

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

Harald Eriksen*,

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

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 teorigrevne 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*

Received: January, 2017; Accepted: September, 2017; Published: November, 2017

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,

*Correspondence to: Harald Eriksen, Postboks 1099, Blindern, 0317 Oslo. Email: harald.eriksen@ils.uio.no

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 'Vurdering for læring' 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 igjennom 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 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 eleven 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 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 feste 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 grunnlag til å formulere følgende hypotese: Elevers oppfatning av lærers klargjøring av kriterier i vurderingsarbeidet er positivt statistisk assosiert med elevenes opplevde nytte (hypotese 1, heretter 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 morsmåls-læreres evne til å gi formative tilbakemeldinger på et skriftlig arbeid, med elevenes resultater fra en nasjonal skriveprøve i førstespråket. 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, og 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 fremkommer 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 muligheter for å levere tekster på nytt for elevene 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 elevers 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 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). 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 (Havnæs 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 sjans for at elevene nyttiggjør seg tilbakemeldingene om læreren legger til rette. 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 fra 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. Derfor er det 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 etter direkte forespørsel fra forsker. Skolene har ulik profil med hensyn til inntakssnitt, elevsammensetning, studieretninger, antatt sosioøkonomisk status for nærområdet og etnisk sammensetning av elevgruppa. 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 bortfallskjevhet 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 samsvarte 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 om 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 «enkeltitems» 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 altså 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 = gjennomsnittskåre, SD = standardavvik. Cronbachs alfa måler den interne konsistensen mellom indikatorene som ligger til grunn for den latente variabelen.

Tabell 2 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

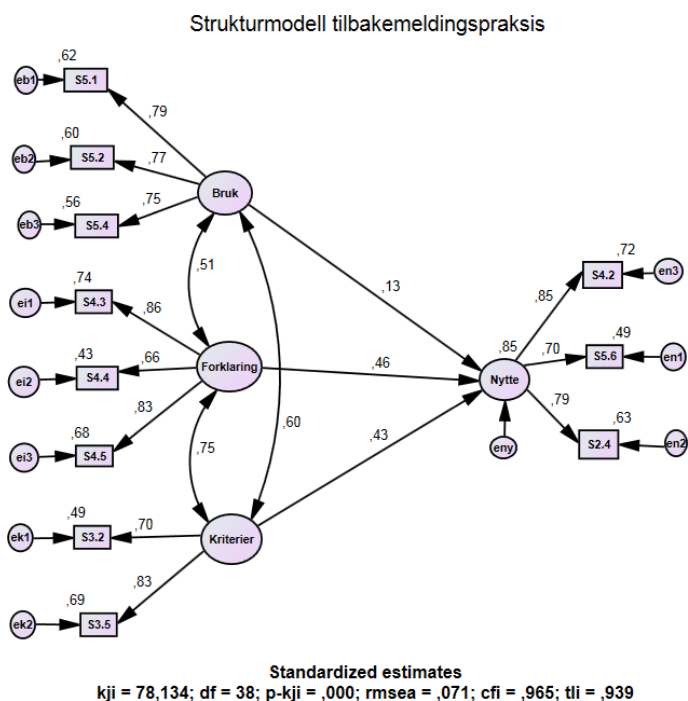
opplevd nytte og ligger på 1,8–1,9. Dette kan tolkes som et uttrykk for at elevene oppfatter at lærerne i ulik grad legger til rette for bruk, det vil si at noen gjør det i større grad enn andre og at færre har svart det nøytrale alternativet «4».

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 på mellom 4,0 og 4,6, noe som indikerer at elevene dels har en nøytral, dels 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-verdier 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



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

Tabell 3. Resultater av hypotesetesting

Hypotese	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 at lærer bruker kriterier, og at elever opplever tilbakemeldinger som nyttige. Det er likevel usikkerhet knyttet til funnet, og dette blir diskutert videre under.

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 elevenes opplevelse av at læreren gir konkrete råd og elevenes 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).

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

Tabell 4. Konfirmerende faktoranalyse: Relevante fit-opplysninger

Faktorer	Antall Ind.	Faktorladninger		Kji-kvadrat	Frihets-grader	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

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 ($K_j = 47,873$, $df = 29$, $p\text{-}k_j = 0,015$, $RMSEA = 0,055$, $CFI = 0,981$, $TLI = 0,964$). 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 tilbakemeldingspraksis til sine seks lærere mer enn at de speiler en større populasjon. For å undersøke om dette er tilfelle, ble det regnet ut intraklassekorrelasjonskoeffisient-skårer (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. Intraklassekorrelasjonskoeffisient-skårer, 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 undersøker tre hypoteser, og i det følgende vil resultater, valget av variabler og mulige forklaringer på funnene bli diskutert.

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 det er 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 skrivetekniske 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 (Kvithylid & 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 på 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 det 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 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 Kriterier, rokker likevel ikke dette ved slutningen om at bruk av vurderingskriterier er gunstig. Dette resonnementet betinger at ytterligere validering av instrumentet lykkes og må følgeses som tentativt.

I en studie av Parr og Timperley (2010) kobles morsmåslæreres evne til å gi gode tilbakemeldinger, definert innenfor et VFL-rammeverk, med elevers 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 dypereliggende 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 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 læreres 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 være andre forhold som også ligger under, f.eks. 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 tilbakemelding, har de ofte begynt å jobbe med noe nytt, med nye skriveoppgaver og nytt fagstoff, altså at skolekonteksten er sånn at elevene har en jevn strøm av oppgaver de skal gjøre, også vurderinger som karaktersettes i ulike fag, og at det dermed er 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). Dermed 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 som 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 skrivingen 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 er 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 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 noen 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, men det er vanskelig å avklare dette uten eksperimentelle tilnærminger. Kontrollerte eksperimenter har vært utenfor rekkevidde for å iverksette denne studien, men det ville ideelt sett være 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 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). Men det må her 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 vil være 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 det forsøkt å løse problemet 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 eksempel 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 bygger ikke på internasjonalt validerte instrumenter. Grunnen er at det ikke har vært mulig å finne slike, og undersøkelsen er derfor utviklet for denne studien spesielt. 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, 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, og det 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 elever 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 være 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 denne praksisen har et betydelig potensial for å informere 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 større utvalg, og dessuten elevers 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 eleven følger opp på en dedikert måte. Her ser det ut til å være et betydelig forbedringspotensial. Studier av lærerens 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 B. Lawson, S. Ryan & W.R. Winterrowd (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
- 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
- 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. (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. (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
- 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: Akademia 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
- 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
- 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
- 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
- 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
- Hopenbeck, 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
- 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
- Jönsson, A. (2013). Facilitating productive use of feedback in higher education. *Active learning in higher education*, 14(1), 63–76. 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: gymnasisisters berättelser om lärares 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

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.

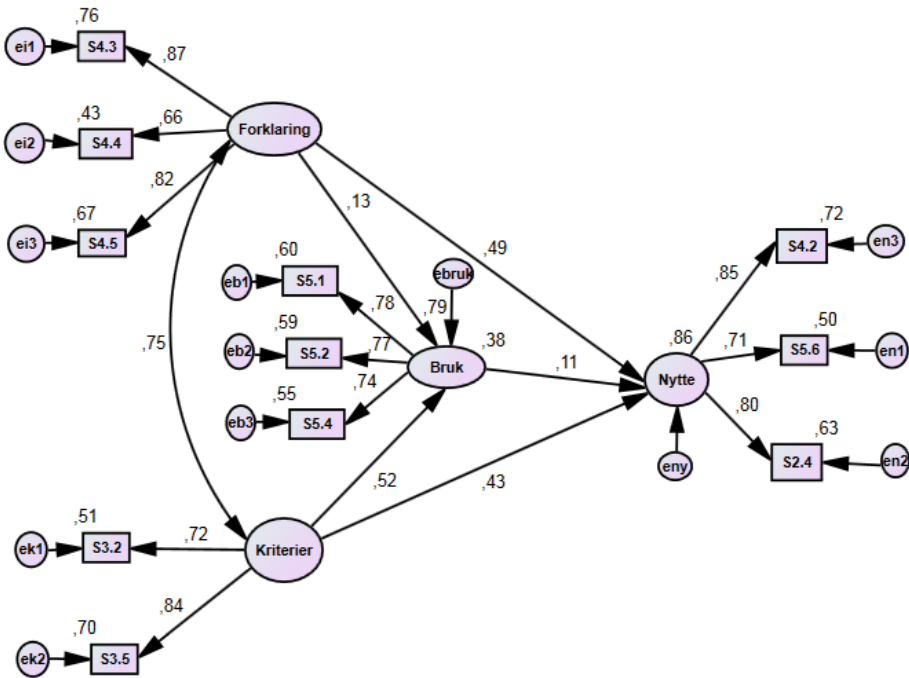
Standardized Total Effects (Group number 1 - Default model)

	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

Modellen bygger på resonnementet 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

Figur 2. Strukturmodell med bruk som medierende variabel

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 $b(\text{Forklaring} \rightarrow \text{Bruk}) = ,13$ og støtter ikke hypotesen. Assosiasjonen mellom elevs 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.