkning for analysene vi gjennomfører i denne sluttrapporten, da fokus er på hvordan utfallsmålet endrer seg over tid, ikke hvilket nivå det ligger på. Med andre ord har vi kjørt analyser med følgende utfallsmål:

- 1) Gjennomsnitt av de tre standpunktkarakterene
- 2) Kun eksamenskarakter i norsk skriftlig hovedmål
- 3) Kun eksamenskarakter i norsk skriftlig sidemål
- 4) Kun standpunktkarakter i skriftlig sidemål.

Alle analyser er dokumentert i vedleggstabellene.

Dataene inneholder også en rekke bakgrunnsvariabler som er viktige i analysene, slik som karakterer fra avsluttet grunnskole (både gjennomsnittskarakter i form av grunnskolepoeng og de tre standpunktkarakterene i norsk hovedmål, norsk sidemål og muntlig norsk), kjønn og type utdanningsprogram (studiespesialisering, idrettsfag eller musikk, dans og drama). I tillegg har vi også tatt inn variabler på skolenivå, slik som deltakelse i forsøket og om deltakende skoler har valgt en ordning med én eller to karakterer i norsk.

Filen vi fikk fra Utdanningsdirektoratet var organisert som elev-fag-observasjoner, hvor hver observasjon var et fag som en elev hadde deltatt i, med tilhørende informasjon om karakterer. Opprinnelig inneholdt filen nesten 2,4 millioner elev-fag-observasjoner og omtrent 293 000 unike elever, hvorav ca. 270 000 elever har observasjoner av karakterer på Vg3, som er vårt utfallsmål.

Karakterene er registrert som tallkarakterer (1-6), men det finnes også andre verdier i variablene for standpunktkarakter. Her brukes mange forskjellige koder, men vanligst er D, som står for «deltatt», IV for «ikke vurderingsgrunnlag» og F for «fritatt». Noen svært få (mellom 10 og 30) har fått karakterene B for «bestått», IG for «ikke godkjent» og ID for «ikke deltatt». I hovedanalysene benytter vi oss av de elevene som har fått en karakter med en tallverdi (1-6), for å beregne gjennomsnittskarakter i norskfaget. Vi har tilleggsanalyser der vi undersøker om andelen som får ikke vurderingsgrunnlag endrer seg for elever som har vært en del av forsøket sammenlignet med de som ikke har vært med. For eksamenskarakterene forholder vi oss bare til elevene som har fått en karakter med verdi mellom 1 og 6.

Elever som ikke har fulgt skoleløpet gjennom hele videregående (for eksempel på grunn av at vedkommende var utvekslingselev i Vg2), er holdt utenfor analysene. Videre er også elever som har byttet skole holdt utenfor analysene, siden skolebytte gjør det vanskelig å kontrollere for skolefaste effekter. Det er kun karakterer som er oppnådd i det normerte løpet som blir benyttet i analysen. Det vil si at vi har sett bort fra de karakterene som er et resultat av at elever har tatt opp fag etter at de egentlig skulle ha vært ferdig med skoleløpet (forbedring av karakterer), slik at kun de opprinnelige karakterene er beholdt.

I den opprinnelige filen hadde også elever som gikk andre og tredje året kommet med for det første kullet, 2007-kullet. Disse er ikke omfattet av forsøket og holdes derfor utenfor analysene. I den opprinnelige filen var det to variabler for karakter, der den første kun skulle være for Vg1-fag og Vg2-fag, mens den andre kun skulle være standpunktkarakterer for Vg3-fag. Imidlertid fant vi tilfeller der det var skrevet inn karakterer for begge disse variablene for Vg3-fag. I de fleste tilfeller var det samme karakter i begge variablene, men i de tilfellene der det var forskjellig karakter, ble det vanskelig å avgjøre hvilken som faktisk var standpunktkarakter. I disse tilfellene fjernet vi elev-observasjonene (149 elever).

Da forsøket ble implementert valgte de fleste skoler enten én karakter eller to karakterer. Noen få skoler, totalt 15 i antall, valgte å innføre kombinasjonsvarianter, enten at man hadde én karakter i Vg1 og to karakterer i Vg2, eller at de endret variant underveis (fra én karakter til to karakterer, eller fra to karakterer til én karakter), eller at de trakk seg fra forsøket underveis. Elever fra disse 15 skolene er holdt utenfor analysene, da de verken kan være med i tiltaksgruppen eller sammenligningsgruppen.

For å kunne analysere materialet har vi aggregert dataene på individnivå, det vil si skapt en fil der hver person er én observasjon med tilhørende verdier på standpunktkarakterer, eksamenskarakter og andre nevnte kontrollvariablene. Etter å ha definert utvalget som skal være med i analysen, har vi 197 217 elevobservasjoner fordelt på 332 skoler. Dette antallet omfatter både skoler som var innenfor og skoler som var utenfor forsøket. Om lag 18 prosent av utvalget var med i forsøket. De

resterende gikk enten på en skole som ikke var del av forsøket eller gikk på en tiltaksskole før forsøket startet.

4.1.1 Utfallsvariabel

I analysene opererer vi med flere typer utfallsvariabler, men felles for dem alle er at de tar utgangspunkt i karakterer i norsk oppnådd i Vg3, og som føres på elevens vitnemål, det vil si standpunktkarakter og eksamenskarakter. I hoveddelen av analysene er utfallsvariabelen gjennomsnittet av de tre standpunktkarakterene, og det er disse analysene som danner grunnlag for teksten og for figurer som blir presentert. Imidlertid, siden vi har informasjon om eksamenskarakter i norsk hovedmål for alle elever, har vi også kontrollert at vi ikke får et annet resultat dersom vi bruker eksamenskarakteren i norsk hovedmål som utfallsvariabel. Vi har også kjørt analyser på standpunktkarakter og eksamenskarakter i norsk sidemål.

4.2 Forventinger til forsøket

Basert på de positive holdningene som kommer frem fra lærerintervjuene i delrapport 2 fra evalueringen (Seland m.fl. 2016), kan man forvente at det å ha færre karakterer på Vg1 og Vg2 muligens kan gi en positiv effekt på karakternivået i Vg3, der elevene blir vurdert med tre standpunktkarakterer, samt eksamenskarakter i hovedmål. Denne forventningen er basert på at lærerne antar at færre karakterer og færre prøver i Vg1 og Vg2 vil bidra til å øke elevenes motivasjon for faget. Bøyum og Norendal (2011) argumenterer for at en reduksjon av antallet karakterer i norskfaget har en gunstig effekt på elevenes motivasjon for læring i faget (se kapittel 1 i denne rapporten).

Samtidig vet vi fra et forsøk med valgfritt sidemål i Oslo kommune at det å få færre karakterer i norsk (ved at elevene ikke fikk vurdering i sidemål, selv om de fikk undervisning i sidemål) ikke hadde særlig stor effekt på holdningen til norskfaget (Vibe & Borgen 2007). Med utgangspunkt i denne påstanden kan vi like godt anta at forsøket ikke vil ha noen nevneverdig effekt på karaktergjennomsnittet i norsk i Vg3.

Vi vil i analysene teste om forskjellen mellom de som er i tiltaksgruppen og i sammenligningsgruppen er signifikant forskjellig fra null. Det vi kaller null-hypotesen vår er derfor at forskjellen i gjennomsnittskarakteren i norskfaget ikke er signifikant forskjellig fra null. Vi godtar et signifikansnivå på 0,1, som vil si at vi godtar at sannsynligheten for å få et resultat som viser at forskjellen ikke er null, selv om den likevel er null, er 10 prosent.

Den alternative hypotesen er at forsøket har ført til en forskjell i karakternivået mellom elever som har deltatt i forsøket og elever som ikke har deltatt i forsøket.

4.3 Forskjeller på tiltaksskoler og sammenligningsskoler

Før vi begynner analysene vil vi se nærmere på om det er forskjeller mellom skoler som har vært en del av forsøket og de som ikke var det (sammenligningsskolene). Denne deskriptive framstillingen viser noen mønstre knyttet til hva slags skoler som har valgt å delta i forsøket. Tabell 4.1 viser at det er en relativt jevn fordeling på skoler, dersom vi ser på alle skoler. Halvparten av skolene er ikke med i forsøket (sammenligningsskoler), 4 prosent er holdt utenfor fordi de har deltatt med ulike varianter eller avbrutt forsøket, og resten er forholdsvis jevnt fordelt på forsøk med én karakter eller to karakterer. Dersom vi sammenligner allmennfaglige skoler med kombinerte skoler (skoler som også tilbyr yrkesfaglige utdanningsprogrammer) ser vi dels at allmennfaglige skoler deltar i forsøket i litt større grad enn kombinerte skoler, samt at det er en litt større andel av de allmennfaglige skolene som har valgt å være med i forsøket med én karakter.

Tabell 4.1: Fordeling på tiltaksskoler og sammenligningsskoler, etter allmennfaglig/kombinert skole

	Allmennfaglig skole	Kombinert skole	Alle skoler
Én karakter	31 %	19 %	21 %
To karakterer	21 %	26 %	25 %
Ikke med	45 %	51 %	50 %
Tatt ut av analysen	3 %	4 %	4 %
N=100%	62	278	340

Tabell 4.2 viser fordelingen på offentlige og private videregående skoler. Hovedforskjellen mellom de to typene skoler er at private skoler har deltatt i forsøket i mye mindre grad enn offentlige skoler.

Tabell 4.2: Fordeling på tiltaksskoler og sammenligningsskoler, etter offentlig/privat skole

	Offentlige skoler	Private skoler
Én karakter	24 %	12 %
To karakterer	27 %	18 %
Ikke med	49 %	70 %
N=100%	290	50

Den siste variabelen vi skal undersøke er skolestørrelse, der vi har delt skolene i tre omtrent like store grupper og skiller mellom små skoler (300 elever eller færre), mellomstore skoler (301-600 elever) og store skoler (601 elever eller flere), for å se om det er forskjell i deltakelse i forsøket. Tabell 4.3 viser at jo større en skole er, desto større er sannsynligheten for å ha deltatt i forsøket. Blant de minste skolene, med færre enn 300 elever er det kun 21 prosent som deltar i forsøket, mens det er 55 prosent av de mellomstore skolene som deltar og to tredeler

av de store skolene som deltar i forsøket. Generelt er det en jevn fordeling på hvilken type forsøk de som deltar er med på, om de har valgt én eller to karakterer. Det eneste som skiller seg ut er at det er en litt større andel av de store skolene som har valgt et forsøk med to karakterer.

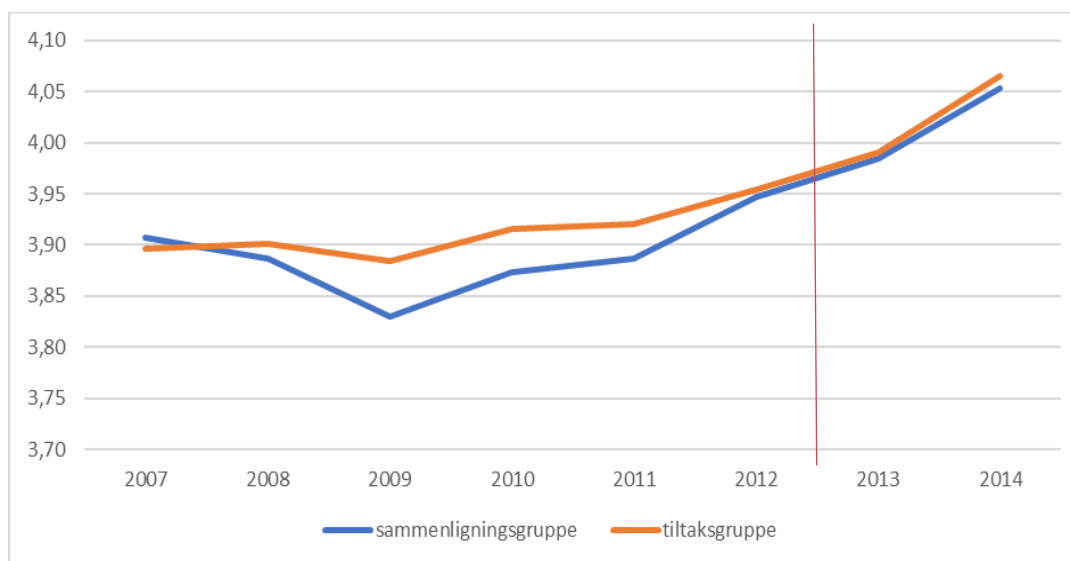
Tabell 4.3: Fordeling på tiltaksskoler og sammenligningsskoler, etter skolestørrelse

	Små skoler (- 300)	Mellomstore skoler (301-600)	Store skoler (601+)
Én karakter	10 %	27 %	30 %
To karakter	11 %	28 %	38 %
Ikke med	79 %	45 %	32 %
N=100%	110	118	112

Samlet sett ser vi at det er noen små forskjeller mellom skolene. Offentlige skoler og skoler med allmennfaglig innretning deltar oftere enn private skoler og kombinerte skoler, og sannsynligheten for å være med i forsøket øker jo større skolen er. Alle disse variablene kontrollerer vi for i den videre analysen fordi vi kan kontrollerer for skolefaste effekter. Dette innebærer at vi tar hensyn til alle skolekarakteristika som ikke endrer seg over tid.

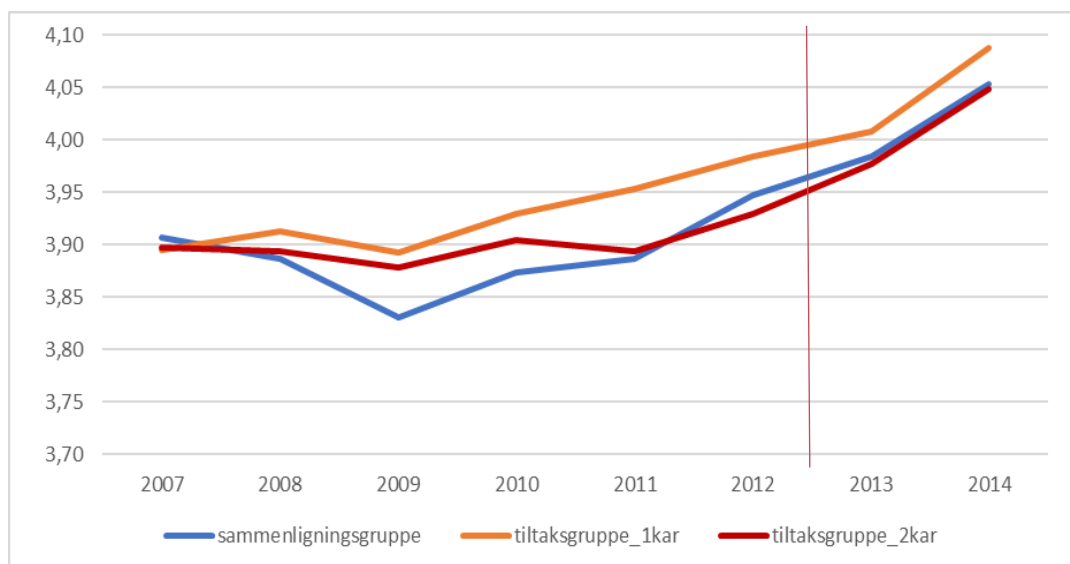
4.4 Grafiske analyser

Vi starter med å se på sammenhengen mellom utfallsmålet og deltakelse i forsøket i en figur. Utfallsmålet er gjennomsnittlig standpunktkarakter i norsk, basert på skriftlig hovedmål, skriftlig sidemål samt muntlig karakter. Figur 4.1 illustrerer utviklingen over tid i gjennomsnittskarakteren i norskfaget for tiltaksgruppen (rød) og sammenligningsgruppen (blå) der tiltaksgruppen inkluderer alle skolene som har vært med i forsøket. Det fremgår av figuren at det ikke er en stor forskjell i utfallsvariabelen mellom tiltaksskolene og sammenligningsskolene. Figuren viser at det var noe variasjon før tiltaket ble satt i verk, men at de to linjene i stor grad følger hverandre, samt at det ikke er noen forskjell etter forsøket ble satt i verk. Vi ser også en svak generell stigning i gjennomsnittskarakteren i norskfaget over tid.



Figur 4.1: Grafisk fremstilling av forskjell-i-forskjeller analysene på gjennomsnittlig standpunkt-karakter i norsk, for tiltaksgruppe og sammenligningsgruppe

I figur 4.2 sammenligner vi tiltaksskoler som har vært med i forsøket med én karakter, de som har deltatt med to karakterer, og de som ikke har vært med i forsøket. Vi observerer at alle tre grafene følger hverandre noenlunde likt i perioden før forsøket ble satt i verk, en bekreftelse på at gruppene antakelig er sammenlignbare. Vi observerer også at forskjellen mellom grafene verken ser ut til å økes eller reduseres i etterkant av forsøket, som illustrerer at forsøket sannsynligvis ikke har hatt en effekt på karaktersnittet i norsk.



Figur 4.2: Grafisk fremstilling av forskjell-i-forskjeller analysene på gjennomsnittlig standpunkt-karakter i norsk, for tiltaksgruppe med en karakter, tiltaksgruppe med to karakterer og sammenligningsgruppe

Med andre ord indikerer de grafiske analysene at det ikke er noen forskjell i karaktergjennomsnitt mellom tiltaksskoler og sammenligningsskoler, og at det heller ikke er noen forskjell når vi skiller mellom de som har hatt forsøk med én karakter eller med to karakterer, sammenlignet med å fortsette som før med tre karakterer.

4.5 Regresjonsanalyse med forskjell-i-forskjeller metode

Selv om de grafiske analysene ikke viser noen tydelig sammenheng mellom deltakelse i forsøket og karakternivå, kan det likevel være at forsøket har hatt en effekt på dette utfallet. For å undersøke dette har vi benyttet oss av forskjell-i-forskjeller metoden. Vi inkluderer skolefaste effekter som tar hensyn til skolekjennetegn som ikke endrer seg over tid. Vi kontrollerer også for årsfaste effekter, noe som vil si at vi tar hensyn til trenden vi har sett over tid, hvor karakterer jevnt over stiger i hele perioden vi undersøker.

I den første modellen (Modell 1 i tabell V1) har vi kun kontrollert for om eleven har fritak i sidemål. Dette er først og fremst fordi et fritak vil si at gjennomsnittskarakteren i norskfaget er basert på to og ikke tre karakterer. Regresjonsresultatet viser ingen effekt av forsøket i denne modellen. Samtidig finner vi positive signifikante årseffekter fra 2011, som bekrefter mønsteret vi så i de grafiske analysene som vil si at karakternivået stiger noe i perioden vi ser på (men uavhengig av forsøket).

I den andre modellen (Modell 2 i tabell V1) har vi inkludert grunnskolepoeng som kontrollvariabel, hvilket det er god forskningsmessig dekning for at er sterkt korrelert med karakterer i videregående (Byrhagen, Falch & Strøm 2006, Markusen, Flatø & Reiling 2017). Når vi inkluderer grunnskolepoeng er det om lag 25 000 elevobservasjoner som faller ut fordi vi ikke har data om grunnskolepoeng knyttet til disse elevene. Dette er også i all hovedsak elever som startet på videregående i 2007. Resultatet av denne regresjonen viser fortsatt ingen effekt av forsøket på elevers snittkarakter i norskfaget. Samtidig finner vi at jo flere grunnskolepoeng, desto høyere karaktersnitt i norskfaget i Vg3, som er i tråd med tidligere forskning. I tillegg øker forklaringskraften til regresjonsanalysen betraktelig når vi inkluderer grunnskolepoengvariabelen, hvilket gir en sterk indikasjon på at denne modellen er mer presis enn en modell uten denne kontrollvariabelen.

I den tredje modellen (Modell 3 i tabell V1) har vi i tillegg inkludert utdanningsprogram og kjønn. Utdanningsprogram faller under tre kategorier: Idrettsfag, Musikk, dans og drama og Studiespesialisering. Vi finner fortsatt ingen effekt av forsøket på snittkarakterer i norskfaget. Samtidig finner vi at gutter har signifikant lavere karaktersnitt enn jenter, i tråd med etablert forskning på dette området. Både elever fra Musikk, dans, drama og Studiespesialisering har høyere karaktersnitt

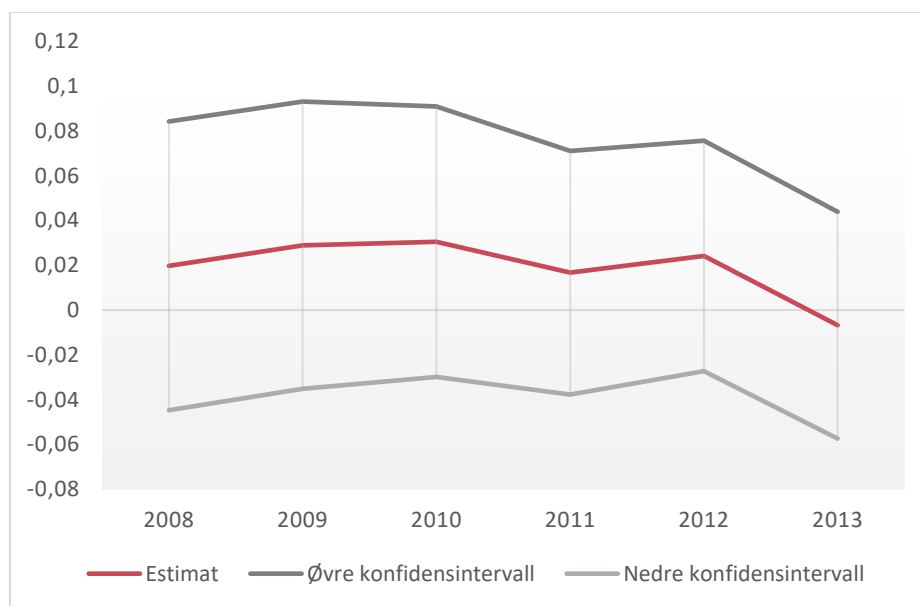
enn elever på Idrettsfag. Årseffekter er fortsatt positive og også de andre kontrollvariablene holder seg stabile når man sammenligner med modell 2.

Vi har i tillegg gjort tre subsample analyser. I den første av disse har vi undersøkt effekten av forsøket på kun gutter eller kun jenter (Modell 4.1 og 4.2 i tabell V1), for å se om det eventuelt kan være at forsøket har hatt ulik effekt på de to gruppene. I den andre subsample analysen har vi undersøkt effekten av å ha kun én karakter og å ha to karakterer (Modell 5.1 og 5.2 i tabell V1). I disse analysene finner vi heller ikke noen effekt av forsøket på karaktersnittet. Hverken gutter eller jenter har hatt noen signifikant effekt av forsøket, og det går heller ikke å identifisere om én av de to karaktersettingene har vært bedre på å øke prestasjonsnivået til elevene.

I den tredje subsample analysen har vi delt opp elever i fire kvartiler etter hvor mange grunnskolepoeng disse har, og undersøkt effekten av forsøket på hver av de fire undergruppene (Modell 6.1-6.4 i V2). Vi har også sett kun på elevene som var blant de 10 prosent lavest presterende i forhold til grunnskolepoeng (Modell 7 i V2). Slik kan vi undersøke om det er enkelte grupper elever som i større grad har hatt en effekt av å delta i forsøket. Vi klarer ikke å finne en effekt av forsøket for noen av disse undergruppene.

Vi har også gjort alle disse foregående analysene med kun eksamen i hovedmål, kun eksamen i sidemål, samt kun standpunkt i sidemål som utfallsvariabel, for å se om forsøket eventuelt har hatt en effekt på ett av disse målene (Tabell V3 til tabell V7). Vi finner ingen effekt av forsøket på noen av disse karakterene.

Funnet av at det ikke er noen signifikant effekt av forsøket kan også illustreres grafisk. Figur 4.3 viser en grafisk illustrasjon av effekten av forsøket. Den midterste linjen indikerer estimatet for de ulike årene, mens de grå linjene på hver side viser konfidensintervallet. Siden estimatet med konfidensintervall dekker området rundt 0 (null) viser figuren at det ikke er noen signifikant effekt på gjennomsnittlig standpunktkarakter i norsk å være en del av forsøket, sammenlignet med å ikke være det.



Figur 4.3: Grafisk visualisering av resultatet av forskjell-i-forskjeller analysen, effekt av å være en del av forsøket på gjennomsnittlig standpunktkarakter i norsk i Vg3. Estimert basert på modell 1 i vedleggstabell V1.

4.6 Oppsummering og konklusjon

Analysene av karakterdata viser at det ikke er noen forskjell i karaktergjennomsnitt i norsk i Vg3 mellom tiltaksskoler og skoler som ikke har deltatt i forsøket, verken når vi ser på sammenhengen grafisk eller når vi gjennomfører regresjonsanalyser der vi kan ta hensyn til ulike kontrollvariabler. Dette indikerer at forsøket med én eller to karakterer i stedet for tre karakterer i første og andre klasse av videregående opplæring ikke har noen effekt på karaktersnittet i norsk som elevene oppnår i Vg3.

Vi har også gjennomført analyser separat for gutter og jenter, samt sammenlignet de best presterende (målt i grunnskolepoeng) med de svakest presterende elevene og finner da heller ikke noen signifikant effekt av å ha vært med i forsøket. Analysene er også kjørt med standpunktkarakter i sidemål, samt eksamenskarakter i hovedmål og sidemål som utfallsmål. Ingen av disse analysene viser en signifikant effekt av å ha deltatt i forsøket.

5 Skoleledernivåets erfaringer fra forsøket

I dette kapittelet ser vi nærmere på hvordan skoleledere vurderer forsøket med redusert antall karakterer ved halvårsvurdering i norsk og hva slags erfaringer de har gjort seg underveis. Til dette formålet har vi hentet data fra to kilder: 1) en spørreundersøkelse blant skoleledere i videregående skole, 2) intervjuer med et utvalg representanter for skoleledelsen ved fire videregående skoler og to ungdomsskoler som har deltatt i forsøket.

Skoleledernes forventninger til forsøket ble analysert i delrapport 2 i denne evalueringen (Seland m.fl. 2016). I dette kapittelet sammenligner vi skoleledernes forventninger med hvordan de vurderer utbyttet av forsøket. I delrapport 2 brukte vi også forskningsintervjuer med representanter fra skoleledelsen på fire videregående skoler til å belyse hvordan skoler har lagt til rette for forsøket samt hvilke erfaringer de hadde gjort seg på tidspunktet da forsøket hadde vart i omlag tre år.⁹ I dette kapittelet supplerer vi intervjudata om erfaringene fra videregående skoler med nye forskningsintervjuer i ungdomsskolen, gjennomført våren 2017.

Fremgangsmåte i spørreundersøkelse og intervju med skoleledere er presentert i kapittel 1 i denne rapporten.

5.1 Utvalg og svarprosent

For å hente inn data om skoleledernes erfaringer med forsøket med redusert antall karakterer ved halvårsvurdering i norsk har vi brukt Utdanningsdirektoratets faste spørreundersøkelse «Spørsmål til Skole-Norge» (se kapittel 1 for nærmere presentasjon).

«Spørsmål til Skole-Norge» høsten 2015 ble sendt til 130 videregående skoler, og deltakelsen var på 74 prosent. Representativiteten for videregående skoler ble

⁹ Skoler over hele landet har kommet inn i forsøket på ulike tidspunkt. Intervjuene i vårt utvalg av skoler ble gjennomført etter at forsøket i disse skolene var i avslutningen av sitt tredje år, slik at kullet som startet med forsøket Vg1 det første året, var kommet til avslutningen av Vg3 på intervjutidspunktet. Dette for at skoleledelse så vel som lærere og elever kunne ha mulighet til å reflektere over erfaringer og sammenligne vurdering i og utenfor forsøket.

vurdert som god (Gjerustad og Waagene 2015). Våren 2017 var svarprosenten blant videregående skoler 65 prosent, og representativiteten ble også da vurdert som god (Federici, Gjerustad, Waagene, Larsen m.fl. 2017). Det var 82 skoleledere i videregående som svarte på spørsmålet om skolen deltok i forsøket med utprøving av én eller to karakterer i halvårsvurdering i norskfaget, og 38 av disse (46 prosent) svarte ja. Det var også fire skoler som svarte at de ikke deltok nå, men hadde deltatt tidligere. Disse valgte å ikke svare på spørsmål om hvorfor de avsluttet forsøket.

Videregående skoler som i «Spørsmål til Skole-Norge» våren 2017 svarte at de deltok i forsøket, danner grunnlaget for resultatene vi presenterer i dette kapitlet. I tillegg har vi koblet dataene fra 2015 med dataene fra 2017 for å undersøke i hvor stor grad det er de samme skolene som deltar i forsøket med én karakter i norsk som har svart på begge undersøkelsene. 33 av de 38 skolene fra 2017 svarte også i 2015. Det vil si at når vi sammenligner svarene fra 2015 med 2017, er det i stor grad de samme skolene som har svart på begge undersøkelsene.

Utvalgskriterier for valg av skoler hvor vi har gjort forskningsintervjuer er presentert i kapittel 1 i denne rapporten. Ved alle skolene sendte vi vår henvendelse med forespørsel om intervju til rektor. Ved alle de fire videregående skolene ble vår henvendelse videresendt til et faglig ledernivå med ansvar for norskfaget ved skolen. Det var også representanter for dette ledernivået vi fikk intervjuet i de fire videregående skolene. Ved de to ungdomsskolene har vi intervjuet rektorene, i det ene tilfellet i par med assisterende rektor. Når vi bruker materiale fra disse intervjuene omtaler vi alle disse skoleledelsesrepresentantene som «ledere» uten å skille mellom deres formelle posisjon ved skolen.

5.2 Implementering av én eller to karakterer ved halvårsvurdering

I 2015 fikk skolelederne i videregående skoler spørsmål om de hadde én eller to karakterer ved halvårsvurderingen i norskfaget. Basert på disse svarene og en oppdatert oversikt vet vi at 20 av skolene som svarte på undersøkelsen i 2017, har gjennomført forsøket med én karakter i norsk, mens 14 av skolene har hatt to karakterer i forsøksperioden. Tre skoler som har svart på spørreundersøkelsen har hatt en kombinasjon av disse to alternativene. Ved én skole har vi ikke klart å finne informasjon om hva slags alternativ som var valgt under forsøket.

Alle skoleledere som i 2017 svarte at de deltok i forsøket med én eller to karakterer, fikk spørsmål om de i løpet av pilotperioden hadde *endret* antallet karakterer som elevene i Vg1 og Vg2 får ved halvårsvurderingen. Av disse svarte 23 skoleledere ja, mens 15 svarte nei på dette spørsmålet. De 23 som svarte ja, fikk videre spørsmål om å utdype hva denne endringen besto i. I alt 16 ledere benyttet denne muligheten.

Vi har sammenlignet svarene disse 16 lederne gav om endring av antall karakterer, med hvordan lederne i 2015 hadde svart på spørsmål om hva slags vurderingsform de hadde valgt for halvårsvurdering. Denne sammenligningen viser at det er kun én videregående blant disse 16 som oppgir et svar som tyder på at de har endret vurderingsform i løpet av pilotperioden. Denne skolen har gått fra to karakterer til én karakter. De andre som har beskrevet endring i antall karakterer, har oppgitt samme antall som i 2015. På tross av spørreundersøkelsens spørsmål har vi derfor ikke valide svar på hvor hyppig slike endringer kan sies å ha forekommet under forsøket og hva som i tilfelle har begrunnet slike endringer.

Fra Utdanningsdirektoratet kjenner vi til at det også er andre videregående skoler som har endret antall karakterer under forsøket, men disse inngår ikke i utvalget i spørreundersøkelsen.

5.3 Forventninger og resultater av forsøket

I forskningsintervjuene med representanter fra skoleledernivået ved fire videregående skoler i forsøket understreket disse representantene hvordan redusert antall karakterer reduserer behovet for vurderingssituasjoner i faget. Lederne forklarte at denne situasjonen gjorde det mulig for lærerne å gå i dybden av undervisningsstoffet og være mer kreative i møte med elevene. Alle de fire lederne fulgte opp disse utsagnene ved å si at slike rammer for undervisningen gjør det mulig for lærerne å øke omfanget av vurderinger i faget. Hvis dette skal tolkes bokstavelig (reduisert antall vurderingssituasjoner gir mulighet for å øke omfanget av vurdering), fremstår det unektelig som et paradoks. Dette utsagnet, som vi også finner igjen hos lærerne vi har intervjuet (se senere i denne rapporten), er imidlertid blant våre hovedfunn i den kvalitative datainnsamlingen i denne evalueringen: norsklærne vi har intervjuet og deres faglige ledere trekker et konsistent skille mellom formelle vurderingssituasjoner med karakter og de løpende tilbakemeldinger som lærer gir elev som del av undervisningen. I intervjuene benevner imidlertid norsklærerne og deres ledere begge situasjoner som «vurdering». De ønsker seg redusert omfang av formelle vurderingssituasjoner hvor eleven får en begrunnet og dokumentert tallkarakter, og de vil bruke tiden som på denne måten frigis til å gi eleven flere tilbakemeldinger på arbeid som er i utvikling gjennom den løpende undervisningen. Disse tilbakemeldingene gis uten karakter, og har for lærerne form av «fremovermeldinger» om hva eleven kan gjøre bedre.

Vi fikk lignende svar i de to ungdomsskolene vi besøkte. Ved en av disse skolene fortalte lederen at lærerne hadde erfart at det å få karakter noen ganger kunne bli et hinder for læring for eleven. «[Elevene] får en karakter, og så blir de bare opptatt av denne, og ikke alle de gode tilbakemeldingene som læreren gir,» sa denne lederen. Forsøket med redusert antall karakterer ved halvårsvurdering gikk ved de to

ungdomsskolene parallelt med en lokal utprøving av å gi elevene færre vurderinger med karakterer i norsk gjennom skoleåret, og lederne så dette som to sider av samme sak.

For å kunne forstå hva det var norsklærerne ville *bort fra* gjennom forsøket, spurte vi lederne deres om hvordan disse skolene organiserte vurdering i norskfaget forut for forsøket med redusert antall karakterer. Vi skal kort se hva de svarte.

Ved den ene videregående skolen hadde de seks formelle vurderingssituasjoner hvert semester i norskfaget forut for forsøket, det vil si to vurderinger per disiplin i faget, som denne lederen uttrykte det. «Dette er en arv som ligger i kulturen her på skolen, sånn har vi alltid gjort det her,» la hun til. En av de andre skolene hadde tidligere fire formelle vurderingssituasjoner hvert halvår, som under forsøket var redusert til tre. En tredje skole hadde som generelt krav at i alle fag som vurderes med karakter, skulle lærer sørge for tre formelle vurderingssituasjoner hvert halvår. «Og i norsk, med tre karakterer, gir jo dette ni [vurderingssituasjoner],» sa lederen. Den ene ungdomsskolen vi besøkte hadde tilsvarende krav. Lederen sa her: «Du bør ha tre vurderingsprodukter som har såpass mange kompetansemål i faget i seg at du kan være trygg på at den karakteren du setter er riktigst mulig.»

Det var i hovedsak skolene som hadde bestemt hvor mange formelle vurderingssituasjoner de la til grunn for halvårskarakteren, i noen tilfeller etter råd fra skoleeier. Ved et par av de videregående skolene utdypet lederne skoleeiers krav til dokumentasjon av underveisvurderinger og halvårsvurderinger. Det vitner om et omfattende arbeid:

Vi er bedt om å dokumentere underveisvurdering, selv om det er borte fra forskriften. Det er dette som brukes hvis det blir klager. Skoleeier har bedt om det, og rektor har formidlet det videre. Alle underveisvurderinger legges inn i [elektronisk læringsplattform], og halvårskarakteren også. Så skriver vi en linje om hva som er gjort, for eksempel at vi har holdt en fagsamtale, for at eleven har fått den informasjonen hun trenger før hun får den karakteren. For vi pleier jo å ha samtale med elever før vi deler ut karakter.

Leder skole 3

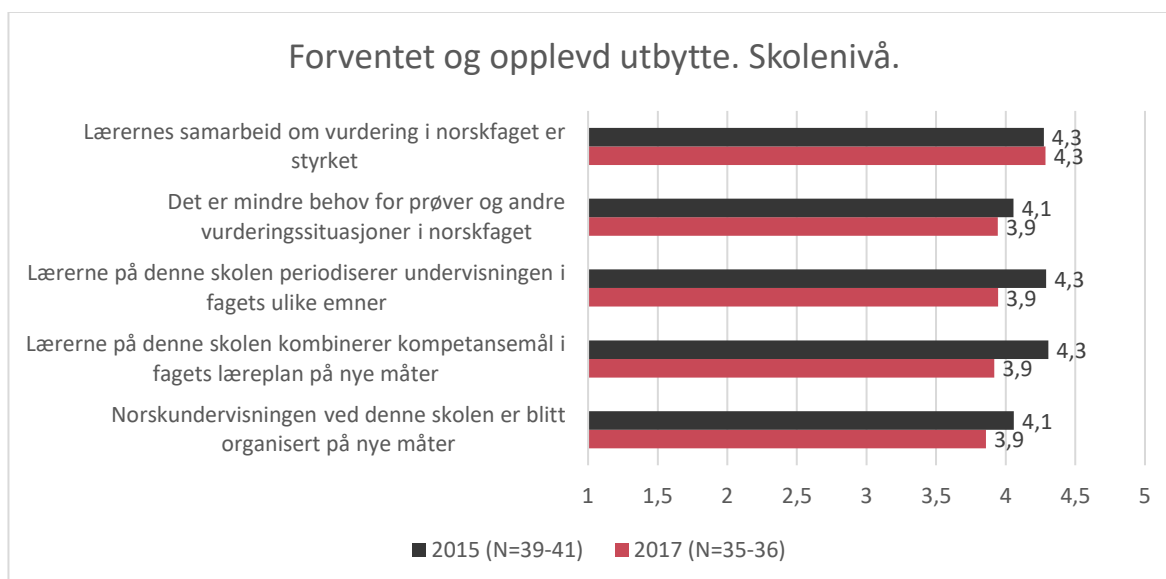
På bakgrunn av intervjuene med disse lederne tolker vi det som at man ved formelt å redusere antall karakterer ved halvårsvurdering, automatisk kan redusere antallet vurderingssituasjoner med karakter i faget *uten å rokke ved skolens opprinnelige bestemmelser og normer om antallet av slike vurderingssituasjoner*. Den videregående skolen som hadde ni vurderingssituasjoner med karakter i norsk hvert halvår forut for forsøket, kunne med én karakter i norsk redusere dette kravet med over 60 prosent uten å endre andre elementer i sitt regelverk. Resultatet var ifølge lederen betydelig økt tid og handlingsrom i undervisningen.

Redusert antall vurderingssituasjoner med karakter gir derfor ifølge disse lederne lærerne tid til å drive grundigere og mindre oppstykket undervisning. Integrert i denne undervisningen er en rekke situasjoner hvor eleven får vist frem arbeidet sitt til læreren, læreren får gi eleven tilbakemelding, og eleven kan (hvis hun er interessert og læreren etterspør det) få prøve seg på nytt og presumtivt bli dyktigere i faget uten at arbeidet nødvendigvis blir bedømt med en karakter. «Slik kan man kjøpe seg litt tid som kan brukes til å gi større variasjon,» sa lederen ved den ene ungdomsskolen. «Du trenger ikke disse stopp-punktene for å kunne sette en karakter, og du kan ha mer fokus på grunnleggende ferdigheter i faget.»

5.3.1 Utsagn om forsøkets virkning på skolenivå

I "Spørsmål til skole-Norge" i 2015 fikk skolelederne i videregående skole en rekke spørsmål om hva de forventet at den nye vurderingsordningen ville føre til. Da ble de spurt om hvor enige de var i at redusert antall karakterer ved halvårsvurdering i norsk ville føre til endringer i organisering og gjennomføring av undervisning og vurdering, samt elevers utbytte. I 2017 fulgte vi opp i spørreundersøkelsen med å spørre om redusert antall karakterer faktisk hadde ført til slike endringer. Dette gjør det mulig å undersøke forskjeller mellom forventet og opplevd utbytte, ved å sammenligne svarene knyttet til *opplevd* utbytte fra 2017 med svarene knyttet til *forventet* utbytte i 2015.

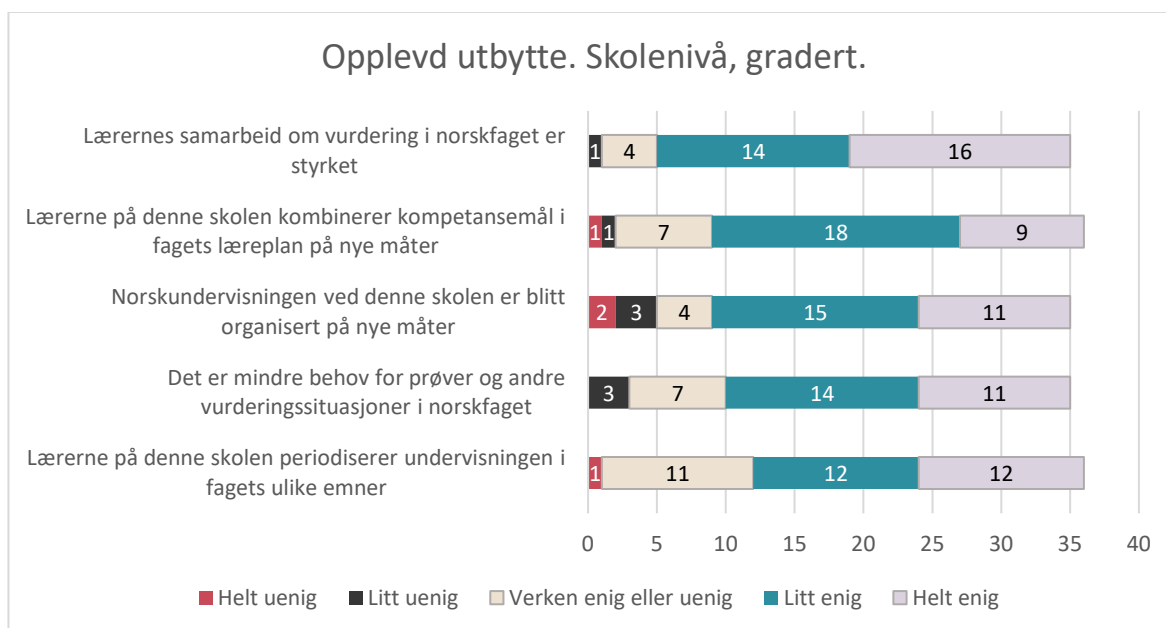
I figur 5.1 og 5.2 sammenligner vi svarene fra 2015 med 2017 ved å vise gjennomsnittlig skåre for hvert utsagn. Skalaen har blitt gitt følgende verdier: 1 «Helt uenig», 2 «Litt uenig», 3 «Verken enig eller uenig», 4 «Litt enig», 5 «Helt enig». Videre i analysen vil vi gå mer detaljert inn på svarfordelingene fra spørringene i 2017 angående opplevd utbytte, for å undersøke nærmere spredning i svarene. Figur 5.3 og 5.4 viser disse fordelingene.



Figur 5.1: Betydning av forsøket for organisering og gjennomføring av undervisning og vurdering. Gjennomsnittlig skåre.

Figur 5.1 viser resultatene av forventet og opplevd utbytte knyttet til organisering og gjennomføring av undervisning og vurdering. I figuren ser vi at lederne opplevde utbytte ved forsøksperiodens slutt ligger godt opp mot hva de forventet i 2015. Særlig ser vi at lederne opplever at lærernes samarbeid om vurdering i norskfaget har blitt styrket som resultat av den nye vurderingsformen. Det lave antallet respondenter gjør at ingen av forskjellene vi ser mellom forventet og opplevd utbytte er statistisk signifikante, men kan likevel tolkes som en indikasjon på hva skolelederne er fornøyde med i forsøket.

Intervjuene med i alt seks representanter fra skoleledernivået ved like mange skoler forteller ikke om særlig tilrettelegging eller organisatoriske endringer ved fagseksjon for norsk ved skolene. Ved disse skolene fikk vi inntrykk av at norsklærerne selv bestemte hvorvidt og i tilfelle hvordan de ville bruke fellestid til å snakke om spørsmål som vedrører forsøket, men at det i liten grad var gjenstand for eksplisitt faglig diskusjon mellom lærerne. Lærernes samtaler og samarbeid ville naturlig dreie seg om undervisning og vurdering, men ikke om forsøket med redusert antall karakterer i norsk i seg selv. En av disse lederne sa at redusert antall karakterer i norsk på et vis «presset frem» økt samarbeid mellom norsklærerne, fordi læreboka ikke lenger kunne styre verken undervisning og vurdering i samme grad som lærerne kanskje var vant til. Et par andre lot det skinne gjennom i intervjuet at lærerne fikk arbeide som før også under forsøket, det vil si at noen av dem var interessert i og fulgte opp i tett samarbeid med andre, mens noen lærere trivdes best med å jobbe alene. Forsøket i seg selv hadde altså ikke endret organisering eller samarbeidsformer mellom lærerne ved disse skolene i en grad som gjorde det naturlig å snakke om forskjeller «før» og «under» forsøket.



Figur 5.2: Betydning av forsøket for organisering og gjennomføring av undervisning og vurdering. Fordeling i antall. N=35-36. Fra spørringene våren 2017

Figur 5.2 viser resultater fra spørsmål om opplevd utbytte slik dette ble stilt i spørreundersøkelsen i 2017. Her ser vi hvordan lederne svar på opplevd utbytte fordeles seg over skalaen for utsagnene som knytter seg til organisering og gjennomføring av undervisning og vurdering. Vi ser at det er stor grad av enighet innenfor hvert av utsagnene, hvor de fleste har svart at de er litt eller helt enig i påstandene. Det er få som har svart at de er litt eller helt uenig i utsagnene, og vi kan heller ikke finne signifikante forskjeller i andel som har sagt seg enten helt eller litt enig i de ulike utsagnene. Dette tyder på at de fleste ledere er litt eller helt enig i at forsøket har gitt utbytte på alle disse områdene.

I intervjuene stilte vi lederne spørsmål for om mulig å få bedre innsikt i hvordan redusert antall karakterer ved halvårsvurdering i norsk påvirker undervisningen i faget. Et av spørsmålene dreide seg om hvordan norsklærerne arbeider med læreplanens kompetansemål for faget. Lederne vi intervjuet mente at det ikke var antallet karakterer i faget som styrte lærernes arbeid med kompetansemålene verken før eller under forsøket, fordi de fleste av kompetansemålene uansett vil berøre hovedmål og sidemål, skriftlig og muntlig. Derfor har det alltid vært nødvendig for lærerne å kombinere kompetansemålene i undervisning og vurdering. Som en av lederne sa: «Det er jo ikke sånn at du [som lærer] deler opp kompetansemål og sier at 'det skal jeg vurdere i sidemål og det skal jeg vurdere i hovedmål.' Problemet med kompetansemål er jo like gjerne at ett mål kan være kjempesvært, som at det er mange mål.»

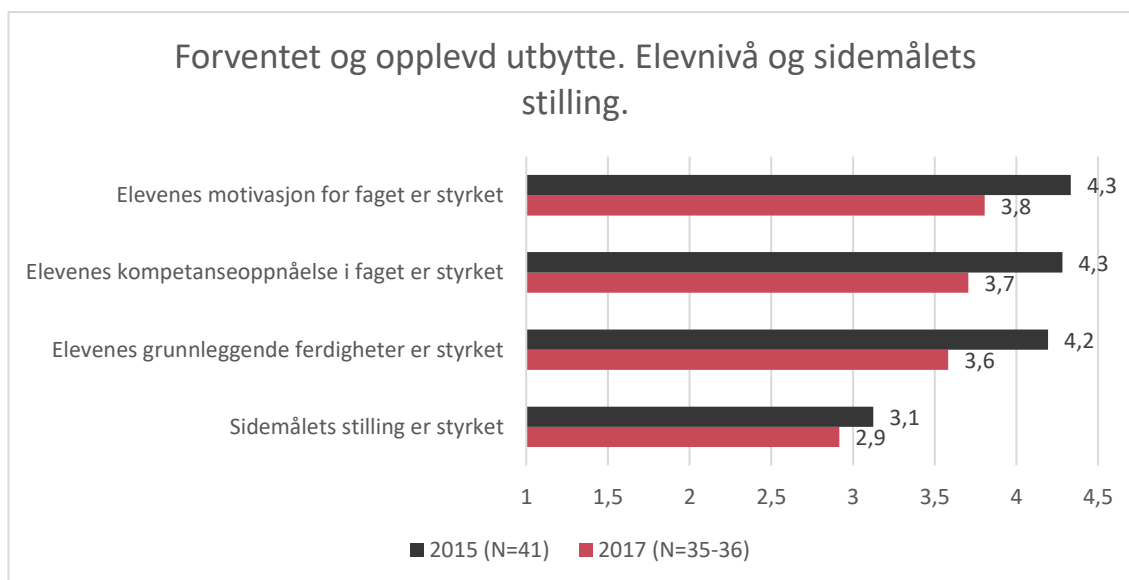
En annen leder fortalte levende om en norsklærers samarbeid med historielæreren, hvor skriving i norsk, litteraturhistorien i norskfaget og undervisning om en

særsilt periode i historie ble kombinert på nye og kreative måter. Det var ifølge denne lederen tiden som ble frigjort i undervisningen i norskfaget under forsøket som gjorde det mulig for lærerne å arbeide tverrfaglig på denne måten. Ulike kompetansemål og vurderingskrav på tvers av fag (ikke bare innen norskfaget) gjorde det ifølge lederen generelt vanskelig for lærere å samarbeide på fruktbare måter.

Lederne ga også eksempler på at lærerne i større grad kunne periodisere emner i norskundervisningen når de ikke lenger måtte bruke tid på å skaffe formelt vurderingsgrunnlag i samme utstrekning som tidligere. En av de videregående skolene og begge ungdomsskolene brukte periodisering for å konsentrere opplæringen i målformene, men ellers var det særlig skriving som ble nevnt når lederne snakket om periodisering. I disse utsagnene ble periodisering av emner koblet direkte til endring i vurdering, fordi læreren nå kunne kommentere og gi eleven tilbakemeldinger i prosessen frem mot ferdig produkt i en ubrutt arbeidsperiode. «Når vi nå har kuttet en vurderingssituasjon [med karakter], har vi økt på med revidering [av elevtekster],» sa en av lederne. «Før sa vi jo at elevene måtte revidere, og så gikk vi rundt og hjalp dem litt, men det var det. Vi har ikke mer retting enn før, men mer av det vi retter, er revisjoner.»

5.3.2 Forsøkets virkning på elevnivå og sidemålets stilling i faget

Disse ledernes utsagn om undervisning og vurdering glir naturlig over i hva lederne håper eller mener kommer ut av forsøket med hensyn til arbeidsmengde, motivasjon og læring. I figur 5.3 viser vi hvordan skoleledere i videregående skole fordeler seg i svar på fire spørsmål om elevers motivasjon og læring under forsøket med redusert antall karakterer i norsk



Figur 5.3: Elevenes utbytte og sidemålets stilling. Gjennomsnittlig skåre.

I figur 5.3 viser vi resultatet av skoleledernes forventninger og opplevd utbytte av forsøket knyttet til elevenes resultater og sidemålets stilling i faget. Vi ser at for utsagnene som handler om elevenes motivasjon, kompetanseoppnåelse og grunnleggende ferdigheter er skoleledernes vurdering av utbytte større enn hva som er tilfelle for skoleledernes vurdering av hvorvidt sidemålets stilling er styrket i norskfaget. Vi ser at dette samsvarer godt med hva slags forventninger skoleledere hadde ved begynnelsen av forsøket. Dette tyder på at skolelederne i videregående skole i liten grad opplever at forsøket med redusert antall karakterer ved halvårsvurdering i norskfaget som et tiltak som styrker sidemålets stilling, og at dette heller ikke i stor grad var et forventet utfall av ordningen da ledere ble spurt om dette i 2015.

Syn på forsøkets betydning for sidemålets stilling viste seg å være helt annerledes i de to ungdomsskolene. Et år tidligere hadde vi som forskere erfart at det var vanskelig å få ledere og lærere i videregående skole i tale om sidemålet under forsøket (se delrapport 2, Seland m.fl. 2016). Her var forskjellene mellom de to skoleslagene betydelig og derfor interessant. I intervjuene fortalte lederne ved begge ungdomsskolene at forsøket passet godt med den faktiske situasjonen for sidemålsundervisningen ved skolen, idet det ved disse ungdomsskolene hadde vært vanlig å vente med å introdusere elevene for sidemål. Det vil si at disse elevene ikke hadde undervisning i sidemål første halvår på ungdomsskolen, men at dette ble et større og tydeligere innslag i undervisningen gjennom andre halvdel av 8. klasse. Forut for forsøket ga derfor én av disse skolene elevene sine én skriftlig karakter og én muntlig karakter i norsk det første halvåret, og ventet med halvårskarakteren i sidemål til avslutningen av 8. trinn. Ved den andre skolen ble ikke elevene introdusert for sidemål og halvårsvurdering i sidemål før på 9. trinn. Forsøket med to karakterer i halvårsvurdering i norsk videreførte og legitimerte derfor denne praksisen ved disse skolene.

Det er ikke opp til oss å vurdere om denne tidligere praksisen ved de to ungdomsskolene har vært i samsvar med lovverket. Relevant for denne evalueringen er det imidlertid at begge disse lederne mente at én skriftlig og én muntlig karakter ved halvårsvurdering i norsk ga lærerne mulighet til en slik forsinket introduksjon av sidemål i undervisningen (og vurderingen), og at dette gagnet elevene i oppstarten av ungdomsskolen og i tilvenningen til det å få karakterer. Den ene av disse lederne begrunnet dette ytterligere med at norskfaget i seg selv er så stort og omfattende. «Elevene har vel hatt noe sidemål på barneskolen, men det er vel ikke noen stor del av norskfaget der,» sa denne lederen. «Så det å kunne starte [det første] halvåret med å kunne komme inn i ting uten at vi hadde trengt å tenke på nynorsken – hvis du spør norsklærerne, tror jeg de ville sagt at det hadde vært greit.» Denne lederen la til at elevene likevel kunne introduseres for sidemål på et

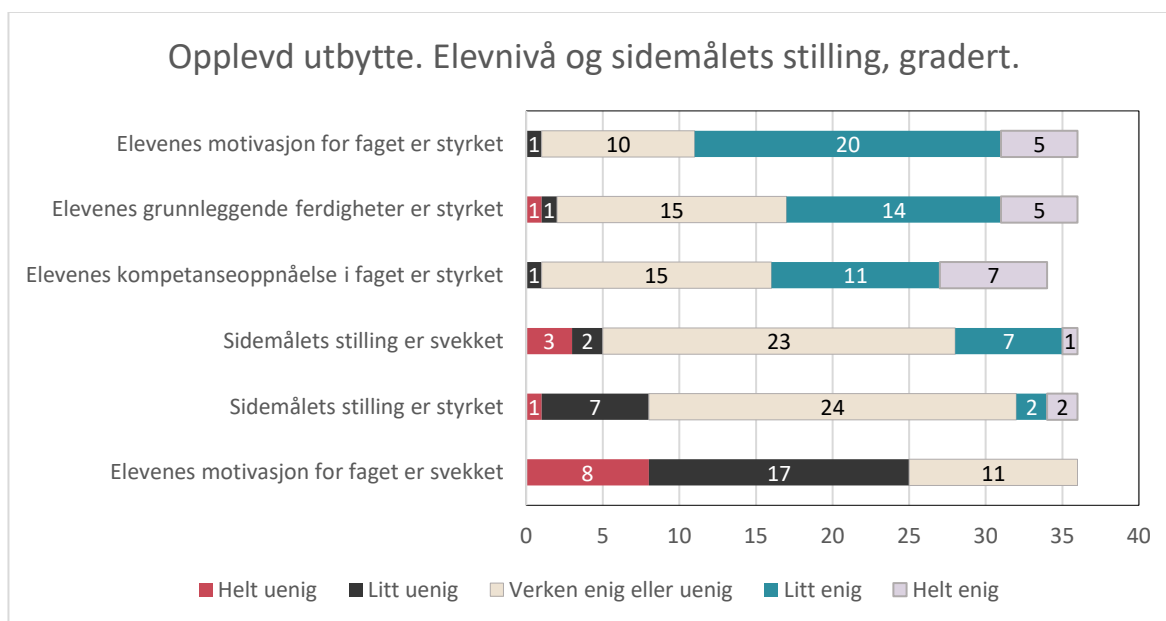
tidlig tidspunkt, for eksempel gjennom å lese tekster på nynorsk. Dette kunne læreren føle seg frem til selv, når han eller hun ikke trengte vurderingsgrunnlag for å sette en egen halvårskarakter i sidemål.

Vi spurte disse to skolelederne hvordan denne praksisen sikrer at alle kompetansemål i sidemålsundervisningen blir dekket. Den ene av disse lederne viste til skolens fagplan for norskfaget, og sa at forsøket hadde aktualisert fagplanen og alle norsklærernes ansvar for at kompetansemålene i faget ble dekket. «Vi tok en ordentlig diskusjon om hvordan man skal kunne få [sidemålet] inn i faget og heve elevene,» sa denne lederen. Ved den andre skolen fremhevet lederen at det å skrive naturlig nok foregår i begge målformer, og at økt trening på skrivning og innsats for det å heve elevenes skrivekompetanse følgelig ville kunne bedre elevenes læringsutbytte både i bokmål og sidemål. Lærerne på denne skolen kombinerte målformene i flere av vurderingssituasjonene, slik at elevene skulle besvare én oppgave på bokmål og én oppgave på sidemål. Denne lederen utdypet:

Jeg synes lærerne har vært flinke til å se på elevenes skrivekompetanse. Om du uttrykker deg på sidemål eller på bokmål, så finnes det fellesnevnerne som forteller om elevens skrivekompetanse. (...) Det å være opptatt av det å skrive – at man ser på innhold og oppbygning og ikke bare terper på sterke og svake verb og substantivbøyning.

Leder, skole 5

Samtidig nevnte denne lederen at en lærer som hadde overtatt norskundervisning på 10. trinn for elever som hadde deltatt i dette forsøket, hadde formidlet at elevene ikke var på det ferdighetsnivået i sidemål som denne læreren hadde forventet. Lederen selv understreket at alle norsklærerne ved skolen var bevisste på at sidemålet ikke skulle være noe man «la bort,» og at alle var enige i at det skulle leses og skrives i begge målformer. Spørsmålet denne lederen likevel stilte seg selv i intervjuet, var om det nå under forsøket ble skrevet like mye på sidemålet i 9. som før forsøket. «Så lærerne er spent på eksamensresultatene,» sa denne lederen.



Figur 5.4: Elevenes utbytte og sidemålets stilling. Fordeling i antall. N=34-36. Fra spørringene våren 2017.

I figur 5.4 viser vi svarfordelingen til utsagnene om utbytte som knytter seg til elevene, slik skolelederne i spørreundersøkelsen våren 2017. Her er også to nye utsagn om erfaringer med forsøket inkludert. Skolelederne ble bedt om å svare på om de mener at elevenes motivasjon for faget er *svekket* under forsøket, og tilsvarende om sidemålets stilling i faget er svekket i samme periode. Disse påstandene ble inkludert i spørreundersøkelsen for å fange opp hvorvidt uenighet i om forsøket har styrket elevenes motivasjon og sidemålets stilling betyr at det er enighet i at det motsatte har skjedd, altså en svekkelse. Det er liten grunn til å tro at dette er tilfelle hva angår elevers motivasjon. Her er de fleste litt eller helt enig i at forsøket har styrket elevers motivasjon, mens ingen er enige i at motivasjonen har blitt svekket som følge av forsøket.

Derimot ser vi at for utsagnene knyttet til hvorvidt sidemålet er styrket eller svekket, er fordelingene ganske like. De som er uenige i at sidemålet er styrket, er også enige i at sidemålet er svekket som følge av forsøket. De fleste er likevel enten uenige eller svarer at de er «Verken enig eller uenig». Dette tyder på at ved enkelte skoler opplever ledere at redusert antall karakterer svekker sidemålets stilling, men at de fleste skoleledere har en opplevelse av at forsøket sidemålets stilling verken er styrket eller svekket som følge av forsøket.

Litt under halvparten av lederne har også svart at de verken er enig eller uenig i at elevenes grunnleggende ferdigheter og kompetanseoppnåelse i faget er styrket, mens noen flere er litt eller helt enig i dette.

Vi har også undersøkt om det er signifikante forskjeller mellom svarene til ledere ved skoler som har én karakter med de som har to karakterer ved halvårsvurdering i norsk. Vi finner ingen signifikante forskjeller i hvordan ledere har svart på noen av spørsmålene om opplevd utbytte av ordningen.

Hvis vi sammenligner svarene om opplevd utbytte knyttet til organisering, vurdering og gjennomføring av undervisning med opplevd utbytte hos elevene, ser vi et mønster der forsøket i stor grad vurderes som betydningsfull for lærernes undervisning, samarbeid og vurdering, mens det oppleves i noe mindre grad å ha hatt betydning for elevers motivasjon og faglig utvikling. Dette ser ut til å samsvare godt overens med forventningene ledere hadde ved begynnelsen av pilotperioden.

Til sist i spørreundersøkelsen ble skoleledere spurt om det var noe de ønsket å legge til i forbindelse med forsøket. Det var ti ledere som valgte å gi en kommentar, og tabell 5.1 viser en oversikt over disse svarene.

Tabell 5.1: Lederes kommentar om forsøket. N=10.

Er det noe annet dere ønsker å legge til når det gjelder resultater av forsøket?
Det fungerer veldig bra!
Vi ønsker dette som en varig ordning.
Ynskjer å halde fram. Mange gode didaktiske føremoner ved forsøket
Det er positivt med forsøk, men flere synes det er dumt at det må være egen karakter i sidemål i 3. klasse.
Det bør kun være én karakter i norsk til standpunkt og mulighet til å trekke ut hovedmål, sidemål og muntlig til eksamen.
Vi tror at Vurdering for læring i større grad vil bli integrert i norskfaget ved at elevene kun får en skriftlig karakter der sidemålet inngår og en muntlig karakter.
Sidemål burde inn på vg1 og vg2, og avsluttes på vg2
Det har skapt litt mindre "stress" i kvardagen for faget, seier nokre lærarar, medan nokre sier at dei får for lite skrivetrening ved at elevane har litt mindre innleveringar og dermed mindre skrivetrening.
Det har vært en utfordring å få alle lærere til å ta i bruk mulighetene forsøket har gitt.

Tabell 5.1 viser at flere av dem som har kommentert, uttrykker at forsøket har vært positivt og/eller at de ønsker at dette skal bli en varig ordning. Enkelte har gitt spesifikke forslag til hvordan de ønsker at karakterordningen skal videreføres. Et av disse utsagnene gjelder standpunkt- og eksamensordningen, som ikke omfattes av forsøket. To av utsagnene gjelder bestemmelser for sidemålet, men disse utsagnene omtaler heller ikke forhold som berøres av forsøket. Også disse utsagnene peker imidlertid mot hva skoleledernes erfaringer fra forsøket kan bety for hvordan resultatene av forsøket skal tolkes i lys av de samlede vurderingsordningene i norskfaget.

Av kommentarer som omhandler selve forsøket, viser to av dem til utfordringer ved gjennomføringen. En skoleleder skriver her at noen norsklærere er bekymret for at elevene får for lite skrivetrening når det er færre innleveringer, mens den andre lederen forteller at det har vært utfordrende å få alle lærere til å utnytte mulighetene forsøket har gitt.

5.4 Oppsummering

I dette kapitlet har vi sett på hvordan skoleledere vurderer utbytte av ordningen med én eller to karakterer i norsk ved midtveisvurderingen. Vi har sammenlignet hvordan ledere har svart på opplevd utbytte av forsøket i 2017 med hva de forventet av dette i 2015.

Vi har vist at de fleste skolelederne som har svart på spørsmål om forsøket i Utdanningsdirektoratets faste spørreundersøkelse er enige i at ordningen med to karakterer i halvårsvurderingen i norsk har gjort det enklere å organisere og gjennomføre undervisning og vurdering i faget. Vi har også vist at litt over halvparten av disse skolelederne er enige i at forsøket har ført til at elevenes motivasjon, grunnleggende ferdigheter og kompetanseoppnåelse i norskfaget er styrket under forsøket, men her er det også flere skoleledere som gir uttrykk for at de er i tvil om at forsøket har ført til endring.

Enkelte skoleledere mener at sidemålets stilling i norskfaget har blitt svekket som følge av forsøket med én eller to karakterer. Flertallet av skolelederne som har svart på spørreundersøkelsen, har imidlertid ikke noen sterk mening om dette. Intervjuer med skoleledere i ungdomsskoler viser at forsøket med to karakterer i halvårsvurdering i norsk samstemmer med hvordan disse skolene praktiserte undervisning og vurdering i sidemål på 8. trinn også forut for forsøket.

Blant skolelederne som har benyttet muligheten til å kommentere på resultater av forsøket, uttrykker de fleste en tilfredshet og ønske om å videreføre ordningen.

6 Norsk lærernes erfaringer fra forsøket

Som ledd i evalueringen har vi besøkt fire videregående skoler og to ungdomsskoler som deltok i forsøket med redusert antall karakterer i norsk. Ved hver skole har vi intervjuet en gruppe norsklærere for å finne svar på følgende forskningsspørsmål:

- Hvordan arbeider lærerne ved forsøksskolene med kompetansemålene i læreplanen sammenlignet med dagens ordning med tre karakterer?
- Hvordan vektlegger lærerne ved forsøksskolene arbeidet med skriftlig sidemål, skriftlig hovedmål og muntlig norsk?
- Fører færre karakterer i norsk til endret handlingsrom for lærerne i deres mulighet til å organisere undervisningen og vurderingen på en hensiktsmessig måte?
- I hvilken grad og på hvilken måte utnyttes et eventuelt økt handlingsrom?
- Fører færre karakterer ved halvårsvurdering i norsk til at lærerne opplever det som lettere eller vanskeligere å organisere undervisning og vurdering i faget?
- Fører færre karakterer ved halvårsvurdering i norsk til at lærere opplever dette faget som mer eller mindre motiverende?

I delrapport 2 fra evalueringen (Seland m.fl. 2016) presenterte vi lærerintervjuene fra videregående skole i full bredde. Dette kapittelet er en noe forkortet versjon av intervjuanalysen fra delrapporten, supplert av lærerintervjuene fra ungdomsskolene med særlig vekt på sidemålets stilling i norskfaget under forsøket.

Ungdomsskolene og de videregående skolene er besøkt på forskjellige tidspunkt. Vi gjennomførte intervjuene på de videregående skolene våren 2016, da disse skolene hadde vært del av forsøket i minst to hele skoleår og var i ferd med å avslutte det tredje skoleåret med eksamen for elevene som hadde vært del av forsøket siden Vg1. Ungdomsskolene kom i gang med forsøket først i 2014, og vi planla derfor tidspunktet for intervjuene i ungdomsskolene til våren 2017 for å kunne snakke med lærerne der om tilsvarende erfaringer med forsøket i avslutningen av tredje skoleår. Ved den ene av disse ungdomsskolene viste det seg imidlertid at de hadde kommet i gang med forsøket enda et år senere, slik at det første kullet av

elever i forsøket der var i ferd med å avslutte 9.trinn da vi kom på besøk. Det ga likevel god mening å intervju disse lærerne om erfaringer med forsøket, da de hadde gjort seg mange tanker om dette. Ikke minst gjaldt disse refleksjonene forsøkets betydning for opplæringen og vurderingen i sidemål.

Grunnlaget for valget av de seks skolene er beskrevet i kapittel 1, hvor vi også har gjort rede for gruppeintervju som metode for datainnsamling. Til sammen har vi møtt og intervjuet 20 norsklærere ved disse skolene, fordelt på grupper som besto av mellom to og seks personer. Intervjuguiden som er brukt under intervjuene, ble trykt som vedlegg til delrapport 2 fra evalueringen.

6.1 Vurdering *for* læring og vurdering *av* læring i norskfaget

Intervjuguidene som ligger til grunn for datainnsamlingen i denne rapporten, skiller mellom undervisning og vurdering i norskfaget. Slik vi først oppdaget gjennom intervjuene med faglederne i norsk (se tidligere i denne rapporten), gjelder ikke dette skillet like sterkt for norsklærerne selv. I samtaler med norsklærerne gled utsagnene om undervisning og vurdering over i hverandre fra første stund, og illustrerer hvordan disse to aktivitetene integreres med hverandre gjennom lærernes daglige møter med elevene i klasserommene.

Til gjengjeld finnes det i hvert eneste av de seks gruppeintervjuene et innforstått skille mellom bestemte typer av vurderingsformer. Med «innforstått» mener vi at alle norsklærerne i intervjuene bruker begrepet «vurdering» om alt som dreier seg om å gi eleven tilbakemeldinger – med og uten karakter – med og uten innlevering – forberedt og uforberedt – til tross for at de i sine *forståelser* av begrepet opererer med skarpe skiller mellom ulike typer vurderingssituasjoner.

Ved alle de seks skolene hadde det gjennom flere år pågått et arbeid for å utvikle lærernes vurderingspraksis enten inspirert av eller konkret i tråd med prinsippene for «Vurdering for læring». Ikke alle skolene hadde formelt deltatt i Utdanningsdirektoratets satsing med samme navn, men i alle gruppeintervjuene kom lærerne med utsagn som bekreftet for oss at det blant lærerne eksisterte en oppmerksomhet om vurdering og en mer eller mindre systematisk innsats for utvikling av det faktiske vurderingsarbeidet ved skolen.

Alle lærerne vi snakket med, viste i intervjuene at de var oppmerksomme på et skille mellom vurdering *for* læring og vurdering *av* læring. Enkelte av dem snakket også om vurdering *som* læring. Dette er nyanser og forståelser av vurderingsbegrepet som er knyttet til prinsippene for god vurdering i satsingen «Vurdering for læring». Særlig skilte lærerne ved de fire videregående skolene i sine utsagn klart mellom vurdering *for* læring og vurdering *som* læring på den ene siden, og vurdering *av* læring på den andre siden. I undervisningen vil de drive vurdering *for* og *som* læring, mens vurdering *av* læring ifølge flere av disse lærerne har mindre reell

verdi i norskfaget på Vg1 og Vg2 i studieforberedende utdanningsprogram.¹⁰ Vurdering *av* læring på disse årstrinnene legger ifølge lærerne beslag på knappe tidsressurser i opplæringen,¹¹ som går på bekostning av tid til vurdering *for* og vurdering *som* læring. Derfor vil lærerne gjerne at vurderingssituasjoner *av* læring finner et nødvendig minimum. Vi møtte lignende utsagn fra lærerne i ungdomsskolene vi besøkte.

Det lærerne uttrykker gjennom gruppeintervjuene, er at for dem er det å gi eleven såkalte «fremovermeldinger» en naturlig, nødvendig og ønsket del av undervisningen time for time og dag for dag i klasserommet. «Fremovermeldinger» er et begrep som særlig er knyttet til satsingen «Vurdering for læring»¹² og prinsipper for bedre læring basert på å gi eleven tilbakemeldinger underveis i arbeidet fremfor å vurdere elevens resultater isolert fra denne arbeidsprosessen. «Fremovermeldinger» peker derfor ut over hva eleven faktisk har prestert på et gitt tidspunkt i opplæringen, og inneholder en anbefaling om hvordan dette arbeidet kan utvikles med tanke på ytterligere måloppnåelse i læreplanen. Dette forklarer hvorfor undervisning og vurdering omtales som to sider av samme sak i lærerintervjuene. Slik vi vil komme tilbake til i kapittelet om elevnivået senere i denne rapporten, var imidlertid ikke lærerne konsekvente med hensyn til hva de kalte denne typen vurderinger. Ved en av ungdomsskolene ble en slik vurdering for eksempel kalt for «respons».

Hovedfunnet i lærerintervjuene uttrykkes enkelt ved et utsagn fra en lærer ved en av de videregående skolene, som sa at forsøket med redusert antall karakterer ved halvårsvurdering i norsk har redusert presset på antall vurderingssituasjoner som man må ha hvis man skal produsere en gyldig karakter. «Skal du ha et troverdig grunnlag for vurdering med karakter i hver disiplin, må du ha minst to vurderingssituasjoner, helst må du ha tre. Det er veldig låst, det gir lite frihet og lite ro,» sa en norsklærer ved en annen av de videregående skolene. Også ved ungdomsskolene hersket tilsvarende normer for hvor mange vurderingssituasjoner man må ha gjennom semesteret for å kunne sette gyldig karakter. Ved en av disse skolene beskrev norsklærerne om hvordan de måtte «kjempe om plassen» i elevenes prøveplan, hvor hensynet til elevene hele tiden måtte veies mot alle lærernes behov for å få nok vurderinger i alle fag. Når antallet vurderingssituasjoner med karakter

¹⁰ Til gjengjeld mener mange av disse norsklærerne at vurdering *av* læring i norsk har stor verdi for elevene på Vg3, fordi slik vurdering også underveis i skoleåret skal bidra i elevenes forberedelse til eksamen.

¹¹ På Vg1 og Vg2 i utdanningsprogram for studiespesialisering er det fire uketimer i norsk, mens det i Vg3 er seks uketimer i faget.

¹² Se Seland, Lødding og Prøitz (2015) *Delrapport 1 fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk* for en bredere presentasjon av "Vurdering for læring" i relasjon til forsøket.

gjennom semesteret i ett av fagene på denne måten reduseres, frigjøres tid i prøveplanen og i uketimen som er satt av til faget. Denne tiden kan ifølge lærerne brukes til tettere integrasjon mellom undervisning og «fremovermeldinger» i faget.

6.1.1 Eksempler på integrasjon mellom undervisning og vurdering i norskfaget

Når norsklærerne i intervjuene snakker om undervisning og vurdering i norskfaget, er det nesten alltid skriving de bruker som eksempel. God skriveopplæring krever ifølge disse lærerne en kontinuerlig prosess med mange øvingssituasjoner for elevene, og som gir dem mulighet til å trene på å skrive varierte typer tekst.

Øving i skriveopplæring i norskfaget består av lærerens undervisning om teksttyper, elevenes tilnærming til teksttypene gjennom å skrive, det å gi elevene tilbakemeldinger på tekstutkastene under arbeidet og elevenes revidering av tekstutkastene for eventuelle nye tilbakemeldinger fra lærer eller medelever. «Jeg lar min Vg2-klasse skrive ganske mye uten at jeg gir dem karakterer,» sa en lærer. «Det er rom for å levere flere utkast, rom for at elevene leser hverandres tekster, til at elevene gjør egenvurdering av arbeidet,» sa en kollega på samme skole. Lærerne på de øvrige skolene hadde lignende utsagn og ambisjoner for skriveopplæringen.

Gjennomgående fortalte lærerne at de under forsøket kan la elevene arbeide med skriving over lengre tidsperioder enn hva de kunne med tre karakterer ved halvårsvurdering i faget. De kan la elevene prøve seg på flere typer tekst flere ganger og med flere mindre innleveringer enn hva som tidligere var mulig, og de kan variere mer mellom kortsvar- og langsvarsoppgaver. Lærere fortalte også at de under forsøket i større grad kan la elevene øve på avgrensede aspekter ved skriveprosessen, for eksempel det å lage gode avsnitt eller det å lage en god innledning.

Lærere på flere av skolene snakket om at skriveopplæring og skrivetrening krever større oppmerksomhet i undervisningen nå enn før. Dette forklarer de med at dagens elever både leser og skriver mindre enn hva som var tilfelle for 10-15 år siden. «[Elevene] har ikke referansene som kreves for å lese en tekst, og det gjør at vi må nærme oss teksten på en annen måte. Vi må snakke om andre ting [enn vi gjorde før] når elevene skriver,» sa en lærer ved en videregående skole. En kollega av denne læreren fulgte opp: «Elevene sier mye oftere i dag at 'jeg leser aldri.' Mens noen sier at de leser mye. Det er et stort skille der.» Lærerne i ungdomsskolen var i likhet med lederne deres (se forrige kapittel) opptatt av at for elever som begynner på ungdomsskolen, er det svært mye nytt. Ikke minst er det å få karakterer noe disse elevene aldri har prøvd. For undervisningen i norsk på særlig 8. trinn var derfor disse lærerne opptatt av at elevene skulle få en rolig tilvenning til det å bli jobbe med faget på måter som leder frem mot vurdering. Det å la elevene få lese tekster

var for ungdomsskolelærerne viktig som introduksjon til det å skrive. En lærer på en av de videregående understreket det samme poenget da hun sa i intervjuet at all skriving har utgangspunkt i tekst.

Gjennom særlig ett av intervjuene på en videregående skole fikk vi en levende fortelling om hva som kan skje når det lærerne opplever som press om antall vurderingssituasjoner tas bort som følge av at antall karakterer ved halvårsvurdering reduseres. På denne skolen fortalte læreren om forfatterbesøk i klasserommet og elevenes «fortsettelsesskriving» på forfatterens nye romanidé. Elevene skrev dikt og fikk diktene vurdert i «Diktkammeret» i Dagbladet. Elevene laget «fotostory» om barnebokforfattere og en filmfortelling. I et tverrfaglig prosjekt med historielæreren hadde elevene arbeidet med mediakompetansemål i norsk og en bestemt historisk hendelse. Elevene laget avisforsider som fortalte om hendelsen i «sann-tid». Historiske perioder ble også gjort relevant da elevene skulle skrive om situasjoner fra egen hverdag og bruke språklige elementer fra tekster som er bevart fra denne perioden. «Kunne du gjennomført dette hvis dere hadde hatt tre karakterer ved halvårsvurdering i norsk?» spurte vi denne læreren i intervjuet. Læreren svarte:

Å nei. [...] Det å ha tre karakterer [ved halvårsvurdering], det er som en tvangstrøye for å få nok vurderingssituasjoner. Da må du tenke vurderingssituasjoner med én gang. Nå kan undervisningsopplegget utvikle seg på en mer naturlig måte.

Norsklærer, skole 4

Forfatterbesøket, «fortsettelsesskrivingen», diktforsøkene og avisforsidene som nettopp er beskrevet, ble brukt av læreren for å vurdere og bidra til elevers forståelse, faglige utvikling og kompetanseoppnåelse. Gjennom disse undervisningsoppleggene produserte elevene store mengder tekst, men i alle disse situasjonene hadde læreren samtidig mange muligheter for å gjøre vurderinger av elevenes muntlige kompetanse.

Ved begge ungdomsskolene kunne norsklærerne fortelle at forsøket med redusert antall karakterer i norsk ga dem mulighet til å vente med å introdusere elevene på 8. trinn for sidemålet i norskfaget. Hvis elevene skulle hatt tre karakterer ved halvårsvurdering, måtte elevene hatt undervisning og vurdering i sidemålet gjennom høstsemesteret på lik linje med norsk hovedmål og norsk muntlig. Når disse skolene gjennomførte forsøket med henholdsvis én karakter i norsk skriftlig (hovedmål og sidemål) og én karakter i norsk muntlig, ble det opp til lærerne å velge når elevene skulle arbeide med sidemålet. Løsningen disse lærerne hadde valgt var å gi elevene en «myk» start på norskfaget på 8. trinn, og deretter innføre sidemålet gradvis i undervisningen og i vurderingssituasjonene enten på våren på 8. trinn eller først på høsten på 9. trinn. En av disse lærerne sa:

Fordelen [med forsøket] er at vi kan ta nynorsken litt mer gradvis. Vi forventer ikke at elevene skal kunne skrive en perfekt tekst på nynorsk på 8. trinn. Til [halvårsvurdering til] jul talte bokmål 90 prosent [av skriftligkarakteren] og sidemål 10 prosent. (...) Nå til sommeren spiller nynorsk en litt større rolle.

Norsklærer, skole 6

Årsakene til at disse ungdomsskolelærerne ønsker å vente med å introdusere elevene for sidemål, er sammensatt. Dels peker disse lærerne på at når elevene begynner på ungdomsskolen, representerer dette en stor overgang for dem sammenlignet med barneskolen. Mange av dem opplever at de får flere lekser, og det er vanlig at elevene har to formelle vurderingssituasjoner i løpet av en uke – for eksempel en prøve og en innlevering. Lærerne er oppmerksomme på at det kan ta tid for elevene å venne seg til de nye arbeidskravene, og hvis de kan lette denne tilvenningen med å vente med å innføre vurdering i sidemål, mener de at dette er til elevenes beste. Samtidig satte enkelte av lærerne spørsmålstegn ved elevenes forkunnskaper i sidemål basert på hva de hadde lært på barneskolen. De visste at sidemålet også inngår i kompetansemål for faget på 7. trinn, men hadde ofte opplevd at elevene ikke kjente til sidemål når de begynte på ungdomsskolen. Løsningen for disse lærerne ble derfor å la elevene nærme seg sidemålet gjennom å lese tekster uten at sidemålet ble gjort til gjenstand for vurdering den første tiden på 8. trinn.

På begge ungdomsskolene (8. trinn) og for en eller to Vg1-klasser på en av de videregående skolene falt forsøket sammen med en norm om å ikke gi tallkarakterer ved underveisvurdering før halvårsvurderingen. Det å gi vurdering uten karakter gjennom et helt semester er ikke et element ved selve forsøket, men kan likevel ses som en parallell eller en utvidelse av det handlingsrommet som lærerne forteller at forsøket gir dem. Når lærerne velger vurdering uten karakter, kan dette imidlertid også være av hensyn til elevene og et ønske om at elevene skal legge mindre vekt på tallkarakterer og mer vekt på lærernes «fremovermeldinger». Lærerne mener at dette kan være gunstig for elevene i en oppstarts- og tilvenningsfase, som 8. årstrinn eller Vg1.

6.1.2 Lærernes arbeid med kompetansemål i faget

Oppdragsgiver ønsker å vite hvordan lærerne arbeider med kompetansemålene i norsk under forsøket med redusert antall karakterer, og oppdragsgiver ønsker også å vite hvordan forsøket påvirker opplæringen i hovedmål og sidemål og balansen mellom skriftlig og muntlig som egne disipliner i faget. I det følgende behandler vi sidemål/hovedmål og skriftlig/muntlig i forlengelse av hvordan lærerne arbeidet med kompetansemålene i faget. Dette gjør vi fordi det i intervjuene blir klart at lærerne er mer opptatt av sammenhengene mellom disse disiplinene og læreplanen enn sammenhengene mellom disiplinene og selve forsøket. Dette gir

også mening idet læreplanen er den samme uansett hvor mange karakterer lærerne gir i norsk.

Alle lærerne fikk i intervjuene spørsmål om hvordan de arbeider med kompetansemål og med vektingen av sidemål/hovedmål og balansen mellom skriftlig/muntlig, men vi har ikke fått særlig utfyllende data om dette. Det kan skyldes at vi ikke var tilstrekkelig pågående i intervjusituasjonen, og at vi kunne ha fått lærerne til å fortelle mer inngående om dette ved å stille dem flere spørsmål eller å stille spørsmålene på en annen måte. En annen forklaring er imidlertid også mulig, og det er at lærerne under forsøket arbeider slik de dypest sett alltid har gjort, med unntak av at de har færre formelle vurderingssituasjoner av læring. Alle lærerne i utvalget vårt gir uttrykk for at de så hyppig som mulig inkluderer vurdering for læring i undervisningen ved å gi elever fortrinnsvis muntlige og løpende tilbakemeldinger. Vi forstår det som at de kan øke omfanget av slike tilbakemeldinger uten at dette i seg selv må planlegges i en separat prosess basert på kompetansemålene. Vi skal se på eksempler fra intervjuene som har ledet oss frem til en slik hypotese.

På den første videregående skolen spurte vi lærerne hvordan de arbeidet med kompetansemål under forsøket. I intervjuet ga vi lærerne utskrifter av kompetansemål for skriftlig kommunikasjon i Vg2, og vi hadde håpet at lærerne ville ta for seg ett og ett kompetansemål og forklare hvordan de planla undervisning i tråd med dette målet (se også kapittel 1, hvor vi beskriver og kommenterer bruken av disse utskriftene av læreplanmål som element i kvalitativ metode). Allerede ved innledningen av dette intervjuet ble det imidlertid klart at bare et par av lærerne som var til stede underviste på Vg2, mens de øvrige lærerne forholdt seg til kompetansemål på «sine» trinn, som var Vg1 og Vg3. Lignende sammensetninger av lærergruppene opptrådte i alle intervjuene vi gjennomførte i videregående skole, og Vg2-læreplanmålenes manglende relevans for samtalen ble understreket av at så godt som alle lærerne vi snakket med, hadde sitt nærmeste faglige samarbeid med kolleger på samme trinn. Derfor gjorde utskriftene våre lite nytte i seg selv i intervjuene, men bidro i en viss grad til å pense lærerne inn på emnet «kompetansemål» i samtalen.

På denne første videregående skolen svarte derfor en av lærerne at arbeidet med kompetansemål ikke endres nevneverdig av forsøket fordi kompetansemål ikke bare henger sammen med vurdering. «Det er jo hvordan man jobber med kompetansemålene i det daglige også, og der er det kanskje ikke så stor endring,» sa denne læreren. En kollega fulgte opp:

Det er jo ikke sånn at vi jobber med ett og ett kompetansemål, det er jo flere kompetansemål som går inn i en og samme oppgave, så det er vanskelig å si at før jobbet vi sånn og nå jobber vi slik. Eller, jeg ser ikke noen forskjell i hvordan jeg jobbet med kompetansemål [før og nå], for vi jobber med kompetansemålene både

med tre karakterer og med én karakter, men vi slipper å stresse den karaktersettinga like mye som før.

Norsklærer, skole 1

På neste skole fikk vi utsagn som minnet om det over, nemlig at forsøket i seg selv ikke endrer arbeidet med kompetansemålene. «En oppgavetype kan du prøve på alle disipliner, muntlig, skriftlig og sidemål,» sa en lærer på denne skolen, «det er ikke noen motsetning i dette.»

Læreren som hadde gitt oss en rekke eksempler på undervisningsopplegg (se over), utdypet for oss hvilke kompetansemål som ble dekket eller også kombinert i de ulike oppleggene, og forklarte: «Det er vurderingssituasjonen som legger lista for hvor hvilke metoder du kan bruke, så metodefriheten [i forsøket] gir oss muligheter til å trekke inn kompetansemålene på nye måter.» Denne læreren viste primært til læreboka for hvordan kompetansemål ble dekket innenfor de ulike undervisningsoppleggene, men understreket samtidig i intervjuet at det er læreplanen som styrer dette.

På en av de videregående skolene var det en lærer som overfor oss og kollegene utdypet hvordan læreplanmål og karaktersystem kan sies å tilhøre to tidsaldre i utvikling av norskfaget i skolen, og brukte dette som forklaring på hvorfor det synes å være en svak forbindelse mellom læreres arbeid med kompetansemålene og antallet karakterer som skal brukes ved halvårsvurdering i faget. Karaktersystemet basert på eksamensordningen er eldst av disse, forklarte denne læreren, og forholder seg i prinsippet til norsk som et ferdighetsfag, det vil si å bedømme hvor dyktig eleven er skriftlig og muntlig. Læreplanen følger ikke eksakt den samme inndelingen av kompetansemål i skriftlig, muntlig og sidemål, og gjør at man for eksempel tester mål fra kulturdelen av læreplanen som grunnlag for muntlig karakter, men også ved skriftlige prøver. Arbeid med kompetansemål og antall karakterer eksisterer derfor som parallelle idealer i vurderingsarbeidet, men blir forsøkt «knadd sammen», slik denne læreren uttrykte det.

I intervjuet fikk ikke denne læreren respons på dette utsagnet fra kollegene, og vi hørte ikke lignende utlegninger fra lærere på de andre skolene vi besøkte. Vi gjengir likevel denne forklaringen her, fordi vi synes det er en interessant observasjon, og at observasjonen peker på et mulig misforhold mellom læreplans kompetansemål og eksamensordning som idealer for vurdering i faget. Vi har imidlertid et spinkelt datagrunnlag for å utvide akkurat dette bidraget til diskusjonen, annet enn at de fleste av norsklærerne vi intervjuet, benyttet spørsmålet om arbeid med kompetansemål til å si at hovedproblemet her var at det er for mange kompetansemål i faget. I praksis, sa disse lærerne, blir ikke alle kompetansemålene gjenstand for vurdering. Dette ble ikke forandret av forsøket med færre karakterer ved halvårsvurdering, mente lærerne. Ved en av ungdomsskolene sa en av lærerne under intervjuet:

Jeg vil at vi skal jobbe grundigere med færre ting. Det er så mye vi skal gjennom, du får aldri tid til å bli god til noe. Elevene kan få dikttolkning til eksamen, men vi har bare hatt om det i to uker. Vi skulle nemlig også ha tid til å skrive en fagartikkel og en debattartikkel og en novelle og lære om reklame før vi hadde to uker til lyrikken.

Norsklærer, skole 5

Også spørsmål om vekting av hovedmål og sidemål og balanse mellom skriftlig og muntlig karakter ble i intervjuene på videregående skole besvart med noe distanse til forsøket med redusert antall karakterer i norsk. Dette til tross for at hvis forsøket skal gjennomføres med to karakterer, skal disse fremkomme som én karakter i skriftlige ferdigheter og én karakter i muntlige ferdigheter. Delingen mellom skriftlig og muntlig karakter ble av lærerne på skoler med to karakterer oppfattet som det naturlige i norskfaget, og ble sjelden utdypet i intervjuene. Unntaket var på en av skolene som praktiserer én samlekarakter i norsk, og hvor lærerne bemerket at vektingen mellom skriftlig og muntlig nok kunne variere fra lærer til lærer. I særlig ett av gruppeintervjuene på videregående skole var det flere kommentarer om sidemålet, som enten inngår i én samlekarakter i norsk, eller i karakteren som skal vise skriftlige ferdigheter. Denne skolen praktiserte kombinasjonen av én karakter på Vg1 og to karakterer på Vg2. Skolens norm var at sidemålet skulle utgjøre 50 prosent av skriftligkarakteren (på Vg2), men lærerne hadde i realiteten ikke verktøy for å sikre at det ble slik. Det ville nok bli opp til den enkelte lærer, sa disse lærerne i intervjuet. De fortalte samtidig at de foretrakk å gi elevene kontinuerlige «drypp» av hovedmål og sidemål fremfor å periodisere undervisning i en av målformene.

På den videregående skolen som hadde to karakterer i både Vg1 og Vg2, mente lærerne at når antallet formelle vurderingssituasjoner gikk ned, ga dette også mer tid til å arbeide grundigere med sidemålet. Også ved denne skolen var bokmål hovedmålet i undervisningen. «Det blir mer ro over skrivingen i begge målformer, og det blir tid til å bearbeide, følge opp og se en utvikling både for oss og for eleven,» sa en av lærerne ved denne skolen. «Når vi har bare to karakterer, kan ikke eleven 'blåse' i sidemålet, fordi de må arbeide med nynorsk for å få en god karakter i norsk skriftlig,» sa en kollega. «Jeg vil si at nynorsken blir oppgradert med to karakterer,» sa en tredje lærer som deltok i dette intervjuet, og utdypet:

Jeg føler at vi kan slappe litt av i undervisningen og få gjort ting ordentlig. Jeg behøver ikke å stresse dem for at de skal få en nynorsk karakter, de kan skrive noe på nynorsk uten å få karakter, og vi kan lese nynorske tekster uten at de skal få en egen karakter på det. Allikevel gjør vi jo begge deler, det er ikke slik at man sluntrer unna. Det blir bare mer ro over situasjonen.

Norsklærer, skole 3

I dette sitatet trekker læreren en direkte forbindelse mellom forsøket og sidemålsopplæringen, fordi det blir mindre press på vurderingssituasjonene for å frembringe en egen sidemålsskarakter. En annen lærer ved samme skole bemerket at det var først under forsøket det så ut til å ha gått opp for elevene at bokmål og sidemål ikke er to forskjellige språk. Dette kan ha sammenheng med hvordan læreren driver norskundervisningen når det ikke lenger skal settes separate karakterer i de to målformene.

I utvalget av videregående skoler som vi har besøkt, er det en skole hvor majoriteten av elevene har nynorsk som hovedmål. Dette syntes både oppdragsgiver og vi som forskere var et fornuftig utvalgskriterium med tanke på å undersøke sidemålets stilling i forsøket. Da vi snakket med lærerne på denne skolen, viste det seg likevel at grunnlaget for å diskutere nettopp dette spørsmålet i grunnen falt bort, fordi «sidemålets stilling» i norskfaget nødvendigvis må forstås som nynorskens stilling i faget under forsøket. Ved denne skolen foregikk som sagt all undervisning i hovedmålet på nynorsk, selv om enkeltelever i tråd med lovverket kunne velge seg bokmål som hovedmål og få dette bedømt i tilhørende vurderingssituasjoner. Det var derfor ingen ved skolen som stilte spørsmål om nynorskens posisjon verken forut for eller under forsøket, og undervisningen i bokmål som sidemål foregikk slik den alltid hadde gjort. Det betyr at ved denne skolen ble det lagt generelt mindre arbeid i sidemålsundervisningen enn hva som var tilfelle ved skoler der nynorsk er sidemål. Dette fordi elevene ved «nynorskskolen» ifølge lærerne ble tett eksponert for bokmål i alle mulige sammenhenger utenfor skolen, og behersket dette vel så godt som nynorsk. Opplæring i hovedmålet så vel som sidemål gikk derfor som før ved denne skolen, og hvis forsøket gagnet en av målformene i særlig grad, må det være nynorsk. Dette fordi elevene i likhet med de andre skolene fikk flere øvingssituasjoner på å skrive i norskfaget.

De to ungdomsskolene vi har besøkt som ledd i denne evalueringen, ligger begge i bokmålskommuner.¹³ Ved begge disse skolene brukte lærerne mye tid under intervjuet på å snakke om hvordan forsøket gjorde det mulig for dem å vente med å introdusere elevene for sidemålet og at sidemålet på denne måten kunne innføres i norskfaget på en «mykere» og mer interessevekkende måte. Disse lærerne understreket alle hvilken verdi de selv så i nynorsk og i det å gi *opplæring* i sidemålet, men flere av dem (vi har intervjuet i alt fem ungdomsskolelærere) mente at de ikke så samme verdi i å gi elevene *vurdering* i sidemålet. Dette hang i intervjuene sammen med at disse lærerne så nynorsk som et fag hvor elevene gjør mye feil (rettskriving, grammatikk, vokabular, setningsbygning) og hvor elevene utviser lav motivasjon. I tillegg mente de at mange av elevene deres har liten forkunnskap eller

¹³ Det totale utvalget nynorskkommuner blant skoleeierne som har ungdomsskoler som deltar i forsøket, er så lite at vi ikke kunne ha sikret disse skolenes anonymitet på forsvarlig måte dersom vi hadde besøkt dem for å gjøre intervjuer.

trening fra sidemålsopplæring når de starter i ungdomsskolen (se også tidligere i kapittelet). Disse lærerne mente derfor at den tilretteleggingen forsøket ga dem i å gi elevene en mer forsiktig innføring i sidemål, gagnet elevene i møte med sidemålsundervisningen.

Denne saken hadde imidlertid en annen side i de to ungdomsskolene, og denne siden handler om kompetansemålene i sidemålsopplæringen. Ved den ene ungdomsskolen hadde det vært fast praksis å vente med sidemålsopplæringen til våren på 8. trinn eller høsten på 9. trinn. På den andre ungdomsskolen fortalte en av lærerne at dette hadde vært en praksis blant *noen* lærere for en tid tilbake, men at forsøket nå legitimerte en slik praksis ved skolen for alle som ønsket det. Selv om lærerne, både dem som kjente slik praksis fra før og dem som ikke gjorde det, syntes ordningen hadde mye for seg, undret de seg samtidig om dette gjorde at elevene kanskje ikke fikk nok opplæring i sidemål, sett i lys av kompetansemålene. Samtidig argumenterte disse med at læreplanen ikke opererer med et sett helt adskilte kompetansemål for de to målformene, slik at økt trening i det å skrive kan gagne elevene i begge målformer.

Ved den ene av disse ungdomsskolene mente lærerne at forsøket hadde bidratt til at praksisen med å vente med sidemål til våren på 8. trinn fikk tiltrengt legitimitet. Samtidig opplevde de på denne skolen at flere foreldre var skeptiske med hensyn til at sidemålsskarakteren «forsvant» under forsøket. Lærerne fortalte videre at fagseksjonen ved skolen hadde brukt forsøket til å ta frem og revidere den lokale norskløplanen med særlig vekt på sidemålsundervisningen. En av lærerne ved denne skolen sa at arbeidet med planen hadde gitt henne en forsikring om at undervisningen nå var mer lik for alle lærere. Når hun tidligere spurte kolleger om hvordan de startet nynorskundervisningen, fikk hun inntrykk av at praksis var svært ulik. Mens lærerne ved denne skolen tidligere hadde fordelt kompetansemålene i sidemål over 9. og 10. trinn, ble målene nå fordelt over alle tre årstrinn i ungdomsskolen. «Da fikk vi også snakket om hvor det er smart å begynne, hvordan vi kan få motivert elevene,» sa denne læreren. Nå begynte lærerne med språkhistorien for å gi elevene forståelse av hvorfor de skal lære sidemålet. «Jeg er veldig glad for dette, for språkhistorien har nok fått 'svi litt',» fortsatte denne læreren. «Vi har tatt det hvis vi har hatt tid, hvis det har vært mulighet. Men for noen elever går det nå opp et lys.»

Denne lærerens kollega hadde også en refleksjon om muntligdelen av norskfaget, som lærerne generelt var mindre opptatt av i intervjuene. De var imidlertid enige om at mange av elevene får vist større kunnskap muntlig enn skriftlig. Samtidig, sa denne læreren, er det å sette muntligkarakter noe av det vanskeligste. Målene er vage, mente hun, og det er veldig mange kompetansemål som skal dekkes, samtidig som det er tidkrevende å ha vurderinger i dette. «Så det er kanskje den

delen av norskfaget som jeg sliter mest med,» sa hun, «det å finne gode vurderingsmåter i muntlig.» Denne skolen gjennomførte forsøket med to karakterer, mens denne læreren mente at hvis man skulle ha én karakter ved halvårsvurdering i norsk, må man også gjøre noe med kompetansemålene.

6.1.3 Lærere om motivasjon og læring under forsøket

Flere av lærerne vi snakket med, fortalte hvordan de selv fikk øket motivasjon som lærere under forsøket fordi bortfall av krav om et visst antall vurderingssituasjoner som grunnlag for karakter ved halvårsvurdering reduserte opplevelsen av stress i arbeidshverdagen. Lærerne mente at også elevene ble mer motiverte for faget når antallet formelle prøver og innleveringer gikk ned, og at det ble mindre «skummelt» for elevene å levere inn tekst når lærerens tilbakemelding ikke automatisk ble gjenspeilet i halvårskarakteren. Slik mente enkeltlærere i intervjuene at elever nå i større grad forsto forskjellen mellom vurdering med henblikk på øving og vurdering med henblikk på resultat, og at elevene i det hele tatt syntes noe mindre opptatt av karakterer.

Mer ro og mindre stress i undervisningen kunne også skape rom for flere diskusjoner og mer spontant engasjement i timene, noe mange elever så ut til å like, sa lærerne. Mulighetene til å snakke og diskutere var ikke lenger bare forbeholdt samfunnsfag eller religion og etikk, men kunne oppstå i timen for eksempel på grunn av noe en elev hadde lest.

En ulempe for elevene, sa flere lærere, er likevel at basert på en samlekarakter kan det være vanskelig for dem å se om man er svak i nynorsk og god i muntlig, og i praksis ender slike sprik i måloppnåelse til en «midt på treet»-karakter tilsvarende 3 eller 4. Forskjellene mellom disiplinene skal likevel fremgå av elevsamtalene forut for halvårsvurderingen. Noen elever var ifølge lærerne misfornøyde med dette, fordi de ønsket å vise frem at de kunne få virkelig gode karakterer i for eksempel hovedmål (skriftligkarakteren) eller norsk muntlig. Motsatt fortalte andre lærere om elever som var takknemlige for at de unngikk strykkarakter eller laveste karakter i for eksempel nynorsk, og at denne ble «skjult» i en 3er i samlekarakter i norsk.

I hovedsak var det likevel reduksjon i rettebyrde og lettere arbeidsbetingelser lærerne viste til når de snakket om konsekvensene av forsøket i norsk. Eksempelene på undervisningsopplegg fra læreren som fikk antall formelle vurderingssituasjoner redusert fra ni til tre per semester, viser at denne endringen i arbeidsbetingelser kan lede til stor kreativitet. Denne læreren sa:

Om elevene har blitt mer motiverte av dette karakterforsøket, det vet vi jo ikke, men jeg tror kanskje det. De vil jo føle at det er mindre tvang til vurdering, at vi kan ha et opplegg på forhånd som bygger opp til en vurderingssituasjon. For de

skal jo lære ting på forhånd, de skal lære om en teksttype, de skal lære å jobbe seg frem mot den. Hvis du bare har vurderinger, får du ikke gjort dette, og du føler aldri at du ble ferdig med det du skulle ha gjort. Nå får du liksom gjort jobben. (...) Jeg synes det er viktig at elevene blir oppglødde. At de synes det er kjekt, samtidig som de lærer noe.

Norsklærer, skole 4

En kollega av denne læreren la i intervjuet til at norskfaget kan sies å ha «slitt litt med imaget.» Norsk er ikke faget som elevene gleder seg mest til i skolehverdagen. Å slippe til kreativitet i undervisning og la dem ha færre formelle vurderingssituasjoner, kan ifølge denne læreren få elevene til å glede seg litt mer til norsktimene. Ungdomsskolelærerne som vi har intervjuet, viste som tidligere nevnt til elevenes manglende motivasjon for sidemål og den lettere tilvenning til sidemålet og dermed mulighet for å styrke elevenes motivasjon som forsøket ga dem.

Læreren som er sitert over, understreket selv betydningen av at også læreren opplever å ha det trivelig sammen med elevene i timene. Trivselen ble styrket ved at læreren nå under forsøket kunne bestemme mer over tiden selv, det var færre pålegg om hva som skulle skje i faget og når det skulle skje i løpet av semesteret. Det var flere av lærere i videregående skole som i intervjuene sa at de opplevde at de hadde blitt bedre lærere under forsøket.

6.2 Skolekonteksten

Oppdragsgiver har ikke spurt spesielt etter hva slags systemer, normer og kultur for samarbeid som omgir lærerne i forsøket med redusert antall karakterer i halvårsvurdering i norsk. Vi har heller ikke stilt direkte spørsmål til lærerne i intervjuene om hva slags kultur for samarbeid som preger forsøket ved de ulike skolene, bortsett fra et enkelt spørsmål om hvor ofte norsklærerne møtes og om hvordan samarbeidet er organisert. Gjennom gruppesamtalene kom det likevel frem en rekke opplysninger som ga oss mer kunnskap og flere ideer om hvilke faktorer som muligens påvirker utformingen av forsøket – og dermed arbeidet med undervisning og vurdering – lokalt. Denne kunnskapen og disse ideene vil vi presentere her mot slutten av kapittelet.

Alle norsklærerne vi har snakket med, fortalte at de har møter med kolleger på samme trinn og avsatt tid til samarbeid omtrent hver fjortende dag eller minimum en gang i måneden. De møter alle kolleger i faget et mindre antall ganger i løpet av et semester, og på enkelte av skolene virket det som om dette overgripende fagsamarbeidet var tilgodesett med lite tid og fant sted litt tilfeldig.

Noen av lærerne på videregående skole fortalte riktignok at de var involvert i et tett, nært og fagmessig fruktbart kollegasamarbeid, med utspring i felles arbeids-

rom for en mindre gruppe norsklærere. Kolleger av disse lærerne sa samtidig i intervjuet at de ikke opplevde å ha tilgang til tilsvarende strukturer for samarbeid, fordi disse strukturene hadde en uformell form ved skolen. Andre lærere sa i intervju at de ved oppstarten av forsøket hadde løst vurdering akkurat slik som de alltid hadde drevet vurdering i norsk, altså uten å redusere antall formelle vurderingssituasjoner som skulle ligge til grunn for halvårskarakteren. Mange av våre spørsmål som gjaldt arbeid med kompetansemål og vekting av skriftlig/muntlig og hovedmål/sidemål (se over), ble besvart med utsagn som «det er vel egentlig opp til den enkelte lærer.»

På to av de fire videregående skolene utviklet gruppeintervjuene seg til meningsutvekslinger mellom lærerne om et eventuelt behov for styring av norskfaget generelt og forsøket i norskfaget spesielt ved skolen. Det første kunne ifølge lærerne oppnås ved at ledelsen fikk lærerne til å jobbe mer helhetlig med opplæringen i faget gjennom de tre årene i videregående skole, det vil si ikke bare konsentrere seg om kompetansemålene som tilhører hvert enkelt årstrinn. Vi fikk inntrykk av at organiseringen av lærersamarbeid på trinn nå forsterket denne oppstykkningen av faget, noe som ble understreket av hvordan lærerne i intervjuet ikke kunne bruke Vg2-kompetansemål til å diskutere planlegging av undervisning på Vg1 eller Vg3 (se tidligere i dette kapittelet). Lærersamarbeid finner naturligvis sted på disse skolene, men kan tilsynelatende formes mer av relasjoner mellom kolleger som liker å arbeide sammen og har gjort det over tid, enn en helhetlig plan for hvordan slikt samarbeid kan styrke undervisningen ved skolen.

På den andre videregående skolen som diskuterte lærersamarbeid, var det delte meninger om hvorvidt dette arbeidet var tilstrekkelig og oppfylte de målsettingene som man mente burde finnes for samarbeidet. Enkeltlærere ved denne skolen deltok i et tett kollegasamarbeid om forberedelse, utvikling av oppgaver og vurderingskriterier med påfølgende bytte av rettebunker, mens andre enkeltlærere ved samme skole ikke hadde del i disse strukturene.

På den ene av ungdomsskolene vi har besøkt, hadde forsøket som nevnt ført til at norsklærerne i samarbeid tok frem skolens norskplan. Prosessen førte til at kompetansemålene i sidemål ble fordelt over alle tre skoleårene, og ikke bare over 9. og 10. trinn slik praksis ved skolen hadde vært tidligere. Lærerne ble også enige om i hvilken rekkefølge de skulle arbeide med de ulike delene i læreplanen. En lærer ved denne skolen uttrykte at hun nå kjente seg trygg på at lærerne hadde mer lik praksis i undervisningen. Ved den andre ungdomsskolen hadde ikke alle lærerne på tilsvarende måte bygget forsøket på felles samarbeid. Læreren som der fortalte at han lot sidemål telle kun 10 prosent av elevenes skriftligkarakter til halvårsvurdering til jul på 8. trinn for deretter å øke denne andelen frem mot sommeren (se tidligere i dette kapittelet), ble i intervjuet møtt av en kollega som spontant

utbrøt: «det hørtes veldig lurt ut!» Denne responsen indikerer at lærerne ikke tidligere hadde snakket sammen om denne løsningen, men at praksis utvikler seg hos den enkelte lærer og eventuelt i samarbeid med den eller dem av kollegene som denne læreren har mest kontakt med.

Slik kan vi si at forsøket med redusert antall karakterer ved halvårsvurdering i norsk medieres og formes av strukturene som omgir lærerne i form av administrativ ledelse (tid til samarbeid), pedagogisk ledelse (føringer for samarbeid, for samarbeidets innhold og målsettinger), kultur som preger kollegiet ved den enkelte skole og tradisjoner som sier «slik gjør vi det her». I tillegg kommer lærernes metodefrihet, erfaringer og ambisjoner for faget og for elevene gjennom undervisningen, og oppmerksomhet om elevresultater og vurdering i offentligheten. Forsøket gir lærerne frihet, men de kan også velge å begrense friheten fordi de ikke kjenner seg trygge på at det de gjør, gagnar kolleger som underviser på andre trinn, og dermed elevenes samlede utbytte av faget når de kommer til Vg3 eller 10. trinn. Gjennom lærernes fortellinger kan det se ut til at forsøket i noen grad har blitt individualisert av de samme lærerne til å utvikle og spisse de delene av undervisningen og opplæringen som lærerne synes er viktigst og som fortjener mest oppmerksomhet. Samtidig har vi også i datamaterialet eksempler på at forsøket ifølge lærernes utsagn har bidratt til å styrke en skolekultur for vurdering og en interesse for vurderingsarbeid. Ved alle de fire videregående skolene fikk vi likevel også svar i intervjuene som fastslo at «det er vel opp til den enkelte lærer.»

6.3 Oppsummering

I dette kapittelet har vi sett at færre karakterer ved halvårsvurdering i norsk gjør at norsklærerne kan redusere antallet vurderingssituasjoner som inneholder vurdering av elevenes læring understøttet av en tallkarakter. Dette har vi kalt for formelle vurderingssituasjoner. Disse vurderingssituasjonene ligger til grunn for karakteren som eleven får ved halvårsvurderingen. Skolene og kollegiene har etablert tradisjoner og felles oppfatninger av hvor mange formaliserte vurderingssituasjoner som skal inngå i en halvårskarakter i et fag, og lærerne har etablert systemer for å dokumentere og grunngi denne karakteren ved hjelp av de forutgående vurderingssituasjonene. Med færre karakterer ved halvårsvurdering kan lærerne redusere antall formaliserte vurderingssituasjoner som skal dokumentere halvårsvurderingen.

Derfor er det interessant at lærerne forklarer at de kan vurdere elevene *mye mer* når de samtidig skal gi elevene *færre* halvårskarakterer. Handlingsrommet oppstår fordi til tross for at skolene beholder antall krav til formelle vurderingssituasjoner som skal ligge til grunn for hver enkelt halvårskarakter, reduseres omfanget av disse vurderingssituasjonene helt enkelt fordi det gis færre halvårskarakterer. De

fleste (men ikke alle) lærerne vi har snakket med, benytter denne situasjonen til å gjøre to ting. For det første endrer mange lærere vektingen fra vurdering av elevenes resultater til vurdering av elevenes læringsprosess mens de er underveis i semesteret. Dette kan føre til at elevene i praksis får mange flere, mindre formaliserte vurderinger fra læreren, fra medelever og hvor de må vurdere sitt eget arbeid. For det andre benytter mange lærere bortfallet av skolens krav om seks til ni formaliserte prøver eller innleveringer til å planlegge undervisningen og vurderingen på nye måter. Dette kan de gjøre fordi de har mer tid til rådighet gjennom semesteret. Planleggingen kan dreie seg om innhold, prosess eller tidsbruk på emner – og naturligvis også kombinasjoner av dette.

Vi har i vårt datamateriale fått flest eksempler på hvordan lærere utnytter økt handlingsrom i undervisning under forsøket til utvide skriveopplæringen. Dette er et emne og en læringsprosess som krever mye tid, der mange og mindre formaliserte vurderingssituasjoner kan bidra til å utvikle elevenes skriveferdigheter. Med færre formaliserte vurderingssituasjoner med vurdering av læringsresultat kan lærerne la elevene arbeide med en teksttype over lengre tid enn hva som ifølge lærerne var mulig med tre karakterer ved halvårsvurdering i faget. Eleven kan levere inn flere skriftlige arbeider for hver teksttype som gjennomgås i undervisningen gjennom semesteret, og med hyppigere tilbakemeldinger får eleven øvd seg mer på å skrive i ulike sjangre.

Lærerne forteller også om hvordan muntligdelen av faget kan få en større, mer spontan plass i timene, fordi klassen nå har tid til å diskutere spørsmål på bakgrunn av aktuelle saker i for eksempel media eller tema de selv har støtt på i tekst. Opp-læringen i norsk sidemål kan ha fått mer tid og flere øvingssituasjoner under forsøket fordi elevene får en samlet karakter i skriftlig norsk. Samtidig er det lærere ved flere av skolene som på våre spørsmål om vekting av hovedmål/sidemål og skriftlig/muntlig svarer «det er vel opp til den enkelte lærer.»

Svært få av lærerne mener at deres arbeid med kompetansemålene i faget er påvirket av forsøket. Vi vil likevel nevne en av ungdomsskolene, som har gjort et gjennomgående arbeid for å sikre at kompetansemålene i sidemålet dekkes på lik måte av alle lærerne. Det overrasker oss heller ikke at lærernes arbeid med kompetansemålene fremstår som nokså uendret under forsøket, fordi læreplanen, kompetansemålene og rutinene for standpunkt-karaktersetting i avgangsåret er de samme uavhengig av om skolen deltar i forsøket eller ikke.

I intervjuene har vi sett mange eksempler på at lærerne bare i liten grad er samordnet og organisert i forsøket fra fagfellesskapet eller skoleledelsens side. Forsøket har derfor gitt dem frihet, men det kan synes som om denne friheten samtidig er individualisert. Forsøket har dermed ikke nødvendigvis utviklet kollegiet samlede vurderingspraksis, selv om vi mener at vi også har sett klare indikasjoner på dette særlig på en av de videregående skolene og en av ungdomsskolene som vi har

besøkt. På bakgrunn av datamaterialet stiller vi derfor opp en hypotese om at selv om lærerne er tilfreds med forsøket, kan det også ha bidratt til å bygge opp under læreres individualiserte vurderingspraksis der hvor fagfellesskap eller skoleledelse ikke på forhånd og under forsøket har opprettholdt innsats for å gjøre slik praksis kollektiv.

7 Elevers erfaringer med norskfaget og med forsøket

For gruppeintervjuene med elever i denne evalueringen har den overordnede problemstillingen vært elevenes syn på betydningen av norskfaget og syn på norskfaget som skolefag. Denne har vært ytterligere konkretisert i følgende spørsmål:

- Hvordan opplever elevene lærernes undervisning og vurdering i norskfaget?
- Fører færre karakterer ved halvårsvurdering i norskfaget til at elever opplever dette faget som mer eller mindre motiverende?
- Hvordan opplever elevene overgangen fra færre karakterer til normalt karakterregime i Vg3?
- Hvordan opplever elevene at forsøket eventuelt legger bedre til rette for deres læring i faget?

I kapittel 5 i delrapport 2 (Seland m.fl. 2016) er problemstillingene belyst med utgangspunkt i gruppeintervjuer med elever i *studieforberedende utdanningsprogram ved fire videregående skoler*. I dette kapitlet formidler vi analyser av ytterligere to gruppeintervjuer med *10.trinnselever ved to ungdomsskoler* (jf. tabell 1.1 i kapittel 1).

Vi har valgt å fremheve noen sentrale resultater fra intervjuene med videregåendelever innledningsvis, for å klargjøre hva som sammenlignes i analysen av elevutsagn fra de to utdanningsnivåene. Vi omtaler først elevenes utsagn om norskfaget som skolefag og deres erfaring med vurdering. Betydningen av at det er elever på avgangstrinnet vi har snakket med, kan sammenlignes på de to utdanningsnivåene. Vi undersøker også hvilket antall karakterer elevene ville ha foretrukket hvis det var de som fikk bestemme dette. Likeledes ønsker vi å være oppmerksomme på hvordan karakterregimet virker inn på motivasjon for skolearbeid og læring, slik elever uttaler seg om dette. Til slutt i kapitlet henter vi igjen frem materialet fra det videregående nivået når vi belyser hvert enkelt forskningsspørsmål gjengitt ovenfor, med tydeliggjøring av hovedinntrykk fra det videregående nivået og hva som er kommet frem fra ungdomstrinnet. Aller først vil vi imidlertid beskrive noen valg og avveiiinger som er gjort i prosessen frem til formidlingen av resultatene i dette kapitlet.

7.1 Tolkning og koding av elevmaterialet fra ungdomsskolen

Den generelle problemstillingen og første forskningsspørsmål kan betraktes som relativt overordnet forsøket, mens de neste tre forskningsspørsmålene er tettere koblet til forsøket. I samtaleene har temaene motivasjon og læring (andre og fjerde spørsmål) delvis glidd over i hverandre, og de presenteres derfor samlet. Opplevelse av overgangen fra forsøk til ordinært regime (tredje spørsmål) har ikke hatt samme relevans for de to skolene. Ved Skole 5 hadde 10.trinnselevne ikke hatt færre karakterer i norsk på de foregående trinnene, men elevene ved Skole 6 hadde slik erfaring (se også kapittelet om skoleledererfaringer og lærererfaringer for ytterligere kommentarer til dette forholdet).

Etter kodingen er forekomst og omfang av tekst som angår vurdering, relativt omfattende, mens tekst som angår læring har et betydelig mindre omfang. Med stikkordet vurdering har vi blant annet kodet utsagn om vurdering i form av tallkarakterer versus skriftlig eller muntlig vurdering formulert med lærerens egne ord. En kan mene at det er nettopp dette forsøket med færre karakterer handler om, hvilket tilsier at denne forskjellen bør være et viktig omdreiningspunkt når vi undersøker elevers erfaringer med forsøket. En kan også forstå vurdering som overordnet forsøket; det er noe elever har erfaring med uavhengig av om de har vært eksponert for forsøket eller ikke. Vi har valgt å omtale elevenes erfaring med vurdering både overordnet og gjennom forsøket.

Det gir ikke nødvendigvis god informasjon om en stiller elever spørsmål om hvordan de vurderer betydningen av norskfaget eller hvordan de ser på norskfaget som skolefag, selv om vi også har konkretisert dette blant annet som et spørsmål om de liker norskfaget. Det kan være mer informativt å stille andre og mer indirekte spørsmål, for eksempel om hvilke typer tekster de foretrekker å skrive. Ved hjelp av slike spørsmål får vi frem informasjon om hva elever mener det er viktig å lære, hva som oppleves som vanskelig og hvordan de oppfatter vurderingskriterier innenfor norskfaget.

7.2 Hovedinntrykk fra videregående materialet

Elevenes syn på betydningen av norskfaget og syn på norskfaget som skolefag, er en problemstilling som det er svært vanskelig avgrense, ikke minst med tanke på hvor omfattende faget er, for eksempel bedømt ut fra omfanget av kompetansemål. En første konkretisering kan være spørsmålet: hva synes elevene om norskfaget.

Blant videregående elevenes uttrykk for at faget engasjerer dem var poengteringen av at det er interessant «å kjenne norsk historie for å forstå hvorfor vi tenker og skriver som vi gjør nå». En glede ved å skrive en langsvarsoppgave på eget målføre var også nevnt. Mer negative oppfatninger gikk ut på at det er mye repetisjon, at historien er mindre interessant enn nåtiden og at det ikke alltid er åpenbart

hvorfor «man lærer om veldig mye om og om igjen, som man kanskje ikke trenger å gå så dypt inn i».

7.2.1 Et taktskifte mot eksamen

Elevene i videregående, som alle gikk i Vg3 i studieforbereende utdanningsprogram, var gjennomgående opptatt av at det nærmet seg eksamenstid da vi intervjuet dem i mars–april 2016. Dette gjorde at det ikke var så lett for elevene å skille overgangen fra en eller to til tre standpunktkarakterer i norsk fra en mer generell opplevelse av at antallet vurderingssituasjoner var økt, ettersom mange lærere skulle ha grunnlag for å sette standpunktkarakterer samtidig som alle fag nå var avgangsfag.

For Vg3-elevene vi snakket med, synes det som tallkarakterene gir veldig sterke signaler om deres måloppnåelse og at de virker skjerpende for arbeidsinnsatsen. Uttalelsene er også preget av at et høyt karaktersnitt har vært veldig viktig for mange elever med tanke på å komme inn på ønskede studier.

7.2.2 Uenighet om det optimale antallet karakterer

Det var betydelig variasjon i hvordan elevgruppene ved de videregående skolene vurderte hva som var det optimale antallet karakterer i norskfaget. De samme argumentene, eksempelvis at norsk er et omfattende fag med mange kompetansemål, ble brukt for diametralt forskjellige konklusjoner. Heller ikke elevenes prestasjonsnivå, bedømt ut fra den enkelte elevs øvrige uttalelser, ser ut til å sammenfalle med hvordan de konkluderte på spørsmålet om det er best med tre eller færre karakterer.

7.2.3 Karakterer som motivasjonsfaktor

Videregåendeelevene har vært ganske eksplisitte på at de motiveres av karakterer. Resonnementet kunne elevene selv se i sammenheng med hva de hadde lært om forskjeller mellom indre og ytre motivasjon; de fastholdt at karakterer, som ytre-motivering, er en viktig faktor.

I elevkapitlet i delrapporten som vi her refererer (Seland m.fl. 2016), har vi tatt forbehold om at vi kanskje ikke har klart å inkludere elever som strever med norskfaget, eventuelt at slike elever ikke har kommet frem med alternative synspunkter på hvordan karakteren virker inn på innsatsen. Dæhlen m.fl. (2011) finner en klar vektlegging av karakterer blant ungdomsskoleelever som snakket om skolemotivasjon i gruppeintervjuer. Hovdhaugen m.fl. (2014) drøfter det samme basert på

fokusgruppeintervjuer blant elever i Vg3 innenfor studieforbereidende utdanningsprogram. En tenke seg at karakterer virker motiverende for høyt presterende elever og demotiverende for elever som stadig får dårlige karakterer. I likhet med Dæhlen og hennes kolleger må vi altså ta forbehold om at vi kanskje ikke har klart å inkludere elever som strever med norskfaget, eventuelt at slike elever ikke kom frem med alternative synspunkter på hvordan karakteren virker inn på innsatsen.

7.3 Norskfaget

7.3.1 Elevers vurderinger av norskfaget

Fra intervjuene på ungdomstrinnet kan vi rapportere at på spørsmål til den ene elevgruppen om hva de syntes om norskfaget, ble det raskt sagt at «det er litt tørt og kjedelig». Med forsøk på å få dette utdypet, var spørsmålet om innholdet er krevende, hvorpå en gutt insisterte «Nei, ikke krevende, bare kjedelig». Dette ble forklart med at de opplevde at de gjør det samme hele tiden, og at det føles unyttig. Nynorsk var raskt nevnt på spørsmålet om hva som oppleves som unyttig, en annen sa: «Diktanalyse, bildeanalyse og så videre. Jeg ser poenget, men jeg føler ikke at jeg får så mye ut av det.»

I den andre elevgruppen ble det sagt at det er viktig å ha gode karakterer i matematikk og dessuten i språk for å være godt forberedt på videregående skole. På direkte spørsmål ble det samme sagt om norskfaget: «Jo bedre karakterer du har, jo bedre forberedt er du jo». Vi kan altså merke oss at karakterer tolkes som signal om hvor godt forberedt en er for videre utdanning.

Dette er hva vi har fått rede på om de intervjuede elevenes syn på norskfaget med bruk av direkte spørsmål. Det er imidlertid vanskelig å vite nøyaktig hva elevene sikter til når de påstår at faget er kjedelig eller unyttig. Med et ønske om å komme bak slike uttalelser forsøkte vi å stille mer indirekte spørsmål for å få elevene til å snakke om sine erfaringer uten at de skulle karakterisere norskfaget eller rangere fagets viktighet i forhold til andre fag.

Som omtalt i første delrapport i denne evalueringen, er kompetanse for deltakelse i et demokrati en overordnet målsetning i Kunnskapsløftet og i norskfaget, men det har vært debattert hvorvidt en når målet gjennom styrking av sakprosaens plass, eventuelt på bekostning av klassiske skjønnlitterære tekster (se også Seland m.fl. 2015, kapittel 3). Uansett, for å komme tettere på elevenes erfaringer og opplevelser i norskfaget, stilte vi spørsmål om hvorvidt de ville foretrekke å skrive kreative fremfor argumenterende tekster, om de spesialiserte seg eller om de oppfordres til å øve på å skrive alle typer tekster i norskfaget. Ifølge læreplanen skal elevene kunne skrive kreative, informative, argumenterende og reflekterende tekster.

En av guttene i intervjuet ved Skole 6 (ungdomsskole) mente at det er viktigere å trene på å skrive argumenterende tekster enn på å skrive kreative tekster. «Vi har mer bruk for det», hevdet han. Argumenterende tekster hadde vært vanskeligere i 8. klasse enn nå, og tidligere skrev de mer kreativt: «Da skrev vi jo bare historier». En annen gutt mente at for ham var det vanskelig å dikte opp, da kom han ikke i gang, mens en jente tilføyde at når en først kommer i gang, er det enklere, men at det som er vanskelig er å finne en god ide eller handling. En annen jente tilkjennega at hun syntes «strukturen på kreative tekster er litt mer skjult enn i en argumenterende tekst. Så det er lettere å skåre på det i argumenterende tekster». Alle mente at de oppfordres til å trene på alt, men en gutt fremhevet at det er lettere å komme i gang når en skal skrive en argumenterende tekst, fordi en kan ta et standpunkt. En jente syntes det var lettere med informative tekster, fordi «da har du alt lagt ut foran deg, du kan bare lage en tekst basert på det du finner».

Ved Skole 5 (ungdomsskole) mente også elevene at de oppfordres til å øve på alle typer tekster. Her var to elever inne på at de var best med kreative tekster, mens en annen gutt mente han foretrakk å skrive debattartikler og fagartikler, og at det var dette han fikk best karakter på. Han pekte på at det også finnes de som får toppkarakter på kreative tekster. Dette mente han var avhengig av eleven, men også læreren: «det kommer an på hva de trekker fram».

Med denne tilnærmingen har vi fått vite litt om hvordan elevene vi har snakket med, arbeider i skriftlig norsk, hva de har fremhevet som mer eller mindre vanskelig: å finne og sortere stoffet man trenger for å skrive en informativ tekst, å komme i gang med skrivingen eksemplifisert som å ta et standpunkt eller å konstruere en historie, og om hvordan strukturen i ulike typer tekster er mer eller mindre vanskelig å identifisere.

7.3.2 Elevers erfaringer med vurdering i norskfaget

Elevene vi snakket med ved Skole 5 hadde erfaring for at norsklærerne gjorde det klart for dem nøyaktig hva de jobbet med til enhver tid og hvordan de ble evaluert. For eksempel kunne de få høre at nå var det fokus på verb og substantiv. Selv om alt ville bli rettet, var det dette som var fokuset og som ville telle mest. En gutt mente at det han særlig likte var at det var lett å få en konstruktiv tilbakemelding fra læreren. Læreren ville kunne se ikke bare hvor godt de mestret det de hadde fokus på, men hvor godt de kunne implementere det i tekster, fremholdt han. Denne gutten mente også at norsk var faget hvor han opplevde mest nyttige tilbakemeldinger. Han mente at han alltid fikk tydelig beskjed om hva som kan forbedres, og han visste alltid hva han ble vurdert på. Han forklarte at de fikk vite hva slags prøve de skulle ha i slutten av en periode for eksempel med nynorsk. De fikk

først vite hva læringsmålene var, hva det ville bli fokus på og hva de ville bli evaluert på. Han mente det var lett å forstå hva de ble vurdert på:

Ja. Veldig. Norsklæreren min, hun er veldig opptatt av å bygge videre fra emne til emne. Så hun er veldig glad i å ta en samtale med eleven etter viktige prøver. Hun passer på at elevene forstår hvorfor de fikk karakteren de fikk og å gi dem noe å bygge videre på.

De andre elevene som deltok i intervjuet og som hadde en annen norsklærer, fikk vurdering i form av lærerens skriftlige vurdering i tillegg til tallkarakter. En jente medga at det fantes situasjoner hvor hun ikke forstod hva læreren mente, for eksempel hvis beskjeden var at hun måtte jobbe med verb, uten at hun gadd å spørre om det. Gutten vi har referert ovenfor, vurderte det slik at det ville telle om de viste interesse, derfor kviet ikke han seg for å spørre.

Lignende erfaringer kom frem i det andre gruppeintervjuet, og ved denne skolen ble det ofte brukt et skjema for avkryssing av hvordan de lå an. Etter tentamen hadde de også brukt slike skjemaer for å vurdere sine egne besvarelser.

7.4 Overgangen til flere karakterer i norsk

Sammenlignet med elevintervjuene i Vg3 rapportert i Seland m.fl. (2016), er det et hovedinntrykk at forberedelser til eksamen ikke har like stor oppmerksomhet på 10. trinn. Elevene var intervjuet i den samme perioden på de to trinnene, det vil si mars–april. I motsetning til videregåendelevne, var de ungdomsskoleelevne vi snakket med, litt mer tilbaketrukket med tanke på å komme inn der de ønsket. Vi kan ikke anta at dette gjelder generelt for avslutningen av ungdomstrinnet; det gjelder neppe i områder hvor det er konkurranse om plasser ved enkelte, attraktive videregående skoler. Likevel kan denne forskjellen være verdt å merke seg når vi her sammenligner materialet fra de to forskjellige nivåene.

7.4.1 Det optimale antall karakterer i norskfaget

Ved en av de to ungdomsskolene vi besøkte, hadde ikke elevene hatt færre karakterer i norsk på de foregående trinnene. På intervjutidspunktet var forsøket med færre karakterer blitt implementert på 8. og 9. trinn, men de 10. trinnselevne hadde altså ikke erfaring med noen overgang på dette området. Ikke desto mindre hadde de mange meninger om hva som var best: en, to eller tre karakterer i norsk.

En jente mente at det ofte er nynorsk-karakteren som trekker ned når en får karakter for både hovedmål og sidemål, derfor mente hun, blir det «et mer korrekt standpunkt med separate karakterer». En annen elevs uttalelse peker på at de

fleste ikke bryr seg så mye om nynorsk karakteren. Logikken med separate karakterer blir da å kunne skille fra hverandre prestasjonen i de to målformene i stedet for at de blir slått sammen til et snitt, som følgelig ikke sier så mye om hvor den beste prestasjonen ligger.

En fordel med færre karakterer ville være at de fikk «mindre prøver og sånt, mindre innleveringer og stress», mente en av guttene. Også for lærerne ville det være en fordel, de ville ikke trenge å lage tre–fire prøver for hver karakter, men kanskje bare fem–seks til sammen. En fare med færre karakterer kunne være at nynorsk ble ganske «bortglemt», fremholdt en av jentene.

I den gruppen hvor de hadde erfaringer med færre karakterer, var alle enige om at tre karakterer er det beste. Begrunnelsene de ga var at det er bedre å vite hvordan man ligger an og følgelig som en jente fremhevet: «om du trenger å jobbe mer, konsentrere mer på det faget eller om du bare må holde deg der du er». Ettersom flere av disse elevene mente de foretrakk tre karakterer fordi det ga dem mer informasjon om hvordan de lå an, ble de spurt om de kunne tenke seg flere enn tre karakterer. Til dette svarte de nektende, og en gutt sa: «Nei, da blir du vurdert i hjel». Bedømt ut fra denne guttens videre utdanningsplaner er det nærliggende å anta at han ikke hadde et veldig sterke karakterer.

1.1.1 Nynorsk som nytt «fag»

De elevene som hadde erfaring med færre karakterer, oppsummerte for hverandre hva som hadde skjedd i 10. trinn. En gutt oppfattet det slik: «Før fikk vi norsk muntlig og skriftlig. Nå får vi nynorsk skriftlig også». Uten separat karakter i nynorsk hadde en av jentene følt hvordan hun lå an, men med karakter visste hun det nå bedre enn hun gjorde tidligere. En annen jente mente at hun tenkte mer på det nå med tanke på standpunkt ettersom det var en egen karakter, enn hun gjorde tidligere.

Andre uttalelser i intervjuene peker på at det er vanskelig at det elevene oppfatter som et nytt fag, det vil si nynorsk, introduseres på åttende trinn når det tar så lang tid før de får karakter i dette faget. Det er dette elevene ved Skole 6 oppfatter har skjedd når de har hatt færre karakterer på tidligere trinn; Nynorsk introduseres, men det tar to år før de får karakter. Hvorvidt denne oppfatningen er imøtegått av lærere, har vi ikke kjennskap til.

Disse elevene ble spurt om de visste gjennom 8. og 9. trinn når de hadde henholdsvis hovedmål og sidemål, og dette bekreftet de. De hadde fått beskjed av lærerne om at de skulle ha en bestemt målform innenfor en bestemt periode. De mente likevel at de ikke hadde skrevet så mye på nynorsk tidligere. Før hadde det vært fokusert på grammatikk og ordbøying, mens de nå skrev mer, og nå ville de også få tentamen i sidemål.

1.2 Motivasjon og læring gjennom forsøket

1.2.1 Motivasjon ved hjelp av karakterer

Ved skole 6 fortalte elevene at de hadde hatt en stor diskusjon med lærerne året i forveien om færre karakterer som et premiss i forsøket. Både elever og foreldre var kritiske til at det ikke skulle gis tallkarakterer. For lærerne var det en utfordring å få elevene til å forstå og innrette seg etter de vurderingene lærerne ga uten bruk av tallkarakterer. Blant de fire elevene vi snakket med, pekte en av jentene på at karakterer var hennes viktigste kilde til motivasjon:

Jeg personlig får mer motivasjon av å vite karakteren. Da vet jeg hvor jeg ligger og at jeg må jobbe mer. Men som nå når vi ikke har fått vite karakter utenom på slutten av halvåret, det liker jeg ikke. Jeg hadde hatt mer motivasjon om jeg fikk vite den.

Læreres tilnærming var å gi elevene det de kalte en respons som vurdering av en skriftlig innlevering. En av elevene vi snakket med ga følgende forklaring på hva en respons er: «Du får jo ikke vite hvor du ligger hen da, bare hva du må gjøre bedre og hva du har gjort bra». Da må vi selv finne ut hvor vi ligger, mente en annen elev, som sa «det er bedre å få et tall, for da er man sikker». I intervjuet ble de utfordret gjennom en påpekning av at noen ser for seg at karakteren kan stå i veien for at man involverer seg i læringen. Til dette svarte en elev at hun så den siden også, men tilføyde: «personlig foretrekker jeg å vite».

Det kom frem at hvorvidt elever gjør bruk av informasjonen i en respons, kan være et spørsmål om en er fornøyd eller ikke, mente en gutt: «Hvis du er fornøyd så blåser du i det, jeg fikk det som står der ingen grunn til å jobbe videre». På et spørsmål om det kan være greit å slippe noe av presset på å stadig måtte tenke på karakterer og poengsummen, bekreftet imidlertid alle elevene. En av guttene sa seg enig og pekte på at «foreldre vil jo at du skal nå topp». Den jenta som hadde forklart oss hva en respons var, mente at det var både og. Vi oppfatter at hun fastholdt sin oppfatning om at det er en fordel å få en tallkarakter fremfor en vurdering med ord.

En konklusjon vi kan trekke av denne samtalen er at elevene synes å være vel informert om fordeler ved færre karakterer og de kjenner til hensikten med en respons. De erfarer også at en god tallkarakter kan avlede fra å forsøke å sette seg inn i hva læreren formidler gjennom responsen. Vi vet at diskusjonen har foregått på skolen og med uttalt motstand fra elever og foreldre mot færre karakterer i norskfaget slik forsøket legger opp til, og at i alle fall noen elever likevel vil foretrekke tallkarakterer som indikasjon på eget prestasjonsnivå.

En annen konklusjon er at elevene synes å mene at det alltid vil være mulig å vurdere deres prestasjoner med tallkarakterer. Når dette ikke gjøres, må de gjette, og et tall ville gjøre dem sikrere. På denne måten tilbakeholdes informasjon som

elevene oppfatter som viktig. En tallkarakter fremstår som en mulig tilbakemelding, men som likevel holdes tilbake. Dermed blir den en hemmelig størrelse, som de bare kan gjette seg frem til, og dette med betydelig grad av usikkerhet. En tolkning av noen av elevenes uttalelser i denne diskusjonen er altså at de later til å tro at karakteren finnes eller at den er mulig å sette, men at den holdes skjult for dem.

7.4.2 Nynorsk som et «eksplosivt fag»

I begge elevgruppene på ungdomsskolene kunne de huske at de startet opp med sidemål i 8. klasse. For en av guttene var dette altfor sent. Han mente at ungdomsskolen er et veldig sent stadium for å introdusere nynorsk som et nytt skriftspråk (sic), spesielt ettersom de også hadde valgspråk (sic) å bygge videre på. Han mente nynorsk ble implementert ganske plutselig, og han savnet klarere linjer for evaluering innenfor nynorsk. Om karaktersetting sa han følgende:

I flere fag så vil jeg jo si at karakter er en motivasjon. Spesielt fag der man er god for da kan man sikte litt høyt. Men i et så eksplosivt fag [som nynorsk] så vil jeg si det er mer en stresskilde enn noe annet.

Med eksplosivt fag, oppfatter vi at han sikter til at nynorsk var implementert så plutselig og uten at de hadde opparbeidet seg forkunnskaper eller hadde kunnet forberede seg. Vi ser her at denne eleven selv skiller mellom karakterer som motiverende og som en stressfaktor avhengig av hans eget opplevde mestringnivå.

7.4.3 Læring gjennom forsøket

En kan tenke seg at læring skjer gjennom at elever får arbeide med en tekst over tid, at de får tilbakemeldinger på et utkast for så å arbeide videre med teksten, slik lærere har fremhevet at de har bedre muligheter til på grunn av forsøket. I begge elevgruppene ble det bekreftet at de har hatt slike skriveprosesser, men vi fikk ikke frem klare meningsytringer om hvordan elevene synes det er å arbeide slik.

Etter mange forsøk på å utfordre elevenes syn på verdien av å få tallkarakterer og om det å være opptatt av karakterer kan distrahere dem fra å lære, konkluderte en jente som følger:

Ja, altså karakterer legger jo en del press på det, men jeg føler ikke det går i veien for selve faget. Det er mer det at for mange vurderinger gjør at istedenfor å lære om for eksempel norsk litteraturhistorie så er det mye repetisjon.

Dette er det nærmeste vi kommer en bekreftelse innenfor dette intervjumaterialet fra ungdomstrinnet på at orientering om karakterer og orientering om å lære og

fordype seg kan stå i et motsetningsforhold. Når lærere trenger mange vurderings-situasjoner for å kunne sette en standpunktkarakter, betyr det at mye repetisjon tvinger seg frem, slik elever erfarer det.

7.5 Oppsummering fra de to nivåene

Vi vil i det følgende belyse problemstillingene som er gjengitt innledningsvis i kapitlet. Dette gjør vi gjennom å gjengi konklusjonene fra analysene av intervjumaterialet fra elever i fire videregående skoler som er presentert tidligere, og grunnlaget for disse konklusjonene finnes i andre delrapport fra denne evalueringen (Seland m.fl. 2016: kapittel 5). For hver problemstilling supplerer vi med konklusjoner fra presentasjonen i dette kapitlet av elevmaterialet fra de to ungdomsskolene.

7.5.1 Hvordan opplever elevene lærernes undervisning og vurdering i norskfaget?

Videregåendeelevene

Selv om elever var forbeholdne i sin begeistring for norskfaget, kunne de fortelle om hvordan spesifikke arbeidsmåter har vært til hjelp for å forstå hva som skal til for å vise høy måloppnåelse. Å lære om hvordan sensor leser en tekst, var for enkelte elever veldig nyttig og lærerikt. Å se flere fag i sammenheng var godt likt av noen elever. En av fordelene var at dette gir færre prøvesituasjoner i en hektisk tid. Å skrive for eksempel en historisk fagtekst kunne oppleves som mer meningsfullt enn å gjennomføre flere mindre prøver. Det er likevel slik at slike arbeidsformer kanskje ikke har engasjert alle like mye.

En annen nokså utbredt oppfatning var at norskfaget er svært omfattende og sammensatt av mange ulike ferdighets- og kunnskapsområder. Det ble beskrevet som et veldig stort fag. For noen fremstår norskfaget som ganske ensformig om enn mer avansert over tid fra ungdomsskolen og gjennom videregående skole; det er det samme pensumet, de samme måtene å arbeide på og det samme innholdet. Når mange lærere vil utvikle noe nytt, slik elevene har erfart gjennom forsøket, viser de seg å utvikle det samme, var en uttrykt oppfatning.

Vi har også hørt betraktninger om at undervisningen var mer variert på foregående trinn i videregående, da de hadde færre karakterer. Den gang var det mer diskusjon, timene var ikke så ensformige som de er blitt i Vg3, og de hadde bedre tid, poengterte elever.

Ungdomsskoleelevene

Variasjon i undervisning og arbeidsformer har ikke vært et fremtredende tema i samtalen med ungdomsskoleelevene, og de har selvfølgelig ikke så lang erfaring med norskfaget som videregåendeelevene. Med fokus på skriving har ungdomsskoleelevene i overveiende grad gitt uttrykk for hva de opplever som krevende ved å skrive kreative tekster sammenlignet med å skrive argumenterende og informative tekster. Flere av elevene har understreket at læringsmål og vurderingskriterier gjøres tydelige for dem i norskfaget.

1.2.2 Hvordan opplever elevene overgangen fra færre karakterer til normalt karakterregime i Vg3?

Videregåendeelevene

Innholdet og arbeidsmåtene i faget ble gjennomgående beskrevet som ganske like til tross for overgangen til tre karakterer. Et synspunkt er at tre karakterer gir bedre pekepinn om hvor de er på vei ved halvårsvurderingen. Karakterer fungerer som et veldig sterkt signal om elevenes måloppnåelse og skjerpende for deres oppmerksomhet.

Antallet prøvesituasjoner i norskfaget ble oppfattet som økt i Vg3 sammenlignet med de foregående trinnene, samtidig som også lærere i andre fag trengte et tilstrekkelig antall vurderingssituasjoner for å kunne sette standpunktkarakter. Noen elever mente fokus på forestående eksamen måtte være årsaken til at de trente gjennom flere skriftlige innleveringer enn de var vant til tidligere, men andre syntes å være overbevist om at det var det økte antallet karakterer som tvang frem flere prøvesituasjoner.

Vi kunne ikke skille antall karakterer fra andre forhold som endrer seg parallelt i overgangen til Vg3. Konsentrasjonen om eksamen innebar at elever måtte trene mer på å produsere tekst. Faget var blitt et avgangsfag med i alle fall én viktig og obligatorisk eksamen. Noen av uttalelsene antydte at norskfaget og spesielt sidemålet ble forsømt under presset fra avgangsfag på foregående trinn. Ved en annen skole fortalte elevene om stor grad av systematikk i vekslingen mellom målformer i skriveoppgavene. Dette illustrerer den variasjonen som med stor sannsynlighet også finnes mellom skolene i hvordan forsøket med færre karakterer har vært praktisert.

Ungdomsskoleelevene

Sammenlignet med uttalelsene fra elevene i videregående, har vårt materiale fra ungdomsskolen ikke på langt nær så mye fokus på prøver og forestående eksame-

ner. Ved den ene skolen hvor de har opplevd overgang til flere standpunktkarakterer i norsk, oppfatter elevene at de har fått introdusert nynorsk på 8. trinn, men at det tar to år før de får karakter i faget. Når de etter hvert får karakterer i sidemål, ser vi at dette oppleves som et stressmoment eller som motiverende, avhengig av elevens opplevde mestringsnivå.

Diskusjonen med ungdomsskoleelevene som har opplevd et regime med en eller to karakterer i norskfaget, tyder på at de oppfatter det vesentligste med forsøket som tilbakeholdelse av informasjon om tallkarakterer. Selv med tilbakemeldinger fra læreren om hva den enkelte eleven har gjort bra og hva som kan gjøres bedre på en skriftlig innlevering, er elevene mer interessert i å vite hvor på skalaen deres prestasjon ville bli plassert. Uten denne informasjonen blir de nødt til å gjette, og de frustreres over usikkerheten ved dette.

7.5.2 Hvordan opplever elevene at forsøket eventuelt legger bedre til rette for deres læring i faget?

Videregåendeelevene

Det var ingen enighet blant videregåendeelevene om hvor mange karakterer som er det optimale i faget. Uenighetene syntes å gå på tvers av prestasjonsnivå; de som presterer vesentlig bedre i ett område (sidemål eller muntlig) falt ned på ulike konklusjoner om hvorvidt det er best med tre karakterer eller om det er like greit med en eller to karakterer som reflekterer et snitt. Vi så også at antallet kompetansemål eller bredden i faget anvendes i argumentasjon for flere så vel som færre karakterer.

Ungdomsskoleelevene

Fra samtalene med ungdomsskoleelever har vi vært inne på at mange vurderingssituasjoner resulterer i mye repetisjon, og at dette går på bekostning av å lære noe nytt. Det er ellers sparsomt hva ungdomsskoleelevene har fortalt om læreprosesser. Materialet viser en mer entydig oppslutning om tre karakterer som det beste for norskfaget. Ved den ene skolen hvor de ikke har opplevd noe annet enn tre standpunktkarakterer i norsk, mener de at dette er det beste regimet. Ved den andre skolen er synliggjøring av den beste tallkarakteren og hvordan denne skiller seg fra vurderingen i sidemål et viktig argument for tre karakterer.

7.5.3 Blir elevene mer motivert for norskfaget med færre karakterer ved halvårsvurdering?

Videregåendeelevene

Frustrasjon over å tvinges til å måtte prioritere fag eller kompetansemål når arbeidspresset blir stort, kom til uttrykk fra en gutt som også fortalte om gleden han opplevde når han skrev langsvarsoppgaver. Slike uttrykk for motivasjon er meget sjeldne i materialet. Et flertall av elevene mente de ble motivert av tanken på karakterene og karaktersnittet. Selv med kjennskap til teorier om karakteren som ytre-motivasjon, ble oppfatningen av karakterer som en sterk motivasjonsfaktor opprettholdt. Disse elevene var også bevisste krav om høye karaktersnitt for å komme inn på studier de ønsker seg, men mente også at motivasjonen kan svekkes dersom en ikke vet hva en skal gjøre etter videregående.

I seg selv er dette neppe et oppsiktsvekkende funn. Dæhlen m.fl. viser at karakterer virker ulikt i ulike prestasjonssjikt: de kan være en kilde til innsats for elever som lykkes, mens de virker motsatt for elever som ut fra erfaringer mener at det ikke nytter uansett (ibid.: 83). Når vi gjennom intervjuene med videregåendeelevene kom til at karakterene utgjør en så viktig motivasjonsfaktor, måtte vi ta forbehold om at vi kanskje ikke hadde klart å inkludere elever som strever med norskfaget, eventuelt at slike elever ikke kom frem med alternative synspunkter på hvordan karakteren virker inn på innsatsen. Vi må altså ta forbehold om at gruppeintervjuer kanskje ikke så lett bringer frem alternative synspunkter på karakterer som motivasjonsfaktor.

Motivasjon er nært forbundet med prestasjoner og prestasjonserfaringer (Eccles & Wang 2012). Andre tilnærminger i denne evalueringen, spesielt sammenligning av prestasjonsutvikling i og utenfor forsøket, kan gi bedre svar på i hvilken grad karakterer styrker elevenes læring, og i hvilke prestasjonssjikt det i så fall skjer. Resultater fra disse analysene er presentert tidligere i denne rapporten. Når de viser at det ikke er forskjell i karakterutvikling mellom elever som har deltatt versus elever som ikke har deltatt i forsøket, heller ikke blant svakt presterende, har vi ikke grunnlag for å hevde at et redusert antall tallkarakterer styrker læringen, i alle fall ikke over et tidsspenn på to til tre år.

Ungdomsskoleelevene

Også ungdomsskoleelevene har vært veldig eksplisitte på hvordan de motiveres av å få sitt arbeid bedømt ved hjelp av tallkarakter. I samtalene har vi forsøkt å utfordre dem og belyse spørsmålet fra flere sider. Selv om karakterer øker presset og kan virke stressende, fastholder elever fordelene med karakterer med den begrunnelsen at det er bedre å vite. Vi kan også merke oss at de fleste av de elevene vi har

snakket med, kjenner argumentene fra en tidligere diskusjon om forsøket ved skolen, hvor også foreldre tok til orde for flere tallkarakterer.

Gjennom grundig lesing av intervjuutskriftene har vi imidlertid oppdaget to tilfeller til at elever også kan oppleve karakterer som noe negativt. Det gjelder den gutten som hevdet at med mer enn tre karakterer i norsk, ville de bli vurdert i hjel. Vi har også sett hvordan en gutt skiller mellom når karakterer virker motiverende og når de utgjør en stressfaktor. Det avgjørende er hvordan han erfarer at han mestrer det som han blir vurdert på.

8 Videre arbeid med å utvikle vurdering i norskfaget

I denne evalueringen har vi studert organisering og gjennomføring av forsøk med én eller to karakterer ved halvårsvurdering i norsk, vi har undersøkt skolelederens, læreres og elevers erfaringer med og synspunkter på forsøket, og vi har sammenlignet standpunktkarakterer for elever som har vært del av forsøket med elever som ikke har deltatt i forsøket.

Hovedproblemstillingen som denne evalueringen skal besvare, er som følger:

Fører forsøket med færre karakterer ved halvårsvurdering i norsk til at skolene kan utvikle mer hensiktsmessig organisering av undervisning og vurdering i norskfaget?

Vårt hovedfunn i denne rapporten er at elevenes standpunktkarakterer ikke viser systematiske forskjeller avhengig av om de har deltatt i forsøket eller ikke. På denne bakgrunnen diskuterer vi i dette kapitlet på et overordnet nivå hvordan vi kan forstå *hensiktsmessig organisering av undervisning og vurdering*. Diskusjonen utvides deretter med spørsmålet om det er elementer i vår studie som peker frem mot en mer spesifisert utprøving av forsøk med én eller to karakterer ved halvårsvurdering i norsk.

8.1 Ingen systematiske forskjeller i elevenes standpunktkarakterer

Evalueringsoppdraget ber oss om å undersøke om det er mulig å identifisere systematiske forskjeller i elevenes læring som følge av deltakelse i forsøket (se kapittel 1). I dette arbeidet har vi brukt elevers standpunktkarakterer ved utgangen av Vg3 som et mål på læring i norskfaget.

Om standpunktkarakteren gjelder det at denne karakteren skal gi informasjon om elevens kompetanse ved avslutningen av opplæringen. Standpunktkarakteren settes av elevens faglærer på bakgrunn av en rekke forutgående vurderingssituasjoner med og uten karakter gjennom avgangsåret i videregående skole, hvor hel-

dagsprøven i vårterminen tillegges stor vekt. Standpunktkarakteren føres på elevens vitnemål og er sammen med eksamenskarakteren det grunnlaget som brukes for å vurdere elevens kvalifikasjoner og egnethet ved søkning til høyere utdanning.

Samtidig betyr det å sette likhetstegn mellom elevens læring i norskfaget og elevens standpunktkarakter i norskfaget at elevens læring *avgrenses* til denne ene karakteren. Denne karakteren gir dermed et forenklet bilde av hva eleven faktisk kan. Siden standpunktkarakteren i norsk også hviler tungt på én heldagsprøve mot slutten av semesteret, vil den kompetansen eleven får vist på denne prøven, også avgrense målet på hva eleven faktisk har lært i faget gjennom alle årene i grunnopplæringen. Litteraturen vi har gjennomgått i forbindelse med denne evalueringen viser at norskfaget er et svært innholdsrikt og mangefasettert fag, som favner ideer og idealer om blant annet skriftkyndighet og grunnleggende ferdigheter, kulturforståelse, kommunikasjon, dannelselse og identitetsutvikling. Dette gjør at elevens evne til å uttrykke seg ved å skrive en tekst under en heldagsprøve ved avslutningen av skoleåret, umulig kan demonstrere den fulle bredden og dybden i alt hva eleven virkelig har tilegnet seg fra faget og hvordan han eller hun har utviklet seg som menneske gjennom skoletiden.

I delrapport 2 fra denne evalueringen (Seland m.fl. 2016) refererte vi hva norsk-lærere *ønsket* at elevene skulle sitte igjen med etter norskundervisningen når de var ferdige med videregående skole. Lærerne svarte oss at de håpet elevene hadde blitt gode til å lese og til å oppfatte budskap, til å forstå referanser i språket og derigjennom bli bedre til å skrive. Denne praktiske forståelsen og bruken av språket gled i lærernes refleksjoner over i det bredere dannelses- og deltakelsesaspektet ved opplæringen. På en skole hvor lærerne først hadde vært innom alle disse temaene i den innledende delen av intervjuet, så de på hverandre og smilte. «Ja, det er en veldig kompleks prosess,» sa en av dem. «Vi håper jo at alle disse enkeltdelene skal bli en helhet hos elevene etter hvert.»

Dette kan illustrere hvordan norskfaget er et fag som rommer svært vidtrek-kende ambisjoner, muligheter og potensial for utvikling hos elevene. Det er åpenbart at dette er fenomener, egenskaper, ferdigheter og forståelser som kan bruke lang tid på å manifestere seg hos den enkelte, og som det kan være vanskelig, for ikke å si umulig å måle ved hjelp av en enkelt karakter og sidestille dette med elevens læring i faget.

Til tross for dette brukes standpunktkarakteren like fullt som mål på læring i norsk utdanningsforvaltning. I denne evalueringen har vi brukt dette målet som grunnlag for en hypotese om at dersom forsøket har skapt systematiske forskjeller i elevenes læring, vil vi kunne se dette i form av endring i karaktergjennomsnitt for elever som har deltatt i forsøket sammenlignet med elever som ikke har deltatt i forsøket. Analysen av standpunktkarakterer for 197 000 elever viser at det ikke er

noen forskjell i standpunktkarakter for elevene som har deltatt i forsøket sammenlignet med elevene som ikke har deltatt i forsøket. Det er ingen forskjell i dette resultatet når vi i analysen nyanserer mellom skoler som har gjennomført forsøket med én karakter og skoler som har gjennomført forsøket med to karakterer ved halvårsvurdering. Det er heller ingen forskjeller i dette resultatet når vi sammenligner elever som har et lavere eller et høyere antall grunnskolepoeng, og som vi kunne betegnet som henholdsvis lavt presterende eller høyt presterende elever. Det er heller ingen forskjeller i dette resultatet når vi sammenligner elever på bakgrunn av kjønn, skolestørrelse eller skolens plassering i landet. Vår hypotese om at deltakelse i forsøket ville endre elevenes læring slik dette er målt i standpunktkarakter, kan dermed avvises.

I vitenskapen betegnes et slikt funn som et 0-funn. Funnet forteller om fravær av effekt av et tiltak beregnet på grunnlag av bestemte mål og prosedyrer som er satt for å måle denne effekten. Målene og prosedyrene som vi har brukt for å komme frem til dette funnet, har vi gjort rede for tidligere i denne rapporten. Når man står overfor et 0-funn er det viktig å være klar over to ting. Det første er at fravær av målt effekt ikke er det samme som å slutte seg til at tiltaket må kategoriseres som mislykket. Vi har rike, sammensatte og nyanserte kvalitative data fra implementeringen av forsøket med redusert antall halvårskarakterer i norsk i seks skoler. I det følgende skal vi på nytt vende tilbake til denne datakilden for å nyansere 0-funnet ved å spørre *for hvem* forsøket kan sies å ha hatt nytte. Det andre man bør gjøre når man i forskningen støter på et 0-funn, er å spørre seg om man kunne ha kommet til et annet resultat dersom man hadde nærmet seg problemstillingen på en annen måte. Vi kommer tilbake til dette spørsmålet i avslutningen av kapitlet.

8.2 For hvem har forsøket vært hensiktsmessig?

Utdanningsdirektoratet (2013a) skriver at reduksjon av antall karakterer ved halvårsvurdering i norsk vil gi bedre rom for prioritering av opplæring i grunnleggende ferdigheter, større mulighet til konsentrasjon om emner samt mindre fokus på skriftlige innleveringer og dokumentasjon for å sikre godt nok grunnlag for vurdering. Vi oppfatter disse momentene som ønsker og mål for utviklingen av undervisningen i faget, og gjør dette til utgangspunkt for vår drøfting av hensiktsmessigheten i forsøket.

En positiv formulering av 0-funnet om forsøkets effekt på elevenes standpunktkarakter i norsk er at *færre karakterer ved halvårsvurdering ikke fører til dårligere karakter ved standpunktvurdering*. Når vi ser på hele elevgruppen under ett, taper

elevene ingenting på at lærerne deres setter én eller to karakterer ved halvårsvurdering istedenfor tre karakterer. Vi kan derfor på samme måte som over av vise at forsøket har påvirket elevenes læring på *negativ* måte.

Fra analysen av de kvalitative intervjuene med skoleledere og lærere vet vi at de fleste av disse lærerne har endret praksis under forsøket ved at de nå gir elevene færre vurderinger hvor det settes karakter, og i stedet gir elevene flere og mer avgrensede tilbakemeldinger som ledd i den løpende undervisningen. En mulig tolkning av resultatet er derfor at *med gjeldende læreplan i norskfaget har lærerne i realiteten større frihet til å bestemme over undervisning og vurdering enn hva de selv har oppfattet at de hadde før forsøket*, og slik de gjennom vår evaluering har gitt uttrykk for. Dette gjelder uavhengig av hvorvidt lærerne skal sette tre, to eller én karakter ved halvårsvurdering. I arbeidet med evalueringen har vi blitt gjort kjent med at normer og krav til antall karaktergivende prøver som skal ligge bak hver halvårsvurdering, er bestemt av lokale regler, sedvane og tradisjoner. Vi kan spørre oss om ikke disse lokale reglene og normene kunne vært endret uten et slikt forsøk. Samtidig gir forsøket en indikasjon på at det kan rokkes ved slike lokale tradisjoner uten at elevenes læringsresultater endres, i det minste når vi ser på alle elevenes standpunkt-karakterer under ett.

Gjennom denne evalueringen har vi sett at lærerne verdsetter disse mulighetene både for seg selv og på vegne av elevene. Når omfanget av vurderingssituasjoner med karakter går ned, sier lærerne at arbeidssituasjon deres endres på to måter. Den første endringen innebærer at norsklærerne opplever at rettebyrden minker. Dette gjør at de kan bruke arbeidstiden på å planlegge og utvikle undervisningen for elevene, noe vi i denne evalueringen har sett er viktig for at lærerne skal oppleve at de selv er motivert for jobben sin og kjenner mestring og glede gjennom yrkesutøvelsen. Den andre endringen innebærer at lærerne kan prioritere tiden i undervisningen annerledes når det ikke lenger skal repeteres lærestoff til like mange prøver, og lærerne ikke trenger like mye tid til gjennomgang av rettede prøver sammen med elevene i klasserommet. Vi understreker samtidig at vi ved å snakke med bare et utvalg lærere kun kan få *indikasjoner* på at slike endringer kan skje i lærernes arbeidshverdag, men ingen sikkerhet om at det skjer for alle lærere og ved alle skoler i forsøket. En alternativ hypotese som vårt evalueringsdesign ikke har gitt oss mulighet til å undersøke, er derfor at mange lærere *ikke* har endret sin undervisnings- og vurderingspraksis under forsøket, men bare gitt elever færre karakterer ved halvårsvurdering.

En lærer, som vi har gjengitt i delrapport 2 i denne evalueringen, bekreftet imidlertid at hun hadde gjort slike endringer i undervisnings- og vurderingspraksis under forsøket, og uttrykte at hun opplevde seg selv som en bedre lærer (Seland m.fl. 2016). Forsøket og dette «frirommet» som forsøket skaper i undervisningen, kan derfor i realiteten ha gitt disse lærerne mulighet til å praktisere det de selv lærte

under utdanningen, og som vi kan forutsette var del av deres opprinnelige motivasjon for å bli lærere. Hvis dette er riktig, kan vi si at *forsøket kan ha bidratt til å styrke læreres opplevelse av egen faglighet og mening i læreryrket gjennom det at behovet for karaktergivende prøver gjennom semesteret er blitt redusert*. Vi mener at dette elementet må stå sentralt i en vurdering av hensiktsmessigheten i forsøket.

I en diskusjon av forsøkets hensiktsmessighet må vi imidlertid også lytte til elevene. Vi har allerede sett at elevenes læringsresultater ikke påvirkes av forsøket, men dette er bare målt ved elevenes standpunktkarakterer. Det elevene forteller i intervjuene, gir også informasjon om deres læring og deres interaksjon med læreren gjennom undervisningen.

I noen av skolene vi har besøkt, og særlig i de to ungdomsskolene, har det i norskfaget vært en utvikling som peker mot en generell norm for færre eller ingen vurderinger med karakter underveis i semesteret. Slik vi har sett i denne evalueringen, kan forsøket med færre karakterer ved halvårsvurdering legge til rette for en slik utvikling av vurderingsordningene i norskfaget. Mot denne bakgrunnen er det viktig å legge merke til hvordan enkeltelever vi har snakket med, gir uttrykk for at de under forsøket ikke får informasjon om egen læring som *erstatte* den informasjonen de selv mener at de ville ha fått gjennom karakterer.

På dette punktet kan lærerens forståelse være forskjellig fra elevens forståelse. Et eksempel fra en av ungdomsskolene var tilbakemeldingen som ble kalt for «en respons». Fra lærerens side var dette en beskrivelse av elevens måloppnåelse. For en av elevene syntes det imidlertid som det viktigste kjennetegnet ved en respons var fraværet av en tallkarakter. Dette resulterer i at elevene må gjette på «hvor de ligger», noe som er mer usikkert enn om man får en tallkarakter. Dette synspunktet ble også utfordret gjennom intervjuerens påpekning av at noen ser for seg at karakteren kan stå i veien for at man involverer seg i læring. Til dette svarte en elev at hun også forsto dette argumentet, men hun tilføyde likevel at «personlig foretrekker jeg å vite.»

Vi kan forstå dette som at elevene oppfatter at deres prestasjoner alltid kan vurderes med en karakter, og når de ikke får denne tallkarakteren, tolker de det som om informasjon holdes tilbake. Følgelig blir forskjellen mellom det å være i forsøket og det å ikke være i forsøket bare at karakteren formidles eller den holdes skjult. Elever i forsøket kan følgelig fortsette å «jakte» på informasjonen som de mener at en karakter ville gitt dem, og kan bli frustrerte når de ikke får vite hvordan de «ligger an» fordi karakteren ikke er gjort kjent for dem.

Dette får oss til å spørre om det som skjer med undervisningen i klasserommene når karakterene blir borte, også skaper nye utfordringer for læreren. Lærerne sier at de under forsøket kan bruke et «nytt språk» for å kommunisere med elevene om læring og progresjon, men vi oppfatter det slik at elevene ikke nødvendigvis snakker dette nye «vurderingsspråket» like flytende som lærerne. *Vi vil anta at hvis ikke*

karakterer som uttrykk for elevers læring i faget underveis i semesteret, erstattes av noe annet som er kvalitativt bedre for å hjelpe elevene i læringsprosessen, skjer det sannsynligvis heller ikke noen kvalitativ utvikling av elevens læring. Hvis lærerens nye «vurderingsspråk» ikke blir mottatt, forstått og omsatt i egen utvikling hos elevene, får ikke elevene hva vi kan kalle for et «metablikk» på egen læring. Vi sier ikke med dette at det å få vurdering med karakter er det samme som at elevens læring styrkes, men karakterskalaen er tross alt et språk som elevene kjenner og som de møter i de andre fagene.

Lærere forteller i intervjuene at det å ha færre vurderingssituasjoner med karakter i løpet av semesteret legger til rette for at elevene får bedre tid til å fordype seg i norskfaget. De får tid til å gjøre læringen grundigere og bruke mer tid på de ulike emnene og oppgavene. I vårt materiale har vi utsagn fra lærere om at det å få karakterer på enkeltprøver ikke nødvendigvis er til beste verken for eleven eller for elevens læring. Lærernes argument er at karakteren flytter elevens fokus fra den egentlige læringen. Elevene er inne på det samme argumentet når de diskuterer seg imellom i intervjuene vi har gjort, men for dem ville en tallkarakter være et tydeligere signal om deres læring enn lærerens vurdering av hva de har gjort bra og hva som kan gjøres bedre. For elevene fremstår dermed tallkarakteren, i kraft av sitt fravær, som en hemmelighet som de i beste fall kan gjette seg frem til.

Lærernes og elevenes ulike interesser i det å sette/det å få karakterer settes på spissen når elevene snakker om det å få karakter i sidemål. Da elevene ble spurt om hva som representerte den største overgangen for dem når de i avgangsåret fikk tre karakterer ved halvårsvurdering, var svaret at de nå fikk karakter i sidemål. For noen av dem var dette en brå oppvåkning, mens andre uttrykte misnøye med at de ikke fikk kjennskap til sidemålskarakteren før det var gått mer enn to år av opplæringen. Lærerne på sin side, og særlig lærerne på ungdomsskolene, viser i intervjuene mye omtanke i å gi elevene en forsiktig tilnærming til sidemålet. Enkelte blant lærerne mener også at det å få en egen karakter i sidemålet tidlig i opplæringsløpet kan svekke elevens motivasjon for norskfaget.

Elevutsagnene om disse forholdene kan ha vært påvirket av metoden vi har brukt for å intervjuer elevene, idet spørsmål fra en voksen intervjuer kan gjøre at elevene svarer det de tror at intervjueren vil ha bekreftet. Gruppedynamikken i intervjuene kan også skape en tendens til at hver enkelt elev erklærer seg «enig med siste taler». Samtidig er elevene i en fase av livet hvor det sammenligne seg med andre er viktig for dem, og oppfatninger og synspunkter er i rask endring. *Mest av alt illustrerer elevutsagnene at det å ikke få karakterer er en sak som har flere sider for dem.* Et budskap vi ønsker å formidle gjennom denne sluttrapporten, er at denne flersidigheten ved å ikke få karakterer ved vurderinger, må veies mot fordelene som lærerne sier at forsøket skaper for dem i undervisningen og vurderingen.

Gjennom lærerintervjuene som ble analysert i delrapport 2 (Seland m.fl. 2016) og referert i denne rapporten, har vi fått inntrykk av at lærerne ikke nødvendigvis samarbeider mer systematisk om vurdering under forsøket enn hva de gjorde tidligere. Forsøket gir lærerne større frihet, men dette resulterer altså ikke automatisk i at de gjør ting mer likt. Dette betyr i tilfelle at norsklærerne innenfor en og samme skole kan ha ulik praksis i hvordan de gir elevene tilbakemeldinger uten karakter, noe som skaper forskjeller for elever som går i ulike klasser. Vi kan imidlertid se for oss at forsøket skaper større forskjeller mellom hvordan norsklærerne gir elevene tilbakemeldinger og hvordan lærere med ansvar for andre fag gir elevene tilbakemeldinger etter prøver eller presentasjoner. Kanskje får norsklærerne dermed en annen praksis for hvordan de gir tilbakemeldinger enn hva som er tilfelle for læreren som har de samme elevene i matematikk, naturfag eller KRLE. *Vi kan se for oss at slike ulikheter i læreres vurderingspraksiser – både innenfor samme fag og mellom fag – kan være vanskelige for elevene å orientere seg mot og forstå.*

8.3 Hva styrer undervisning og vurdering i norskfaget?

I kapittel 1 i denne rapporten stilte vi opp fire beskrivelser av sammenhenger mellom vurdering og undervisning. Beskrivelsene ble gjort på bakgrunn av litteraturen som vi har gjennomgått i forbindelse med evalueringen (se også Seland m.fl. 2015). Vi skal her ta for oss de ulike beskrivelsene og diskutere dem i lys av hva evalueringen viser.

Utdanningsdirektoratet utlyste evalueringsoppdraget med bakgrunn i norsklæreres uttalte opplevelser av at vurdering i norskfaget påvirker hvordan lærerne kan drive undervisning i faget. Dette formidlet lærerne til direktoratet gjennom høringsrunden for siste læreplanrevisjon i faget.

I denne evalueringen har vi sett at denne opplevelsen i stor grad samsvarer med de betingelser som de intervjuede lærerne oppfatter at de har for å drive undervisning i norskfaget, idet de mener seg bundet til å forberede, rette, dokumentere og gjennomgå mellom to og tre karaktergivende prøver (eller innleveringer eller elevpresentasjoner) for hver halvårskarakter som skal settes. Tre halvårskarakterer gir dermed lærerne en arbeidsbyrde tilsvarende mellom seks og ni slike karaktergivende prøver per semester, og tilsvarende redusert tid til å drive variert og læringsfremmende undervisning.

Det at lærerne har kjent seg bundet av bestemte arbeidskrav i vurderingen, beror på tolkning av lokale regler og tradisjoner. Det nasjonale regelverket for vurdering setter ingen krav til antall karaktergivende prøver som skal ligge bak hver halvårskarakter. *Antall halvårskarakterer i seg selv kan derfor ikke sies å ha noen innvirkning på undervisningen, men får likevel slik innvirkning gjennom lokal tolkning og håndhevelse av regelverket.*

En beskrivelse av sammenhengen mellom vurdering og undervisning som vi oppfatter som mer nyansert i denne situasjonen, tar derfor hensyn til sammenhengen mellom læreplan og vurdering, slik som vi finner dette omtalt i NOU 2014: 7. Her omtales også aktørenes fortolkning av læreplanen. Dette er inkorporert i forskningsdesignet for denne evalueringen, og etter vårt syn gir dette et bredt og nyansert bilde av hvordan sammenhengen mellom vurdering og undervisning i norskfaget bør forstås. Vår evaluering bekrefter langt på vei at det er disse sammenhengene som er i spill i forsøket med redusert antall karakterer ved halvårsvurdering i norsk. NOU 2014: 7 viser med dette samtidig hvilke faktorer man kan se nærmere på dersom man vil forandre på implementering av læreplanen og vurderingssystemet, altså i hovedsak lærernes undervisning. Dette dreier seg om a) arbeidet med læreplanens form og innhold, og b) aktørenes forståelse og fortolkning av læreplanen og vurderingssystemet.

Vi skal imidlertid ikke slippe Kunnskapsløftet av syne og denne reformens kunnskapssyn, læringssyn og politiske mål for opplæringen. Betydningen av utdanningsreformers kunnskapssyn og læringssyn omtales av Forsberg og Lindberg (2010). Ballet og Kelchtermans (2009) har på sin side vist til hvordan politiske mål for opplæringen kan påvirke lærernes arbeidssituasjon, som igjen kan påvirke deres arbeid med og fortolkning av læreplanen. I norsk sammenheng kan dette forstås som at krav som stilles til elevenes læring og til dokumentering av elevenes læring underveis i opplæringsløpet, påvirker lærernes arbeid med vurdering og dermed deres undervisning. Disse kravene er nedfelt i Kunnskapsløftet forstått både som læreplanreform og styringsreform, som er innrettet med hensikt om å øke og styrke elevenes læring. Med tanke på vårt funn om at lærerne kjenner seg bundet av lokale regler og tradisjoner i antall vurderingssituasjoner som skal ligge til grunn for hver halvårskarakter, tror vi at vi nærmer oss en god virkelighetsbeskrivelse her. Aktørenes forståelse og fortolkning av læreplanen, slik utvalget bak NOU 2014: 7 gjør dette til en sentral faktor for implementering av læreplanen, vil nemlig være sterkt influert av slike regler eller signaler.

Vi kan oppsummere dette som at evalueringen viser at lærernes undervisning påvirkes av lærernes og skoleledernes forståelse og fortolkning av læreplan og regelverket for vurdering, understøttet av fylkesmannens råd, veiledning og eventuelle krav til dokumentasjon og behandling i klagesaker. Slik vil lærernes arbeid med vurdering påvirke deres undervisning, men dette skjer gjennom en fortolkning av hvilke krav læreplanen og styringsdokumentene for Kunnskapsløftet stiller til vurdering og dokumentasjon av vurdering, som igjen har rot i et overgripende syn på læring og styring av aktiviteten i skolen. Dette er svært komplekse sammenhenger. Når vi gjennom denne evalueringen har fått demonstrert at elevenes standpunkt-karakterer ikke påvirkes av å ha færre karakterer ved halvårsvurdering, synes de

øvrige sammenhengene som her er beskrevet, å ligge åpne for debatt og videre utvikling.

8.4 Anbefalinger for videre utvikling av vurdering i norskfaget

Norskfaget er, som tidligere beskrevet, et stort, innholdsrikt og på mange måter svært fordrende fag. Vi ser ikke at forsøket med redusert antall karakterer ved halvårsvurdering har bidratt til å løse årsakene til utfordringene som synes innebygget i norskfaget, og som vi har beskrevet i kapittel 1 i denne rapporten og dessuten i delrapport 1 fra evalueringen (Seland m.fl. 2015). Ifølge Forsberg og Lindberg (2010: 30) er det nødvendig å gå tilbake ikke bare til læreplanen, men også til kunnskapssyn og læringssyn som ligger til grunn for læreplanen hvis man skal diskutere vurderingsordningen i et fag. Dette ligger utenfor evalueringens mandat.

Hvis vår analyse av elevers standpunktkarakterer hadde vist signifikante forskjeller i resultater mellom elever som har deltatt og ikke deltatt i forsøket, eller som har hatt én kontra to karakterer i halvårsvurdering i forsøksperioden, ville dette dannet utgangspunkt for en videre anbefaling av en mer systematisk utprøving av forsøket. Når vi ikke kan vise til slike resultater av forsøket, må våre anbefalinger bli rent kvalitative. Vi skal derfor gå nærmere inn i to slike forhold som evalueringen gir oss grunnlag for å diskutere, nemlig 1) fortolkning av krav til dokumentasjon av elevenes læring og 2) lærernes kommunikasjon med elevene om vurdering underveis i opplæringen.

Det at såpass mange av de intervjuede norsklærerne mener at forsøket har gjort det mulig for dem å endre sin vurderings- og dermed også undervisningspraksis, gjør det nærliggende å stille spørsmålet om det i utgangspunktet er nødvendig med to til tre karaktergivende prøver (eller innleveringer eller presentasjoner) for hvert delområde i faget? Delområdene i norskfaget i læreplanen svarer til tredelingen i fagets eksamensordning, som gjenspeiles i ordningen med standpunktkarakterer i faget. Forsøket har ikke omfattet elevenes avgangår, så de skulle verken ha standpunktkarakterer eller eksamen de to årene de deltok i ordningen. (Ingen lærere vi har snakket med, tok til orde for å utvide ordningen til også å gjelde avgangsåret.) Lærerne forteller at forsøket ikke endrer deres arbeid med kompetansemålene i faget, noe som bør tilsi at elevene får den undervisningen de har krav på – og undervises i det samme uavhengig av antall karakterer ved halvårsvurdering.

Vi har riktignok sett at enkeltlærere har bekymret seg for om elevene får tilstrekkelig undervisning i sidemål under forsøket, fordi forsøket gjør det mulig å «komprimere» og skyve på dette delområdet i undervisningen til læreren mener at elevene er mer modne og mottakelige. Lærernes bekymring kan imidlertid også peke på at det er andre utfordringer med undervisningen i sidemål enn bare at

dette skal vurderes med egen halvårskarakter, slik vi også har vært inne på i kapittel 1 i denne rapporten. Hvis sidemålsundervisningen dreier seg om å ha én norsktime i uka om bøyningsformer, vokabular og setningsbygning i sidemål, blir dette åpenbart til «noe annet» enn opplæringen i hovedmålet. Dette er kanskje en medvirkende forklaring på at lærere mener at elever må «skjermes» for sidemålet tidlig i ungdomsskolen. Lærere i begge skoleslag har imidlertid også fortalt at mer tid til skriving i faget gir dem mulighet til å integrere hovedmålet og sidemålet i undervisning og vurdering på en for dem ønskelig og gunstig måte.

En av våre to anbefalinger på grunnlag av denne evalueringen er derfor at utdanningsmyndighetene, utdanningsforvaltningen og lærerprofesjonen går dypere inn i spørsmålet om hvor mange vurderingssituasjoner man virkelig trenger for å sette halvårskarakter. Dette kan diskuteres uavhengig av fag og uavhengig hvor mange halvårskarakterer som skal settes, fordi vi har sett i evalueringen at det kun er dette antallet og ikke læreplanen eller kompetansemålene som avgjør hvor mange vurderingssituasjoner med karakter som læreren vil gjennomføre pr. semester. Dette spørsmålet synes derfor å kunne avgjøres rent juridisk, men må samtidig ta hensyn til de krav og avveininger som Kunnskapsløftet setter for elevers rettigheter og dokumentasjon av at rettighetene er oppfylt.

Det nasjonale regelverket for vurdering oppgir ingen norm for antall vurderingssituasjoner med karakter som skal ligge til grunn for halvårskarakteren. Vi antar derfor at skoleledernes og lærernes oppfatning om at det eksisterer slike normer kommer fra skoleeier eller også fra fylkesmannens veiledning og behandling av klagesaker. Derfor spør vi: kan man se for seg andre, valide og for elevene lett forståelige fremgangsmåter for å samle og formidle data om elevers læringsprosess gjennom semesteret som i større grad kan erstatte de karaktergivende prøvene og innleveringene? Og: kan disse metodene og fremgangsmåtene gi personlig, faglig, profesjonell og juridisk trygghet for lærere, elever, foresatte, skoleledere, skoleeiere, fylkesmannen og de nasjonale utdanningsmyndighetene om kvaliteten i de samme læringsprosessene?

Vi tror det er her innsatsen i den fremtidige utviklingen av vurderingsordninger i norskfaget må ligge. Prosessene som omfatter og knytter sammen undervisning, vurdering og læring synes å være langt mer komplekse enn hva et forenklet spørsmål om hva som er best av én, to eller tre halvårskarakterer kan besvare. Snarere synes *hvordan man faktisk vurderer* og *hvordan man snakker om vurdering* å være det som kan skape endring i elevers læring. Hvis disse problemstillingene anses som de viktigste for elevenes læring, skyves krav om antall vurderingssituasjoner med karakter i bakgrunnen. Vi står da igjen med det overordnede spørsmålet om hva slags læringssituasjon vi ønsker for elevene.

Dette bringer oss inn på vår andre anbefaling for utvikling av vurderingsordningene i norsk, og som dreier seg om kvaliteten i det vurderingsfaglige språket

som lærerne benytter i møte med elevene, dersom karakterer underveis i semesteret skal erstattes av noe annet. Vi ser behovet for at det arbeides mer med å utvikle et vurderingsspråk som elevene forstår, og som dermed representerer et bedre redskap for både lærer og elev i elevens læringsprosess. Dette vurderingsspråket kan ikke være et pseudokaraktersystem med alternative skalaer eller nivåinndelinger, men gi elevene det vi har kalt et metablikk på egen læring.

Et av forskningsspørsmålene for denne evalueringen har vært om forsøket fører til at elevene motiveres mer eller mindre for norskfaget. Vi vet at det er en vekselvirkning mellom elevers motivasjon og prestasjoner i et fag. Når vi ikke ser en utvikling i elevers prestasjoner, gjør dette det vanskelig å argumentere for at elevers motivasjon for norskfaget er styrket under forsøket. Dette betyr ikke at vi skal utelukke at forsøket, med et potensiale for tid til å benytte nye arbeidsformer og mulighet til fordypning, kan ha bygget opp elevers interesse for faget eller måter å arbeide på, det å fordype seg – på sikt. Det er sannsynlig at det kan skje, og kanskje gagnar det elevene som går videre til høyere utdanning hvor det vil kreves at man jobber langsiktig.

Selv om elevene ikke er umiddelbart enige, tror vi på lærerne som forteller at vurdering med karakter flytter elevenes oppmerksomhet fra selve læringsprosessen og over til læringsprosessens resultat. Våre samtaler med lærere og elever viser samtidig at det kan være vanskelig å få elevene til å tenke annerledes, å få dem til å «lytte» til noe annet enn en tallkarakter. Dette gjør at vi stiller spørsmålet om lærerne under forsøket virkelig har gitt elevene noe annet og kvalitativt bedre redskap til å utvikle læringen når de samtidig har redusert løpende vurderinger med karakterer?

Man kan forstå elevene. I intervjuene stadfester de at de orienterer seg om avgangsvitnemål og muligheter i fremtiden. Alle signaler de oppfatter i og utenfor skolen, tilsier at karakterer er betydningsfulle og noe de må jobbe hardt med. For elever som skal søke seg til høyere utdanning, har også eksamens- og standpunkt-karakterer en helt reell betydning og verdi. Vi har også sett av intervjuene med elevene at foreldres forventninger om gode prestasjoner spiller inn på elevenes opplevelse av karakterer som svært betydningsfullt, og at det også kan medføre opplevelser av press og stress for elevene. Kanskje er det først når samfunnets syn på karakterens betydning underveis i opplæringsløpet endres, at elevene får slappe av og konsentrere seg om å lære underveis og spare prestasjonene til slutt?

8.5 Avslutning

I denne rapporten har vi ved hjelp av intervjuer med særlig lærere, men også med flere elever, gitt mange argumenter for at elevenes læringsresultater kan bli bedre når undervisningen i norskfaget endres som følge av at lærer har bruk for færre vurderingssituasjoner med karakter i løpet av semesteret. Likevel viser analysen

av karakterdata at det ikke er noen forskjell i læringsresultater mellom elever ved skoler som har deltatt i forsøket sammenlignet med elever ved skoler som ikke har vært del av forsøket.

Dette kan tyde på at forsøket har avdekket noen «blindsoner» i vurderingsarbeidet som forsøket i seg selv ikke har bidratt til å løse. Vi har i dette kapittelet omtalt disse blindsonene som mangler i kunnskap om hvordan man snakker om vurdering når man reduserer bruken av karakterer for å gi eleven signal om nivå og eventuell progresjon i opplæringen, og vi anbefaler at dette spørsmålet blir gjenstand for eventuell fremtidig forskning om vurdering i norskfaget og i andre skolefag.

Hvis målet med forsøket har vært å gi lærerne en enklere og mer motiverende arbeidshverdag, synes dette på den annen side å ha lyktes. Vi setter imidlertid spørsmålstegn ved hvorvidt det var nødvendig med et slikt forsøk for at lærerne skulle redusere omfanget av vurderingssituasjoner med karakter underveis i semesteret. Normen som disse lærerne endrer på, synes å være begrunnet med behov for dokumentasjon i behandling av eventuelle klager, og ikke pedagogiske hensyn.

Et tankekors i avslutningen av denne diskusjonen er at det ikke synes å være noen enkel fremgangsmåte for å endre på de signalene og forestillingene om karakterens betydning som elevene bærer med seg inn i ungdomsskolen og gjennom årene i videregående skole. Arbeidet med å utvikle vurderingsordningene bør derfor ikke bare gjelde norskfaget, men inkorporere alle skolefagene og derfor alle lærerne.

Referanser

- Angrist, Joshua D. & Jörn-Steffen Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton, NJ: Princeton University Press
- Apple, M. V. (1986). *Teachers and Texts. A political economy of class and gender relations in educations*. London: Routledge.
- Baird, J.; Hopfenbeck, T.; Newton, P.; Stobart, G. & Steen-Utheim, A. (2014). *State of the Field Review Assessment and Learning*. Case nr. 13/4697, Norwegian Knowledge Centre for Education. Oslo: Kunnskapssenter for utdannning.
- Bakke, J. O. og Kverndokken, K. (2010). "Å vurdere muntlighet – fra et norskfaglig ståsted" I: Dobson, S.; Engh, R. og Lied, S. (red.) *Vurdering for læring i fag*. S. 60-74. Kristiansand: Høgskoleforlaget.s
- Ballet, K. & Kelchtermans, G. (2009). «Struggling with workload. Primary teachers' experience of intensification.» *Teaching and Teacher Education*, 25(2009), 1150-1157. DOI: 10.1016/j.tate.2009.02.012
- Berge, K. L. (2014). "Fra eksamensskrivning i norsk til skriving som grunnleggende ferdighet – to vurderingsformer." I: Hvistendahl, R. E. og Roe, A. (red.) *Alle tiders norsksdidaktiker. Festskrift til Frøydis Hertzberg på 70-årsdagen*. Oslo: Novus forlag.
- Berge, K. L. (2005). "Skriving som grunnleggende ferdighet og som nasjonal prøve – ideologi og strategier." I: Aasen, A. J. og Nome, S. (red.) *Det nye norskfaget*. Bergen: Fagbokforlaget og Landslaget for norskundervisning.
- Berge, K. L.; Evensen, L. S.; Hertzberg, F. og Vagle, W. (2007). "Det umulige er nesten mulig. Om norskeksamen på ungdomsskolen og utvikling av vurderingsnormer" I: Tveit, S. (red.) *Elevvurdering i skolen. Grunnlag for kulturendring*. S. 74-84. Oslo: Universitetsforlaget.
- Black, P.; Harrison, C.; Lee, C. S.; Marshall, B & Wiliam, D. (2003). *Assessment for learning, putting it into practice*. Berkshire: Open University Press.
- Black, P. J. & Wiliam, D. (2003) 'In praise of educational research': formative assessment, *British Educational Research Journal*, 29(5), 623–637.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Bransford, J. D.; Brown, A. L. & Cocking, R. R. (2000). *How people learn*. Washington D. C.: National Academy Press.

- Breivik, T. (2011). «Haldningar til eige fag – ei undersøking blant norsklærarar.» *Norsklæreren* nr. 1, s. 42-43.
- Buland, T. (2014). «Det avhenger mye av læreren – elevers opplevelse av vurdering og læring». I: Sandvik, L. V. og Buland, T. (red.) *Vurdering i skolen. Utvikling av kompetanse og fellesskap*. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS). Kapittel 5, s. 73-86. Trondheim: NTNU Program for lærerutdanning og SINTEF Teknologi og samfunn.
- Byrhagen, Karen N., Torberg Falch & Bjarne Strøm (2006): *Frafall i videregående opplæring: Betydningen av grunnskolekarakterer, studieretninger og fylke*. SØF-rapport nr. 08/06, Trondheim: Senter for Økonomisk Forskning AS
- Bøyum, B. og Norendal, L. (2011). «Tre karakterer i norsk – en anakronisme?» *Norsklæreren*, nr. 1, s. 27-29.
- Columbia University (2018): Difference-in-Difference Estimation. Columbia University Mailman School of Public Health. Tilgjengelig på nettet: <https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation>
- Dobson, S.; Engh, R. og Lied S. (red.) (2010). *Vurdering for læring i fag*. Kristiansand: Høyskoleforlaget.
- Duncan, C. R. & Noonan, B. (2007). "Factors Affecting Teachers' Grading and Assessment Practices." *The Alberta Journal of Educational Research*, 53(1), 1-21.
- Dæhlen, M.; Smette, I. og Strandbu Å. (2011) *Ungdomsskoleelevers meninger om skolemotivasjon*. NOVA-rapport 4/11. Oslo: NOVA.
- Eccles, J. & Wang, M.-T. (2012). "So What Is Student Engagement Anyway?" I: Christenson, S.L., Reschly, A.L. & Wylie, C. (eds.) *Handbook on Research on Student Engagement*. New York: Springer.
- Eggen, A. E. (2004). *Alfa and Omega in Student Assessment; Exploring Identities of Secondary School Science Teachers*. PhD thesis, Department of Teacher Education and School Research, University of Oslo. Oslo: University of Oslo.
- Eide, O. (2012). «Norskfaget = norsk språk i bruk. Om rammer for eit heilskapeleg norskfag» *Norsklæreren*, nr. 1/2012, s. 60-64.
- Evensen, L. S. (2014). "Men kan det komme noe godt fra lærerværelset? Om validitet i læreres vurdering" I: Hvistendahl, R. E. og Roe, A. (red.) *Alle tiders norskdidaktiker. Festskrift til Frødis Hertzberg på 70-årsdagen*. Oslo: Novus forlag.
- Evensen, L. S. (2003). *Kvalitetssikring av læringsutbyttet i norsk skriftlig, KAL-prosjektet. Sammendragsrapport*. Trondheim: NTNU.
- Federici, R. A.; Gjerustad, C.; Vaagland, K.; Larsen, E. H.; Rønsen, E. og Hovdhaugen, E. (2017). *Spørsmål til Skole-Norge våren 2017*. NIFU-rapport 2017/12. Oslo: NIFU

- Forsberg, E. og Lindberg, V. (2010). *Svensk forskning om bedömning – en kartläggning*. Rapport 2: 2010. Stockholm: Vetenskapsrådet.
- Fjørtoft, H. (2013). "Vurderingspraksiser i norskfaget". I: Sandvik, L. V. og Buland, T. (red.) *Vurdering i skolen. Operasjonalisering og praksier*. Delrapport 2 fra prosjektet Forskning på individuell vurdering i skolen. S. 99-118. Trondheim: NTNU Skole- og lærerforskning Program for lærerutdanning og SINTEF Teknologi og samfunn.
- Fjørtoft, H. (2010). "Vurdering og læring I norsk skriftlig. Mot en kriteriebasert skrivepedagogikk.» I: Dobson, S.; Engh, R. og Lied, S. (red.) *Vurdering for læring i fag*. Kapittel 4, s. 49-59. Kristiansand: Høgskoleforlaget.
- Gjerustad, C. og Waagene, E. (2015). *Spørsmål til Skole-Norge høsten 2015. Resultater og analyser av Utdanningsdirektoratets spørreundersøkelse blant skoleledere og skoleeiere*. NIFU-rapport 42/2015. Oslo: NIFU.
- Gough, D. & Thomas, J. (2012). «Commonality and diversity in reviews». I: Gough, D.; Oliver, S. & Thomas, J. (eds.) *An introduction to systematic reviews*, pp. 25-65. London: Sage.
- Gubrium, J. F. & Holstein, J. A. (1997). *The New Language of Qualitative Method*. New York: Oxford University Press.
- Gaare, O. (2006). "Norskfaget og nasjonallitteraturen." *Norsklæreren*, nr. 2/2006, s. 22-24.
- Halkier, B. (2012). «Fokusgrupper» I: Brinkmann, S. og Tanggaard, L. (red.) *Kvalitative metoder. Empiri og teoriutvikling*. Kapittel 5, s. 133-152. Oslo: Gyldendal Akademisk.
- Hamre, P. (2004) *Vurdering og pedagogisk utviklingsarbeid i norskfaget*, Rapport frå PLUTO-prosjektet ved Høgskulen i Volda, Volda: Arbeidsrapport 154.
- Harlen, W. & James, M. (1997) Assessment and learning: differences and relationships between formative and summative assessment, *Assessment in Education: principles, policy and practice*, 4(3), 365–379.
- Hertzberg, F. (2008). "Assessment in writing in Norway: A case of balancing dilemmas" I: Havnes, A. & McDowell, L. (eds.) *Balancing dilemmas in assessment and learning in contemporary education*. S. 51-60. New York: Routledge.
- Hertzberg, F. (2007). "Vurderingprosjekt i norskfaget – kommentar fra KAL-forskere" *Norsklæreren*, nr. 2/2007. S. 12-23.
- Hodgson, J.; Rønning, W.; Skogvold, A. S. og Tomlinson, P. (2010). *Vurdering under Kunnskapsløftet. Læreres begrepsforståelse og deres rapporterte og faktiske vurderingspraksis*. NF-rapport nr. 17/2010. Bodø: Nordlandsforskning.
- Hopfenbeck, T.; Tolo, A.; FLopez, T. & El Masri, Y. (2013). *Balancing Trust and Accountability? The Assessment for Learning Programme in Norway. A Governing Complex Education Systems Case Study*. Paris: OECD.

- Hovdhaugen, E.; Seland, I.; Lødding, B.; Prøitz, T. S. og Vibe, N. (2014). *Karakterer i offentlige og private videregående skoler. En analyse av eksamens- og standpunktkarakterer i norsk og matematikk og rutiner for standpunktvurdering i offentlige og private videregående skoler*. NIFU-rapport 24/2014. Oslo: NIFU.
- Hvistendal, R. E. og Roe, A. (2014). *Alle tiders norsksdidaktiker. Festskrift til Frøydis Hertzberg på 70-årsdagen*. Oslo: Novus forlag.
- Haaland, Ø. (2009). "Dannelsesbegrepet og den postmoderne erfaring." I: *Norsk pedagogisk tidsskrift*, 93(5), s. 415-421.
- Iglund, M.-A. (2009). *Mens teksten blir til: Ein kasusstudie av lærarkommentarar til utkast*. Avhandling ph.d., Institutt for lærerutdanning og skoleutvikling, Det utdanningsvitenskapelige fakultet, Universitetet i Oslo. Oslo: Universitetet i Oslo.
- Iversen, H. og Otnes, H. (2011). «Grammatikken i bruk – i arbeid med tekster» *Norsklæreren*, nr. 3/2011, s. 18-25.
- Karseth, B. og Engelsen, B. U. (2013). «Læreplanen for Kunnskapsløftet: Velkjente tråkk og nye spor». I: Karseth, B.; Møller, J. og Aasen, P. (red.) *Reformtakter. Om fornyelse og stabilitet i grunnopplæringen*. Kapittel 3, s. 43-60. Oslo: Universitetsforlaget.
- Kjelen, H. (2013). *Litteraturundervisning i ungdomsskolen. Kanon, danning og kompetanse*. Avhandling ph.d., Institutt for språk og litteratur, Det humanistiske fakultet, NTNU. Trondheim: Norges Tekniske og Naturvitenskapelige Universitet (NTNU).
- Kringstad, T. og Kvithyld, T. (2013). «Kva er fagskriving i norskfaget?» *Norsklæreren*, nr. 4/2013, s. 11-36.
- Kunnskapsdepartementet (2013). *Læreplan i norsk*.
<https://www.udir.no/Upload/larerplaner/Norsk/200613-Lareplan-norsk-fastsatt.pdf>
- Lechner, Michael (2010): The Estimation of Causal Effects by Difference-in-Difference Methods, *Foundations and Trends in Econometrics* 4(3):165-224
- Lekholm, A. K. & Cliffordson, C. (2009). «Effects of Student Characteristics on Grades in Compulsory School.» *Educational Research and Evaluation*, 15(1), 1-23.
- Lekholm, A. K. & Cliffordson, C. (2008). "Discrepancies between School Grades in Compulsory School." *Educational Research and Evaluation*, 15(1), 1-23.
- Lervaag, N. J. (2010). *Muntlige framføringer i et retorisk perspektiv. En undersøkelse om undervisning i retorikk og elevenes evne til å bruke retorikken i arbeid med muntlige framføringer*. Masteroppgave i norsksdidaktikk. Institutt for lærerutdanning og skoleutvikling, Utdanningsvitenskapelig fakultet, Universitetet i Oslo. Oslo: Universitetet i Oslo.

- Markussen, Eifred, Martin Flatø & Rune Borgan Reiling (2017):
Skolebidragsindikatorer i videregående skole: Indikatorer for skolars bidrag til fullføring, gjennomføring og gjennomsnittskarakterer for skoleåret 2014-2015 og kullet som begynte i videregående skole høsten 2012. NIFU-rapport 2017:7. Oslo: NIFU.
- Mausethagen, S. og Kostøl, A. (2009). *Relasjonen mellom lærer og elev og lærerens undervisningspraksis. En casestudie av læreres forståelser, klasseromdiskurs og elevperspektivet i fire klasserom på 7. trinn.* Rapport nr. 18-2009. Elverum: Høgskolen i Hedmark.
- McMillan, J. H. (2003). "Understanding and Improving Teachers' Classroom Assessment Decision-Making: Implications for Theory and Practice." *Educational Measurement: Issues and Practice*, 22(4), 34-43.
- Mjåtveit, A. og Jakobsen, A. (2013). "Grunnskolelæreres undervisningspraksis." I: Reinertsen, A. B.; Groven, B.; Knutas, A. og Holm, A. (red.) *FoU i praksis. Artikkelsamling fra konferanse om praksisrettet FoU i lærerutdanning.* Side 184-191. Oslo: Akademika forlag.
- Moe, T. O. (2013). *Norsklæreren om norskfaget. En kvalitativ diskursanalyse.* Masteroppgave i norskdidaktikk, Høgskolen i Sør-Trøndelag, avdeling for lærer- og tolkeutdanning. Trondheim: HiST.
- Moslet, I. (2009) «Norsklærer». I Smidt, J. (red.) *Norskdidaktikk – ei grunnbok*. 3. utgave. S. 23-36. Oslo: Universitetsforlaget.
- Møller, J.; Prøitz, T. S.; Rye, E. og Aasen, P. (2013). «Kunnskapsløftet som styringsreform» I: Karseth, B.; Møller, J. og Aasen, P. (red.) *Reformtakter. Om fornyelse og stabilitet i grunnopplæringen.* Kapittel 2, s. 23-41. Oslo: Universitetsforlaget.
- Måseide, A. (2012). «Norsk språk og danning.» *Norsklæreren* nr. 2/2013, s. 39-44.
- Nasjonalt kompetansemiljø i vurdering (2012). «Teoretisk bakgrunnsdokument for arbeid med vurdering for læring på ungdomstrinnet.» Høgskolen i Vestfold, Høgskolen i Oslo og Akershus, Høgskolen i Telemark, Høgskolen i Lillehammer, NTNU og Høgskolen i Volda.
https://www.udir.no/PageFiles/35149/Ungdomstrinnet_Bakgrunnsdokument_vurdering_for_laring_vedlegg_4.pdf
- Nicolaysen, B. K. (2005). «Tilgang på deltaking. Nynorsken si rolle for norsk demokrati.» I: Nordal, A. S. (red.) *Didaktiske perspektiv på nynorskopplæring. Skrifter frå Nasjonalt senter for nynorsk i opplæringa.* Nr. 1. Volda: Nasjonalt senter for nynorsk i opplæringa.
- Nordenbo, S. E.; Larsen, M. S.; Tiftikci, N.; Wendt, R. E. & Østergaard, S. (2008). *Lærerkompetanser og elevenes læring i førskole og skole.* København: Danish Clearinghouse for Educational Research.

- NOU 2014: 7 *Elevenes læring i fremtidens skole. Et kunnskapsgrunnlag*. Utredning fra offentlig utvalg oppnevnt ved kongelig resolusjon 21. juni 2013, avgitt til Kunnskapsdepartementet 3. september 2014.
- Nusche, D.; Earl, L.; Maxwell, W. & Shewbridge, C. (2011). *OECD Reviews og Evaluation and Assessment in Education: Norway*. Paris: OECD.
- Nyen, T.; Jordfald, B. og Seip, Å. A. (2009). *Tidstyvene. En beskrivelse av lærernes arbeidssituasjon*. Fafo-rapport 2009:23. Oslo: Fafo.
- Nygård, J. og Sture, V. (2013). «Kva skal vi med sidemål?» *Syn og segn*, nr. 1/2013, s. 40-46.
- Nyheim, I.-L. (2012). «Norskfaget slankes» *Ordet: kvartalsskrift for norsk språk og kultur*, nr. 3/2012, s. 5.
- Nyheim, I.-L. (2011). «Uro i norskfaget» *Ordet: kvartalsskrift for norsk språk og kultur*, nr. 3/2011, s. 7.
- Parr, S. (2012). "Classroom assessment in writing" I: McMillan, J. H. (ed.) *SAGE Handbook of Research on Classroom Assessment*. London: Sage Publications
- Penne, S. og Hertzberg, F. (2008). *Muntlige tekster i klasserommet*. Oslo: Universitetsforlaget.
- Prøitz, T. S. (2013). «Variations in grading practice – subjects matter.» *Education Inquiry*, 4(3), 1-22.
- Prøitz, T. S. og Borgen, J. S. (2010). *Rettferdig standpunktbedømming. Det (u)muliges kunst?* NIFU STEP-rapport 16/2010. Oslo: NIFU STEP.
- Resh, N. (2009). «Justice in Grades Allocation: Teachers' Perspective.» *Social Psychology of Education*, 12(3), 315-325.
- Rise, S. E. (2007). *Motivasjon for nynorsk. En studie av klasseromslederes, skolelederes og elevers egen motivasjon for undervisning og læring i nynorsk*. Masteroppgave ved program for utdanningsledelse, Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo. Oslo: Universitetet i Oslo.
- Røed, D. F. (2010). *Mål og mening. Didaktiske perspektiv på nynorsk med fokus på elevhaldningar*. Masteroppgave i kultur- og språkfagenes didaktikk, fordypning i norsk. Høgskolen i Hedmark.
- Rønning, W. (2012). *Fra karakterjag til læring – fra dommer til trener. Rapport fra et forsøk med én karakter i norsk ved to videregående skoler*. Arbeidsnotat nr. 1005/2012. Bodø: Nordlandsforskning.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144. doi:10.1007/BF00117714
- Sandvik, L. V. (2014). «Hva innebærer validitet i vurdering?» I: Sandvik, L. V. og Buland, T. (red.) *Vurdering i skolen. Utvikling av kompetanse og fellesskap*. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS).

- Kapittel 3, s. 43-58. Trondheim: NTNU Program for lærerutdanning og SINTEF Teknologi og samfunn.
- Sandvik, L. V. og Buland, T. (red.) (2014) *Vurdering i skolen. Utvikling av kompetanse og fellesskap*. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS). Trondheim: NTNU Program for lærerutdanning og SINTEF Teknologi og samfunn.
- Sandvik, L. V.; Engvik, G.; Fjørtoft, H.; Langseth, I. D.; Aaslid, B. E.; Mordal, S. og Buland, T. (2012). *Vurdering i skolen. Intensjoner og forståelser. Delrapport 1 fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)*. Trondheim: NTNU Program for lærerutdanning og SINTEF Teknologi og samfunn.
- Sannrud, A. W. (2010). «Norsk lektorlag og arbeidsbyrden i norskfaget». *Norsklæreren*, nr. 2/2010, s. 74-75.
- Seland, I.; Gjerustad, C. og Lødding, B. (2016) Delrapport 2 fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk. NIFU-rapport 29/2016. Oslo: NIFU.
- Seland, I.; Lødding, B. og Prøitz, T. S. (2015). *Delrapport 1 fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk. Litteraturstudie*. NIFU-rapport 33/2015. Oslo: NIFU.
- Skedsmo, G. (2011). Formulation and realisation of evaluation policy: inconsistencies and problematic issues. *Educational Assessment, Evaluation and Accountability*, 23(1), 5–20.
- Skirbekk, G. (2013a). «Norskfaget – kva skal vi med det?» *Norsklæreren*, nr. 1/2013, s. 53-59.
- Skirbekk, G. (2013b). «»Eit norskfag i krise» *Syn og segn*, nr. 1/2013, s. 53-59.
- Steinfeld, T. (2012a). «Norskplanen som verneplan? Om kulturarv i norskfaget» I: *Norsklæreren*, nr. 2/2012 s. 29-34.
- Steinfeld, T. (2012b). «Umåtelig eller umistelig – norskfaget under revisjon». *Bedre Skole*, nr. 3, 2012, s. 83-87.
- Stensby, M. (2011). «Hva skal ut?» *Norsklæreren*, nr. 1/2011, s. 18-19.
- St.meld. nr. 35 (2007-2008) *Mål og mening – ein heilskapeleg norsk språkpolitikk*. Oslo: Kultur- og kirkedepartementet.
- St.meld. nr. 16 (2006-2007) *...og ingen sto igjen. Tidlig innsats for livslang læring*. Oslo: Kunnskapsdepartementet.
- St.meld. nr. 30 (2003-2004) *Kultur for læring*. Oslo: Utdannings- og forskningsdepartementet.
- Stobart, G. (2008). *Testing times. The uses and abuses of assessment*. London: Routledge.
- Stock, James H. & Mark W. Watson (2015). *Introduction to Econometrics*, Third Edition. Boston: Pearson

- Svenkerud, S. (2013). *Opplæring i muntlige ferdigheter på 9. trinn*. Avhandling ph.d. ved Det utdanningsvitenskapelige fakultet, Oslo, Universitetet i Oslo.
- Svenkerud, S.; Klette, K. og Hertzberg, F. (2012). "Teaching oral skills" *Studies in Education*, vol. 32, pp. 35-49.
- Søyland, A. (2012). "Sidemål = språklig styrke". *Syn og Segn*, nr. 1/2012, s. 4-9.
- Såheim, A. S. (2012). *Lærernes skriftlige prosessrespons, slik elevene forstår den. Verbale protokoller med elevreaksjoner på lærerkommentarer*. Masteroppgave i norskdidaktikk ved Institutt for lærerutdanning og skoleutvikling, Det utdanningsvitenskapelige fakultet, Universitetet i Oslo. Oslo: Universitetet i Oslo.
- Taras, M. (2005) Assessment—summative and formative—some theoretical reflections, *British Journal of Educational Studies*, 53(4), 466–478.
- Titlestad, A. (2010). "Hvorfor kan de ikke få velge?" *Norsklæreren*, nr. 2/2010.
- TNS Gallup (2006). *Kartlegging av sidemålsundervisningen på 10. trinn og VKI allmennfaglig studieretning – resultater fra en nasjonal undersøkelse blant elever, norsklærere og rektorer*. Oslo: TNS Gallup.
- Tveit, S. (2014). «Profiles of education assessment systems worldwide: Educational assessment in Norway». *Assessment in Education*, 21(2), 221-237.
- Utdanningsdirektoratet (2016). Individuell vurdering Udir-5-2016 (merknader til regelverket om vurdering i forskrift til opplæringsloven kapittel 3) <https://www.udir.no/regelverk-og-tilsyn/finn-regelverk/etter-tema/Vurdering/individuell-vurdering-udir-5-2016/>
- Utdanningsdirektoratet (2013a). «Oppsummeringer av høringsuttalelser og tilråding til endringer i læreplanen i norsk» https://www.udir.no/Upload/larerplaner/Utkast/gjennomgaende/forslag_KD_100413/Vedlegg_4-norsk_oppsummering.pdf?epslanguage=no
- Utdanningsdirektoratet (2013b). «Norsk – veiledning til læreplan». Publisert 20. august 2013, sist oppdatert 19. august 2015. <https://www.udir.no/laring-og-trivsel/lareplanverket/veiledning-til-lp/norsk---veiledning-til-lareplan/>
- Utdanningsdirektoratet (2011). *Norsk landrapport til OECD*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2006). *Framtidas norskfag. Språk og kultur i eit fleirkulturelt samfunn*. Rapport. Arbeidsgruppe nedsatt av Utdannings- og forskningsdepartementet. Oslo: Utdanningsdirektoratet.
- Wyatt-Smith, C. & Klenowski, V. (2012). Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assessment in Education. Principles, Policy & Practice*, 20(1), 35–52.
- Aase, L. (2005). «Skolefagenes ulike formål – danning og nytte.» I: Bjørhaug, K.; Fenner, A. B. og Aase, L. (red.) *Fagenes begrunnelser. Skolens fag og arbeidsmåter i et dannelsesperspektiv*. Bergen: Fagbokforlaget.

Aasen, P.; Møller, J.; Rye, E.; Ottesen, E; Prøitz, T. S. og Hertzberg, F. (2012).
*Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes
og institusjonenes rolle i implementeringen av reformen.* NIFU-rapport 20/2012.
Oslo: NIFU og Institutt for lærerutdanning og skoleforskning (ILS),
Universitetet i Oslo.

Vedlegg

Følgende er vedlegg til rapporten:

1. Oversikt over forskningsspørsmål i evalueringen
2. Regresjonstabeller for analyse av standpunktkarakterer i kapittel 4

Vedlegg 1: Oversikt over forskningsspørsmål i evalueringen

Operasjonaliserte problemstillinger til videre undersøkelse
(Tidligere gjengitt i kapittel 1 i delrapport 2, Seland et al 2016)

Forskernes kunnskapsinteresse	0. Norskfagets formål og oppgaver i opplæringen
Skoleeiernivå	
Skoleledernivå	
Lærere	Lærernes syn på norskfaget og idealene for undervisningen i norsk (lærernes syn på fagets funksjon og egen rolle), herunder syn på utviklingen i faget etter Kunnskapsløftet.
Elever	Elevenes syn på betydningen av norskfaget og syn på norskfaget som skolefag
Utdanningsdirektoratets kunnskapsinteresse	1. Hvordan arbeider lærerne ved forsøksskolene med kompetansemålene i læreplanen sammenlignet med dagens ordning med tre karakterer?
Skoleeiernivå	Skoleeiers målsetting og plan for forsøket i lys av Kunnskapsløftet Skoleeiers tilrettelegging av forsøket
Skoleledernivå	Skolens målsetting og plan for forsøket i lys av Kunnskapsløftet Skolens tilrettelegging for forsøket Omfanget, graden og den opplevde verdien av lokalt læreplanarbeid med operasjonalisering av kompetansemål gjennom utarbeidelse av læringsmål og vurderingskriterier. Hva ligger i begrepet «halvårsvurdering» utdypet gjennom forståelse av begrepene «underveisvurdering» og «sluttvurdering», inkludert skolens og/eller skoleeiers bestemmelser for skriftlig dokumentasjon av halvårsvurderingen

	Hvorvidt skolen tidligere har vært omfattet av satsingene Vurdering for læring, og/eller kompetansehevingstiltak i norsk og i vurdering (inkludert sensorskolering) og erfaringene med disse satsingene. Arbeidet for å utvikle en vurderingskultur ved skolen.
Lærere	Skolens plan for forsøket Skolens tilrettelegging for forsøket Omfanget, graden og den opplevde verdien av lokalt læreplanarbeid med operasjonalisering av kompetansenivå, utarbeidelse av læringsmål og vurderingskriterier. Hva ligger i begrepet «halvårsvurdering» utdypet gjennom forståelse av begrepene «underveisvurdering» og «sluttvurdering», inkludert skolens og/eller skoleeiers bestemmelser for skriftlig dokumentasjon av halvårsvurderingen Hvorvidt skolen tidligere har vært omfattet av satsingene Vurdering for læring, og/eller kompetansehevingstiltak i norsk og i vurdering (inkludert sensorskolering) og erfaringene med disse satsingene. Arbeidet for å utvikle en vurderingskultur ved skolen, herunder lærernes opplevelse av støtte fra overordnet, kultur for lærersamarbeid, støtte og tillit i kollegiet generelt.
Elever	
Utdanningsdirektoratets kunnskapsinteresse	2. Hvordan vektlegger lærerne ved forsøksskolene arbeidet med skriftlig sidemål, skriftlig hovedmål og muntlig norsk? 3. Fører færre karakterer ved halvårsvurdering i norsk til at lærerne opplever det som lettere eller som vanskeligere å organisere undervisning og vurdering i faget? 4. Fører færre karakterer i norsk til endret handlingsrom for lærerne i deres mulighet til å organisere undervisningen og vurderingen på en hensiktsmessig måte?

	5. I hvilken grad og på hvilken måte utnyttes et eventuelt økt handlingsrom?
Skoleeiernivå	Hvordan argumenterer skoleeier for at forsøket skal tilrettelegge for lærernes arbeid?
Skoleledernivå	Hvordan argumenterer skoleleder for at forsøket skal tilrettelegge for lærernes arbeid?
Lærere	Læreplanens betydning for undervisning og opplevelse av disponering av tid til emner i norskfaget, eventuelt mediert gjennom bruk av læreverk. Hvordan lærerne i forsøksordningene faktisk arbeider med undervisning i norskfaget sammenlignet med hvordan de oppgir å ha arbeidet før forsøket startet, (hvordan sikrer lærerne i forsøket at de gir undervisning knyttet til alle kompetansemålene?) Her må de mulige motsetningene muntlig norsk – skriftlig norsk og hovedmål – sidemål behandles spesielt. Hvordan lærerne i forsøksordningene faktisk arbeider med vurdering i norskfaget sammenlignet med hvordan de oppgir å ha arbeidet før forsøket startet, (hvordan sikrer lærerne i forsøket at de gir vurdering knyttet til alle kompetansemålene?) Her må de mulige motsetningene muntlig norsk – skriftlig norsk og hovedmål – sidemål behandles spesielt. Lærernes egne synspunkter og refleksjoner over hvorvidt det å redusere antall karakterer ved halvårsvurdering i norsk bøter på uttalte problemer med disponering av undervisning og/eller tid i og utenfor faget, og hvordan de begrunner dette. Her er det også viktig å få frem eventuelle ulemper eller problemer med redusert antall karakterer i faget.
Elever	Hvordan opplever elevene lærernes undervisning og vurdering i norskfaget?
Utdanningsdirektoratets kunnskapsinteresse	6. Fører færre karakterer ved halvårsvurdering i norskfaget til at elever og lærere opplever dette faget som mer eller mindre motiverende?
Skoleeiernivå	Hvordan argumenterer skoleeier for at forsøket eventuelt skaper økt motivasjon for faget?
Skoleledernivå	Hvordan argumenterer skoleleder for at forsøket eventuelt skaper økt motivasjon for faget?
Lærere	Fører færre karakterer ved halvårsvurdering i norskfaget til at lærere opplever dette faget som mer eller mindre motiverende?

Elever	Fører færre karakterer ved halvårsvurdering i norskfaget til at elever opplever dette faget som mer eller mindre motiverende?
Utdanningsdirektoratets kunnskapsinteresse	7. Hvordan opplever elevene overgangen fra færre karakterer til normalt karakterregime i Vg3?
Skoleeiernivå	
Skoleledernivå	Hvilke inntrykk har skoleleder av elevenes opplevelse av overgangen fra færre karakterer til normalt karakterregime i Vg3?
Lærere	Hvilke inntrykk har lærerne av elevenes opplevelse av overgangen fra færre karakterer til normalt karakterregime i Vg3?
Elever	Hvordan opplever elevene overgangen fra færre karakterer til normalt karakterregime i Vg3?
Forskernes kunnskapsinteresse	8. Hvordan tilrettelegger forsøket for elevenes læring?
Skoleeiernivå	Hvordan argumenterer skoleeier for at forsøket legger bedre til rette for elevenes læring i faget?
Skoleledernivå	Hvordan argumenterer skoleleder for at forsøket legger bedre til rette for elevenes læring i faget?
Lærernivå	Hvordan argumenterer lærerne for at forsøket legger bedre til rette for elevenes læring i faget?
Elevnivå	Hvordan opplever elevene at forsøket eventuelt legger bedre til rette for deres læring i faget?

Vedlegg 2: Regresjonstabeller for analyse av karakterer i kapittel 4

Tabell V1: Regresjonsmodell DID, med gjennomsnitt standpunktkarakter som utfall. Modell 1 til 5.2

	Modell 1	Modell 2	Modell 3	Modell 4.1 Gutter	Modell 4.2 Jenter	Modell 5.1 Én karakter	Modell 5.2 To karakterer
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	0.004	-0.025	-0.022	-0.03	-0.014	-0.029	-0.026
	-0.022	-0.017	-0.017	-0.02	-0.018	-0.024	-0.02
2007 (ref modell 1)	0						
	(.)						
2008 (ref modell 2-5)	-0.009	0	0	0	0	0	0
	-0.016	(.)	(.)	(.)	(.)	(.)	(.)
2009	-0.040**	-0.01	-0.011	-0.001	-0.017	-0.012	-0.01
	-0.017	-0.014	-0.014	-0.016	-0.014	-0.017	-0.018
2010	-0.01	0.021	0.021	0.029*	0.015	0.023	0.025
	-0.016	-0.014	-0.014	-0.017	-0.015	-0.016	-0.018
2011	-0.007	0.028**	0.028**	0.030*	0.026*	0.040**	0.016
	-0.018	-0.014	-0.014	-0.017	-0.015	-0.016	-0.018
2012	0.034*	0.055***	0.054***	0.053***	0.057***	0.071***	0.048***
	-0.018	-0.015	-0.015	-0.018	-0.015	-0.017	-0.018
2013	0.061***	0.090***	0.086***	0.074***	0.096***	0.101***	0.087***
	-0.022	-0.017	-0.017	-0.021	-0.018	-0.02	-0.021
2014	0.136***	0.163***	0.159***	0.154***	0.163***	0.164***	0.159***
	-0.024	-0.019	-0.019	-0.023	-0.02	-0.022	-0.022
Fritatt sidemål	-0.250***	0.167***	0.165***	0.115*	0.214***	0.174***	0.184***
	-0.067	-0.05	-0.049	-0.063	-0.057	-0.047	-0.057
Grskolep		0.111***	0.107***	0.109***	0.106***	0.108***	0.108***
		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
Idrett (ref)			0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)
MDD			-0.053*	-0.057	-0.066**	-0.070**	-0.053
			-0.027	-0.035	-0.028	-0.034	-0.037
Studiespes			0.067***	0.082***	0.043**	0.053**	0.073***
			-0.018	-0.02	-0.02	-0.024	-0.02
Gutt			-0.174***			-0.177***	-0.176***
			-0.005			-0.007	-0.007
_cons	3.915***	-1.106***	-0.930***	-1.172***	-0.852***	-0.921***	-0.953***
	-0.013	-0.03	-0.035	-0.039	-0.038	-0.049	-0.038
N	195767	170003	170003	76227	93776	94568	110511
R ²	0.004	0.488	0.499	0.466	0.474	0.498	0.497

Tabell V2: Regresjonsmodell DID, med gjennomsnitt standpunkt karakter som utfall. Modell 6.1 til 7

	Modell 6.1 1. Kvartil	Modell 6.2 2. Kvartil	Modell 6.3 3. Kvartil	Modell 6.4 4. Kvartil	Modell 7 Laveste 10 prosentil
	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	-0.013	0.018	-0.003	-0.004	-0.005
	-0.031	-0.023	-0.021	-0.019	-0.043
2008 (ref)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
2009	0	-0.002	-0.02	-0.017	-0.018
	-0.021	-0.02	-0.018	-0.016	-0.028
2010	0.035*	0.026	0.024	-0.009	0.026
	-0.02	-0.018	-0.018	-0.017	-0.026
2011	0.052**	0.051***	0.016	-0.014	0.038
	-0.022	-0.019	-0.018	-0.016	-0.029
2012	0.089***	0.082***	0.032*	0.01	0.087***
	-0.021	-0.02	-0.018	-0.016	-0.028
2013	0.127***	0.100***	0.066***	0.035*	0.126***
	-0.026	-0.023	-0.022	-0.021	-0.034
2014	0.227***	0.186***	0.127***	0.082***	0.212***
	-0.031	-0.027	-0.024	-0.022	-0.041
Fritatt sidemål	0.197***	0.180***	0.052	0.048	0.119*
	-0.07	-0.062	-0.075	-0.103	-0.067
Grskolepoeng	0.080***	0.107***	0.119***	0.121***	0.064***
	-0.001	-0.003	-0.003	-0.001	-0.003
Idrett (ref)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
MDD	0.034	-0.067**	-0.094***	-0.114***	0.072
	-0.051	-0.029	-0.032	-0.029	-0.061
Studiespes	0.111***	0.029	0.043**	0.037*	0.130***
	-0.022	-0.02	-0.021	-0.021	-0.027
Gutt	-0.220***	-0.185***	-0.151***	-0.104***	-0.220***
	-0.009	-0.009	-0.008	-0.009	-0.013
_cons	0.096	-0.916***	-1.484***	-1.582***	0.604***
	-0.058	-0.125	-0.148	-0.07	-0.094
N	41470	41918	41428	45187	19281
R ²	0.169	0.073	0.064	0.211	0.11

Tabell V3: Regresjonsmodell DID, med eksamenskarakter i hovedmål som utfall. Modell 1 til 5.2

	Modell 1	Modell 2	Modell 3	Modell 4.1 Gutter	Modell 4.2 Jenter	Modell 5.1 En karakter	Modell 5.2 To karakterer
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	0.035	0.007	0.009	0.021	0.002	-0.005	0.025
	-0.023	-0.017	-0.017	-0.022	-0.02	-0.023	-0.021
2007 (ref modell 1)	0						
	(.)						
2008 (ref modell 2-5)	0.045**	0	0	0	0	0	0
	-0.018	(.)	(.)	(.)	(.)	(.)	(.)
2009	0.054***	0.028*	0.027*	0.021	0.034**	0.027	0.017
	-0.018	-0.016	-0.016	-0.019	-0.017	-0.022	-0.018
2010	0.176***	0.148***	0.148***	0.158***	0.139***	0.121***	0.142***
	-0.017	-0.014	-0.014	-0.018	-0.016	-0.018	-0.019
2011	0.127***	0.100***	0.099***	0.110***	0.090***	0.099***	0.084***
	-0.019	-0.014	-0.014	-0.016	-0.016	-0.019	-0.017
2012	0.162***	0.123***	0.121***	0.091***	0.147***	0.098***	0.111***
	-0.02	-0.016	-0.016	-0.019	-0.018	-0.021	-0.02
2013	0.159***	0.134***	0.130***	0.087***	0.163***	0.112***	0.122***
	-0.026	-0.019	-0.02	-0.025	-0.022	-0.024	-0.023
2014	0.205***	0.177***	0.173***	0.134***	0.203***	0.167***	0.161***
	-0.026	-0.019	-0.019	-0.024	-0.022	-0.023	-0.021
Fritatt sidemål	-0.571***	-0.262***	-0.265***	-0.265***	-0.269***	-0.275***	-0.247***
	-0.066	-0.044	-0.045	-0.069	-0.063	-0.058	-0.041
Grskolep		0.096***	0.093***	0.094***	0.093***	0.094***	0.093***
		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
Idrett (ref)			0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)
MDD			0.037*	0.055*	0.040*	0.038	0.032
			-0.02	-0.029	-0.022	-0.025	-0.026
Studiespes			0.161***	0.148***	0.182***	0.149***	0.165***
			-0.013	-0.015	-0.016	-0.018	-0.016
Gutt			-0.107***			-0.104***	-0.109***
			-0.005			-0.007	-0.006
_cons	3.294***	-1.026***	-0.991***	-1.081***	-1.012***	-0.988***	-0.974***
	-0.014	-0.026	-0.03	-0.036	-0.035	-0.039	-0.035
N	191173	166927	166927	74720	92207	92516	108471
R ²	0.006	0.282	0.288	0.274	0.264	0.285	0.288

Tabell V4: Regresjonsmodell DID, med eksamenskarakter i hovedmål som utfall. Modell 6.1 til 7

	Modell 6.1 1. Kvartil	Modell 6.2 2. Kvartil	Modell 6.3 3. Kvartil	Modell 6.4 4. Kvartil	Modell 7 Laveste 10 prosentil
	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	0.03	0.039	0.043	0.036	0.028
	-0.032	-0.024	-0.028	-0.026	-0.046
2008 (ref)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
2009	0.139***	0.068***	-0.002	-0.085***	0.166***
	-0.023	-0.021	-0.021	-0.022	-0.03
2010	0.245***	0.183***	0.134***	0.033	0.260***
	-0.024	-0.019	-0.021	-0.02	-0.028
2011	0.163***	0.154***	0.087***	-0.001	0.191***
	-0.021	-0.019	-0.02	-0.019	-0.029
2012	0.198***	0.182***	0.104***	0.009	0.231***
	-0.023	-0.021	-0.023	-0.021	-0.031
2013	0.251***	0.198***	0.092***	-0.018	0.292***
	-0.027	-0.025	-0.03	-0.028	-0.039
2014	0.282***	0.231***	0.129***	0.046*	0.289***
	-0.029	-0.025	-0.028	-0.027	-0.041
Fritatt sidemål	-0.206***	-0.339***	-0.359***	-0.266*	-0.217***
	-0.056	-0.097	-0.093	-0.144	-0.069
Grskolepoeng	0.072***	0.093***	0.096***	0.110***	0.060***
	-0.002	-0.003	-0.004	-0.002	-0.004
Idrett (ref)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
MDD	0.03	0.009	0.044	0.033	0.069*
	-0.029	-0.029	-0.034	-0.03	-0.037
Studiespes	0.179***	0.131***	0.170***	0.159***	0.187***
	-0.015	-0.016	-0.021	-0.025	-0.019
Gutt	-0.134***	-0.093***	-0.088***	-0.068***	-0.148***
	-0.011	-0.01	-0.01	-0.011	-0.014
_cons	-0.276***	-1.034***	-1.127***	-1.707***	0.118
	-0.069	-0.152	-0.187	-0.109	-0.125
N	39753	41327	40988	44859	18047
R ²	0.097	0.033	0.026	0.088	0.07

Tabell V5: Regresjonsmodell DID, med standpunktkarakter i sidemål som utfall. Modell 1 til 5.2

	Modell 1	Modell 2	Modell 3	Modell 4.1 Gutter	Modell 4.2 Jenter	Modell 5.1 Én karakter	Modell 5.2 To karakterer
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	-0.003	-0.017	-0.014	-0.023	-0.007	-0.03	-0.023
	-0.021	-0.017	-0.017	-0.021	-0.018	-0.024	-0.021
2007 (ref modell 1)	0 (.)						
2008 (ref modell 2-5)	-0.01	0	0	0	0	0	0
	-0.018	(.)	(.)	(.)	(.)	(.)	(.)
2009	-0.044**	-0.009	-0.011	-0.01	-0.01	-0.021	0.009
	-0.019	-0.017	-0.017	-0.02	-0.019	-0.021	-0.022
2010	-0.009	0.019	0.019	0.018	0.019	0.013	0.027
	-0.018	-0.016	-0.016	-0.02	-0.017	-0.02	-0.021
2011	-0.005	0.026	0.025	0.016	0.033*	0.040**	0.018
	-0.018	-0.016	-0.016	-0.019	-0.018	-0.019	-0.022
2012	0.028	0.050***	0.049***	0.045**	0.052***	0.069***	0.046**
	-0.019	-0.017	-0.017	-0.021	-0.018	-0.02	-0.022
2013	0.050**	0.075***	0.071***	0.052**	0.087***	0.095***	0.082***
	-0.022	-0.019	-0.019	-0.024	-0.02	-0.022	-0.024
2014	0.129***	0.148***	0.143***	0.133***	0.151***	0.146***	0.155***
	-0.025	-0.021	-0.021	-0.025	-0.022	-0.023	-0.024
Grskolepoeng		0.114***	0.111***	0.110***	0.111***	0.111***	0.111***
		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
Idrett (ref)			0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)
MDD			-0.062**	-0.075**	-0.065**	-0.061*	-0.056*
			-0.026	-0.035	-0.027	-0.034	-0.031
Studiespes			0.064***	0.073***	0.051**	0.068***	0.073***
			-0.019	-0.021	-0.02	-0.026	-0.021
Gutt			-0.163***			-0.167***	-0.164***
			-0.005			-0.007	-0.007
_cons	3.631***	-	-1.419***	-1.563***	-1.420***	-1.437***	-1.435***
	-0.014	1.588***	-0.039	-0.044	-0.043	-0.056	-0.042
N	177740	155941	155941	68953	86988	86128	100934
R ²	0.003	0.417	0.425	0.398	0.398	0.426	0.424

Tabell V6: Regresjonsmodell DID, med standpunkt karakter i sidemål som utfall. Modell 6.1 til 7

	Modell 6.1 1. Kvartil	Modell 6.2 2. Kvartil	Modell 6.3 3. Kvartil	Modell 6.4 4. Kvartil	Modell 7 Laveste 10 prosentil
	b/se	b/se	b/se	b/se	b/se
Effekt av å være i forsøket	0.002	0.03	-0.002	-0.002	0.014
	-0.029	-0.025	-0.024	-0.023	-0.039
2008 (ref modell 2-5)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
2009	0.01	-0.025	-0.007	-0.027	0.032
	-0.024	-0.023	-0.022	-0.021	-0.032
2010	0.034	0.012	0.028	-0.007	0.063**
	-0.022	-0.02	-0.022	-0.021	-0.031
2011	0.051**	0.037*	0.024	-0.016	0.059*
	-0.022	-0.021	-0.021	-0.02	-0.031
2012	0.081***	0.065***	0.034	0.004	0.093***
	-0.022	-0.023	-0.022	-0.021	-0.029
2013	0.102***	0.084***	0.058**	0.022	0.099***
	-0.025	-0.024	-0.025	-0.025	-0.032
2014	0.213***	0.165***	0.127***	0.063**	0.225***
	-0.03	-0.027	-0.028	-0.025	-0.039
Grskolepoeng	0.078***	0.109***	0.125***	0.132***	0.063***
	-0.002	-0.003	-0.004	-0.002	-0.003
Idrett (ref)	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
MDD	-0.007	-0.062**	-0.075**	-0.111***	0.047
	-0.043	-0.027	-0.035	-0.035	-0.049
Studiespes	0.094***	0.038*	0.051**	0.046*	0.108***
	-0.023	-0.021	-0.024	-0.026	-0.024
Gutt	-0.190***	-0.172***	-0.151***	-0.114***	-0.177***
	-0.01	-0.01	-0.01	-0.01	-0.016
_cons	-0.185***	-1.391***	-2.143***	-2.450***	0.308***
	-0.064	-0.147	-0.17	-0.086	-0.108
N	34087	38466	39361	44027	14581
R ²	0.122	0.056	0.049	0.168	0.078

Tabell V7: Regresjonsmodell DID, med eksamenskarakter i sidemål som utfall. Modell 3 (full modell)

	Modell 3
	b/se
Effekt av å være i forsøket	-0,002
	-0,023
2008 (ref modell 2-5)	0
	(.)
2009	0.088***
	-0,02
2010	0.158***
	-0,02
2011	0.110***
	-0,021
2012	0.076***
	-0,02
2013	0.206***
	-0,024
2014	0.139***
	-0,028
Grskolepoeng	0.099***
	-0,001
Idrett (ref)	0
	(.)
MDD	0.075***
	-0,023
Studiespes	0.162***
	-0,015
Gutt	-0.108***
	-0,006
_cons	-1.442***
	-0,036
N	81594
R ²	0,325

Tabelloversikt

Tabell 1.1: Utvalg av skoler til forskningsintervjuer, fordeling med antall karakterer i forsøket og hovedmål for majoriteten av elevene	38
Tabell 3.1: Kilder norsklærere vektlegger i arbeid med standpunktkaraktersetting (bearbeidet utdrag fra NIFU rapport 16-2010).....	78
Tabell 4.1: Fordeling på tiltaksskoler og sammenligningsskoler, etter allmennfaglig/kombinert skole.....	89
Tabell 4.2: Fordeling på tiltaksskoler og sammenligningsskoler, etter offentlig/ privat skole	89
Tabell 4.3: Fordeling på tiltaksskoler og sammenligningsskoler, etter skolestørrelse.....	90
Tabell 5.1: Lederes kommentar om forsøket. N=10.	106
Tabell V1: Regresjonsmodell DID, med gjennomsnitt standpunktkarakter som utfall. Modell 1 til 5.2.....	165
Tabell V2: Regresjonsmodell DID, med gjennomsnitt standpunktkarakter som utfall. Modell 6.1 til 7.....	166
Tabell V3: Regresjonsmodell DID, med eksamenskarakter i hovedmål som utfall. Modell 1 til 5.2.....	167
Tabell V4: Regresjonsmodell DID, med eksamenskarakter i hovedmål som utfall. Modell 6.1 til 7.....	168
Tabell V5: Regresjonsmodell DID, med standpunktkarakter i sidemål som utfall. Modell 1 til 5.2.....	169
Tabell V6: Regresjonsmodell DID, med standpunktkarakter i sidemål som utfall. Modell 6.1 til 7.....	170
Tabell V7: Regresjonsmodell DID, med eksamenskarakter i sidemål som utfall. Modell 3 (full modell)	171

Figuroversikt

Figur 1.1: Evalueringens oppbygning illustrert ved emnenes innbyrdes forbindelser og grenseflater	27
Figur 1.2: Forskjell-i-forskjeller, grafisk forklaring (Columbia University, 2018).....	36
Figur 4.1: Grafisk fremstilling av forskjell-i-forskjeller analysene på gjennomsnittlig standpunkt-karakter i norsk, for tiltaksgruppe og sammenligningsgruppe.....	91
Figur 4.2: Grafisk fremstilling av forskjell-i-forskjeller analysene på gjennomsnittlig standpunkt-karakter i norsk, for tiltaksgruppe med en karakter, tiltaksgruppe med to karakterer og sammenligningsgruppe.....	91
Figur 4.3: Grafisk visualisering av resultatet av forskjell-i-forskjeller analysen, effekt av å være en del av forsøket på gjennomsnittlig standpunkt-karakter i norsk i Vg3. Estimat basert på modell 1 i vedleggstabell V1.....	94
Figur 5.1: Betydning av forsøket for organisering og gjennomføring av undervisning og vurdering. Gjennomsnittlig skåre.	100
Figur 5.2: Betydning av forsøket for organisering og gjennomføring av undervisning og vurdering. Fordeling i antall. N=35-36. Fra spørningene våren 2017.....	101
Figur 5.3: Elevenes utbytte og sidemålets stilling. Gjennomsnittlig skåre.	102
Figur 5.4: Elevenes utbytte og sidemålets stilling. Fordeling i antall. N=34-36. Fra spørningene våren 2017.....	105

Nordisk institutt for studier av
innovasjon, forskning og utdanning

Nordic institute for Studies in
Innovation, Research and Education

www.nifu.no