

Læringsfremmende vurderingspraksis – med blikk på norskfaget i to skoler

Harald Eriksen



Avhandling levert til vurdering for graden Philosophiae Doctor

Institutt for lærerutdanning og skoleforskning
Det utdanningsvitenskapelige fakultet

Universitetet i Oslo

2017

© Harald Eriksen, 2018

*Doktoravhandlingar forsvart ved
Det utdanningsvitenskapelige fakultet, Universitetet i Oslo.
Nr. 286*

ISSN 1501-8962

Det må ikke kopieres fra denne boka i strid med åndsverkloven eller med avtaler om kopiering inngått med Kopinor, interesseorgan for rettighetshavere til åndsverk.

Omslag: Hanne Baadsgaard Utigard.
Grafisk produksjon: Representralen, Universitetet i Oslo.

Takk

Takk til alle som har bidratt med ulike former for hjelp og støtte i arbeidet med avhandlingen. Jeg vil først og mest takke veilederne professor Eyvind Elstad og førsteamanuensis Gustaf Bernard Uno Skar ved NTNU for mange oppmuntringer og konstruktiv kritikk underveis i prosessen.

Takk til sluttleser professor Kari Smith ved NTNU som ga mange formative tilbakemeldinger og demonstrerte hvordan teori kan omsettes i praksis.

Takk til kollegaer som har bidratt med innspill, lesing av artikkelutkast, delkapitler, kurspaper, konferanseframlegg og ikke minst som diskusjonspartnere: Knut-Andreas Christophersen, Dijana Tiplić, Eli Lejonberg, Inga Staal Jensen, Kathrine Nesje, Nora Elise Hesby Mathé, Jonas Bakken, Vanja Røkkum Jansen, Ida Katrine Riksaasen Hatlevik, Drita Saliu Abdulahi, Kirsti Engelién, Marthe Blikstad Balas, Lisbeth Brevik, Tonje Stenseth, Frøydis Hertzberg, Anne-Kathrine Woldsnes, Hege Hermansen, Anniken Furberg og dessuten Line Ingulfsen og resten av medlemmene i TEPECs stipendiatforum.

For språkvask og korrekturlesing vil jeg særlig takke Henny Bjerke som har bidratt med solide innspill til både artikler og kappe og jeg takker Kirsten Halvorsen for hjelp med artikkel 1.

Takk til forskergruppen TEPEC som tok godt imot meg fra start, og som har kommet med mange gode innspill på artikkelutkast og mye mer. Takk også til forskergruppen SISCO som lot meg være assosiert medlem i mer enn to år, og hvor jeg fikk mange konstruktive innspill. Det gjelder også ansatte ved Skrivesenteret i Trondheim.

For kollegial støtte og godt samarbeid vil jeg jeg takke Eva Thue Vold, Ole Andreas Kvamme, Janne Thoralsvdatter Scheie, Karsten Korbøl, Kristin Helstad, Eline Wiese, Kari Anne Rødnes, Christian Brandmo, Glenn Ole Hellekjær, Jeff Hall, Jon Magne Vestøl og dessuten den kompetente og hyggelige administrasjonen ved ILS.

Institutt for lærerutdanning og skoleforskning har lagt forholdene godt til rette for skrivingen av

avhandlingen og har vært en god arbeidsplass de fire årene fra 2013 til 2017. Balansen mellom undervisning, forskning og dessuten støtte til utviklingsprosjektet med undervisningsvurdering (høsten 2015) har betydd mye.

Tidligere kollegaer fra Elvebakken videregående skole, inkludert ledelsen, og særlig norskseksjonen fortjener en takk for å ha stilt opp til pilot-intervjuer og for å ha gitt bort undervisningstid til testing av spørreskjemaundersøkelse. En spesiell takk går til Anne Lise Jomisko og Lars Greger Granlund som bidro i arbeidet med å reliabilitetsteste kodesystemet for tilbakemeldinger.

Støtte på hjemmebane er også viktig, og det er mange som har vært både oppmuntrende og som dessuten har holdt ut perioder med selvsentrert kandidat. Knut og Svanhild Eriksen og Elin Therese Eriksen har støttet meg hele veien, det samme har Pauline Munch Brudal, og Lisbeth og Paul Brudal som også har vært interesserte, lest og diskutert.

Spesiell takk til Sigurd og Vebjørn Brudal Eriksen som har måttet holde ut perioder med mye jobbing. Dere har gitt mye glede, og nå skal dere få litt tilbake. Fotballtur blir første stopp. Billetter er bestilt.

Oslo, november 2017

Sammendrag av avhandlingen *Læringsfremmende vurderingspraksis – med blikk på norskfaget i to skoler*

Målet med avhandlingen er å undersøke norsklæreres vurderings- og tilbakemeldingspraksis i et Vurdering for læring (VfL)-perspektiv (Black & Wiliam, 2009), etter som internasjonal forskning framhever økt læringsutbytte som en potensiell konsekvens av at lærere jobber med vurdering på læringsfremmende måter. I norsk skole er en slik praksis dessuten lovpålagt (Udir, 2010b). Det overordnede forskningsspørsmålet er: Hva kjennetegner norsklæreres skriftlige vurderings- og tilbakemeldingspraksis i to kollegier?

Jeg undersøkte fenomenet, lærernes vurderings- og tilbakemeldingspraksis, gjennom ulike metoder og fra ulike aktørperspektiver, og det metodiske designet kan kalles triangulering (Creswell, 2013) og er dessuten inspirert av Maxwells (2005) begrep interaktivt design. Metodene som er brukt, er dokumentanalyse av tilbakemeldinger og dokumenter med relevans for læreres vurderingspraksis, tematisk analyse av intervjudata og strukturell likningsmodellering i analyse av svar på en spørreskjemaundersøkelse som kartla elevens oppfatninger av lærers tilbakemeldingspraksis. Deltakere var seks lærere ved to videregående skoler og 12 norskklasser på studiespesialiserende og yrkesfaglige studieretninger. Data ble samlet inn gjennom et skoleår og besto av 1504 elevtekster med lærers tilbakemeldinger, intervjudata fra 6 intervjuer, og 213 elevs svar på en spørreskjemaundersøkelse.

Avhandlingens teoretiske fundament (Black & Wiliam, 2009) definerer en læringsfremmende vurderingspraksis ved at elevene kjenner målet med opplæringen, at de får tilbakemelding om hvordan de har prestert sammenliknet med målet og råd om hvordan de skal komme nærmere målet og at informasjon fra vurdering (i dette tilfellet tilbakemeldingen) blir brukt i videre læringsarbeid, og denne definisjonen legges også til grunn i avhandlingen.

Sentrale funn er at lærerne skrev mange tilbakemeldinger til egne elever som kunne kodes som læringsfremmende. Tilbakemeldingene var skrevet til en heldagsprøve, omtrent 9 uker før

eksamen. Analysene viser at omtrent halvparten av tilbakemeldingene kunne kodes som VfL, og at lærerne ved en skole hadde mer samstemt praksis enn den andre. Ved skolen med svakest elevgrunnlag og høy grad av lærerautonomi var lærernes praksis mest: En lærer skrev mange VfL-tilbakemeldinger og hadde organisert heldagsprøven som en prosesskriveøkt slik at elevene skulle bruke læreres tilbakemelding og forbedre egen tekst før levering – det hadde kun én elev gjort. Praksis til de andre lærerne var ulik: en skrev svært mange tilbakemeldinger til elevene, og det diskuteres i hvilken grad elevenes er i stand til å nyttiggjøre seg så mange tilbakemeldinger. Den tredje skrev svært få VfL-kommentarer og markerte språkfeil. I artikkel 2 er hovedfunn at det er en tydelig statistisk assosiasjon mellom elevs opplevelse av at lærer gjennom tilbakemeldinger forklarer hva eleven kan gjøre bedre, og opplevd nytte (avhengig variabel). At lærer legger til rette for bruk av tilbakemeldinger, er derimot ikke signifikant statistisk assosiert med opplevd nytte. Dette funnet er i strid med det en fra et teoretisk utgangspunkt kunne forvente, og funnet blir diskutert i avhandlingen. I spørsmålet om hva som kjennetegner vurderings- og tilbakemeldingspraksis i to kollegier er et hovedfunn at lærerne på skolen med sterkest samarbeidskultur var mer samstemte i sin skriftlige tilbakemeldingspraksis, og at de hadde flere muligheter til å legge til rette for bruk av tilbakemeldinger i videre læringsarbeid enn ved skolen med mer autonome lærere. Ved begge skoler oppfattet lærerne at elevene hadde vansker med å anvende vurderingskriterier i læringsøyemed, og ved skole A mente lærerne at elever hadde store problemer med å forstå kriterier.

Implikasjoner av funnene er at lærere kan ha mye å vinne på å gi tilbakemeldinger når det trengs mest, nemlig når elevene er i en prosess. Organiseringen av dette arbeidet bør ikke overlates til hver enkelt lærer, men støttes av ledelsen ved hver skole, og også mer langsiktig og overordnet på politisk nivå.

Innholdsfortegnelse

II	Takkebrev
IV	Sammendrag
VI	Innholdsfortegnelse
1	Innledning
2	1. Avklaring av teoretiske nivåer i avhandlingen
2	1.2 Kappens struktur
3	2 Avhandlingens teoretiske grunnlag
5	2.1. Bærekraftig vurdering og vurdering er læring – en internasjonal diskusjon om vurdering
6	2.2 Vurdering og forholdet mellom summativ og formativ vurdering
6	2.3 Hvordan kan en læringsfremmende vurderingspraksis defineres?
9	2.4 Vurdering for læring i Norge: den norske utdanningskonteksten
11	2.5 Oppsummering
12	3 Litteraturgjennomgang
12	3.1 Hva sier forskningen om VfL i Norge?
13	3.2 Norskfaget i videregående skole
14	3.2.1 Hvilke ideologiske og teoretiske perspektiver ligger til grunn for skriftlig norsk?
20	3.3 Tilbakemeldinger og tilbakemeldinger som kan bidra til læring
21	3.3.1 Forskning på tilbakemeldinger generelt
22	3.3.2 Tilbakemeldinger på skriving og hvordan de kan bidra til læring
23	3.3.3 Elevers- og studenters bruk av lærers tilbakemeldinger på skriving
26	3.4 Kriterier og bruk av kriterier
27	3.5 Kritikk av formativ vurdering
28	3.6 Lærerprofesjonalitet, lærersamarbeid og elevprestasjoner
30	3.7 Oppsummering og forskningsspørsmål
32	4 Forskningsdesign og metoder
32	4.1 Deltakere og datagrunnlag i avhandlingen
34	4.2 Oversikt over analysemetoder i artiklene
35	4.3 Dokumentanalyse
37	4.4 Strukturell likningsmodellering
40	4.4.1 Eksplorerende faktoranalyse
41	4.5 Intervju
44	4.5.1 Intervju i kombinasjon med andre data
45	4.6 Validitet og reliabilitet

48	4.7 Forskningsetikk
48	4.7.1 Ivaretagelse av informantene – deltakelse og læreren som «den andre»
50	4.7.2 Sinnelagsetikk og konsekvenser av forskningen
50	4.7.3 Anonymitet
51	4.7.4 Bruk av studien – og språklig framstilling
53	5 Sammendrag av artikler (Oversikt over PhD-avhandlingens tre artikler)
54	5.1 Sammendrag av artikkel 1
55	5.2 Sammendrag av artikkel 2
57	5.3 Sammendrag av artikkel 3
59	6 Diskusjon og forskningsbidrag
64	Forskningsbidrag
64	Metodisk
66	Empirisk
67	Teoretisk
68	Konklusjoner – svar på forskningsspørsmålet
69	Videre forskning
70	Implikasjoner for praksis
71	Referanser
78	Vedlegg 1 – Modell
79	Vedlegg 2 – Totaloversikt over data
91	Vedlegg 3 – Eksempel på tematisk analyse
94	Del 2
95	Artikkel 1. Eriksen, Harald (2017). Vurdering for læring i norskfaget: Om lærerens skriftlige tilbakemeldinger på skriftlige tekster. <i>Acta Didactica Norge</i> , Vol. 11, Nr. 1. DOI: http://dx.doi.org/10.5617/adno.4078
116	Artikkel 2. Eriksen, Harald (2017). Elevers oppfatning av lærers tilbakemeldingspraksis: Om sammenheng mellom graden av VfL-praksis og elevenes opplevde nytte av lærers skriftlige tilbakemeldinger i skriftlig norsk. <i>Nordic Journal of Literacy Research</i> , Vol. 3, Nr.2. DOI: http://dx.doi.org/10.23865/njlr.v3.683
141	Artikkel 3. Eriksen, Harald: Norsklæreres erfaringer med læringsfremmende vurdering av elevtekster i to kollegier. (i revisjonsprosess etter krav om “major revisions” fra tidsskrift).

1 Innledning

Formålet med denne studien er å undersøke norsklæreres vurderings- og tilbakemeldingspraksis i skriveopplæringen i et Vurdering for læring (VfL)-perspektiv. Fenomenet som er undersøkt, er norsklæreres skriftlige tilbakemeldinger til egne elever, lærernes oppfatninger og refleksjoner over egen tilbakemeldingspraksis og elevers oppfatninger av denne praksisen. Ettersom praksisen er situert i skoler, der lærere jobber sammen i kollegier, er også lærerprofesjonalitet, og mer spesifikt lærersamarbeid undersøkt. For å utforske fenomenet, ble data samlet inn ved to ulike skoler, der seks lærere underviste i tolv norskklasser. Datagrunnlaget omfattet også elevers oppfatninger og ble analysert med ulike metoder som samlet sett gir et mangefasettert bilde av lærernes praksis.

Flere forhold har motivert valget av fenomen, teoribakgrunn og analysemåter. Først og fremst er den langvarige satsingen på en mer læringsfremmende vurderingspraksis både internasjonalt og i Norge en viktig årsak til at jeg har ønsket å utforske feltet: Forskningen (bl.a. Black & Wiliam, 1998) viser at gevinstene kan være betydelige i form av økt læringsutbytte hvis lærere kan bidra til å realisere en mer læringsfremmende vurderingspraksis, og tilbakemeldinger er en nøkkelkomponent når man i utdanningskontekster jobber med vurdering på læringsfremmende måter (Sadler, 1998).

For det andre framhever toneangivende forskere som Black og Wiliam (2009, s. 7) at lærere og elever må ses som partnere i læringsprosessen, og det er en viktig motivasjon for å inkludere både lærer- og elevperspektiver i studien. Elever som er involvert i eget vurderings- og læringsarbeid, slik forskningen på VfL viser (Sadler, 2010), er dessuten bedre rustet for livslang læring, noe ethvert samfunn vil ha nytte av (Hayward, 2015).

For det tredje viser forskningen på VfL at potensialet for økt læring er betydelig hvis man har systemer som sikrer at informasjon fra vurderingsarbeidet blir brukt i videre læringsarbeid (Wiliam, Lee, Harrison & Black, 2004). For å kunne iverksette tiltak som kan bidra til å virkeliggjøre en læringsfremmende vurderingspraksis, trenger man derfor kunnskap om hvordan lærere jobber sammen i kollegier, fordi forskning viser at læreres faglige samarbeid kan være relatert til økte elevprestasjoner (Goddard, Goddard & Tschannen-Moran, 2007). I en slik kontekst spiller også skoleledelsens tilrettelegging for samarbeid en viktig rolle.

Siden forskningen som ligger til grunn for retningslinjer for læringsfremmende vurderingspraksis (f.eks. Black & Wiliam, 1998) er generelt formulert, trengs kunnskap om hvordan satsingen manifesterer seg i ulike fag (Bennett, 2011). Norskfaget i videregående skole er valgt fordi det er sentralt i opplæringen, og det faktum at mange norsklærere bruker mye tid på å lese elevtekster og skrive tilbakemeldinger er også en motivasjonsfaktor å utforske nettopp dette faget i et VfL-perspektiv. På et overordnet nivå blir derfor forskningsspørsmålet: Hva kjennetegner norsklæreres skriftlige vurderings- og tilbakemeldingspraksis i to kollegier?

1.0 Avklaring av teoretiske nivåer i avhandlingen

I avhandlingen ses læringsteorier og hva som kjennetegner læringsfremmende vurdering i lys av en «nested model». Det vil si at sammenhengen mellom læring og vurdering betraktes som fenomener som er gjensidig avhengige av hverandre, og dette synet speiler framstillingen i avhandlingen som helhet. For å presisere dette poenget ser jeg læring på et overordnet nivå og vurdering ses som en del av læringsprosesser. For at vurdering skal ha større sjanse for å bidra til læring må den ha spesielle kjennetegn og i en praktisk virkelighet skal dessuten VfL realiseres av lærere i møte med elever.

1.2 Kappens struktur

Nedenfor følger en redegjørelse for VfLs teoretiske fundament, som speiler nivåene i en «nested model»: Jeg gjør først greie for læringsteorier, så knytter jeg vurdering til læring og gjør greie for hva som kjennetegner en læringsfremmende vurderingspraksis. Siden vurdering i norskfagets skriveopplæring er fenomenet som utforskes, gjør jeg greie for kontekstuelle forhold som innvirker på norskfaget i skolen og hva som kjennetegner skriving i faget. Videre knytter jeg VfL til vurdering av skriftlige arbeider og her spiller tilbakemeldinger en avgjørende rolle. Jeg gjør rede for forskning på tilbakemeldinger på generelt nivå før jeg knytter an til tilbakemeldinger på skriving og til slutt til forskning om lærerprofesjonalitet og lærersamarbeid. Litteraturgjennomgangen fører fram mot en begrunnet presentasjon av forskningsspørsmålet. Deretter kommer et eget kapittel om metodene som er brukt, inkludert forskningsetikk og diskusjon av validitets- og reliabilitetsspørsmål som også tar opp i seg avhandlingens begrensninger. Sammendrag av de tre artiklene leder til en diskusjon av hovedfunn og avhandlingens forskningsbidrag, implikasjoner for praksis og konklusjoner.

2 Avhandlingens teoretiske grunnlag

Et premiss som ligger til grunn for denne avhandlingen, er at vurdering har et potensial for å bidra til læring hvis den organiseres på bestemte måter. Dermed er det også nødvendig å avklare hvordan begrepet læring kan forstås. Säljö (2009) viser hvordan de tre hovedtradisjonene i læringsteorien – behaviorismen, kognitiv teori og teorier som ser læring som situert i en sosial sammenheng – bygger på ulike analytiske enheter og at de derfor kan være svært krevende å forene. Følgelig vil svaret på spørsmålet om hvordan læring kan forstås i denne avhandlingen avhenge av hvilke analytiske enheter det empiriske grunnlaget er basert på, og hvilken epistemologisk status jeg tillegger disse enhetene. Å skulle observere læring direkte er nærmest umulig, derfor trenger vi teorier om hva læring er og hvordan den kan måles (Säljö, 2009).

Avhandlingens empiriske grunnlag er læreres handlinger, deres refleksjoner over egne handlinger, og elevenes oppfatninger av lærernes handlinger. Det er altså flere analytiske enheter som samlet sett utgjør det empiriske datagrunnlaget. Perspektiver jeg bruker i analysene av empirien, bygger på forskning om hvordan man kan jobbe med vurdering så sjansen for at den kan bidra til læring øker. Forskere i feltet VfL (Black & Wiliam, 1998) har analysert en rekke studier ut fra dette perspektivet, de har syntetisert funnene og utledet noen kjennetegn på vurderingspraksis som har større sjanse for å bidra til læring. Disse kjennetegnene kan ses som en type verktøy lærerne kan bruke, og som predikerer økt sannsynlighet for læring hvis de anvendes. Slik sett kan spørsmålet om hva som er en læringsfremmende vurderingspraksis, tillegges teoristatus. I det følgende vil jeg knytte lærers tilbakemeldingspraksis til de tre hovedteoriene om læring for å synliggjøre hvordan en læringsfremmende vurderingspraksis kan forbindes med disse.

I behaviorismen kan lærers tilbakemelding ses som en form for stimulus som har til hensikt å fremkalle en bestemt respons (Säljö, 2009). Hvis stimulus er utformet på bestemte måter, etter anbefalingene fra VfL-forskningen (Black & Wiliam, 2009), så sier denne forskningen at det er større sjanse for at læring finner sted. Læring blir da forstått som endring i observerbar atferd, for eksempel at elever skriver bedre gjennom at de mestrer delferdigheter de tidligere ikke har mestret, og lærers tilbakemelding kan ses som et redskap til forsterking av ønsket eller reduksjon av uønsket atferd (feil tegnsetting, f.eks.). Avhandlingens empiriske grunnlag gjør det imidlertid krevende å benytte det behavioristiske perspektivet fordi det metodiske designet ikke er eksperimentinnrettet.

I kognitiv læringsteori kan en tilbakemelding ses som et resultat av at en lærer har anvendt kognitive ressurser til å vurdere en elevtekst og formulere en respons på en bestemt måte. Hvis den er utformet for å fremme læring, er den forventet å sette i gang en kognitiv prosess hos eleven slik at denne får økt grad av forståelse for hva som kan kjennetegne en god tekst. Jeg legger, med utgangspunkt i Anderson (1982), til grunn at tilbakemeldinger i høy grad blir gitt av lærerne for å bidra til elevenes utvikling som skrivere, og at et viktig siktemål for lærerens tilbakemeldings- og vurderingspraksis er å påvirke elevens læringsarbeid (Brophy, 1983). Anderson (1982) ser læring som en endring i individets mentale strukturer i retning av mer kunnskap og høyere grad av mestring av ferdigheter. Ferdighetstilegnelsen er prosessen der en ferdighet går fra et deklarativt nivå til et proseduralt – det vil si at den lærende gjennom underprosesser som generalisering, diskriminering og evne til å slå sammen delprosesser får kunnskap i å anvende kognitive ferdigheter i møte med en oppgave. Anderson (1982, s. 370) framhever at verbal mediering er en viktig del av den innledende fasen av læringsprosessen. Dette poenget synliggjør at lærers tilbakemeldingspraksis kan ses som en måte å gi eleven informasjon, mediert skriftlig eller muntlig, om hvordan hun/han kan mestre en skriveoppgave bedre. Eleven må i sin tur så prosessere denne informasjonen og omsette den i endret og forhåpentligvis forbedret evne til å skrive en tekst. At dette er en komplisert oppgave, der lærerens støtte er viktig, kan illustreres av at Anderson (1982, s. 369) framhever at det vanligvis vil ligge minst 100 timers øving bak tilegnelsen av en kognitiv ferdighet.

De kognitive prosessene vil alltid foregå i en kontekst og i samspill med andre, blant annet ved at den som skriver, konstituerer seg som en modellskriver i dialog med en modelleser (Lotmann, 1990). Et perspektiv som dermed kan utvide den teoretiske forståelsen av fenomenet, er å legge vekt på at tilbakemeldingen har et dialogisk og derfor sosialt preg. I sosiokulturell teori kan en vektlegge at tilbakemeldingen skrives av lærer i rollen som «den signifikante andre» (Vygotskij, 1978) som søker å få eleven til å mestre på et høyere nivå enn hun/han ville fått til på egen hånd, ved at tilbakemeldingen har potensial til å bevege eleven inn i «den proksimale sonen». Samhandlingen mellom lærer og elev er det sentrale her. Avhandlingen har imidlertid ikke undersøkt denne samhandlingen direkte, og det legges derfor mindre vekt på sosiokulturell læringsteori.

Siden VFL er et relativt nytt konsept, kan det legges til grunn at også lærere er eller har vært i en læringssituasjon i arbeidet med å utvikle en mer læringsfremmende vurderingspraksis, og gjennom samarbeid og samhandling kan de hjelpe hverandre i det Lave og Wenger (1991) kaller «Communities of practice». Et slikt fellesskap kjennetegnes av at deltakerne har felles mål,

engasjement i et sosialt fellesskap og et delt repertoar av ressurser for å nå målene. Et kollegium, slik det er undersøkt i denne avhandlingen, kan ses som eksempel på et slikt fellesskap. Lærernes tilbakemeldinger kan dessuten bidra til at elever går fra å være «noviser» til «legitime deltakere» (Lave & Wenger, 1991) i skolens skrivepraksis.

Kompleksiteten i menneskelig kognisjon og samhandling trer ytterligere fram om en også inkluderer motivasjonsmessige og mer personlighetsorienterte perspektiver; det er meningen at tilbakemeldingen skal påvirke eleven til å bli en bedre skriver. For at dette skal skje, må eleven ønske å ta til seg tilbakemeldingen, ha motivasjon (Pintrich, 2003) og selvdisciplin (Duckworth & Seligman, 2005). I tillegg må det være strategisk fornuftig (Elstad, 2006) å gjøre arbeidet med å f.eks. forbedre en tekst. Elevens mestringsforventninger (Bandura, 1977; 1997) kan også antas å påvirke en mulig effekt av tilbakemeldingen og evnen til å omsette innholdet i en tilbakemelding til bedret skriveatferd. Forhold som kan påvirke mestringsforventninger, er blant annet verbal påvirkning («persuasion»), som kan ha ulike former. Teorien til Bandura (1977) viser derfor at lærers tilbakemelding, sett som en verbal påvirkning, kan predikere økt innsats og utholdenhet hos elever i en skriveprosess, og dermed også øke muligheten for at eleven lykkes med å utvikle sine skriveferdigheter. På dette grunnlaget er personrettet tilbakemelding inkludert i analysene i denne studien.

Gjennomgangen av lærings- og motivasjonsorienterte perspektiver på vurdering peker mot en svært sammensatt praksis, og jeg velger å inkludere flere perspektiver for å fastholde kompleksitet og for å søke å gi et nyansert bilde av fenomenet. Jeg legger en hovedvekt på kognisjonsteoretiske perspektiver, og inkluderer dessuten synet på skolen som et «community of practice». Disse valgene er begrunnet av avhandlingens empiri og av den teoretiske bakgrunnen. Ettersom et grunnleggende premiss for avhandlingen er at vurdering har potensial til å bidra til læring, følger det nå en kort presentasjon av den internasjonale diskusjonen om vurdering, en definisjon av vurdering, en teoretisk avklaring av skillet mellom formativ og summativ vurdering, etterfulgt av en forskningsbasert definisjon av formativ vurdering.

2.1. Bærekraftig vurdering og vurdering *er* læring – en internasjonal diskusjon om vurdering

I et internasjonalt perspektiv er det verdt å merke seg en uttalt motsetning mellom forventningen om at lærerne skal bruke resultater fra vurdering til å fremme læring, basert på sitt faglige skjønn,

og «accountability»-perspektiver som står sentralt i Storbritannia og USA (Hayward, 2015). Med begrepet «sustainable assessment» tar Boud (2000) til orde for å styrke læringsaspektet, blant annet gjennom å involvere egen- og hverandrevurdering (Boud & Falchikov, 2006; Boud, 2000). En litt annen, men beslektet innfallsvinkel finnes hos Hayward (2015), som argumenterer for at vurdering *er* læring ved å fokusere på at alle nivåer i utdanningssystemet bør/må bruke resultater fra vurdering, enten det er i klasserommet, eller man ser på langsiktig utvikling i eksamensresultater på nasjonalt nivå til å utvikle og forbedre hele utdanningssystemet. I Norge i 2017 er det imidlertid begrepet VfL som dominerer og det er derfor vektlagt i denne studien.

2.2 Vurdering og forholdet mellom summativ og formativ vurdering

Vurdering er i denne avhandlingen forstått med utgangspunkt i Taras (2010). Hun sier:

Assessment is a judgement: parameters i.e. standards, criteria, context etc are required to make this judgement. These parameters may be implicit i.e. in the assessor's head, or they may be explicit and shared with others (2010, s. 3020).

Læreres vurdering av i hvilken grad en elevprestasjon møter en faglig standard, kaller Taras (2010) en summativ vurdering. Spørsmålet videre er om vurderingen skal munne ut i en kvantifisering, som karakter på et sluttprodukt, eller om den skal bidra til læring. Ofte vil lærer, i hvert fall i en kontekst der elevene gis karakter, søke å oppfylle begge hensikter. Taras advarer mot å se hensiktene som dikotomier. Dette synspunktet deler hun med Newton (2007), som også ser summativ og formativ vurdering som et spørsmål om hva informasjon om en prestasjon skal brukes til, og at overgangen mellom dimensjonene må ses som et kontinuum. Slik er de også forstått i denne avhandlingen.

Det er flere studier som advarer mot at den summative dimensjonen i vurdering lett kan dominere (Butler, 1987; 1988). Krogh og Jensen (2008, s. 66) skriver om mappevurdering: «Når elever får karakter og kommentar samtidig, er det svært vanskelig å få dem til å bruke den formativt tenkte veiledningen som ligger i lærerkommentaren.» Dette er altså en problematikk som lærerne i studien må forholde seg til, og den generelle delen av læreplanen, lansert høsten 2017 (Udir, 2017c), advarer mot bruk av summativ vurdering som ikke følges opp – det framholdes at slik bruk av måling og karaktersetning kan hemme læring.

2.3 Hvordan kan en læringsfremmende vurderingspraksis defineres?

Tanken at vurdering skal bidra til læring, har lange røtter. Connors og Lunsford (1993) trekker linjer tilbake til 1880-årene og enda videre bakover i historien når de hevder at lærere nærmest

til alle tider har gitt tilbakemeldinger på elevers og studenters arbeider i den hensikt å bidra til utvikling og læring. I en slik læringsfremmende vurdering spiller tilbakemelding en avgjørende rolle, og Bloom (1968) var en viktig premissleverandør i utviklingen av forståelsen for viktigheten av tilbakemeldinger i læringsarbeid. Benjamin Bloom og hans medforfattere tok til orde for en mer individualisert form for undervisning innenfor rammene av masseutdanningssystemet, der læreren kan identifisere behov hos, og gi tilbakemelding til, hver enkelt elev. At en vurderingspraksis skal bidra til læring, henger tett sammen med begrepet formativ vurdering, som stammer fra Scriven (1967). Begrepet er brukt om vurdering som har til hensikt å skape endring, f.eks. i en læringsprosess. Ramaprasad (1983) understreket poenget med læringsaspektet da han med utgangspunkt i organisasjonsteori søkte å komme fram til en definisjon av tilbakemelding (feedback), på grunn av en på det tidspunktet, manglende felles forståelse:

Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way (Ramaprasad 1983, s. 4).

Kjernepunktet hos Ramaprasad (1983) ligger i vektleggingen av at tilbakemeldingen må brukes om den skal kunne kalles formativ. Denne definisjonen er ofte sitert, og er dessuten videreutviklet av mange bidragsyttere (Black & Wiliam, 2009). Definisjonen mangler imidlertid en kobling til en utdanningskontekst. Sadler (1989) bygger videre på Ramaprasad (1983) når han sier at

The information about the gap between actual and reference level is considered as feedback only when it is used to alter the gap. If the information is simply recorded, passed to a third part who lacks either the knowledge or the power to change the outcome, or is too deeply coded (for example, as a summary grade given by the teacher) to lead to appropriate action, the control loop cannot be closed (Sadler, 1989, s. 121).

Denne definisjonen, som også er mye sitert, sier lite om *hva* f.eks. læreren skal gjøre. Dette er derimot utforsket i en omfattende metastudie av Black og Wiliam (1998), som rapporterer om effekten av «formative assessment» fra 250 studier mellom 1988 og 1998. Forfatterne fant stor effekt av intervensjoner der formativ vurdering ble brukt for å fremme læring. De ga råd om hvordan dette feltet kunne utvikles videre, både praktisk og forskningsmessig, men et problem var en mangetydig bruk av «formativ vurdering» i ulike fagfelt. Assessment Reform Group (2002) lanserte derfor begrepet «Assessment for learning» (AfL) som erstatning for formativ vurdering. Dette begrepet er tatt opp av norske utdanningsmyndigheter og er oversatt til vurdering for læring. Wiliam (2011) sier om denne definisjonen i en reviewartikkel som oppsummerer forskningen innen «formative assessment»-feltet at vurdering for læring er målet, og formativ vurdering er et middel til å nå dette målet.

Som sentrale aktører i AfL-feltet søkte Black og Wiliam i 2009 å definere begrepet formativ vurdering på bakgrunn av tidligere definisjoner, og de knyttet det direkte an til en skole- og opplæringskontekst. Dette perspektivet er mest relevant for denne avhandlingen. De skrev:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (2009, s. 9).

I sitatet er formativ vurdering situert i en klasseromsituasjon der læreren gir elevene oppgaver og leder ulike prosesser som skal bidra til læring hos elevene. Denne måten å organisere undervisningen framheves av Shulman (2005) som en type internasjonalt kanonisert «signature pedagogy», der elever eller studenter utforsker et læringsmål introdusert av en lærer, etterfulgt av elevarbeid, individuelt eller i grupper, og der timen konsolideres med en lærerstyrt helklassesamtale. Denne arbeidsmåten skiller seg på flere måter fra den situasjonen en norsklærer står overfor når elevene skal gis opplæring i skriving og få tilbakemeldinger på skriftlige arbeider. Forskning på «classroom assessment» er likevel inkludert i framstillingen siden det ikke går an å trekke skarpe skiller mellom en «classroom assessment»-praksis og den praksisen en lærer utøver når hun skal gi en skriftlig tilbakemelding til en elev på et skriftlig arbeid i den hensikt å hjelpe eleven til å utvikle en skriveferdighet.

Black og Wiliam (2009) framholder at formativ vurdering er en praksis som befinner seg i en svært kompleks kontekst og med lærer, elever og medelevene (peers) som aktører som har og kan ha ulike roller, men det er likevel noen fellesnevnerer for hva som kjennetegner en formativ vurderingspraksis. Det er i Black og Williams (2009, s. 8) teoretiske system at lærer og elever deler mål eller kriterier for å finne fram til hva som er målet, at læreren finner informasjon om hva elevene vet og kan når læringsøkten begynner og sørger for tilbakemeldinger som gjør at elevene kommer nærmere målet. Dessuten er det viktig at elevene er involvert i eget vurderings- og læringsarbeid. Ettersom disse trekkene ved en læringsfremmende vurderingspraksis er i tråd med mye av den mest sentrale forskningen på formativ vurdering eller VfL, danner de også et viktig grunnlag for min forståelse av hva VfL er i denne avhandlingen.¹

Hovedpunktene i forskningen (Black & Wiliam, 1998; Black & Wiliam, 2009; Sadler 1989) ligger tett på retningslinjene til Utdanningsdirektoratet (2010b), nemlig 1) at elevene kjenner målet med

¹ Begrepene formativ vurdering, læringsfremmende vurdering og vurdering for læring er overlappede og brukes til dels synonymt i avhandlingen av hensyn til språklig variasjon.

opplæringen, 2) at de får informasjon om hvordan de har prestert i forhold til målet, 3) at de får informasjon (tilbakemeldinger) som bringer dem nærmere målet og 4) at de er involvert i eget læringsarbeid ved å vurdere eget arbeid og utvikling (Hopfenbeck, Tolo, Florez & El Masri, 2013; Udir, 2010b). Disse generelt formulerte retningslinjene er operasjonalisert på ulike måter i denne avhandlingen.

I det videre følger en presentasjon av den norske utdanningskonteksten, der en læringsfremmende praksis skal realiseres.

2.4 Vurdering for læring i Norge: den norske utdanningskonteksten

Utdanningssektoren i Norge har en hierarkisk struktur i sin oppbygning og påvirkes av mange faktorer. På et overordnet nivå kan en framheve at transnasjonale trender når norsk offentlighet og bidrar til å påvirke norsk skole. Eksempler på slike påvirkningsfaktorer kan internasjonale storskalaundersøkelser, som TIMSS, PIRLS og PISA sies å representere for å nevne tre som har fått stor medieoppmerksomhet. Sammenlikningene er ment å gi det enkelte lands utdanningsmyndigheter et empirisk grunnlag for benchmarking. En annen slik trend er satsingen på en læringsfremmende vurderingspraksis, som blant annet OECD har gitt råd om hvordan kan videreutvikles i Norge (Hopfenbeck et al., 2013).

På et nasjonalt nivå har denne påvirkningen innen utdanningssektoren manifestert seg på flere plan, og det er verdt å merke seg lovendringer, som befinner seg på øverste politiske nivå, i regjeringen og på Stortinget. Ifølge institusjonell teori er det grunn til å regne med at slike endringer, etter en viss tid, vil manifestere seg i praksis (Kennedy, 1997). At denne veien fra lov- og politikkendring til endringer i den enkelte lærers praksis kan være lang, illustreres av Hayward, Priestely og Young (2004) som viser hvordan policyendringer innen vurderingsfeltet i Skottland bidro til endringer av praksis i skolene etter to til tre år, men at det også var utfordringer med implementeringen.

På nivået under det nasjonale er kommuner og fylkeskommuner (for videregående skole) ansvarlige for å følge opp skolene i rollen som skoleeier. Skolene har en tredelt hierarkisk struktur – skolens ledelse, lærere og elever (Ingersoll & Kralik, 2004). I norske skoler vil rektor stå ansvarlig overfor skoleeier, og mellomledere/avdelingsledere være i spennet mellom to hierarkiske nivåer, dels kan de være kollegaer med andre lærere, dels kan de ha og ikke ha personalansvar delegert av skolens ledelse. Skolene kan videre ses som institusjoner som er underlagt nasjonale føringer, skoleeiers føringer og det er skolens rektor og ledelsens oppgave å sørge for å ivareta og implementere disse

i praksis. Det er i samhandlingen mellom lærer og elev i hvert klasserom, i dialogen om elevenes skriving, lærers vurdering av denne og tilbakemeldingspraksisen at føringene skal manifestere seg i handlinger. Som aktører i denne konteksten skal dessuten lærer og elev agere i tråd med sine roller, og det er klare forventninger til begge (Sugai & Horner, 2002).

De nasjonale retningslinjene er generelt formulert, og det innebærer at det enkelte kollegium og videre den enkelte lærer må omsette dem til praktisk undervisnings- og vurderingspraksis i sitt fag. Teori om lærersamarbeid og lærerprofesjonalitet er anvendt i denne studien fordi forskning finner at grad av lærersamarbeid korrelerer med elevprestasjoner (Goddard, Hoy & Hoy, 2000; 2004; Goddard, Goddard & Tschannen-Moran, 2007).

Siden begynnelsen av 2000-tallet har det i Norge vært rettet oppmerksomhet og ressurser mot å utvikle en læringsfremmende vurderingspraksis, blant annet fikk Black og Williams' studie fra 1998 stor oppmerksomhet. I Stortingsmelding nr. 30 (2003–2004) framheves det at vurderingskulturen i Norge har forbedringspotensial, og at arbeid med å utvikle vurderingskriterier i alle fag skulle være en prioritert oppgave. Mortimore, Field og Pont (2006) framhevet i en OECD-rapport om likeverd i utdanningen (*Equity in Education*) Norges forbedringsmuligheter når det gjaldt oppfølging og vurdering av elevers læringsarbeid. Stortingsmelding nr. 16 (2006–2007) legger også vekt på at arbeid med vurdering må styrkes for å øke elevenes læringsmuligheter, og dette følges opp i Stortingsmelding nr. 31 (2007–2008). Satsingen *Bedre vurderingspraksis 2006–2009* (Udir, 2009) kan ses som et viktig startpunkt for det konkrete arbeidet med å dreie vurderingspraksis mer mot elevers læring. Formålet var å utvikle lokale vurderingskriterier for alle fag. Kriteriene skulle både bidra til at målene for opplæringen ble klarere for lærere og elever, og til at grunnlaget for karakterer kom tydeligere fram enn før. En endring i opplæringsloven (Lovdata, 2006/2009) slo i forlengelsen fast at målet med opplæringen skal være kjent for elever og lærlinger, og at vurderingen gjennom året skal bidra til læring. *Bedre vurderingspraksis* ble fulgt opp av *Vurdering for læring 2010–2014*, der 184 kommuner deltok. Hensikten med dette prosjektet var å utvikle en vurderingspraksis som fremmer læring, og siden 2014 har satsingen fortsatt over hele landet (Udir, 2014a; 2016).

Statusen til VfL i Norge er spesiell ved at føringene er nedfelt i lovverket (Hopfenbeck et al., 2013), og det har vært slik i over ti år når denne avhandlingen skrives. Når satsingen er så omfattende, i form av investerte ressurser over lang tid, er det rimelig å forvente at den kommer til uttrykk i læreres praksis på skolenivå og i deres mer autonome vurderingspraksis (Hayward, et al. 2004). Lovfestingen bidrar til å styrke denne antakelsen.

Det er først og fremst lærerne som skal realisere en mer læringsfremmende vurderingspraksis i praktisk virkelighet. Dermed kan en slik overordnet satsing ses i lys av en økende grad av sentralisering og dels målstyring av norsk skole. Stortingsmelding nr. 11 (2008–2009) gir lærerne et klart definert ansvar for elevenes læring, og sammenholdt med signaler fra KS (Lejonberg & Elstad, 2014) kan retningslinjene forstås som uttrykk for en internasjonal trend (Ingersoll & Kralik, 2004) henimot økende grad av målstyring av lærernes arbeid i skolen, og også mot økende grad av profesjonalisering av læreryrket (Evetts, 2013; Sachs, 2016). Slik kan det hevdes av «accountability»-tenkningen i hvert fall indirekte påvirker norske lærere.

2.5 Oppsummering

Så langt er det gjort rede for avhandlingens grunnlag med hensyn til valg av fenomen, av det teoretiske grunnlaget for fenomenet og definisjoner som viser hvordan det er forstått. Det er videre gjort rede for en norsk utdanningskontekst der VfL skal realiseres. Jeg har også synliggjort hvordan norske myndigheter har satset på VfL. Avhandlingen er et forskningsbasert bidrag som kan øke forståelsen av hvordan satsingen på VfL kan komme til uttrykk i norskfagets vurderingspraksis. Spørsmål som følger i kjølvannet, er: Hvilken forskning ligger til grunn for anbefalingene av en VfL-orientert praksis innen feltene jeg har undersøkt?

3 Litteraturgjennomgang

Ettersom forskningsfeltene avhandlingen bygger på er så omfattende, vil jeg vektlegge studier som er mye sitert og anses som viktige, blant annet ved at de er framhevet i den internasjonale forskningen om formativ vurdering (Gough et al., 2012), og ved at denne forskningen ligger til grunn for retningslinjer for læringsfremmende vurderingspraksis som er fremmet av norske myndigheter. Litteraturgjennomgangen av tilbakemeldingsforskning på et overordnet plan, av elevers oppfatninger av lærers tilbakemeldinger og av læreres faktiske tilbakemeldinger i ulike skrivekontekster vektlegger sentrale studier, fordi forskningsfeltene er omfattende. Det samme gjelder forskning på lærerprofesjonalitet og lærersamarbeid. Litteraturgjennomgangen er i hovedsak rettet mot forskningsfunn.

I en norsk kontekst er det mulig å ha en mer «aggregative» tilnærming (Gough et al., 2012, s. 3), det vil si at forskning på VfL i skriftlig norsk i videregående skole i større grad kan oppsummeres, og at både kvantitative og kvalitative undersøkelser er tatt med fordi feltet er noenlunde avgrenset.

I det videre følger en gjennomgang av sentral forskning innen VfL i Norge, en presentasjon av norskfaget og spesielt skriveopplæringen i videregående skole, der også teoretiske perspektiver er inkludert fordi teori om skriving og en beskrivelse av skriving og skriveopplæring i en norsk utdanningskontekst er tett sammenvevd. I tillegg presenteres forskning på tilbakemeldinger som er en essensiell del av VfL, og tilbakemeldinger på skriving mer spesielt, som snevrer inn perspektivet i tråd med avhandlingens forskningsspørsmål. Videre er det lagt vekt på elevers oppfatninger og bruk av tilbakemeldinger og kriterier, som ses som et uttrykk for målet med opplæringen. I tillegg vil forskning som stiller seg kritisk til VfL, inkluderes i litteraturgjennomgangen. Deretter følger en redegjørelse for lærerprofesjonalitet og lærersamarbeid.

3.1 Hva sier forskningen om VfL i Norge?

Underveis i satsingen mot en mer læringsfremmende vurderingspraksis har det vært enkelte evalueringer. Særlig er FIVIS-rapportene sentrale, der vurderingspraksis i flere fag på åtte skoler ble kartlagt via blant annet observasjon og intervjudata. Sluttrapporten (Sandvik & Buland, 2014) framholder at det er langt igjen til myndighetenes satsing på en læringsfremmende

vurderingspraksis manifesterer seg. Det finnes også annen forskning som både framhever at satsingen er viktig på mange skoler (Hertzberg & Roe, 2016) og at det er mangler knyttet til gjennomføringen, men i et stort perspektiv er det likevel lite forskning som har undersøkt hvordan den store satsingen viser seg i norskfaget i videregående skole.

Det finnes imidlertid studier som kartlegger hvordan VfL praktiseres i andre fagfelt. Gamlem (2014) orienterer sin forskning mot grunnskolen og læreres tilbakemeldinger i klasserommet og finner at det er liten grad av læringsfremmende tilbakemeldinger, i likhet med Saliu-Abdulah, Hellekjær og Hertzberg (2017) som har studert læreres tilbakemeldinger i engelskfaget i et VfL-perspektiv. Også Burner (2016) finner problemer med implementeringen av en VfL-praksis i engelskfaget ved at elever og lærere har ulik og dels motstridende forståelse av fenomenet. Havnes et al. (2012) har sett på grunnskolen i en større studie som dekker flere fag og finner, i likhet med studiene som er nevnt, utfordringer ved implementeringen av VfL i skolen. Bueie (2016; 2015) har undersøkt elevens oppfatninger av lærers skriftlige tilbakemeldinger i norskfaget i videregående skole, og finner blant annet at elever leser og verdsetter lærers tilbakemeldinger, og at de har klare preferanser på hva de setter pris på. Alt i alt viser gjennomgangen at forskningen på VfL i videregående skole er forholdsvis begrenset, og at forskerne finner implementeringen av VfL mangelfull, selv om det også er funn som peker i motsatt retning.

3.2 Norskfaget i videregående skole

Fordi avhandlingens empiriske grunnlag er norskfaget i videregående skole, og særlig Studiespesialiserende studieretning (heretter ST), legges det i den videre redegjørelsen hovedvekt på hvordan faget framtrer i videregående skole. I noen tilfeller er helheten i faget også trukket inn.

Læreplanen i norsk (Udir, 2013) fastslår at fagets formål er å fremme kulturforståelse, kommunikasjon, dannelse og identitetsutvikling hos elevene. De skal videre utvikle en mangfoldig språkkompetanse, og det framheves at skrivekompetanse som utvikles i faget, er et nødvendig grunnlag for elevenes læring og forståelse på alle trinn og i alle fag. Elevene skal dessuten møte et bredt spekter av tekster, produsere ulike teksttyper og tilpasse språket mottaker og skrivesituasjon. Læreplanen vektlegger også både historiske og internasjonale perspektiver. Av de grunnleggende ferdighetene tillegges norskfaget et særlig ansvar for opplæringen i muntlige og skriftlige ferdigheter. Faget er inndelt i tre hovedområder – muntlig kommunikasjon, skriftlig

kommunikasjon, språk, litteratur og kultur – med tilhørende kompetansemål. Ferdigheter som vektlegges i kompetansemålene på tvers av årskurs og studieprogrammer er bruk av fagbegreper, f.eks. om retoriske appellformer, bruke og redegjøre for språklige virkemidler, tolke og vurdere sammenheng mellom innhold, form og formål i ulike tekster, vise til og bruke kilder. Dette er bare et utvalg av mål som går over flere år av opplæringen.

Faget har mange timer, sammenliknet med andre fag, og elevene får tre karakterer etter tiende klasse og etter å ha fullført studiespesialiserende studieretning på Vg3 (hovedmål, sidemål, muntlig). Når en legger til eksamenskarakterer, kan en uteksaminert elev etter videregående skole ha seks norskkarakterer på vitnemålet, i tillegg til karakterene fra grunnskolen (minst tre, fire eller fem hvis eleven trekkes ut til eksamen skriftlig og muntlig). Summen av karakterer viser at faget tillegges stor vekt av myndighetene, og det å lykkes i faget vil følgelig kunne bety mye for en elevs framtidsmuligheter.

Også for fylker, kommuner og skoler har resultatene på norskeksamen betydning. Resultatene er offentlig tilgjengelige og kan sammenholdes med standpunkt karakterer og andre fag ned på skolenivå (f.eks. i statistikkportalen.udir.no). Slik kan de brukes til offentlige rangeringer av skoler, noe som kan ha betydning for elevers skolevalg, og følgelig for inntakssnitt for hver enkelt skole. Videre kan statistikken brukes til å kartlegge om noen skoler har varige og tydelige avvik mellom standpunkt og eksamenskarakterer. Eksamen i norsk fungerer slik som en potensiell målestokk for skolekvalitet.

Om en oppsummerer aspektene ved faget som er nevnt over, viser det seg at faget har stor betydning, både for skoleeier, skolens ledelse, lærerne og enda mer for elevene. Når en også ser at faget tillegges et hovedansvar for skriveopplæringen og at denne anses som et nødvendig grunnlag for at elevene skal lykkes i alle fag, så understreker det viktigheten av å ha forskningsbasert kunnskap om skriveopplæringen i norskfaget.

En viktig del av studien er å undersøke hva slags tilbakemeldinger elever får av lærere i skriftlig norsk i videregående skole. Noen sentrale spørsmål jeg vil søke å belyse, er: Hvilke ideologiske og teoretiske perspektiver ligger til grunn for skriftlig norsk – i et historisk perspektiv og i dag?

3.2.1 Hvilke ideologiske og teoretiske perspektiver ligger til grunn for skriftlig norsk?

Skriveforsker Kjell Lars Berge (2005b, s. 16) definerer skolens oppgave som skriveopplæringsinstitusjon som dette:

I vårt samfunn er det skolens oppgave å sørge for at barn og ungdommer kan utvikle semi-otiske meningsskappingsressurser (...). Det betyr (...) at skolen skal legge forholdene til rette for å utvikle elevenes tekstferdigheter på et nivå som setter dem i stand til å skape tekster (...) som kan fungere på relevante måter i ulike tekstkulturer.

Som en presisering og eksemplifisering legger han til at en elev ikke er forventet å skulle kunne skrive et essay etter høykulturelle normer, men at de skal mestre det å skrive et essay «som innenfor skolens tekstkulturer blir forstått som essays» (Berge et al., 2005b, s. 16). Når disse normene i tilstrekkelig grad er oppfylt, får elevene status som deltakere i tekstkulturen. Hva som defineres som tekst vil variere innen ulike kultursfærer, men en deltaker er en person som er i stand til å produsere noe som de andre deltakerne anerkjenner som en tekst (f.eks. vil graffiti på en t-banevogn i noen kulturkontekster defineres som tekst, i andre som hærverk). Dette synet på eleven som aktør eller deltaker i en kulturkontekst er forankret i flere teoriretninger, som jeg gjør rede for i det følgende.

For å definere skriving i skolen i et teoriperspektiv er det relevant å se på den internasjonale IEA-undersøkelsen som kartla «teaching and learning of written composition» (Gorman, Purves & Degenhart, 1988). IEA-forskerne analyserte seg fram til fire ideologiske posisjoner som lå til grunn for skriveprøvene i medlemslandene (Saari & Purves, 1992, s. 51): 1) en personlig-vekstorientert skriveideologi («personal growth approach»), 2) en kunnskaps- og kulturarvorientert teori (kalt «heritage» eller «knowlegde about writing-approach»), 3) en ferdighetsorientert ideologi («competence/skills in writing») og 4) en verdiorientert skriveideologi, kalt «academic writing as humanistic value».

I sin analyse av skriveprøven i videregående skole framholder Berge (1996) at det er en kunnskaps- og kulturarvorientert ideologi som dominerer morsmåslæreres «doxa», det vil si deres faglige ideologier og forestillinger. Det har den gjort i den såkalt lærde skolen siden 1800-tallet. Noe av essensen er at elevene skal tilegne seg kunnskap og kulturarv gjennom lesing av kanoniserte verk, og at de skriver etter mønster fra modelltekster, vanligvis hentet fra en nasjonal (skjønnlitterær) kanon. I tillegg kommer opplæring i grammatikk og andre tilgrensende emner som i sum støtter opp under den grunnleggende ideologien. Denne skrivemåten har dominert i skolen, særlig på det vi i dag kjenner som videregåendenivå, fram til 1980-årene, mens den allerede fra 1950-årene vek plassen for en mer funksjonell, samfunnsorientert nytteideologi (Øygarden, 2001) i lavere klassetrinn. Etter innføringen av K06 og revisjonen av læreplanverket for norskfaget i 2013 kan en si at denne tendensen, eller bevegelsen bort fra kunnskaps- og kulturarvideologien, har blitt enda mer framtrædende. En kan fortsatt se at kunnskap om litteraturhistorie er etterspurt

på norskeksamen (f.eks. våren 2013 og 2016), men det er elevenes tekst-, tolknings-, skrive- og refleksjonskompetanse som står i sentrum. At det våren 2014 ble innført en korttekstoppgave på norskeksamen, kan ses som uttrykk for dette. Kortteksten, som alle må svare på, har så langt vært utformet slik at elevene må bruke ett eller flere tekstvedlegg og slik vise tekstkompetanse. I tillegg må de vise kunnskaper om ulike fagfelt, ferdigheter i teksttolkning og anvendelse av fagbegreper.

En grunnleggende antakelse i IEA-undersøkelsen var at sjanger var en bærende konstruksjon i elevers skrivearbeid (Berge, 2005b, s. 20). I KAL-studiene av elevers avgangstekster i grunnskolen 1998–2001 fant forfatterne at svært få sensorer opptrådte som formalister med hensyn til formelle ferdigheter eller med hensyn til sjanger (Berge, 2005a). Det som i stedet dominerte, var at elevene ble sett som relativt autonome skrivere, og at skrivesituasjonen som eksamen representerer, var påfallende symmetrisk (Berge, 2005b, s. 17) i strid med det en kunne forvente seg, ettersom elevene ble (og blir) vurdert av en sensor og tildelt en karakter med betydning for deres videre muligheter. En kan si at elever og sensorer forhandler om meningsdannelse og kvalitet i skrivingen. Noe som underbygger dette perspektivet, er en forholdsvis stor forekomst av det Andersen og Hertzberg kaller risikotekster (2005b, s. 300). At så mye som 40 % av elevene var villige til å ta en risiko på eksamen, tolkes av Andersen og Hertzberg som et uttrykk for at skolens skriveopplæring preges av samme symmetriske tendens, og at elevenes individuelle skriveprosjekt står i sentrum i opplæringen – ikke ulike varianter av formalisme. Dette viser seg også i data fra Normprosjektet (Skar & Aasen, 2016, s. 15), der forskerne finner at når elever løser en skriveoppgave til dels svært ulikt, kan bedømmere trenge eksemplifiseringer av hvordan kriterier skal anvendes. Lærerens tilbakemeldinger til egne elever kan derfor være en arena for en slik forhandling og et startpunkt for videreutvikling av teksten. Dermed gir innsikt i utforming av lærerkommentarer også en pekepinn om at lærerne i 2015 orienterte seg mot en dialog med elevene.

Et sentralt poeng i skriveideologien som har dominert i Norge de siste tiårene, er at teksten skal vurderes som en helhet, ikke som ulike bestanddeler av formal karakter. Summen av elevenes prestasjon – forstått som en kommuniserende helhet, er det sentrale (Udir, 2017b). Selv om sensorskjemaer som brukes til eksamen er inndelt i kategorier, skal disse oppsummeres i en holistisk vurdering, og ikke være et produkt av en kvantifisering av enkeltdelene. Den samme holdningen finnes i sensorveiledninger i KAL-materialet, der instruksjonen helt klart gikk imot en sjangerformalisme (Berge, 2005b, s. 25–26).

Det står dessuten helt sentralt at alle elevtekster vurderes som unike. «De er et umiddelbart uttrykk for tekstskaperens forsøk på å lage en samlet forståelse av kunnskaper og mentale og sanselige erfaringer gjennom en tekst i en særskilt skrivesituasjon» (Berge, 2005b, s. 12). Han fortsetter med å si: «Samtidig er alle skapt innenfor en situasjons- og kulturkontekst der det gjelder visse normer for hvordan ytringer skal utføres for å gjelde som tekster, hvilke variasjonsmuligheter som gjelder for utføring, og hvordan de er bygd opp.»

Dette siste peker mot en kulturemiotisk forståelse, der Lotmann (1990) ser kulturen som et åpent hierarki av tekster, der noen av tekstene fungerer som modelltekster fordi de fastholder distinktive trekk i kulturen som også definerer hva som ikke kan regnes som tekster. Elevtekster vil måtte kvalifisere som tekster innenfor en kontekst som både er definert av kulturen i samfunnet og av skolekulturen. For å få til dette vil eleven konstituere seg i skriveprosessen som en modellsender som skriver for en modelleser. Gjennom å tilegne seg tekstnormene som gjelder i skolen, og i neste omgang i samfunnet utenfor, vil eleven som nevnt i innledningen kvalifisere seg som deltaker i det Luhmann (1986, s. 179) ser som en mangefasettert verden der kommunikasjon er et bærende aspekt ved individets samfunnsrolle, der det gjennom språklig interaksjon med andre kvalifiserer seg som «legitim deltaker» (Lave & Wenger, 1991). Dette impliserer også kognitive prosesser som griper ut over den opplagte situasjonen i en skolesammenheng og modellen som ligger til grunn for norskeksamen, slik det er presentert i KAL-materialet og kommer til uttrykk blant annet i «skrivehjulet» (Fix, 2008; Sandvik, 2012, s. 159). Der står eleven som aktør i sentrum, og hjulet synliggjør en rekke valg eleven kan foreta, på basis av sosiale, kognitive og diskursive ressurser som alle kan bidra til å oppfylle skriveaktens formål. Dette knytter seg også til forståelsen av elevens tekst som et meningspotensial (Halliday, 1994) som utvikler seg i en kulturkontekst, og der teksten tolkes i en skrivesituasjon som Halliday kaller en situasjonskontekst. Denne vektleggingen av språklig samhandling leder til forståelsen av skrivingen som en kontekstuell aktivitet som også står i forbindelse med Vygotskijs syn på læring som sosiokulturelt betinget. Vygotskij (1978, s. 117) sier blant annet at for at elevene skal lære å skrive, må en ta med i betraktningen at skriving er en aktivitet som også er å forstå som en kulturell ferdighet.

Som vist kan skriving ses i et sosialt, kommuniserende perspektiv, men de prosessene som en kan legge til grunn at foregår i skriverens bevissthet, har også betydning. I et banebrytende verk lanserte Bereiter og Scardamalia (1987) to modeller som beskriver underliggende kognitive prosesser som aktiveres i skriveprosesser og er eksempler på hvordan skriving som en problemløsende aktiviteten

kan forstås. Den ene modellen (1987, s. 8) beskriver kunnskapsgjengivelse («The knowledge-telling model»), og viser hvordan emne- og diskurskunnskaper aktiveres og revideres under prosessen fram mot en ferdig tekst. Den andre modellen («The knowledge-transforming model», 1987, s. 12) gjengir en mer kompleks modell, som beskriver prosesser der kunnskap om både emne og diskurs utvikles i skriveprosessen og understøtter tenking på et mer avansert nivå enn kunnskapsgjengivelsesmodellen. De understreker likevel at også avanserte tekster kan komme ut av kunnskapsgjengivelsesmodellen, gitt at tankeprosessene i bearbeidningen av fagstoffet har foregått på forhånd. Forfatterne underbygger forståelsen av de to modellene med omfattende empirisk forskning som undersøker ulike stadier i skriveprosessen, og denne forskningen er også forankret i kognitiv (lærings)teori.

For at skriveopplæringen skal fungere best mulig, må skrivingen være relevant for elevenes liv ved at den framstår som meningsfull og at skrivesituasjonene er mest mulig autentiske, et poeng blant annet framhevet av Smidt (2002). Et forsøk på å operasjonalisere disse tre faktorene kan en hevde at finnes i prosessorientert skrivepedagogikk. Norsk skole er blitt påvirket av prosessorientert skrivepedagogikk, utviklet blant annet av Britton (1975) og videreutviklet i Norge i 1980-årene, blant annet av Dysthe (1987). Ved at elevene utvikler og forbedrer egne tekster over tid og presenterer et eget utvalg som leveres til vurdering, kan en si at eleven både styrkes som aktør og at skrivesituasjonen søkes gjort mer meningsfull. Det vil i tillegg kunne foregå en utstrakt grad av samtale, tilbakemelding og bearbeidning av elevenes tekster i samspill med både medelever og lærere, et poeng som også framheves av Krogh og Jensen (2008) og Hertzberg og Roe (2016).

Et sentralt moment i forståelsen av kommunikasjon i dette perspektivet, er at kommunikasjonen ses som en forhandling om mening, og at den forutsetter både samarbeid (Grice, 1975) og dragkamp (Nystrand & Wiemelt, 1991). Skriveforsker Lars S. Evensen (2005, s. 194) tar utgangspunkt i sosialpsykologen Rommetveits kommunikasjonsprinsipper og sier at kommunikasjonen preges av 1) gjensidighet, 2) en felles avgrenset språkbruk og 3) empatisk språkbruk. Når disse faktorene trer i kraft, oppstår noe han kaller «temporarily shared social reality» (Rommetveit, 1974). I denne fenomenologiske sammenhengen ses elever (og lærere) som aktører som handler ut fra egne interesser og formål, og opptrer i sosialt definerte roller i en livsverden (der begrepet kan synes å ligge nært Luhmanns (1986) forståelse av eleven som deltaker i en kulturell og sosial verden).

Denne vekten på dialogisme og på prosesskriving leder over til at skrivingen i større grad ses i et diskursperspektiv, der begreper som aktør, prosess og deltakelse står i sentrum (Evensen, 2005, s.

195) og ikke ytre normer, f.eks. gitt av sjangerkrav eller formelle trekk ved skrivingen. Et grunnlag for denne forståelsen finnes i symbolsk interaksjonisme (Blumer 1969; Charon 1995; Mead 1934), der en mener at identiteten skapes i samspillet med andre. En elev vil i dette perspektivet romme et subjektivt, individuelt «jeg» som påvirkes av og samspiller med et kulturelt konstruert «meg». En kan tenke seg at en elev som konstituerer seg som modellskriver, definerer seg selv i forhold til nettopp dette kulturelt betingete «meg». Skriveren vil ha tilgang til en rekke konvensjoner som fungerer som ferdig etablerte rammer, slik at ingen skriveprosess begynner helt forfra, uten en felles gitt forforståelse. Sjangerer kan forstås som en slik type ramme, av Bakthin (1986) definert som en erfaringsbasert og generalisert (konvensjonell) diskurskontrakt, som finner sted innenfor en tolkningsramme og en forventningshorisont for lesing og skriving (Evensen 2005, s. 200). Berge (2005b, s. 54) sier samtidig at i KAL-materialet er sjangrene som er angitt i skriveoppgavene å anse som *tilbud* til elevene, og ikke som påbud om faste skrevenormer etter modelltekster. Retningslinjer for norskeksamen fra Utdanningsdirektoratet (2017b) viser at betegnelsene som erstattet tidligere sjangerer, er informative, argumenterende, resonnerende og kreative tekster. At tradisjonelle sjangerer er erstattet, kan tjene som argument for at eleven blir sett som en selvstendig skriver som utnytter de ressursene hun/han har, i tråd med det som er framhevet over. Det kan også ses som et uttrykk for at en ferdighetsorientert ideologi er til stede i forståelsen av norskfaget, slik den kommer til uttrykk i sammenheng med eksamen.

Det understøtter ytterligere synet på skriving i norsk skole som skriving som i noen grad, eller kanskje ideelt sett, foregår på elevenes premisser, der de ses som selvstendige, tolkende individer som gjennom sin tekstproduksjon skal skape tekster som anerkjennes i kulturkonteksten, og at lærer følgelig mer er å se som en dialogpartner i en forholdsvis symmetrisk relasjon, slik Andersen og Hertzberg (2005) peker på. Samtidig er det langt på vei gjennom lærerens tilbakemelding at elevene blir kjent med og sosialisert inn i normene for skriveopplæringen, enten det er kortsiktige delmål eller det Marshall (2004) kaller horisontmodellen, der helhetsforståelse og mestring av tekstnormer står sentralt. Lærers vurdering, kommunisert gjennom tilbakemeldingen, bygger en bro mellom elevens forståelse og skolens, der lærer får en rolle som normbærer.

Redegjørelsen over viser at det er vide rammer for hva som kan regnes som en tekst i skolens fellesskap. Det blir dermed relevant å definere målet for opplæringen ut fra en horisont-modell (Marshall, 2004), det vil si en holistisk forståelse. I denne modellen er en viktig del av lærers arbeid å klargjøre mål og forventninger. Disse er mangfoldige og beskrives derfor av metaforen horisont

heller enn at begrepet mål anvendes. Det kan eksemplifiseres med at elever som bedrer seg på noen områder skriftlig, kanskje som en følge av lærers tilbakemelding, opplever at karakterer ikke bedrer seg nevneverdig fordi det å mestre en skriveoppgave krever så mange og sammensatte ferdigheter. Hvis eleven lærer seg å mestre kommatering, kan det eksempelvis hjelpe lite om tekster som leveres inn til vurdering har store strukturproblemer eller mangler fagspråk. Lærer må likevel søke å tydeliggjøre hva som er mål og kriterier. Kriterier spiller en rolle i operasjonalisering av målet, men de fanger bare opp aspekter ved målet og kan ikke være uttømmende uten å samtidig miste praktisk betydning (Sadler, 2010), noe som bidrar til å gjøre det krevende å klargjøre mål for læreren. I en studie viser Myhill, Jones, Lines og Watson (2012) at gode fagkunnskaper hos læreren ser ut til å spille en viktig rolle for elevenes læring i grammatikk, et emnefelt som er nedfelt i norsk læreplan. Det er grunn til å anta at det samme gjelder for andre aspekter av skriveopplæringen, også evnen til å gi tilbakemeldinger.

3.3 Tilbakemeldinger og tilbakemeldinger som kan bidra til læring

Ettersom forskningen på tilbakemelding (feedback) er svært mangfoldig, er den videre gjennomgangen orientert mot læringskontekster med vekt på noen få innflytelsesrike studier. Gjennomgangen vil rette seg mot tilbakemelding generelt, tilbakemelding på skrijving og deretter tilbakemelding på skrijving i norskfaget i et VFL-perspektiv, i den grad det er mulig å finne forskning om det siste.

Tilbakemeldingen spiller en helt sentral rolle i en læringsfremmende vurderingspraksis (Black & Wiliam, 2009). Begrepet defineres av Hattie og Timperley (2007, s. 81) som «information provided by an agent (e.g. teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding.»

Denne definisjonen ser tilbakemelding på et overordnet nivå, og Shute (2008, s. 154) tilfører en ekstra dimensjon ved å rette fokus mot formativ tilbakemelding: «Formative feedback is defined (...) as information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning.»

Sitatet understreker at tilbakemeldinger har potensial til å påvirke læring. Et viktig spørsmål blir derfor hva forskningen sier om tilbakemeldinger.

3.3.1 Forskning på tilbakemeldinger generelt

Forskningen som ligger til grunn for anbefalingene, er i stor grad bygd på effektstudier som er brukt av forskerne til å utlede råd om god praksis. I det følgende blir tre sentrale, det vil si mye siterte og innflytelsesrike studier, omtalt spesielt. Den første er Kluger og DeNisis (1996) metastudie som undersøkte 131 studier innen ulike fagfelt på tilbakemeldingsintervensjoner i forskning som strekker seg tilbake til begynnelsen av 1900-tallet. Opprinnelige feedbackdefinisjoner stammer fra tekniske fagfelt (Kluger & DeNisi, 1996), der feedback er informasjon som gis en komponent, for eksempel en termostat, som i sin tur for eksempel sørger for at en varmeovn skrur seg på slik at referansetemperaturen kan oppnås. At termostaten reagerer på informasjonen den får ved å sende signal til varmeovnen, er det essensielle punktet og det som skiller feedback fra informasjon. Denne dimensjonen er også framhevet som essensiell i VfL, som vist i kapittel 2.

Noen særlig viktige funn hos Kluger og DeNisi (1996) er at tilbakemeldinger kan fremme læring gjennom å rette oppmerksomheten mot hva som er korrekt svar heller enn hva som er feil, og at påvirkningskraften er større når det er klare mål for oppgaven. De finner videre at ros er en lite effektiv tilbakemelding når den ikke er rettet mot trekk ved oppgaven, og at tilbakemelding kan føre til redusert prestasjon hvis den inneholder kritikk som oppleves som truende for selvet (1996, s. 273). Kluger og DeNisis funn er omtalt i to andre metaanalyser som også skal presenteres her. Den første er Hattie og Timperleys (2007) «The power of feedback». Her har forfatterne blant annet oppsummert 12 tidligere metaanalyser som undersøker effekten av tilbakemeldingsintervensjoner, kontrollert for minst 14 faktorer (2007, s. 84), og på bakgrunn av tidligere forskning presenterer de en modell (2007, s. 87) for effektive tilbakemeldinger, det vil si tilbakemeldinger som får den lærende til å prestere bedre. Hovedtrekkene i modellen er at formålet med tilbakemeldingen er å redusere avstanden mellom faktisk prestasjon og ønsket prestasjon. For å oppnå dette kan elever gjøre ekstra innsats eller bruke mer effektive strategier (eller gi opp/ redusere målet for egen del), mens læreren kan angi passende vanskelige mål og gi elevene støtte gjennom å lære dem strategier for å løse en oppgave og gi tilbakemelding. Tilbakemeldingen bør svare på tre spørsmål som kaster lys over hvordan eleven har prestert i forhold til målet og hva eleven bør gjøre for å komme nærmere målet (Where am I going?/How am I going?/Where to next?). I tillegg virker hvert spørsmål på fire nivåer: oppgavenivå, prosessnivå, selvreguleringsnivå og selvnivå. De mest effektive tilbakemeldingene retter seg mot oppgaven eller prosessen, eller selvreguleringsnivåene, mens de som retter seg mot selvet i mindre grad er effektive fordi de tar oppmerksomheten bort fra oppgaven.

Shute (2008) har i den nevnte reviewartikkelen undersøkt forskning på formativ tilbakemelding spesielt, og den representerer dermed en innsnevring av perspektivene fra Kluger og DeNisi (1996) og Hattie og Timperley (2007), selv om mange perspektiver er felles. Shute (2008) viser at det ikke noe enkelt svar på spørsmålet om hva som bidrar til læring, men nevner en rekke forhold som kan tas med i betraktning når en lærer eller medelev skal gi tilbakemelding. Disse inkluderer blant annet at ulike former for tilbakemeldinger har størst sjanse for å påvirke elevens læring når læreren tilpasser tilbakemeldingen til tidspunkt (umiddelbart/forsinket), hvem den lærende er (faglig sterk/svak) og kjennetegn ved oppgaven (enkel/kompleks). I tillegg bør lærer ha bevissthet om hva man bør unngå (normativ/sammenliknende tilbakemelding, karakterer sammen med formativ tilbakemelding, personrettet ros, demotiverende kritikk, gi rett svar, omfattende feilanalyse, og motsatt, hva som kan bidra til læring på et generelt nivå (rette fokus mot oppgave, utdype poenger, ikke gi for mye informasjon, være tydelig, knytte til mål, formulere den skriftlig og gi informasjon etter at den lærende selv har forøkt å løse oppgaven) (Shute, 2008, s. 177–181). I tillegg bør eleven eller studenten være motivert, ha mulighet til å bruke tilbakemeldingen og dessuten faktisk greie og ønske å bruke den (2008, s. 175).

3.3.2 Tilbakemeldinger på skriving og hvordan de kan bidra til læring

Det finnes mange studier av læreres tilbakemeldingspraksiser innen språkopplæring. Særlig innenfor andrespråk- og fremmedspråkopplæringen (Ferris, 2004; Lee, 2004; Truscott, 1996), men mange studier retter seg også mot høyere utdanning (Dysthe, 2011; Sadler, 1989; 2010) eller mot grunnskolen. Mange studier av læreres tilbakemeldinger i skriveopplæringen retter seg dessuten mot grammatikk (Ferris, 2004; Lee, 2004) og ulike nivåer av språklig korrigerings (Bergman-Claeson, 2003; Lee, 2007), eller de har et helhetsperspektiv der de har kategorisert ulike talehandlinger (Magno & Amarles, 2011; Tunstall & Gipps, 1996). Mye sitert er også Sommers (1982) som fant at studenter ofte fikk vage tilbakemeldinger, noe som delvis støttes av Kronholm-Cederberg (2009), som fant at finske gymnaselever hadde problemer med å forstå hva læreren skrev til dem. Dette funnet ga meg motivasjon til å undersøke hva norsklærere skriver til egne elever.

Studier av MacArthur (2012) og Graham et al., (2011) viser at lærers tilbakemeldinger kan ha positiv effekt på elevers skriveutvikling, og forskning på respons på skriving har overlappende funn med forskning innen VfL, der bruk av informasjon fra tilbakemeldinger i videre læringsarbeid er et kjernepunkt. I en sentral newzealandsk studie (Parr & Timperley, 2010) analyseres 59

grunnskolelæreres tilbakemeldinger på en elevtekst i en VfL-kontekst. Forskerne fant at de lærerne som best viste evne til å gi VfL-tilbakemeldinger, hadde elevene med størst læringsframgang, målt i score på en nasjonal skriveprøve i førstespråket. Parr og Timperleys (2010) studie har et eksperimentelt design, og resultatene har derfor bare indirekte overføringsverdi fordi jeg undersøker lærernes tilbakemeldinger på en heldagsprøve i norsk – en lang og kompleks tekst sammenliknet med den Parr og Timperley (2010) bruker i sin studie. Likevel er den relevant, fordi den kombinerer skriveopplæring med VfL gjennom å kategorisere førstespråklæreres tilbakemeldinger på en elevtekst. Den er også metodisk overførbar ved at kategoriene forskerne har utviklet, har relevans i norsk skriveopplæringssammenheng, noe som blir gjort rede for i metodedelen.

I en studie som kartlegger elevers oppfatninger av tilbakemeldinger i klasserommet finner Gamlem og Smith (2013) at lærere i ungdomsskolen sjelden legger til rette for bruk av tilbakemeldinger, og at klassen ofte er i gang med noe nytt når elevene får tilbakemelding på et arbeid. De framhever også at det ofte er frivillig om elevene skal jobbe videre med tilbakemeldinger eller ikke, og at det er større sjanse for at elevene nyttiggjør seg tilbakemeldingene om læreren legger til rette for bruk. Dette skjer, ifølge Gamlem og Smith, sjelden. Dette funnet er i samsvar med Bueie (2014; 2015; 2016), som i flere studier har utforsket elevers oppfatninger av norsklærers tilbakemeldingspraksis innenfor et VfL-rammeverk. Her kommer elevperspektivet fram på individnivå på bakgrunn av både kvalitative intervjudata og kvantitative data. Bueie (2015) finner blant annet at elever stort sett forstår hva læreren skriver til dem, men at de bare jobber videre med en tekst om de får eksplisitt instruks om det.

Dysthe (2011) argumenterer i en studie av tilbakemeldinger til masterstudenter for at tilbakemeldinger kan ha læringspotensial om de retter seg mot dypereleggende forhold i teksten, selv om teksten ikke skal revideres. Dysthes studie (2011) har relevans for skolen fordi tekstlige forhold også i norskfaget har likhetstrekk og dermed overføringsverdi fra gang til gang, slik det er gjort rede for i kapittelet om norskfaget i skolen. At lærers tilbakemeldinger kan rette oppmerksomhet mot tekstforhold som har overføringsverdi til framtidige skrivesituasjoner, understrekes også av skriveforskeren MacArthur (2012). En forutsetning er at den aktuelle informasjonen gjøres forståelig og tilgjengelig for elevene.

3.3.3 Elever- og studenter bruk av lærers tilbakemeldinger på skriving

Spørsmålet om hva som kan bidra til å forklare studenters bruk eller manglende bruk av tilbakemeldinger, er kartlagt i en rekke studier. Jönsson (2013) framhever i en reviewartikkel ulike

forhold som kan bidra til å forklare hvorfor bruk av tilbakemeldinger i skriveopplæring synes å representere en utfordring. Han finner at studenter må oppfatte tilbakemeldingen som nyttig, og her er «timing» av betydning. Dette funnet samsvarer med Lea og Street (1998), som finner at det er et problem med tidspunktet for når tilbakemeldingen gis. En grunn til dette er at studentene ofte jobber med annet fagstoff når de får tilbakemelding. Jönsson (2013) finner dessuten at det kan være et misforhold mellom hva studenter ønsker seg og hva de får.

Straub (1997) framhever i en eldre reviewartikkel om studenters oppfatninger av tilbakemeldinger at studenter i skriveopplæringen foretrekker tilbakemeldinger som framstår som klare. Klare tilbakemeldinger kan bidra til å utvikle teksten særlig på lokalt nivå gjennom å være spesifikke og å ha informasjon om hvordan teksten kan forbedres, et poeng som også framheves av Sadler (2010, s. 538). Straub (1997) finner videre at elever kan oppleve en tilbakemelding som nyttig hvis den begrunner karakteren på en måte som oppfattes som tydelig og rettferdig.

Straub (1997) viser i forlengelsen av dette at en rekke forhold knyttet til kontekst, som klasseromsklima, tonen i tilbakemeldinger, tidspunkt og i hvilken grad læreren er spesifikk og konkret, har betydning for hvordan tilbakemeldingen oppfattes, et poeng som også framheves av Gamlem og Smith (2013). Studenter misliker at tonen i tilbakemeldingen oppfattes som negativ, ironisk eller som et direktiv som ikke tar hensyn til studentens hensikt med en tekst. Dette er forhold som også Smidt (2002) har utforsket i en norsk sammenheng med liknende konklusjoner. På den andre siden er det fellesnevnerne for hva studenter og elever foretrekker av tilbakemeldinger. Ferris (2003) finner at studenter liker konkrete forbedringsforslag, Beach (1989) at de liker forklaringer på hvorfor noe er bra eller ikke og Straub (1997) framhever at tilbakemeldinger kan være kritiske, så lenge de ikke utfordrer tekstens overordnede ideer.

Gibbs og Simpson (2004) finner at flere forhold influerer studenters bruk av tilbakemeldinger i videre arbeid. De peker blant annet på at tilbakemeldingen bør være i samsvar med studentens opplevelse av oppgavens innhold, kravene i faget og i oppgaven, og at de har lært teknikker som gjør at de er oppmerksomme på og kan ta imot innholdet i tilbakemeldingen. Annen forskning på feltet finner blant annet at en autoritativ tone minsker muligheten for at studenter vil ta til seg tilbakemeldingen (Cowie, 2005), og at de mangler strategier for bruk (Jönsson, 2013). At bruk av tilbakemeldinger er et komplekst felt i skriveopplæring, illustreres også av Calhoon-Dillahunt og Forrest (2013), som i en pilotstudie utforsker hvilke holdninger skrivere har til tilbakemeldinger i en opplæringsfase og hvordan de bruker tilbakemeldinger. Forskerne finner at studenter i en

skriveprosess i liten grad endrer tekstene sine i tråd med lærernes anbefalinger, selv om de tydelig uttrykker at de setter pris på tilbakemeldingene. Videre finner forskerne at studenter ønsker seg tilbakemelding fra lærer, men at de reagerer ulikt etter at de har fått respons, et poeng som understøttes av Hyland (1998), som finner at det i noen grad er personavhengig hvordan en tilbakemelding oppfattes.

Tidligere forskning (Horvath, 1984) finner at lærere antar at elever og studenter forstår og bruker tilbakemeldinger, et funn som altså samsvarer med Calhoun-Dillahunt og Forrest (2013) og de konkluderer derfor med at studenter må lære å bruke tilbakemeldinger for å utnytte læringspotensialet som ligger i dem. Beach og Friedrich (2006, s. 227) legger i forlengelsen av dette poenget vekt på det holdningsskapende arbeidet som må ligge til grunn for at tilbakemeldinger skal være læringsfremmende.

Studenters oppfatninger henger som vist tett sammen med spørsmålet om bruk av tilbakemeldinger. Jeg skal kort skissere egen- og hverandrevurdering som perspektiver på bruk av tilbakemeldinger fordi dette er et kjerneelement i VfL-teorien, og fordi Jönsson (2013) i den nevnte review artikkelen finner at studenter mangler strategier for bruk av tilbakemeldinger. Sadler (2010) skisserer hverandrevurdering («peer assessment») som en potensiell løsning fordi denne typen arbeid gir studenter øvelse i å vurdere tekst på samme måte som lærere, og at de følgelig utvikler kompetanse gjennom å både gi og motta tilbakemeldinger blant annet fordi de får tilgang til det som ofte er taus kunnskap om tekstkvalitet. Topping (2009) er en sentral forsker innen dette feltet. Han framhever at når studenter eller elever får opplæring i å gi respons på tekster, så kan kvaliteten være like god som når eleven får respons fra lærer. Det er med andre ord et stort potensial for å lette lærers arbeid med å gi tilbakemeldinger hvis elever får øvelse i å bruke hverandrevurdering i skriveprosesser. Med tilstrekkelig opplæring og klare kriterier kan også egenvurdering ha en positiv effekt (Sadler, 1998; Topping, 2009).

Det er nedfelt i vurderingsforskriften at elever skal være involvert i vurdering av eget læringsarbeid (Udir, 2010b), noe som kan ses som en speiling av definisjonen av VfL, som vektlegger det samme. At elever får opplæring i å vurdere eget arbeid, etter klare retningslinjer eller kriterier slik Sadler (2005; 2010) tar til orde for, kan derfor ses som en sentral komponent for en læringsorientert tilbakemeldingspraksis. I Norge har Hoel (1995) tatt til orde for en liknende praksis i en doktorgradsavhandling om elevrespons på skrivning.

3.4 Kriterier og bruk av kriterier

Som vist spiller klare mål og kriterier en viktig rolle i tilbakemeldingsforskningen. En artikkel av Wiliam, et al. (2004, s. 54) framstiller en intervensjonsstudie der det å utvikle og integrere kriterier i undervisningsløpet er uttrykk for en operasjonalisering av felles mål for en læringsøkt. Forfatterne framhever, i likhet med Black og Wiliams' artikkel fra 1998, at felles mål er en sentral del av VFL.

Målet med skriveopplæringen er svært komplekst; Sadler (1989) viser hvordan det kan ligge opptil 50 latente kriterier til grunn for en god tekst – det vil si forhold ved teksten som leseren blir oppmerksom på i det skriveren ikke mestrer dem, men som en ikke ville festet seg ved om eleven hadde mestret dem. Når en skal kommunisere hva som er målet med opplæringen i en skrivepedagogisk kontekst, bør derfor læreren velge ut noen kriterier eller forhold som det fokuseres særlig på, da det er vanskelig for en elev å forholde seg til et stort mangfold av potensielle kriterier samtidig (Kvithyld & Aasen, 2011).² Sadler (2010) viser samtidig at kriterier i seg selv ikke bidrar til klarhet, blant annet fordi lærere og studenter kan ha ulike oppfatninger av innholdet eller at kriteriene ikke fanger opp forhold som er viktige for studentens oppgave. Bueie (2015) finner dessuten at elever i liten grad forholder seg til vurderingskriterier. Det legges likevel til grunn at det er bedre at læreren har kriterier enn at kriterier mangler fordi eleven dermed har tilgang til en skriftliggjøring av gjeldende normer som ellers ville være en del av lærers tause kunnskap (Sadler, 2010, s. 546). Sett i lys av Marshalls (2004) horisontmetafor for skriveopplæring, og at elevene har en forholdsvis selvstendig rolle som skrivere i videregående skole og at myndighetene siden 2007 har satset på utvikling av kriterier for fag, er det mest relevant å se skriveopplæringen i et både-og-perspektiv. Kriterier kan hjelpe elever, studenter eller lærere til å rette oppmerksomhet mot viktige deler av teksten og dermed gjøre det enklere å gi tilbakemelding.

Det er satset på å utvikle standarder for skriving i Normprosjektet (Berge & Skar, 2015; Evensen et al., 2016), der forskere og lærere sammen har jobbet mot utforming av nivåbeskrivelser og forskningsmessig holdbare typebeskrivelser av elevtekst på ulike alderstrinn, og der normbeskrivelser utarbeidet av forskere er prøvd ut i praksis av lærere i flere omganger. Resultatet er et instrument som kan bidra til å klargjøre hva en kan forvente av elevers skriveferdigheter på ulike alderstrinn, noe som igjen kan bidra til å samordne og forbedre instruksjonen av elevenes skriving. Dette har igjen til hensikt å gjøre elevene til bedre skrivere. Eksempelet illustrerer at det

² På grunnlag av denne antakelsen er derfor det komplekse forholdet «målet med opplæringen» i avhandlingen søkt operasjonalisert gjennom spørsmål som retter seg mot bruk av kriterier i artikkel 2 og 3, og kriterier er dels brukt som grunnlag for kategorien «mål» i artikkel 1, ettersom kriteriene kan ses som et aspekt ved lærerens operasjonalisering av målene med skriveopplæringen.

etter *Bedre Vurderingspraksis* (2007–2009) fortsatt investeres betydelige ressurser i utviklingen av vurderingskriterier og typebeskrivelser som blant annet skal bidra til mer profesjonelle bedømmere – det vil si at de følger standarder og finner tekstbaserte kvalitetskriterier framfor en mer taus kunnskap, slik faglig skjønn også kan komme til uttrykk. Selv om Normprosjektet er uten direkte forbindelse til VfL, er det likevel felles kjennetegn i vektleggingen av skriving som en prosess der elevens vekst står sentralt. At lærere dessuten er invitert inn i arbeidet med å arbeide fram typebeskrivelsene, og at disse er ment som hjelpemidler i vurderingsarbeid og skriveopplæring, kan forstås som uttrykk for en økende grad av profesjonalisering av norske skriveopplærere.

3.5 Kritikk av formativ vurdering

Bennett (2011) tar til orde for å videreutvikle definisjonen av formativ vurdering blant annet fordi det er vanskelig å måle effekten av et konstrukt det ikke er enighet om grensene for – altså at det både kan oppstå et begrepsvaliditetsproblem, og at det i forlengelsen kan oppstå et reliabilitetsproblem knyttet til måling av et slikt fenomen. I likhet med Black og Wiliam (2009) framhever Bennett (2011) kontekstens, som også kan bety fagkontekstens, betydning for effekten av å innføre en formativ vurderingspraksis. I forlengelsen av dette poenget vil det derfor lett oppstå uklarhet om hva VfL er, når det er opptil hver enkelt lærer å oversette de generelle prinsippene til praktiske handlinger i sitt eget klasserom, i sitt fag, med egne elever.

Det har også fremkommet kritikk av gjengivelsen av effektstudiene hos f.eks. Black og Wiliam (Bennett, 2011). Elwood (2006) mener eksempelvis at effektstørrelsene er overdrevne fordi de i mange tilfeller baserer seg på måling av svaktpresterende elevers læringsframgang, fordi slike elever har stor sjanse for å gjøre det bedre hvis man setter i gang et forbedringstiltak. Det synes likevel å være enighet om at de grunnleggende prinsippene i VfL bidrar til økt grad av læring når de praktiseres. Det betyr også at lærers tilbakemelding, som et ledd i en læringsfremmende praksis, har dette potensialet, gitt at den er knyttet til målet for opplæringen og inneholder informasjon til eleven om hvordan hun/han skal komme nærmere målet i form av en instruks eller forklaring. I tillegg må læreren legge til rette for bruk av tilbakemeldingen. At det å virkeliggjøre disse prinsippene kan være en svært kompleks prosess, illustreres blant annet av Svanes og Skagen (2016), som viser hvordan Hattie og Timperley (2007) i sin innflytelsesrike studie framholder at god tilbakemelding må oppfylle inntil ni simultane kriterier for å ha optimal effekt. Når man opererer med en så kompleks modell for god tilbakemeldingspraksis, vil en forsker lett kunne finne at lærere i liten grad

følger modellens anbefalinger langs alle disse dimensjonene. Black og Wiliam (2009, s. 24) viser til noe av det samme når de framhever at lærers tilbakemelding må tilpasses elevens faglige nivå, emosjonelle tilstand og motivasjonsnivå, og at dette i tillegg skal utformes i en klasseromskontekst, der det gjerne er 30 elever, og at læreren følgelig har sekunder på seg til å avgjøre hva som er en læringsfremmende tilbakemelding. Kanskje er det noe enklere å skrive tilbakemeldinger på skriftlige arbeider fordi læreren har bedre tid, og eleven har bedre tid til å forholde seg til tilbakemeldingen (Shute, 2008). Men også dette er en kompleks oppgave som dessuten bærer i seg utfordringen med å få eleven til å engasjere seg videre i teksten, slik at læring eller utvikling av skriveferdighet kan finne sted.

3.6 Lærerprofesjonalitet, lærersamarbeid og elevprestasjoner

Profesjonalitet er et begrep som ifølge Evetts (2013) er i stadig endring og som tolkes ulikt innen ulike yrkesgrupper. Et felles kjennetegn er at begrepet ofte brukes om sammenhenger der yrkesutøvelse er basert på kunnskaper og former for ekspertise som leger, jurister, lærere eller sosialarbeidere kan besitte. Tradisjonelt har bruk av skjønnerkunnskap basert på dømmekraft (Biesta, 2015) blitt framhevet som viktig attributt ved profesjonell yrkesutøvelse, og lærere er tradisjonelt betraktet som autonome yrkesutøvere. Et eksempel på dette finnes hos Lortie (1975), som gjennom en eggkartongmetafor beskriver at lærere jobber i atskilte sfærer, selv om de naturligvis også hører til en skole og et kollegium. Gjennom flere tiår er samtidig læreres jobb satt under press med tanke på resultatansvar eller «accountability», et begrep som ikke ofte brukes i en norsk kontekst, men som er sentralt i angloamerikansk skolekontekst (Glatter, 2012; Hanushek & Raymond, 2005), der lærere i økende grad er kommet under press fra myndigheter som ønsker sterkere kontroll med skolen³ (Ingersoll & Kralik, 2004).

I forlengelsen viser Sachs (2016, s. 421) hvordan lærerprofesjonalitet også er et mangetydig begrep og tegner opp en modell med fire dimensjoner hvor samarbeidende profesjonalitet («collaborative professionalism») er en av de fire. Dette ses som en form for organisatorisk profesjonalitet som også

3 I denne studien brukes begrepet som inngang til å forstå framveksten av et økende krav om at lærere skal arbeide kollektivt. Å fremme lærerprofesjonalitet har blitt holdt fram som ønskelig utvikling fra politisk hold, slik det blant annet uttrykkes i St.meld. nr. 11 (2008–2009). Poenget er også blitt vektlagt i utdanningsforskning (Sachs, 2016), og det er en tendens at nasjonale utdanningsmyndigheter (Mausethagen & Granlund, 2012) og KS (Lejonberg & Elstad, 2014) i sterkere grad enn lærernes interesseorganisasjoner har vektlagt deltakelse i kollektive læringsprosesser i skoleorganisasjonen.

er tett forbundet med lærernes og skolens verdier, og hvor fornyelse, refleksjon og forbedring, både på skole- og individnivå, står sentralt. Denne dimensjonen er mest aktuell i min studie, som ser på læreres samarbeidskulturer ved to skoler.

For at lærersamarbeidet skal fungere godt, er det viktig at det på ledelsesnivå legges til rette for at styrkene ved lærersamarbeid kan vokse fram. Leithwood og Riehl (2003) framhever at skoleledere har mye å vinne på å fasilitere framvekst av felles mål, ha høye forventninger, støtte opp under gruppeaksept av mål og følge med på organisasjonens prestasjoner. Den enkelte medarbeider trenger også støtte, intellektuell stimulans og utfordring, og kan dra nytte av å ha modeller å måle egen prestasjon og atferd mot. Også andre forhold spiller inn, men det er en sterk vektlegging av at kollektive prosesser, utvikling og ikke minst framvekst av kollegialitet, en opplevelse av samhørighet og sosial inkludering (Keltchermans, 2006), er faktorer som kan bidra til å øke læreres opplevelse av egen mestring og indirekte virke til å bedre elevenes prestasjoner.

Selv om økt grad av lærersamarbeid kan kobles med økte elevprestasjoner (Goddard, Goddard & Tschannen-Moran, 2007; Moolenaar, Slegers & Daly, 2012), er det samtidig vanskelig å påvise noen direkte kobling mellom samarbeid og kvalitet i skolen. Vangrieken et al., (2015) viser i en reviewartikkel mange fordeler med lærersamarbeid, som samstemming av mål, mulighet for arbeidslettelse gjennom felles planlegging, økt tro på egne evner, intern støtte og læring, økt effektivitet og mer, men også mulige ulemper som opplevelse av autonomitap, konformitetspress, økt arbeidsbyrde, økt konfliktnivå og intern konkurranse. Det vil ofte være nødvendig for ledelsen ved en skole å finne en balanse mellom autonomi og samarbeid, slik at hver lærer og fagseksjon fungerer best mulig (Hargreaves, 1994; Leithwood & Riehl, 2003).

At lærersamarbeid kan ha positiv innvirkning på elevresultater, viser som nevnt en studie av Goddard et al., (2007, s. 878) som undersøker sammenhenger mellom grad av lærersamarbeid om læreplanforståelse, «instruction» og profesjonell utvikling ved 47 skoler og elevers resultater i lesing og matematikk. I denne sammenhengen kan en trekke en parallell til arbeidet med å utvikle detaljerte, omfattende vurderingskriterier og målformuleringer, som har likhetstrekk med å jobbe med læreplaner, som i utgangspunktet er overordnede, ofte abstrakt formulerte styringsinstrumenter. En studie som peker på et annet aspekt ved denne dimensjonen av lærersamarbeid, ble utført av Dobbie og Fryer (2011), som fant at kollektivt arbeid med å utvikle lokale læreplanmål ikke var statistisk relatert til økte elevprestasjoner. Samlet sett viser de ulike studiene hvordan forbindelsene mellom samarbeid og lærerprofesjonalitet, kvalitet i skolen og elevprestasjoner er kompleks.

I litteraturgjennomgangen har jeg vektlagt studier som kan gi informasjon om hva som kjennetegner en læringsfremmende vurderingspraksis fra ulike aktørperspektiver og studier som presenterer funn om sammenhenger mellom ulike aspekter ved VfL og læring. Dette valget gjør at andre forhold, som metode, effekt eller ideologisk/teoretisk bakgrunn for de nevnte studiene er mindre vektlagt.

3.7 Oppsummering og forskningsspørsmål

Presentasjonen fram til hit viser at den norske staten har satset aktivt på å fremme en mer læringsfremmende vurderingspraksis i skolen siden 2007 gjennom en rekke tiltak. Grunnlaget for disse er en omfattende internasjonal forskning innen feltet VfL (Black & Wiliam, 1998; 2009; Wiliam, 2011) og tilbakemeldinger (Hattie & Timperley, 2007; Shute, 2008) som dokumenterer at en VfL-orientert praksis gir bedring i læringsresultater når de implementeres. Forskingen er imidlertid hentet fra ulike land, fra ulike nivåer i utdanningssystemene og fra mange fagfelt, og kan derfor ikke gi noen direkte kunnskap om hvordan norske lærere løser den lovpålagte oppgaven å jobbe med læringsfremmende vurdering.

Det er viktig å understreke at studieobjektet i denne avhandlingen ikke er effekten av tilbakemeldinger, men utformingen av den, lærernes oppfatninger av god vurderingspraksis i skriftlig norsk og elevenes oppfatninger av lærernes skriftlige tilbakemeldingspraksis. Dette er alle vilkår som har betydelig potensial for å påvirke den faktiske effekten av tilbakemeldinger, men, som Black og Wiliam (2009) påpeker, er det knapt mulig å fastslå at én type tilbakemelding bidrar til læring, særlig fordi kontekstforhold, som f.eks. hvordan eleven har det følelsesmessig, læringsmotivasjon (Dweck, 2000; Pintrich, 2003) eller mestringsforventninger («self-efficacy», Bandura, 1977; 1997) kan ha stor betydning. Dette er noe som vil variere mellom elever og klasser, og fra tidspunkt til tidspunkt. Black og Wiliam (2009, s. 7) understreker dessuten at både lærere og elever har et kollektivt ansvar, og at de må ses som partnere i læringsprosessen hvis den skal ha større mulighet for å bidra til læring. I denne komplekse praksisen er altså tilbakemeldingen en viktig brikke.

Dermed følger det at det er viktig å bygge ut kunnskapen om lærerens tilbakemeldingspraksis på skriftlige tekster, og at tilbakemeldingene må ses i et større perspektiv, der både lærers handlinger og refleksjonene som ligger bak disse, tas med i betraktning. Videre er elevers oppfatninger av lærers tilbakemeldings- og vurderingspraksis viktig fordi en nøkkel til effektiv bruk av

tilbakemeldinger er betinget av at elevene ønsker å ta dem imot, et forhold som også kan påvirkes av skolekonteksten. I sum leder dette til det overordnede forskningsspørsmålet, som inkluderer aktørperspektiver, handlinger og oppfatninger: Hva kjennetegner norsklærerens skriftlige tilbakemeldings- og vurderingspraksis i to kollegier?

4 Forskningsdesign og metoder

Fenomenet som er undersøkt i avhandlingen, er læreres skriftlige vurderings- og tilbakemeldingspraksis til egne elever, og både læreres og elevers oppfatninger av denne praksisen. Ettersom fenomenet er situert i en spesifikk skolekontekst og involverer en rekke ulike aktører, er det også valgt et undersøkelsesdesign som benytter flere metoder, da disse frambringer ulike former for data.

Undersøkelsesdesignet er inspirert av Maxwells (2005) begrep «interaktivt design». Det kan framstilles som en forenklet modell (vedlegg 1), og viser hvordan det er en interaktiv sammenheng mellom formålet med studien, forskningsspørsmål, teorier og konseptuelt rammeverk, analysemåter og validitetsavveininger. Eksempelvis er statistiske analysemåter, som ble utviklet for denne studien spesielt, avhengige av akseptabel statistisk validitet for å kunne bidra til ny kunnskap. Arbeidet med utviklingen av studien innebærer på denne måten at ulike deler av designet kontinuerlig vurderes mot hverandre.

Samlet sett gir et slikt design et mer nyansert, detaljert og komplekst bilde av fenomenet enn én metode vil gjøre alene (Johnson & Christensen, 2014). Designet kan dessuten knyttes til begrepet triangulering (Creswell, 2009), der et hovedpoeng er at flere datakilder kan gi et mer helhetlig og nyansert innblikk i et fenomen enn bare en kilde – i dette tilfellet vil det si hvordan en VfL-praksis manifesterer seg i en skolekontekst der viktige aktører (elev, lærer, skoleledelse) er gitt en stemme, og der det i tillegg er frembrakt data som i noen grad er generaliserbare til andre skoler på samme nivå i utdanningssystemet.

Undersøkelsesdesignet er også inspirert av Lijparts (1971) idé om et komparativt design. Det vil si skoler jeg antok hadde ulikheter langs flere dimensjoner, ble forespurt om å delta for å øke diversitet i datagrunnlaget.

4.1 Deltakere og datagrunnlag i avhandlingen

Datagrunnlaget er hentet fra to videregående skoler i skoleåret 2014–2015. Tre norsklærere deltok fra hver skole med til sammen 12 klasser fra både studiespesialiserende og yrkesfaglige studieretninger. I tabell 1 presenteres en oversikt over lærere og klasser.

Tabell 1 Deltakere: lærere, klasser og antall elever

Skole A	Lærer/ pseudonym	Utdanningsnivå	Erfaring	Klasse	Antall elever
	Anne	Master	1 år	ST Vg3, ST Vg1, YF Vg2	78
	Berit	Bachelor	3 år	ST Vg3, YF Vg2	46
	Carina	Master	8 år	ST Vg2, YF Vg1	47
Skole B	Dina	Master	24 år	ST Vg3	31
	Eva	Master	5 år	ST Vg3, ST Vg2	58
	Frøy	Master	5 år	ST Vg3, ST Vg1	59

Skolene er av middels størrelse, med rundt 600 elever. Skole A hadde elever med høy grad av innvandrerbakgrunn, og den hadde fritt inntak, det vil si at det ikke var konkurranse om å komme inn på skolen og at en kan legge til grunn et betydelig svakere karaktergrunnlag for elevene ved denne skolen enn ved skole B. Elevene ved skole B kom inn med et snitt på 4,0 (høsten 2014) og hadde en mer homogen elevgruppe enn hva som synes tilfellet ved skole A.

Jeg ble gitt tilgang til hver av lærernes fagrom i læringsplattformen skolene brukte. Elevene lastet selv opp tekster i innleveringsmapper gjennom året, og de ble kommentert av læreren i læringsplattformen. I tillegg ble undervisningsmateriale, vurderingsskjemaer, uke-, måneds- og halvårsplaner og annet lagt inn av læreren gjennom året. Dette materialet lastet jeg ned til forskningsformål når det ble ansett å kunne kontekstualisere og/eller gi et mer nyansert bilde av lærernes vurderings- og tilbakemeldingspraksis. Jeg lastet ned og leste gjennom 1504 elevtekster med lærerkommentarer fra til sammen 319 elever. Det utgjorde cirka 75 % av alt materiale fra fagrommet og praktisk talt alle elevtekster. Som grunnlag for artikkel 1 ble 171 elevtekster fra en heldagsprøve valgt ut til analyse, og lærer hadde skrevet tilbakemeldinger til alle elevene. Det ble hentet data fra en klasse på studiespesialiserende studieretning per lærer.

I artikkel 2 ble resultater fra en spørreskjemaundersøkelse analysert (N = 213).

Spørreskjemaundersøkelsen ble besvart på papir og distribuert av elevenes egne lærere på slutten av skoleåret. Elevene ble informert om hva det innebar å delta muntlig. De fikk et spesialtilpasset skriv som presenterte forskningsprosjektet (som var utarbeidet i samråd med NSD), og informasjon om at det ikke ville få noen konsekvenser av noe slag om de ikke ville delta; ingen valgte å reservere seg mot deltakelse. Utvalget består dermed av de elevene som var til stede på skolen den aktuelle dagen. Det foreligger ikke grunn til å anta noen form for utvalgsskjevhet. Undersøkelsen var utformet for å kartlegge elevers oppfattelse av lærers vurderings- og tilbakemeldingspraksis i norskfaget.

I artikkel 3 er datagrunnlaget 6 intervjuer, slik det er gjort rede for i tabell 2. I tillegg danner også

dokumenter fra læringsplattformen, slik det er beskrevet over, og mer nøyaktig i artikkel 3, et grunnlag for å sammenholde informasjon fra intervjuene med lærernes faktiske praksis gjennom året. En representant fra ledelsen og en lærer i rollen som fagkoordinator, eller som hadde en ledende rolle i skolens fagsamarbeid i norskseksjonen, ble intervjuet noen måneder etter at øvrige data var samlet inn.

Tabell 2 Oversikt over datamaterialet for artikkel 3

Skole	Lærere	Type data	Varighet	Bakgrunns-data	Type data	Varighet	Type data	Varighet
A	Anne Berit Carina	Gruppeintervju, transskribert	0.45	Oppgaver, skjemaer, planer, ressurser for undervisningen	Intervju med ledelse, feltnotater	0.20	Intervju med lærer, feltnotater	0.20
B	Dina Eva Frøy	Gruppeintervju, transskribert	1.15		Intervju med ledelse, feltnotater	0.20	Intervju med lærer, feltnotater	0.20

4.2 Oversikt over analysemetoder i artiklene

I artikkel 1 er lærernes tilbakemeldinger analysert ved hjelp av dokumentanalyse i en kvalitativ studie. Den har samtidig en kvantitativ komponent ved at tilbakemeldingene er tallfestet. Dette er hovedsakelig gjort for å få oversikt over et omfattende materiale. Studien er kvalitativ fordi tolkningen av en lærers tilbakemelding er basert på faglig skjønn og er forankret i en kontekst der en koder trenger ekspertkunnskap for å kunne bruke et kodeskjema spesielt utviklet for denne studien. Designet er her betinget av et ønske om å operasjonalisere VFL i møte med en praktisk skolevirkelighet. Den teoretiske inngangen bidrar samtidig til at det er aspekter ved lærernes tilbakemeldinger som ikke er analysert, men kodesystemet trengte ikke noen restkategori for ytringer som ikke passet inn i systemet, noe som kan ses som en indikasjon på at kodesystemets begrepsvaliditet var akseptabel.

I artikkel 2 er elevenes innstillinger til lærers (skriftlige) tilbakemeldingspraksis undersøkt gjennom en spørreskjemaundersøkelse som ble analysert ved hjelp av strukturell likningsmodellering. Valget av metode var betinget av flere forhold. Ett var at hver lærer deltok med to klasser hver, til sammen tolv klasser, og det å intervjuer elever fra hver klasse i tilstrekkelig antall ville gitt svært mye intervjudata, som likevel bare med noen grad av sikkerhet kunne sies å være representative for klassene. Et annet forhold som lå til grunn for valget av metode, var et ønske om å utforske statistiske

relasjoner mellom lærers bruk av kriterier, instruksjoner og tilrettelegging for bruk (slik teorifeltet framhever at gir økt læring), og elevens opplevde nytte. Denne kunnskapen kunne ikke kommet fram i en kvalitativ undersøkelse, i hvert fall ikke på den måten som artikkel 2 gir innblikk i.

I artikkel 3 bidrar intervju som metode til å gi innblikk i hvilke begrunnelser, erfaringer og opplevelser som ligger til grunn for lærernes tilbakemeldingspraksis. Denne innsikten er viktig fordi den kan bidra til å informere praksisfeltet om forhold som kan lette innføringen av en lovpålagt praksis, som er generelt utformet og må finne uttrykk i det enkelte fagfeltet. Holdninger er dessuten viktige faktorer som har betydning for handling (Delvaux et al., 2013). For å øke innsikten i kontekstens betydning ble også representanter for ledelsen ved de to skolene intervjuet. En bred oversikt over datamaterialet finnes i vedlegg 2.

I det videre følger en utdypende redegjørelse for metodene som er anvendt, inkludert analysemåter og dataenes epistemologiske status.

4.3 Dokumentanalyse

Lærernes tilbakemeldinger til egne elever er undersøkt gjennom en form for dokumentanalyse, etter Scotts (2007) definisjon, der et dokument kjennetegnes av at det er et skriftlig produkt. Analysen av dokumentene er kvalitativ ettersom kodingen av materialet bygger på skjønn (Hsieh & Shannon, 2005, s. 1278) og målet er å finne mening i teksten. Analysen bygger dermed på en grunnantakelse om at tekster er meningsbærende, at meningen er stabil, slik at den kan finnes av uavhengige kodere, at teksten er skrevet med en hensikt og at denne hensikten kan spores. Selv om datamaterialet er summert opp og frekvens er løftet fram som en dimensjon, er ikke opptelling noe som i seg selv bærer mening i analysen. Opptelling er gjort for å få oversikt over materialet.

Teori om hvordan vurdering kan bidra til læring, danner det teoretiske utgangspunktet for analysene og er gjort rede for i kapittelet som definerer VfL (2.3). Antakelsen som ligger til grunn for kodingen, er at de teoretiske begrepene lar seg operasjonalisere. VfL er samtidig en teoretisk linse som snevrer inn lesingen av tilbakemeldingene, og dette perspektivet avgjør hva jeg har funnet. Jeg ser VfL og tilbakemeldingen innen dette systemet som et teoretisk konstrukt (Suppe, 1989). Konstruktet hevder, basert på omfattende empirisk forskning (Black & Wiliam, 1998), at hvis en tilbakemelding har visse kjennetegn, og hvis den brukes aktivt i videre arbeid med et fagstoff, øker sjansene for at læring finner sted. Dette konstruktet, som er generelt formulert, danner grunnlaget for kodesystemet som er brukt i studien som ligger til grunn for artikkel 1.

Det er mulig å se Hattie og Timperleys (2007) vilkår for en læringsfremmende tilbakemelding som en idealtipe (Weber, 2000, s. 178). Det vil si at en slik tilbakemeldingspraksis vanskelig kan oppfylles av noen. Svanes og Skagen (2016) framhever nettopp dette poenget i sin kritikk av hvordan VfL og tilbakemeldingspraksis spesielt holdes fram som en form for 'beste praksis'. Når dette poenget er understreket, er det likevel klart at innholdet i og bruken av lærerens tilbakemelding har betydning, noe omfattende forskning innen ulike fagfelt finner (Shute, 2008). Selv om en ikke kan forvente å finne en idealpraksis, kan retningslinjer bidra til å styre praksis i ønsket retning. Det vil si at elever får tilbakemeldinger som retter seg mot målet med opplæringen, at de får informasjon om hvordan de kan komme nærmere målet via konkrete forklaringer og at de selv er involvert i eget læringsarbeid (Hopfenbeck et al., 2013).

Kodene som ble utviklet, er navngitt som 1) mål, 2) instruks, 3) korrigerende og 4) relasjonsbyggende kommentar. Hver enhet er en semantisk, altså meningsbærende enhet, som i dette tilfellet vil si et utsagn som har en grense mot et annet utsagn, definert som at det andre utsagnet har et annet innhold. Kodebeskrivelsene er utledet fra en kombinasjon av skriveforskning (Parr & Timperley, 2010), læreplaner, skjema for måloppnåelse som brukes på norskeksamen på videregående skole, og fra fagkunnskap knyttet til norskfaget i skolen. Det er ikke noen hierarkisk forskjell mellom kodene.

Ettersom den kvalitative kodingen er bygd på faglig skjønn, kan en ikke forvente full reliabilitet mellom kodere. Bergman-Claeson (2003) framhever i en studie av læreres tilbakemeldinger til elever ved et svensk gymnas at det ikke er mulig å etablere koder som er gjensidig utelukkende, helt ut. Dette forholdet kan funderes i Rommetveits (1972) sosialpsykologiske tilnærming til språket som mangetydig, en grunnantakelse som også deles av hermeneutikken (Gadamer, 2008). Når kodene er utledet av et teoretisk ståsted, altså deduktivt, er det videre arbeidet med å utarbeide grenser for hver kode informert av «directed approach» (Hsieh & Shannon, 2005), det vil si at forsker har teoriinformerte forventninger i møtet med det empiriske grunnlaget for studien, og forventningene kan justeres i løpet av kodeprosessen.

De semantiske enhetene som ligger til grunn for kodene har ikke helt klare grenser. Hvor de går, avgjøres ved en skjønnsmessig vurdering, og dette er en utfordring når kodesystemet skal operasjonaliseres. Utfordringen ble søkt løst ved at de semantiske analyseenhetene ble avgrenset da jeg gjorde datasettet klart for koding. Slik ble det etablert gjensidig utelukkende kategorier som igjen kunne brukes til å reliabilitetsteste kodesystemet.

I en artikkel om innholdsanalyse, som diskuterer dette fenomenet fra et metodisk perspektiv, tar Strijbos et al. (2006) til orde for at i koding av språklige meningsenheter (forankret i analyse av kommunikasjonsenheter i digital kommunikasjon), skal tilfredsstillende reliabilitet kunne sikres av godt utarbeidete kodebeskrivelser. Forfatterne legger derimot liten vekt på fagkontekst, og kodingen er ikke tydelig informert av et teoretisk rammeverk. Hvis en legger til grunn at observasjoner alltid er ladet med teori (Hanson, 1958), er det vanskelig å se at en koder uten kjennskap til et spesialisert fagfelt skal kunne oppnå like god reliabilitetsskåre som en faglig ekspert.

Fenomenet som er studert i artikkel 1, er norsklæreres skriftlige tilbakemeldinger til egne elever i skriftlig norsk på videregående-skole-nivå. En lærer trenger faglig spesialisering for å kunne undervise på dette nivået, følgelig ble kodere med inngående kjennskap til fagfeltet valgt til å reliabilitetsteste kodesystemet. I en innledende runde ble kodesystemet prøvd ut mot to ekspertkodere, men uten at tilfredsstillende reliabilitet ble oppnådd. Deretter ble kodebeskrivelsene tydeliggjort og 20 prosent av materialet ble kodet på nytt av en annen ekspertkoder, denne gangen med tilfredsstillende resultat. Da oppnådde studiens kodesystem en Cohens kappa-skåre på ,76, som av Fleiss (1981) og av Landis og Koch (1977) regnes som henholdsvis «substantial» og «excellent». På det grunnlaget anses kodesystemet i studien som akseptabelt.

Cohens kappa brukes som en måleenhet på interrater-reliabilitet. Fordelen med å bruke en slik metode er at den fanger opp systematiske målefeil i kodesystemet, som ikke nødvendigvis vil avdekkes dersom man kun beregner det prosentmessige sammenfallet i koding mellom to kodere (Lombard et al., 2002). Cohens kappa tar også høyde for at tilfeldigheter kan påvirke sluttresultatet og justerer skåren etter det. Å regne ut tilfeldighetens betydning som feilkilde er en fast del av statistisk metode (Christophersen, 2013). Grensene for hva som kan regnes som akseptabel skåre, er basert på en faglig konsensus.

4.4 Strukturell likningsmodellering

I artikkel 2 bruker jeg strukturell likningsmodellering (SEM) som metode. SEM er en statistisk metode som brukes til å undersøke samsvaret mellom en teoretisk modell som uttrykker teoretiske antakelser og empiri. Datasettet er basert på en spørreskjemaundersøkelse der indikatorenes utforming og tematiske fokus er informert av teorigrunnlaget for VfL (Black & Wiliam, 2009). Det ble lagt vekt på indikatorer som tapper i hvilken grad elevene oppfatter at de kjenner målet med opplæringen, operasjonalisert gjennom indikatorer om kriterier som i sin tur kan knyttes til målet

med opplæringen. Videre er det indikatorer som måler i hvilken grad elever oppfatter at lærer legger til rette for bruk av tilbakemeldinger, som er en essensiell del av VfL. Til slutt er det indikatorer som måler i hvilken grad elevene oppfatter at de får tilbakemeldinger av læreren som inneholder en forklaring som kan bringe dem nærmere målet. Den empiriske forskningen som danner grunnlaget for VfL som teorikonstrukt, viser som nevnt at det er større sjanse for at vurderingspraksis bidrar til læring om disse forholdene er oppfylt (Black & Wiliam, 2009). I modellen er det lagt til grunn at elevene vil oppleve at lærers tilbakemeldingspraksis er mer nyttig hvis de oppfatter at lærer har en mer VfL-orientert praksis. Nytte er derfor avhengig variabel. Det gir til sammen en modell med tre uavhengige variabler – kriterier, bruk og forklaring, og en avhengig variabel – nytte.

De fire latente variablene er alle uttrykk for komplekse underliggende forhold, slik teorigjennomgangen har vist. Fra et epistemologisk ståsted er det ikke mulig å få svar på i hvilken grad elevene f.eks. oppfatter at tilbakemeldinger fra lærer på skriftlige arbeider blir brukt i undervisningen gjennom enkeltindikatorer. Det er fullt mulig å spørre for eksempel om tilbakemeldingene blir brukt av elev eller lærer, i hvilken grad det er lagt til rette for bruk eller i hvilken grad elevene oppfatter at de bruker tilbakemeldinger uten lærers innblanding, hva det innebærer at elevene bruker tilbakemeldingen og mer. Denne rekken av spørsmål illustrerer hvordan ulike indikatorer i sum kan bidra til å operasjonalisere aspekter ved det fenomenet en vil måle. Fenomenet uttrykkes dermed gjennom en latent variabel som består av fellesvariansen i de observerte variablene (Brown, 2006, s. 13). Den latente variabelen er altså en hypotetisk, bakenforliggende variabel, og den er her utledet av VfL-teori.

Det ble gjennomført en konfirmerende faktoranalyse av relasjonene mellom indikatorer og latent variabel. De fire faktorene ble i den innledende faktoranalysen vurdert, og basert på verdiene til hver observerte variabel ble enkeltindikatorer med lav faktorladning tatt ut av modellen. Det er ikke uproblematisk å ta ut enkeltindikatorer av en modell, men basert på teoretisk og faglig grunnlag⁴ er det samtidig rimelig å justere en modell som en del av en valideringsprosess. Slik får konfirmerende faktoranalyse et eksplorerende preg (Christophersen, 2013).

For å ytterligere undersøke forholdet mellom indikatorene og de latente variablene ble det gjennomført en konfirmerende faktoranalyse, der de latente variablene med tilhørende indikatorer ble undersøkt to og to. Resultatene er gjengitt detaljert i artikkel 2. Hovedfunnene var at de

4 Informasjon fra professor Frank Goldhammer (Deutsches Institut für Internationale Pädagogische Forschung), 16. desember, 2016.

latente variablene – forklaring, bruk og nytte – var velspesifiserte, mens kriterier hadde svakere føyningsmål. En slutning av disse analysene er at det må tas forbehold når sammenhenger mellom kriterier og de øvrige variablene blir presentert, og instrumentet vil på dette punktet måtte utvikles i en videre forskningsprosess.

Ettersom dataene er samlet inn ved to skoler, er de sterkt klynget. Det ble derfor gjennomført en intraklassekorrelasjonstest av hver indikator. Resultatene som er gjengitt i artikkel 2, viser at fem indikatorer (av 11) har verdier som tilsier at verdiene er påvirket av konteksten, altså skole eller lærer. Da dataene er samlet inn ved kun to skoler, er flernivåanalyse krevende, ettersom dette vanligvis gjøres med 30 nivå-2-enheter, og sjelden med færre enn 10 (Christophersen, 2013). En naturlig slutning må være at i framtidig forskning vil en oppfølgingsstudie som baserer seg på empiri fra flere skoler, bidra til å styrke den ytre validiteten til min studie.

I modellen er stikoeffisienter numeriske uttrykk for de teoretisk anlagte kausalforklaringene som teorien angir. Det er likevel ikke mulig å si hvilken kausalretning som faktisk kommer til uttrykk i modellen fordi kausaliteten er teoretisk begrunnet, ikke statistisk. Statistisk analyse kan i hovedsak bare tallfeste styrken på relasjonen. Styrken på assosiasjonene mellom de latente variablene varierer, og resultatene diskuteres i artikkel 2. Generelt regnes en estimert effekt på 0,1 som lav, 0,3 som middels og 0,6 som høy (Kline, 2005). Det er verdt å merke seg at samlet predikert varians til avhengig variabel i modellen er høy (0,86), og at dette kan være uttrykk for noen grad av lineær avhengighet i modellen. Høy predikert varians indikerer at en eller flere uavhengige variabler korrelerer høyt med avhengig variabel, det vil si at de er delvis samme latente variabel. Det er rimelig å anta, på teoretisk grunnlag, at det er en nær sammenheng mellom de latente variablene. De utgjør samlet sett en sum av faktorer som VFL-teorien framholder bør være til stede for å øke mulighetene for at vurdering skal bidra til læring (Black & Wiliam, 2009). Fravær av en faktor vil bidra til at modellens teoretiske basis, uttrykt gjennom de latente variablene, er svakere operasjonalisert.

Samsvaret mellom den teoretiske modellen og datamaterialet er uttrykt gjennom ulike føyningsmål som samlet sett angir hvilken grad av avvik det er mellom den hypotetiske modellen og den observerte. Det er benyttet ulike mål for hvor godt hypotetisk modell passer med observerte verdier. En kjikvadrattest bygger på sammenheng mellom antall frihetsgrader og kjikvadratverdien og p-kjiverdien ($p > 0,000$) regnes som for lav, men det er ikke nok til å forkaste modellen. «Root mean square error approximation» (RMSEA) tar hensyn til at ingen modell vil kunne passe perfekt.

Verdier under 0,05 anses som tilfredsstillende, og mellom 0,05 og 0,08 som brukbare, mens høyere verdier anses som uegnet (Kline, 2005). Her viser verdiene brukbar skåre (mindre enn 0,08). «Tucker Lewis-index» (TLI) sammenlikner forholdet mellom kjikvadratverdiene i hypotetisk modell med observert. TLI straffer komplekse modeller, og bør ha verdier over 0,9. «Comparative fit index» (CFI) sammenlikner modellen med den observerte når alle residualverdier er satt til 0. Her bør verdiene være større enn 0,95 (Kline, 2005). I begge tilfeller er vilkårene for brukbar føyning oppfylt for den estimerte modellen.

4.4.1 Eksplorerende faktoranalyse

I forlengelsen av spørsmålet om en mulig lineær avhengighet i modellen, ble det gjennomført en eksplorerende faktoranalyse for å utforske sammenhengene mellom antall indikatorer og antall underliggende faktorer. Resultater av analysene viser at indikatorene i hovedsak fordeler seg på tre underliggende faktorer. Teoretisk sett kan dette forklares ved at en læringsfremmende vurderingspraksis, slik den er forstått og operasjonalisert i denne avhandlingen, består av tre forhold som har et potensial når de opptrer samlet, nemlig at lærer og elev har en delt forståelse for målet med opplæringen, at elevene får tilbakemeldinger og at disse brukes i videre læringsarbeid. Det er knapt mulig å ta ut en faktor uten at helheten i modellen svekkes, slik det er argumentert for ovenfor.

Tabell 3 viser at tre faktorer forklarer 54 % av total varians i indikatorene som danner grunnlaget for modellen ved ortogonal rotasjon, med «varimax» metode, som søker å spre varians på flest mulig faktorer, noe som gjør det enklere å tolke innledende resultater av analysene (Field, 2013). Fordelingen på tre faktorer viser seg også i den grafiske framstillingen i et «scree-plot» av variansfordeling.

Tabell 3 Resultater av eksplorerende faktoranalyse

Factor	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,109	44,433	44,433	6,723	42,016	42,016	3,642	22,760	22,760
2	1,767	11,046	55,479	1,351	8,444	50,460	2,734	17,090	39,851
3	1,167	7,292	62,771	,595	3,720	54,180	2,293	14,329	54,180
4	,962	6,013	68,784						
5	,836	5,222	74,006						
6	,686	4,290	78,296						
7	,541	3,379	81,675						
8	,503	3,144	84,819						
9	,456	2,851	87,670						
10	,390	2,439	90,109						
11	,373	2,333	92,442						
12	,328	2,053	94,495						

4.5 Intervju

I artikkel 3 ble intervju benyttet som hovedmetode i datainnhenting. Det ble gjennomført to typer intervjuer: semistrukturert gruppeintervju og intervjuer med en representant for ledelsen og en for lærerne ved hver skole. Til sammen ble det gjennomført seks intervjuer.

Epistemologisk sett kan intervjuer benyttes til å få informasjon om respondentenes oppfatninger, erfaringer og holdninger slik de kommer til uttrykk i intervjusituasjonen. Säljö (1997) framholder at intervjudata mer enn noe annet er uttrykk for hvordan respondenter velger å opptre i den diskursive situasjonen. I dette tilfellet er lærerne intervjuet som representanter for en rolle. Det er ingen spørsmål som retter seg mot dem som privatpersoner, og en kan legge til grunn at lærerne kjenner til og utagerer en måte å snakke om vurdering og tilbakemelding på som er institusjonalisert. Datagrunnlaget er først og fremst en slik samtalediskurs som ikke har noen direkte lenke til lærernes faktiske erfaringer fordi en med Säljö (1997) kan argumentere for at bevisstheten om disse erfaringene nettopp kommer fram i den faglige samtalen som oppstår i intervjusituasjonen, og at den ikke er uttrykk for ferdig uttenkte opplevelser og erfaringer. For å utdype dette poenget videre kan en legge til grunn at fenomenet som er i sentrum for intervjusituasjonen aktiverer fagkunnskaper, pedagogiske kunnskaper, kunnskaper om elevene og skolekonteksten, følelser overfor elevene, bevissthet om hva statlige retningslinjer sier, føringer fra kollegium og skoleledelse, sosiale hensyn, egen rolleforståelse, forståelse for selve intervjusituasjonen, der spørsmålet om hva som er relevant å si spiller en sentral rolle, og ikke minst hvilke ideer lærerne hadde om forskeren, dennes hensikter, opptreden og kunnskaper. Det er også mulig å anta at intervjuobjektene uttaler seg slik at de ikke skal tape anseelse.

For å eksemplifisere poenget kan det påpekes at ingen av lærerne på noe punkt uttalte seg negativt om enkeltelever, selv om en ut fra både innsamlet bakgrunnsdata og lærernes utsagn i intervjusituasjonen kan se at en god del elever ikke gjorde som læreren ba dem om. Det er nærliggende å tenke at denne typen elevatferd vil kunne avstedkomme en negativ følelsesmessig respons hos lærer, som altså har lest elevens tekst, brukt tid og ressurser på å gi tilbakemelding, og opplever at denne innsatsen neglisjeres av eleven. Når det ikke kom fram noen negative utsagn om enkeltelever, kan det henge sammen med at lærerne ikke hadde slike følelser. Men det kan også være at de ikke ville gi uttrykk for dem fordi de opplevde at det enten ikke var relevant, ikke var akseptabelt eller at det ikke ble spurt om eksplisitt. Eksempelet illustrerer altså at det lærerne sier i intervjusituasjonen, kan ses som sosialt akseptable holdninger, erfaringer og oppfatninger, og at det

først og fremst er de sidene ved arbeidet deres som tematiseres i intervjuguiden som kommer fram i intervjusituasjonen. En kan i forlengelsen, på bakgrunn av den argumentasjonen som er presentert her, også si at datamaterialet i noen grad speiler lærernes rolleforståelse.

Det er ellers mange intervjustudier (f.eks. Matre & Solheim, 2014) som ser respondenters utsagn i en intervjusituasjon som uttrykk for deres verdier, holdninger og meninger, og at det kvalitative intervjuets styrke ligger i å nettopp la respondentene komme til orde med sine erfaringer og sette ord på sentrale aspekter av sin opplevelse av intervjuets tematikk. I tillegg kan eventuelle uklårheter tydeliggjøres med oppfølgingsspørsmål og aspekter ved den tematikken som intervjuguiden legger føringer for, kan utdypes på måter som ikke var intendert av forsker. Det er med Säljö (1997) mulig å tenke seg at spørsmålene fra intervjuguiden dessuten speilet andre forhold eller fenomener for lærerne enn det som var intendert av meg som forsker. Et slikt ståsted innebærer ikke nødvendigvis noen motsetning til Säljös (1997) kritikk, spørsmålet er i utgangspunktet mer hvilken status en tillegger intervjudataene. Når lærere snakker om sin egen vurderingspraksis, er det et fenomen som i denne konteksten antas å være kjent for alle involverte parter. Jeg velger derfor å utdype hvordan en felles forståelse ble etterstrebet, og ser Säljös (1997) kritikk mot å oppfatte intervjudata fenomenografisk som en advarsel mot å anta at utsagn om erfaringer nødvendigvis er å sette ord på reelle erfaringer.

Tre hensyn som har innvirkning på hvordan intervjudataene ble til, er forskers rolle, rammen for intervjuet og utformingen av intervjuguiden og jeg skal kort gjøre rede for disse. Säljö (1997) viser til et eksempel på en situasjon der en forsker, i en sosialantropologisk sammenheng, intervjuer en person om navn på fugler og oppdager etter hvert at forsker og intervjuobjekt ser fuglene med helt ulike perspektiver og tillegger dem en vidt forskjellig ontologisk status. Min rolle som forsker i intervjusituasjonen var at jeg hadde jobbet som lektor i samme skoletype (videregående skole) og i samme fag, altså norskfaget, sammenhengende i mer enn 13 år. Dette bidro til jeg kom nærmere et innsiderperspektiv (Kvernbekk, 2005) i intervjusituasjonen som økte sjansene for at nyanser som kunne trenge oppfølging eller utdyping, ble fulgt opp med ekstra spørsmål. Min mulighet til å innta en innsiderposisjon ble svekket av at jeg kom til skolene i rollen som forsker; jeg intervjuet dem ikke som kollega. En utfordring med innsiderposisjonen kan være at egne oppfatninger om fenomenet som er i fokus kunne gjort at jeg uintentert søkte å styre samtalen etter egne preferanser heller enn å fange opp respondentenes. Kvernbekk (2005) inntar en radikal innstilling til dette spørsmålet når hun problematiserer om det i hele tatt er mulig å innta et innsiderperspektiv for andre

enn seg selv, og dette perspektivet kan tjene som en advarsel mot å anta at man som forsker har kunnskaper om kontekst eller deltakere som det ikke er dekning for. Min innsidernære posisjon var basert på erfaringer som lå mindre enn to år tilbake i tid da intervjuene ble gjennomført, men fra en annen skolekontekst.

Intervjuene foregikk over et tidsrom på mellom 45 minutter (skole B) og 75 minutter (skole A). Ingen av intervjuene var preget av tidspress, og det ble satt av tid til å snakke om forhold ved vurderings- og tilbakemeldingspraksis som intervjuguiden ikke hadde tatt høyde for. Ved begge skolene ønsket lærere å legge til informasjon.

For å undersøke aspekter ved vurderingspraksis som kan knyttes til VfL, ble det utviklet en intervjuguide (vedlagt artikkel 3). Guiden er utformet med spørsmål som søker å kartlegge ulike aspekter ved en VfL-orientert praksis. Den stiller spørsmål om ulike sider ved bruk av tilbakemeldinger, bruk av kriterier, hensikten med vurdering og mer. Intervjuguiden ble pilotert i fire gruppeintervjuer med 18 lærere høsten 2013, og utformingen ble dessuten diskutert med erfarne forskere. Lærerne ble ikke spurt direkte om de satset på VfL eller jobbet etter VfL-prinsipper, ettersom jeg tok det for gitt at de visste at opplæringsloven krever denne typen vurderingspraksis og at de derfor vanskelig kunne si noe annet enn at de gjorde det. På bakgrunn av de definisjonene som er brukt i denne studien, skal intervjuguidens begrepsvaliditet være holdbar.

Intervjuet som metode kan i denne studien knyttes til Kvale og Brinkmanns (2010) gruvearbeidermetafor der hensikten er å frambringe utdypende kunnskap om et avgrenset felt. Intervjuguidens operasjonaliseringer springer ut av en kombinasjon av forskerens erfaringer og fagkunnskaper innen feltet, og forskning om hvordan vurdering kan bidra til læring, i tillegg til føringer fra norske myndigheter (Udir, 2010b). Den informasjonen som kommer fram, er altså et uttrykk for en kombinasjon av forsknings- og erfaringsbaserte spørsmål og gir svar på hvordan lærerne, som en respons på denne operasjonaliseringen, snakket om erfaringer med og oppfatninger av sin egen og skolens vurderings- og tilbakemeldingspraksis.

I artikkel 3 er intervjuet benyttet som metode for å komplementere data som er utviklet fra andre deler av studien som samlet sett utgjør avhandlingen. Det vil si at hva slags tilbakemeldinger lærerne skrev til egne elever i skriftlig norsk, som nevnt, ble analysert ved hjelp av dokumentanalyse, og elevenes oppfatninger av aspekter ved lærers skriftlige tilbakemeldingspraksis (som kan knyttes til VfL som konstrukt) ble undersøkt via en spørreskjemaundersøkelse som ble analysert med SEM. Disse to undersøkelsene ga samlet sett et bilde av hva lærerne gjorde og

hvordan elevene deres oppfattet denne praksisen, men undersøkelsene viste ikke hvorfor lærerne gjorde som de gjorde. De ga heller ikke innblikk i hvilke utfordringer og muligheter lærerne så i sammenhengen mellom egen vurderingspraksis og elevenes læringsprosesser.

Analysene av intervjuene ble gjennomført ved at de først ble transkribert, og å gjøre denne jobben var et ledd i prosessen med å gjøre seg kjent med datamaterialet. De transkriberte intervjuene ble så lastet inn i NVivo11, et program som gjør det enklere å analysere større mengder tekstdata (og annen type kvalitativ data) ved at tekstelementer som er kodet innenfor en kategori (node), enkelt kan leses under ett. Det ble gjennomført en tematisk analyse av transkripsjonene (Braun & Clarke, 2006) og i prosessen ble en rekke tematiske koder identifisert og definert, slik det framgår av artikkel 3. Et eksempel på hvordan koding er gjennomført, finnes i vedlegg 3.

Som tiltak for å sikre reliabilitet ble intervjuene med tilhørende koder også lest av to forskere med kjennskap til feltet og kodingen jeg hadde gjennomført ble vurdert etter «face value»-validitetskriterier (Schmidt & Rader, 1999), som vil si at hvis kodingen framsto som umiddelbart fornuftig, var det en indikator på at så var tilfelle. I tråd med Braun og Clarke (2006) ble det utarbeidet relativt åpne kategorier som er definert av det tematiske innholdet.

4.5.1 Intervju i kombinasjon med andre data

Det forskningsmetodologiske designet til artikkel 3 søkte å frambringe kunnskap som i noen grad var egnet til å gi innsikt i lærernes faktiske praksis. Dette ble gjort ved at jeg som forsker hadde innblikk i lærernes skriftlige praksis, slik den kom til uttrykk i den elektroniske læringsplattformen. Innsikten i lærernes praksis var på denne måten i noen grad til stede i intervjusituasjonen, og det lærerne sa at de gjorde, var også etterprøvbart ved at bakgrunnsdatamaterialet kunne undersøkes på nytt. Dette datamaterialet ble i hovedsak generert ved at jeg leste og kvantifiserte elevers deltakelse i skriveprosesser. I tillegg sammenliknet jeg vurderingskriterier på tvers av klasser og skoler, og det samme gjorde jeg med oppgavesett, slik at graden av samsvar kunne kartlegges.

Det må samtidig understrekes at det er et begrenset innblikk i lærernes praksis som kommer fram gjennom kombinasjonen av intervjuene og bakgrunnsdata ettersom observasjoner fra klasserommet ikke er en del av studien, og at en helhetlig oversikt over et fenomen som tilbakemeldingspraksis heller ikke lar seg undersøke fullt ut sett fra et vitenskapsfilosofisk ståsted (Suppe, 1989).

4.6 Validitet og reliabilitet

Gitt at det å operasjonalisere en tilbakemeldingspraksis er en svært kompleks oppgave, er det i denne avhandlingen valgt å gå ut fra en vid definisjon av VfL. Dette valget er begrunnet av at en smal definisjon, slik Svanes og Skagen (2016) argumenterer for, lett vil føre til at forskningen finner at lærerens praksis kommer til kort, og at denne vurderingen vil bygge på en teoretisk forståelse som ikke tar hensyn til kontekst og kompleksitet, en tilnærming også Black og Wiliam (2009) advarer mot.

Jeg utviklet derfor et kodesystem for å sikre en valid og reliabel behandling av datasettet. Andre forskere har benyttet Hattie og Timperleys (2007) feedback-system, som det ble gjort rede for i kapittel 3. De framhever som nevnt at en tilbakemelding som bidrar til læring, ofte svarer på de tre spørsmålene «Where am I going?/How am I going?/Where to next?», som kan være rettet mot fire nivåer: oppgaven, prosessen, selvregulering og selvet. Forskere som Gamlem (2014) og Bueie (2016) har benyttet Hattie og Timperleys (2007) system, men jeg mener det kan være et problem i møte med lærers konkrete tilbakemeldinger. Eksempel: Skal man kode en tilbakemelding fra en lærer, så vil et utsagn fra Eva, sitert fra artikkel 1, kunne se ut som dette: «Husk at du alltid må ha med navn på forfatter og tittel i innledningen.» Hvis en ser dette utsagnet opp mot de tre spørsmålene, kan en argumentere for at det fyller alle funksjoner på likt: Tilbakemeldingen sier noe om prestasjonen sammenliknet med målet, at eleven har glemt to viktige deler av en innledning til en analyse av en skjønnlitterær tekst, og sier samtidig hva som skal være med, nemlig forfatternavn og tittel. Tilbakemeldingen svarer også på det siste spørsmålet: Neste gang hun skal skrive en analyse av en tekst, skjønnlitterær eller saktekst, gjør eleven lurt i å huske å ta med forfatternavn og tittel på teksten som skal analyseres. Eksempelet illustrerer at et kodesystem basert på Hattie og Timperleys (2007) anbefalinger kan være krevende å operasjonalisere ettersom et kodesystem helst skal ha gjensidig utelukkende kategorier. Det er også verdt å merke seg at Hattie og Timperleys (2007) spørsmål ikke er kontekstavhengige, men i prinsippet kan gjelde som retningslinjer for tilbakemeldinger i alle fag og på alle nivåer, fra første klasse i grunnskolen til masternivå eller høyere. Hattie og Timperley (2007, s. 103) peker dessuten selv på at grensene mellom spørsmålene, når de skal operasjonaliseres, er «fuzzy».

Ettersom VfL-kodesystemet som ble utviklet, oppnådde akseptable Cohens kappa-skårer, bør ikke begrepsvaliditet representere noen alvorlig trussel for analysene i artikkel 1. Imidlertid bør spørsmålet om artikkelens og avhandlingens økologiske validitet diskuteres. I artikkel 1 er den

analyserte skrivesituasjonen en eksamensliknende heldagsprøve. Spørsmålet er i hvilken grad skriveprøven kan ses som uttrykk for en mer helhetlig praksis. Prøven består av et sett oppgaver og tekster som er organisert slik at alle må svare på en korttekstoppgave, som har et analytisk preg, og så velge én av de andre oppgavene og skrive en lengre tekst, der det å benytte seg av en eller flere av de vedlagte tekstene i oppgaveheftet er et eksplisitt krav (formulert i de tilhørende vurderingskriteriene). Selv om elevene valgte ulike sjangrer og en lærer (Carina) hadde organisert heldagsprøven som en prosesskriveøkt, så er heldagsprøven bare en form for skriving i skolen. Det må derfor legges til grunn at artikkel 1, trass i et kritisk case-design (Patton, 2014), først og fremst gir informasjon om lærernes tilbakemeldinger på én type skrivesituasjon. Det er mulig at lærerne ga andre former for tilbakemeldinger i andre skrivesituasjoner. Datamaterialet kunne gi svar på slike spørsmål, men det ligger et omfattende arbeid til grunn for analyse av de 171 elevtekstene med lærers tilbakemeldinger, og stipendperiodens begrensninger gjør det vanskelig å gjøre bredere anlagte analyser.

For analysene i artikkel 2 er det relevant å anvende Shadish, Cook og Campells (2002) validitetssystem for kvantitative undersøkelser som opererer med fire typer validitet. Statistisk validitet betinger tilstrekkelig antall respondenter, noe som er tilfellet i min studie: Christophersen (2013) peker på at som en tommelfingerregel krever bivariat analyse minst 50 respondenter og at en ekstra latent variabel vil kreve 15 ekstra respondenter. Det vil si at en fire-faktormodell tilsvarende den som presenteres i artikkel 2 vil kreve minst 90 respondenter, og antallet respondenter var 213. Et annet validitetskrav er at estimerte sammenhenger mellom avhengig og uavhengige variabler er systematiske, altså ikke basert på tilfeldighet eller samplingsfeil. For modellen som helhet synes dette å være tilfelle, men delen av modellen som knytter seg til konstruert kriterier synes å ha svakere validitet. Funn som beskriver dette konstruets sammenheng med øvrige variabler, må ses som mer usikre enn de øvrige. Kriterier har likevel deskriptive verdier som gjør at indikatorene kan brukes til å si noe om elevenes oppfattelse av om lærer har og bruker kriterier.

De latente variablene er uttrykk for komplekse begreper, og en kunne anta at det også ville vært behov for mange indikatorer for å fange opp bredden i begrepene. I en studie av Gogol et al. (2014) fant forskerne i en sammenlikning av måleinstrumenter at korrelasjonsmønstre i modeller med få (ett til tre) og flere indikatorer ga de samme korrelasjonsmønstrene. Dette kan tas til inntekt for at en modell med få indikatorer kan fungere for å måle også komplekse begreper, slik tilfellet er i den estimerte modellen.

Modellen jeg benytter, er ikke hentet fra veletablerte måleinstrumenter. Det burde de ideelt sett vært, men det var ikke mulig å finne forskning som testet ut akkurat denne teoretiske konstruksjonen, selv om teorifeltet har vært utviklet over lang tid. Dette kan ses som en svakhet ved undersøkelsen, ettersom spørsmålene som er stilt ikke er reliabilitetstestet på forhånd. Gitt det overordnede temaet for avhandlingen er dette samtidig en måte å utvikle ny kunnskap på, og føyningsmålene er, som vist i artikkel 2, akseptable. Validitetstruslene bør imøtegås ved ytterligere testing av instrumentene i framtidig forskning.

Indre validitet handler om styrken på de teoretisk anlagte kausalsammenhengene. Hvorvidt kausaliteten er slik som antatt, blir diskutert i artikkelen. Det fins ulike måter å måle indre reliabilitet i instrumentet på. I dette arbeidet er Cronbachs alfa benyttet, ettersom det er en svært vanlig måleenhet å bruke for å teste dette aspektet ved de latente variablene, og fordi måleenheten Cronbachs alfa viser om de indikatorene som er ment å uttrykke ulike aspekter ved samme latente variabel, har oppnådd et svarmønster hos respondentene som tilsier at det finnes en slik indre konsistens. Estimatene som er oppgitt, er en funksjon av korrelasjonen mellom de enkelte spørsmålene og mellom antall spørsmål som er med. Generelt sett regnes skårer på 0,9 som svært høye, og skårer over 0,7 som tilstrekkelige (Kline, 2005, s. 59). Verdier nær 1 vil samtidig innebære at indikatorene uttrykker samme forhold, statistisk sett. Alle verdiene i de indikatorene som er benyttet i modellen, er innenfor verdiene 0,75–0,89. Ettersom det ikke finnes noen egen test for begrepsvaliditet, og teoretiske begreper og sammenhenger ikke lar seg observere direkte, er de skårene som er oppgitt her, indikasjoner på at modellen og de latente variablene er operasjonalisert på en måte som tilfredsstillende krav til begrepsvaliditet. Noe av begrepsvaliditeten er også fanget opp av «face validity» (Bryman, 2012, s. 53), det vil si hvor godt indikatorene gir et umiddelbart inntrykk av å fange opp sentrale sider ved begrepet de er ment å måle, basert på skjønnerkunnskap. Alle indikatorene er utviklet i samarbeid med erfarne forskere som arbeider innenfor samme felt.

Andre validitetstrusler er at det er mulig å tenke seg at lærerne, gjennom å delta i studien, har justert sin praksis, så den er mer i tråd med gjeldende retningslinjer, og at det bildet som tegnes i denne studien ikke er representativt for gjeldende praksis hos disse lærerne, fenomener som av Lund (2002) betegnes som evalueringsforståelse. Det vil si at lærerne spiller en rolle de mener stiller dem i et godt lys. Hypotese-gjetting vil si at de gjetter hva som er formålet med datainnsamlingen og oppfører seg annerledes enn vanlig som en konsekvens. Dette er rett nok spørsmål om begrepsvaliditet, men de kan også påvirke spørsmålet om hvor typisk den observerte praksisen er

for disse lærerne. Noe som taler imot denne validitetstrusselen, er at det er samlet data fra et helt år, og hver lærer har skrevet tilbakemeldinger til rundt 300 elevtekster gjennom denne perioden. Det kan også presiseres at ingen av lærerne kunne vite hvilken av vurderingssituasjonene som skulle velges ut for analyse. Det er heller ingen av lærerne som i utgangspunktet meldte seg til å delta i studien – de ble med etter forespørsel.

Bortfall av respondenter representerer en validitetstrussel. 319 elever kunne deltatt i spørreundersøkelsen (tallet er litt usikkert; det er nok litt lavere på grunn av elever som sluttet i løpet av året, noe som hendte ved begge skoler), men 221 deltok, og av disse ble 213 beholdt i datasettet som ble analysert. Data ble samlet inn på begge skoler mot slutten av skoleåret, i mai, og de som ikke har svart, var borte den aktuelle dagen. Det er imidlertid ikke noen grunn til å anta at andre grunner enn at tilfeldig og usystematisk fravær er årsak til bortfallet.

Antallet respondenter og geografiske forhold begrenser muligheten for generaliserbarhet til en større populasjon (Shadish, Cook & Campbell, 2002), men funn som er gjort kan følges opp av videre forskning med tanke på å finne støtte for studiens ytre validitet. Intraklassekorrelasjonsskårene som presenteres i artikkel 2 viser at fem indikatorer har verdier over 0,05 og at verdiene dermed i noen grad er påvirket av hvilken nivå 2-enhet (skole) elevene gikk på. Det må følgelig vises varsomhet med å generalisere funnene til en total populasjon av elever i videregående skole, men det er samtidig rimelig å anta at elever ved liknende skoleslag kan ha samme svarmønstre.

4.7 Forskningsetikk

4.7.1 Ivaretagelse av informantene – deltakelse og læreren som «den andre»

Å skrive tilbakemeldinger på elevtekster i videregående skole er en praksis som kan oppfattes som nesten privat siden den foregår i et lukket rom der lærer gir tilbakemelding til én spesifikk elev. Den er ikke formulert med tanke på at den skal kunne deles med hele klassen, der den peker på både styrker og svakheter i elevenes tekster, gir råd og begrunner karakterer. Selv om tilbakemeldinger kan bli lest av flere enn eleven (medelever, foresatte, kanskje kollegaer), er det likevel oftest eleven selv som velger om tilbakemeldingen skal leses av andre eller ikke, i hvert fall når de går i videregående skole. Det er grunn til å tro at dette også preger den som skriver. At jeg som forsker kommer inn i denne lærer-elev-sfæren, er derfor på flere måter som å se inn i et univers som ikke er ment for alles øyne.

Da jeg søkte informanter, tenkte jeg at det å få lærere til å stille opp i studien, når de verken kjenner forskeren eller prosjektet, kunne bli vanskelig, særlig når det å finne to ulike skolekontekster var viktig. Jeg brukte derfor kontaktene mine fra årene i videregående skole for å få tilgang til potensielle informanter – mennesker jeg ikke kjente, men som tidligere kollegaer kjente. Da jeg hadde fått foreslått noen navn, henvendte jeg meg til skolens ledelse og så videre derfra til informantene, hvor jeg i en e-post henvendte meg direkte til dem (ikke i form av masseutsendt e-post til hele norskavdelingen) og hilste fra kontaktene mine. Dette har helt sikkert bidratt til at jeg temmelig raskt fikk tak i de informantene jeg trengte. Spørsmålet er i hvilken grad de har følt seg presset til å delta av den direkte henvendelsen. På den ene siden involveres en kollega og kanskje venn i avveiningen om en skal si ja eller nei. Dette kan innkodes som et press. På den andre siden tok jeg kontakt via e-post, et medium som ga mer distanse, lærerne slapp å si nei ansikt til ansikt eller på telefon, og en del unnlot å svare, uten at jeg la noe videre press på dem. Jeg tør derfor, så langt, å anta at informantene deltok på frivillig basis, selv om jeg strengt tatt ikke helt ut kjenner motivene deres. De ble også gitt informasjon om muligheten til å trekke seg via informasjonsskrivet som er godkjent av NSD.

I den videre datainnsamlingsprosessen var det relevant å ha med seg et relasjonsetisk perspektiv, der en sentral målsetting hele veien var å ivareta informantene som «den andre» (Levinas, 2004). Ved at jeg møtte lærerne og at de viste meg tillit, ville jeg, ifølge den danske teologen og filosofen Løgstrup (1956), samtidig stå overfor det som kalles en *etisk fordring*, som her kan forstås som at jeg får en viss makt over dem. Dette gjelder nesten alle menneskemøter, og makten er forstått som at en gjør seg sårbar overfor andre når en viser dem tillit. Som stipendiat var jeg i en posisjon der jeg antakeligvis kunne tillegges ekspertinnsikt (om hva som er god praksis), som jeg igjen kunne bruke til å kritisere lærerne og deres måter å gi tilbakemeldinger på. Jeg kunne også stille dem i et uheldig lys overfor kollegaer, ledelse og det miljøet de er en del av. Informantene vil kunne gjenkjenne seg selv dersom jeg bruker deres utsagn i min avhandling, noe jeg også gjør. Dermed vil de også ha tilgang til min «dom» over noe som ellers hører til deres personlige lærerpraksis. En sentral oppgave var derfor å opptre slik at informantene opplevde at de kunne stole på meg og i forlengelsen av det føle seg trygge. En måte å få dette til på, var å være nøye med å opptre i tråd med etiske retningslinjer gitt av NESH (etikkom.no, s. 11). Under punkt 6, «Krav om respekt for integritet, frihet og medbestemmelse», står det at: «Individet har behov for selv å kunne påvirke hva som skjer med det på viktige områder av livet. Å være underlagt andres observasjon og tolkning kan oppleves som nedverdiggende. Aktsomhet er særlig påkrevet (...) når individet aktivt bistår med å skaffe

informasjon til veie, for eksempel ved å la seg observere eller intervju» og «når individet kan identifiseres, for eksempel når personer og miljøer kan gjenkjennes i forskningsrapportene.» Dette siste punktet vil jeg komme nærmere inn på når jeg diskuterer anonymiseringen av informantene.

4.7.2 Sinnelagsetikk og konsekvenser av forskningen

I møte med informantene skal deres menneskeverd og integritet ivaretas. Her synes det relevant å framheve perspektivet som blir framsatt av Kant i *Morallov og frihet*: «Alle fornuftige vesener er nemlig underlagt den lov at enhver av dem aldri skal behandle seg selv eller noen annen *bare som middel*, men alltid *tillike som formål i seg selv*.» (Kant, 1970, s. 46). Lærerne var ikke midler for min datainnsamling, de var heller ikke midler som skulle sørge for at vurderingssituasjoner var i tråd med gjeldende forskrifter, men mennesker som er mål i seg selv. Det er også viktig å huske at tilbakemeldingene de ga ikke var ment for forskning, men ble skrevet for elevene.

Elevene er en annen gruppe informanter som skal ivaretas. Det som først og fremst var viktig, var å tydelig formidle at deltakelse var frivillig, og at de var informert om hva det innebar å delta i studien. Jeg sørget for at det forelå alternative oppgaver når de skulle svare på spørreundersøkelsen, i tråd med NSDs anbefalinger. Elevene ble i tillegg informert muntlig og via et skriv, der det framgikk at jeg samlet inn noen av deres arbeider fordi lærers tilbakemeldinger ofte var skrevet rett inn i elevarbeidene. Hvis noen ikke ønsket å være med på dette, ville de lett kunne tas ut av studien, men ingen reservert seg mot deltakelse.

4.7.3 Anonymitet

Et sentralt spørsmål om ivaretagelsen av informantene er hvordan de blir framstilt i rapporteringen av forskningen. Det er nødvendig å veie hensynet til transparens i forskningsframstillingen mot hensynet til informantene, men hensynet til informantenes integritet og anonymitet kommer foran hensynet til mest mulig presis forskningsformidling. Siden jeg samlet inn data fra tre lærere på hver skole, var det viktig å skrive om funnene på en slik måte at deltakerne i studien ikke kunne identifiseres. Jeg legger til grunn at kollegaene deres kjenner til at de deltok i studien. I forlengelsen søkte jeg å etablere et datasett der jeg anonymiserte tekster, kommentarer og annet materiale i framstillingen. Det ble heller ikke løftet fram informasjon i artiklene som i særlig grad var egnet til å identifisere lærerne. Her var det naturlig og formålstjenlig å anonymisere funnene i så stor grad

at det i hvert fall ikke var mulig i skjelne de tre lærerne ved hver skole fra hverandre, et arbeid som ble lettere fordi avhandlingen undersøker en praksis fra et teoretisk perspektiv – den søker ikke å beskrive lærernes personlige stil. Det kan også legges til grunn at lærerne deltok frivillig i studien, og ga et ettertrykkelig, informert samtykke til å delta. En viss grad av identifiserbarhet, for eksempel hvis noen har spesifikk kjennskap til akkurat hvilke lærere som deltok i studien i det aktuelle skoleåret, vil jeg derfor mene er innenfor rammene av hva en må kunne forvente.

4.7.4 Bruk av studien – og språklig framstilling

I *Om friheten* fra 1859 framhever John Stuart Mill (2010) et utilitaristisk etikkssyn, der han tar til orde for å maksimere nytten av det en gjør, i mitt tilfelle forskningen, så den i størst mulig grad kommer samfunnet til gode. Som tidligere lektor i skolen er det personlig viktig for meg å etterstrebe en avhandling som ivaretar informantenes perspektiver og viser forståelse for kompleksiteten i deres hverdag. De retningslinjene som er gitt av forskningsfeltet om hva som bidrar til læring i læreres vurderings- og tilbakemeldingspraksis, danner mye av grunnlaget for norske myndigheters reguleringer av dette feltet. Samtidig er det viktig å ha i mente at det ikke er klart hvordan de generelle rådene om god praksis skal omsettes til praktisk handling i klasserommet. Som Wiliam et al. (2004, s. 51) påpeker, er det ikke gitt at forskning uten videre lar seg overføre til praksis:

This difficulty of ‘putting research into practice’ is not the fault of the teacher. But nor is it a failing in the research. Because our understanding of the theoretical principles underlying successful classroom action is weak, research can never tell teachers what to do. Indeed, given the complexity of classrooms, it seems likely, that the positivist dream of an effective theory of teacher action – which would spell out the ‘best’ course of action given certain conditions – is not just difficult and a long way off, but impossible in principle (Wiliam, 2003).

En konsekvens av Wiliams synspunkt er at jeg, som forsker, må være svært varsom med å kritisere lærernes praksis, særlig siden jeg til enhver tid har en ufullstendig innsikt i den komplekse konteksten de virker i.

I forlengelsen av dette har jeg søkt å følge retningslinjene gitt av den amerikanske psykologforeningen APA, og unngå å dele inn informantene i enkle, og derfor unyanserte, kategorier («avoid labeling»), som kan brukes av andre fordringsholdere, som politikere, byråkratiet eller pressen til å framsette forenklete påstander av typen «lærerne gir ikke elevene gode nok

tilbakemeldinger». Forskningsformidlingen må makte å framstille den kompleksiteten som en norsklærers tilbakemeldingspraksis befinner seg innenfor, og samtidig tilføre samfunnet ny kunnskap om denne praksisen, ved at prosjektet gjennomføres i tråd med anerkjente og transparente retningslinjer for god forskning. Et middel for å unngå uheldig bruk av forskning er å unngå å sette poenger fra funnene på spissen for å redusere muligheten for at de kan feiltolkes eller misbrukes. Fra Mills utilitaristiske perspektiv vil også nytten av arbeidet mitt være større om den øker forståelsen for lærernes arbeidssituasjon hos andre grupper i samfunnet, noe som stiller krav til størst mulig grad av klarhet i framstillingen.

5 Sammendrag av artikler

Oversikt over PhD-avhandlingens tre artikler

	Tittel og publiseringsstatus	Forskningsspørsmål	Data	Metoder	Hovedfunn
Artikkel 1	Vurdering for læring i norskfaget: Om lærerens skriftlige tilbakemeldinger på skriftlige tekster. Publisert 2017 i Acta Didactica Norge.	I hvilken grad gir lærerne skriftlige tilbakemeldinger på skriftlige prøver i norskfaget som er i tråd med premisene for formative tilbakemeldinger, slik de er uttrykt i internasjonal forskning?	171 tilbakemeldinger skrevet av 6 lærere til egne elever ved 2 skoler (A og B) på en heldagsprøve. 6 klasser.	Dokumentanalyse, kategorisering av læreres tilbakemeldinger til egne elever.	Mer enn halvparten av tilbakemeldingene er utformet etter VfL-retningslinjer. Det er betydelig forskjell på lærerne, høyest forekomst er ca. 65 %, lavest er 12 %. Det er forskjell på skole A og B.
Artikkel 2	Elevers oppfatning av lærers tilbakemeldingspraksis: om sammenheng mellom graden av VfL-praksis og elevenes opplevde nytte av lærerens skriftlige tilbakemeldinger i skriftlig norsk. Publisert 2017 i Nordic Journal of Literacy Research.	H1: Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte. H2: Elevers oppfatning av lærers ferdighet til å forklare fagstoffet er positivt statistisk assosiert med opplevd nytte. H3: Elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte.	Spørreskjemaundersøkelse, N = 213. Respondenter er elever fra 2 skoler, 12 klasser, ST og YF.	Strukturell likningsmodellering (SEM).	Elevene oppfatter at hvis lærer har og bruker kriterier og forklarer hvordan de kan forbedre seg er dette positivt statistisk assosiert med opplevd nytte, mens hypotese 3 ikke finner støtte. Elevene oppfatter ikke at lærers tilrettelegging for bruk av tilbakemeldinger er positivt assosiert med nytte.
Artikkel 3	Norsklæreres erfaring med læringsfremmende vurdering av elevtekster i to lærerkollegier. I revisjonsprosess – etter krav om «major revisions» fra tidsskrift.	Hva kjennetegner norsklæreres vurderingsarbeid ved to samarbeidskulturer?	Intervjudata av lærere og ledelse ved skole A og B: 2 semistrukturerte gruppeintervjuer, ett intervju med representant for ledelse ved hver skole, ett med representant for lærerne. Oppgaver, kriterier, undervisningsressurser samlet inn over skoleåret 2014–2015 fra 6 lærere i 12 klasser.	Tematisk analyse av transkriberte intervjuer med NVivo11.	Lærerne på skolen med sterkest samarbeidskultur uttrykte at de hadde flere muligheter til å legge til rette for en VfL-orientert praksis. Ved begge skoler var det utfordringer med å få elever til å revidere tekst på grunnlag av lærers tilbakemeldinger, særlig ved skole A. Lærerne ved begge skoler oppfattet at elever hadde vansker med å forstå vurderingskriterier

5.1 Sammendrag av artikkel 1

Vurdering for læring i norskfaget: Om lærerens skriftlige tilbakemeldinger på skriftlige tekster. Artikkelen analyserer norsklæreres tilbakemeldinger på elevtekster skrevet på en heldagsprøve omtrent ni uker før eksamen. Bakgrunnen for å undersøke denne praksisen er den samme i alle tre artiklene, og forskningsfeltet som informerer denne undersøkelsen er gjort rede for i litteraturgjennomgangen.

Undersøkelsen av tilbakemeldinger er forankret i skriveopplæringen i norskfaget. Et sentralt premiss for studien er at lærers skriftlige tilbakemelding til elevene på skriftlige innleveringer er en viktig brikke i skriveopplæringen ved at den har potensial til å hjelpe elevene til å forbedre seg, gitt at den er utformet på en slik måte at den har større sjanse for å fremme læring og spesielt om den er utformet etter prinsippene for VfL, som er kommunisert av norske utdanningsmyndigheter som norm. Litteraturgjennomgangen i avhandlingen (og artikkelen) viser at tilbakemeldinger er et komplekst felt og at forbindelsen mellom tilbakemelding og læring er usikker, men at anbefalinger likevel lar seg utarbeide.

Hvorvidt man finner en slik praksis, vil blant annet avhenge av forskningsspørsmål og teoretisk inngang. Forskningsspørsmålet er: I hvilken grad gir lærerne skriftlige tilbakemeldinger på skriftlige prøver i norskfaget som er i tråd med premissene for formative tilbakemeldinger, slik de er uttrykt i internasjonal forskning?

Metoden er dokumentanalyse av tilbakemeldinger, som dermed er definert som meningsbærende, analyserbare enheter. Tilbakemeldingene er skrevet inn i elevtekster og har et dialogisk forhold til disse: de er kategorisert ut fra både innhold og utforming og i en del tilfeller ut fra forbindelsen til elevenes ytringer. Metoden er gjort mer utdypende greie for i kappens metodedel. 171 elevtekster (en heldagsprøve) med læreres tilbakemeldinger ble analysert ved hjelp av et kodeskjema utarbeidet for studien spesielt. Kodeskjemaet hadde fire hovedkategorier. De to første var knyttet til VfL, og er kalt «mål» og «instruks». De to andre er kalt «korrigering» og «relasjonsbyggende kommentar».

Sentrale funn var at de seks lærerne samlet sett ga omtrent 50 % tilbakemeldinger som kunne kodes som VfL. Det var likevel betydelige forskjeller. Ved skole A hadde lærerne størst forskjeller og tydeligst særpreg. Pseudonymet Anne ga svært mange tilbakemeldinger – noen elever fikk over 70 og omtrent halvparten kunne knyttes til VfL. Berit ga omtrent 12 % VfL-kommentarer og hadde mange språklige rettinger. Carina ga omtrent 65 % VfL-kommentarer, hadde få språklige rettinger

og hadde lagt opp til at elevene kunne jobbe videre med teksten og levere på nytt; dette hadde bare én elev gjort. Ved skole B hadde Dina og Eva henholdsvis 42 % og 59 % Vfl-kommentarer, med litt ulike særpreg, mens Frøy hadde over 65 % og nesten ingen språklige rettinger. Frøy kom nærmest en Vfl-orientert praksis, mens Berit var lengst unna. Bare Dina hadde skrevet et fåtall karakterbegrunnende kommentarer.

I diskusjonen vektlegges at en heldagsprøve kan ses som en summativt orientert vurderingssituasjon, men en teoretisk forståelse som sier at bruken av informasjonen fra vurderingen avgjør om den kan kalles summativ eller formativ er lagt til grunn, og dessuten at dikotomien har et kunstig preg ettersom flere teoretikere (Newton, 2007; Taras, 2010) heller framhever denne dimensjonen som et kontinuum. I tillegg ble retting av språkfeil diskutert ettersom det er omstridt om de bidrar til læring eller ikke, mengden tilbakemeldinger – som i Annes tilfelle, tidspunktet for tilbakemelding – den er gitt etter selve skriveøkten, og at det er uvisst om den skal brukes i videre læringsarbeid. Her legges det til grunn at bruk er viktig, men at innholdet også har stor betydning.

Oppsummert viser undersøkelsen at lærerne i store trekk skrev tilbakemeldinger på heldagsprøvebesvarelsene med elevenes læring i fokus, og at det var lite å se av en eksplisitt summativt orientert praksis.

5.2 Sammendrag av artikkel 2

Artikkelen «Elevs oppfatning av lærers tilbakemeldingspraksis: Om sammenheng mellom graden av Vfl-praksis og elevenes opplevde nytte av lærers skriftlige tilbakemeldinger i skriftlig norsk» har til hensikt å undersøke elevs oppfatninger av norsklærers skriftlige tilbakemeldingspraksis i skriftlig norsk. For at elevene skal ha størst mulig sjanse for å nyttiggjøre seg slike tilbakemeldinger, er det viktig å kjenne til hvordan de oppfatter denne praksisen. Artikkelens viktigste bidrag er at den gir innblikk i elevs oppfatninger av sentrale sider ved det som av forskningen antas å være en læringsfremmende vurderingspraksis.

Det teoretiske grunnlaget for artikkelen er omfattende internasjonal forskning innen feltet Vfl som er gjort rede for i kappens litteraturgjennomgang. Metoden som ble brukt, var strukturell likningsmodellering (SEM) av resultater fra en spørreskjemaundersøkelse (N = 213). Metoden ble valgt for å kartlegge hvordan elevene oppfattet et komplekst forhold som en Vfl-orientert praksis,

der det var et poeng å se sammenheng mellom flere variabler. Elevene tok stilling til påstander, gradert på en sjupunkts skala med navngitte ytterpunkter.

Forskningsspørsmål er formulert som tre hypoteser: 1) Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte. 2) Elevers oppfatning av lærers ferdighet i å forklare fagstoffet er positivt statistisk assosiert med opplevd nytte. 3) Elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte.

Hypotesene er undersøkt med en modell med fire latente variabler der tre operasjonaliserer en VfL-praksis, mens den fjerde, avhengige variabelen, er opplevd nytte. Valget er begrunnet i forskning (Jönsson, 2013) som undersøker hva som gjør at studenter bruker, eller lar være å bruke, tilbakemelding fra lærer i videre arbeid, og der opplevd nytte ses som en sentral faktor. De øvrige variablene var også informert fra forskningsfeltet. Målet med opplæringen ble operasjonalisert gjennom indikatorer som tapper om elevene oppfatter at læreren har og bruker kriterier. Kriterier er ikke det samme som målet (Sadler, 2010), men er likevel sett som en operasjonalisering av sider ved målet for opplæringen. Forklaring er den neste VfL-variabelen, og den er søkt tappet gjennom indikatorer som fokuserer på om elevene opplever at lærer forklarer fagstoffet godt, om de oppfatter at lærer modellerer arbeidsmåter og om lærer er konkret i tilbakemeldingen. Bruk er den siste uavhengige variabelen og bygger på en grunnantakelse om at elever må bruke lærers tilbakemeldinger til å forbedre egen tekst om de skal ha størst sjanse for å bidra til læring (Black & Wiliam, 2009). Variabelen er operasjonalisert gjennom indikatorer som handler om elevene oppfatter at lærer ber elevene jobbe med tilbakemeldinger i timen, om de kan levere tekster på nytt og om lærer krever denne typen arbeid av elevene.

Deskriptiv statistikk viser en gjennomsnittskåre på indikatorer for alle variabler på mellom 4 og 5, noe som kunne indikere at elevene hadde en nøytral holdning til lærernes tilbakemeldingspraksis, men et standardavvik mellom 1,5 og 1,9 kan ses som et uttrykk for at datamaterialet har en betydelig spredning, noe som igjen kan tolkes som at det er betydelig forskjell mellom hvordan elevene oppfatter lærers vurderingspraksis. Det kan tyde på at det er klare forskjeller mellom lærerne.

Resultatene viser at hypotese 1 fikk empirisk støtte ved at stikoeffisienten mellom opplevd klargjøring av kriterier og opplevd nytte var signifikant og forholdsvis sterk $b(\text{Kriterier} \rightarrow \text{Nytte}) = ,43$. Hypotese 2 fikk også støtte i empirien: $b(\text{Forklaring} \rightarrow \text{Nytte}) = ,46$. Hypotese 3 fikk ikke empirisk støtte: $b(\text{Bruk} \rightarrow \text{Nytte}) = ,13$ og var ikke signifikant ($p > 0,12$).

Føyningsmål som ble brukt for å estimere strukturmodellen var kji-kvadrat 78,134, df 38, p-kji ,000, RMSEA ,71, CFI 9,65 og TLI 9,39, og sett under ett indikerer de en akseptabel føyning. Det ble gjort konfirmerende faktoranalyse for å undersøke hvordan de ulike delene av modellen var spesifisert, og analysen viste at kriterier var minst velspesifisert mens forklaring og bruk var velspesifisert. Dette indikerer igjen at hovedfunn i studien kan tillegges empirisk varians og at sjansen for at de er framkommet ved målefeil, er redusert. Funntil knyttet til sammenhenger mellom kriterier og opplevd nytte trenger å utforskes videre.

Det mest oppsiktsvekkende funnet sett i lys av teorien, var mangel på signifikant statistisk assosiasjon mellom opplevd tilrettelegging for bruk og opplevd nytte. Forklaringer kan være tidspunkt for tilbakemeldingen – at elevene jobber med andre emner og andre tekster i faget når de får tilbakemelding, eller finnes i innholdet i tilbakemeldingene.

5.3 Sammendrag av artikkel 3

Artikkelen «Norsklæreres erfaring med læringsfremmende vurdering av elevtekster i to lærerkollegier» bygger på teori om vurdering og vurdering for læring i likhet med de to andre artiklene. Studiens hovedbidrag er å gi innsikt i hvordan lærere i ulike samarbeidskulturer reflekterer over og opplever det å jobbe med læringsfremmende vurdering i skriftlig norsk, og i noen kjerneproblemer i operasjonaliseringen av VfL i skriveopplæringen.

For å belyse vurderingsarbeidet er forskning om læringsfremmende vurdering lagt til grunn, for eksempel Black og Wiliam (2009) i tillegg til forskning på respons på skriving (MacArthur, 2012). Dessuten er forskning på lærersamarbeid (Vangrieken et al., 2015) brukt som rammeverk til analysene og diskusjonene av implikasjoner ved det å jobbe i ulike samarbeidskulturer.

Forskningsspørsmålet er: Hva kjennetegner norsklæreres vurderingsarbeid ved to samarbeidskulturer?

Metoden som ble brukt, var intervju og tematisk analyse av intervjudata, som i tillegg ble sammenholdt med bakgrunnsdata som oppgaver, elevbesvarelser, undervisningsressurser og vurderingskriterier samlet inn over ett skoleår. De transskriberte intervjuene ble analysert tematisk (Braun & Clarke, 2006) ved hjelp av NVivo11. Etter en kodeprosess ble tre temaer vektlagt i videre analyse: skolens fagkultur i norskseksjonen, bruk og oppfølging av tilbakemeldinger på skriftlige oppgaver – inkludert hverandre- og egenvurdering og vurderingskriterier.

Intervjuene med ledelse og representant for norsklærerne viste at det ble lagt til rette for samarbeid over timeplanen på skole B, men ikke på skole A, og det ble beskrevet en høyere grad av lærerautonomi ved skole A, som blant annet viste seg i at lærerne ved skole B brukte felles vurderingskriterier, mens lærerne ved skole A hadde ulike.

Sentrale funn er at lærerne ved skole A ga uttrykk for at de fant det krevende å få elevene til å bruke tilbakemeldinger til å forbedre tekster, noe som også viste seg i bakgrunnsdata, der for eksempel læreren Carina organiserte skriveopplæringen som mappevurdering og der bare 9 av 29 elever leverte mappen ved årsslutt. Lærerne på skole A uttrykte ikke noen samlet praksis for bruk av tilbakemeldinger, men oppfattet det som tilfeldig hva de gjorde; de la også vekt på elevforutsetninger som en årsak til at det var vanskelig å jobbe med revisjon av tekster i klassene. Selv om lærerne ved skole B ikke tematiserte at det var utfordrende å få elever til å slutte opp om skriveprosesser, viste data at dette var et problem ved begge skoler.

Lærerne ved begge skoler oppfattet det som at elever fant det vanskelig å bruke vurderingskriterier i læringsarbeid og ved skole A framhevet lærerne at faglig svakere elever rett og slett hadde problemer med å forstå kriterier.

Noen konklusjoner er at begge skoler hadde utfordringer med å få elevene til å jobbe med tekstrevisjon, og at det kan se ut til å være mye å vinne på å la elevene jobbe mer med formative prosesser og tilbakemeldinger på heldagsprøver, siden dette var en skriveform som hadde stor oppslutning blant elevene ved begge skoler, og at hverandrevurdering kan være en måte å sørge for at alle elever får tilbakemelding på egen tekst i prosess, slik forskningen anbefaler, men som er lite benyttet. Det er samtidig av avgjørende betydning at elevene får god opplæring og klare retningslinjer i både å gi, motta og bruke informasjon fra slike prosesser (Topping, 2009), og at ledelsen legger organisatorisk til rette for at lærerne kan veilede egne elever i denne situasjonen.

6 Diskusjon og forskningsbidrag

Avhandlingens overordnede forskningsspørsmål er: Hva kjennetegner norsklæreres skriftlige vurderings- og tilbakemeldingspraksis i to kollegier? Funn fra de tre artiklene blir diskutert i lys av VfL-teori og VfL-forskning. Jeg legger vekt på at innholdet i lærernes tilbakemeldinger er viktig, og at hvordan de brukes og vilkårene for bruk også kan være av avgjørende betydning for realiseringen av en VfL-orientert praksis i norskfagets skriveopplæring, slik myndighetene på ulike vis har lagt til rette for og dessuten krever.

Et sentralt empirisk funn i avhandlingen er at det ikke er noen signifikant statistisk assosiasjon mellom elevers opplevde nytte av lærers skriftlige tilbakemeldingspraksis, og at lærer legger til rette for bruk av tilbakemeldingen. Enten lærer ber elevene jobbe med teksten i timen, etter at den er gitt tilbake, at de får denne typen arbeid i lekse eller at de har mulighet til å levere en forbedret versjon, så erfarer elevene ikke at denne typen arbeid eller muligheter er assosiert med nytte. Når en ser på teorien om hvordan vurdering kan bidra til læring, så er bruk av innholdet i tilbakemeldingsinformasjon kjernen i tidlige definisjoner, noe som også legges til grunn i Black og Williams definisjon fra 2009: Informasjon fra vurdering må brukes i videre læringsarbeid for at vi skal kunne snakke om VfL, teoretisk sett. Det vil si at når informasjon om hva elevene har fått til, eller hvordan en tekst kan bli bedre, blir brukt til faktisk forbedring, da øker mulighetene for at læring finner sted. I skriveforskningen finnes det knapt noen som hevder at den gode teksten blir til i en rettlinjet skriveøkt der det ikke gjøres revisjoner. Tvert imot framhever f.eks. Hertzberg og Roe (2016), Dysthe (2011) eller MacArthur (2012) at tekstproduksjon er en prosess (Bereiter & Scardamalia, 1987), og at de beste resultatene oppnås når elever gjennom faglig kompetent veiledning får muligheten til å jobbe videre med egne tekster (Sadler, 2010). Når forskningen fra to felt er så klar på dette punktet, er det viktig å undersøke mulige årsaksforklaringer.

En forklaring kan ligge i innholdet i lærers tilbakemeldinger. Parr og Timperley (2010) viser at lærers tilbakemeldinger kan være av ulik kvalitet, og at en uklar eller negativ tilbakemelding kan bidra til å gjøre en elevtekst dårligere når eleven søker å revidere den på grunnlag av det hun/han tror læreren har ment. Dette er konsistent med forskning på tilbakemeldinger, slik det f.eks. er vist av Kluger og DeNisi (1996), av Weaver (2006), Shute (2008) eller Straub (1997). Imidlertid viser funn i artikkel 1 at lærerne på den siste heldagsprøven før eksamen, med variasjoner, ga elevene

tilbakemeldinger som langt på vei var i tråd med anbefalingene fra tilbakemeldingsforskningen. Om en ser heldagsprøven som en kritisk case (Patton, 2014), og vektlegger at på dette tidspunktet i opplæringsløpet – ni uker før eksamen på slutten av 13 års skolegang – står det å gi en god begrunnelse for karakteren i sentrum, viste altså analysene at lærerne ikke eksplisitt begrunnet karakterer. Ut fra dette er det mulig å anta at lærerne også ellers i året gir elevene tilbakemeldinger de kan lære av, slik det er definert i studien – at de retter seg mot målet med opplæringen, inneholder forklaringer på hvordan elevene kan komme nærmere målet, at mange dessuten er skrevet i teksten, slik Parr og Timperley (2010) anbefaler, og at det er mye faglig og personlig ros eller oppmuntring i tekstene, men svært lite kritikk av elevene, noe Straub (1997) eller Smidt (2002) framhever som viktig for at elever skal ønske å engasjere seg i videre arbeid med en tekst. Gjennomgangen av de 1504 elevtekstene med tilbakemeldinger fra læreren er i samsvar med dette bildet, selv om det ikke er gjort en systematisk analyse av mer enn 171. For å oppsummere dette poenget så framstår det som lite plausibelt at elevene ikke vil finne det nyttig at lærer legger til rette for bruk av tilbakemeldinger til tekstrevisjon på grunn av innholdet i tilbakemeldingene. Tvert imot skulle en kunne forvente det motsatte, ettersom rundt halvparten av lærerens tilbakemeldinger kunne knyttes til en anbefalt VFL-praksis.

Det sentrale med dette funnet er at en viktig brikke i grunnlaget for en mulig VFL-orientert praksis kan se ut til å være til stede hos fem av seks lærere. Det betyr at det er større mulighet for disse lærernes elever til å bruke tilbakemeldingen til å utvikle seg som skrivere. Intervjudataene fra artikkel 3 understøtter bildet av lærere som setter elevenes læring først i sin vurderingspraksis. Det må samtidig understrekes at det ikke er noe grunnlag for å hevde at dette faktisk skjer, og det er teoretisk mulig at elevene til den læreren som gir færrest VFL-kommentarer likevel lærer mye, ettersom spørsmålet om hva som kan bidra til læring er svært sammensatt. Satt på spissen kan en kommentar gitt på riktig tidspunkt, bidra til mer læring enn alle kommentarene gitt til en annen elev hvis hun/han ikke er klar for å ta den imot, et forhold som også kan henge sammen med det Elstad (2006) kaller et strategisk valg, og som innebærer at det må være rasjonelt for eleven å engasjere seg i teksten på nytt. Her kan kontekstuelle forhold spille en avgjørende rolle.

Tidspunktet for tilbakemeldingen kan synes som en nærliggende forklaring på manglende opplevd nytte av tilrettelegging for bruk av tilbakemeldingsinformasjon. Eksempelvis legger Hattie og Timperley (2007) til grunn at tilbakemeldingen har større sjans for å bidra til læring om den kommer *i* prosessen, og ikke etter. Dette svaret kan nok være en del av bildet, men det er likevel

ikke fullt ut tilfredsstillende som forklaring. Alle lærerne uttrykker i intervjuet at prosesskriving er best, og alle hadde i noen grad lagt til rette for gjennomføring av denne typen skriveopplæring i løpet av skoleåret. Men på begge skoler var elevopplutningen til dels svak. Skole A hadde svakest grad av opplutning, og her kan det synes nærliggende å vektlegge elevforutsetninger som en forklaring, men situasjonen var ikke så forskjellig ved skole B, der elevene kom inn med klart høyere inntakssnitt.

En annen mulig forklaring på elevers manglende opplevde nytte av å bruke lærers tilbakemeldinger til å forbedre tekst kan ligge i at elevene oppfatter seg som autonome skrivere. Andersen og Hertzberg (2005) og Skar og Aasen (2016) finner i ulike skrivekontekster at en del elever tar sjanser når de skriver. De utfordrer den mest opplagte løsningen på en oppgave og velger personlige og originale innfallsvinkler. Om en ser på læreplanen, er dette helt legitimt – elevene skal øve seg i kreativ skriving og kritisk og selvstendig tenkning. Forskerne ser elevatferden som et uttrykk for at skriveopplæringen i den norske utdanningskonteksten er preget av et forhold mellom skriver og leser som ikke er veldig asymmetrisk. Med Lotmann (1990) kan en si at den modelleseren eleven konstituerer i sin skriveprosess ikke er en formalist eller streng dommer, men en interessert leser, en som vil kunne sette pris på en kreativ innfallsvinkel. Her legger jeg til grunn at elever mot slutten av videregående skole ikke er radikalt forskjellige fra studenter, og at forskning på studenters syn på lærerkommentarer har overføringsverdi til en skolekontekst. At elevene ser seg selv om selvstendige skrivere, kan eksempelet til Calhoon-Dillahunt og Forrest (2013) også illustrere. De fant at studenter satte pris på lærers tilbakemeldinger på tekstutkast, men at de ofte unnlot å følge lærers råd. Straub (1997) fant også at studenter kan sette pris på kritikk på tekstutkast, så lenge kritikken ikke utfordrer tekstens grunnleggende ideer. Gitt at elevene ser seg selv som autonome skrivere, så kan det bidra til å forklare at de ikke finner det nyttig å skulle bruke lærers tilbakemeldinger i videre arbeid hvis de opplever at læreren utfordrer tekstens ideer for mye, eller at de ser læreren mer som leser enn dommer. Om en ser skoleskrivingen i lys av Bereiter og Scardamalias (1987) kognitive modell for tekstproduksjon, kan det tas til støtte for synet på den autonome eleven, som gjerne forbindes med selvregulert-læring-tenkningen (Duckworth & Seligman, 2005), siden forskerne framhever at det er eleven som anvender sine mentale ressurser og evner i en problemløsningsprosess (Anderson, 1982) der hun/han investerer energi og tid, og at eleven følgelig har et eierskap til egen tekst som hun/han ikke uten videre gir fra seg til en leser. Dermed kan noe av forklaringen ligge i skrivingens kognitive dynamikk – eleven eier egen tekst, læreren er en leser som har en kombinert rolle som modellmottaker av teksten og dommer som skal sette karakter.

Om en ser funnet i sammenheng med en lærers erfaring, slik den er beskrevet i artikkel 1, at omtrent en tredel av elevene leverte mappe på slutten av året, mens resten lot det være, selv om det var et krav, så kan det tjene som et eksempel på en konkret situasjon der elevens manglende opplevde nytte kommer til uttrykk. Her fikk de mulighet til å levere tekster på nytt etter å ha fått tilbakemelding og karakter – og de hadde mulighet til å få bedre karakter, men de valgte å ikke bruke muligheten. Her kan det synes som om forestillingen om eleven som en autonom skriver blir utfordret, eller at forklaringen ikke er tilstrekkelig. Noe av forklaringen kan ligge i skrivingens kognitive dynamikk; det er krevende å gå inn i en tekst man har skrevet tidligere, og elevene må være tilstrekkelig motiverte for å gjøre dette arbeidet (Pintrich, 2003). Elevene ved skole A var det man kan kalle faglig svake, om en legger det lave inntakstallet til grunn, og i artikkel 3 uttrykker lærerne at det er svært krevende å få en del av elevene til å jobbe med tekstproduksjon i hele tatt, og at det følgelig er vanskelig å få dem til å gjøre noe annet med en tekst enn å rette opp skrivefeil. Det kan også være at rammevilkårene ikke i tilstrekkelig grad la til rette for at elevene investerte i arbeidet med å forbedre egen tekst, og at faglig svake elever vil trenge andre insentiver for å jobbe med tekst enn faglig sterke. I forlengelsen av dette poenget kan det også tenkes at elevene ikke var sikre på hvordan de skulle bruke lærers tilbakemeldinger.

Om en trekker inn kontekstens betydning, slik den er beskrevet i artikkel 3, så kan det virke som om elevene ved skole A ikke fikk noen systematisk opplæring i bruk av tilbakemeldinger. En slik opplæring er nødvendig, ifølge både Jönsson (2013) og Sadler (2010). I intervjuene ved skole A beskrev lærerne at de faktisk satte av tid i timene eller på andre måter la til rette for bruk av tekstkommentarene, men at dette skjedde på ulike måter og at elevene var motvillige til å gjøre annet enn å rette språkfeil. Ved skole B uttalte lærerne at de hadde ulike strategier for å legge til rette for bruk, men ingen av skolene hadde noen systematisk praksis, i hvert fall ble ikke det framhevet i intervjusituasjonen. Ut ifra de empiriske funnene i artikkel 1 og 3 er det derfor ikke grunn til å anta at elevene fikk noen systematisert opplæring i bruk av tekstkommentarer. Det kan være et bidrag til en forklaring på den manglende statistiske assosiasjonen i artikkel 2.

Elevene har andre avgangsfag enn norsk på ST Vg2, så det er ikke urimelig at de vil velge å prioritere disse fagene ettersom karakterer i norsk på ST Vg2 ikke vil komme på vitnemålet ved avsluttet skolegang. Dette kontekstuelle forholdet kan bidra til å forklare elevenes manglende oppslutning om skriveprosesser, slik Carina opplevde, selv om hun altså både skrev gode tilbakemeldinger, slik artikkel 1 viser, og hadde systemer for bruk, slik det framkommer i artikkel

3 – begge deler er forhold som VFL-forskningen vektlegger. Det er imidlertid mindre sannsynlig som forklaring på manglende oppslutning om skriveprosesser på ST Vg3. En annen forklaring kan derfor ligge i at tilbakemeldingene kommer for sent (Hattie & Timperley, 2007; Shute, 2008), altså at elevene oppfatter at de er ferdige med oppgaven. Hvis en ser på Dysthes artikkel (2011), som undersøker hva som kan være hensikten med tilbakemeldinger når det ikke er intendert at studenter skal revidere tekst, så konkluderer hun med at tilbakemeldingen kan rette seg mot overordnede forhold, som struktur eller andre faktorer som har overføringsverdi fra gang til gang. MacArthur (2012) deler dette synet på overføringsverdi i skriveopplæringen. Det vil altså si at hvis elevene fikk slike tilbakemeldinger, kunne de dra nytte av dem til å forbedre de generelle skriveferdighetene sine og ikke bare lykkes bedre i norskfaget, men også i andre fag. Eksempelvis skulle elevene våren 2015 normalt opp til 3 skriftlige eksamener, og det trengs neppe ytterligere argumentasjon for å hevde at gode skriveferdigheter øker sjansene for å lykkes med disse eksamenene, selv om fagtradisjoner kan variere. Artikkel 1 viste da også at elevene faktisk fikk mange tilbakemeldinger som de kunne ha gode muligheter til å lære av. Hvorfor opplever de da at lærers tilrettelegging for bruk er lite nyttig?

I spørsmålet om kontekstens betydning for arbeidet med å gi elevene tilbakemeldinger i skriftlig norsk, kan det se ut som om artikkel 1 kan gi en form for svar på spørsmålet. Lærerne viste i store trekk at de evnet å gi elevene sine gode tilbakemeldinger – elevene fant en klar positiv sammenheng mellom lærers evne til å forklare fagstoffet i tilbakemeldingene og opplevd nytte, og noen av dem fant at lærer la til rette for bruk. Men Carina og de andre lærerne ved skole A opplevde at selv om de la til rette for bruk og organiserte arbeidskrevende prosesskrivinger, så støttet ikke elevene opp om dette arbeidet. På skole B hadde de derimot en felles organisering av heldagsprøver der læreren var sammen med egne elever hele (eller deler av) dagen og kunne gi dem tilbakemelding og hjelp, slik Dina beskrev i intervjuet. Når lærerne er autonome og jobber hver for seg, er det mer vanskelig å få til en slik praksis. Timeplanene skal romme omtrent 18 timer undervisning for en norsklærer i heltidsstilling i videregående skole, og det følger at en lærer sjelden har mulighet til å ha en dag med norskklassen som har heldagsprøve uten at det er lagt til rette for det organisatorisk. I en praksis der lærer har egen klasse på heldagsprøven og aktivt veileder, vil dessuten spørsmålet om karakter (Butler, 1987) ha redusert betydning, siden den nødvendigvis må komme som en følge av vurderingsarbeidet etterpå. Elevene vil kunne få tilbakemelding akkurat når de trenger det, slik flere studier anbefaler (f.eks. Hattie & Timperley, 2007), noe som muliggjør en VFL-praksis fordi bruk av informasjon fra vurdering (lærer leser elevens tekstutkast der og da, eller deler av det) kan brukes

i elevens videre skriveprosess (Black & Wiliam, 2009). Videre viste bakgrunnsdata ved begge skoler at langvarige skriveprosesser, som fordypningsemnet i norsk, en oppgave lærerne ved skole B fant spesielt vellykket, og mappevurderingen ved skole A, hadde en variabel oppslutning blant elevene. Det var i alle klasser opp mot 10 elever som ikke leverte utkast eller var med i prosessene. Derimot viste elevene ved begge skoler nærmere 100 % oppslutning om heldagsprøver i faget. En implikasjon for praksis blir dermed at det kan se ut som om heldagsprøven i norsk er en gunstig arena for å drive skriveopplæring i videregående skole, gitt at den organiseres som en læringsøkt mer enn som en summativ prøve.

Et poeng er også at en slik organisering ikke innebærer vesentlig merbelastning for lærerne, noe som er viktig, siden tids- og arbeidspress er et opplevd problem for lærere både i Norge og internasjonalt (Skaalvik & Skaalvik, 2008). Rett nok viser noen studier (Shute, 2008) at høytpresterende elever som jobber med komplekse oppgaver kan ha større nytte av forsinket, skriftlig respons, men dette forutsetter også andre forhold – som at elevene er motiverte for denne typen arbeid og at det er rasjonelt å engasjere seg i det – ut over at det kan bidra til læring på lang sikt. Det er grunn til å spørre hvor rasjonelt det er å jobbe videre med et arbeid som er avsluttet, hvis man for eksempel har en matematikkprøve dagen ettersom også krever forberedelser (Skaalvik & Skaalvik, 2008). Tilrettelegging for en formativt organisert heldagsprøve kan slik sett tjene som ett eksempel på en måte skolen kan legge til rette for en mer læringsfremmende vurderingspraksis. Et premiss for at det skal fungere, sett i lys av elevenes manglende oppslutning om skriveprosesser, er at det er fornuftig for elevene å slutte opp om denne typen arbeid (Elstad, 2006). Slik både artikkel 2 og 3 viser, er det ikke arbeid med tekster elevene opplever som avsluttede som synes å fungere best.

Forskningsbidrag

Metodisk

I artikkel 1 ble læreres tilbakemeldinger kodet etter et kodeskjema utviklet for å si noe om graden av VfL-praksis. Det var ikke mulig å finne andre studier som gjorde akkurat dette. Metoden med kategorisering av innholdet i tilbakemeldinger bidrar til kartleggingen av VfL-satsingen. Heller enn å operasjonalisere ulike sider ved lærers tilbakemeldingspraksis, som er gjort før, ville jeg altså spesifikt undersøke en VfL-tilbakemeldingspraksis. Dette var gjort av Parr og Timperley (2010), men av årsaker som er gjort rede for i artikkelen, lot ikke deres kodesystem seg overføre uten

endringer til en norsk kontekst. Gitt at andre forskere kan oppnå akseptable Cohens kappa-skårer, kan kodesystemet være et metodisk bidrag til kartleggingen av VfL-praksis i andre studier.

Designet med å velge ut et kritisk case (Patton, 2014) bidro, med forbehold, til å øke muligheten for å få innblikk i et større materiale som ellers er vanskelig tilgjengelig på grunn av omfanget. I forlengelsen av dette ble det i artikkel 2 brukt statistiske metoder for å undersøke elevenes oppfatning av uavhengige variabler, som samlet sett operasjonaliserer sentrale aspekter ved en VfL-praksis og kobler denne til opplevd nytte fordi dette aspektet ved elevenes oppfatninger kan bidra til å predikere i hvilken grad en VfL-praksis kan være effektiv. De latente variablene ble utviklet for studien spesielt, av samme årsaker som for artikkel 1, da det ikke finnes ferdig utviklede instrumenter som kartlegger elevers oppfatninger av lærers tilbakemeldingspraksis i lys av VfL-teori. Strukturell likningsmodellering kan gi informasjon som vanskelig kan komme fram på andre måter, fordi sammenhenger i komplekse datasett kan analyseres ved hjelp av metoden. Med begrensninger gir metoden informasjon om oppfatningene til populasjonen – elever i videregående skole, i liknende skoler. Innen norskfaget er kvalitative dybdestudier mest vanlig, og artikkelen tilfører fagtradisjonen nye aspekter.

Artikkel 3 bruker tematisk koding av transskriberte intervjudata som metode. Dette er ikke spesielt i forskning på norskfaget, men innholdet i intervjuene er sammenholdt med lesing av andre typer data, hovedsakelig dokumenter, som er sentrale i lærers praksis. Heller ikke dette er nytt. Det er mange studier som bruker flere metoder for å kartlegge et fenomen (f.eks. Bueie, 2015; 2016, Gamlem, 2014; Sandvik & Buland, 2014), men de gjør det i andre kontekster enn de som er undersøkt i denne avhandlingen. At det er samlet data fra seks lærere gjennom et helt skoleår, er også mer enn mange andre studier. Tidsaspektet og innsamlingen og lesingen av dokumenter fra et helt skoleår tilfører mer enn intervjudata alene. Säljö (1997) framholder at intervjuer må ses som uttrykk for intensjonalitet og diskurskompetanse. Det vil si at intervjuene gir innblikk i hvordan lærerne oppfatter at det er hensiktsmessig å snakke om vurderingspraksis med kollegaer og forsker, og at de ikke kan si noe om lærernes faktiske arbeid. Ved å sammenholde intervjudataene med andre dokumenter kan studien, med klare begrensninger, overskride denne ulempen og gi mer informasjon.

Metoden i artikkel 3 er inspirert av Lijparts (1971) prinsipp om komparativt design for å undersøke to kulturer i den hensikten å kunne få innsikt i forskjeller mellom dem. Metodevalget kan slik bidra til å gi innsikt i hvordan kontekstforhold kan påvirke mulighetene for å ha en VfL-praksis,

og samlet sett bidrar avhandlingen til å gi innblikk i ulike aktørperspektiver, seks læreres faktiske tilbakemeldingspraksis, deres refleksjoner over denne praksisen og mer generaliserbar kunnskap om hvordan ulike sider ved VfL-satsingen praktiseres og oppfattes i skolen.

Empirisk

Hvis vi vet mer om hva slags tilbakemeldinger lærere skriver, er det lettere å ha en informert oppfatning av hvordan satsingen på VfL lykkes og hvordan lærere kan støttes i sitt arbeid. Artikkel 1 gir derfor et informert grunnlag for videre diskusjon av både praktiske implikasjoner og spørsmålet om videre forskning.

Artikkel 2 gir kunnskap om elevers oppfatning av sammenhenger i en VfL-informert praksis. Hvis VfL skal ha et optimalt potensial i norsk skole, trenger man innblikk i hvordan elever oppfatter en slik praksis, og denne studien bidrar med slik kunnskap. Elevene opplever at hvis lærer har og bruker kriterier og forklarer fagstoffet, er det positivt statistisk assosiert med opplevd nytte. Det er ikke en statistisk signifikant assosiasjon mellom opplevd tilrettelegging for bruk og opplevd nytte, noe som kan tolkes som at elevene ikke opplever at lærers tilrettelegging for bruk av informasjon fra tilbakemeldinger til å forbedre tekst er nyttig. Ettersom bruk av informasjon fra vurdering i videre læringsarbeid er et kjernepunkt i VfL, er dette et svært sentralt funn med praktiske implikasjoner.

Artikkel 3 får fram informasjon om hvordan lærere opplever at det er å jobbe med en vurderingspraksis som skal bidra til læring i to kollegier. Et sentralt funn er at skolen som har en tydelig formulert kultur for samarbeid, er mer fagorientert, på den måten at de er mer samstemte enn et mer autonomt orientert kollegium. Dette er ikke overraskende, men kan likevel fungere som et argument for mer faglig samarbeid mellom norsklærere, fordi det ser ut til å gi lærerne flere muligheter til å implementere en VfL-orientert vurderingspraksis. Samtidig må samarbeidet ha klare grenser, blant annet slik at en opplevelse av autonomitap og andre negative konsekvenser av tett samarbeid kan unngås (Vangrieken et al., 2015). Studien avdekker også hvordan lærerne opplever det å engasjere elever i arbeid med å nyttiggjøre seg informasjon fra tilbakemeldinger i skriveprosesser. Lærerne finner dette utfordrende, særlig ved skole A, og artikkel 3 gir slik mer informasjon om et hovedfunn i artikkel 2.

Teoretisk

I utviklingen av kodesystemet som danner grunnlaget for artikkel 1, viser det seg at det ikke er mulig, i lærerens tilbakemelding, å skille mellom at 1) elevene blir kjent med målet for opplæringen, og 2) hvordan de skal komme seg nærmere målet. Innsikten bidrar til å synliggjøre et skille mellom tilbakemeldinger i en skriveopplæringskontekst med komplekse mål og den forskningen som ofte siteres av norske utdanningsmyndigheter, f.eks. Black og Wiliam (1998) eller Wiliam et al. (2004), som baserer seg på fenomenet «classroom assessment», der studieobjektet ofte er grunnskoleelevers læring i klasserommet, operasjonalisert i det Shulman (2005) kaller en internasjonalt anerkjent «mønstertime», hvor lærer klargjør mål, forklarer fagstoff, aktiverer elevene og oppsummerer til slutt, og der alt er rammet inn i en helklassekontekst. En konsekvens av denne teoretiske innsikten er at det bør gis mer målrettede anbefalinger til ulike typer lærere og for ulike kontekster, og at de overordnede prinsippene risikerer å tape mening hvis de kommuniseres på et svært generelt grunnlag. Dette poenget framheves av Bennett (2011), men uten at det er knyttet til den konteksten som er undersøkt i denne avhandlingen.

Ettersom mye forskning på VfL springer ut fra «classroom assessment-forskning» og funnene fra denne forskningen er overført til andre utdanningskontekster i form av generelle råd, så bidrar denne studien med innsikt i hvordan elever i en skriveopplæringskontekst, som både er lik – og ulik – klasseromssituasjonen, oppfatter et kjerneelement i VfL. Det er ikke sikkert de rådene som den generelle VfL-teorien gir, slik de er kommunisert av norske myndigheter, uten videre lar seg omsette i praksis i andre kontekster – de må oversettes. Denne studien bidrar med innsikt i hvorfor denne oversettelsen må finne sted, nemlig at elever ikke finner arbeid med revisjon av tekster de er ferdige med nyttig, mens dette er et premiss for mye av norsklærernes tilbakemeldingspraksis.

Sammenhengen mellom lærersamarbeid og VfL er lite utforsket, og artikkel 3 gir derfor innblikk i hvordan mulighetene for å implementere en VfL-orientert praksis kan framtre i ulike skolekontekster og i ulike samarbeidskulturer, der autonome kulturer ser ut til å ha færre muligheter for å implementere en VfL-orientert praksis enn en skole med mer samarbeid som er organisatorisk tilrettelagt for av ledelsen. Teoretisk sett bidrar dette til å utvide kunnskapen om hvordan en kan jobbe med vurdering som bidrar til læring. Siden de store metastudiene som anbefalingene fra norske myndigheter tar utgangspunkt i er svært generelt formulert (Black & Wiliam, 1998), og teoretiske bidrag til feltet (Black & Wiliam, 2009; Sadler, 2010; Taras, 2010) også munner ut i generelle råd, bidrar avhandlingen med å konkretisere og tydeliggjøre de teoretiske utgangspunktene – noe som er nødvendig om teorien skal kunne omsettes i praksis (Bennett, 2011).

Kombinasjonen av empiriske funn viser at selv om lærerne gir gode tilbakemeldinger, definert fra VfL-perspektiv, er det ikke nok til å engasjere elevene i videre arbeid med egne tekster, og det er heller ikke nok at lærerne faktisk legger til rette for bruk av tilbakemeldingene, et poeng som også står svært sentralt i definisjonen av en læringsfremmende praksis. Spørsmålet blir da hva som er en god tilbakemelding til en elev som skriver en kompleks tekst, og hvordan man kan engasjere eleven i videre skrivearbeid.

Konklusjoner – svar på forskningsspørsmålet

Hva kjennetegner norsklæreres skriftlige vurderings- og tilbakemeldingspraksis i to kollegier? Jeg besvarer det overordnede forskningsspørsmålet ved å først se på delstudiene i de tre artiklene hver for seg. I artikkel 1 var forskningsspørsmålet: «I hvilken grad gir lærerne skriftlige tilbakemeldinger i norskfaget som er i tråd med premissene for formative tilbakemeldinger, slik de er uttrykt i internasjonal forskning?» Her fant jeg at lærerne i forholdsvis høy grad ga slike tilbakemeldinger, et funn jeg velger å tolke positivt. Lærerne skrev med elevenes læring som mål. Det var imidlertid klare forskjeller mellom lærerne. Særlig ved skolen med høyest grad av lærerautonomi var det variasjoner: En lærer hadde en praksis som lå svært tett på en anbefalt praksis – hun ga mange læringsfremmende kommentarer, og la til rette for arbeid med tilbakemeldingene. En annen lærer hadde svært få tilbakemeldinger, og hennes praksis kan beskrives som en korrekturlesers, med noen få kommentarer der et lavt antall av disse var utformet slik at de hadde større sjanse for å fremme læring. Den tredje læreren ga elevene svært mange tilbakemeldinger. Selv om hun skrev gode tilbakemeldinger, og mer enn halvparten kunne kodes som VfL, så var mengden av tilbakemeldinger så stor (mer enn 2000 på 32 elevtekster) at det er vanskelig å se hvordan en elev skal kunne skille viktig fra uviktig og bruke dem i videre arbeid. Lærerne ved skolen med en mer omforent praksis lå tettere på en VfL-praksis samlet sett. Omtrent halvparten av kommentarene deres hadde en VfL-orientert utforming.

I artikkel 2 ble tre hypoteser utforsket, som i sum kan ses som uttrykk for en operasjonalisering av en VfL-praksis. Hypotesen om at elevers oppfatning av lærers ferdighet til å forklare fagstoffet er positivt statistisk assosiert med elevenes opplevde nytte av tilbakemeldingen, fant empirisk støtte. Dette var ut fra forskningslitteraturen lite overraskende. Hypotesen at elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte, fant ikke støtte. Stikoeffisienten var ikke signifikant, og dette funnet var overraskende, ettersom

forskningen så tydelig framhever bruk av tilbakemeldinger som et kjernepunkt i en VfL-orientert praksis. Den siste hypotesen var at elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte. Denne hypotesen fant støtte, men den statistiske validiteten krever mer forskning for å kunne vektlegges.

I artikkel 3 var forskningsspørsmålet: Hva kjennetegner norsklæreres vurderingsarbeid ved to samarbeidskulturer? Undersøkelsen viste at lærerne ved skolen med tette fagsamarbeid syntes å ha flere muligheter til å realisere en læringsfremmende vurderingspraksis, eksempelvis gjennom samstemming av mål på tvers av klasser og økt mulighet for å organisere veiledete skriveprosesser, der elever kan bruke lærers tilbakemeldinger. Datamaterialet viste også at lærerne ved begge skoler satte elevenes læring i sentrum, men at det kunne være krevende å få elevene til å jobbe med tekster som hadde fått tilbakemelding fra lærer, særlig ved skolen med svakest elevforutsetninger, men også ved skolen med faglig sterkere elever og tettere samarbeid. Et annet funn var lærerne oppfattet at elever, uavhengig av skolekontekst, synes å ha vanskelig for se læringsfremmende aspekter ved vurderingskriterier eller rett og slett forstå dem. Gitt satsingen på utvikling av kriterier i norsk skole (Udir, 2009), er dette et viktig funn.

Om en sammenholder de tre artiklene, kan det se ut som om lærerne, uavhengig av skoler, langt på vei skrev gode tilbakemeldinger, men at det var krevende å få elevene til å bruke tilbakemeldingene til videre læringsarbeid. Det sistnevnte (bruk av tilbakemeldinger) vektlegger forskningen som et kjernepunkt i VfL. En mulig forklaring på manglende bruk kan være å finne i elevenes manglende opplevde nytte av slikt arbeid. Det er også en konklusjon at lærerne på skolen med størst grad av norskfaglig samarbeid og organisatorisk tilrettelegging for samarbeid, synes å ha større muligheter for å få til en VfL-orientert praksis, men at denne koblingen samtidig er kompleks.

Videre forskning

Studier som retter seg mot elevens motivasjon (Dweck, 2000) slik den måles fra ulike teoretiske ståsteder, elevens selvregulering (Pintrich & Zusho, 2002), elevens «volition» og selvdisiplin (Baumeister, 2002) kunne også vært aktuelle innfallsvinkler til en utforskning av læringsfremmende tilbakemeldinger i norsk skole. Videre kan endringer i rammevilkårene, ifølge Elstad (2006), f.eks. ved at elever i større grad pålegges å arbeide videre med tekster som har fått tilbakemelding, påvirke elevenes innsats i slikt arbeid. Elstad (2008) viser også at lærere har større sjanse for å lykkes med denne typen arbeid om de står sammen som kollegium. Om en kontekst av klarere krav

stilt av et samlet kollegium kunne påvirke elevens oppfatning av lærers tilbakemeldingspraksis, er følgelig interessant for videre forskning.

Implikasjoner for praksis

Implikasjoner av funnene i denne avhandlingen for lærernes praksis er at de har mye å vinne på å redusere omfanget av kommentarer på sluttprodukter, fordi det å bruke disse kommentarene ikke forbindes med opplevd nytte av elevene. Lærerne legger i intervjuene til grunn at kommentarene er skrevet for læring, men elevene oppfatter ikke noen nytte av å jobbe med dem. Her er det en tydelig diskrepans som bør få konsekvenser for lærernes arbeidsmåter og organisering av skriveopplæringen. Forskning av Straub (1997) og Jönsson (2013) viser at elever og studenter setter pris på tilbakemeldinger, men det er ikke det samme som at de lærer av dem. Lærerne bør, i tråd med vurderingsforskriften, bruke kreftene sine på å gi tilbakemeldinger når elevene har størst nytte av det, det vil si når de er i en skriveprosess.

For at dette skal være mulig, må skolen legge til rette for at en læringsfremmende vurderingspraksis kan finne sted. Dette er et ansvar enkeltlærere vanskelig kan påta seg alene. Funnene fra avhandlingen illustrerer at skolen der det er lagt til rette for samarbeid i norskseksjonen, har kommet lengst i å lykkes med en læringsfremmende praksis. Lærerne på skolen der det ikke er tilrettelagt, legger til dels ned en stor innsats både organisatorisk og i antall arbeidstimer brukt på å skrive tilbakemeldinger, men som gruppe viser de seg mer ujevne enn ved skolen der en slik tilrettelegging er på plass.

På et policynivå kan funnene i avhandlingen ses som et insentiv til å fortsette å styrke arbeidet med å fremme en læringsfremmende vurderingspraksis. Funn av Hayward (2015), som evaluerte den skotske satsingen «Assessment is for learning», viser at når videre utvikling av læringsfremmende praksis overlates til lokalt nivå, så går den tilbake. Utdanningssystemet er dels bygd opp mot en sortering av elever, og karakterer er adgangsbilletter til videre karrierer for mange av elevene, så sikring av en reliabel sluttvurdering er viktig. Det er derimot ikke viktig at elevene får karakterer hele veien gjennom skoleåret. At den generelle delen av læreplanen (Udir, 2017c) understreker at testing av elever uten oppfølging kan være skadelig, er derfor et positivt signal. At elevene først og fremst går på skolen for å lære, kan synes opplagt, og da bør også satsingen på en vurderingspraksis som bidrar til læring fortsette, med støtte fra sentralt hold.

Referanser

- Andersen S. og Hertzberg, F. (2005): Å ta en risk. «Utfordrerne» blant eksamensskriverne I: Berge, K. L., et al., *Ungdommers skrivekompetanse, bind II, Norskeksamen som tekst*, s. 275–301. Oslo: Universitetsforlaget.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), s. 369–406.
- ARG (2002). *Assessment for Learning: 10 Principles*, hentet fra <http://www.assessment-reform-group.org.uk/>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), s 191–215.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford og New York: Oxford University Press.
- Bakhtin, M. (1986). *Speech genres and other late essays*. Austin: University of Texas Press.
- Baumeister, R. F. (2002). Yielding to temptation: Self-control failure, impulsive purchasing, and consumer behavior. *Journal of Consumer Research*, 28(4), s. 670–676.
- Beach, R. (1989). Evaluating writing to learn: Responding to journals. I Lawson, B. Ryan, S. og Winterowd, W. R. (red.), *Encountering Student Texts: Interpretive Issues in Reading Student Writing*, s. 183–197. Urbana, IL: National Council of Teachers of English.
- Beach, R. og Friedrich, T. (2006). Respons to Writing. I MacArthur, C. A., Graham, S. og Fitzgerald, J. (red.), *Handbook of Writing Research*, s. 222–234. New York og London: The Guilford Press.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), s. 5–25.
- Bereiter, C. og Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale: Lawrence Erlbaum Associates Publishers.
- Berge, K. L. (1996). *Norskensorenes tekstnormer og doxa. En kultursemiotisk og sosiotekstologisk analyse*. Trondheim: NTNU.
- Berge, K. L., Evensen, L. S., Hertzberg, F. og Wagle, W. (2005a). *Ungdommers skrivekompetanse – bind I, Norskensuren som kvalitetsvurdering*. Oslo: Universitetsforlaget.
- Berge, K. L., Evensen, L. S., Hertzberg, F. og Wagle, W. (2005b). *Ungdommers skrivekompetanse – bind II, Norskeksamen som tekst*. Oslo: Universitetsforlaget.
- Berge, K. L. og Skar, G. (2015). *Ble elevene bedre skrivere? Intervensjonseffekter på elevers skriveferdigheter og skriveutvikling*. Trondheim: Høgskolen i Sør-Trøndelag.
- Bergman-Claeson, G. (2003). *Tre lärare – tre världar: lärarkommentarer till elevtexter i tre gymnasieklasser*. Doktorgradsavhandling. Uppsala universitet.
- Biesta, G. J. J. (2016). ICT and education beyond learning: a framework for analysis, development and critique. I Elstad, E. (red.). *Digital expectations and experiences in education*, s. 29–43. Rotterdam/Boston/Taipei: Sense Publishers.
- Black, P. og Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), s. 7–74.
- Black, P. og Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* (formerly: *Journal of Personnel Evaluation in Education*), 21(1), s. 5–31.
- Bland, J. M. og Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314 (7080), s. 572.
- Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), s. 1–12.
- Blumer, H. (1969). *Symbolic Interactionism. Perspective and Method*. Berkley: University of California Press.
- Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), s. 151–167.
- Boud, D. og Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), s. 399–413.

- Braun, V. og Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), s. 77–101.
- Brinkmann, S. (2014). Interview. I *Encyclopedia of Critical Psychology*, s. 1008–1010. New York: Springer.
- Britton, J. L. (1975). *The Development of Writing Abilities (11-18)*. London: Schools Council Publications.
- Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75(5), s. 631–661.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York og London: Guilford Publications.
- Bryman, A. (2012). *Social research methods (4th edition)*. Oxford: University of Oxford Press.
- Bueie, A. A. (2014). «Jeg synes det er ett fett om det er ‘klarte’ eller ‘hadde’.» Om elevers forståelse og reaksjon på lærers tekstrespons i norskfaget. I A. J. Aasen, et al., (red.) *Skriv! Les!* (Bind 2) s. 109–132. Bergen: Fagbokforlaget.
- Bueie, A. A. (2015). Summativ vurdering i formativ drakt – elevperspektiv på tilbakemelding fra heldagsprøver i norsk. *Acta Didactica Norge*, 9(1), Art-4, s. 1–21.
- Bueie, A. A. (2016). Nyttige og mindre nyttige lærerkommentarer – slik elevene ser det. *Nordic Journal of Literacy Research*, 2(1), s. 1–22.
- Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, 60(6), s. 626–648.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79(4), s. 474–482.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational psychology*, 5(1), s. 1–14.
- Calhoun-Dillahunt, C. og Forrest, D. (2013). Conversing in Marginal Spaces: Developmental Writers' Responses to Teacher Comments. *Teaching English in the Two Year College*, 40(3), s. 230–247.
- Charon, J. M. (1995). *Symbolic interactionism*. Englewood Cliffs, NJ: Prentice Hall.
- Christoffersen, K. A. (2013). *Introduksjon til statistisk analyse. Regresjonsbaserte metoder og anvendelse*. Oslo: Gyldendal Akademisk.
- Connors, R. J. og Lunsford, A. A. (1993). Teachers' rhetorical comments on student papers. *College Composition and Communication*, 44(2), s. 200–223.
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education*, 27(2), s. 199–214.
- Creswell, J. W. (2013). *Qualitative inquiry and research designs. Choosing among five approaches*. (3. utgave). Los Angeles: SAGE publications.
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G. og Van Petegem, P. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education*, 36, s. 1–11.
- Dobbie, W. og Fryer Jr, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4), s. 28–60.
- Duckworth, A. L. og Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16(12), s. 939–944.
- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- Dysthe, O. (1987). *Ord på nye spor – Innføring i prosessorientert skrivepedagogikk*. Oslo: Det Norske Samlaget.
- Dysthe, O. (2011). 'What is the Purpose of Feedback when Revision is not Expected?' A Case Study of Feedback Quality and Study Design in a First Year Master's Programme. *Journal of Academic Writing*, 1(1), s. 135–142.
- Elstad, E. (2006). Understanding the nature of accountability failures in the technology-filled classroom:

- disaffected students and teachers who give in. *Journal of Curriculum Studies*, 38(4), s. 459–481.
- Elstad, E. (2008). Towards a theory of mutual dependency between school administrators and teachers. Bargaining theory as research heuristic. *Educational Management. Administration and Leadership*, 36(3), s. 383–404.
- Elwood, J. (2006). Formative assessment: possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13(2), s. 215–232.
- Evensen, L. S., Berge, K. L., Thygesen, R., Matre, S. og Solheim, R. (2016). Standards as a tool for teaching and assessing cross-curricular writing. *Curriculum Journal*, 27(2), s. 229–245.
- Evensen, L. S. (2005). Perspektiv på innhold? Relieff i ungdomsskoleelevers eksamensskrivning. I: Berge, K. L., et al., *Ungdommers skrivekompetanse bind II, Norskeksamen som tekst*, s. 191–236. Oslo: Universitetsforlaget.
- Evetts, J. (2013). Professionalism: Value and ideology. *Current Sociology*, 61(5-6), s. 778–796.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. London, Mahwah og New Jersey: Routledge.
- Ferris, D. R. (2004). The “grammar correction” debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...?). *Journal of Second Language Writing*, 13(1), s. 49–62.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.
- Fix, M. (2008). *Texte schreiben. Schreibprozesse im Deutschunterricht*. Paderborn: Verlag Ferdinand Schöningh.
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), s. 105–112.
- Gadamer, H. G. (2008). *Philosophical hermeneutics*. Berkley og Los Angeles: University of California Press.
- Gamlem, S. T. M. (2014). *Tilbakemelding Som Støtte for Læring På Ungdomssteget*. Universitetet i Stavanger. Det humanistiske fakultet.
- Gamlem, S. T. M. og Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), s. 150–169.
- Gibbs, G. og Simpson, C. (2004). Does your assessment support your students’ learning? *Journal of Teaching and Learning in Higher Education*, 1(1), s. 3–31.
- Glatter, R. (2012). Persistent preoccupations: The Rise and Rise of School Autonomy and Accountability in England. *Educational Management Administration & Leadership*, 40(5), s. 559–575.
- Goddard, Y. L., Goddard, R. D. og Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), s. 877–896.
- Goddard, R. D., Hoy, W. K. og Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, 37(2), s. 479–507.
- Goddard, R. D., Hoy, W. K. og Hoy, A. W. (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher* 33(3), s. 3–13.
- Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S. Keller, U. Fischbach, A. og Preckel, F. (2014). «My Questionnaire is Too Long!» The assessment of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology* 39(3), s. 188–205.
- Gorman, T. P., Purves, A. C. og Degenhart, R. E. (red.) (1988). *The IEA Study of Written Composition I: The International Writing Tasks and Scoring Scales*. New York: Pergamon Press.
- Gough, D., Thomas, J. og Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic reviews*, 1(1), s. 1–28.
- Graham, S., Harris, K. og Hebert, M. (2011). *Informing writing: The benefits of formative assessment*. Washington DC: Alliance for Excellent Education.
- Grice, H. P. (1975). Logic and conversation i: Cole, P. og Morgan, J. (red.): *Syntax and Semantics: Speech Acts. Vol. 3*. New York: Academic Press.
- Halliday, M. A. L. (1994). *An introduction to functional grammar*. London: Arnold.

- Hanson, N. R. (1958): *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*. Cambridge: Cambridge University Press.
- Hanushek, E. A. og Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), s. 297–327.
- Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age*. New York: Teachers College Press.
- Hattie, J. og Timperley, H. (2007). The power of feedback. *Review of educational research* 77(1), s. 81–112.
- Havnes, A., Smith, K., Dysthe, O. og Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1), s. 21–27.
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, 22(1), s. 27–43.
- Hayward, L., Priestley, M. og Young, M. (2004). Ruffling the calm of the ocean floor: merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30(3), s. 397–415.
- Hertzberg, F. og Roe, A. (2016). Writing in the content areas: a Norwegian case study. *Reading and writing*, 29(9), s. 555–576.
- Hoel, T. L. (1995): *Elevsamtaler om skriving i vidaregåande skole. Responsgrupper i teori og praksis*. Doktorgradsavhandling, Universitetet i Trondheim.
- Hopfenbeck, T., Tolo, A., Florez, T. og El Masri, Y. (2013). Balancing Trust and Accountability? The Assessment for Learning Programme in Norway: A governing complex education systems case study. Oslo: OECD.
- Horvath, B., K. (1984). The components of written response: a practical synthesis of current views. *Rhetoric Review* 2(2), s. 136–156.
- Hsieh, H.-F. og Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), s. 1277–1288.
- Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of second language writing*, 7(3), s. 255–286.
- Johnson, B. og Christensen, L. (2014). *Educational Research: Quantitative, Qualitative, and Mixed Approaches 5th Edition*. Los Angeles: Sage Publications.
- Igland, M. A. (2008). *Mens teksten blir til: Ein kasusstudie av lærarkommentarar til utkast*. Oslo: Unipub.
- Ingersoll, R. og Kralik, J. M. (2004). *Teaching quality*. Education Commission of the States.
- Jönsson, A. (2013). Facilitating productive use of feedback in higher education. *Active learning in higher education*, 14(1), s. 63–76.
- Kant, E. (1970): *Morallo og frihet*. Oslo: Gyldendal.
- Keltchermans, G. (2006). Teacher collaboration and collegiality as workplace conditions. A review. *Zeitschrift für Pädagogik*, 52(2), s. 220–237.
- Kennedy, M. M. (1997). The connection between research and practice. *Educational Researcher*, 26(7), s. 4–12.
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling. 2nd edition*. New York og London: The Guilford Press.
- Kluger, A. N. og DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), s. 254–284.
- Krogh, E. og Juul Jensen, M. (2008). *Portfolioevaluering og portfoliodidaktik*. København: Dansk-lærerforeningens forlag.
- Kronholm-Cederberg, A. (2009). *Skolans responskultur som skriftpraktik: gymnasisisters berättelser om lärares skriftliga respons på uppsatsen*. Åbo: Åbo akademis forlag.
- Kvale, S. og Brinkmann, S. (2010). *Det kvalitative forskningsintervjuet*. Oslo: Gyldendal Akademisk.
- Kvernbekk, T. (2005). *Pedagogisk teoridannelse: insidere, teoriformer og praksis*. Bergen: Fagbokforlaget.
- Kvithyld, T. og Aasen, A. J. (2011). Fem teser om funksjonell respons på elevtekster. *Tidsskriftet Viden om*

læsning, 5, s.10–16.

Landis, J. R. og Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), s. 159–174.

Lave, J. og Wenger, E. (1991): *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lea, M. R. og Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in higher education*, 23(2), s. 157–172.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, 13(4), s. 285–312.

Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, 12(3), s. 180–198.

Leithwood, K. A. og Riehl, C. (2003). *What we know about successful school leadership*. Nottingham: National College for School Leadership.

Lejonberg, E. og Elstad, E. (2014). Bruk av profesjonsbegrepet i kampen om utdanningspolitikkenes innhold, i Elstad, E. og Helstad, K. (2014). *Profesjonsutvikling i skolen*. Oslo: Universitetsforlaget.

Lévinas, E. (2004). *Den annens humanisme (3. utg.)*. Oslo: Aschehoug.

Lijphart, A. (1971). Comparative Politics and the Comparative Method. *American Political Science Review*, 65(3), s. 682–693.

Lombard, M., Snyder-Duch, J. og Bracken, C. C. (2002). Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research*, 28(4), s. 587–604.

Lortie, D. C. (1975). *School teacher: A sociological inquiry*. Chicago: University of Chicago Press.

Lotmann, J. (1990). *Universe of the Mind*. New York og London: Tauris.

Luhmann, N. (1986). The autopoiesis of social systems. *Sociocybernetic paradoxes*, 6(2), s. 172–192.

Lund, T. (2002). *Innføring i forskningsmetodologi*. Oslo: Unipub forlag.

Løgstrup, K. E. (1956): *Den etiske fordring*. København: Gyldendal.

MacArthur, C. A. (2012). Evaluation and revision processes in writing. I Berninger, V. W. (red.). *Past, present, and future contributions of cognitive writing research to cognitive psychology*, s. 561–483. London: Psychology Press.

Magno, C. og Amarles, A. (2011). Teachers' feedback practices in second language academic writing classrooms. *The International Journal of Educational and Psychological Assessment*, 6(2), s. 21–30.

Marshall, B. (2004). Goals or horizons – the conundrum of progression in English: or a possible way of understanding formative assessment in English. *Curriculum Journal*, 15(2), s. 101–113.

Matre, S. og Solheim, R. (2014). Lærersamtalar om elevtekstar – mot eit felles fagspråk om skriving og vurdering. I Hvistendahl R. og Roe, A. (red.), *Alle tiders norskdidaktiker*, s. 219–244. Oslo: Novus.

Mausethagen, S. og Granlund, L. (2012). Contested discourses of teacher professionalism: current tensions between education policy and teachers' union. *Journal of Education Policy* 27(6), s. 815–833.

Maxwell, J. A. (2005). *Qualitative Research Design. An interactive approach*. London: Sage Publications.

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Mill, J. S. (2010): *Om friheten*. Oslo: Vidarforlaget.

Moolenaar, N. M., Slegers, P. J. og Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education*, 28(2), s. 251–262.

Mortimore, P., Field, S. og Pont, B. (2006). Equity in education (country note). Thematic review. Norway. Oslo: OECD.

Myhill, D. A., Jones, S. M., Lines, H. og Watson, A. (2012). Re-thinking grammar: the impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education*, 27(2), s. 139–166.

- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), s. 149–170.
- Nystrand, M. og Wiemelt, J. (1991). When is a text explicit? Formalist and dialogical conceptions, *Text II*(1), s. 25–41.
- Ongstad, S. (2009). Kan skriveutvikling planlegges? Enkelteleven i møte med skolens oppgavekultur. I Haugaløkken, O., Evensen, L. S., Hertzberg F. og Otnes H. (red.), *Tekstvurdering som didaktisk utfordring*, s. 74–83. Oslo: Universitetsforlaget.
- Parr, J. M. og Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing writing*, 15(2), s. 68–85.
- Patton, M. Q. (2014). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. Thousand Oaks: Sage Publications.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), s. 667–686.
- Pintrich, P. R. og Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. I Wigfield, A. og Eccles, J. S. (red.). *Development of achievement motivation*, s. 249–284. San Diego: Academic Press.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), s. 4–13.
- Rommetveit, R. (1972). *Språk, tanke og kommunikasjon*. Oslo: Universitetsforlaget.
- Rommetveit, R. (1974). *On message structure: A framework for the study of language and communication*. London: Wiley.
- Saari, H. og Purves, A. C. (1992). The curriculum in mother-tongue and written composition. I A.C Purves: *The IEA Study of Written Composition 2: Education and performance in fourteen countries*. Oxford: Pergamon Press, s. 37–86.
- Sachs, J. (2016). Teacher professionalism: why are we still talking about it? *Teachers and Teaching*, 22(4), s. 413–425.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), s. 119–144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), s. 77–84.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), s. 175–194.
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), s. 535–550.
- Saliu-Abdulah, D., Hellekjær, G. O. og Hertzberg, F. (2017). Teachers' (Formative) Feedback Practices in EFL Writing Classes in Norway. *Journal of Response to Writing*, 3(1), s. 31–55.
- Sandvik, L. V. (2012). Skrivekompetanse i fremmedspråk – hva innebærer det? *Norsk pedagogisk tidsskrift*, 2(96), s. 154–165.
- Sandvik, L. V. og Buland, T. H. (2014). *Vurdering i skolen. Utvikling av kompetanse og fellesskap*. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS).
- Schmidt, F. L. og Rader, M. (1999). Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52(2), s. 445–464.
- Scott, D. (2007). Resolving the quantitative-qualitative dilemma: a critical realist approach. *International Journal of Research & Method in Education*, 30(1), s. 3–17.
- Scriven, M. (1967). The Methodology of Evaluation. I Tyler, R. W., Gagne, R. M, Scriven, M. (red.): *Perspectives of Curriculum Evaluation: Book The Methodology of Evaluation*. Chicago: Rand McNally.
- Säljö, R. (2009). Learning, Theories of Learning, and Units of Analysis in Research. *Educational Psychologist*, 44(3), s. 202–208.
- Säljö, R. (1997). Talk as Data and Practice – a critical look at phenomenographic inquiry and the appeal to

- experience. *Higher Education Research & Development*, 16(2), s. 173–190.
- Shadish, W. R., Cook, T. D. og Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston og New York: Wadsworth Cengage learning.
- Shulman, L. (2005). Signature pedagogies in the professions. *Daedalus*, 134(3), s. 52–59.
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), s. 153–189.
- Skar, G. B. U. og Aasen, A. J. (2016). Risikotrekk og skjulte kvaliteter i elevtekstar. *Nordic Journal of Literacy Research*, (2), s. 1–20.
- Skaalvik, E. M. og Skaalvik, S. (2008). Trivsel, utbrenthet og mestringsforventning hos lærere: En utfordring for skoleledere. I Andreassen R-A., Skaalvik, E. og Irgens E. (red.). *Skoleledelse. Betingelser for læring og ledelse i skolen*, s. 141–152. Trondheim: Tapir akademisk.
- Smidt, J. (2002). Double Histories in Multivocal Classrooms Notes Toward an Ecological Account of Writing. *Written Communication*, 19(3), s. 414–443.
- Schmidt, F. L. and Rader, M. (1999), Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52, 445±64
- Schmidt, F. L. and Rader, M. (1999), Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52, 445±64
- Sommers, N. (1982). Responding to student writing. *College composition and communication*, 33(2), s. 148–156.
- Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English*, 31(1), s. 91–119.
- Strijbos, J.-W., Martens, R. L., Prins, F. J. og Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education*, 46(1), s. 29–48.
- Stortingsmelding nr. 11 (2008–2009) *Læreren Rollen og utdanningen*. Oslo: Kunnskapsdepartementet. Hentet fra <https://www.regjeringen.no/no/dokumenter/stmeld-nr-11-2008-2009/id544920/>
- Stortingsmelding nr. 31 (2007–2008). *Kvalitet i skolen*. Oslo: Kunnskapsdepartementet. Hentet fra <https://www.regjeringen.no/no/dokumenter/stmeld-nr-31-2007-2008/id516853/>
- Stortingsmelding nr. 16 (2006–2007). ... og ingen to igjen. *Tidlig innsats for livslang læring*. Oslo: Kunnskapsdepartementet. Hentet fra <https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007/id441395/>
- Stortingsmelding nr. 30 (2003–2004). *Kultur for læring*. Oslo: Kunnskapsdepartementet. Hentet fra <https://www.regjeringen.no/no/dokumenter/stmeld-nr-030-2003-2004/id404433/>
- Sugai, G. og Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, 24(1–2), s. 23–50.
- Suppe, F. (1989). The Received View versus the Semantic Conception. I *The semantic conception of theories and scientific realism*, s. 62–72. Urbana og Chicago: University of Illinois Press.
- Svanes, I. K. og Skagen, K. (2016). Connecting feedback, classroom research and Didaktik perspectives. *Journal of Curriculum Studies*, 1–18.
- Taras, M. (2010). Assessment for learning: assessing the theory and evidence. *Procedia-Social and Behavioral Sciences*, 2(2), s. 3015–3022.
- Topping, K. J. (2009). Peer assessment. *Theory into practice*, 48(1), s. 20–27.
- Truscott, J. (1996). The Case Against Grammar Correction in L2 Writing Classes. *Language learning*, 46(2), s. 327–369.
- Tunstall, P. og Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, 22(4), s. 389–404.
- Utdanningsdirektoratet (2017a): <https://statistikkportalen.udir.no/Pages/default.aspx>
- Utdanningsdirektoratet (2017b). *Sensorveiledning, norsk hovedmål*. Hentet fra <https://www.udir.no/eksamen-og-prover/eksamen/sensurere-eksamen/>
- Utdanningsdirektoratet (2017c). Ny generell del av læreplanverket. <https://www.udir.no/laring-og-trivsel/lar>

[eplanverket/fagfornyelsen/ny-generell-del-av-lareplanen/](#)

Utdanningsdirektoratet (2016). Vurdering for læring, 6. samling for pulje 5. Hentet fra

https://www.udir.no/globalassets/filer/vurdering/vfl/presentasjoner/pulje-5/udir-siste-samling_dag_1.pdf

Utdanningsdirektoratet (2014a). *Videreføring av satsingen «Vurdering for læring» 2014–2017*. Hentet fra <https://www.udir.no/Upload/Vurdering/Grunnlagdokument-2014-2017.pdf>

Utdanningsdirektoratet (2014b). Eksamensforskrift. Hentet fra [http://www.udir.no/Vurdering/Eksamen-videregaende/Endringer-og-overgangsordninger/Endringer/Eksamensoppgaver-i-norsk-2014-for-Vg3-og-pabygging-til-generell-studiekompetanse-/](http://www.udir.no/Vurdering/Eksamen-videregaende/Endringer-og-overgangsordninger/Endringer/Eksamensoppgaver-i-norsk-2014-for-Vg3-og-pabygging-til-generell-studiekompetanse/)

Utdanningsdirektoratet (2013). Læreplan i norsk. Hentet fra <https://www.udir.no/kl06/NOR1-05/Hele>

Utdanningsdirektoratet (2010a). Nasjonal satsing på vurdering for læring (2010–2014). Hentet fra <http://www.udir.no/Vurdering-for-laring/Nasjonal-satsing1/Nasjonal-satsing-pa-Vurdering-for-laring>

Utdanningsdirektoratet (2010b). *Individuell vurdering*. (Rundskriv Udir-1-2010). Hentet fra <http://www.udir.no/Regelverk/Finn-regelverk-for-opplaring/Finn-regelverk-etter-tema/Vurdering/Udir-1-2010-Individuell-vurdering>

Utdanningsdirektoratet (2009). *Bedre vurderingspraksis*. Sluttrapport fra Utdanningsdirektoratet 2009. Hentet fra <http://www.udir.no/Tilstand/Forskning/Rapporter/Utdanningsdirektoratet/Bedre-vurderingspraksis-sluttrapport-fra-Utdanningsdirektoratet-2009/>

Vangrieken, K., Dochy, F., Raes, E. og Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, s. 17–40.

Vähäpassi, A. (1988). The domain of the school writing and the development of the writing tasks. I Gorman, T. P., Purves, A. C. og Degenhart, R. E. (red.). *The IEA Study of Written Composition 1: The International Writing Tasks and Scoring Scales*. New York: Pergamon Press.

Vygotsky, L. S. (1978). *Mind in Society. The Development of Higher Psychological Processes*. Cambridge, MA.: Harvard University Press.

Weber, M. (2000). *Makt og byråkrati. Essays om politikk og klasse, samfunnsforskning og verdier*. Oslo: Gyldendal Norsk Forlag.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, 31(3), s. 379–394.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), s. 3–14.

Wiliam, D., Lee, C., Harrison, C. og Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education* 11(1), s. 49–65.

Øygarden, B. (2001). Kvar vart det av stilen? Om skriftlege sjanrar ved avgangsprøva i norsk i grunnskolen 1962 til 1999. I Kulbrandstad, L. I. og Sjølie, G. (red.). *På Hamar med norsk*. Rapport frå konferansen «På norsk med mellomtrinnet» 18.–19. januar 2001. Del 1: Skrivning og lesing. Høgskolen i Hedmark, rapport nr. 11–2001, s. 52–72.

Internettreferanser

Etiske retningslinjer for de nasjonale, forskningsetiske komiteene, hentet fra

<https://www.etikkom.no/Forskningsetikk/Etiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>

Etiske retningslinjer for American Psychological Association, hentet fra

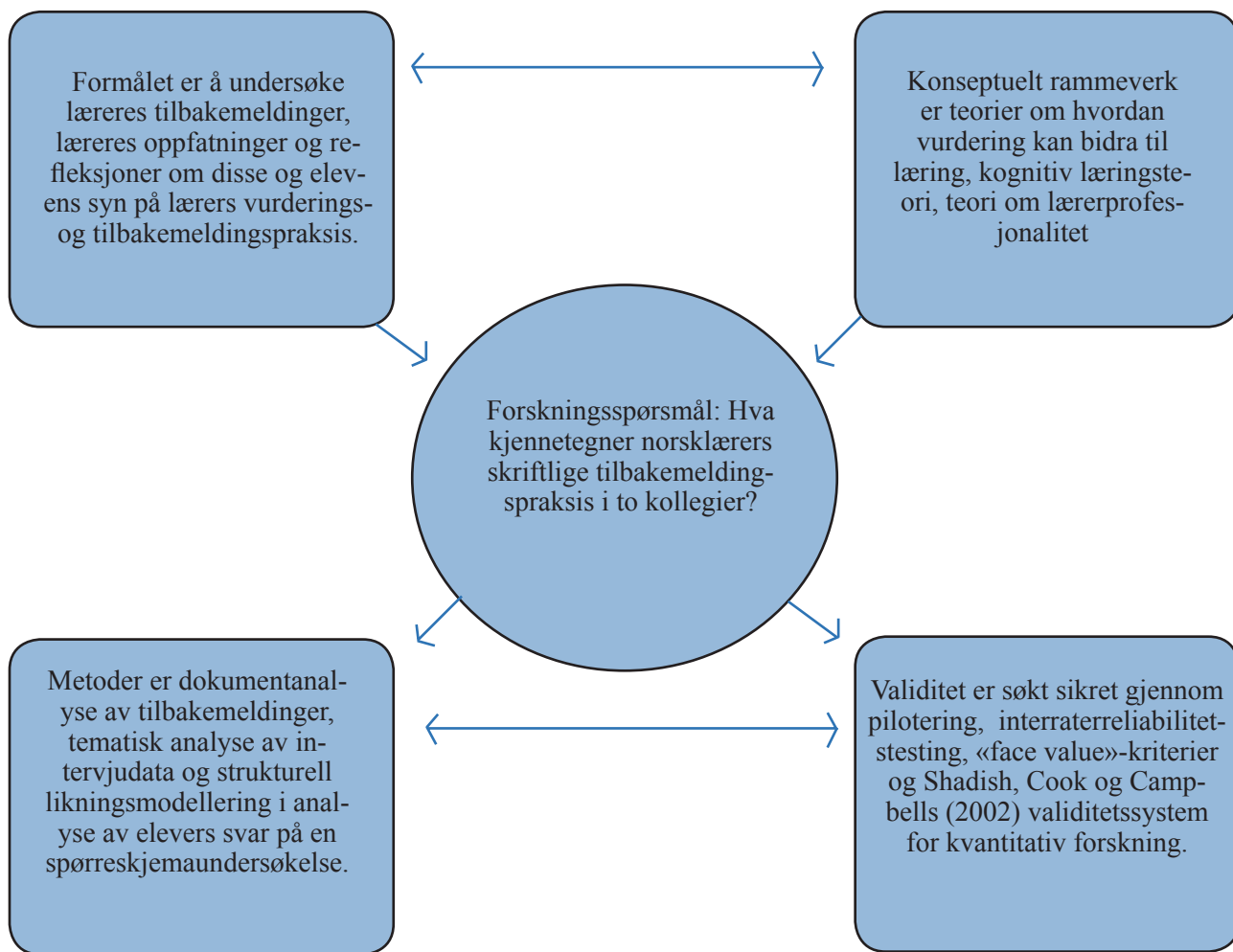
<http://www.apa.org/ethics/code/index.aspx?item=11>

Forskrift til opplæringslova. Hentet fra Lovdata: https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_4#KAPITTEL_4

Vedlegg 1

Metodisk design – etter Maxwell (2005) – en interaktiv modell for forskningsdesign

Mellom formål og metoder og konseptuelt rammeverk og validitet kan en også tenke seg piler som uttrykker en gjensidig og interaktiv forbindelse.



Vedlegg 2. Oversikt over datamateriale og datainnsamling

Tidslinje for datainnsamling

- Høst 2013 og vår 2014: kontakt med skoler, avtaler med skoler og lærere. Godkjenning av prosjektet innhentet fra NSD.
- Høst 2014–vår 2015: datainnsamling via læringsplattform for lærere og klasser.
- Vår 2015, mai: Elevene svarer på spørreskjemaundersøkelse, intervjuer med lærerne.
- Desember 2015–januar 2016: intervju med representant fra ledelse og lærere ved hver skole om arbeidet i fagseksjonen for norsk ved begge skoler.

Tabell 4 Deltakere: lærere og klasser

Skole A	Lærer/pseudonym	Utdanningsnivå	Erfaring	Klasse
	Anne	Master	1 år	ST Vg3, ST Vg1, YF Vg2
	Berit	Bachelor	3 år	ST Vg3, YF Vg2
	Carina	Master	8 år	ST Vg2, YF Vg1
Skole B	Dina	Master	24 år	ST Vg3
	Eva	Master	5 år	ST Vg3, ST Vg2
	Frøy	Master	5 år	ST Vg3, ST Vg1
Sum	6 lærere			12 klasser

Tabell 5 Antall elevtekster med tilbakemeldinger fra lærer totalt

Skole A	Klasse	Antall elever ved skolestart	Antall vurderings-situasjoner lastet ned	Antall elevtekster totalt
Anne	YF Vg2	17	3	23
	ST Vg1	29	7	137
	ST Vg3	31	8	114
Berit	YF Vg2	14	4	45
	ST Vg3	32	6	165
Carina	YF Vg2	19	3	42
	ST Vg2	28	6	121
Skole B				
Dina	ST Vg3	31	5	148
Eva	ST Vg2	30	6	194
	ST Vg3	28	5	129
Frøy	ST Vg1	30	6	150
	ST Vg3	29	9(5)*	236
Sum		319	64	1504

* Frøy lot elever levere flere utkast, slik at tekster ble tatt inn og fikk tilbakemelding 2 ganger i 4 av

5 vurderingssituasjoner.

Tabell 6 Oversikt over datamateriale for artikkel 1 – tekster valg for analyse

Skole A	Lærer/pseudonym	Elever/klasse	Elevtekster med tilbakemeldinger
	Anne	Vg3	32
	Berit	Vg3	32
	Carina	Vg2	24
Skole B	Dina	Vg3	30
	Eva	Vg3	25
	Frøy	Vg3	28
Sum	6 lærere	6 klasser	171 tekster

Totaloversikt over kodefordeling etter analyse av læreres tilbakemeldinger i artikkel 1

Forklaringer: Tekst = i teksten, Slutt = i sluttkommentar, N = antall elevsvar med lærerkommentarer, VFL er Vurdering for læring. Tall indikerer frekvens. Under kategoriene mål, instruks, relasjon, korrigerende av innhold, korrigerende av språk er det også notert spredningstall per lærer.

Tabell 7, skole A

Lærer	Tekst	Slutt	Sum	Mål		Instruks		Relasjon		Korr innhold		Korr språk	
				T	S	T	S	T	S	T	S	T	S
Anne				T	S	T	S	T	S	T	S	T	S
N=32	724	396	1120	82	129	260	99	9	2	92	82	281	84
	VFL	VFL	VFL	0-9	0-9	1-13	0-4	0-3	0-1	0-6	0-5	1-21	1-4
	342	228	570										
Berit	728	119	847	2	30	19	47	0	1	39	30	668	11
N=32	VFL	VFL	VFL		0-3	0-3	0-4			2-6	0-4	7-63	0-1
	21	77	98										
Carina	236	115	351	22	45	111	51	37	3	51	9	15	7
N=24	VFL	VFL	VFL	0-4	0-8	0-6	0-4	0-8	0-1	0-6	0-4	0-2	0-2
	133	96	229										

Tabell 8, skole B

Lærer	Tekst	Slutt	Sum	Mål		Instruks		Relasjon		Korr innhold		Korr språk	
				T	S	T	S	T	S	T	S	T	S
Dina N=30	603 VFL 198	248 VFL 159	851 VFL 357	49 0-6	118 2-8	149 0-7	41 0-2	136 1-9	2 0-1	68 0-8	46 0-4	201 0-19	41 0-3
Eva N=25	577 VFL 318	144 VFL 106	721 VFL 424	131 2-14	73 1-6	187 1-9	33 0-3	37 0-5	3 0-1	107 1-9	6 0-1	115 0-18	29 0-2
Frøy N=28	541 VFL 357	92 VFL 61	633 VFL 418	157 3-12	31 0-4	200 0-7	30 0-3	75 0-4	12 0-2	92 0-7	11 0-2	17 0-3	8 0-2

Tabell 9, Oversikt over datamaterialet for artikkel 2

Skole A	Klasse	Antall elever ved skolestart	Antall svar – disse elevene var til stede den aktuelle skoledagen	Respondenter tatt ut etter mye manglende data/ svarmønster som bryter med reverserte indikatorer
Anne	YF Vg2	17	11	1
Anne	ST Vg1	29	21	1
Anne	ST Vg3	31	24	1
Berit	YF Vg2	14	12	1
Berit	ST Vg3	32	17	1
Carina	YF Vg2	19	6	1
Carina	ST Vg2	28	22	
Skole B				
Dina	ST Vg3	31	19	
Eva	ST Vg2	30	28	1
Eva	ST Vg3	28	20	1
Frøy	ST Vg1	30	14	
Frøy	ST Vg3	29	26	
Sum	12 klasser	319	221	8

Manglende data kan forklares av to forhold: 1) elever som sluttet igjennom skoleåret. Dette forekom ved begge skoler, men jeg hadde ikke tilgang til totalt antall. 2) fravær fra skolen den aktuelle dagen.

Tabell 10, oversikt over datamaterialet for artikkel 3

Skole	Lærere	Type data	Varighet	Bakgrunns-data	Type data	Varighet	Type data	Varighet
A	Anne Berit Carina	Gruppeintervju, transskribert	0.45	Oppgaver, skjemaer, planer, ressurser for undervisningen	Intervju med ledelse, feltnotater	0.20	Intervju med lærer, felt-notater	0.20
B	Dina Eva Frøy	Gruppeintervju, transskribert	1.15		Intervju med ledelse, feltnotater	0.20	Intervju med lærer, felt-notater	0.20

Bakgrunnsdata fra artikkel 3

Tabell 11 viser elevopplutning om utvalgte skriveoppgaver som har et element av prosesskriving ved skole A. IFO betyr individuell fordypningsoppgave, en stor oppgave i norskfaget på ST Vg3.

Vurdering angir type skriveoppgave som elevene hadde fått tilbakemelding på.

En kortere versjon av denne oversikten er også vedlagt artikkel 3.

Tabell 11, oversikt over skriveprosesser ved skole A

Lærer	Klasse	Antall elever i klassen	Vurdering	Antall levert førsteutkast	Andreutkast	Endelig versjon
Anne	3ST	32	IFO	17	4	10
Anne	1ST	29	Kåseri	6		25
Berit	3ST	32	IFO	22 (problemstilling)		23
Carina	2ST	29	Mappe			9

Fellestrekk ved skole A

Skolen har felles heldagsprøver på ST Vg3, de to siste før eksamen. Begge klassene på ST Vg3 skrev individuell fordypningsoppgave, oppgavene var ikke identisk utformet.

Anne

I Annes kåseriprosess unnlot 4 elever å levere tekst, og 5 ble tatt for plagiat. 20 av 29 leverte som de skulle. Det var nesten 100 % oppslutning om heldagsprøver, men klart lavere på prosesskrivingene.

Berit

Berit hadde lite prosesskriving på ST. Den individuelle fordypningsoppgaven var organisert med en skriveoppgave på skolen med oppgaver til bok 1 – denne ble karactersatt. Her leverte 29 av 32.

På YF hadde hun prosesskriving i en klasse med 14 elever. Her leverte 6 førsteutkast og 2 endelig versjon.

Carina

Carina hadde heldagsprøve med førsteutkast og andre utkast på slutten av Vg2 – bare én leverte to ganger. Hun hadde mappevurdering, som innebar at elevene skulle samle 4 tekster i en mappe, tekster som skulle velges ut av det de hadde skrevet gjennom året, de hadde fått tilbakemelding og karakter av lærer, og skulle rette opp/videreutvikle og levere på nytt. Av tabellen ser man at det kun var 9 som leverte.

På YF hadde hun et system for at elevene leverte inn rettet versjon av tekster de hadde fått karakter og tilbakemelding på – cirka 60 % gjorde dette.

Oversikt over bakgrunnsdata, skole B

Tabell 12 viser elevoppslutning om utvalgte skriveoppgaver som har et element av prosesskriving ved skole B. IFO betyr individuell fordypningsoppgave. Vurdering angir hvilken type skriveoppgave elevene fikk, og hvilken de fikk tilbakemelding på.

Tabell 12, oversikt over skriveprosesser ved skole B

Lærer	Klasse	Elever	Vurdering		Antall levert utkast	Antall endelig	Antall rettet
Dina	ST Vg3	31	IFO		16	27	
Eva	ST Vg3	27	IFO			20	
Eva	ST Vg1	31	Skrivedag på skolen			28	11
Eva	ST Vg1	31	Novelleanalyse		29	18	
Frøy	ST Vg3	28	IFO		Ulikt, se under	27	
Frøy	ST Vg1	30	Skrivedag på skolen		22	28	

Fellestrekk ved skole B

Skolen hadde felles heldagsprøver, ST Vg3 (3 klasser), og på ST Vg1 (to klasser), felles struktur på skoleåret for ST Vg3 med tanke på vurderingers innhold og rekkefølge. Individuell fordypningsoppgave hadde veldig likt utformet oppgaveformulering, det vitner om samarbeid.

Dina

Hadde organisert IFO som en skriveprosess, tallene viser til innleveringsmapper i læringsplattformen: 1 problemstilling – halve klassen leverte, lærer kommenterte, 2

handlingsreferat, 3 analyse – lærer kommenterte ikke noen av disse, det kan se ut som innleveringen hadde en kontrollfunksjon, 4 endelig versjon med karakter og kommentarer.

Dina hadde ikke klare systemer for levering av ny versjon av oppgaver. Men i likhet med Frøy hadde hun retorisk analyse som en tretrinns-innleveringsoppgave, som var supplert med en muntlig vurdering til slutt.

Eva

Hadde system for andregangslevering av tekst. Elevene skrev på skolen (28 leverte), de fikk tilbakemelding og skulle levert på nytt, og 11 av 28 gjorde det. Elevene fikk ikke karakter på første versjon, men nivå ble gullet ut – samme som Frøy.

IFO var ikke organisert som en direkte skriveprosess, men det var krav om delrapportering gjennom prosessen, som viste seg som innleveringsmapper. Her var elevenes oppslutning varierende. 20 av 27 leverte endelig versjon i innleveringsmappa.

Novelleanalyse viser samme mønster: 29 leverte førsteutkast, 18 leverte rettet versjon.

Evas og Frøys klasser på ST Vg1 var likt organisert med tanke på vurderingers utforming og tidspunkt.

Frøy

Hun hadde to tilfeller av prosesskriving: Den ene av retorisk analyse med utkast og endelig versjon. 17 elever av 28 leverte to versjoner, 7 bare en sluttversjon, 4 leverte ikke i hele tatt.

IFO var organisert som en spesielt omfattende skriveprosess. Elevene fikk tilbakemelding på alle deloppgaver, som ble levert inn gjennom høsten (tallene etter = angir antall levert) 1. versjon = 20, 2. versjon = 16, 3. versjon = 18, 4. versjon = 15, endelig versjon = 27.

Frøy la i utstrakt grad til rette for bruk av informasjon fra tilbakemelding for å forbedre tekst. På tradisjonelle heldagsprøver ga hun teksten tilbake uten karakter – denne ble holdt tilbake til eleven hadde rettet etter lærers tilbakemelding.

På ST Vg1 var 3 vurderinger skrevet i prosess.

Hun ga ikke karakter direkte, men limte inn skjema med kjennetegn på måloppnåelse i elevenes besvarelse, og gullet ut hvordan eleven lå an på ulike delmål, så de kunne få et bilde av karakteren eller i hvert fall nivået på den måten.

Øvrige dokumenter

På grunn av omfang og av hensyn til anonymisering lar ikke dette materialet seg presentere med full transparens.

Fordi det er lagt ved artikkel 1 og anvendt i analysene i artikkel 3 legges det ved eksempler på vurderingsskjemaer med kjennetegn på måloppnåelse fra begge skoler. Disse omtales også som kriterier, men er altså strengt tatt vurderingsskjemaer med kjennetegn på måloppnåelse som antas å fungere som vurderingskriterier. Lærerne omtaler disse skjemaene som kriterier i intervjudataene, slik jeg forstår det.

Skjema med kjennetegn på måloppnåelse, benyttet ved skole B, ST Vg3 (alle lærere)

Skole	Gruppe	Eksaminandnr		
	1	Karakter 2	Karakter 3 – 4	Karakter 5 – 6
		Eksamenssvaret	Eksamenssvaret	Eksamenssvaret
<i>Hoved- inntrykk</i>		er i noen grad et relevant svar på kortsvarsoppgaven er i noen grad et relevant svar på langsvarsoppgaven viser noe forståelse og bruk av de vedlagte tekstene er enkelt eller uklart formulert	er et stort sett relevant svar på kortsvarsoppgaven er et stort sett relevant svar på langsvarsoppgaven viser at eleven forstår og bruker de vedlagte tekstene er stort sett godt formulert, med hensiktsmessig ordforråd	er et presist og relevant svar på kortsvarsoppgaven er et utfyllende og relevant svar på langsvarsoppgaven viser god forståelse og bruk av de vedlagte tekstene er gjennomgående godt formulert, med nyansert ordforråd

<p>Språk og formelle ferdigheter</p>	<p>har sammenbinding mellom setninger og avsnitt</p> <p>viser at eleven mestrer hovedformene i formverket til en viss grad</p> <p>viser at eleven mestrer ortografi og tegnsetting til en viss grad</p> <p>skiller mellom sitat og egenprodusert tekst og oppgir kilder</p>	<p>har nokså god/god sammenbinding mellom setninger og avsnitt</p> <p>viser at eleven stort sett mestrer formverket</p> <p>har stort sett korrekt ortografi og tegnsetting</p> <p>markerer sitat og viser til kilder på en ryddig måte</p>	<p>har god sammenbinding mellom setninger og avsnitt</p> <p>viser at eleven mestrer formverket</p> <p>har gjennomgående korrekt ortografi og tegnsetting</p> <p>markerer sitat og viser til kilder ryddig og nøyaktig</p>
<p>Kortsvaret</p>	<p>har en viss struktur og sammenheng</p> <p>viser noe faglig kunnskap</p> <p>kommenterer eller argumenterer på en enkel måte</p> <p>peker på noen karakteristiske trekk ved tekster</p> <p>har noe fagspråk</p>	<p>har stort sett hensiktsmessig struktur og god sammenheng</p> <p>viser nokså god/god faglig kunnskap</p> <p>kommenterer eller argumenterer på en klar måte</p> <p>gjør greie for karakteristiske trekk ved tekster</p> <p>har formålstjenlig fagspråk</p>	<p>har hensiktsmessig struktur og god sammenheng</p> <p>viser meget god faglig kunnskap</p> <p>kommenterer eller argumenterer på en presis måte</p> <p>gjør godt greie for karakteristiske trekk ved tekster</p> <p>har presist og formålstjenlig fagspråk</p>

	1	Karakter 2	Karakter 3 – 4	karakter 5 – 6
		Den aktuelle elevteksten	Den aktuelle elevteksten	Den aktuelle elevteksten
Kreative tekster		<p>har en viss struktur og sammenheng</p> <p>har med noen relevante momenter og/eller eksempler</p> <p>viser at eleven har valgt noen innfallsvinkler og språklige virkemidler</p> <p>reflekterer i noen grad</p> <p>bruker noen enkle litterære virkemidler i utformingen av teksten</p> <p>viser i noen grad til språklig/kulturhistorisk kunnskap dersom det er relevant</p> <p>bruker vedlagte tekster på en enkel måte</p>	<p>har stort sett hensiktsmessig struktur og god sammenheng</p> <p>har relevante momenter og/eller eksempler</p> <p>viser noe kreativitet i valg av innfallsvinkler og språklige virkemidler</p> <p>reflekterer, prøver ut tanker</p> <p>bruker litterære virkemidler i utformingen av teksten</p> <p>viser til og bruker språklig/kulturhistorisk kunnskap dersom det er relevant</p> <p>bruker vedlagte tekster på en selvstendig måte</p>	<p>har hensiktsmessig struktur og god sammenheng</p> <p>har relevante momenter og/eller eksempler, med god bredde</p> <p>viser kreativitet i valg av innfallsvinkler og språklige virkemidler</p> <p>reflekterer, prøver ut tanker på en selvstendig måte</p> <p>bruker litterære virkemidler på en gjennomført måte</p> <p>viser til og reflekterer over språklig/kulturhistorisk kunnskap dersom det er relevant</p> <p>bruker vedlagte tekster på en kreativ og selvstendig måte</p>
Samlet vurdering av kortsvar og langsvaer		<p>Samlet sett viser elevsvaret</p> <p>lav kompetanse i faget</p>	<p>Samlet sett viser elevsvaret</p> <p>nokså god kompetanse i faget</p> <p>eller</p> <p>god kompetanse i faget</p>	<p>Samlet sett viser elevsvaret</p> <p>meget god kompetanse i faget eller</p> <p>framifrå kompetanse i faget</p>

Eksempel på vurderingsskjema, benyttet ved skole A (Berit)

Bildet er minimalt beskåret for å bevare anonymitet.

Bruk de delene av skjemaet som passer til oppgaven.

Analyser, tolkninger, sammenlikninger		Karakter 2 Den aktuelle elevenksten	Karakter 3 - 4 Den aktuelle elevenksten	Karakter 5 - 6 Den aktuelle elevenksten	Kommentar
Analyser, tolkninger og sammenlikninger	har en viss struktur og sammenheng	<input checked="" type="checkbox"/>	har stort sett hensiktsmessig struktur og god sammenheng	har hensiktsmessig struktur og god sammenheng	Da må jeg fram at du vurderer om det er riktig at folk vurderer d. eleven sin eller hvem de snakker i eller hvor de er fra. Det kommer ikke fgs fram at du har hva teksten handler
	har med noen relevante momenter og/eller eksempler	<input checked="" type="checkbox"/>	har relevante momenter og/eller eksempler	har relevante momenter og/eller eksempler, med god bredde	
	bruker noe fagspråk	<input checked="" type="checkbox"/>	bruker relevant fagspråk	bruker presist og relevant fagspråk	
	peker på noen virkemidler i tekster	<input checked="" type="checkbox"/>	gjør greie for noen virkemidler og viser i noen grad hvordan de fungerer i teksten	gjør greie for virkemidler og viser hvordan de fungerer i teksten	
	peker på noen sider ved argumentasjon/bruk av appellformer	<input checked="" type="checkbox"/>	gjør greie for hovedtrekk i argumentasjon/bruk av appellformer og funksjonen de har	gjør presist greie for argumentasjon/bruk av appellformer og funksjonen de har	
peker på noen sider ved budskap/hovedsynspunkt	<input checked="" type="checkbox"/>	gjør greie for tema/budskap/hovedsynspunkt	gjør greie for tema/budskap/hovedsynspunkt på en selvstendig måte		
reflekterer over tekster på en enkel måte, har noen vurderinger	<input checked="" type="checkbox"/>	reflekterer over tekster på en faglig relevant måte, har vurderinger og grunnleggende synspunkter	reflekterer over tekster på en selvstendig måte, har vurderinger og grunnleggende synspunkter		
sammenlikner noen sider ved tekster	<input checked="" type="checkbox"/>	sammenlikner tekster på en oversikkelig måte	sammenlikner tekster på en oversikkelig og god måte		
viser en viss kjennskap til den kulturhistoriske sammenheng	<input checked="" type="checkbox"/>	viser at eleven kan plassere tekster kulturhistorisk, med noe forståelse for sammenheng	viser at eleven kan plassere tekster kulturhistorisk, med forståelse for sammenheng		
bruker eventuelle kilder på en delvis relevant måte	<input checked="" type="checkbox"/>	bruker eventuelle kilder på en relevant og etterprøvbart måte	bruker eventuelle kilder relevant, etterprøvbart og selvstendig		
Samlet vurdering av kort-svar og langesvar	Samlet sett viser eleven var et lav kompetanse i faget	Samlet sett viser eleven var et nokså god kompetanse i faget eller god kompetanse i faget	Samlet sett viser eleven var et meget god kompetanse i faget eller framfrå kompetanse i faget	Karakterforslag 2	

Eksempel på forenklet skjema, benyttet ved skole A, ST Vg2 (Carina)

Konkretisering av kompetansemålene			
Kjennetegn for måloppnåelse og vurderingskriterier.			
Kriterier:	Lav grad av måloppnåelse (1-2) Svaret viser at eleven kan	Middels grad av måloppnåelse (3-4) Svaret viser at eleven kan	Høy grad av måloppnåelse (5-6) Svaret viser at eleven kan
Innhold	Teksten svarer delvis på oppgaven, men bruker i liten grad relevante virkemidler innenfor valgt sjanger. Teksten viser i liten grad relevant fagkunnskap for oppgaven. Eks. Litteraturhistorie i oppg. 3.	Teksten svarer på oppgaven og bruker noen relevante virkemidler innenfor valgt sjanger. Teksten viser relevant fagkunnskap for oppgaven.	Teksten besvarer oppgaven på en hensiktsmessig måte og bruker relevante virkemidler for valgte sjanger. Teksten bruker fagkunnskap for å utdype eller fremheve budskapet i teksten.
Oppbygging	Teksten følger i liten grad struktur for valgt sjanger.	Teksten følger til en viss grad struktur for valgt sjanger	Teksten har en god og hensiktsmessig struktur for valgt sjanger.
Språk og formelle ferdigheter	-mestre hovedformene i formverket til en viss grad og skrive sammenhengende -til en viss grad mestre ortografi og tegnsetting -bruke et enkelt ordforråd og i noen grad kunne variere setningskonstruksjoner	-mestre hovedformene i formverket, ha stort sett god tekstbinding -stort sett mestre ortografi og tegnsetting -bruke et variert ordforråd og varierte setningskonstruksjoner	-mestre formverket godt, ha god tekstbinding -mestre ortografi og tegnsetting -bruke et stort og variert ordforråd og varierte setnings-konstruksjoner
Kilder	- skille mellom sitat fra kilder og egenprodusert tekst. Dersom dette mangler er det ikke grunnlag for vurdering av hele eller deler av teksten!		
Hovedinntrykk	-skrive en tekst som kommuniserer og fungerer til en liten grad som en helhet Teksten svarer lite relevant på oppgaven som er gitt.	-skrive en tekst som stort sett fungerer som en helhet Teksten svarer stort sett på det oppgaven spør etter.	-skrive en tekst som fungerer svært godt som en helhet med god balanse mellom de ulike delene. Teksten er et utfyllende, relevant og kreativt svar på det oppgaven spør etter.

Vedlegg 3

Eksempel på tematisk analyse anvendt i artikkel 3

Bildet under viser 'noder' som er utledet av den tematiske analysen. Nedenfor er «bruk av tilbakemelding» vist som et eksempel på hvordan tekstavsnitt fra transskriberte intervjuer er kodet inn i de ulike nodene. Nodene er som beskrevet i artikkel 3 utviklet på grunnlag av både intervjuguide og av funn som ble gjort underveis i kodeprosessen. Eksempelen illustrerer hvordan den overordnede kategorien etter hvert ble delt inn i underkategorier som beskriver ulike dimensjoner ved bruk av informasjon i tilbakemeldinger. Underkategoriene er dels utledet av intervjuguiden (egen- og kameratvurdering og tidspunkt for tilbakemelding) og en utviklet underveis (prosess).

Name	Sources	References
Bruk av fagbegreper	0	0
Bruk av tilbakemelding	2	15
Egenvurdering	2	2
Kameratvurdering	2	5
Prosess	2	11
Tidspunkt for tilbakemelding	2	3
Elevforutsetninger	2	15
Karakterer	2	18
Kriterier	2	6
Lærers arbeidsmengde	2	9
Personer	0	0
Anne	1	35
Berit	1	43
Carina	1	24
Dina	1	34
Eva	1	37
Frøy	1	42
Produsere tekst	1	4
Samarbeidskultur i norskseksjonen	2	8
Tilbakemld på personnivå	2	4

«Sources» viser til de transskriberte intervjuene – de kommer enten fra skole A eller B, og markeringene som står som 1 viser at kodene bare kommer fra en skole, 2 fra begge. «References» viser til antall tekstavsnitt som er kodet inn i en node. Noen tekstavsnitt kan handle om flere ting for eksempel sammenhengen mellom bruk av tilbakemelding og karakterer, og da er tekstavsnittet kodet inn begge steder.

Nedenfor vises eksempelutsagn knyttet til kategorien bruk av tilbakemelding – tekstavsnittene legges under hverandre til høyre, her er bruk av tilbakemeldinger det som er framhevet, slik fanen høyt på siden viser.

Nodes		
Name	Sources	References
Bruk av fagbegreper	0	0
Bruk av tilbakemelding	2	15
Egenvurdering	2	2
Kameratvurdering	2	5
Prosess	2	11
Tidspunkt for tilbakemelding	2	3
Elevforutsetninger	2	15
Karakterer	2	18
Kriterier	2	6
Lærers arbeidsmengde	2	9
Personer	0	0
Produsere tekst	1	4
Samarbeidskultur i norskseksjonen	2	8
Tilbakemld på personnivå	2	4

00:22:55.02] Interviewer: hvilket inntrykk har dere, dere har i og for seg svart litt på det, men hvilket inntrykk har dere av hvor flinke elevene er til å bruke eller gjøre nytte av de tilbakemeldingene dere gir dem?

[00:23:11.01] Speaker Name: Dina vi må gjøre noe med dem for at de skal dra nytte av dem, vi må også sånn som disse vekstpunktene vi gir, da kan vi - sier ofte - ta utgangspunkt i vekstpunktene du fikk i forrige tekst, når du skriver neste, og gjerne ha dem med deg hvis du er på skolen, på eget ark og se at dette her må jeg forbedre, men også ikke bare vekstpunktene men også de sterke sidene må du også ta med deg og det du gjorde bra sist må du også gjøre på samme måte denne gangen sånn at det ikke bare blir fokus på det negative, det blir jo veldig feil det også, men hvis ikke.. vi må gjøre noe med det for at de skal gjøre noe med det, hvis vi ikke snakker om de veldig sterke, arbeidssomme elevene som gjør denne jobben alene. ja eller at det er en annen motivasjonsfaktor inne i bildet f.eks. hvis det er en vipposituasjon mot slutten av skoleåret nå så har jeg opplevd at flere som har gått tilbake på de tidligere vurderingene for å se hva hun at jeg trengte da? for å få den karakteren jeg ønsker og så har de selv oppsøkt dem på den måten [00:24:12.04] Speaker Name: Eva je ger helt enig i at det veldig opp til oss hva vi gjør med dem og det er få elever som klarere å nyttegjøre seg dem på egenhånd, og jeg tenker at det er også viktig i neste runde at vi gjentar det, forrige gang så var du ganske dårlig på å skrive innledning, men den innledningen du har skrevet nå, den fungerer veldig bra for her har du gjort sånn og sånn og det og det, og det krever jo litt av oss, å holde oversikt over hva den enkelte får til i den enkelte vurderingssituasjon, men noen ting husker vi jo i hvert fall hvis det er elever vi har hatt i flere år, så vi kjenner dem godt etter hvert [00:24:47.22] Speaker Name: Dina så noen ganger går vi jo tilbake til vekstpunktet i forrige tekst, og kommenterer vekst eller ikke vekst i forhold til forrige tekst du skrev [00:24:52.26] Speaker Name: Frøy de trenger jo hjelp til å se sin egen utvikling og - jeg har jo faktisk blitt mye bedre til å skrive innledninger - det er jo da det er viktig å gjøre de samme type øvelsene, det er litt som på trening, hvis du ikke har, hvis du driver og pumper jern (ler) så jobber du med kondisjon så er det ikke så rart at du ikke blir sterkere neste gang du pumper jern, men hvis en får tilbakemeldinger på det samme så tror jeg man blir mer motivert.

Del 2

Artikkel 1

Vurdering for læring i norskfaget: Om lærerens skriftlige tilbakemeldinger på skriftlige tekster

Sammendrag

Vurdering for læring (VfL) er et stadig voksende forskningsfelt, og det har hatt betydelig innflytelse på norsk utdanningspolitikk gjennom både lovendringer og større satsinger der det er investert mye i undervisningsressurser, etterutdanning av lærere og lokalt forankrete utviklingsprosjekter. Det er imidlertid lite forskning på hvordan denne satsingen manifesterer seg i praksis, og spesielt med hensyn til VfL i norskfaget. I denne artikkelen presenteres resultater av en undersøkelse av seks norsklæreres skriftlige tilbakemeldinger til egne elever på 171 skriftlige heldagsprøver i norsk, gitt tilbake til elevene med cirka ni uker igjen av skoleåret. Tilbakemeldingene er kvalitativt analysert og kategorisert i en eksplorerende studie. Funnene indikerer at lærerne langt på vei gir elevene tilbakemeldinger som er i tråd med anbefalingene fra sentral VfL-forskning om hva som kjennetegner formative tilbakemeldinger, men at det samtidig er betydelig variasjon.

Nøkkelord: Vurdering for læring, vurdering av læring, skriftlig norsk, lærers tilbakemelding, formativ vurdering, skriveopplæring

Bakgrunn

Vurdering for læring (VfL) er et konsept som er utviklet over tid, og sentrale forskere, som Ramaprasad (1983) og Black og Wiliam (1998; 2009), legger avgjørende vekt på at kunnskap om læring som en tilegner seg gjennom vurdering, brukes i det videre læringsarbeidet. I tillegg er en sentral komponent at den lærende får tilbakemelding om hvordan han/hun står i forhold til målene for opplæringen (Sadler, 1989). Tilbakemeldingen bør utformes slik at den bidrar til å redusere avstanden mellom ønsket mål og det den lærende har prestert. Hattie og Timperley (2007) legger blant annet vekt på Sadlers (1989) forståelse når de definerer hva som kjennetegner tilbakemeldinger av høy kvalitet, og Parr og Timperley (2010, s. 70) har i forlengelsen operasjonalisert kjennetegn på formative tilbakemeldinger innenfor rammene av en VfL-kontekst. De formulerer det slik: «Formative assessment has been reframed as a social, collaborative activity, aligned more with learning. The emphasis has shifted to the teacher and the students,

working in partnership to enhance student learning». Denne forståelsen, som er forankret i en skriveopplæringskontekst, legges også til grunn her. Begrunnelsen for den foreliggende studien er i forlengelsen av dette at det har betydning å studere *innholdet* i tilbakemeldinger fordi de har et læringsfremmende potensial, og at sjansen for at tilbakemelding kan brukes i videre læringsarbeid er større hvis den er utformet etter prinsippene for formativ vurdering. Hvordan disse prinsippene er operasjonalisert, ut fra Parr og Timperley (2010), blir gjort rede for i metodedelen.

Studier av læreres tilbakemeldingspraksiser innen språkopplæring er det mange av, særlig innenfor andrespråk- og fremmedspråkopplæringen (Ferris, 2004; Lee, 2004; Truscott, 1996), men mange studier retter seg mot høyere utdanning (Dysthe, 2011; Sadler, 1989) eller mot grunnskolen.

Forskningen på læreres tilbakemeldinger i skriveopplæringen har ofte hatt et fokus på grammatikk (Ferris, 2004; Lee, 2004) eller de har hatt et helhetsperspektiv der man har kategorisert ulike talehandlinger (Magno & Amarles, 2011) og ulike nivåer av språklig korrigering (Bergman-Claeson, 2003; Lee, 2007). Det er derimot færre studier som retter seg mot videregående skole og mot vurdering for læring i førstespråkopplæringen.

Utformingen av tilbakemeldingen er av betydning fordi en uklar tilbakemelding i liten grad har potensial til å kunne brukes i læringsarbeid (Weaver, 2006). I forskning på læreres skriftlige tilbakemeldinger fant for eksempel Sommers (1982) at studenter ofte fikk vage tilbakemeldinger, noe som delvis bekreftes av Kronholm-Cederberg (2009), som fant at finske gymnaselever hadde problemer med å forstå hva læreren skrev til dem. Også på dette grunnlaget er det viktig å få kunnskap om hva norsklærere skriver til egne elever. I en studie av tilbakemeldinger til masterstudenter argumenterer Dysthe (2011) for at tilbakemeldinger kan ha læringspotensial om de retter seg mot dypereliggende forhold i teksten, selv om teksten ikke skal revideres. Dysthes studie (2011) har relevans for norskfaget i skolen fordi lærernes tilbakemeldinger kan være et redskap i skriveopplæringen, og tekstlige trekk har likhetstrekk og dermed overføringsverdi fra gang til gang, hvis denne informasjonen gjøres forståelig og tilgjengelig for elevene. I den foreliggende studien er det omtrent ni uker igjen av skoleåret for elevene, og de har dermed potensielt sett tid til å nyttiggjøre seg informasjon fra læreren. Om de faktisk har gjort det, er imidlertid ikke undersøkt.

VfL knyttet til skriveopplæring

I en sentral newzealandsk studie (Parr & Timperley, 2010) analyseres 59 grunnskolelæreres tilbakemeldinger på en elevtekst i en VfL-kontekst. Forskerne finner at de lærerne som best

viser evne til å gi VfL-tilbakemeldinger, har elevene med størst læringsframgang, målt i score på en nasjonal skriveprøve i førstespråket. Parr og Timperleys (2010) studie har et eksperimentelt design, og resultatene har derfor bare indirekte overføringsverdi fordi det ikke er lærernes daglige praksis som er analyseobjekt. Likevel er den relevant, fordi den kombinerer skriveopplæring med VfL gjennom å kategorisere førstespråklæreres tilbakemeldinger til en elevtekst. Den er også metodisk overførbar ved at kategoriene forskerne har utviklet, har relevans i norsk skriveopplæringsammenheng, noe som blir gjort rede for i metodedelen.

En mer læringsfremmende vurderingspraksis har vært et satsingsområde for norske myndigheter over en lengre periode, men sluttrapporten til Forskning på individuell vurdering i skolen (FIVIS) framhever at det ikke synes å være noen klar, etablert forståelse i skolen av hvordan vurdering kan bidra til økt mestring hos elevene (Sandvik & Buland, 2014). I lys av dette funnet er norsklæreres skriftlige tilbakemeldinger til egne elever i den foreliggende studien undersøkt i et dobbelt perspektiv. Det ene er innenfor et VfL-rammeverk som i en norsk kontekst blant annet har kommet til uttrykk gjennom prosjekter som Bedre vurderingspraksis (2007–2009) og Vurdering for læring (2010–2014). Prinsippene for en mer læringsfremmende vurderingspraksis, som fikk lovstatus i 2010 (Udir, 2010), er at 1) elevene kjenner målet med opplæringen, 2) de får informasjon om hvordan de har prestert i forhold til målet, 3) de får informasjon om hvordan de skal komme seg nærmere målet og 4) de er involvert i eget læringsarbeid ved å vurdere eget arbeid og utvikling (jf. Hopfenbeck, Tolo, Florez & El Masri, 2013). I min studie undersøkes særlig om elevene får informasjon om hvordan de har prestert og om hvordan de skal komme nærmere målet. Det legges til grunn at tilbakemeldingen fra læreren er en av de viktigste kildene til denne informasjonen. Et sentralt spørsmål er derfor hvordan denne tilbakemeldingen er utformet. Det andre perspektivet kaster lys over hva som ellers kjennetegner skriftlige tilbakemeldinger lærere gir til elevene. Denne siden av undersøkelsen bygger på tidligere studier av læreres skriftlige tilbakemeldinger (Bergman-Claeson, 2003; Lee, 2007; Magno & Amarles, 2011).

I norsk kontekst er det lite forskning på norsklæreres faktiske tilbakemeldinger til egne elever i videregående skole. I stedet rapporterer f.eks. Sandvik et al. (2014) om læreres oppfatninger av skolens og deres egen vurderingspraksis på bakgrunn av dokumenter, observasjons- og intervjudata, men ikke faktiske tilbakemeldinger. Det er ikke opplagt at lærernes egenrapportering og skriftlige praksis samsvarer. Hertzberg og Roe (2016) framhever at VfL er en viktig satsing på mange skoler, men det er ikke tilstrekkelig kunnskap om hvordan VfL er operasjonalisert i de enkelte fagene og

heller ikke i hvilken grad dette skjer på den enkelte skole. Hayward, Priestley og Young (2004) viser at forskningsbaserte utdanningsreformer har en kompleks vei fra initiering til praksis.

Et eksempel på et grunnskolerettet tiltak innen norsk språkopplæring er Norm-prosjektet (Matre & Solheim, 2015; Berge, Evensen & Thygesen, 2016; Evensen, Berge, Thygesen, Matre & Solheim, 2016), som fra ulike perspektiver har jobbet med å utvikle tekstnormer og standarder for reliabilitetstesting av vurdering av skriving (Jølle, 2014; Berge & Skar, 2015). Det finnes også casestudier av enkeltlæreres skrivepraksis (Igland, 2008) og av elevers oppfatninger av læreres tilbakemeldingspraksiser i ungdomskolen (Bueie, 2015a; 2015b; 2016), men ingen av disse ser spesifikt på VfL i norskfaget i videregående skole.

Min studie søker å gi svar på hvordan norsklærere håndterer oppgaven med å utforme skriftlige VfL-tilbakemeldinger i norskfaget. På denne måten er den med på å gi innsikt i hvordan en VfL-komponent, tilbakemeldingen, kan operasjonaliseres innen et fagfelt og bidra med innsikt i spørsmålet om hvordan norsklærere, i sitt daglige virke, løser oppgaven med å gi tilbakemeldinger til egne elever. De overordnede prinsippene for hvordan vurdering kan bidra til læring er generelt formulert, og de bør derfor forankres i konkrete fag- og skolekontekster (Bennett, 2011; Black & Wiliam, 1998).

Parr og Timperleys (2010) forståelse av VfL ligger til grunn for operasjonaliseringen, som også er i samsvar med opplæringslovens definisjon av underveisvurdering og de endringer som Stortingsmelding nr. 28 (2015–2016) løfter fram, og som vektlegger at formativ vurdering er aktiviteter som skal bidra til læring. Tilbakemeldingen er her forstått som en slik aktivitet, og det er i denne betydningen begrepet brukes her. Forskningsspørsmålet blir derfor: I hvilken grad gir lærerne skriftlige tilbakemeldinger på skriftlige prøver i norskfaget som er i tråd med premissene for formative tilbakemeldinger, slik de er uttrykt i internasjonal forskning?

Metode

Forskingsspørsmålet er belyst gjennom en kvalitativ, eksplorerende studie, som gjennom dokumentanalyse kategoriserer lærernes tilbakemeldinger og kvantifiserer dataene for å gi oversikt over et komplekst materiale.

Deltakere og utvalg

Studien er basert på data fra to videregående skoler på Østlandet. Skole A kombinerer yrkesfaglige og studiespesialiserende studieretninger. Den har cirka 600 elever og er geografisk plassert i et område med antatt lav sosioøkonomisk status. Det var ingen nedre grense for inntak på studiespesialiserende studieretning høsten 2014, det vil si at alle som søkte, kom inn. Skole B har cirka 650 elever og ligger i et område med antatt høy sosioøkonomisk status. Den har en mer homogen elevgruppe enn skole A. Den har studiespesialiserende studieretning med et inntakssnitt på 4,0 høsten 2014.

Den siste heldagsprøven på bokmål på slutten av skoleåret ble valgt som analyseobjekt (studiespesialiserende studieretning). Utvalget ble gjort med tanke på å få inn sammenliknbare data, det vil si noenlunde korresponderende oppgaver, teksttyper og tekstlengder på tvers av skoler og klasser. Prøven egner seg til analyse fordi den er gitt på slutten av opplæringen for fem av klassene og slutten av skoleåret for én klasse, og det er derfor stor sjanse for at den har betydning for elevenes standpunktkarakter, noe alle klassene har felles. Dermed kan en tenke seg at læreren i sine skriftlige tilbakemeldinger vil vektlegge begrunnelse av karakteren, og at det kan være vanskeligere å finne forekomster av tilbakemeldinger utformet for å bidra til læring. Spørsmålet er da om lærerne prioriterer å gi denne typen tilbakemeldinger, og en kan anta at hvis de gjør det på dette tidspunktet, er det sannsynlig at de også gjør det ellers i året. Prøvesituasjonen kan således fungere som en kritisk case (Patton, 2014), det vil si at den kan kaste lys over lærerens skriftlige tilbakemeldingspraksis gjennom året. Heldagsprøvene er for øvrig en del av et større datamateriale samlet inn i skoleåret 2014–2015.

Valget av heldagsprøven reiser spørsmålet om forholdet mellom formativ og summativ vurdering. Forståelsen er teoretisk forankret i Newton (2007), som framhever at begrepet summativ vurdering viser til en vurderingsform som oppsummerer resultater, mens formativ vurdering viser til spørsmålet om hva resultatene skal brukes til. I dette tilfellet undersøkes om lærernes tilbakemelding er utformet på en slik måte at de har større potensiale for å bidra til læring.

Oversikt over datagrunnlaget for studien (Tabell 1)

Skole A	Lærer/pseudonym	Utdanningsnivå	Erfaring	Elever/klasse	Elevtekster med tilbakemeldinger
	Anne	Master	1 år	Vg3	32
	Berit	Bachelor	3 år	Vg3	32
	Carina	Master	8 år	Vg2	24
Skole B	Dina	Master	24 år	Vg3	30
	Eva	Master	5 år	Vg3	25
	Frøy	Master	5 år	Vg3	28

Kodeskjema

Et sentralt spørsmål er hvordan et teoretisk konstrukt som en læringsfremmende tilbakemelding kan operasjonaliseres. Et svar finnes i studien til Parr og Timperley (2010), som har kategorier som de knytter til kvaliteten på tilbakemeldingene. Disse danner grunnlag for det kodeskjemaet som er utarbeidet for å kartlegge forekomsten av VfL-tilbakemeldinger også i denne studien. I tillegg bygger skjemaet på studier av Bergman-Claeson (2003) og Magno og Amarles (2011), for å kunne gi et mer helhetlig og nyansert bilde av lærernes skriftlige tilbakemeldingspraksis. To av lærerne har dessuten et fargekodesystem for språkfeil (se appendiks A).

Kodeskjemaet kartlegger i tillegg hvor lærerens kommentar er plassert (Bergman-Claeson, 2003); en sluttkommentar vil kunne få et mer indirekte preg enn en kommentar skrevet rett inn i elevteksten. Muligheten for at eleven misforstår lærers tilbakemelding blir større hvis den er skrevet til slutt (Parr & Timperley, 2010, s. 76), og derfor kan denne dimensjonen si noe om i hvilken grad en tilbakemelding har potensiale for å kunne bidra til læring.

I det følgende presenteres en utdyping av hvilket rasjonale som ligger til grunn for hver kategori.

Kategori 1 – Mål

Parr og Timperley (2010) operasjonaliserer kategorien ved å undersøke om lærerens tilbakemelding kommenterer om eleven har nådd målet med skriveøkten og om tilbakemeldingen er direkte relatert til målet eller om den er generelt formulert. I hvilken grad tilbakemeldingen bruker evidens fra elevteksten til å si noe om prestasjonen med tanke på formålet med skrivesituasjonen i en retorisk kontekst, er Parr og Timperleys (2010) siste målrelaterte kategori. Skillet mellom mål og hvordan eleven har prestert i forhold til målet, er her slått sammen fordi en påpeking av hva som er målet, som er skrevet inn i elevens tekst, vil synliggjøre avstanden mellom prestasjon og ønsket prestasjon.

Målet med opplæringen er komplekst. Sadler (1989) viser til 50 mulige kriterier for hva som kan regnes som en god tekst, og da er ikke tekstens innhold en del av bildet. For å operasjonalisere denne kategorien er det blant annet tatt utgangspunkt i vurderingskriterier som er knyttet til oppgavesettene elevene har skrevet ut fra, og som har klare likhetstrekk med sensurskjemaet for norskeksamen (se vedlegg 2 og 3). I tillegg er henvisninger til fagbegreper og forhold som har overføringsverdi til senere skrivesituasjoner (Dysthe, 2011), kodet inn her. Kategorien korresponderer med de overordnede prinsippene for hvordan tilbakemeldinger kan bidra til læring (jf. Hopfenbeck et al., 2013) ved at det vektlegges at de må relatere til målet med opplæringen i den hensikt å redusere avstanden mellom prestasjon og ønsket prestasjon (Sadler, 1989).

En typisk kommentar kan vise til fagbegrepet hovedsyn, som står sentralt i analyse av saktekster, en svært vanlig oppgavetype på eksamen: «Tenker du at dette er hovedsynet? Det må du i tilfelle si litt klarere.» (sitert Eva). Et annet eksempel viser til fagbegrepet metafor: «Eksempel på metafor for døden som du bør kommentere!» (sitert Carina). Å kjenne til og kunne bruke fagbegreper er et sentralt mål i faget, slik det er uttrykt i både læreplan og skjemaet med kjennetegn på måloppnåelse som elevene ble vurdert ut fra i den aktuelle skrivesituasjonen. Begge tilbakemeldingene har overføringsverdi til senere skrivesituasjoner fordi de peker på forhold som står sentralt neste gang eleven skal skrive en liknende oppgave.

Kategori 2 – Instruks

Den andre kategorien rommer konkrete henvisninger til hvordan eleven kan forbedre teksten, det vil si komme nærmere målet med opplæringen. Hos Parr og Timperley (2010) regnes tilbakemeldingen å ha høyere kvalitet jo mer spesifikk den er. På bakgrunn av dette er modellering og forklaringer fra lærerens side kodet som instruks når de gir informasjon om *hvordan* eleven kan utvikle ferdigheter, bruke fagbegreper og håndtere tekstmessige forhold som det er sannsynlig at elevene møter igjen. En instruks kan se ut som dette: «I en innledning skal du ALLTID presentere tittel, forfatter og hvor og når teksten ble publisert.» (sitert Eva).

Det er også lagt til grunn at en instruks inneholder en implisitt oppfordring til bruk, slik at imperativer er oppfattet som instruks («husk at du må ...»), og det samme gjelder spørsmål som i mange tilfeller inviterer eleven til en dialog om tekstens innhold eller form. Samtidig kan spørsmål også ha korrigerende funksjoner, så kodeplasseringen er vurdert ut fra kontekst.

Kategori 3 – Relasjonsbyggende kommentar

Som en egen kategori er person- eller prestasjonsrettet ros tatt med. Den kaster lys over lærerens relasjonsbyggende arbeid (Skaalvik & Skaalvik, 2011). Det legges til grunn at denne formen for tilbakemelding er skrevet for å bidra til å motivere eleven i det videre læringsarbeidet (Dweck, 2002), siden relasjonen mellom lærer og elev regnes som viktig for elevens arbeidsinnsats og prestasjoner. Denne kategorien faller utenfor rammene av VfL, fordi den inneholder for lite informasjon om hvordan eleven skal komme nærmere målet, men det kan godt tenkes at den bidrar positivt i læringsarbeidet fordi den kan påvirke elevens evne til selvregulering. Den kan imidlertid også bidra til å trekke oppmerksomhet bort fra oppgaven og mot person, noe som reduserer muligheten for læring (Hattie & Timperley, 2007).

Kategori 4 – Korrigering

Språklige korrigeringer er kodet i en egen kategori. Det er lagt til grunn at eleven må jobbe med språkkorrigeringene for at de skal bidra til læring, og at det er av begrenset verdi å vite at man gjorde feil i forrige skriveoppgave. Det bygger ikke på en oppfatning om at språkkorrigeringer ikke kan bidra til læring, men at de i liten grad er utformet etter anbefalingene om hva som kjennetegner en god tilbakemelding fordi de inneholder lite informasjon til eleven.

Det samme gjelder innholdskorrigeringer knyttet til den lokale skrivesituasjonen; det er mindre sjanse for at eleven lærer, fordi oppgaven er et tilbakelagt stadium når eleven får tilbakemelding av læreren. Med mindre eleven skal jobbe med den samme oppgaven igjen, blir slike tilbakemeldinger lite effektive. Hvis eleven derimot skal jobbe videre med den, vil en slik tilbakemelding kunne bidra til læring (Hattie & Timperley, 2007).

Et eksempel på en innholdskorrigering er at læreren skriver: «Du må forklare litt bedre hvordan en oljeplattform symboliserer at Nord-Norge er øverst og Sør-Norge nederst. Jeg ser ikke sammenhengen» (sitert Eva). Her legges det til grunn at eleven ikke vil få noen ny oppgave som går ut på å analysere den samme sammensatte teksten, altså med det samme innholdet, og responsen retter seg derfor mot stoff som eleven er ferdig med.

En forenklet versjon av analyseskjemaet er presentert i tabell 2. Eksemplene på tilbakemeldinger ble skrevet inn i elevens tekst som merknader, slik at det foreligger en direkte kobling mellom elevens ytring og lærers tilbakemelding. For en mer nyansert versjon av skjemaet og oversikt over funn, se appendiks A og B.

Kjennetegn på kategorier og eksempler på operasjonalisering (Tabell 2)

	VfL	VfL	Ikke VfL	Ikke VfL
	1 Mål	2 Instruks	3 Relasjonsbyggende	4 Korreksjon
Kjennetegn på kategori	Viser til kriterier, fagbegreper eller analysemetoder med overføringsverdi til senere skrivesituasjoner	Forklaringer på hvordan elevene skal rette opp feil språklig eller i innhold med overføringsverdi. Modellering Oppfordringer til bruk av tilbakemeldingen	Uspesifikk ros rettet mot person eller prestasjon	Retting uten forklaring, påpeking av faglige feil som er tilknyttet den spesifikke skrivesituasjonen
Eksempel	«Svært god formulering av hovedsynet.»	Her må du i stedet si at «Avsender har lånt autoritet fra»	«Bra!» «Godt poeng!»	«Gene Dalby er en mann.»
Begrunnelse	Plasseringen er begrunnet med lærers bruk av det fagspesifikke uttrykket «hovedsyn» som står sentralt i analyse av saktekster	Tilbakemeldingen peker på en uheldig formulering og modellerer en bedre formulering. Den kan brukes av eleven i en liknende situasjon fordi retorisk analyse er vanlig på eksamen	Tilbakemeldingen inneholder ikke informasjon om hva som er bra eller hvorfor poenget er godt	Kommentaren har begrenset overføringsverdi til en ny skrivesituasjon fordi det er lite sannsynlig at eleven får en tekst av forfatteren Gene Dalby også på eksamen

Kvantifisering av kvalitative data

Tilbakemeldingene ble kvalitativt kodet i tråd med operasjonaliseringen av hver kategori og resultatene ble videre kvantifisert ved opptelling av forekomst. Dette er en tilnærming som annen forskning på lærerrespons har brukt (Bergman-Claeson, 2003; Bueie, 2015b; Lee, 2008).

Kvantifiseringen er gjort med tanke på å finne tendenser og sentrale trekk som samlet sett gir en oversikt over lærerens skriftlige tilbakemeldingspraksis, og som dermed kaster lys over forskningsspørsmålet. Uten en kvantifisering er det vanskelig å få overblikk over materialet.

Samtidig gir den kvalitative kodingen innblikk i hver enkelt lærers særegne praksis. Metoden ligger tett opp til tematisk koding (Gibbs, 2008).

Analyse

En lærers tilbakemelding ble analysert i semantiske/meningsbærende enheter, heller enn som hele tekstblokker fordi lengre, flerleddete tilbakemeldinger, som vist i tabell 3, ikke vil kunne passe inn i bare en kategori. Eksempelet er sitert Frøy ved skole B:

Eksempel på operasjonalisering av kodesystemet (Tabell 3)

Lærerkommentar	Kategori
Dette er en god innledning til en retorisk analyse fordi du presenterer reklamens mål og hovedsyn.	Mål
Bra!	Relasjon
Til eksamen bør du få med tittelen i startsetningen: Reklamen «Alle er opptatt av oljeprisen, men ikke prisen for oljen, tar opp ...» slik at jeg vet hvilken Amnesty-reklame det er snakk om.	Instruks (Lærerforklaring med modellering)

Her er første setning kodet i kategorien mål, fordi den bruker fagbegrepet hovedsyn og dessuten viser at eleven skal presentere analyseobjektet for leseren i begynnelsen av teksten. Dette poenget har en potensiell overføringsverdi til senere skriving. «Bra» er kodet som en relasjonsbyggende kommentar, og utsagnet forsterker den faglige rosen. Setningen «Til eksamen ...» uttrykker eksplisitt en forventning om bruk av informasjonen i tilbakemeldingen, og følges opp av en instruks i form av både modellering, som kan brukes i en senere skrivesituasjon, og en utdypende forklaring på hvorfor den måten å formulere seg på er bedre enn elevens egen. Her peker også læreren indirekte på konseptet mottakerbevissthet, som er sentralt i skriveopplæringen (Evensen et al., 2016), og hun viser også på en indirekte måte til en avstand mellom ønsket prestasjon og faktisk prestasjon.

Tiltak for å sikre reliabilitet

En ekstern ekspertkoder har kodet 21 % av materialet. Et estimat av Cohens kappa-skåre viser et sammenfall på .77, som av Landis og Koch (1977, s. 165) betegnes som «substantial» (om verdier mellom .61 og .80) og ligger helt i øvre sjikt av denne skalaen. Fleiss (1981, s. 218) betegner verdier fra .75 som «excellent», og det er den høyeste kategorien på denne skalaen. Sett under ett indikerer skåren derfor at reliabiliteten til kodeskjemaet er akseptabel.

Studiens begrensninger

I hvilken grad tilbakemeldingspraksisen til en lærer faktisk bidrar til læring, sier skjemaet ikke noe om. Det er mulig å tenke seg at én tilbakemelding kan bidra til mer læring enn 10 som ikke treffer eleven av ulike grunner. Derfor er det ikke noen direkte forbindelse mellom antallet VFL-

kommentarer og elevens læring. Skjemaet sier i forlengelsen av dette heller ikke noe om i hvilken grad en språklig ytring som er kodet som relasjonsbyggende, faktisk er det for mottaker. Utsagn kan tolkes svært ulikt, ut fra kontekst (Rommetveit, 1972). Denne mangetydigheten er en del av norskfagets egenart (Garmannslund, Elstad & Christophersen, 2015) og kan også ses som uttrykk for den mangetydigheten som preger synet på teksttolkning i et hermeneutisk perspektiv (Gadamer, 2008).

Kategoriene er ikke gjensidig utelukkende og kan ikke være det helt ut (Bergman-Claeson, 2003). Det er likevel mulig å gi et forholdsvis presist bilde av tilbakemeldingspraksis, slik det er gjort i tidligere forskning (Bergman-Claeson, 2003; Lee, 2008; Magno & Amarles, 2011; Parr & Timperley 2010), gitt at interrater-reliabiliteten er akseptabel.

Kommentarene speiler ikke helheten i lærerens tilbakemeldingspraksis. Denne består også av muntlige tilbakemeldinger, svar på e-poster, uformelle samtaler og annet. Skriftlig tilbakemelding har likevel en sentral plass i norsklærerens virke, både fordi den kan leses og brukes i senere læringsarbeid, og fordi den er tilgjengelig for eleven i etterkant.

Tonen i tilbakemeldingen er ikke undersøkt. Ifølge Smidt (2002) er det en faktor som kan bety mye for om en elev engasjerer seg i videre skrivearbeid eller ikke.

Siden lærerne i studien frivillig har meldt seg til å delta, er seleksjonsskjevhet en potensiell validitetstrussel. En kan tenke seg at dette er lærere som for eksempel er spesielt trygge på sin egen praksis.

Resultater

Resultatene blir presentert etter skole og for hver lærer. En tolkning og sammenfatning av funnene danner utgangspunkt for diskusjonsdelen. Når det i det følgende står tilbakemelding, menes semantiske tilbakemeldingsenheter, slik det er gjort rede for over. Med per elev/ elevtekst/elevsvar menes hele elevens besvarelse på heldagsprøven.

Ved skole A ble det samlet data fra lærerne Anne, Berit og Carina. Anne skrev fyldige tilbakemeldinger til 32 elever, både underveis i teksten og med en sluttkommentar på litt under en side per besvarelse. Store deler av sluttkommentaren er felles for alle elevene, og den forklarer hva som er hovedtrekkene ved den gjeldende oppgaven og hvordan elevene har prestert i forhold til

den. Anne har totalt skrevet inn 1120 tilbakemeldinger og har i tillegg 964 fargekodemarkeringer med en variasjonsbredde fra 7 til 50. 724 av tekstkommentarene er skrevet inn i teksten og av disse er 342 kodet som instruks eller mål (heretter VfL), et gjennomsnitt på 11 per elevsvar. En typisk kommentar fra Anne var en klar oppfordring eller instruks: «Bruk etternavn eller hele navnet. Fornavn bruker vi om personer vi kjenner godt.» I sum er 570 av 1120 tilbakemeldinger utformet i tråd med VfL-anbefalingene, 51 %. Anne er den læreren som har gitt flest tilbakemeldinger, og inkluderer en fargekodene blir summen 2084, i snitt 65 per elevtekst.

Berit skrev inn 21 VfL-kommentarer til 32 elever i løpende tekst, resten er lokale, språklige korrigeringer (668). I sluttkommentarene (119) er 77 i tråd med anbefalingene. I snitt har elevene fått 2,9 VfL-tilbakemeldinger og 26,5 språklige korrigeringer, med en variasjonsbredde fra 2 til 63. 12 % av de totale tilbakemeldingene kan knyttes til VfL. En typisk kommentar fra Berit har et generelt preg, i en sluttkommentar skriver hun for eksempel: «(...) skriv mer i siste avsnitt og vis at du kan reflektere mer.» Det er ikke relasjonsbyggende kommentarer, men 43 tilfeller av plusstegn eller spørsmålstegn i tekstene. I sum er Berit den læreren som har gitt færrest skriftlige tilbakemeldinger, hennes praksis i denne vurderingssituasjonen består i hovedsak av språklig korrigerings av elevenes tekster.

Carina ga tilbakemelding til 24 elever i en vg2-klasse. Elevene hadde skrive dag på skolen (heldagsprøve) basert på et eksamensliknende oppgavesett, men fikk mulighet til å levere en ferdig versjon av teksten noen dager senere etter å ha fått tilbakemelding fra læreren. Muligheten ble med ett unntak neglisjert av elevene. Carina har gitt 236 kommentarer i løpende tekst, av disse kan 133 knyttes til VfL. Av 115 sluttkommentarer er 96 av samme type. Hun har 39 relasjonsbyggende kommentarer, og det er 673 fargekodemarkeringer av språkfeil, i snitt 26 per svar, med en variasjonsbredde fra 4 til 60. En tilbakemelding fra Carina kan se ut som dette: «Langteksten din er litt kort (bør være på to sider), men har en grei struktur. Her kunne du også skilt ut innledningen, der du presenterte tekstene, forfatterne og si noe om hvordan du skal løse oppgaven. Da blir teksten enda ryddigere for leseren.» Tilbakemeldingene inneholder altså ofte forklaringer eller instruksjoner. Av Carinas tilbakemeldinger er det en klar overvekt som kan knyttes til VfL, 229 av totalt 351, det vil si 65 %. Hun har også skrevet langt færre tilbakemeldinger enn kollegaene ved samme skole og er den på skolen som i størst grad har lagt opp til at elevene skal jobbe videre med teksten.

På skole B deltok lærerne Dina, Eva og Frøy. Alle lærerne brukte skjema med kjennetegn på måloppnåelse da de ga tilbakemelding. Dina (30 elever) skrev 603 tilbakemeldinger i løpende tekst,

og av disse kan 198 knyttes til VfL. Sluttkommentarene er nummerert og knyttet til skjemaet med kjennetegn på måloppnåelse. Av totalt 248 sluttkommentarer er 159 kodet som VfL. Hun har 136 relasjonsbyggende kommentarer i teksten, flest av alle. Dina kombinerer ofte ros med korrigerende. Til en elev skriver hun først: «Fra landsmålet kom på 1850-tallet; det er ikke *mange* hundre år» (innholdskorrigerende) og rett under: «God innledning!» I sum kan 42 % av tilbakemeldingene knyttes til VfL.

Eva skrev 721 tilbakemeldingsenheter til 25 elever, av disse er 577 i løpende tekst og 114 som sluttkommentar. I løpende tekst kan 318 knyttes til VfL, 59 %, og for sluttkommentarene er forholdet 106 av 144, 73 %. Det er altså en klar overvekt av VfL-tilbakemeldinger. Et særtrekk ved hennes praksis er at hun stiller mange spørsmål til elevene, og de kan blant annet fungere som instruksjoner og korrigerende. Til en elev skriver hun: «I stedet for å si “budskapet”, bør du her skrive tolkningen din av budskapet. Hva er ordspillet? Forklar.» Hun har få språklige korrigerende, i snitt 5,3 per elevsvar.

Frøy ga 28 elever 541 tilbakemeldinger i løpende tekst og av disse kan 357, 65 %, knyttes til VfL, og i sluttkommentarene, som bare er gitt til noen elever, er 61 av totalt 92 kommentarer av denne typen (66 %). Det er 87 relasjonsbyggende kommentarer. I sum har Frøy mange VfL-kommentarer (418) og relativt få korrigerende (103). Hun gir mye relasjonell støtte og modellerer og utdyper ofte poengene sine: Til en elev skriver hun: «God innledning fordi du presenterer teksten og deretter gjør rede for formålet med kampanjen.» Frøy har nesten ikke språklige korrigerende i tekstene, bare cirka en per elev.

Oversikt over funn

Funnene er her presentert i en enkel oversikt i tabell 4, se appendiks A og B for detaljert oversikt over hvordan tilbakemeldingene fordeler seg på kategoriene.

Sammenfatning av funn (Tabell 4)

Skole A	Lærer	Totalt	VfL	%	Spredning	Ikke VfL	%	Spredning
	Anne	1120	570	51 %	0–13	550	49 %	0–21
	Berit	847	98	12 %	0–4	749	88 %	0–63
	Carina	351	229	65 %	0–8	122	35 %	0–8
Sum				36 %			64 %	
Skole B	Dina	851	357	42 %	0–8	494	58 %	0–19
	Eva	721	424	59 %	1–14	297	41 %	0–18
	Frøy	633	418	66 %	0–12	215	34 %	0–7
Sum				54 %			47 %	

Alle elever har fått tilbakemeldinger, i snitt er det 26 per elevsvar. Spredningstallene viser samtidig at det særlig innen språklige korrigeringer er store variasjoner mellom lærerne, fra 0 til 3 per elevsvar (Frøy ved skole B) og fra 0 til 63 hos Berit ved skole A. De øvrige kategoriene kjennetegnes av at lærerne har gitt et sted mellom 0 og 10 tilbakemeldinger per elevsvar.

Oppsummering av funn

Svaret på forskningsspørsmålet – i hvilken grad lærerne utformer skriftlige tilbakemeldinger på skriftlige prøver i norskfaget som er i tråd med premisene for formative tilbakemeldinger slik de er uttrykt i internasjonal forskning, er i første rekke at lærerne gir elevene mange formative tilbakemeldinger, de fleste rundt halvparten eller mer, én lærer er unntaket (12 %). Oversikten viser også at lærerne vanligvis gir elevene mange tilbakemeldinger på en skriftlig tekst, både i form av skriftlige kommentarer og språklige korrigeringer i teksten og til slutt. Her er unntaket Frøy, som nesten ikke har språklige korrigeringer. Samtlige lærere har gitt flest tilbakemeldinger i løpende tekst, i tråd med anbefalingene til Parr og Timperley (2010).

Selv om dette er en vurderingssituasjon der det å sette karakter står sentralt, er det likevel mange tilfeller av tilbakemeldinger som kan knyttes til VfL, og nesten ingen eksplisitt karakterbegrunnende kommentarer. Denne hensikten kan derimot summen av korrigeringer til dels fylle.

Et sentralt funn er også at det er en tydelig variasjonsbredde i hvordan lærerne gir tilbakemeldinger.

Diskusjon

Et sentralt formål med studien var å undersøke i hvilken grad gir lærerne skriftlige tilbakemeldinger som retter seg mot målet med opplæringen og i hvilken grad de inneholder konkrete råd og instruksjoner som kan hjelpe elevene til å komme nærmere målet. Helhetsbildet av lærernes tilbakemeldingspraksis er at alle elever fikk skriftlige tilbakemeldinger på heldagsprøven, og at de fleste lærerne ga forholdsvis mange VFL-kommentarer. Det er likevel grunn til å nyansere bildet.

Heldagsprøven er en skrivesituasjon som likner eksamen på mange måter (Bueie, 2015b), og som aktiviserer ferdigheter og kunnskaper hos elevene som speiler deres helhetlige kompetanse på slutten av opplæringen. Dysthe (2011) argumenterer for at tilbakemeldinger i skriveopplæring kan rettes mot dybdeforhold i teksten og på den måten ha overføringsverdi til neste skrivesituasjon selv om den aktuelle teksten ikke skal revideres. Det er forventet på eksamen at elevene kan skrive en godt strukturert tekst med et faglig relevant innhold, at de behersker analyseteknikker som er spesifikke for norskfaget (f.eks. diktanalyse og retorisk analyse), at de bruker fagbegreper og at de bruker tekstene som ligger til grunn for oppgavesettet på en relevant måte. Dette er forhold som *ikke* endrer seg fra oppgave til oppgave, men er noe lærere og elever jobber med kontinuerlig gjennom året. Det betyr at det også er en klar forbindelse mellom heldagsprøven og eksamen (Bueie, 2015b) og at tilbakemeldinger som retter seg mot de nevnte forholdene, vil kunne være til nytte i elevenes læringsprosess nettopp fordi de skal aktiviseres igjen i en nært forestående skrivesituasjon som er grunnleggende lik heldagsprøven.

En faktor som også tilsier at lærerne bør vektlegge læring i denne skrivesituasjonen, er at norskeksamen er den eneste nesten alle elever på studiespesialiserende studieretning må ha, og at den er en viktig målestokk for både skolens og den enkelte lærerens grad av suksess med opplæringen. At lærerne derfor har en sterk interesse i å gjøre det de kan, også gjennom tilbakemeldingene sine, for at elevene skal lykkes på eksamen, vil i hvert fall være rasjonell atferd. Elevene er i tillegg nær slutten av skoleåret og har mulighet til å bruke informasjon fra læreren til å bedre skrivekompetansen sin, og kanskje også standpunkt karakteren. Undersøkelsen viser at de fleste lærerne utformer mange av tilbakemeldingene på en slik måte at de kan bidra til læring, og forskning viser klart at utformingen av tilbakemeldingen har betydning for dens potensial til å bidra til læring. Tilbakemeldinger som peker mot sentrale trekk ved målet for oppgaven og gir elevene konkrete råd om hvordan de skal komme nærmere målet (Parr & Timperley, 2010), har et større potensial for å bidra til læring og utvikling enn tilbakemeldinger som er vage (Sommers, 1982;

Kronholm-Cederberg, 2009), for kritiske (Askew & Lodge, 2000) eller retter seg mot selvet heller enn mot oppgaven (Hattie & Timperley, 2007).

Et viktig spørsmål er videre om en kan snakke om vurdering for læring når det er uvisst om elevene skal jobbe videre med heldagsprøveteksten eller de ikke har gjort det. Hvis en legger definisjoner av vurdering for læring til grunn som baserer seg på en klasseromspraksis (Black & Wiliam, 2009) eller som legger avgjørende vekt på bruk av læringsresultater (Ramaprasad, 1983), så kan en ikke det. Da er det mest nærliggende å si at heldagsprøven i norsk, gitt i april, er vurdering av læring. Dette er også understreket av blant annet Bueie (2015b), som finner at lærere kanskje gir elever tilbakemeldinger som er utformet med tanke på videre læring, men at denne muligheten langt på vei kan gå tapt når en ikke gir elevene muligheter til å benytte seg av informasjonen fra tilbakemeldingen til å jobbe videre med teksten (jf. også Fjørtoft, 2013). Tidligere forskning har vist at kommentarer ofte har kommet etter at elever er ferdige med en oppgave (Fjørtoft, 2013; Hattie & Timperley, 2007; Hillocks, 1987). Da er det vanskelig å bruke tilbakemeldingsinformasjonen formativt, og læringspotensial kan gå tapt. Det kan også løftes fram som et problem at tilbakemeldingen gis med karakter, noe blant annet Leahy et al. (2005) legger vekt på. Samtidig må det altså foreligge en tilbakemelding til elevene som de kan forstå og potensielt sett bruke.

Et sentralt aspekt ved tidspunktet for å undersøke om lærerne gir VfL-tilbakemeldinger (på heldagsprøven), er at en må kunne forvente at det å begrunne karakteren står sterkt. Det er imidlertid bare en av lærerne (Dina) som skriver eksplisitt karakterbegrunnende kommentarer. Når en ser lærernes tilbakemeldinger under ett, viser altså undersøkelsen at mange av dem formulerer seg på måter som kan bidra til læring, og at hensikten, i tråd med Newton (2007), heller synes å være læring enn karaktersetning, ettersom karakterbegrunnelser ser ut til å komme fram på en indirekte måte.

Funnene viser samtidig at Carina faktisk legger opp til videre arbeid med teksten, og at Frøy holder tilbake karakter, noe som peker mot en VfL-orientert praksis også i denne situasjonen. For Carinas del var det slik at elevene hadde anledning til å jobbe videre med heldagsprøven på bakgrunn av tilbakemeldinger, men at de ikke benyttet muligheten. Legger en definisjonen for VfL fra Parr og Timperley (2010) til grunn, kan en her si at det er elevene som ikke følger opp sin del av arbeidet. Hva som er årsakene til denne manglende innsatsen sier imidlertid ikke datamaterialet noe om.

Av VfL-kategoriene er det flest instruksjer og færre målkommentarer (se appendiks A og B). En forklaring kan ligge i målets kompleksitet, der hva som er en god tekst er svært mangesidig. Sadler

(1989) viser til rundt 50 mulige kriterier for hva som kan være en god tekst. Lærerne peker på faktorer som mottakerbevissthet, de bruker analysebegreper med overføringsverdi til neste skrivesituasjon og kan peke på strukturelle trekk som avsnittsinndeling. En del tilbakemeldinger tar også opp i seg uttrykk fra vurderingsskjemaet med kjennetegn på måloppnåelse som følger med skriveoppgaven, særlig gjelder dette ved skole B. At målkategorien er underrepresentert, kan også henge sammen med at elevene gjøres kjent med målet for opplæringen gjennom undervisningen, noe lærerne er pålagt etter opplæringsloven (Udir, 2010), slik at det i mindre grad uttrykkes skriftlig av lærer når vurderingssituasjonen er et tilbakelagt stadium. Noe som taler imot denne antakelsen er at flere av lærerne skriver utdypende forklaringer, modellerer språklige formuleringer og oppfordrer til bruk av tilbakemeldingsinformasjonen. Det framgår slik sett at de på mange måter forutsetter at elevene leser tilbakemeldinger, legger vekt på innholdet og i neste skrivesituasjon vil nyttiggjøre seg informasjonen. At elevene også kan tenkes å dele denne oppfatningen, framheves blant annet av Bueie (2015a), som har intervjuet elever om hvordan de oppfatter og nyttiggjør seg lærers skriftlige tilbakemeldinger på en tekst de er ferdige med. Hvis en ser summen av VfL-kommentarer som uttrykk for en intensjonalitet, framgår det at lærerne langt på vei skriver med elevenes læring som mål.

Et annet sentralt trekk er at lærerne korrigerer elevenes tekster. Det er omstridt om korrigerende av språkfeil bidrar til læring (Ferris, 2004; Truscott, 1996), men korrigeringen kan tenkes å bidra til å begrunne karakter og dermed ha summative funksjoner. Det er likevel verdt å merke seg at en lærer (Frøy) som i høy grad gir VfL-tilbakemeldinger, i liten grad korrigerer elevenes tekster språklig. Den læreren som har gitt flest tilbakemeldinger (Anne ved skole B), har også flest språklige korrigeringer, og det kan tenkes at elevene vil finne det krevende å forstå hva de skal fokusere på i videre læringsarbeid, særlig siden mange av dem presterer på middels nivå eller lavere. En slik praksis vil dessuten kreve et omfattende etterarbeid av elevene. Derfor kan en si at denne praksisen, trass i at det er mange VfL-kommentarer, er vanskelig for elevene å omsette i læring.

Hvis en ser de seks lærerne under ett, kan det altså se ut som om satsingen på å utvikle en vurderingspraksis i skolen manifesterer seg i at norsklærerne utformer mange tilbakemeldinger på en slik måte at de legger til rette for læring, også i en situasjon der det å begrunne karakter må stå sentralt, ut fra et rasjonelt perspektiv.

Konklusjoner

Rundt halvparten av tilbakemeldingene som er gitt, er utformet i tråd med anbefalingene, om en ser de seks lærerne under ett. Dette viser seg klart på skole B, men lærerne på begge skoler har mange tilbakemeldinger som kan knyttes til VfL. Eksplisitt karakterbegrunnende kommentarer er det svært få av. Slik sett kan det hevdes at lærerne, selv i denne skrivesituasjonen som er viktig for standpunkt karakteren, gir tilbakemeldinger utformet med tanke på elevenes læring. Dette tegner et bilde av lærere som har egne elevers læring i sentrum. At VfL-kommentarer i mange tilfeller kombineres med ros, forsterker bildet.

Implikasjoner for videre forskning

Det trengs mer forskning for å få kunnskap om hvor representative disse lærerne er for en større populasjon. Det vil også være av interesse å undersøke i hvilken grad lærerne gir elevene muligheter til å bruke tilbakemeldinger. Fjørtoft et al. rapporterte om dette i 2013 og fant at mulighetene var få, men siden da har satsingen på VfL fortsatt. Derfor er det viktig å få mer kunnskap om hvordan lærere jobber med vurdering og vurdering for læring i dag.

Implikasjoner for praksis

Når en legger til grunn at mange av lærerne i studien skriver en forholdsvis høy andel VfL-tilbakemeldinger, så kan det se ut som om en sentral utfordring er å fremme måter å bruke tilbakemeldingene på i den videre læringsprosessen. Det kan også se ut som elevene i større grad bør kunne involveres i prosessen og kanskje at de pålegges å bruke lærers tilbakemeldinger, særlig når en ser på Carinas tilfelle, der kun en elev, i en hel klasse, har benyttet denne muligheten. Å gjøre en slik praksis til et krav kan være en løsning, gitt at den ikke øker lærernes arbeidsbyrde, ettersom tids- og arbeidspress allerede er et betydelig problem i skolen (Skaalvik & Skaalvik, 2013).

Det er diskrepans mellom lærerne i studien og minst en av lærerne gir få VfL-kommentarer. Dette impliserer at det kan være mye å vinne på å gi lærere opplæring i hvilke former for tilbakemeldinger som har størst sjanse for å bidra til framgang for elevene i skriveopplæringen.

Referanser

- Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops – linking feedback and learning. I Askew, S. (red.) *Feedback for learning*, 1-17. London. Routledge.
- Bergman-Claeson, G. (2003). *Tre lärare-tre världar: lärarkommentarer till elevtexter i tre gymnasieklasser*. Uppsala universitet.
- Berge, K. L., Evensen, L. S., & Thygesen, R. (2016). The Wheel of Writing: a model of the writing domain for the teaching and assessing of writing as a key competency. *The Curriculum Journal*, 27.2 1-18.
- Berge, K. L., & Skar, G. (2015). Ble elevene bedre skrivere? Trondheim. Høgskolen i Sør-Trøndelag.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education*, 5(1), 7-74.
- Bueie, A.A. (2015)a. «Hvis lærer sier at vi skal rette, så gjør jeg det» - en studie av elevers forståelse og bruk av lærerkommentarer på heldagsprøver i norsk. *Acta Didactica Norge*, 9(1), Art-4.
- Bueie, A.A. (2015)b. Summativ vurdering i formativ drakt – elevperspektiv på tilbakemelding fra heldagsprøver i norsk. *Acta Didactica Norge*, 9(1), Art-4.
- Bueie, A. A. (2016). Nyttige og mindre nyttige lærerkommentarer–slik elevene ser det. *Nordic Journal of Literacy Research*, 2(1).1–28.
- Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways), i Aronson J. (ed.): *Improving academic achievement: Impact of psychological factors on education*, 37-69. San Diego, Academic Press.
- Dysthe, O. (2011). 'What is the Purpose of Feedback when Revision is not Expected?' A Case Study of Feedback Quality and Study Design in a First Year Master's Programme. *Journal of Academic Writing*, 1(1), 135-142.
- Evensen, L. S., Berge, K. L., Thygesen, R., Matre, S., & Solheim, R. (2016). Standards as a tool for teaching and assessing cross-curricular writing. *The Curriculum Journal*, 27.2 1-17.
- Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...?). *Journal of second language writing*, 13(1), 49-62.
- Fjørtoft, H. (2013). Vurderingspraksiser i norskfaget. I Sandvik, L.V., og Buland, T. (Red.), *Vurdering i skolen. Operasjonaliseringer og praksiser* (s. 99-118). Trondheim: NTNU
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), 105-112.
- Gadamer, H. G. (2008). *Philosophical hermeneutics*. University of California Press.
- Garmannslund, P. E., Elstad, E. & Christophersen, K. (2015). Norsk lærers innflytelse på egne elevers motivasjon, selvdisiplin, innsats og norskprestasjon: Empiriske sammenhenger hos førsteårselever i videregående skole. *FOU i praksis* 9(1), 9–32.
- Gibbs, G. R. (2008). *Analysing qualitative data*. Los Angeles. Sage.
- Hattie, John, and Helen Timperley (2007). The power of feedback. *Review of educational research* 77.1, 81–112.
- Hayward, L., Priestley, M., & Young, M. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30(3), 397–415.

- Hertzberg, F. & Roe, A. (2016). Writing in the content areas: a Norwegian case study. *Reading and writing*, 29(9), 555–576.
- Hillocks Jr, G. (1987). Synthesis of Research on Teaching Writing. *Educational Leadership*, 44(8), 71–76, 78, 80–82.
- Hopfenbeck, T., Tolo, A., Florez, T., & El Masri, Y. (2013). *Balancing Trust and Accountability? The Assessment for Learning Programme in Norway*. Oslo. OECD.
- Iglund, M. A (2008). *Mens teksten blir til: Ein kasusstudie av lærarkommentarar til utkast*. Oslo, Unipub.
- Jølle, L. (2014). Pair assessment of pupil writing: A dialogic approach for studying the development of rater competence. *Assessing Writing*, 20, 37-52.
- Kronholm-Cederberg, A. (2009). *Skolans responskultur som skriftpraktik: gymnasiseters berättelser om lärarens skriftliga respons på uppsatsen*. Åbo akademis forlag.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159-174.
- Leahy S., Lyon, C., Thompson, M. & Wiliam, D. (2005). Classroom assessment: Minute by minute and day-by-day. *Educational leadership*, 63 (3), 18–24.
- Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(2), 69-85.
- Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, 12(3), 180-198.
- Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, 13(4), 285-312.
- Magno, C., & Amarles, A. (2011). Teachers' feedback practices in second language academic writing classrooms. *The International Journal of Educational and Psychological Assessment*, 6(2), 21-30
- Matre, S., & Solheim, R. (2015). Writing education and assessment in Norway: Towards shared understanding, shared language and shared responsibility. *L1-Educational Studies in Language and Literature*, 15, 1-33.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), 149-170.
- Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing writing*, 15(2), 68-85.
- Patton, M., Q. (2014). *Qualitative Research & Evaluation Methods: Integrative Theory and Practice*. Sage Publications.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13.
- Rommetveit, R. (1972). *Språk, tanke og kommunikasjon*. Oslo. Universitetsforlaget.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), 119-144.
- Sandvik, L. V. & Buland, T. (2014). *Vurdering i skolen. Utvikling av kompetanse og fellesskap. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)*. Trondheim. NTNU.
- Skaalvik E.M., & Skaalvik, S. (2013). *Lærerrollen sett fra lærernes ståsted*. Trondheim. NTNU. Samfunnsforskning AS.
- Skaalvik, E. M., & Skaalvik, S. (2011). *Motivasjon for skolearbeid*. Trondheim. Tapir akademisk.
- Smidt, J. (2002). Double Histories in Multivocal Classrooms Notes Toward an Ecological Account of Writing. *Written Communication*, 19(3), 414-443.

- Sommers, N. (1982). Responding to student writing. *College composition and communication*, 33(2), 148-156.
- Stortingsmelding nr. 28 (2015–2016). Fag – fordypning – forståelse – en fornyelse av kunnskapsløftet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language learning*, 46(2), 327-369.
- Utdanningsdirektoratet (2014). Nasjonal satsing på vurdering for læring (2010–2014). <http://www.udir.no/Vurdering-for-laring/Nasjonal-satsing1/Nasjonal-satsing-pa-Vurdering-for-laring>
- Utdanningsdirektoratet (2010). Rundskriv Udir-1-2010. Individuell vurdering. <http://www.udir.no/Regelverk/Finn-regelverk-for-opplaring/Finn-regelverk-etter-tema/Vurdering/Udir-1-2010-Individuell-vurdering>
- Utdanningsdirektoratet (2009). Bedre vurderingspraksis, sluttrapport fra Utdanningsdirektoratet 2009. <http://www.udir.no/Tilstand/Forskning/Rapporter/Utdanningsdirektoratet/Bedre-vurderingspraksis-sluttrapport-fra-Utdanningsdirektoratet-2009/>
- Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, 31(3), 379-394.

Artikkel 2

Elevers oppfatning av lærers tilbakemeldingspraksis: Om sammenheng mellom graden av VfL-praksis og elevenes opplevde nytte av lærers skriftlige tilbakemeldinger i skriftlig norsk

Sammendrag

Formålet med denne studien var å undersøke elevers oppfatninger av norsklærerens tilbakemeldingspraksis i skriftlig norsk. På bakgrunn av omfattende internasjonal forskning innen feltet 'Vurdering for læring' (VfL) er det i denne studien utviklet teoridrevne hypoteser som testes ved å analysere et datasett fra to videregående skoler. Hypotesene er at elevene oppfatter lærers tilbakemelding som mer nyttig hvis den 1) retter seg mot målet med opplæringen, 2) gir elevene informasjon via en forklaring eller instruks om hvordan de skal komme nærmere målet og 3) at lærer legger til rette for bruk av tilbakemeldingen.

Strukturell likningsmodellering (SEM) av resultater fra en spørreundersøkelse blant elever (N = 213) er brukt som metode til å estimere forholdet mellom de fire latente variablene – kriterier, forklaring, bruk og nytte (avhengig variabel). SEM er brukt fordi metoden egner seg til å undersøke sammenhenger i komplekse datasett. Funn er at elevene oppfatter at hvis lærer forklarer hvordan de kan forbedre seg, er dette positivt statistisk assosiert med opplevd nytte. Et mer overraskende funn er at lærerens tilrettelegging for bruk av tilbakemeldinger ikke er signifikant assosiert med opplevd nytte. Resultatene blir diskutert i artikkelen.

Nøkkelord: Vurdering for læring, lærers tilbakemelding, skriveopplæring i norskfaget, elevoppfatninger av tilbakemelding, bruk av tilbakemelding

Introduksjon

Siden midten av 2000-årene har det vært en økende grad av satsing på en vurderingspraksis i norsk skole som har til formål å fremme læring (Utdanningsdirektoratet, 2010; Hopfenbeck et al., 2013). Satsingen bygger blant annet på internasjonal forskning innen feltet 'Vurdering for læring' (heretter forkortet VfL), særlig står studier av Black og Wiliam (1998; 2009) sentralt. Blant de mest sentrale funnene er noen grunnleggende prinsipper for hva som skal til for at vurdering skal

ha større sjanse for å bidra til læring. Prinsippene er 1) at elevene kjenner målet med opplæringen, 2) at de får informasjon om hvordan de har prestert i forhold til målet, 3) at de får tilbakemelding om hva de skal gjøre for å komme nærmere målet og 4) at denne informasjonen brukes i det videre læringsarbeidet (Hopfenbeck et al., 2013). Disse fire anbefalingene danner grunnlaget for hvordan VfL er forstått i denne studien.

Videre framhever Hattie og Timperley (2007) at hvordan læreren utformer tilbakemeldinger, har betydning for om elever eller studenter kan bruke dem i videre læringsarbeid, et poeng som også framheves av Parr og Timperley (2010). I tillegg er bruken av tilbakemeldinger et av kjernepunktene i VfL som teorifelt (Black & Wiliam, 2009). Samlet sett begrunner disse studiene hvordan tilbakemeldinger er en svært viktig del av en læringsfremmende vurderingspraksis, og det er også en del av grunnlaget for valget av studieobjekt her.

Siden de sentrale prinsippene er generelt formulert, må de forankres i konkrete fag- og skolekontekster (Bennett, 2011). I denne studien forankres VfL i norskfaget i videregående skole, nærmere bestemt i norskfagets skriveopplæring. Denne vinklingen er blant annet valgt fordi faget er sentralt i skolen, og fordi skriving er en av de grunnleggende ferdighetene som det er satset spesielt på gjennom innføringen av skolereformen Kunnskapsløftet (Hertzberg & Roe, 2016).

Hvordan en formativ tilbakemeldingspraksis er operasjonalisert i norske klasserom, er blant annet vurdert i FIVIS-rapportene (Sandvik & Buland, 2014). De framhever at skolen har langt igjen før en VfL-praksis er etablert. Havnes et al. (2012) fant i en toårig intervensjonsstudie i videregående skole blant annet at elever og lærere hadde ulike oppfatninger av tilbakemeldingspraksis, men ingen av de nevnte studiene vektlegger spesielt om og i hvilken grad VfL operasjonaliseres i en skriveopplæringskontekst.

Formålet med denne studien er å undersøke elevens oppfatning av lærerens tilbakemeldingspraksis i skriftlig norsk. I det følgende vil jeg redegjøre for det teoretiske og forskningsmessige grunnlaget for denne studiens fire latente variabler, som operasjonaliserer viktige sider ved VfL. Denne operasjonaliseringen er foretatt med spesielt fokus på en formativ tilbakemeldingspraksis.

Tilbakemeldingens nytte

Den avhengige variabelen i denne studien er elevens oppfatning av nytte, og grunnlaget er følgende: I en skolesammenheng er en stor del av hensikten med å gi tilbakemelding gjennom

skoleåret å bidra til læring. Black og Wiliam (1998) framhever i en mye sitert metastudie at en vurderingspraksis som er i tråd med VfL, gir bedre læringsresultater enn en mer summativt orientert praksis som har til hensikt å måle prestasjon eller nivå på et gitt tidspunkt (Taras, 2005). Nytte er dessuten er en sentral faktor i en reviewartikkel av Jönsson (2013) om studenters oppfatning av tilbakemeldinger, der det legges til grunn at studenter har større sjanse for å jobbe videre på bakgrunn av tilbakemeldinger hvis de oppfatter dem som nyttige.

At en tilbakemelding kan fylle flere funksjoner, er samtidig et poeng Newton (2007) peker på, og som understreker at tilbakemeldingspraksis er et komplekst fenomen med flere hensikter. Det kan for eksempel ikke utelukkes at elever kan oppleve en tilbakemelding som nyttig hvis den begrunner karakteren på en måte som elevene finner tydelig og rettferdig – et poeng framhevet av Straub (1997) i en studie som undersøkte studenters preferanser når det gjelder tilbakemeldinger. Det blir derfor viktig hvordan variabelen nytte i dette tilfellet er operasjonalisert. Et eksempel er: «Når jeg får tilbakemelding fra norsklæreren på et skriftlig arbeid, får jeg et klarere bilde av hva som forventes av meg.» Operasjonaliseringen framgår ellers av tabell 2.

Kriterier

Et sentralt punkt sier at elevene må kjenne målet med opplæringen (Hopfenbeck et al., 2013, Udir, 2010). I skriveopplæring er målet svært komplekst; Sadler (1989) viser hvordan det kan ligge opptil 50 latente kriterier til grunn for en god tekst – det vil si forhold ved teksten som leseren blir oppmerksom på i det skriveren ikke mestrer dem, men som en ellers ikke ville festet seg ved. Når en skal kommunisere hva som er målet med opplæringen i en skrivepedagogisk kontekst, bør derfor læreren velge ut noen kriterier eller forhold som det fokuseres særlig på (Kvithyld & Aasen, 2011). På grunnlag av denne antakelsen er derfor det komplekse forholdet «målet med opplæringen» søkt operasjonalisert gjennom spørsmål som retter seg mot bruk av kriterier, ettersom kriteriene kan ses som et aspekt ved lærerens operasjonalisering av målene med skriveopplæringen. Sadler (2010) viser samtidig i en studie av studenters oppfatninger av kriterier at kriterier i seg selv ikke bidrar til klarhet, blant annet fordi lærere og studenter kan ha ulike oppfatninger av innholdet eller at kriteriene ikke fanger opp forhold som er viktige for studentens oppgave. Dette viser at spørsmålet om kriterier er komplekst, og hvorvidt det er sammenheng mellom opplevd nytte og bruk av kriterier som viser til målet med opplæringen, er et empirisk spørsmål. Bueie (2015) finner dessuten i en studie av ungdomsskoleelevers syn på lærers tilbakemeldinger i norskfaget at elever i

liten grad forholder seg til vurderingskriterier. Det legges likevel til grunn at det er bedre at læreren har kriterier enn at slike mangler, fordi eleven dermed har tilgang til en skriftliggjøring av gjeldende normer (Sadler, 2010, s 546). Dette gir meg bakgrunn for å formulere følgende hypotese: Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte (hypotese 1, H1). Et eksempel på en indikator er: «Tilbakemelding fra norsklæreren er knyttet til kriterier som er spesielle for hver enkelt oppgave.»

Forklaring

Elevene trenger informasjon om hvordan de skal komme seg nærmere målet med opplæringen (Sadler, 1989). Dette forholdet er søkt operasjonalisert gjennom indikatorer som undersøker om og i hvilken grad elevene oppfatter at læreren forklarer, modellerer og er konkret i sin tilbakemelding på skriftlige arbeider, slik at elevene kan forbedre seg, det vil si komme nærmere målet. Det kan de gjøre ved å jobbe videre med den aktuelle teksten, et poeng som står sentralt i en mye sitert metastudie av Hattie og Timperley (2007). Dysthe (2011) framhever videre at tilbakemeldinger kan være effektive hvis de retter seg mot et overordnet nivå i en tekst, det vil si forhold som går igjen fra tekst til tekst, selv om studenten ikke skal bearbeide den aktuelle teksten. Et eksempel fra norskfaget på overordnede forhold kan være tolkning av skjønnlitterære tekster, som innebærer at elevene benytter analyseteknikker og begrepsapparat som likner fra gang til gang, og som elevene øver på gjennom skolegangen.

I en studie av Parr og Timperley (2010) har forskerne sammenholdt 52 morsmålslæreres evne til å gi formative tilbakemeldinger på et skriftlig arbeid, med elevenes resultater fra en nasjonal skriveprøve i førstespråket. De lærerne som i høy grad gir formative tilbakemeldinger, har elevene med størst læringsframgang. Formative tilbakemeldinger kjennetegnes i studien ved at de retter seg mot målet med opplæringen, at de inneholder en konkret forklaring og er forankret i elevens tekst.

Straub (1997) framhever at studenter i skriveopplæringen foretrekker tilbakemeldinger som framstår som klare. Klare tilbakemeldinger kan bidra til å utvikle teksten særlig på lokalt nivå gjennom å være spesifikke og å ha informasjon om hvordan teksten kan forbedres (Sadler, 2010, s 538). Disse aspektene har med lærerens ferdighet til å forklare fagstoffet å gjøre, og derfor deduserer jeg følgende hypotese: Elevers oppfatning av lærers ferdighet til å forklare fagstoffet er positivt statistisk assosiert med opplevd nytte (H2). Et eksempel på en indikator er: «Norsklæreren forklarer i tilbakemeldingene hva jeg skal gjøre annerledes neste gang.»

Bruk

Black og Wiliam (2009) framhever i en studie som oppsummerer det teoretiske grunnlaget for VfL at informasjon som framkommer gjennom vurdering, bør brukes av lærer og elever i det videre arbeidet. Om – og på hvilken måte – elevene oppfatter at læreren legger til rette for bruk av tilbakemelding, er undersøkt gjennom indikatorer som søker å måle om det settes av tid i timene til å jobbe med tilbakemeldinger, om det er mulig for elevene å levere tekster på nytt og om læreren krever denne typen aktivitet av elevene. Premisset som ligger til grunn for variabelen, er at lærerens tilrettelegging for bruk påvirker hvorvidt elevene nyttiggjør seg informasjonen i tilbakemeldingen. Dette framhever Bueie (2015) som viktig for elevens videre arbeid. At lærer legger til rette, antas derfor å være viktig, men er ikke det eneste som kan bidra til å forklare om elever vil nyttiggjøre seg tilbakemeldinger.

Spørsmålet om hva som kan bidra til å forklare studenters bruk eller manglende bruk av tilbakemeldinger, er kartlagt i en rekke studier. Jönsson (2013) framhever ulike forhold som kan bidra til å forklare hvorfor bruk av tilbakemeldinger i skriveopplæringen synes å representere en utfordring. Han finner at studenter må oppfatte tilbakemeldingen som nyttig, og at «timing» er av betydning her. Dette funnet samsvarer med Lea og Street (1998). De finner i en studie som undersøker studenters og læreres oppfatninger av tilbakemeldingspraksis at det er et problem med tidspunktet for når tilbakemeldingen gis. En grunn til dette er at studentene ofte jobber med annet fagstoff når de får tilbakemelding. Jönsson (2013) finner dessuten at det kan være et misforhold mellom hva studenter ønsker seg og hva de får.

Annen forskning på feltet finner blant annet at en autoritativ tone minsker muligheten for at studenter vil ta til seg tilbakemeldingen (Cowie, 2005), at de mangler strategier for bruk og at en del ikke forstår hva læreren skriver til dem (Kronholm-Cederberg, 2009). Straub (1997) framhever i forlengelsen av dette at en rekke forhold knyttet til kontekst – tonen i tilbakemeldinger, tidspunkt og i hvilken grad læreren er spesifikk og konkret – har betydning for hvordan tilbakemeldingen oppfattes. Det er likevel fellesnevner for hva studenter og elever foretrekker av tilbakemeldinger. Ferris (2003) finner at studenter liker konkrete forbedringsforslag, Beach (1989) at de liker forklaringer på hvorfor noe er bra eller ikke og Straub (1997) framhever at tilbakemeldinger kan være kritiske, men at tilbakemeldinger som utfordrer tekstens overordnede ideer i mindre grad settes pris på. I forlengelsen finner han at studenter misliker at tilbakemeldingen oppfattes som negativ, ironisk eller som et direktiv som ikke tar hensyn til studentens hensikt eller idé med en tekst.

I tillegg er det i noen grad personavhengig hvordan den samme tilbakemeldingen oppfattes (Hyland, 1998). Gibbs og Simpson (2004) finner i en studie som undersøker hvordan vurdering kan bidra til økte studentprestasjoner at flere forhold influerer studenters bruk av tilbakemeldinger i videre arbeid. De peker blant annet på at tilbakemeldingen bør være i samsvar med studentens opplevelse av oppgavens innhold, kravene i faget og i oppgaven, og at de har lært teknikker som gjør at de er oppmerksomme på og kan ta imot innholdet i tilbakemeldingen. At bruk av tilbakemeldinger er et komplekst felt i skriveopplæring, illustreres også av Calhoon-Dillahunt og Forrest (2013), som finner at studenter i en skriveprosess i liten grad endrer tekstene sine i tråd med lærernes anbefalinger, selv om de tydelig uttrykker at de setter pris på tilbakemeldingene.

I en norsk kontekst finnes også forskning på elevers og studenters syn på tilbakemeldinger på et generelt nivå, men mindre i skriveopplæring spesielt (Havnes et al., 2012). Gamlem og Smith (2013) finner, i en studie som undersøker ungdomsskoleelevers oppfatninger av lærers tilbakemeldinger i klasserommet, blant annet at lærere sjelden legger til rette for bruk av tilbakemeldinger, og at klassen ofte er i gang med noe nytt når elevene får tilbakemelding på et arbeid. De framhever også at det ofte er frivillig om elevene skal jobbe videre med tilbakemeldinger eller ikke, og at det er større sjanse for at elevene nyttiggjør seg tilbakemeldingene om læreren legger til rette for det. Dette skjer, ifølge Gamlem og Smith, sjelden. Dette funnet er i samsvar med Bueie (2014; 2015; 2016), som i en rekke studier har utforsket elevers oppfatninger av lærers tilbakemeldingspraksis innenfor et Vfl-rammeverk. Her kommer elevperspektivet fram på individnivå på bakgrunn av både kvalitative intervjudata og kvantitative data. Bueie (2015) finner blant annet at elever stort sett forstår hva læreren skriver til dem, men at de bare jobber videre med en tekst om de får eksplisitt instruks om det. På bakgrunn av dette er derfor særlig lærers aktive tilrettelegging for bruk søkt tappet i undersøkelsen, fordi dette synes å være et viktig startpunkt for om elever handler på bakgrunn av informasjon i tilbakemeldingen eller ikke. Et eksempel på en indikator er: «Vi bruker tid i timene til å jobbe videre med tekster som har fått tilbakemelding.»

Hypotese 3 (H3) er at elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte, og hypotesen er hovedsakelig fundert på Vfl-teori, slik den er formulert av Black og Wiliam (2009, s 9): Når informasjon fra vurdering brukes i videre læringsarbeid, øker sjansene for læring. At lærere legger til rette for at elevene jobber videre med en tekst på bakgrunn av informasjon i en tilbakemelding, ses i min studie som å bruke informasjon fra vurdering til videre arbeid. Som vist, er det motstridende funn innen feltet og dessuten lite forskning

på skriving i en VfL-kontekst i Norge. Det er derfor et formål med denne studien å utforske hvordan og i hvilken grad elever oppfatter at norsklærerens vurderingspraksis samsvarer med sentrale sider ved en VfL-praksis, slik de er gjort rede for over, og i hvilken grad elevene oppfatter denne praksisen som nyttig. Undersøkelsens forskningsspørsmål er derfor om det er mulig å finne empirisk støtte for de teoridrevne hypotesene.

Tabell 1

Oversikt over hypoteser

H1	Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte.
H2	Elevers oppfatning av lærers ferdighet til å forklare fagstoffet er positivt statistisk assosiert med opplevd nytte.
H3	Elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte.

Metode

Utvalg, operasjonalisering, beskrivende statistikk

For å utforske styrken i statistiske assosiasjoner er det gjennomført en analyse av en spørreskjemaundersøkelse. Datasettet omfatter 213 elevrespondenter fra 12 klasser ved to videregående skoler på Østlandet, våren 2015. Valget av akkurat disse skolene var betinget av et ønske om diversitet i datagrunnlaget, og de deltok på direkte forespørsel fra forsker. Skolene har ulik profil med hensyn til inntakssnitt, elevsammensetning, studieretninger, antatt sosioøkonomisk status for nærområdet og etnisk sammensetning av elevgruppen. Utvalget er sammensatt av elever fra Vg1, Vg2 og Vg3 fra både studiespesialiserende og yrkesfaglige studieretninger og fanger en viss diversitet av elever i videregående skoler. Skolene deltok på tidspunktet for datainnsamlingen i et forskningsprosjekt som undersøkte seks læreres tilbakemeldingspraksiser, og denne delen av undersøkelsen ser på elevenes perspektiv. Det var frivillig for elevene å svare på undersøkelsen, og de var sikret full anonymitet. Hensikten var ivaretagelse av eleven som respondent i tråd med etiske retningslinjer. Forskningsprosjektet var tilrådet av NSD.

Elevene svarte på en papirversjon av undersøkelsen i en skoletime, distribuert av deres egne lærere. Ingen av elevene motsatte seg å delta. Med andre ord er det potensielle problemet med bortfallsskjevhet som skyldes at elever ikke ønsker å delta, ikke til stede. Elevene ble bedt om å

merke av på en sjupunktsskala med navngitte ytterpunkter hvilket synspunkt som samsvarer best med deres egen overbevisning. De tok stilling til påstander som er gjengitt i oversikten tabell 2.

Spørreskjemaundersøkelsen hadde til hensikt å kartlegge hvordan elevene oppfatter lærerens vurderingspraksis generelt og skriftlige tilbakemeldingspraksis spesielt på aggregert nivå, og hadde indikatorer om flere forhold enn det som er rapportert her, men som ikke er analysert av hensyn til denne artikkelens forskningsfokus. Variablene som søker å måle elevenes oppfatninger av ulike aspekter ved lærerens tilbakemeldingspraksis, er alle uttrykk for komplekse forhold, som vanskelig kan måles gjennom «enkelt-items» i en spørreundersøkelse. Derfor er det her tatt i bruk «multi-items» som en forskningstilnærming, det vil si at jeg har brukt flere indikatorer som måler et underliggende begrep.

Datasettet er tilrettelagt for analyse ved «full information maximum likelihood»-utregning (Kline, 2005). Totalt sett er problemet med bortfall neglisjerbart.

Analysemåte

Først gjennomførte jeg en beskrivende statistisk analyse (tabell 2). Deretter ble strukturell likningsmodellering brukt for å analysere relasjonene mellom variablene (figur 1). Metoden egner seg til å analysere forholdet mellom flere variabler ved hjelp av samme modell; en innføring i metoden finnes for eksempel hos Kline (2005). Modellen er basert på teori slik den er definert i artikkelen, og den viser sammenheng mellom vurderingsteoretiske forhold som er forankret i norskfaget. Siden dataene er samlet inn ved to skoler, kan de være sterkt klynget, og det ble derfor regnet ut en ICC-skåre (intraklassekorrelasjonskoeffisient) for hver indikator for å undersøke om klynging er et problem for undersøkelsens validitet. Resultatene er gjengitt i tabell 5.

I vurderingen av tilpassing mellom modell og datasett ble følgende mål brukt: «root mean square error of approximation» (RMSEA), «comparative fit index» (CFI), og «Tucker Lewis-index» (TLI). RMSEA mindre enn 0,05 og TLI og CFI større enn 0,95 indikerer en god føyning, mens RMSEA mindre enn 0,08 og TLI og CFI større enn 0,90 indikerer at det er en akseptabel tilpassing (Kline, 2005). Dette er føyningsmål som vanligvis brukes i SEM og som indikerer i hvilken grad det er samsvar mellom den teoretiske modellen og datasettet.

Strukturmodellen ble estimert med IBM SPSS Amos 24. Ellipsene representerer de latente variablene, sirklene representerer målefeil, og rektanglene representerer de observerte målte

variablene, det vil si påstandene/indikatorerne i spørreskjemaet. Stipilene indikerer en teoretisk anlagt kausalfortolkning, og stikoeffisientene langs stipilene reflekterer den målte styrken i assosiasjonen. Styrken på relasjonene øker med stigende tallverdi.

Verdiene til RMSEA, TLI og CFI indikerer samlet sett at strukturmodellen i figur 1 passer tilfredsstillende til data. P-verdien til kji-kvadrattesten bør ideelt være større enn 0,05 men er her større enn 0,00. Den lave verdien må ses som en svakhet med modellen.

Resultat

Tabell 2

Beskrivende statistikk og indikatorer

Kategorinavn	Indikator	M	SD	Cronbachs alfa
Nytte				
4.2	Tilbakemeldinger fra læreren gjør at jeg skjønner hva jeg skal gjøre for å lykkes bedre med neste oppgave i faget.	4,7	1,6	,847
5.6	Tilbakemeldinger er nyttige fordi de hjelper meg til å gjøre det bedre på neste vurdering.	5,0	1,8	
2.4	Når jeg får tilbakemelding fra norsklæreren på et skriftlig arbeid, får jeg et klarere bilde av hva som forventes av meg.	4,5	1,7	
Bruk				
5.1	Læreren sørger for at elevene jobber videre med tekster etter at de har fått tilbakemelding.	4,6	1,8	,795
5.2	Jeg har mulighet til å levere en ny versjon av en oppgave eller prøve etter at jeg har fått tilbakemelding og karakter.	4,5	1,8	
5.4	Vi bruker tid i timene til å jobbe videre med tekster som har fått tilbakemeldinger.	4	1,9	
Forklaring				
4.3	Norsklæreren forklarer i tilbakemeldingene hva jeg skal gjøre annerledes neste gang.	4,7	1,7	,824
4.4	Læreren viser med eksempler hvordan jeg kunne skrevet på en bedre måte.	4,0	1,9	
4.5	Lærer forklarer helt konkret hva jeg skal gjøre for å bli bedre i tilbakemeldingene.	4,3	1,7	
Kriterier				
3.2	Tilbakemeldingen fra norsklæreren er knyttet til kriterier som er spesielle for hver enkelt oppgave.	4,3	1,6	,749
3.5	Norsklæreren bruker kriteriene for oppgavene i skriftlig norsk når han/hun gir tilbakemelding på skriftlige arbeider.	4,9	1,6	

Latente variabler som måler elevenes oppfatninger av lærers tilbakemeldingspraksis er uthevet. M = gjennomsnittsskåre, SD = standardavvik. Cronbachs alfa måler den interne konsistensen mellom

indikatorene som ligger til grunn for den latente variabelen. Tabellen viser at elevene jevnt over har en nøytral (4 blir her et nøytralt midtpunkt på en sjupunkts skala) til positiv oppfattelse av lærernes tilbakemeldingspraksis, målt med disse indikatorene. På en sjupunktskala kan standardavvik på 1 anses som moderat og 2 som forholdsvis høyt (Christophersen, 2012).

Indikatorer som måler den latente variabelen opplevd nytte, har en gjennomsnittskåre på mellom 4,5 og 5, noe som kan indikere at elevene oppfatter at lærerens tilbakemeldinger alt i alt er moderat nyttige. Samtidig er standardavviket på 1,6–1,8 noe som indikerer en betydelig variasjon i hvor nyttig elevene oppfatter lærerens tilbakemeldinger.

Indikatorer som måler den latente variabelen bruk, viser en gjennomsnittskåre på noe over 4. 4 er et nøytralt midtpunkt og kunne indikert at elevene ikke hadde noen klar formening om dette, men her er imidlertid standardavviket høyere enn for opplevd nytte og ligger på 1,8–1,9. Dette kan tolkes som et uttrykk for at elevene oppfatter at lærerne i ulik grad legger til rette for bruk, det vil si at noen gjør det i større grad enn andre og at færre har svart det nøytrale alternativet «4».

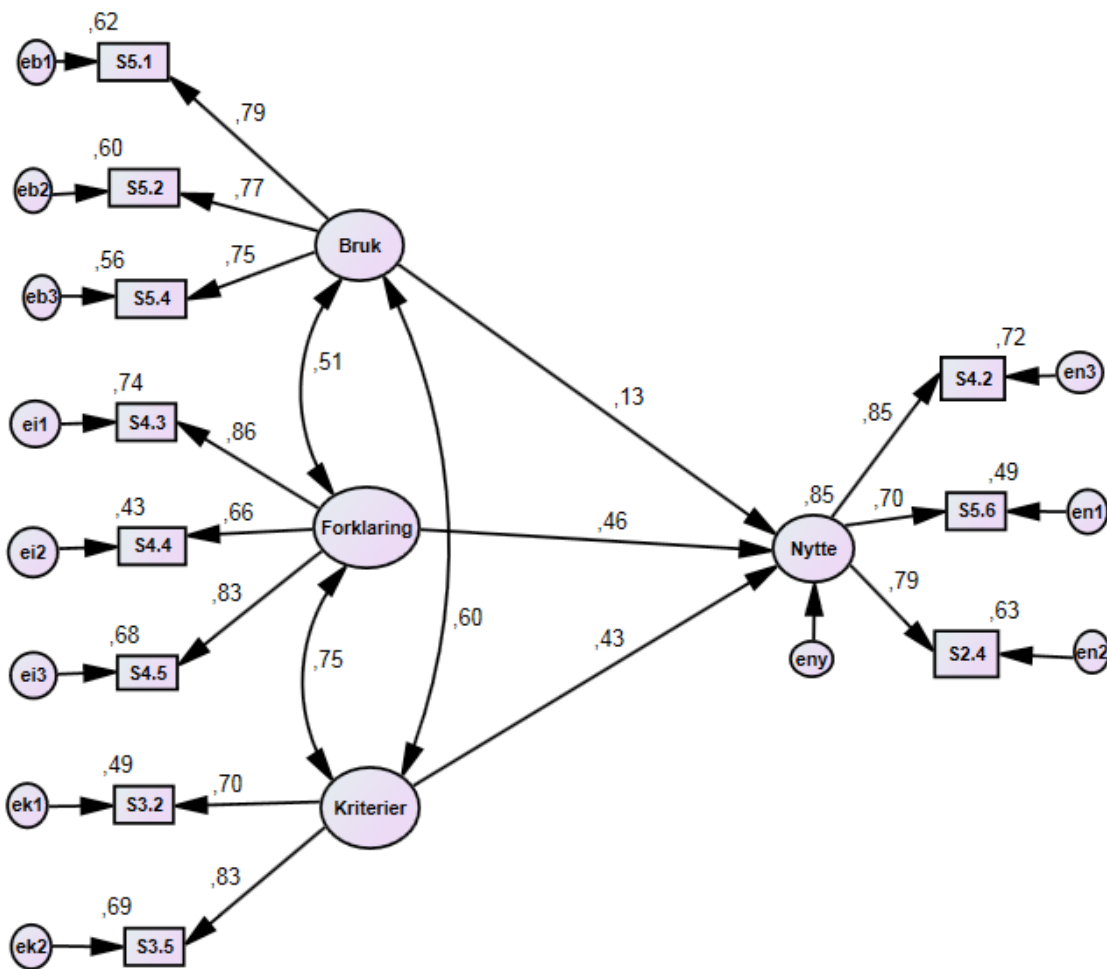
Begrepet forklaring er operasjonalisert gjennom indikatorer som tapper ulike aspekter ved det å gi instruksjon til elevene. At lærer er konkret og modellerer, samt at eleven har en forståelse av hva som skal gjøres bedre neste gang, har et spenn mellom 4,0 og 4,6, noe som indikerer at elevene dels har en nøytral, dels har en svakt positiv oppfatning av hvordan lærer forklarer hvordan de kan forbedre seg og slik komme nærmere målet med opplæringen. Også her er det standardavvik som er forholdsvis høye, mellom 1,7 og 1,9. Med andre ord oppfatter elevene forskjeller i hvordan lærere instruerer.

Begrepet kriterier er en variabel som undersøker hvorvidt elevene oppfatter at læreren har og bruker kriterier når hun gir tilbakemelding på skriftlige oppgaver. Her oppfatter elevene i svakt positiv forstand at læreren har slike kriterier, og standardavviket på 1,6 er noe lavere enn på de øvrige variablene, noe som sier at spredningen er lavere i spørsmål om kriterier

Alt i alt tegner den deskriptive statistikken her et bilde av elever som har en moderat positiv oppfatning av de målte sidene ved lærers vurderingspraksis, men at det er betydelig spredning. Cronbachs alfa-verdiene tyder dessuten på at indikatorene som måler de latente variablene har en akseptabel intern konsistens (Bland & Altman, 1997).

Figur 1 Elevers oppfatning av lærers skriftlige tilbakemeldingspraksis

Strukturmodell tilbakemeldingspraksis



Standardized estimates
 kji = 78,134; df = 38; p-kji = ,000; rmsea = ,071; cfi = ,965; tli = ,939

Figur 1 viser den estimerte strukturmodellen. I det følgende vises resultatene av estimatene av modellen.

Tabell 3**Resultater av hypotesetesting**

yHy- potese	Formulering	Resultat	p-verdier
1	Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte.	Assosiasjonen $b(\text{Kriterier} \rightarrow \text{Nytte}) = ,43$ er signifikant og støtter hypotesen.	0,000
2	Elevenes oppfatning av lærerens evne til å forklare fagstoffet er positivt statistisk assosiert med opplevd nytte.	Assosiasjonen $b(\text{Forklaring} \rightarrow \text{Nytte}) = ,46$ er signifikant og støtter hypotesen.	0,000
3	Elevers opplevde tilrettelegging for bruk av tilbakemeldinger er positivt statistisk assosiert med opplevd nytte.	Assosiasjonen $b(\text{Bruk} \rightarrow \text{Nytte}) = ,13$ er ikke signifikant og støtter ikke hypotesen.	0,127

I figur 1 framkommer det at hypotese nummer 1 finner empirisk støtte, og den positive statistiske assosiasjonen $b(\text{Kriterier} \rightarrow \text{Nytte}) = ,43$ indikerer altså en ganske sterk sammenheng (Kline, 2005, s 122) mellom det at lærer bruker kriterier og at elever opplever tilbakemeldinger som nyttige. Det er likevel usikkerhet knyttet til funnet, og dette blir diskutert videre nedenfor. Hypotese 2 finner empirisk støtte og assosiasjonen er signifikant. Slik variabelen forklaring er operasjonalisert her, vises en sterk sammenheng (Kline 2005, s 122) $b(\text{Forklaring} \rightarrow \text{Nytte}) = ,46$ mellom elevers opplevelse av at læreren gir konkrete råd og elevers opplevde nytte av tilbakemeldingen. Den tredje hypotesen som tilsier at elevenes opplevelse av lærers tilrettelegging for bruk av tilbakemeldingene henger sammen med opplevd nytte, har en sterk forankring i teorien (Black & Wiliam, 2009). I definisjoner av vurdering for læring spiller bruk av resultater fra vurdering til videre læringsarbeid en sentral rolle (Black & Wiliam, 2009). Det er derfor et sentralt funn at elevenes opplevde nytte av lærers tilrettelegging for bruk av tilbakemeldinger ikke er statistisk signifikant.

For å undersøke modellens begrepsvaliditet ble det gjort en konfirmerende faktoranalyse av to og to faktorer, og tabell 4 viser ulik grad av tilpassing mellom modell og data (teori og empiri).

Tabell 4**Konfirmerende faktoranalyse: Relevante fit-opplysninger**

Faktorer	Antall Ind.	Faktorladninger		Kji-kvadrat	Frihetsgrader	p-kji	RMSEA	CFI	TLI
		Min	Max						
Bruk	213								
Nytte		0,67	0,85	25,349	8	0,001	0,101	0,966	0,911
Forklaring		0,69	0,86	10,059	8	0,261	0,035	0,996	0,989
Kriterier		0,71	0,86	18,134	4	0,001	0,129	0,959	0,848
Forklaring	213								
Nytte		0,65	0,89	10,596	8	0,226	0,039	0,996	0,989
Kriterier		0,71	0,87	13,419	4	0,009	0,105	0,978	0,919
Kriterier	213								
Nytte		0,68	0,85	13,419	4	0,009	0,105	0,987	0,919

Testingen av to og to faktorer viser at modellen er velspesifisert i ulik grad, at operasjonaliseringen av kriterier kan diskuteres og at det må tas forbehold når de statistiske sammenhengene knyttet til kriterier vurderes. Her bør en videreutvikling av instrumentet særlig fokusere på operasjonaliseringen av denne latente variabelen.

Teoretisk sett er det helheten i modellen som kan bidra til å forklare samsvaret mellom opplevd VFL-praksis og varians i opplevd nytte. Eksempelvis er det vanskelig å se noen grunn til at elever skulle oppleve samsvar mellom at lærer har kriterier i norskundervisningen og at lærer legger til rette for bruk av tilbakemeldingen isolert sett. For en videreutvikling av modellen vil framtidig forskning kunne avgjøre om andre indikatorer bedre kan operasjonalisere de teoretisk funderte variablene. Det bør også nevnes at modellen passer bedre til datasettet om indikatoren «Vi bruker tid i timene til å jobbe videre med tekster når vi har fått tilbakemeldinger» tas ut ($K_{ji} = 47,873$, $df = 29$, $p\text{-kji} = 0,015$, $RMSEA = 0,055$, $CFI = 0,981$, $TLI = 0,964$). Faglig-substansielt tapper imidlertid denne indikatoren et helt sentralt aspekt ved lærers tilrettelegging for bruk av tilbakemeldinger. Når det skjer i timene, kan læreren svare på spørsmål og følge opp arbeidet på en helt annen måte enn hvis det gis som hjemmearbeid. Som en kan se av figur 1, er fit-skårene noe svakere med denne indikatoren, men faglig-substansielt vurderer jeg at denne modellen er mest relevant.

Siden dataene er samlet på to skoler, kan de være uttrykk for hvordan elevene oppfatter tilbakemeldingspraksisen til sine seks lærere mer enn at de speiler en større populasjon. For å undersøke om dette er tilfellet, ble det regnet ut intraklassekorrelasjonskoeffisiensskårer (ICC) for

hver indikator. Verdier høyere enn 0,05 kan indikere at dataene er klynget (Christophersen, 2012). Resultatene av beregningene er gjengitt i tabell 5.

Tabell 5

Intraklassekorrelasjonskoeffisientskårer, gjengitt per indikator

Indikator	S2.4	S3.2	S3.5	S4.2	S4.3	S4.4	S4.5	S5.1	S5.2	S5.4	S5.6
ICC	,00873	,13052	,00623	,00623	,00838	-	-	,16697	,17360	,04902	,09893

Tabell 5 viser at fem indikatorer har verdier over 0,05, mens fire har under. To indikatorer lot seg ikke estimere. Samlet sett gir ikke ICC-skårene noe klart bilde av om det er et problem at dataene er klynget eller ikke, og det kan derfor være aktuelt å teste ut gyldigheten av resultatene på en større populasjon. Også dette er en oppgave for framtidig forskning.

Diskusjon

Et sentralt formål med strukturmodellen var å utforske sammenhenger mellom latente variabler som står sentralt i teorifeltet som beskriver hvordan vurdering kan bidra til læring. Studien har undersøkt tre hypoteser og resultatene blir diskutert i det følgende.

En hypotese var hvis elevene oppfattet at lærers tilbakemeldingspraksis var i tråd med VfL, var den også statistisk assosiert med graden av opplevd nytte. I et overordnet perspektiv kan det se ut som om elevene oppfattet lærers tilbakemeldingspraksis nøytralt til moderat nyttig ($M=4,5-5$), men at det samtidig er standardavvik på (1,6–1,8) som indikerer at det er en betydelig varians i hvordan elevene har svart. Som vist, er det positive statistiske assosiasjoner mellom forklaring, kriterier og nytte, mens bruk ikke har en statistisk signifikant sammenheng med opplevd nytte. I figur 2 (vedlagt) ble bruk utforsket som medierende variabel, uten at det overordnede bildet endret seg; elevenes oppfatninger av hvor nyttige tilbakemeldingene er, er ikke statistisk assosiert med lærers tilrettelegging for bruk av tilbakemeldingene.

Kriterier er målt med to indikatorer og som analysene viser, er dette konstruert modellens svakest spesifiserte. Sammenhengen mellom at elevene opplever at lærer har og bruker kriterier og opplevd nytte må, med den operasjonaliseringen som foreligger her, vurderes som usikker, og det trengs ytterligere validering av instrumentet for at slutninger om denne sammenhengen skal kunne anses som holdbare. Rasjonale som ligger til grunn for valget av uavhengig variabel, er at

sentrale teoretikere, f.eks. Sadler (1989), i likhet med norske utdanningsmyndigheter (Udir, 2010), framhever at elevene må kjenne til målet med opplæringen for at vurdering skal ha større sjanse for å bidra til læring. På bakgrunn av dette ses kriterier i denne studien som et uttrykk for målet med opplæringen. Sadler (1989) opererer som nevnt med cirka 50 latente kriterier for hva som kan regnes som en god tekst hos førsteårsstudenter på universitetsnivå. Det er nærliggende å tenke at en elev i videregående skole møter omtrent de samme kravene, i hvert fall en avgangselev, og at skrivefaglige kriterier for hva som kan regnes som en god tekst, har store likhetstrekk.

Et problem som følger i kjølvannet av denne kompleksiteten, er hvilken status kriteriene dermed får. De kan ikke uttrykke målet med opplæringen fullt ut, i så fall måtte de være svært omfattende og spesifikke, og dermed bli vanskelig for elevene å ta inn og ikke minst anvende. Sadler (2010) viser dessuten at studenters og læreres tolkning av kriterier kan avvike. En mer hensiktsmessig praksis vil kreve at læreren velger ut noen kriterier eller peker på hvilke det er særlig viktig å være oppmerksom på for å lykkes (Kvithyld & Aasen, 2011). Det er ikke grunnlag for å uttale seg om lærerne faktisk gjør det i denne studien, ettersom jeg ikke presenterer empiri om dette. Kriterier er likevel valgt som indikator på om elevene gjøres kjent med målet med opplæringen fordi det er nærliggende å tenke seg at et fravær av kriterier vil kunne bidra til at elevene er henvist til å gjette hva som skal til for å lykkes med faget, noe som også er i strid med opplæringsloven. Det har dessuten vært en satsing på utvikling av kriterier siden 2007 (Bedre vurderingspraksis) (Udir, 2009). Om denne satsingen har manifestert seg i bedre praksis, er et åpent spørsmål, men elevene i denne studien har i hvert fall svart nøytralt til positivt at de oppfatter at læreren har og bruker kriterier ($M=4,6$; $SD 1,6$) og at de finner kriterier nyttig. Det betyr ikke nødvendigvis at de lærer mer, men kan likevel tjene som grunnlag for å hevde at kriterier i norskopplæringen oppfattes som, og er, positivt. Standardavviket viser samtidig at det er betydelig spredning, noe som kan indikere at elevene oppfatter at lærerne gjør dette i ulik grad.

At det er en positiv statistisk assosiasjon mellom det at elevene oppfatter at læreren har og bruker kriterier og at tilbakemeldingene oppfattes som nyttige, taler for at en fortsatt satsing på utvikling og bruk av vurderingskriterier er gunstig, gitt at kausalretningen er slik modellen foreslår, nemlig at bruk av kriterier fører til en økt opplevelse av nytte. Dersom det er en toveis relasjon mellom nytte og bruk, rokker likevel ikke dette ved slutningen om at bruk av vurderingskriterier er gunstig. Dette resonnementet betinger at ytterligere validering av instrumentet lykkes og må følgelig ses som tentativt.

I en studie av Parr og Timperley (2010) kobles morsmåslæreres evne til å gi gode tilbakemeldinger, definert innenfor et VFL-rammeverk, med elevens læringsframgang. En sentral komponent i den gode tilbakemeldingen, slik den er operasjonalisert av Parr og Timperley (2010), er at den inneholder en instruks eller forklaring som hjelper eleven til å redusere avstanden mellom egen prestasjon og det som er ønsket måloppnåelse. Gitt at målet er komplekst, slik det er argumentert for over, kan en tenke seg at eleven trenger en forklaring fra læreren for å vite hva hun/han skal gjøre for å bli bedre. Dermed kan denne praksisen bidra til at eleven kommer nærmere målet med opplæringen, enten ved å forbedre den aktuelle teksten, noe som av Hattie og Timperley (2007, s 90) omtales som en tilbakemelding på oppgave- eller prosessnivå, den gis underveis i en skriveprosess og har større sjanse for å bidra til læring, eller at den rettes mot dypere liggende tekststrukturer eller faglige forhold som har overføringsverdi fra gang til gang (Dysthe, 2011). I figur 1 framgår det at det er en forholdsvis sterk statistisk assosiasjon $b(\text{Forklaring} \rightarrow \text{Nytte}) = ,46$ mellom det at elevene opplever at læreren forklarer hva de skal gjøre for å bli bedre og opplevd nytte. Dette funnet kan sies å understøtte de antakelsene som ligger nedfelt i teorifeltet. Samtidig er det også her en betydelig spredning, noe som kan indikere at det er markante forskjeller i hvordan elevene opplever lærerens tilbakemeldingspraksis på dette punktet – noen lærere forklarer mye, andre lite, og de gjør det med ulik kvalitet, sett med elevenes øyne. Dette funnet kan ses i sammenheng med f.eks. Sommers (1982) mye siterte studie, som viser til at studenter får vage og generelle tilbakemeldinger. Helhetsbildet her kan likevel sies å være moderat positivt.

Om elevene opplevde at lærer la til rette for bruk av tilbakemeldinger i skriveopplæringen, er den siste latente variabelen i modellen. Som beskrevet over, er bruk en helt sentral faktor i VFL-teorien, og Black og Wiliam (2009) legger en avgjørende vekt på at informasjon om læringsresultater blir brukt i den videre undervisningen. Hvis informasjon om læring ikke brukes til å forbedre resultater eller utvikle ferdigheter, kan potensielle læringsmuligheter gå tapt. Hvordan elevene opplevde lærerens tilrettelegging for bruk av tilbakemeldinger i norskundervisningen, er assosiert med graden av forklaring ($,52$) og kriterier ($,61$). Dette kan tolkes som et uttrykk for at de tre uavhengige variablene, slik teorien også forutsier, har en sammenheng. Strukturmodellen i figur 1 viser en svak positiv sammenheng (Kline, 2005, s 122) mellom bruk og nytte, men denne assosiasjonen er ikke en signifikant statistisk assosiasjon: $b(\text{Bruk} \rightarrow \text{Nytte}) = ,13$, $p > 0,127$. Denne svake assosiasjonen er noe overraskende, sett i lys av den sterke satsningen som har vært i norsk videregående skole for å utvikle en læringsfremmende undervisningspraksis. Det kan ikke utelukkes at det kan henge sammen med operasjonaliseringen av bruk i denne studien, men like fullt gir det en pekepinn om at

det som av flere teoretikere framheves som mest sentralt i VFL-teorien, av en eller flere årsaker ikke henger sammen med opplevd nytte hos elevene. Et premiss for resonnementet er at elevene opplever en tilbakemelding som nyttig når de lærer av den. Det kan også være andre forhold som ligger under, for eksempel at en tilbakemelding framstår som effektiv og tydelig (Straub, 1997) eller at den samsvarer med elevenes forventninger (Gibbs & Simpson, 2004), men alt i alt er det grunn til å anta at de fleste elever er på skolen for å lære, og da framstår funnet som en indikasjon på et urealisert potensial for forbedring.

En mulig forklaring på funnet kan ligge i at det er forholdsvis lav forekomst av tilrettelegging for bruk av tilbakemeldinger. Bueie (2015) finner at elever kun gjør det de blir bedt om av lærer. Sadler (2010) viser også at lærere og studenter kan ha ulike forståelser av oppgaven – av kvalitet og av betydningen av formuleringer av kriterier. Derfor kan det tenkes at elevene ikke forstår lærerens forslag til forbedringer, og at noe av forklaringen dermed ligger i innholdet i tilbakemeldingen. Hvorvidt denne forklaringen er gyldig, bør avklares i videre forskning om dette fenomenet.

Tidspunktet for tilbakemeldingen kan også være en mulig forklaring på funnet. Når elevene får tilbakemeldinger, har de ofte begynt å jobbe med noe nytt, med nye skriveoppgaver og nytt fagstoff; skolekonteksten er sånn at elevene har en jevn strøm av oppgaver de skal gjøre, også vurderinger som karaktersettes i ulike fag, og det er dermed rasjonell atferd å jobbe med de fagene som til enhver tid har nærmeste frist for innlevering, forhold som påpekes av Lea og Street (1998) og Gamlem og Smith (2013). Slik kan tilbakemeldinger som handler om tekster de har skrevet for kanskje flere uker siden, og som handler om et fagstoff de i tillegg anser seg ferdig med, oppleves som lite nyttige å jobbe noe mer med. Noe som kan understøtte en slik opplevelse, er om tilbakemeldinger er rettet mot lokale forhold i teksten, slik at det gir begrenset mening for eleven å rette dem opp eller jobbe videre med dem fordi teksten er et tilbakelagt arbeid. Det kan også tenkes at elevene bare i noen grad ser på skriveingen som en ferdighet de utvikler over tid. De ser ikke noe poeng i å jobbe med arbeidskrevende, strukturelle tekstforhold som avsnittsoppbygging og innledning, eller raffinering av analyse- eller argumentasjonsteknikker, fordi det er krefter investert i noe som ikke resulterer i bedre karakter på den aktuelle oppgaven, selv om det kan gjøre det på lang sikt.

Elevene kan oppleve at tonen i tilbakemeldingen er negativ, og at de derfor ikke blir motiverte til å jobbe videre (Straub, 1997), eller at det er andre sider ved innholdet i tilbakemeldingen som gjør at elevene ikke finner lærers tilrettelegging for arbeid med tilbakemeldingene som nyttig (Hattie & Timperley, 2007; Sadler, 2010). Studiens empiri kan ikke si noe om disse forklaringene er plausible

eller ikke. Det kan også tenkes – ut fra studien til Gibbs og Simpson (2004) – at elevene mangler teknikker for å bruke tilbakemeldinger fra læreren på en konstruktiv måte, eller at de ikke har fått opplæring. Studien viser samtidig at i hvert fall en del av elevene opplever at læreren legger til rette for bruk av tilbakemeldinger – og dette forholdet peker i motsatt retning. En slutning som kan trekkes, er at videre forskning bør følge opp det empiriske funnet som er drøftet her.

Studiens begrensninger

Denne studien er en tverrsnittsundersøkelse. Det betyr at man strengt talt ikke har noe bevis for kausalretning når stikoeffisienter estimeres (Kline, 2005). På den annen side er det plausibelt at for eksempel opplevd nytte kan influere bruk av tilbakemeldinger og motsatt. Med andre ord er det antakelig en toveis sammenheng mellom variablene, som det er vanskelig å avklare uten eksperimentelle tilnæringer. Kontrollerte eksperimenter har vært utenfor rekkevidde for denne studien, men det ville ideelt sett vært verdifullt å gjennomføre slike studier under eksperimentelle betingelser. Det finnes teoretiske begrunnelser for å anta at variablene er relatert, og når empiriske assosiasjoner finnes, er det grunn til å feste lit til at sammenhengene ikke er spuriøse. Det vi kan fastslå, er at den interne validiteten kan anses som tilfredsstillende (Shadish, Cook & Campbell, 2002) selv om det kan hefte usikkerhet om selve kausalretningen.

Utvalget speiler en diversitet av elever i videregående skole. Det betyr at vi kan feste lit til at de empiriske funnene kan ha gyldighet for noe mer enn utvalget (Shadish, Cook & Campbell, 2002). Her må det imidlertid understrekes at det likevel hefter en usikkerhet om utvalget er representativt, og at denne slutningen bygger på skjønn. Det betyr at en ytterligere ekstern validering av denne studiens funn ville vært interessant, noe også ICC-skårene viser. Større utvalg enn det som denne studien er basert på, vil kunne styrke den statistiske konklusjonsvaliditet.

Begrepene er forholdsvis tett forbundet på et teoretisk nivå. Dette gjør den empiriske prøvingen krevende. Her er problemet forsøkt løst ved å ta utgangspunkt i den enkleste modellen man kan tenke seg, for så å øke på med flere forklaringsvariabler. Det kan tenkes at framtidig forskning vil kunne tjene på å forenkle kjernen i min modell for så å øke på med andre, distinkte variabler (for eksempler motivasjon for norskfaget, utøvd selvdisiplin i arbeidet med skriftlige produkter etc.). Dette understreker at forskeren står overfor et avveiningsforhold mellom å prioritere fokus på kjernen i vurdering for læring-problematikk eller å fokusere på distinkt annerledes faktorer som kan tenkes å influere dette.

Undersøkelsen er utviklet for denne studien spesielt og bygger ikke på internasjonalt validerte instrumenter. Grunnen er at det ikke har vært mulig å finne slike. Videre forskning vil kunne bidra til ytterligere validering av instrumentet.

Konklusjoner

Sett under ett kan det se ut som om elever oppfatter at læreren har og bruker kriterier, at hun forklarer hva de skal gjøre for å forbedre seg og i noen grad legger til rette for bruk av tilbakemeldinger, altså slik VfL-teorien og dessuten retningslinjer og lovverk tilsier at de skal. Det framgår også at elevene i store trekk opplever lærers tilbakemeldinger som moderat nyttige, men med variasjon. Dette tegner et bilde av skoler som i det store og hele har tilpasset seg retningslinjene for god vurderingspraksis. Samtidig indikerer høy spredning at det er betydelige forskjeller mellom lærerne, noe som kan tyde på et urealisert potensial i spørsmålet om alle klasser opplever en vurderingspraksis som er i tråd med VfL-anbefalingene. Det mest iøynefallende funnet i studien er at det kan se ut som om elevene oppfatter at lærere på ulike vis legger til rette for bruk av tilbakemeldinger, og at elevene både har muligheter til og krav om å levere tekster på nytt eller i hvert fall jobbe med dem i timen, men at de ikke finner dette nyttig. En forklaring kan være at de er ferdige med det aktuelle fagstoffet og skriveoppgaven, men dette er en antakelse som trenger mer forskning for å kunne besvares.

Videre forskning

I tillegg til det som allerede er nevnt, ville det vært interessant å vite mer om funnene er gyldige også for andre fagkontekster. Det er ikke gitt at funn som gjelder tilbakemeldingspraksis i norskfaget kan overføres til eller har relevans for fag som geografi, naturfag eller samfunnsfag, kanskje heller ikke for fremmedspråkene. Kunnskaper om dette utgjør et betydelig potensiale for informasjon om en praksis som fremmer læring (Hattie & Timperley, 2007), gitt at denne forskningen også kan omformes til gode, konkrete råd som lærere kan anvende i sin undervisning. Et annet relevant felt for videre forskning er å se innholdet i lærers tilbakemelding i sammenheng med elevens oppfatning av den, i et større utvalg, og dessuten elevens oppfatninger av tilbakemeldinger på både ulike teksttyper og på ulike stadier i skriveprosesser.

Implikasjoner for praksis

Studien viser at elever ikke opplever noen signifikant sammenheng mellom lærers tilrettelegging for bruk av tilbakemeldinger og opplevd nytte, og siden teorilitteraturen er så klar i anbefalingen av at tilbakemeldinger brukes i videre læringsarbeid, indikerer dette tapte læringsmuligheter.

Tilbakemeldinger om hvordan elevene kan minske avstanden mellom mål og ønsket mål, har størst mulighet for å bidra til læring når elevene følger opp på en dedikert måte. Her ser det ut til å være et betydelig forbedringspotensial. Studier av lærers arbeid vil kunne si om de trenger tydeligere rammer eller insentiver som kan bidra til mer effektiv bruk av tilbakemeldingsinformasjon i undervisningen.

Referanser

- Beach, R. (1989). Evaluating writing to learn: Responding to journals. In Lawson, B., Ryan, S. & Winterowd, W. R. (Eds.), *Encountering student texts: Interpretive issues in reading student writing* (s. 183–197). Urbana, IL: National Council of Teachers of English.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. [dx.doi.org/10.1080/0969594X.2010.513678](https://doi.org/10.1080/0969594X.2010.513678)
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* (formerly: *Journal of Personnel Evaluation in Education*), 21(1), 5–31. doi:10.1007/s11092-008-9068-5
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education*, 5(1), 7–74. [dx.doi.org/10.1080/0969595980050102](https://doi.org/10.1080/0969595980050102)
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314 (7080), 572. doi.org/10.1136/bmj.314.7080.572
- Bueie, A. A. (2014). «Jeg synes det er ett fett om det er 'klarte' eller 'hadde'.» Om elevers forståelse og reaksjon på lærers tekstrespons i norskfaget. I Aasen, A. J. m. fl. (Red.) *Skriv! Les!* Trondheim: Akademika forlag.
- Bueie, A. A. (2015). Summativ vurdering i formativ drakt – elevperspektiv på tilbakemelding fra heldagsprøver i norsk. *Acta Didactica Norge*, 9(1), Art-4.
- Bueie, A. A. (2016). Nyttige og mindre nyttige lærerkommentarer – slik elevene ser det. *Nordic Journal of Literacy Research*, 2(1). [dx.doi.org/10.17585/njlr.v2.188](https://doi.org/10.17585/njlr.v2.188)
- Calhoun-Dillahunt, C., & Forrest, D. (2013). Conversing in Marginal Spaces: Developmental Writers' Responses to Teacher Comments. *Teaching English in the Two Year College*, 40(3), 230–247.
- Christophersen, K. A. (2012). *IBM SPSS / AMOS. Databehandling og statistisk analyse*. 5. utgave. Oslo: Akademika Forlag.
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education*, 27(2), 199–214. [dx.doi.org/10.1080/0950069042000276721](https://doi.org/10.1080/0950069042000276721)
- Dysthe, O. (2011). 'What is the Purpose of Feedback when Revision is not Expected?' A Case Study of Feedback Quality and Study Design in a First Year Master's Programme. *Journal of Academic Writing*, 1(1), 135–142. [dx.doi.org/10.18552/joaw.v1i1.26](https://doi.org/10.18552/joaw.v1i1.26)
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. New Jersey: Routledge.
- Gamlem, S. M., & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), 150–169. [dx.doi.org/10.1080/0969594X.2012.749212](https://doi.org/10.1080/0969594X.2012.749212)
- Gibbs, G., & Simpson, C. (2004). Does your assessment support your students' learning. *Journal of Teaching and Learning in Higher Education*, 1(1), 3–31.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research* 77(1), 81–112. [10.3102/003465430298487](https://doi.org/10.3102/003465430298487)
- Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1), 21–27. doi.org/10.1016/j.stueduc.2012.04.001
- Hertzberg, F. & Roe, A. (2016). Writing in the content areas: a Norwegian case study. *Reading and Writing*, 29(9). 555–576. DOI:10.1007/s11145-015-9607-7

- Hopfenbeck, T., Tolo, A., Florez, T., & El Masri, Y. (2013). Balancing Trust and Accountability? The Assessment for Learning Programme in Norway. Oslo: OECD. [10.1787/19939019](https://doi.org/10.1787/19939019)
- Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of second language writing*, 7(3), 255–286. [doi.org/10.1016/S1060-3743\(98\)90017-0](https://doi.org/10.1016/S1060-3743(98)90017-0)
- Jönsson, A. (2013). Facilitating productive use of feedback in higher education. *Active learning in higher education*, 14(1), 63–76. [10.1177/1469787412467125](https://doi.org/10.1177/1469787412467125)
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. 2nd edition. New York: The Guilford Press.
- Kronholm-Cederberg, A. (2009). Skolans responskultur som skriftpraktik: gymnasieters berättelser om lärarens skriftliga respons på uppsatsen. Åbo: Åbo akademis forlag.
- Kvithyld, T., & Aasen, A. J. (2011). Fem teser om funksjonell respons på elevtekster. *Tidsskriftet Viden om læsning*, 5, 10–16.
- Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in higher education*, 23(2), 157–172. dx.doi.org/10.1080/03075079812331380364
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), 149–170. dx.doi.org/10.1080/09695940701478321
- Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing writing*, 15(2), 68–85. doi.org/10.1016/j.asw.2010.05.004
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535–550. dx.doi.org/10.1080/02602930903541015
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), 119–144. doi:10.1007/BF00117714
- Sandvik, L. V. & Buland, T. H. (2014). Vurdering i skolen. Utvikling av kompetanse og fellesskap. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS). Trondheim: NTNU.
- Shadish, W.R, Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Wadsworth Cengage learning.
- Sommers, N. (1982). Responding to student writing. *College composition and communication*, 33(2), 148–156. DOI: 10.2307/357622
- Straub, R. (1997). Students' reactions to teacher comments: An Exploratory Study. *Research in the Teaching of English*, 31(1), 91–119. www.jstor.org/stable/40171265
- Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478. Doi: 10.1111/j.1467-8527.2005.00307.x
- Utdanningsdirektoratet (2010). Rundskriv Udir-1-2010. Individuell vurdering. <http://www.udir.no/Regelverk/Finn-regelverk-for-opplaring/Finn-regelverk-etter-tema/Vurdering/Udir-1-2010-Individuell-vurdering>
- Utdanningsdirektoratet (2009). Bedre vurderingspraksis, sluttrapport fra Utdanningsdirektoratet 2009. <http://www.udir.no/Tilstand/Forskning/Rapporter/Utdanningsdirektoratet/Bedre-vurderingspraksis-sluttrapport-fra-Utdanningsdirektoratet-2009/>

Vedlegg

Vedlegg 1

Kovariansmatrise fordelt på latente variabler

Matrisen viser blant annet at de eneste indikatorene som lader på alle de latente variablene er 4.2, 2.4 og 5.6 – som er indikatorene for nytte.

Standardized Total Effects (Group number 1 - Default model)

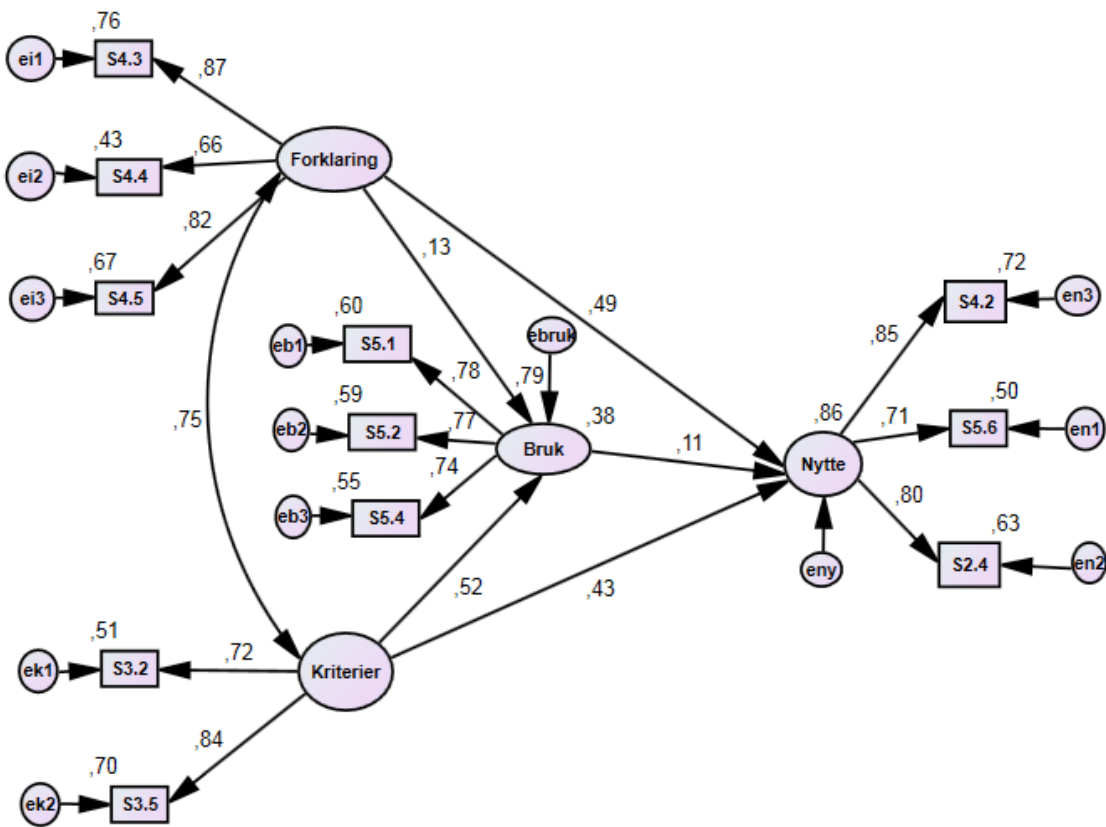
	Kriterier	Forklaring	Bruk	Nytte
Bruk	,516	,130	,000	,000
Nytte	,488	,501	,115	,000
S4.2	,413	,425	,097	,848
S3.5	,837	,000	,000	,000
S3.2	,716	,000	,000	,000
S5.4	,384	,097	,744	,000
S5.2	,398	,100	,771	,000
S5.1	,400	,101	,776	,000
S4.5	,000	,820	,000	,000
S4.4	,000	,656	,000	,000
S2.4	,388	,399	,091	,795
S5.6	,345	,355	,081	,707
S4.3	,000	,870	,000	,000

Vedlegg 2

Figur 2, strukturmodell med bruk som medierende variabel

Modellen bygger på resonnetmentet at Forklaring og Kriterier viser aspekter ved innholdet i tilbakemeldingen, mens Bruk handler om anvendelsen av tilbakemeldingen.

Tilbakemeldingspraksis med bruk som medierende variabel



Standardized estimates

kji = 73,097; df = 38; p-kji = ,001; rmsea = ,068; cfi = ,969; tli = ,954

Resultat

Figur 2 viser at den statistiske sammenhengen mellom elevenes opplevelse av lærers evne til forklaring og opplevd tilrettelegging for bruk ikke er statistisk signifikant $b(\text{Forklaring Bruk}) = ,13$ og støtter ikke hypotesen. Assosiasjonen mellom elevers opplevelse av lærers bruk av kriterier og opplevde systemer for bruk av tilbakemeldinger er statistisk signifikant ($,52$) og forholdsvis høy (Kline, 2005, s 122). Det er altså ikke vesentlige endringer i det totale bildet, selv om status for Bruk endres.

Artikkel 3

Norsklæreres erfaringer med læringsfremmende vurdering av elevtekster i to lærerkollegier

Ingress

Over mer enn et tiår har norske myndigheter satset på utvikling av en mer læringsfremmende vurderingspraksis i skolen. Kjernepunkter i denne satsningen er å fremme bruk av vurderingskriterier, læringsfremmende tilbakemeldinger og at lærerne legger til rette for bruk av informasjon fra tilbakemeldinger i videre læringsarbeid. Vi har imidlertid fortsatt lite forskningsbasert kunnskap om hvordan denne satsingen kommer til uttrykk i norskfagets skriveopplæring i videregående skole og hvordan norsklærere opplever at det er å jobbe med læringsfremmende vurdering. Denne studien utforsker hvordan dette arbeidet oppleves av lærere i to lærerkollegier der det er ulik grad av samarbeid i norskseksjonen. Ved hjelp av tematisk analyse av intervjudata viser studien at grad av samarbeid i norskseksjonen om å vurdere elevtekster og det å gi tilbakemeldinger ser ut til å være relatert til mulighetene for å realisere en læringsfremmende vurderingspraksis.

Introduksjon

Det er et påtrykk fra nasjonale myndigheter (Kunnskapsdepartementet, 2014) og KS (2015) for at lærere bør samarbeide mer om å utvikle læringsmål, bruke kriterier for grad av måloppnåelse og koordinere arbeidet med å utvikle en læringsfremmende vurderingspraksis. I Stortingsmelding nr. 11 (2008–2009) tildeles lærerne et klart definert ansvar for hva elevene lærer. Disse styringssignalene legger føringer for kollektive aspekter ved norsklærerens jobb med vurdere elevenes skriftlige arbeider. I tillegg er flere tiltak for å profesjonalisere vurderingspraksis iverksatt av nasjonale og lokale myndigheter. En kan særlig nevne Bedre vurderingspraksis (2007–2009), der et sentralt formål var å utforme vurderingskriterier for alle fag, og Vurdering for læring 2010–2014 som tok sikte på å utvikle en læringsfremmende vurderingspraksis. Endringer i opplæringsloven (Udir, 2010) stadfester dessuten at vurdering gjennom året skal bidra til læring. Forventninger om lærersamarbeid og læringsfremmende vurderingspraksis uttrykkes med andre ord i samme tidsperiode.

Siden retningslinjene for hva som regnes som god vurderingspraksis (Hopfenbeck et al., 2013) er generelt utformet, er det opptil lærere å tolke anbefalingene og realisere dem. Dette er en kompleks oppgave (Svanes & Skagen, 2016). Bennett (2011) argumenterer for at det trengs studier som kan vise hvordan de generelle prinsippene for god vurdering kan realiseres i ulike skolefaglige kontekster. Denne studien er et bidrag til dette. Norskfaget er sentralt i norsk skole (Hertzberg & Roe, 2016) og er derfor viktig å prioritere i forskning. Jeg undersøker spesielt hvordan læringsfremmende vurderingspraksis (heretter VfL) kan realiseres i norskfagets skriveopplæring fordi kunnskap om hvordan lærere oppfatter og reflekterer over sin egen og skolens vurderingspraksis kan nyttig for at vi bedre skal forstå mulighetene til forbedring av vurderingspraksisen.

Flere studier har på et mer overordnet plan kartlagt implementeringen av VfL-praksis i skolen (f.eks. Bueie, 2015; Burner, 2014; Havnes et al., 2012; Saliu-Abdulahy et al., 2017, Sandvik & Buland, 2014). Disse studiene er relatert til ulike skolefag. Det er imidlertid få studier som har forsket på vurdering i en norskfaglig kontekst i videregående skoler (Bueie, 2016; forfatter, 2017a; 2017b). Forskning om vurderingspraksis og tilbakemeldingspraksis i norskfaget kan være nyttig fordi man bedre kan forstå hvordan styringsintensjoner om VfL og lærersamarbeid kan bli realisert i skoler som i ulik grad har tradisjon for lærersamarbeid. Skolekultur er forhold som i liten grad er belyst i tidligere forskning, og det er relevant å gjøre koplinger mellom vurderingspraksis og skolekultur fordi begge felt er satsinger som bærer i seg et potensial for forbedre skolens praksis.

I det videre vil jeg først presentere sentral forskning om lærersamarbeid og om VfL. I særlig grad vil jeg fokusere på VfL i forbindelse med skriveopplæring. Deretter vil jeg presentere metoder og resultater fra en empirisk undersøkelse. Artikkelen avsluttes med en drøfting av sentrale funn og samt implikasjoner for å utvikle læreres læringsfremmende vurderingspraksis.

Lærersamarbeid

Over tid har det skjedd endringer i lærernes arbeidssituasjon fra en lærerrolle med stor grad av autonomi og individualitet (Lortie, 1975) til en situasjon med mer styring fra myndighetenes side (Kunnskapsdepartementet, 2008), blant annet gjennom krav til hva læreren skal gjøre, kontroll av lærerens arbeid (for eksempel grad av diskrepans mellom standpunkt karakterer og eksamensresultater) osv. Tendensen knyttes gjerne til begrepene profesjonalisering av læreryrket

(Sachs, 2016) og ansvarliggjøring av lærere (Ingersoll, 2011). Tendensen til å styre lærerens arbeid er en pågående, internasjonal trend (Evetts, 2013) der styringsintensjoner, standarder og kontroll vil kunne ha betydning for lærerens arbeidssituasjon.

Vangrieken et al. (2015) viser i en review-artikkel til potensielt gunstige konsekvenser av lærersamarbeid: samstemming av mål, mulighet for arbeidslettelser gjennom felles planlegging, økt tro på egne evner, støtte, læring internt, økt effektivitet og mer. Men lærersamarbeid har ifølge disse forfatterne også mulige ugunstige virkninger som opplevelse av autonomitap, konformitetspress, økt arbeidsbyrde, økt konfliktnivå og intern konkurranse. Det vil ofte være nødvendig for ledelsen ved en skole å finne en klok balanse mellom individuell autonomi og lærersamarbeid (Hargreaves, 1994; Leithwood & Riehl, 2003).

Noen studier viser dessuten at økt grad av lærersamarbeid kan bidra til bedre elevprestasjoner (Goddard, Goddard & Tschannen-Moran, 2007; Moolenaar, Slegers & Daly, 2012). Samarbeid om felles satsinger på en skole behøver imidlertid ikke å være gunstig: en studie av Dobbie og Fryer (2011) viser at kollektivt arbeid med å utvikle lokale læreplanmål ikke er statistisk relatert til bedre elevprestasjoner. Forskingen viser altså at lærersamarbeid er et komplekst felt som både har potensial til å kunne lette arbeidsbyrder og til å styrke lærere, men som også kan ha negative konsekvenser. Konsekvensene av lærersamarbeid er derfor usikre.

Vurdering for læring i norskfagets skriveopplæring

Vurdering for læring som forskningsfelt er svært omfattende. Et viktig, oppsummerende bidrag finnes hos Black og Wiliam (2009) som vektlegger følgende: Lærere bør gjennom sin vurderingspraksis bidra til at elevene kjenner *målet med opplæringen*, de får *tilbakemeldinger om hvordan de skal komme nærmere målet* og informasjon om elevenes læringsframganger *brukes* i det videre arbeidet. Denne forståelsen definerer VfL i min studie. Retningslinjer fra norske myndigheter for en læringsfremmende vurderingspraksis (Udir, 2010) bygger på anbefalingene fra forskningsfeltet (Black & Wiliam, 2009; Wiliam, 2011).

Vurderingskriterier kan ses som uttrykk for målet med opplæringen. Sentrale forskere som Torrance og Pryor (2001) og Sadler (2005) framhever fordeler med å bruke kriterier og vurderingsskjemaer, men begge framhever også at mer individuelt tilpassede læringsmål kommuniserer bedre til både elever og studenter. Sadler (2005; 2010) ser arbeid med kriterier som en måte lærende kan få tilgang

til implisitte tekstnormer. I skriveopplæringen er imidlertid målet med opplæringen svært komplekst (Sadler; 2010) og kjennetegnes av mange latente kriterier. Latente kriterier betyr her forhold ved teksten som blir synlige for leser når skriveren ikke mestrer dem, men som en ikke vil merke seg om de er utført korrekt. Dette viser hvor krevende det kan være å utforme klare kriterier som fanger opp hva som er vesentlig for å kunne skrive en god tekst.

Bruk av informasjon fra vurdering til videre læring er en annen sentral faktor i VfL-teori. I en skriveopplæringskontekst er bruk av læreres tilbakemeldinger til tekstforbedring et viktig uttrykk for denne dimensjonen. Ifølge MacArthur (2013) er revisjonskompetanse et kjennetegn ved en god skriver, og han viser i en reviewartikkel (MacArthur, 2012) at evne til å revidere tekst står sentralt i nær sagt all forskning på skriveopplæring. Her kan man se at VfL-teori (Black & Wiliam, 2009) og MacArthurs (2012) forskning har overlappende perspektiver. Graham et al. (2011) finner dessuten i en annen reviewartikkel om tilbakemeldinger i skriveopplæringen at lærers tilbakemeldinger kan ha positiv innvirkning på skriveferdigheter. Det samme gjelder når lærer legger til rette for at elevene vurderer hverandres tekster (hverandrevurdering). Hverandrevurdering og egenvurdering blir derfor oppfattet som potensielle teknikker som kan anvendes i en VfL-kontekst fordi de muliggjør tilbakemeldinger i prosess. Dette er et poeng Hattie og Timperley (2007) framhever som viktig for å øke mulighetene for læring. Topping (2009) og Sadler (2010) framholder dessuten hverandrevurdering som en effektiv måte å sørge for veiledende tilbakemelding, og de framhever også at elever og studenter kan gi hverandre gode tilbakemeldinger dersom de får god opplæring og klare kriterier å vurdere etter. Sadler (2010) viser også til at en student kan utvikle sin egen tekstkompetanse ved å vurdere andres tekster og gi tilbakemeldinger.

Bueie (2015; 2016) og forfatter (2017a; 2017b) har studert elevers syn på lærers tilbakemeldinger i norskfaget, men de er i mindre grad rettet mot lærers erfaringer med vurderingsarbeid og det å gi tilbakemeldinger. Det samme gjelder Saliu-Abdulahy et al. (2017) som ser på tilbakemeldinger i engelskfaget fra elevperspektivet. Det gjelder også Burner (2016) som fant at elever og lærere hadde ulike oppfatninger av tekstnormer. Elever og lærere må ses som partnere i læringsprosesser (Wiliam, 2011), og følgelig er det viktig å ha med begge perspektiver for å få et helhetlig forskningsperspektiv på fenomenet VfL-praksis.

Oppsummert viser altså forskningen at lærere generelt har opplevd utfordringer med å implementere en læringsfremmende vurderingspraksis, men vi har få studier som undersøker hvordan dette foregår i norskfaget. Gitt at lærere jobber sammen i kollegier, er det også rimelig å anta at

hvordan lærere samarbeider om vurdering i norskfaget kan påvirke mulighetene for å realisere en VfL-praksis. Forskningsspørsmålet i denne studien er derfor: Hva kjennetegner norsklæreres vurderingsarbeid ved to samarbeidskulturer?

Metode

Studien bygger på intervjudata og analyse av ulike dokumenter. Den empiriske undersøkelsen kan ses som en kasusstudie som er inspirert av Lijparts (1971) idé om komparasjon av to kasuser som er like langs mange dimensjoner, men ulike på det området man ønsker å undersøke, nemlig tradisjon for samarbeid mellom lærere som underviser i samme fag. To skoler deltok i studien. Data ble samlet inn gjennom skoleåret 2014–2015. Skole A var en kombinert skole, den hadde cirka 650 elever og hadde meget lave inntakskrav på studiespesialiserende studieretning (heretter ST) høsten 2014. Det betyr at alle som søkte, kom inn. Skole B hadde omtrent 600 elever, bare ST, og inntakssnittet høsten 2014 var 4,0. Denne skolen har en etablert samarbeidskultur blant lærere som i utpreget grad støttes av skolens ledelse (Eggen, 2009). Tre lærere ved hver skole sa seg villige til å delta, og de underviste i 12 norskklasser med til sammen 319 elever. Datamaterialet speiler ulike skolekontekster som i tillegg til forskjellige samarbeidskulturer også hadde ulike elevforutsetninger. Imidlertid hadde skolene likhetstrekk blant annet ved at de begge tilbød ST i videregående opplæring.

Gruppeintervju ble gjennomført ved begge skoler i mai 2015. Utformingen av intervjuguiden var basert på VfL-teori ved at den blant annet vektla kriterier og bruk av tilbakemeldinger i vurderingspraksis som er beskrevet i litteraturgjennomgangen. Eksempler på spørsmål som ble stilt var «Hvor nyttige synes dere kriterier er for elevenes læring?» og «Har du et system for elevenes videre arbeid med tekster, etter at de har fått tilbakemelding?». Et oppfølgingsintervju med en representant fra ledelsen og en av lærerne ble ved årsskiftet 2015–2016 gjennomført for å få mer kunnskap om arbeidet i norskseksjonen. Intervju ble valgt som metode for å få en utdypende og nyansert forståelse av lærernes refleksjoner over og holdninger til egen vurderings- og tilbakemeldingspraksis (Kvale & Brinkmann, 2010). Jeg valgte semistrukturert gruppeintervju fordi jeg ønsket å undersøke lærernes samtale om praksis som en del av et lærerkollegium og en skolekultur. For å søke å skape en atmosfære av tillit og respekt, innledet jeg intervjuene med å be lærerne fortelle om en vurderingssituasjon som de opplevde som spesielt vellykket. Dette fungerte etter hensikten: lærerne tok ordet på eget initiativ og fortalte om både vellykkede situasjoner og utfordringer, samt mangfoldet i egen vurderingspraksis. Intervjuene ble tatt opp og transkribert.

Tabell 1 – Oversikt over intervjudata

(Alle navn er pseudonymer)

Skole	Lærere	Type data	Varighet	Bakgrunns- data	Type data	Varig- het	Type data	Varighet
A	Anne Berit Carina	Gruppeintervju, transskribert	<u>0.45</u>	Se vedlegg 2	Intervju med ledelse, feltnotater	<u>0.20</u>	Intervju med lærer, felt- notater	<u>0.20</u>
B	Dina Eva Frøy	Gruppeintervju, transskribert	1.15		Intervju med ledelse, feltnotater	<u>0.20</u>	Intervju med lærer, felt- notater	<u>0.20</u>

Tematisk analyse (Braun & Clarke, 2006) ble gjennomført av de transkriberte intervjuene ved hjelp av NVivo11. Metoden ble valgt for å kunne kombinere innledende, teoretisk anlagte perspektiver med en analyse av hva som ellers framkommer i datamaterialet. Analysen følger denne sekstrinnsmodellen (Braun & Clarke, 2006): 1) Gjøre seg kjent med data, her ved å selv transkribere intervjuene og så lese dem flere ganger. 2) Generere innledende koder. Dette ble gjort i forlengelsen av første fase og var basert på intervjuguidens føringer. 3) Søke etter temaer. I denne fasen ble temaer som utpekte seg etter den innledende fasen kodet, og ulike tematiske kart ble utviklet fra det komplekse materialet. 4) Revidere de innledende temaene. Dette ble gjort ved å legge vekt på noen sentrale forhold, særlig ulike fasetter ved bruk av tilbakemelding, i tillegg til et kontekstuel forhold som lærersamarbeid. 5) Definere og navngi temaer. Dette ble gjort i forlengelsen av fase fire. 6) Skrive ut funnene. Basert på analysene i trinn 1–5 ble følgende kategorier vektlagt:

- Skolens fagkultur i norskseksjonen
- Kriterier i skriveopplæringen i norskfaget
- Bruk og oppfølging av tilbakemeldinger som en del av veiledning av skriving, herunder hverandrevurdering og egenvurdering

Kategoriene er utledet av teorigrunnlaget, slik det er redegjort for i litteraturgjennomgangen.

Gjennom en læringsplattform ble omtrent 75 % av alle dokumenter de seks norsklærerne publiserte til egne elever gjennom skoleåret lastet ned til forskningsformål, inkludert oppgaver, planer, kriterier, presentasjoner, lærers tilbakemeldinger skrevet inn i elevtekster, karakterer, maler, instruksjoner og lekser. Dokumentene var på ulike vis knyttet til vurderingssituasjonene, og de fungerte som bakgrunnsdata for utformingen av intervjuguide og for analysen av intervjuene. De delene av bakgrunnsdata som er undersøkt og presentert i denne studien, består av en oversikt over utvalgte

vurderingssituasjoner som lærerne organiserte som veiledete skriveprosesser. Det vil si at elevene skulle levere inn utkast, lærer skulle gi tilbakemelding, og eleven skulle jobbe videre med teksten etter å ha fått tilbakemelding. Tabellene viser antall elever per klasse, hvor mange som leverte utkast, hvor mange som bearbeidet egen tekst etter å ha fått tilbakemelding fra lærer og hvor mange som leverte endelig versjon. Det framgår også hva slags vurderingssituasjon eller oppgave det var snakk om. Det framgår dessuten i hvilken grad elevene fulgte opp arbeidet. Disse skrivesituasjonene er framhevet spesielt fordi sentral VFL-teori (Black & Wiliam, 2009; Hattie & Timperley, 2007) legger en avgjørende vekt på at informasjon fra vurdering brukes i videre læringsarbeid, og skriveprosesser er et eksempel på en situasjon der det ofte er lagt til rette for at elever bruker lærers tilbakemeldinger. Dataene gir dessuten innblikk i grad av samarbeid om vurdering i norskfaget ved at oppgaver og vurderingsskjemaer kunne sammenliknes på tvers av klasser og skoler. En oversikt over materialet er presentert i vedlegg 1.

Resultater

I det videre framstiller jeg forskningsfunnene ved å skrive om hver av de tematiske kategoriene og framstiller skolene hver for seg, slik at likheter og forskjeller enklere kan synliggjøres.

Skolens fagkultur i norskseksjonen

Ulike kulturer ved de to skolene viste seg både i gruppe- og oppfølgingsintervjuene. Lederen ved skole A fortalte i oppfølgingsintervjuet at det ikke hadde vært noe systematisk lærersamarbeid over de siste tre–fire årene (det vil si 2011–2015), men fortalte om at skolen hadde et seksjonsmøte hver mandag og at de hadde en fagkoordinator i 10 % stilling. Lederen framhevet at skolen var kompleks. Dette gjorde det vanskelig å etablere en felles vurderingspraksis. Mye av det lederen sa ble støttet av læreren som ble intervjuet, men denne læreren framhevet at seksjonsmøtet som ble avholdt en gang per måned ofte var et møte for hele seksjonen for studiespesialiserende studieretning. I liten grad egnet dette møtet seg til fagspesifikt lærersamarbeid. I gruppeintervjuet ble det vist til eksempler på at forsøk på samarbeid som hadde stoppet opp, blant annet ved at nye lærere som ble ansatt ikke alltid tok del i et etablert samarbeid. Dessuten var det uenighet innad i lærergruppa, f.eks. om hvilke fargekoder som skulle knyttes til hvilke typer språkfeil: «for når det hender at noen får ny lærer, med nye fargekoder, da ble det jo et problem». Dette illustrerer at læreren, som også var

fagkoordinator, ikke oppfattet at skolen hadde noe system for å sette nye lærere inn i hva som var gjeldende praksis, men hun viste også til at lærerne blant annet vurderte sammen på uformell basis og tidfestet sist det hendte. Dette ble framholdt som et eksempel på at samarbeid forekom.

Ved skole B var det satt av tid på timeplanen til fagmøter for norsklærerne, norsklærerne var organisert i «faglag» på trinn, og fagkoordinator hadde en 5 % stilling. Lærerne hadde felles årsplaner, felles vurderinger, en lokalt tilpasset læreplan, og det ble utarbeidet forslag til opplegg for hver periode som lærerne fulgte i ulik grad (opplysningene stammer fra en lærer ved skole B). Flere empiriske kilder viste at lærerne anvendte like vurderingskriterier. I gruppeintervjuet ved skole B ga lærerne uttrykk for en felleskultur i norskseksjonen, som blant annet innebar samretting av elevoppgaver. Frøy sa at «i norskfaget har vi jo en ganske sterk fagkultur, eller vi har flere samlinger hvor vi fellesvurderer. Vi har rett og slett en kultur for å snakke om vurdering». I et intervju med en leder og en lærer om samarbeidet i norskseksjonen ble det bekreftet at samarbeid var en viktig satsing for denne skolen. Dette gjaldt flere fag, men i norskseksjonen var samarbeidet særlig sterkt. Ingen av intervjuene ga noen indikasjoner på om lederen var til stede ved møter i norskseksjonen, men ved skole B oppfattet lærer at de fikk mye relasjonell støtte i form av ros fra ledelsen for sitt arbeid i norskseksjonen.

Vurderingskriterier

På spørsmål om hvor nyttige de fant at kriterier var for elevenes læring, var lærerne ved skole A enige om at elever hadde vanskelig for å forstå generelle vurderingskriterier. Berit mente at det å gjøre elevene kjent med kriteriene for en god tekst over tid var fruktbart, men hun siterte også en elev ved å si: «men jeg skjønner jo ingenting av det som står der». Videre oppfattet lærerne at de kunne frigjøre seg fra kriterier og bruke skjønn når de først hadde gjort seg kjent med kriteriene. De understreket at elevenes læring kom foran en pålitelig vurdering: «en gir ulike elever ulike tilbakemeldinger etter hvert som en blir kjent med dem – og tilpassa den enkelte eleven og ikke nødvendigvis det skjemaet som er tenkt til sluttvurdering» (sitat fra Carina). Anne nyanserte dette bildet ved å si at «samtidig opplever jeg at det er bruk for dem når det er tilpassa helt konkret den teksten de skal skrive (...)». Dette gjaldt spørsmålet om å nyttiggjøre seg dem i skrivearbeidet. De framhevet samtidig at kriterier hadde bidratt til å profesjonalisere vurderingspraksis ved at de ga bedre begrunnelser for karakterene og mer riktige karakterer.

Lærerne ved skole B sa at de aktivt brukte kriterier i vurderingsarbeidet og at det virket «profesjonaliserende» (sitat fra Dina). Samtidig oppfattet de at elever kunne bli forvirret av markering av høy måloppnåelse på mange områder, men at de likevel fikk en svak karakter (for eksempel fordi oppgaven manglet en viktig komponent som drøfting). Eva sa: «da sa jeg, kan du vise meg hvor i teksten din du drøfter? Og det fant han ikke». Her ser det ut til at kriteriene både understøttet karaktersetning og kunne bidra til å klargjøre mål. Frøy framhevet at det vesentlige med vurderingsskjemaer var hvordan de ble brukt. Dette synet delte hun med de øvrige ved skole B.

Alt i alt mente lærerne at vedvarende bruk av kriterier kunne bidra til å kaste lys over hva som kjennetegner en god tekst, men at mange elever, uavhengig av skole, ikke forsto dem og følgelig heller ikke kunne nyttiggjøre seg dem i læringsarbeidet. For å illustrere hva slags kriterier lærerne benyttet, har jeg vektlagt dimensjonen «struktur» som ofte finnes i en eller annen form. Ved skole B brukte lærerne de samme generelle kriteriene på flere oppgaver, utformet for vurdering av sentralgitt eksamen. Det heter blant annet «teksten har en viss struktur og sammenheng/ stort sett formålstjenlig struktur og god sammenheng / formålstjenlig struktur og god sammenheng» for ulike nivåer av måloppnåelse. Ved skole A hadde ikke lærerne felles kriterier, men det er mange eksempler på kriterier som retter seg mot elevtekstens struktur. Anne hadde f.eks. «Oppgaven har en tydelig inndeling i innledning, hoveddel og avslutning. Teksten er delt inn i avsnitt». Her kan en si at Anne har forsøkt å konkretisere dimensjonen «struktur» og at framstillingen er mer tilrettelagt for elevene. Berit brukte eksamenskriterier, og Carina hadde i noen tilfeller spesifikke kriterier for ulike oppgaver, mens andre var formulert på et overordnet nivå, som ved skole B.

Bruk av tilbakemeldinger

I gruppeintervjuet med lærerne ved skole A vektla de bruk av tilbakemeldinger i skriveprosesser som et ideal, og bakgrunnsdata viser at to av dem la aktivt til rette for denne typen arbeid. Alle lærerne kunne beskrive ulike systemer for bruk av tilbakemeldinger, men de hadde ikke en samlet praksis. Carina hadde lagt til rette for at elever kunne få tilbakemelding og jobbe videre med en skriveoppgave som heldagsprøven, men det var det få som gjorde. Hun brukte dessuten skrivemappe i en ST Vg2-klasse igjennom året for å få elevene til å bearbeide egne tekster, og 9 av 29 leverte denne ved årsslutt. Carina uttrykte et aspekt ved denne erfaringen ved å si «dette med å revidere teksten sin, det sitter litt langt inne, annet enn sånn veldig små ting (...)». Hun framhevet videre at «de arbeidsomme elevene klarer å få en del ut av det, men de andre bruker det ikke noe

særlig». I forlengelsen viser hun også til lærernes autonomi i spørsmålet om de skal jobbe med mappevurdering: «Det er ikke alle som bruker mappe, og vi har diskutert hvor mye får vi igjen for det, ikke sant ... Og elevene: hvor mye får de igjen for det?» Det kom også fram at de erfarte vansker med å få elevene til å skrive i hele tatt, og at å jobbe med tilbakemeldingene fra lærer følgelig også var vanskelig: I tilfeller der det ble satt av tid i timen til å jobbe med tilbakemeldinger på skriftlige oppgaver, opplevde Carina situasjonen slik: «så er det jo veldig mange som ikke blir ferdig, mange som kaster bort veldig mye tid (...), så de gjør egentlig ikke så mye av det».

Det kan se ut som om det var litt tilfeldig om elevene fikk innføring i hvordan de skulle bruke lærers tilbakemeldinger. Berit sa at «noen ser på karakteren og kaster den med en gang. Andre ser. Selv om det er fargekoder på Fronter og sånn, var det plutselig noen som – etter jul – som var overraska over at de lå på Fronter». Anne fortalte at «jeg har liksom ikke noe system. Jeg har gjort litt forskjellig fra gang til gang, så noen ganger så har jeg nesten bare ‘nå har dere fått den tilbake, vi bruker ti minutter så dere kan gå, se på kommentaren, og spør meg hvis det’ ... at de ikke har levert noe på nytt egentlig. (...) Jeg føler også på det at det er begrensa hvor mye tid man kan sette av til det i timene, fordi det potensielt er så mye misbrukt tid, og da spesielt for de svakeste».

Eksemplene viser at lærerne hadde en bevissthet om at de burde legge til rette for bruk av tilbakemeldinger på skriftlige arbeider, og at problemet syntes å være å få det til i praksis, både fordi de ikke ga elevene systematisk opplæring, men også fordi elevene var motvillige til å arbeide videre med teksten sin etter at de hadde fått tilbakemelding fra lærer. At Anne satte av tid i timen viser samtidig at hun, etter eget utsagn, og i likhet med Carina, omsatte anbefalt vurderingspraksis i handling.

Ved skole B beskrev lærerne at det å gi tilbakemelding på skriftlige tekster var en del av veiledningen av elevenes skriving, ved at det også skjedde i klasserommet, og at det blant annet kom til uttrykk i hvordan lærerne gjennomførte heldagsprøver i faget. Dina kan tjene som en typisk representant når hun beskrev at hun hadde forsøkt to praksiser: «I år hadde jeg to heldagsprøver hvor den ene var veldig veiledende, på den andre var jeg bare sånn at jeg gikk bort til de som rakk opp hånden. Og det var radikal forskjell på prestasjonen.» Lærerne var samstemte i oppfatningen av at det å gi tilbakemelding i skriveprosessen var mest verdifullt med tanke på læring. Bakgrunnsdata viser at Eva hadde et system for ny innlevering av tekster som hadde fått tilbakemelding, og det framgår at omtrent en tredjedel av elevene leverte en ny versjon. Også Frøy beskrev systemer for arbeid med tilbakemeldinger og begrunnet det med å si at «de må gjøre noe med teksten for å

skjønne hva tilbakemeldingen betyr». Hun beskrev også et system hun hadde brukt der elevene skulle skrive hva som var viktig i den tilbakemeldingen de fikk av henne, og at hun ikke ga elevene karakter før de hadde lest tilbakemeldingen. Ingen av lærerne ved skole B nevnte manglende innsats fra elevenes side i dette arbeidet, selv om bakgrunnsdata viser at det også ved denne skolen kunne være vanskelig å få elevene til å følge opp arbeid med tilbakemeldinger på skriftlige tekster. Det framgår altså at lærerne ved skole B hadde en mer systematisk praksis når det gjaldt bruk av tilbakemeldinger til veiledning av skriving. Samtidig var det tydelige individuelle variasjoner.

Ingen av lærerne ved skole A fortalte om noen systematisk bruk av hverandrevurdering. De hadde forsøkt ulike måter, men uten hell. Berit hadde prøvd at elevene skulle gi hverandre karakter, Anne fortalte om et undervisningsopplegg som gikk ut på å karaktersette anonymiserte tekster («gjett hva slags karakter de fikk»), og lærerne ved skole A mente at elevene «ikke orker, begynner å snakke om andre ting» (sitat fra Carina). Ved skole B beskrev lærerne ulike erfaringer med hverandrevurdering, men framholdt at de sjelden brukte denne arbeidsformen og de hadde ikke noen systematisk praksis.

Diskusjon

De dimensjonene som er framhevet i litteraturgjennomgangen og resultatdelen danner grunnlaget for den videre diskusjonen av forskningsspørsmålet: Hva kjennetegner norsklæreres vurderingsarbeid ved to samarbeidskulturer?

Et premiss for diskusjonen er at skolekulturene i norskseksjonen var ulike, og funnene viser i noen grad hvordan: Ved skole B var det en mer tydelig samarbeidskultur som kom til uttrykk ved organisatorisk tilrettelegging for samarbeid og ved at lærerne oppfattet at de hadde en felles vurderingskultur. En slik kultur kan ha både fordeler og ulemper, om en legger Vangrieken et al. (2015) til grunn: Fordeler kan være arbeidslette i form av felles planlegging, utarbeidelse av felles kriterier og fordeling av arbeidet med å lage oppgaver. En kultur der lærerne snakker om vurdering kan også bidra til å gjøre mål for opplæringen mer tydelige fordi de må gjøres eksplisitte, noe Sadler (2010) framhever som en viktig gevinst ved hverandrevurdering.

På den andre siden er det den individuelle lærer som gir tilbakemeldinger til egne elever og som har mulighet til å tilpasse det hun skriver til denne elevens behov. Dette kan vanskelig være et kollektivt anliggende, særlig siden det kun er norsklæreren som kjenner sine elever i eget fag.

Samarbeidet har altså en grense, og lærernes autonomi må nødvendigvis spille en sentral rolle om de skal kunne tilpasse undervisning og tilbakemeldinger til egne elevers behov. Vangrieken et al. (2015) framhever faren for opplevelse av autonomitap som en mulig ulempe ved tett samarbeid, sammen med konformitetspress og fare for intern konkurranse. At samarbeid kan være krevende for lærerne, kan illustreres ved at det ved begge skoler ble gitt uttrykk for at en tidligere praksis med tett samarbeid delvis var forlatt. Samarbeidet innbefattet utarbeidelse av oppgaver og opplegg ved skole B og et felles fargekodesystem for retting av språkfeil og skriveopplæring med mappevurdering ved skole A. Forskningsfunn understøtter at lærersamarbeid kan ha mange fordeler (Goddard et al., 2007; Vangrieken et al., 2015), men lærere kan trekke seg unna samarbeid som blir for tett. Det kan derfor være fornuftig å finne en balanse mellom autonomi og samarbeidskrav (Hargreaves, 1994). Konsekvenser av dette forholdet blir diskutert videre i to spesifikke kontekster, som er utledet av VFL-teori, nemlig hvordan lærerne jobber med vurderingskriterier og bruk av tilbakemeldinger i læringsarbeid.

Vurderingskriterier

Lærerne ved de to skolene hadde ulik erfaring med bruk av vurderingskriterier. Ved skole A ble bruk av kriterier framstilt som profesjonaliserende på den måten at lærerne følte seg tryggere på egen karaktersetning, men de sa samtidig at elevene hadde vanskelig for å forstå kriteriene. At elever trenger hjelp for å forstå kriterier, er et poeng i Sadlers (2005) forskning, og i min studie kan det se ut som lærerne oppfattet at elevene hadde betydelige problemer med å forstå vurderingskriterier. Legger en Sadler (2005) til grunn, kan det tenkes at det skyldes manglende opplæring. At elevene i to så ulike kontekster som skole A og B hadde det samme problemet, kan samtidig indikere at skjemaer med kjennetegn på måloppnåelse i liten grad egner seg til å klargjøre hva som er målet for opplæringen for elevene. Målet må heller gjøres kjent gjennom opplæringen, mener lærerne. Dette funnet kan sies å samsvare med Dobbie og Fryer (2011) som fant at arbeid med å utvikle lokale læreplanmål ikke var statistisk forbundet med økt læringsutbytte hos elever. At lærerne dessuten erfarte at kriterier kunne villede elever med tanke på grunnlaget for en karakter, illustrerer at det er en lang vei fra mer overordnede mål nedfelt i instrumenter (som vurderingsskjema for norskeksamen er et eksempel på) til å styrke elevens forståelse av hva som er målet med opplæringen. Lærerne var samtidig enige om at det er langt bedre å ha kriterier enn å ikke ha det, og at «magefølelse» som grunnlag for vurdering hadde veket plassen for en mer profesjonell

og pålitelig summativ praksis. Dette kan ses som et positivt aspekt av det Ingersoll (2011) og Evetts (2013) peker på som en internasjonal trend, nemlig at lærernes arbeid underlegges en viss standardisering.

Hvis felles kriterier over tid er en del av en skoles fagkultur i norskfaget (slik det syntes å være tilfelle ved skole B) vil det kunne være lettere elevene å forholde seg til dem i de ulike skrivesituasjonene de står i. Elever og klasser blir dermed mindre sårbare når lærere byttes ut/slutter underveis i året eller når elever skifter klasse. På den andre siden kan vurderingskriterier lett framstå som for abstrakte ideer, slik lærerne ved skole A gir uttrykk for. Her kunne den autonome læreren legge kriteriene til rette for sin klasse og slik øke mulighetene for læring. Sett under ett kan det se ut som om lærerne fant at vurderingskriterier mer enn noe annet egnert seg til å begrunne karakterer, mens forbindelsen til elevenes læring gjennom klargjøring av mål var usikker eller direkte svak – uavhengig av samarbeidskulturen i norskfaget.

Bruk av tilbakemeldinger

Lærerne ved begge skoler var helt klare på at elevene måtte bruke tilbakemeldingene de fikk på skriftlige tekster dersom de skulle lære av dem. Dette er i tråd med retningslinjene fra Udir (2010) og med forskning (Black & Wiliam, 2009; MacArthur, 2013). Lærerne beskrev også at de hadde ulike teknikker for å få dette til. Likevel var det vanskelig å få elevene til å jobbe med tekstrevisjon. Ved skole A knyttet lærerne dette forholdet til elevforutsetninger, det vil si at elevene var så svake skrivere at det å i hele tatt produsere tekst var en stor utfordring. Om en ser på elevopplutning om skriveprosesser ved skolen, støttes denne slutningen: en lærer opplevde at kun 9 av 29 elever leverte mappe ved årsslutt og i en heldagsprøvesituasjon leverte bare 1 av 29 en forbedret tekst etter at lærer hadde gitt tilbakemelding i skriveprosessen.

Samarbeid og felles skolekultur kan se ut til å utgjøre en forskjell med hensyn til elevenes opplutning om arbeid med å forbedre tekster etter å ha fått veiledning fra lærer siden skole B fikk dette til i høyest grad. Dette funnet kan tas til inntekt for at en skole med tydelig samarbeidskultur har større muligheter til å lykkes med å få til en vurderingspraksis som forskning og retningslinjer anbefaler og at den kan ses som et eksempel på en vellykket samarbeidskultur (Vangrieken et al., 2015). Samtidig var det en helt klar tendens ved begge skoler at elevene valgte bort mulighetene til å bedre egen tekst, enten dette var noe læreren krevde av dem eller ikke, og at lærernes forsøk på

å legge til rette for en vurderingspraksis som bidrar til læring altså møtte betydelig motstand hos elevene. Hvordan kan dette funnet forklares?

Et trekk som synes å være felles for disse undervisningssekvensene var at elevene skulle jobbe med tilbakemeldingene på egenhånd på grunnlag av lærers skriftlige tilbakemelding etter at skriveøkten var avsluttet. Rent praktisk framstår dette som en fornuftig ordning, etter som tiden læreren har til hver elev i løpet av en skoletime er kort. Samtidig viser f.eks. Hattie og Timperley (2007) at tilbakemeldinger gitt i en prosess har størst sjanse for å bidra til læring. Dina ved skole B omtalte en vellykket skriveøkt der elevene skrev på skolen og fikk hjelp underveis: alt foregikk innenfor rammene av en dag. Ved begge skolene hadde disse skrivesituasjonene en elevopplutning på nærmere 100 %. For at en slik praksis skal være mulig, må skolen og fagseksjonen ha et samarbeid som gjør at norsklæreren har denne tiden til rådighet. Hvis læreren har timer i andre fag mens elevene skriver, sier det seg selv at det er vanskelig å veilede i en prosess. Denne erfaringen kan også tas til inntekt for at en skole med organisatorisk tilrettelagt samarbeid i norskseksjonen har større muligheter for å realisere en VFL-orientert praksis når det gjelder bruk av tilbakemeldinger i videre læringsarbeid.

Forskningen på egen- og hverandrevurdering (Sadler, 2010; Topping, 2009) viser til positive konsekvenser av slike ordninger i form av økt revisjonskompetanse (Marshall, 2012). Imidlertid hadde ingen av lærerne i denne studien dette som noen fast praksis. Det kan derfor hevdes at denne studien har funnet et urealisert potensial for bruk av tilbakemeldinger i prosess.

Studiens begrensninger

Selv om informasjonen som kommer fram i intervjuene er blitt vurdert sammen med bakgrunnsdataene, er intervjuet en form for selvrapporing. Intervjuet gir derfor først og fremst innsikt i lærernes erfaringer og refleksjoner, slik de kommer til uttrykk i møtet med kollegaer og forsker (Säljö, 1997). En kan ikke vite om dette er faktiske holdninger som manifesterer seg i handlinger. På den annen side studerte jeg også lærernes handlinger gjennom dokumenter og bakgrunnsdata. Selv om selvrapporing via intervju ga begrenset innsikt i hva en lærer gjør i eget klasserom og overfor egne elever, indikerer samsvaret mellom bakgrunnsdata og intervjuer en viss koherens.

Konklusjoner og implikasjoner for praksis og videre forskning

Studien undersøker forhold ved vurderingsarbeidet i norskfaget ved to skolekulturer, og den fokuserer særlig på vurderingskriterier og på bruk av informasjon fra lærers tilbakemeldinger i videre læringsarbeid fordi dette er sentrale dimensjoner ved VfL-orientert praksis. Viktige funn er at lærerne oppfattet det som at vurderingskriterier først og fremst virket som karakterbegrunnende instrumenter, uavhengig av skolekultur, og at faglig svakere elever rett og slett hadde problemer med å forstå kriterier. I spørsmålet om bruk av lærers tilbakemeldinger i videre arbeid med tekstrevisjon viste empirien at lærerne ved begge skoler kunne finne dette krevende. Dette gjaldt i særlig grad skolen med svakest elevgrunnlag. En mulig slutning er at organisatorisk tilrettelegging for samarbeid vil kunne øke mulighetene for å legge til rette for at elever får tilbakemelding i skriveprosesser ved at heldagsprøver/fagdager tilrettelegges som veiledete skriveøkter.

En viktig implikasjon for praksis er at denne studien kan bidra til å rette oppmerksomhet mot hvor viktig det er å finne konkrete, skolebaserte måter å jobbe med bruk av tilbakemeldinger, samt å utvikle læringsmål som kommuniserer godt med elevene (fordi generelle kriterier i liten grad ser ut til å gjøre det). Å utvikle mer kunnskap om hvordan man kan arbeide med kriterier på en mer læringsfremmende måte overfor elever, framstår som en viktig oppgave for framtidig forskning.

Referanser

- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), s. 5–25. <http://dx.doi.org/10.1080/0969594X.2010.513678>
- Berge, K. L. & Skar, G. (2015). *Ble elevene bedre skrivere? Intervensjonseffekter på elevers skriveferdigheter og skriveutvikling*. Trondheim: Høgskolen i Sør-Trøndelag.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), s. 5–31. <http://doi:10.1007/s11092-008-9068-5>
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), s. 77–101. <http://dx.doi.org/10.1191/1478088706qp063oa>
- Bueie, A. A. (2015). Summativ vurdering i formativ drakt – elevperspektiv på tilbakemelding fra heldagsprøver i norsk. *Acta Didactica Norge*, 9(1), Art-4.
- Bueie, A. A. (2016). Nyttige og mindre nyttige lærerkommentarer – slik elevene ser det. *Nordic Journal of Literacy Research*, 2(1). <http://dx.doi.org/10.17585/njlr.v2.188>
- Burner, T. (2014). *Formative Assessment of Writing in English: A school-based study of perceptions, practices and transformations*. Doktorgradsavhandling. Trondheim: NTNU.
- Dobbie, W., & Fryer Jr, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4), s. 28–60. <https://doi.org/10.1257/app.5.4.28>
- Eggen, A. B. (2009). Kvalitativt FoU-arbeid med vekt på vurdering – et møtepunkt mellom ulike praksiser. I: Hovdenak, S. S., Eggen, A. B. & Elstad, E. (red.). *Kunnskapsløftet – fra ord til handling. Rapport fra et skoleutviklingsprosjekt for Utdanningsetaten i Oslo*, s. 48–160. Oslo: Universitetet i Oslo
- Forfatter (2017a) – informasjon utelatt av hensyn til anonymisering
- Forfatter (2017b) – informasjon utelatt av hensyn til anonymisering
- Evetts, J. (2013). Professionalism: Value and ideology. *Current Sociology*, 61(5-6), s. 778–796. <https://doi.org/10.1177/0011392113479316>
- Goddard, Y. L., Goddard, R. D., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), s. 877–896.
- Graham, S., Harris, K. R. & Herbert, M. (2011). *Informing Writing: The benefits of formative assessment*. Washington DC: Alliance for Excellent Education.
- Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age*. New York: Teachers College Press.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), s. 81–112. <https://doi.org/10.3102/003465430298487>
- Hertzberg, F. & Roe, A. (2016). Writing in the content areas: a Norwegian case study. *Reading and Writing*, 29(3), s. 555–576. <http://doi:10.1007/s11145-015-9607-7>
- Hopfenbeck, T., Tolo, A., Florez, T., & El Masri, Y. (2013). Balancing trust and accountability? The assessment for learning programme in Norway: A governing complex education systems case study. Oslo: OECD.
- Ingersoll, R. (2011). Power, Accountability, and the Teacher Quality Problem. Nedlastet fra http://repository.upenn.edu/gse_pubs/236

- Kunnskapsdepartementet (2008). *Læreren. Rollen og utdanningen. St.meld. nr. 11 (2008–2009)* <https://www.regjeringen.no/no/dokumenter/stmeld-nr-11-2008-2009-/id544920/>
- Kvale, S. & Brinkmann, S. (2010). *Det kvalitative forskningsintervjuet*. Oslo: Gyldendal Akademisk Norsk Forlag.
- Leithwood, K. A. & Riehl, C. (2003). *What we know about successful school leadership*. Nottingham: National College for School Leadership.
- Lijphart, A. (1971). Comparative Politics and the Comparative Method. *American Political Science Review*, 65(3), s. 682–693.
- Lortie, D. C. (1975). *School teacher: A sociological inquiry*. Chicago: University of Chicago Press.
- MacArthur, C. (2012). Evaluation and revision processes in writing. I: Berninger, V. W. (red). *Past, present, and future contributions of cognitive writing research to cognitive psychology*, s. 248–262. New York: The Guilford Press.
- MacArthur, C. (2013). Best Practices in Teaching Evaluation and Revision. I: Graham, S., MacArthur, C. A. & Fitzgerald, J. (red.). *Best Practices in Writing Instruction*, s. 215–237. New York: The Guildford Press.
- Matre, S. & Solheim, R. (2014). Lærersamtalar om elevtekstar – mot eit felles fagspråk om skrivning og vurdering. I Hvistendahl, R. & Roe, A. (red.). *Alle tiders norskdidaktiker*, s. 219–244. Oslo: Novus forlag.
- Moolenaar, N. M., Slegers, P. J. & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education*, 28(2), s. 251–262. <http://dx.doi.org/10.1016/j.tate.2011.10.001>
- Sachs, J. (2016). Teacher professionalism: why are we still talking about it? *Teachers and Teaching*, 22(4), s. 413–425. <http://dx.doi.org/10.1080/13540602.2015.1082732>
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), s. 175–194. <http://dx.doi.org/10.1080/0260293042000264262>
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), s. 535–550.
- Saliu-Abdulah, D., Hellekjær, G. O. & Hertzberg, F. (2017). Teachers' (formative) feedback practices in EFL writing classes in Norway. *Journal of Response to Writing*, 3(1), s. 31–55.
- Sandvik, L. V. & Buland, T. H. (2014). *Vurdering i skolen. Utvikling av kompetanse og fellesskap. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)*. Trondheim: NTNU.
- Säljö, R. (1997). Talk as data and practice – a critical look at phenomenographic inquiry and the appeal to experience. *Higher Education Research & Development*, 16(2), s. 173–190.
- Svanes, I. K. & Skagen, K. (2016). Connecting feedback, classroom research and Didaktik perspectives. *Journal of Curriculum Studies*, 49(3), s. 334–351. <http://dx.doi.org/10.1080/00220272.2016.1140810>
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), s. 20–27. <http://dx.doi.org/10.1080/00405840802577569>
- Torrance, H. & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), s. 615–631. <http://dx.doi.org/10.1080/01411920120095780>
- Utdanningsdirektoratet (2014). *Nasjonal satsing på vurdering for læring (2010–2014)*. Oslo: Utdanningsdirektoratet <http://www.udir.no/Vurdering-for-laring/Nasjonal-satsing1/Nasjonal-satsing-pa-Vurdering-for-laring>
- Utdanningsdirektoratet (2010). *Rundskriv Udir-1-2010. Individuell vurdering*. Oslo: Utdanningsdirektoratet

<http://www.udir.no/Regelverk/Finn-regelverk-for-opplaring/Finn-regelverk-etter-tema/Vurdering/Udir-1-2010-Individuell-vurdering>

Utdanningsdirektoratet (2009). *Bedre vurderingspraksis*. Sluttrapport fra Utdanningsdirektoratet 2009. Hentet fra <http://www.udir.no/Tilstand/Forskning/Rapporter/Utdanningsdirektoratet/Bedre-vurderingspraksis-sluttrapport-fra-Utdanningsdirektoratet-2009/>

Vangrieken, K., Dochy, F., Raes, E. & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, s. 17–40. <http://dx.doi.org/10.1016/j.edurev.2015.04.002>

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), s. 3–14. <http://dx.doi.org/10.1016/j.stueduc.2011.03.001>

Vedlegg 1

Intervjuguide, tilbakemeldinger i skriftlig norsk vgs

Demografisk informasjon

1. Lærers utdanning og bakgrunn
_____ BA-grad
_____ MA-grad
Hvis ja: I hvilket fag? _____
_____ Annet, spesifiser _____
_____ Støttefag: _____
2. Undervisningserfaring
_____ 1-3 år
_____ 4-6 år
_____ 7 år eller mer
3. Erfaring som norsklærer (år/måneder): _____
4. Skriver du selv? (Lærebøker, anmeldelser, essays, artikler?)
a. Ja _____ b. nei _____ c. I blant _____
5. Har du vært sensor på hovedmålseksamen?
6. Har du vært sensor på lokalgitt eksamen?
- 7.

A Om skriftlige vurderingssituasjoner (generelt)

- Kan dere beskrive en spesielt god eller vellykket skriftlig vurderingssituasjon i norsk?
- Hva gjorde den vellykket?
- Kan dere fortelle om en vurderingssituasjon dere synes ble mislykket?
- Hvilke hensikter har vurderingssituasjonene i din undervisning?
- Når gir du tilbakemelding? I hvilken grad gir du underveis? Hvor lang tid etter levering?
- Hender det at du bare gir karakter uten tilbakemelding? Hvorfor i så fall?
- Hender du bare gir tilbakemelding og ikke karakter?
- Hva mener du er mest viktig å legge vekt på når du skriver tilbakemelding?
- Følger du opp skriftlige tilbakemeldinger med muntlige?
- En del av dere bruker kriterier/kriterieskjema når dere gir tilbakemeldinger.
- Hvor nyttige synes dere kriterier er for elevenes læring?
- Hvor nyttige synes dere kriterier er for å begrunne karakterene?

B Om bruk av tilbakemeldinger

- Har du et system for elevenes videre arbeid med tekster, etter at de har fått tilbakemelding?
I hvilken grad/hvordan følger du opp dette systemet?
- Hvilket inntrykk har du av elevenes bruk av tilbakemeldinger?

C kameratvurdering og egenvurdering (bruk av tilbakemeldinger)

- Bruker du kameratvurdering? I så fall hvordan?
- Bruker du egenvurdering (elevene vurderer seg selv)? I så fall hvordan?
- Hvordan oppfatter dere at elevene liker/nyttiggjør seg kamerat- og egenvurdering?
- Hvilken effekt mener du at kamerat- og egenvurdering har?

D Avsluttende spørsmål

- I hvilken grad er det forskjell på det å gi tilbakemeldinger i norsk og i andre fag du underviser i?
Hva går forskjellene eventuelt ut på?
- Kan du reflektere over i hvilken grad elevene setter pris på tilbakemeldinger – alt i alt?
- Hvor viktig er det å være støttende og personlig når du gir tilbakemelding?
- Har du noe inntrykk av hvor nyttige elevene opplever at tilbakemeldingene dine er?
- Er du fornøyd med elevenes innsats etter at de har fått tilbakemelding?
- Hvor meningsfullt oppfatter du at det er å skrive tilbakemeldinger til elevene?
- Hvor mye arbeid legger du i å skrive tilbakemeldinger?
- Har du et ideal for hvordan du vil jobbe med tilbakemeldinger?
- Er det noe dere vil legge til, avslutningsvis?

Vedlegg 2

Oversikt over bakgrunnsdata, skole A

Tabell 2 viser elevoppslutning om utvalgte skriveoppgaver som har et element av prosesskriving ved skole A. IFO betyr individuell fordypningsoppgave, en stor oppgave i norskfaget på ST Vg3. Vurdering angir hvilken type skriveoppgave elevene hadde fått tilbakemelding på.

Tabell 2, oversikt over skriveprosesser ved skole A

Lærer	Klasse	Antall elever i klassen	Vurdering	Antall levert førsteutkast	Andreutkast	Endelig versjon
Anne	ST Vg3	32	IFO	17	4	10
Anne	ST Vg1	29	Kåseri	6		25
Berit	ST Vg3	32	IFO	22 (problemstilling)		23
Carina	ST Vg2	29	Mappe			9

Fellestrekk ved skole A

Skolen hadde felles heldagsprøver på ST Vg3, de to siste før eksamen. Begge klassene på ST Vg3 skrev individuell fordypningsoppgave, oppgavene var ikke identisk utformet.

Anne

I Annes kåseriprosess unnlot 4 elever å levere tekst, og 5 ble tatt for plagiat. 20 av 29 leverte som de skulle. Det var nesten 100 % oppslutning om heldagsprøver, men klart lavere på prosesskrivingene.

Berit

Berit hadde lite prosesskriving på ST. Den individuelle fordypningsoppgaven var organisert med en skriveoppgave på skolen med oppgaver til bok 1 – denne ble karactersatt. Her leverte 29 av 32.

På YF hadde hun prosesskriving i en klasse med 14 elever. Her leverte 6 førsteutkast og 2 endelig versjon.

Carina

Carina hadde heldagsprøve med førsteutkast og andre utkast på slutten av Vg2 – bare én leverte to ganger. Hun hadde i tillegg mappevurdering, som innebar at elevene skulle samle 4 tekster i en mappe, tekster som skulle velges ut fra det de hadde skrevet gjennom året; de hadde fått tilbakemelding og karakter av lærer og skulle rette opp/videreutvikle og levere på nytt. Av tabell 2 ser man at det kun var 9 som leverte.

På YF hadde hun et system for at elevene leverte inn rettet versjon av tekster de hadde fått karakter og tilbakemelding på – cirka 60 % gjorde dette.

Oversikt over bakgrunnsdata, skole B

Tabell 3 viser elevoppslutning om utvalgte skriveoppgaver som har et element av prosesskriving ved skole B. IFO betyr individuell fordypningsoppgave. Vurdering angir hvilken type skriveoppgave elevene skrev og fikk tilbakemelding på.

Tabell 3, oversikt over skriveprosesser ved skole B

Lærer	Klasse	Elever	Vurdering	Antall levert utkast	Antall endelig	Antall rettet
Dina	ST Vg3	31	IFO	16	27	
Eva	ST Vg3	27	IFO		20	
Eva	ST Vg1	31	Skrivedag på skolen		28	11
Eva	ST Vg1	31	Novelleanalyse	29	18	
Frøy	ST Vg3	28	IFO	Ulikt, se under	27	
Frøy	ST Vg1	30	Skrivedag på skolen	22	28	

Fellestrekk ved skole B

Skolen hadde felles heldagsprøver, ST Vg3 (3 klasser), og på ST Vg1 (to klasser), felles struktur på skoleåret for ST Vg3 med tanke på vurderingers innhold og rekkefølge. Individuell fordypningsoppgave hadde veldig likt utformet oppgaveformulering, det vitner om samarbeid.

Dina

Hadde organisert IFO som en skriveprosess, tallene viser til innleveringsmapper i læringsplattformen: 1 problemstilling – halve klassen leverte, lærer kommenterte, 2 handlingsreferat, 3 analyse – lærer kommenterte ikke noen av disse, det kan se ut som innleveringen hadde en kontrollfunksjon, 4 endelig versjon med karakter og kommentarer.

Dina hadde ikke klare systemer for levering av ny versjon av oppgaver. Men i likhet med Frøy hadde hun retorisk analyse som en tretrinns innleveringsoppgave, som var supplert med en muntlig vurdering til slutt.

Eva

Hadde system for andregangsl levering av tekst. Elevene skrev på skolen (28 leverte), de fikk tilbakemelding og skulle levere på nytt, og 11 av 28 gjorde det. Elevene fikk ikke karakter på første versjon, men nivå ble gullet ut – samme som hos Frøy.

IFO var ikke organisert som en direkte skriveprosess, men det var krav om delrapportering gjennom prosessen, som viste seg som innleveringsmapper. Her var elevenes oppslutning varierende. 20 av 27 leverte endelig versjon i innleveringsmappa.

Novelleanalyse viser samme mønster: 29 leverte førsteutkast, 18 leverte rettet versjon.

Evas og Frøys klasser på ST Vg1 var likt organisert med tanke på vurderingers utforming og tidspunkt.

Frøy

Hun hadde to tilfeller av prosesskriving: Den ene av retorisk analyse med utkast og endelig versjon. 17 elever av 28 leverte to versjoner, 7 bare en sluttversjon, 4 leverte ikke i hele tatt.

IFO var organisert som en spesielt omfattende skriveprosess. Elevene fikk tilbakemelding på alle deloppgaver, som ble levert inn gjennom høsten (tallene etter = angir antall levert) 1. versjon = 20, 2. versjon = 16, 3. versjon = 18, 4. versjon = 15, endelig versjon = 27.

Frøy la i utstrakt grad til rette for bruk av informasjon fra tilbakemelding for å forbedre tekst. På tradisjonelle heldagsprøver ga hun teksten tilbake uten karakter – denne ble holdt tilbake til eleven hadde rettet etter lærers tilbakemelding.

På ST Vg1 var 3 vurderinger skrevet i prosess.

Hun ga ikke karakter direkte, men limte inn skjema med kjennetegn på måloppnåelse i elevenes besvarelse, og gullet ut hvordan eleven lå an på ulike delmål, så de kunne få et bilde av karakteren eller i hvert fall nivået på den måten.

Errataliste for avhandlingen *Læringsfremmende vurderingspraksis – med blikk på norskfaget i to skoler.*

Del 2 side 116 øverst på siden: korrekt skal være «Artikkel 2».

Oslo, 19.04.2018