

Læringsfremmende vurderingspraksis på ungdomstrinnet: Funn fra TALIS 2018 og LISA-prosjektet

Tone Cecilie Carlsten

Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU)

Inger Throndsen & Kirsti Klette

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Abstract: Assessment practices promoting learning at the lower secondary level: Findings from TALIS 2018 and the LISA project

This chapter presents some of the results from the Norwegian TALIS 2018 survey related to professional knowledge, competence, and practice. Findings from teachers' reporting on their own teaching practices are highlighted, specifically the part of their teaching that deals with assessment practices. Results from the TALIS survey are then interpreted in light of the results from a research project that deals with the same topic with other data sources.

The Linking Instruction and Student Achievement project (LISA) combines video data from teaching practices in language arts and mathematics at the lower secondary level with test results from national tests (reading and arithmetic) from the same classes from which the observation data were taken. LISA is the largest video study of its kind of classroom practices in the Nordic countries and is a suitable frame of interpretation for TALIS 2018 data, as all data materials are linked to the last period of the former national curriculum in Norway. Findings from TALIS 2018 show that there could have been far more teachers who checked that the feedback they provide on written assignments is understood by their students. Results from the LISA project confirms this finding and also show that quite a few teachers are responsible for a large part of the feedback that is regarded to be of high quality. The chapter indicates that there is great potential for working more collectively on

this topic both in the teacher education stage and as part of a well-managed lifelong learning system for teachers.

Keywords: professional learning, assessment practices, classroom studies, video studies

Introduksjon

Resultatene fra TALIS 2018 er, som vi så i kapittel 1, delt inn i fem hovedområder. I dette kapitlet skal vi gå nærmere inn på resultater som hører hjemme under område 1: profesjonskunnskap, kompetanse og praksis. Vi skal spesielt se på læreres rapportering om egen undervisningspraksis, nærmere bestemt den delen av deres undervisning som handler om vurderingspraksis.

Noen av hovedfunnene om læreres vurderingspraksiser er tidligere presentert i den norske kortrapporten *TALIS 2018 – Første hovedfunn fra ungdomstrinnet* (Thronsen et al., 2019). I dette kapitlet skal vi trekke fram noen eksempler når det gjelder å se på TALIS 2018-data om læringsfremmende vurderingspraksis. Vi skal dessuten se på hvordan funn fra TALIS-undersøkelsen kan tolkes i lys av resultater fra et forskningsprosjekt som behandler samme tema med andre datakilder. Til dette formålet har vi valgt å ta utgangspunkt i LISA, et forskningsprosjekt som kombinerer videodata fra undervisning i norsk og matematikk på ungdomstrinnet med testresultater fra framgang på nasjonale prøver (lesing og regning) fra de samme klassene som observasjonsdataene er hentet fra. LISA er et akronym for *Linking Instruction and Student Achievement*,¹ og er den største videostudien på skole i sitt slag fra de nordiske land. Den gir dermed et unikt innblikk i hvordan det undervises i norske klasserom på ungdomstrinnet (Klette et al., 2017). Prosjektet er et godt supplement til å tolke data fra TALIS 2018, ettersom begge datainnsamlinger er knyttet til undervisning etter læreplanene i siste periode av Kunnskapsløftet.

Som nevnt i kapittel 1 og 5 la vi til noen nasjonale tillegsspørsmål i TALIS 2018. Noen av disse var utarbeidet i samarbeid med prosjektleder av

1 <https://www.uv.uio.no/ils/forskning/prosjekter/lisa/>

LISA-studien, og handlet om læreres læringsfremmende vurderingspraksis. Målet med samarbeidet var å styrke spørsmålsstillinger som kunne utdype tidligere hovedfunn fra TALIS 2008 og TALIS 2013. TALIS-funn fra 2008 viste at det var en relativt svakt utviklet faglig oppfølgingskultur fra lærer til elev (Vibe et al., 2009). I analyser av datamaterialet fra 2013 kunne vi se antydninger til at dette forholdet var noe forbedret (Caspersen et al., 2014). Litt av utfordringen med sammenligningene før TALIS 2018 var at spørsmålene var endret mellom 2008 og 2013, og at de i liten grad var presise nok for å tolkes i lys av forskning på vurderingspraksis i den norske konteksten. I dette kapitlet skal vi derfor se nærmere både på hovedfunn og funn fra de nasjonale tilleggsspørsmålene, og tolke dem opp mot noen hovedfunn fra LISA-prosjektet. I LISA er det samlet inn data fra undervisning i matematikk- og norskfaget. For dette kapitlet har vi derfor valgt å bruke utvalgte data for vurdering og tilbakemelding i norsk og matematikk også fra TALIS.

Forskningsspørsmålet vårt er: *Hvilken hyppighet og kvalitet er det på tilbakemeldinger til elever på ungdomstrinnet i norsk- og matematikkfaget ifølge TALIS 2018-data og LISA-data, og hvordan kan disse studiene komplementere hverandre for å gi bedre forståelse for hva som kjennetegner læringsfremmende vurderingspraksis i Norge?*

Kapitlet gir først en begrepsavklaring og en kort status for norsk policyutvikling knyttet til læreres vurderingspraksis. Deretter presenteres mål og utvalgte funn fra TALIS 2018-undersøkelsen for dette temaet for norsk- og matematikkfaget, før vi sammenholder dette med mål og noen relevante resultater fra LISA-studien. Til slutt oppsummerer vi hva disse funnene sier om kvaliteten på fortløpende tilbakemeldinger, og hvordan ulike datakilder kan utfylle hverandre når det gjelder å gi skoleledere, lærere og lærerstudenter kunnskap om hva som kjennetegner god vurderingspraksis på ungdomstrinnet.

Rammer for vurderingspraksis i norsk skole

Vi starter denne delen med en begrepsavklaring. I vurderingssammenheng har det tradisjonelt vært skilt mellom formativ og summativ vurdering. Mens det første er en vurderingsform som har som mål å fremme

læring, skiller den seg fra summativ eller avsluttende vurdering som har som mål å bedømme eller rangere læringsresultater (Scriven, 1967). Senere har forskere ønsket å erstatte begrepet formativ med læringsfremmende, fordi dette på en bedre måte viser til fortløpende praksiser som ikke anses som «mini-summative». Det vil si at hensikten er å vektlegge pedagogiske aspekter ved tilbakemelding til elever som også ligger innenfor deres juridiske rett til medvirkning i vurderingsarbeidet. Dette blir i norsk sammenheng også kalt underveisvurdering. Det handler altså ikke om at lærere skal samle dokumentasjon for framtidig karaktersetning, slik det «mini-summative» også har blitt kritisert for i kjent internasjonal forskning på tilbakemelding i klasserommet (jf. Black & Wiliam, 1998; Hattie & Timperley, 2007; Wiliam, 2010).

I dette kapitlet bruker vi begrepet læringsfremmende vurderingspraksis om den systematiske delen av en lærers undervisning som dreier seg om å gi tilbakemeldinger som har som mål å fremme elevers læring. Vi gjør dette for å understreke at begrepet underveisvurdering, slik det ofte brukes i policyssammenheng, nylig har fått et tillegg til læringsfremmende vurdering i fag: i henhold til forskrift til opplæringslova (2006) er nå også lærelyst tatt inn i formålet med vurdering i fag (Forskrift til opplæringslova, kap. 3, 2020). Dette var ikke en del av forskriften på det tidspunktet TALIS 2018 og LISA-prosjektet ble utviklet og gjennomført, men var likevel en del av det forskningsbaserte grunnlaget for prosjektene. Vi bruker også bevisst begrepet vurderingspraksis framfor vurderingskultur eller oppfølgingskultur, slik det ble gjort i rapporteringer fra TALIS 2008 og TALIS 2013. Dette gjør vi for her å fokusere lærernes egenrapporterte vurderingshandlinger i TALIS 2018. En internasjonal komparativ storskala-spørreundersøkelse som TALIS har heller ikke gode nok data til å si så mye om kvaliteten på det fysiske miljøet, menneskene som inngår i det og selve relasjonene dem imellom på en måte som er sammenlignbart med det man kan observere i data fra LISA-prosjektet.

Vi skal nå ramme inn hvordan vi forstår vurderingspraksis i norsk skole ved å gi en kort gjennomgang av noen sentrale endringer i policyfeltet de siste 15 år. I kapittel 5 skisserte vi hvordan rammene for skoleutvikling på ungdomstrinnet har blitt tydeliggjort i policyssammenheng spesielt de siste 8 årene (Carlsten, 2021). I sammenheng med vurderingsfeltet sett i

lys av skoleutvikling kan vi gå litt lenger tilbake; til St.meld. 16 (2006–2007) ...*og ingen stod igjen. Tidlig innsats for livslang læring*. Her ble utfordringene for elevvurdering oppsummert i fire punkter: 1) Regelverket for individvurdering oppfattes ikke klart nok; 2) både lærerutdanningen og skolen mangler tilstrekkelig kompetanse; 3) det er svak vurderingskultur og vurderingspraksis i skolen; 4) det er forsket lite på individvurdering i Norge. Dette gjaldt ikke bare ungdomstrinnet. Funn fra evalueringen av Reform 97 viste at vurderingspraksis på barnetrinnet blant annet bar preg av fravær av eksplisitte, klare faglige standarder (Klette, 2003). OECD/CERI (2005) etterlyste også en mer systematisk oppfølging av elevers utvikling på ungdomstrinnet. Som en oppfølging av St.meld. 16 (2006–2007) etablerte Utdanningsdirektoratet prosjektet *Bedre vurderingspraksis* i 2007. Sluttrapporten fra forskerne som evaluerte prosjektet anbefalte en videreføring for å styrke vurderingskompetansen i grunnopplæringen (Thronsdén et al., 2009). Høsten 2009 ble ny forskrift om individuell vurdering fastsatt, og disse ga et tydeligere læringsfokus i bestemmelsene om undervisningsvurdering. Det er altså denne forskriften som formelt lå til grunn i Norge da data ble samlet inn i TALIS 2018 og i LISA-prosjektet.

I stortingsmeldingen *Motivasjon – Mestring – Muligheter* ble det pekt på tiltak som burde settes i gang for å øke lærernes undervisningskompetanse og dermed kvaliteten på elevenes opplæring (Meld. St. 22 (2010–2011)). Ett av områdene som ble nevnt spesielt var elevvurdering. Det førte til den fireårige nasjonale satsingen *Vurdering for læring* (VFL) som ble etablert 2010 (Nasjonalt kompetansemiljø i vurdering, 2012). Satsingen VFL ble videreført fra 2014 til 2018, som en oppfølging av Meld. St. 20 (2012–2013) *På rett vei*. Den nasjonale satsingen på læringsfremmende vurdering gjaldt hele skolesektoren. I en rapport om VFL gjennomført som en del av *Spørsmål til Skole-Norge våren 2017*, ble det konkludert med at skoleledere og -eiere som hadde deltatt i satsingen ga uttrykk for et mer omfattende arbeid med vurderingspraksis enn de som ikke hadde deltatt (Larsen et al., 2017). I Utdanningsdirektoratets egen rapportering om erfaringer fra satsingen, basert på flere kilder, framkommer det at det fortsatt er områder som er utfordrende når det gjelder å styrke læringsfremmende vurdering. Ikke minst pekes det på faren ved at vurderingspraksis forenkles og reduseres til «et sett av innøvde prosedyrer og

mekaniske teknikker uten at dette bidrar til reflektert praksis og læring» (Utdanningsdirektoratet, 2019). Dermed vises det til et behov for at lærere har god kjennskap til forskning og grunntanker som ligger til grunn for læringsfremmende vurdering, slik dette kapitlet gir noen eksempler på.

Den nye overordnede delen av læreplanverket fagfornyelsen understreker også betydningen av læringsfremmende vurderingspraksis (Kunnskapsdepartementet, 2017). Blant annet er det en viktig intensjon i oppfølgingen av Meld. St. 28 (2015–2016) *Fag – Fordypning – Forståelse – En fornyelse av Kunnskapsløftet* at læreplanene skal legge til rette for dybdelæring. For å relatere dette til vårt forskningsspørsmål, så vil dermed *hyppighet og kvalitet på læreres tilbakemeldinger* handle om fortløpende tilbakemeldinger som er basert på en balanse mellom forskningsbasert og skjønnsbasert vurderingskompetanse i møte med elevers aktive rolle i læringsprosessen – på en måte som styrker reflekterte sider av praksis og læring også for lærere. En læringsfremmende vurderingspraksis vil i utvidet forstand også handle om å koble det systematiske arbeidet med vurdering for læring bedre til læreplanarbeid som en del av skoleutvikling.

Selv om våre TALIS-data som tidligere nevnt både er begrenset til hyppighet av vurderingshandlinger og en tidsramme som ligger forut for høsten 2020, da det nye læreplanverket trådte i kraft, er dette elementer som nasjonal og internasjonal litteratur innen klasseromsforskning og forskning på lærerprofesjonalitet lenge har pekt på som forutsetninger for god undervisningspraksis i skolen (se f.eks. Ball & Cohen, 1999; Grossman & Loeb, 2010; Grossman & McDonald, 2008; Hattie, 2009; Hiebert et al., 2002; Hiebert & Stiegler 2000; Klette, 2018; Klette et al., 2017; Seidel & Shavelson, 2007; Timperley & Alton-Lee, 2008). Vi skal derfor trekke noen linjer mellom våre funn og gi noen implikasjoner for god læringsfremmende vurderingspraksis helt til slutt i dette kapitlet. Først skal vi imidlertid se nærmere på hva TALIS 2018- og LISA-analysene viser oss om læringsfremmende vurderingspraksis på ungdomstrinnet i Norge.

Læreres vurderingspraksis i TALIS 2018

Når det gjelder mål og metode for TALIS-studien viser vi til kapittel 1, og går her inn på funn som er relevante for dette kapitlet. For å få

kunnskap om vurderingsformene som lærerne på ungdomstrinnet benytter i undervisningen, ble de i TALIS 2018 bedt om å besvare følgende spørsmål: *Hvor ofte bruker du følgende metoder for å vurdere elevenes ferdigheter?* Her skulle lærerne ta utgangspunkt i elevvurderingen sin i en nærmere definert klasse. Tabell 1 viser andelen lærere som underviser i norsk (lesing, skriving og litteratur på morsmålet) og matematikk som svarer «ofte»/«alltid» på de ulike utsagnene om elevvurdering. De andre svaralternativene var «aldri eller nesten aldri» og «noen ganger».

Tabell 1 Prosentandelen lærere som underviser i norsk og matematikk i en aktuell klasse som svarer at de benytter ulike vurderingsformer «ofte»/«alltid».

Utsagn	«Ofte»/«Alltid» Norsk	«Ofte»/«Alltid» Matematikk
Jeg utvikler og gjennomfører egne prøver.	79	70
Jeg gir skriftlig tilbakemelding på elevenes arbeid i tillegg til karakter.	93	77
Jeg lar elevene selv vurdere sin egen framgang.	44	38
Jeg observerer hvordan elevene jobber med bestemte oppgaver og gir dem tilbakemelding.	73	56

Som tidligere nevnt, har det fra nasjonalt hold vært en langvarig satsing på å utvikle lærernes formative vurderingskompetanse. Faglige tilbakemeldinger anses som et viktig element i læringsfremmende vurdering. Som det går fram av tabell 1 svarer hele ni av ti lærere som underviser i norsk at de «ofte»/«alltid» gir skriftlig tilbakemelding på elevenes arbeider i tillegg til å sette karakter. Det er nesten åtte av ti lærere som underviser i matematikk som gjør det samme. Dette resultatet gjør at Norge er blant landene hvor en slik vurderingspraksis forekommer hyppigst (i alt 48 land deltok i TALIS 2018, inklusive alle de nordiske landene). Gode skriftlige tilbakemeldinger har en viktig funksjon ved at de kan gi elevene bedre innsikt i hvor de står faglig, og hva de bør vektlegge i det videre arbeidet med læringsstoffet (Hattie & Timperley, 2007). Fortløpende vurdering av denne typen vil derfor kunne fungere som et redskap for videre læring. På tvers av alle fag var det 82 % av norske lærere som svarte at de «ofte»/«alltid» ga slik tilbakemelding. I TALIS 2013 svarte 75 % av de norske ungdomsskolelærerne at de ga elevene skriftlig tilbakemelding i tillegg til karakter. Det har således vært en positiv økning på dette området siden forrige runde av TALIS-undersøkelsen.

Når det gjelder faglig feedback, svarer hele sju av ti lærere som underviser i norsk at de observerer hvordan elevene jobber med bestemte oppgaver og gir dem tilbakemelding på det de gjør (jf. tabell 1). Slik umiddelbar tilbakemelding kan være en viktig korreksjon eller pekepinn på om elevene er på «feil spor» i forhold til læringsmålet. 56 % av lærerne som underviser i matematikk svarer at de gir slik fortløpende tilbakemelding, så her er det et forbedringspotensiale.

Tabellen viser også at omtrent fire av ti lærere involverer elevene i vurderingen, og at dette er ganske likt mellom norsk og matematikk. Tallet på tvers av fag lå i TALIS 2018 på 39 % for norske lærere. I faglitteraturen blir selvvrdering sammen med tilbakemeldinger løftet fram som sentrale elementer i læringsfremmende vurdering (Black & Wiliam, 2009). For at elever skal kunne nyttiggjøre seg de faglige tilbakemeldingene, er det imidlertid en forutsetning at de forstår hva som ligger i lærerens vurdering. Selvvrdering kan bidra til å fremme slik forståelse. Til sammenligning svarte 29 % av lærerne at de lar elevene vurdere sin egen framgang i TALIS 2013. Det har altså vært en positiv endring på dette området også.

For å undersøke om lærerne er opptatte av å sjekke om tilbakemeldingene de gir på skriftlige oppgaver blir forstått av elevene, ble det i spørreskjemaet lagt til et nasjonalt spørsmål om nettopp dette. Følgende utsagn ble benyttet: *Jeg sjekker at tilbakemeldinger jeg gir på skriftlige oppgaver blir forstått av elevene.* Analyser viser at drøyt halvparten av lærerne (53 %) sjekker «ofte»/«alltid» at elevene forstår tilbakemeldingene som de får på tvers av alle fag. For lærere som underviser i norsk er tallet imidlertid så høyt som 74 %, mens det for lærere som underviser i matematikk er på 54 %. Selv om det er forskjell på oppgaveformer som gis i de to fagene, skulle lærerne her svare på grunnlag av de skriftlige oppgavene som faktisk gis. For at tilbakemeldinger skal være læringsfremmende, er det helt vesentlig at den enkelte elev forstår innholdet i tilbakemeldingen (Black & Wiliam, 2009). På dette området ser det med andre ord ut til at lærernes måter å gi slike tilbakemeldinger på har et forbedringspotensial, spesielt i matematikk.

Videre viser tabell 1 at syv av ti lærere utvikler og gjennomfører egne prøver. Faglige prøver har alltid hatt en sentral plass i undervisningen, og har i stor grad blitt assosiert med summativ vurdering (dvs. vurdering av

læring). Prøver kan imidlertid brukes på en formativ måte (dvs. vurdering for læring) dersom resultatene brukes til å forbedre både lærerens undervisning og elevenes læring. Det avgjørende er med andre ord i hvilken grad elevenes svar på prøver bidrar til at læreren korrigerer og tilpasser undervisningen på bakgrunn av elevenes prestasjoner, samt at de blir brukt som grunnlag for tilbakemelding om hva elevene har forstått og hva de ikke har forstått. Det er dette som avgjør om bruken av prøver kan anses som formativ eller ikke. Ifølge sentrale forskere på området er vurderingen formativ dersom den fører til handlinger som fremmer elevenes læring med utgangspunkt i arbeidet deres (Harlen, 2007; Stobart, 2008).

Forskerorganisasjonen EARLI (European Association for Research on Learning and Instruction) har argumentert for viktigheten av at vurdering for læring integreres i lærernes klasseromsundervisning. Formativ vurdering dreier seg ikke om å benytte en bestemt metode, men læringsfremmende undervisning kjennetegnes ved følgende:

- *Eksplisitte mål* for undervisningen gjøres kjent for elevene, og kriterier illustrerer hvordan mål kan være om de er oppnådd.
- *Læreren stiller «rike spørsmål»* og gir elevene tid til å reflektere før det bes om et svar. Dette gjør at elevene bedre får vist hvor de er i læringsprosessen.
- *Læreren gir tilbakemeldinger* hvor formålet er å hjelpe elevene til å minske gapet mellom det de kan og det de forsøker å strekke seg etter. Tilbakemeldingene skal gi elevene kunnskap om hvordan de skal kunne komme seg til det ønskede læringsmålet.
- *Vurdering av seg selv og andre* bidrar til å øke elevenes forståelse av eget og andres arbeid, særlig når det gjelder kvalitet. Elever som utvikler kunnskap på dette feltet, øker sin bevissthet om egen læring.

TALIS 2018 inneholdt også spørsmål om kjennetegn ved god undervisning, og noen av spørsmålene er relevante i forhold til punktene overfor som dreier seg om karakteriske trekk ved formativ undervisningspraksis. Tabell 2 viser andelen lærere som underviser i norsk (lesing, skriving og litteratur på morsmålet) og matematikk som svarer «ofte»/«alltid» på

spørsmålet «Tenk på undervisningen din, hvor ofte gjør du følgende?» hvor utsagnene går på å sette seg mål og å stille spørsmål til elevene. De andre svaralternativene var «aldri eller nesten aldri» og «av og til».

Tabell 2 Prosentandelen lærere som underviser i norsk og matematikk i en aktuell klasse som svarer «ofte»/«alltid» på utsagn om sin undervisningspraksis

Utsagn	«Ofte»/«Alltid» Norsk	«Ofte»/«Alltid» Matematikk
Jeg starter timen med å presentere klare mål for timen.	77	73
Jeg forklarer hva jeg forventer at elevene skal lære.	87	86
Jeg presenterer oppgaver hvor det ikke er en opplagt løsning.	66	41
Jeg oppfordrer elevene til å velge egne måter å løse komplekse oppgaver på.	49	60

Tabell 2 viser at undervisningstimer i norsk og matematikk har noenlunde lik struktur ved at de har forholdsvis mange eksplisitte mål som framstilles for elevene. Det er en noe ujevn balanse mellom lærerens eksplisitte undervisningsformer og mulighetene som gis elever til selv å bidra i former som er ment å mer eksplisitt styrke dybdeløring.

Som det går fram av denne presentasjonen av TALIS-data, så gir ikke TALIS-undersøkelsen informasjon om kvaliteten i lærernes tilbakemeldinger. Selv om en slik survey er begrenset til å si noe om hyppighet av tilbakemeldinger, så ser vi at valg av spørsmål likevel har en sammenheng med forskning på karakteristiske trekk ved god formativ vurderingspraksis. Funnene viser at det er en del likheter mellom norsk og matematikk når det gjelder spesielt den delen av undervisningspraksis som legger grunnlaget for eksplisitt struktur. Vi finner også at det er en lavere andel lærere som inkluderer former for tilbakemeldinger som trekker inn elevene i vurderingsarbeidet og gir grunnlag for dybdeløring. Som det ble pekt på i Utdanningsdirektoratets rapportering om erfaringer fra satsingen *Vurdering for læring*, er dette områder som er utfordrende når det gjelder å styrke læringsfremmende vurdering. Ikke minst ble det pekt på faren ved at vurderingspraksis forenkles og reduseres til «et sett av innøvde prosedyrer og mekaniske teknikker uten at dette bidrar til reflektert praksis og læring» (Utdanningsdirektoratet, 2019). Former for tilbakemeldinger som gir slik kvalitet handler altså om det vi har vært inne på

i analysen over, som handler mer om den løpende muntlige tilbakemeldingen mellom lærere og elever i selve undervisningen. LISA-prosjektet er designet nettopp med hensikt å finne ut mer om hva som kjennetegner og gir kvalitet i slik praksis, noe vi skal komme inn på i neste avsnitt.

Hvordan vurderes læreres tilbakemeldinger i LISA-prosjektet?

Før vi trekker fram hovedfunn om læreres undervisningspraksis fra LISA-prosjektet som gir innsikt i hva som kjennetegner kvalitet i læringsfremmende tilbakemeldinger, vil vi kort gi en beskrivelse av designet. Prosjektet ble startet opp ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo i 2013, som et fireårig grunnforskningsprosjekt finansiert av Norges forskningsråd. Utgangspunktet for studien var tredelt: For det første var det et mål å forstå hvordan undervisningspraksis hang sammen med elevenes læring i matematikk og norskfaget målt via framgang på nasjonale prøver i lesing og regning. For det andre var det et ønske om å utvikle et forskningsdesign som kunne binde sammen klasseromsdata med data fra elevresultater (her: framgang på nasjonale prøver) i en integrert modell. For det tredje lå det til grunn en intensjon om å teste ut standardiserte kodingsmanualer og instrumenter som kunne måle kvalitet på undervisning i en norsk kontekst. Ved å teste ut og videreutvikle robuste og internasjonale instrumenter komparativt, kunne dette bidra med nye metoder for å få informasjon om lærernes undervisning på tvers av klasserom og kontekster og gi oppdatert kunnskap om dette relevant for både lærerutdanningen og i etter- og videreutdanning av lærere. I dette kapitlet skal vi trekke fram funn som handler om læreres tilbakemeldinger i norsk og matematikk som et eksempel, og ser dermed ikke direkte på elevenes prestasjoner i denne sammenheng.

I LISA-prosjektet har forskerne gjort videoopptak av fire påløpende undervisningstimer i norsk og matematikk på 8. trinn ved tilnærmet 50 ungdomsskoler i Norge ($n = 47$ i matematikk og norsk, totalt 94 klasserom). Ett kamera filmet klassen som helhet og ett kamera har fulgt læreren. En grunn til at videodata er foretrukket framfor klasseromsobservasjoner

innhentet på stedet er at det er en mer presis og dermed systematisk måte å innhente data, der det også er mulig å tydelig skille mellom datainnhenting og datatolkning. Videre kan dataene analysere fra ulike vinkler og perspektiver samt at flere forskere har mulighet til å analysere samme datamateriale – inkludert å diskutere mulige tolkninger. Videre ble det innhentet data om de observerte lærerne, så som utdanningsbakgrunn, faglig skoloring i undervisningsfaget og yrkeserfaring. Elevene har også svart på en spørreundersøkelse, og forskerne har vurdert observasjonsdata opp mot elevenes spørreskjemadata og de samme elevenes framgang på nasjonale prøver i henholdsvis lesing og regning.

I LISA-prosjektet blir videodata fra klasserom på ungdomstrinnet analysert med en kodemanual som kalles PLATO (Grossman et al., 2013). Dette er et akronym for *The Protocol for Language Arts Teaching Observation*. Kodemanualen er utarbeidet av forskere ved Stanford University, og man må være sertifisert for å kunne bruke denne. Grunnen er at de som observerer kan kalibrere og styrke sine tolkninger basert på en felles forskningsbasert og validert forståelse av hva undervisningsaktiviteter som lærernes bruk av tilbakemeldinger, strategiundervisning og tilsvarende innebærer. PLATO-manualen har blitt oversatt til norsk og justert til en norsk kontekst. Den er egnet for å analysere mange viktige sider ved undervisning og vurdering, men adresserer ikke helheten i en skolekultur, slik som emosjonelle aspekter eller lærerens planlegging og refleksjoner tilknyttet faglig arbeid utenfor klasserommet. Den er imidlertid godt egnet for å analysere læringsfremmende vurderingshandlinger både i norsk og andre fag i justert format.

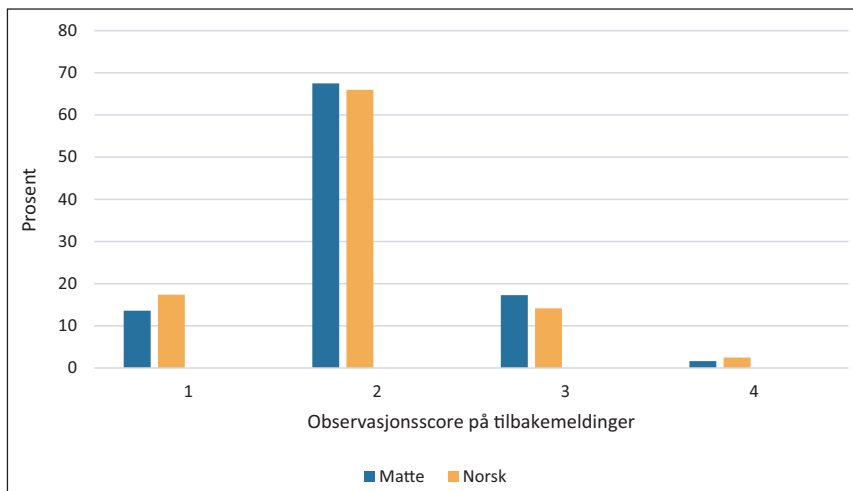
PLATO består av fire hovedområder: 1) undervisningsmessig støtte (scaffolding); 2) faglige krav; 3) bruk og representasjon av fagstoff; 4) klasseromsmiljøet. Disse områdene er igjen delt inn i 12 elementer som skåres av observatører på en skala fra 1–4. I PLATO-manualen er tilbakemelding et element innenfor området «undervisningsmessig støtte». I analysene fokuseres kvaliteten på tilbakemeldinger og især hvordan læreres (og elevenes) tilbakemeldinger kan bistå elevene i deres videre arbeid. Tilbakemelding i denne sammenheng handler altså om hvordan læreren kommenterer kvaliteten på elevenes arbeider, og ikke minst på hvordan de blir veiledet videre for å heve kvaliteten på sitt eget arbeid.

I kodemanualen er gode tilbakemeldinger (score 3 og 4) karakterisert ved at de er spesifikke og fagrettede, og at de bidrar til at elevene forstår hvordan de kan forbedre konkrete sider ved eget arbeid / de ulike aktiviteter. Fravær av tilbakemeldinger og tilbakemeldinger som er generelle og vage gis respektive score 1 og score 2. Score 2-tilbakemeldinger er for eksempel karakterisert ved at de er lite forankret i elevenes konkrete arbeid og der forslag til forbedringer er mer preget av å være innrettet mot selve aktiviteten enn mot hvilke ferdigheter eller kunnskap elevene har og trenger.

Observasjonskodene som benyttes i LISA-prosjektet bygger på forskningsbasert fagdidaktisk kunnskap om undervisning innen hvert av de aktuelle fagområdene (language arts/mathematics), og oppdatert generell litteratur på hva som kjennetegner undervisning av høy kvalitet i klasserom. Kodemanualene er kontinuerlig i en prosess av forskningsbasert utvikling og faglig oppdatering basert på analyser av hundrevis av eksempler på diverse læreres undervisning både i USA og nordiske land.² Likevel er manualene såpass integrert i et internasjonalt fagfellesskap (derav også kravet om sertifisering for å bruke dem) at justeringer ikke skal skje på en slik måte at det går ut over kvalitet og internasjonale komparative dimensjoner.

Hva kan LISA-prosjektet gjennom denne metoden si oss om kvalitet på læreres tilbakemeldinger i klasserommet? Ser vi på tvers av de 95 observerte klasserommene (48 i matematikk og 47 i norsk) er især to trekk framtrædende: Gjennomgående gis det relativt mye og hyppige tilbakemeldinger i disse klasserommene. Hovedvekten av disse tilbakemeldinger er imidlertid hva vi karakteriserte som score 2; det vil si vage, generelle (og ofte positive) tilbakemeldinger. Typiske tilbakemeldinger av denne typen er «Bra, dere!», «Riktig», «Ikke helt», «Kanskje du kan spørre hva læringspartneren din mener?», «Hva med noen flere detaljer?» eller «Her skulle det være avsnitt». Som vi ser av figur 1 forekommer denne type tilbakemeldinger i de fleste observerte klasserom og i begge fag.

2 <https://www.uv.uio.no/quint/english/about/>



Figur 1 Læreres tilbakemeldinger til elever i matematikk og norsk i observerte klasserom på ungdomstrinnet i Norge i LISA-prosjektet.

Tilnærmet 70 % av de observerte tilbakemeldingene er hva vi betegner som generelle, overflatiske og hva vi også vil karakterisere som prosedurale (hvordan løse en oppgave). De gir liten anvisning for hvordan eleven kan jobbe videre for å forbedre eget arbeid. Under 5 % av observasjonene har blitt rubrisert som høykvalitets-tilbakemeldinger. Det vil si tilbakemeldinger som er konkrete, spesifikke og gir anvisninger for hvordan elevene kan bruke lærerens tilbakemelding for å forbedre eget arbeid.

Av de 95 observerte lærerne i LISA-prosjektet var det 9 lærere (fem i norsk og fire i matematikk) som var ansvarlige for alle tilbakemeldinger med høye scorer (score 4). Dette er et interessant funn. Det peker mot at de som har et repertoar for hvordan de kan gi konkrete tilbakemeldinger bruker det – og at de bruker det konsistent. Nærstudier av de fem lærerne som fikk score 4 i norsk viste for eksempel at disse fem ga høykvalitets-tilbakemeldinger på mange og høyst ulike måter. Det som var felles, var at de brukte dem konsistent og på tvers av situasjoner, og at de ofte kombinerte de konkrete og spesifikke tilbakemeldingene (score 4) med mer vage og generelle tilbakemeldinger (score 2) for å skape en skjønnsbasert balanse og flyt i kommunikasjonen med elevene (Dåsvatn, 2016).

Det er positivt at norske lærere på ungdomstrinnet gir oppmuntring og generelle tilbakemeldinger. Det er imidlertid grunn til å være bekymret over at så få av de observerte lærerne hadde et repertoar hvor

de kombinerte tilbakemeldinger som «Fint» og «Bra jobbet!» med spesifikke og konkrete tilbakemeldinger anvisninger som kunne hjelpe elvene i deres videre arbeid. Dette eksemplet viser derfor at det fremdeles er en vei å gå for at norske lærere kan legge bedre til rette for de målene som både policy og forskning peker på som god vurderingspraksis. Funnene fra LISA-prosjektet viser også at det fortsatt er et stykke arbeid som må gjøres før kvaliteten på tilbakemeldinger til elever for å fremme læring blir en del av skolens arbeid med vurderingspraksis som del av en felles skoleutvikling.

Konklusjon

Vi startet dette kapittelet ved å sammenstille sider av læreres undervisningspraksis og vurderingspraksis. Vi har pekt på at disse to sidene ved læreres arbeid ikke alltid har blitt kategorisert på en integrert måte, men at det har skjedd en endring innen både policy- og forskningsfelt som har gått i retning av å se deler av vurderingsarbeidet som en kvalitet ved læreres undervisningspraksis.

Som nevnt har de nasjonale satsingene rundt vurderingsarbeid i skolen satt som mål å gjøre læringsfremmende vurderingspraksis til et systematisk arbeid som også knyttes bedre til læreplanarbeid som en del av en kollektiv skoleutvikling. Utdanningsdirektoratets rapportering om erfaringer fra satsingen *Vurdering for læring* understreker dette kollektive aspektet ved vurderingsarbeid. Ikke minst framheves det at mangel på forskningsbasert kunnskap om ulike sider av dette arbeidet kan føre til at vurderingspraksis forenkles og reduseres på instrumentelt vis, som i liten grad bidrar til reflektert praksis og læring for lærere (Utdanningsdirektoratet, 2019). Dette kapittelet har vist hvordan to ulike forskningsdesign gir komplementerende måter å studere slik praksis, spesielt når det gjelder å forstå hva som kjennetegner en læringsfremmende vurderingspraksis som en integrert del av læreres undervisningspraksis. Vi har spesielt trukket fram eksempler når det gjelder hvordan lærere gir fortløpende skriftlige og muntlige tilbakemeldinger i fagene norsk og matematikk.

TALIS-undersøkelsen skiller på noen måter mellom undervisningspraksis og vurderingspraksis, men har også elementer som knytter disse

sammen, slik vi har vist. TALIS sier lite om kvaliteten på tilbakemeldinger lærere gir sine elever, men lar oss undersøke hyppigheten av tilbakemeldinger knyttet til læreres konkrete undervisningsfag og klasser. LISA-prosjektet har hatt som mål å undersøke både hyppighet og kvalitet på tilbakemeldinger i observerte klasserom, og kan dermed utfylle rom i kunnskapsfeltet om hva som kjennetegner læreres vurderingspraksis på ungdomstrinnet i Norge.

Funn fra TALIS 2018 viser at det har skjedd en forbedring ved flere sider av læreres vurderingspraksis siden TALIS 2013. Vi har imidlertid pekt på at det er noen forskjeller mellom læreres selvrapporteringer om vurderingsformer i norsk og matematikk i TALIS-materialet. Dette gjelder spesielt muntlig faglig feedback, der det er en høyere andel lærere som underviser i norsk som observerer hvordan elevene jobber med bestemte oppgaver og gir dem tilbakemelding på det de gjør. Slik umiddelbar tilbakemelding kan være en viktig korleksjon eller pekepinn på om elevene er på «feil spor» i forhold til læringsmålet.

Det norske tillegget til TALIS-undersøkelsen viser også at det kunne vært langt flere lærere som sjekket at tilbakemeldinger de gir på skriftlige oppgaver blir forstått av elevene. Dette gjelder spesielt i matematikkfaget, selv om vi har påpekt at det kan være noen naturlige forskjeller mellom disse to fagene når det gjelder hvilke didaktiske valg lærere tar. Funn fra LISA-prosjektet indikerer at det innen begge fag er forholdsvis lav kvalitet på tilbakemeldingene, når disse vurderes av forskere som observerer etter gitte kriterier basert på oppdatert forskningsbasert kunnskap. LISA-eksemplet som er tatt med her viser også at det er ganske få lærere som står for en stor del av de tilbakemeldinger som har høy kvalitet. Dette indikerer at det er stort potensiale for å jobbe kollektivt videre med dette tema både i den første lærerutdanningen og som del av en godt ledet livslang læring for lærere.

Referanser

- Ball, D. L. & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. I G. Sykes & L. Darling-Hammond (Red.), *Teaching as the learning profession: Handbook of policy and practice* (s. 3–32). Jossey Bass.

- Black, P. & Wiliam, D. (1998). *Inside the black box. Raising standards through classroom assessment*. King's College London.
- Carlsten, T. C. (2021). Profesjonelle fellesskap på ungdomstrinnet: Om spørsmål og svar i TALIS 2018. (I denne boken).
- Carlsten, T. C., Björnsson, J. K. & Throndsen, I. (2021). *Hovedfunn fra TALIS 2018-undersøkelsen*. (I denne boken).
- Caspersen, J., Aamodt, P. O., Vibe, N. & Carlsten, T. C. (2014). *Kompetanse og praksis blant norske lærere: Resultater fra TALIS-undersøkelsen i 2013* (NIFU-rapport 41/2014). Nordisk institutt for studier av innovasjon, forskning og utdanning. <https://www.nifu.no/publications/1173469/>
- Dalland, C. & Klette, K. (2014). Work-plan heroes: Student strategies in lower-secondary Norwegian classrooms. *Scandinavian Journal of Educational Research*, 58(4), 400–423. <https://doi.org/10.1080/00313831.2012.739200>
- Dåsvatn, M. N. (2016). *Muntlige tilbakemeldinger i norskfaget. En videostudie av fem klasserom* [Masteroppgave, Universitetet i Oslo]. DUO vitenarkiv. <http://urn.nb.no/URN:NBN:no-55660>
- Forskrift til opplæringslova. (2006). *Forskrift til opplæringslova* (FOR-2006-06-23-724). Lovdata. <https://lovdata.no/forskrift/2006-06-23-724>
- Grossman, P. (2010). *PLATO manual*. http://cset.stanford.edu/media/PLATO_Overview.pdf
- Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wyckoff, J. Boyd, D. & Lankford, H. (2010). *Measure for measure: The relationship between measures of instructional practices in middle school English Language Arts and teachers' value added scores* (Working Paper 16015). National Bureau of Economic Research. <https://doi.org/10.3386/w16015>
- Grossman, P. & McDonald, M. (2008). Back to the future: Directions for research in teaching. *American Educational Research Journal*, 45(1), 184–205. <https://doi.org/10.3102/0002831207312906>
- Grossman, P., Loeb, S., Cohen, J. & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education*, 119(3). <https://doi.org/10.1086/669901>
- Harlen, W. (2007). Criteria for evaluating systems for assessment. *Studies in Educational Evaluation*, 33(1), 15–28. <https://doi.org/10.1016/j.stueduc.2007.01.003>
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 22(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hiebert, J., Gallimore, R. & Stiegler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, 31(5), 3–15. <https://doi.org/10.3102/0013189X031005003>

- Hiebert, J. & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS Study. *The Elementary School Journal*, 101(1), 3–20. <https://doi.org/10.1086/499656>
- Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G. C., Sleep, L. & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511. <https://doi.org/10.1080/07370000802177235>
- Klette, K. (2003). Lærernes klasseromsarbeid; Interaksjons- og arbeidsformer i norske klasserom etter Reform 97. I K. Klette (Red.), *Klasserommets praksisformer etter Reform 97* (s. 39–76). Pedagogisk forskningsinstitutt.
- Klette, K. (2018). A new generation classroom studies. I C. Osbeck, Å. Ingerman & S. Claesson (Red.), *Didactic classroom studies: A potential research direction* (s. 225–243). Nordic Academic Press.
- Klette, K., Blikstad-Balas, M. & Roe, A. (2017). Linking instruction and student achievement: Research design for a new generation of classroom studies. *Acta Didactica Norge*, 11(3). <https://doi.org/10.5617/adno.4729>
- Klette, K. & Carlsten, T. C. (2012). Knowledge in teacher learning: New professional challenges. I K. Jensen, L. C. Lahn & M. Nerland (Red.), *Professional learning in the knowledge society*. Sense Publishers.
- Kunnskapsdepartementet. (2017). *Verdier og prinsipper for grunnsopplæringen – overordnet del av læreplanverket*. Regjeringen. <https://www.regjeringen.no/no/dokumenter/verdier-og-prinsipper-for-grunnsopplaringen/id2570003/>
- Larsen, E. H., Vaagland, K. & Federici, R. A. (2017). *Vurdering for læring. Resultater og analyser av skoler og skoleeieres svar på spørsmål om vurderingspraksis* (NIFU-rapport 2017:24). NIFU. <https://www.nifu.no/publications/1504830/>
- Meld. St. 22 (2010–2011). *Motivasjon – Mestring – Muligheter*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld-st-22-2010--2011/id641251/>
- Meld. St. 20 (2012–2013) *På rett vei*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld-st-20-20122013/id717308/>
- Meld. St. 28 (2015–2016). *Fag – Fordypning – Forståelse – En fornyelse av Kunnskapsløftet*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/>
- Nasjonalt kompetansemiljø i vurdering. (2012). *Teoretisk bakgrunnsdokument for arbeid med vurdering for læring på ungdomstrinnet*. https://www.udir.no/globalassets/upload/ungdomstrinnet/rammeverk/ungdomstrinnet_bakgrunnsdokument_vurdering_for_laring_vedlegg_5.pdf
- OECD/CERI. (2005). *Formative assessment. Improving learning in secondary classrooms*. OECD.

- Scriven, M. (1967). The methodology of evaluation. I R. Tyler, R. Gagné & M. Scriven (Red.), *Perspectives of curriculum evaluation* (s. 39–83). Rand McNally.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- St.meld. nr. 16 (2006–2007). ...og ingen sto igjen. Tidlig innsats for livslang læring. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007-/id441395/>
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Taylor and Francis. <https://doi.org/10.4324/9780203930502>
- Thronsdén, I., Hopfenbeck, T. N., Lie, S. & Dale, E. L. (2009). *Bedre vurdering for læring – Rapport fra «Evaluering av modeller for kjennetegn på måloppnåelse i fag»*. Det utdanningsvitenskaplige fakultet, Universitetet i Oslo.
- Thronsdén, I., Carlsten, T. C. & Björnsson, J. K. (2019). *TALIS 2018 – Første hovedfunn fra ungdomstrinnet*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo. <http://hdl.handle.net/11250/2601320>
- Timperley, H. & Alton-Lee, A. (2008). Reframing teacher professional learning. *Review of Research in Education*, 32(1), 328–369. <https://doi.org/10.3102/0091732X07308968>
- Utdanningsdirektoratet. (2019). *Erfaringer fra nasjonal satsing på vurdering for læring (2010–2018)*. <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/erfaringer-fra-nasjonal-satsing-pa-vurdering-for-laring-2010-2018/4.-fortsatt-behov-for-a-videreutvikle-praksis/>
- Vibe, N., Aamodt, P. O. & Carlsten, T. C. (2009). *Å være ungdomsskolelærer i Norge: Resultater fra OECDs internasjonale studie av undervisning og læring (TALIS) (NIFU-rapport 2009:23)*. Nordisk institutt for studier av innovasjon, forskning og utdanning. <https://www.nifu.no/publications/970351/>
- Wiliam, D. (2010). The role of formative assessment in effective learning environments. I H. Dumont, D. Istance & F. Benavides (Red.), *The nature of learning. Using research to inspire practice*. OECD Publishing. <https://dx.doi.org/10.1787/9789264086487-en>