



Rolf V. Olsen
Therese N. Hopfenbeck
Sølvi Lillejord
Astrid Roe

Elevenes læringssituasjon etter innføringen av ny reform

Synteserapport fra evalueringen av Kunnskapsløftet

Elevenes læringssituasjon etter innføringen av ny reform

Synteserapport fra evalueringen av Kunnskapsløftet

Rolf V. Olsen¹, Therese N. Hopfenbeck^{1,2}, Sølvi Lillejord³ og Astrid Roe¹



¹ Institutt for lærerutdanning og skoleforskning, UiO

² Oxford University Centre for Educational Assessment

³ Kunnskapssenteret for utdanning, NFR

© ILS og forfatterne, Oslo, 2012

ISSN: 1891-3075

ISBN: 978-82-90904-98-7

Utgiver og redaksjon for denne webpublikasjonen:

Institutt for lærerutdanning og skoleforskning
Postboks 1099 Blindern
0317 OSLO

www.ils.uio.no

Telefon: (+47) 22 85 50 70

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

Prosjektet er finansiert fra Utdanningsdirektoratet.
Vi takker direktoratet for finansiell støtte og Programstyret for
verdifulle innspill underveis i prosessen. Likevel vil vi påpeke
at forfatterne selvsagt er ansvarlige for alt innholdet i denne rapporten.

Innholdsfortegnelse

Utgangspunktet for syntesen	4
Definisjoner og avgrensninger	4
Avgrensninger mot tre andre synteseområder	4
Definisjon av sentrale begreper i problemstillingene	5
Om kausalitet og effekt	5
Metode	6
Kriterier for inklusjon av studier i syntesen	6
Gjenbeskrivelse og vurdering av litteraturen	7
Organiserende temaer	8
Om Kunnskapsløftet	8
Syntese: Seks organiserende tema.....	9
TEMA 1: Læringsresultater i grunnskolen	9
TEMA 2: Overganger og gjennomføring i videregående skole	12
TEMA 3: Vurdering av og for læring	15
TEMA 4: Elevenes læringsmiljø	17
TEMA 5: Tilpasset opplæring.....	18
TEMA 6: Mangler i evalueringen av K06	22
Mot et system for kontinuerlig evaluering av kvalitet i opplæringen.....	24
Referanser	25

Utgangspunktet for syntesen

Kunnskapsløftet (KL) ble formelt innført fra skoleåret 2006/2007. Tidlig i innføringen ble det lyst ut flere oppdrag om følgeevalueringer av reformimplementeringen. Alle evalueringsprosjektene har fortløpende levert underveisrapporter, og sluttrapportene fra prosjektene er nå også ferdigstilt. Utdanningsdirektoratet har hatt et programstyre som har initiert og fulgt opp alle evalueringene, og de har engasjert fire forskningsmiljøer til å skrive tematiske synteserapporter.

Denne rapporten gir en syntese av temaområdet «Elevenes læringssituasjon etter innføringen av ny reform». I kravspesifikasjonen fra Utdanningsdirektoratet er tre problemstillinger formulert:

- Hvilke effekter har reformen hatt på elevenes læringsutbytte, gjennomføring og formell kompetanseoppnåelse?
- Hvordan håndterer elevene overganger mellom skoleslagene?
- Hvordan bidrar reformen til utvikling av en bedre kultur for læring og bedre tilpasset opplæring for alle?

Dette er spørsmål som implisitt antar at det er mulig å måle effekter av KL fem år etter at reformen ble innført. Spørsmålene identifiserer ikke hvilke reformtiltak som antas å skulle ha gitt effekt. Reformevaluering er i utgangspunktet vanskelig, og det er spesielt vanskelig når man skal evaluere en reform som er sammenvevd med en rekke tiltak som er satt i verk etter innføringen av den formelle reformen. Dette er en utfordring vi vender tilbake til.

De overordnede spørsmålene som denne syntesen skal besvare, ble formulert for spesielt å fange inn funn fra tre av de ti evalueringsstudiene som har fulgt implementeringen av KL:

- NOVA sitt prosjekt «Kunnskapsløftet – også et løft for utjevning av sosial ulikhet i læringsutbytte?» (Bakken 2009, 2010, Bakken og Danielsen 2011, Bakken og Elstad 2012)
- Nordlandsforsknings prosjekt «Sammenhengen mellom undervisning og læring. Arbeidsmåter, utvikling av ferdigheter og læring i norsk, naturfag og samfunnsfag (SMUL)» (Hodgson mfl. 2010a, 2010b, Hodgson mfl. 2012, Rønning mfl. 2008).
- NIFU sitt prosjekt «Tilbudsstruktur, gjennomføring og kompetanseoppnåelse i videregående opplæring» (Frøseth mfl. 2008, 2010, Høst og Evensen 2009, Vibe mfl. 2011, Vibe mfl. 2012)

Denne syntesen var ikke begrenset til å sammenfatte funn kun fra evalueringsstudiene. Vi fikk også i oppdrag å trekke inn funn fra andre relevante studier av norsk skole etter innføringen av KL.

Definisjoner og avgrensninger

For å kunne gjennomføre en syntese er det nødvendig å definere noen begreper som er sentrale i de overordnede syntespørsmålene. Dette bidrar både til å avgrense tilfanget av primærstudier og gjør det mulig å gjenbeskrive primærstudiene slik at vi kan ekstrahere et begrenset antall temaer som kan belyse syntesens problemstillinger.

Avgrensninger mot tre andre synteseområder

De fire synteserapportene har ulike problemstillinger som skal besvares, men det vil også være berøringspunkter mellom dem. Vi har gjort følgende avgrensninger mot de tre andre syntesene for å unngå for mye gjentakelser på tvers av rapportene:

- Læreplaner gir føringer for elevenes læringssituasjon, men dette temaet er ikke sentralt i vår rapport, siden dette ivaretas i en egen syntese (Sivesind 2012). Store deler av

Nordlandsforsknings evaluering handler om implementering av nye læreplaner, og er derfor ikke tatt med i vår synteserapport.

- Det er bestilt en egen synteserapport om yrkesfaglige studieprogram (Aarkrog og Bang 2012). En hovedkilde for den syntesen er NIFUs studie om gjennomstrømming og kompetanseoppnåelse i vgs. Denne studien er sentral også for vårt synteseområde. Vi vil derfor kun inkludere overordnede funn og vurderinger knyttet til videregående skole generelt.
- En «kultur for læring» berører også styring av sektoren. ILS/NIFU sitt prosjekt om reformen, omhandler Kunnskapsløftet som styringsreform, og det er bestilt en egen syntese om dette temaet (Nordenbo 2012). Vi har valgt å inkludere enkelte funn fra denne studien i vår syntese, særlig knyttet til temaet vurdering. Vi har bevisst forsøkt å holde oss til skole- og klasseromsnivået.

Definisjon av sentrale begreper i problemstillingene

Begrepene som er inkludert i de to første delproblemstillingene for vår syntese – «læringsutbytte», «gjennomføring», «formell kompetanseoppnåelse» og «overganger mellom skoleslag» – er ganske presise. Begreper som «læringsresultater», «faglige prestasjoner» og «faglig nivå» vil utover bli brukt synonymt med læringsutbytte. Begrepene «progresjon» og «gjennomstrømming» vil også bli brukt der vi drøfter gjennomføring, formell kompetanseoppnåelse og overganger mellom skoleslag. Ytterligere avgrensninger av disse begrepene vil ikke gjøres innledningsvis, men vil bli drøftet der det er relevant i rapporten. Vi har forsøkt å avgrense begrepene på en måte som er formålstjenlig for syntesen.

Begrepene «elevenes læringssituasjon» i den overordnede problemstillingen og begrepene «kultur for læring» og «tilpasset opplæring» i den tredje delproblemstillingen, er svært åpne og upresise begreper. Alle aspekter ved utdanningssektoren vil skape rammer for elevenes «læringssituasjon», og mange sider ved reformen er direkte eller indirekte en del av satsingen på å skape en «kultur for læring». Elevenes læringssituasjon har vi avgrenset til det noe snevrere begrepet «elevenes læringsmiljø». Vårt bidrag til å si noe om vi er på vei mot en kultur for læring, avgrenses til å oppsummere hva vi vet om vurdering av og for læring. Vi utelukker dermed en forståelse av kultur for læring på overordnet nivå, altså som læring i organisasjoner og på systemnivå. Begrepet «tilpasset opplæring» er i denne syntesen sterkt begrenset av de studiene som faktisk er gjort.

Om kausalitet og effekt

Første og tredje spørsmål som definerer denne syntesens oppdrag, er formulert som spørsmål som etterlyser effekter av Kunnskapsløftet. Dette er spørsmål som i streng forstand vil kreve forskningsdesign som egner seg til å belyse kausale sammenhenger. En foreløpig gjennomgang av studiene som vi skal sammenstille, viser at de bygger på forskningsdesign som ikke gir god støtte til kausale slutninger. To av evalueringsstudiene (NIFU og NOVA) har imidlertid svært gode longitudinelle data hvor de samme kohortene er fulgt over flere år, og denne typen design har størst potensial for å kunne underbygge forsiktige kausale slutninger. Utgangspunktet for arbeidet med vår syntese er derfor at vi ikke forventer å kunne besvare kausaliteten i spørsmålene. Likevel forventer vi å kunne gi gode beskrivelser av endringer over et aktuelt tidsrom for noen sentrale variabler knyttet til elevers læringsutbytte, gjennomføring, formell kompetanseoppnåelse, kultur for læring og tilpasset opplæring. «Effektene» og «bidragene» som henholdsvis den første og tredje

problemstillingen formulerer, vil vi derfor oppfatte som *registrerte endringer i det tidsrommet som de ulike studiene undersøker for de fenomenene som problemstillingene berører*.

Det er altså ikke mulig å slå fast med sikkerhet om effekter over tid skyldes reformen, eller om de skyldes andre endringer som har skjedd i det samme tidsrommet, som for eksempel demografiske endringer i de studerte populasjonene, endringer i lærernes kompetanse (som kan henge sammen med reformer i lærerutdanningen, satsing på etterutdanning etc.) eller andre samfunnsmessige endringer som kan bidra til å påvirke elevers innsats og motivasjon for læring. Vårt utgangspunkt er likevel at det gjennom syntesen vil være mulig å formulere resonnementer som gir grunnlag for å slutte at noen observerte effekter med rimelighet kan antas å være resultater av Kunnskapsløftet. Det er imidlertid verdt å minne om at Haug (2003) i sluttrapporten for evalueringen av Reform 97 også stilte spørsmål ved om alle observerte endringer i skolens praksis faktisk skyldtes implementering av reformen. Det er lite sannsynlig at en reform endrer en hel sektor over natta. De samme lærerne er der med de samme kunnskapene, verdiene og holdningene som de hadde dagen før, skolebyggene står der med de samme klasserommene og utstyret, skolegrensene er der fortsatt og reflekterer langsomme endringer i et sosialt segregert bosetningsmønster, og slik kunne vi fortsatt. Det er for øvrig godt kjent fra andre studier av reformer at endringer i utdanningssektoren tar lang tid (Fullan 2007, Tyack og Cuban 1995), og at det er vanskelig å måle effekt av dem, fordi de er bredt anlagt og forandres underveis i implementeringsprosessen (Sarason 1990).

Effekter kan også forstås som noe annet enn endringer i målbare størrelser. En effekt av K06 kan være at bestemte tiltak gjennomføres, at brukere rapporterer om opplevde endringer eller at praksis endrer seg, uten at man nødvendigvis kan observere endringer i elevenes læringsmiljø eller læringsresultater.

Metode

I de seneste årene har det blitt vanligere å utvikle såkalte kunnskapsoversikter, for eksempel i form av systematiske reviews, for å gi et helhetlig og oppdatert bilde av forskning om et tema. Det finnes flere spesifikke og strenge protokoller for hvordan man går fram for å skape slike systematiske oppsummeringer (se for eksempel Dansk Clearinghouse for Uddannelsesforskning 2007, Gough mfl. 2012, Kunnskapssenteret 2009). Det er ikke mulig med de primærstudiene som foreligger eller spørsmålene som er formulert for syntesen, å følge slike prosedyrer for systematiske kunnskapsoppsummeringer i denne syntesen. Den metodiske tilnærmingen er likevel inspirert av prinsippene for systematiske review. Vi oppfatter altså en syntese som noe mer enn en oppsummering i tradisjonell forstand. Det er en metodisk tilnærming som har som mål å gi mer overordnede og helhetlige beskrivelser av et tema/en problemstilling.

Kriterier for inklusjon av studier i syntesen

I tillegg til de tre evalueringsstudiene som vår syntese i utgangspunktet fikk et spesielt ansvar for å dekke, bygger syntesen på all tilgjengelig norsk forskning av relevans for det overordnede temaet. Med utgangspunkt i de spørsmålene som denne syntesen skal svare på, har vi søkt etter rapporter og andre publikasjoner som kunne tenkes å:

- ha resultater som er generaliserbare på et nasjonalt nivå
- ha data som muliggjør sammenlikning over tid i det aktuelle tidsrommet

- være knyttet til de fenomenene som nevnes i syntesens problemstilling

Dette er en type litteratur som vi ikke typisk forventer å finne i de vitenskapelige tidsskriftene, men heller i form av rapporter og evalueringer, som regel bestilt av direktorat eller departement. Vi søkte derfor i rapportoversiktene for de institusjonene som er hovedleverandører for slike produkter. Dessuten brukte vi Utdanningsdirektoratets egne sider som gir en god oversikt over den forskningen som de har bestilt i perioden etter innføringen av K06. Videre har vi inkludert resultater og tallmateriale som er publisert på ulike nettsider, eksempelvis Skoleporten, Gjennomføringsbarometeret og Statistisk sentralbyrå sine sider for utdanningssektoren.

En syntese er i sitt vesen en sekundæranalyse av resultater fra andre primærstudier, og inneholder vanligvis ikke resultater eller analyser som ikke tidligere er blitt publisert. Vi har avvekiet noe fra et slikt prinsipp ved å inkludere noen få nye framstillingsformer av resultater fra noen av de internasjonale undersøkelsene, og vi har også inkludert noen få resultater som ikke er publisert før.

Til sammen lyktes vi med å identifisere 73 arbeider som inneholder empiriske analyser av norsk skole, og som ut fra tematikk og utgivelsestidspunkt (etter K06) ble vurdert som relevante kilder for syntesen. Etter en første rask lesing ble imidlertid flere av disse utelatt fra syntesen fordi de ikke tilfredstilte ett eller flere av kriteriene for inklusjon nevnt ovenfor. Vi sto til slutt igjen med 50 publikasjoner som i noen grad har tilført informasjon til denne syntesen. Flere av disse studiene har imidlertid ikke informasjon om endring over tid, noe som er helt nødvendig for å kunne svare på problemstillingene som syntesen skal besvare. Det er som regel et begrenset utvalg av resultater i de inkluderte arbeidene som til slutt ble gjenbeskrevet og inkludert som materiale for syntesen. Dette må ikke forstås som en nedvurdering av det som utelates, det er kun arbeidenes relevans og kvalitet i relasjon til spørsmålene for syntesen som har blitt brukt som inklusjonskriterium. Det er flere analyser i evalueringsstudiene (og i andre studier) som gir interessant informasjon om forhold som synes å påvirke eksempelvis læringsutbytte og gjennomstrømming. Men for vår syntese har det kun vært relevant å inkludere disse dersom de har analyser som viser hvordan slike faktorer modereres over tid, eller dersom analysene på annet vis kan sannsynliggjøre at Kunnskapsløftet eller andre sentrale tiltak kan tenkes å påvirke forholdet.

Gjenbeskrivelse og vurdering av litteraturen

I den videre oppsummeringen av den relevante litteraturen valgte vi å bruke et enkelt skjema. Her beskrev vi først de overordnede trekkene ved de enkelte studiene sammen med hovedkonklusjonene. Der det var aktuelt, valgte vi å beskrive enkeltresultater som var relevante for syntesen. I tillegg inneholder gjenbeskrivelsene våre vurderinger av resultatene og de data og metoder som de er basert på. Spørsmål som lå til grunn for vurderingene var: Er utvalget som studien baserer seg på tilstrekkelig beskrevet, og er dette et representativt for en definert populasjon? Er metodene som er brukt tilstrekkelig beskrevet slik at studien er forståelig og transparent? I hvilken grad møter problemstillingen(e) som studien presiserer, en eller flere av de overordnede problemstillingene i kravspesifikasjonen? Er studiedesignet som er valgt, relevant for å finne svar på den/de problemstillingen/e som spesifiseres? I hvilken grad er det rimelig å konkludere med at eventuelle funn i studien kan knyttes til Kunnskapsløftet? I en systematisk review ville man typisk brukt skjemaer som hadde kartlagt svar på disse og andre spørsmål i detalj, men vi har her kun valgt å bruke disse som ledespørsmål for lesingen.

Organiserende temaer

Det finnes ulike tilnærminger til å sammenstille funn på tvers av ulike studier. I noen tilfeller er det mulig å utføre såkalte meta-analyser, som er en formalisert statistisk prosedyre for å estimere spesifiserte effekter på tvers av studier (se for eksempel Hattie 2009). Dette er ikke mulig i denne syntesen. Til det er antallet primærstudier for lite, og bredden i typen av effekter for stor. Vi ser derfor for oss en syntese av mer kvalitativ art, ikke i den strenge meta-etnografiske tradisjonen fra Noblit og Hare (1988), men som systematiserte, deskriptive gjenbeskrivelser som bidrar til å finne mønstre i materialet. Vi hadde på forhånd definert noen temaer som det, gitt problemstillingene, var opplagt at vi skulle inkludere i syntesen, f.eks. resultater om utviklingen i elevenes læringsresultater og læringsmiljø. Men underveis i gjennomlesingen og gjenbeskrivelsen av studiene har vi også hele tiden hatt til hensikt å identifisere andre overordnede temaer som gjør det mulig å drøfte funnene på tvers av studiene.

Om Kunnskapsløftet

Sammenlignet med R94 og L97 kan Kunnskapsløftet på mange måter oppfattes som en mindre gjennomgripende reform. Man kan eksempelvis oppfatte K06 hovedsakelig som en revisjon og justering av både struktur- og innholdsreformen som ble satt i gang med innføringen av R94 (NIFU 2012). Uansett hvordan man velger å se på dette, er det opplagt for oss at en oppfatning av Kunnskapsløftet som en reform som ble iverksatt på et bestemt tidspunkt, ikke er nyttig for å besvare syntesens problemstillinger. Arbeidet med denne syntesen har helt klart vist at K06 er sluttresultatet av prosesser som var satt i gang i god tid før skoleåret 2006/07, som var det formelle startpunktet for reformen.

Omtrent samtidig som M87 ble innført, kom en ekspertgruppe fra OECD med sin review av det norske utdanningssystemet. En sentral konklusjon var at beslutningstakerne i skolesektoren i vårt land ikke kunne vite hvorvidt tiltak virker, fordi det ikke var etablert et system som gjorde det mulig å si noe om tilstanden i skolen (OECD 1989). Dette ble fulgt opp i en rekke utredninger og stortingsmeldinger i tiåret som fulgte. I vår sammenheng er kanskje den viktigste av disse St. mld. nr. 37 (1990-91): *Om organisering og styring i utdanningssektoren*, som introduserte målstyring som styringsform i utdanningssektoren: "Målstyring krever at overordnet myndighet formulerer mål, formidler målene nedover i systemet og analyserer resultatene (...) Oppfølging og resultatvurdering av virksomheten i underliggende organer blir en viktig del av styringen (...) Større grad av målstyring krever et effektivt evaluerings- og tilbakemeldingssystem" (Lillejord 1997). Føringerne i denne meldingen ble fulgt opp gjennom blant annet St. mld. nr. 33 (1991-92): *Kunnskap og kyndighet*, St. mld. nr. 47 (1995-96): *Elevvurdering, skolebasert vurdering og nasjonalt vurderingssystem* (som la grunnlaget for at skolebasert vurdering ble forskriftsfestet) og St. mld. nr. 28 (1998-99): *Mot rikare mål*, som la vekt på at det var en "viktig statlig oppgave å medvirke aktivt til å framskaffe et bedre kunnskapsgrunnlag til bruk i et systematisk fornyings- og kvalitetsutviklingsarbeid i grunnskole og videregående opplæring". Det nasjonale systemet for kvalitetsvurdering i skolen, NKVS, ble foreslått i Kvalitetsutvalgets innstilling "Førsteklasses fra første klasse" (NOU 2002:10). Forslaget ble fulgt opp gjennom tillegg til St. prp. nr. 1 (2002-03) og vedtatt av Stortinget våren 2003. Og til sist ble trådene samlet i St. mld. nr. 30 (2003-04): *Kultur for læring*, og K06 var med det et faktum.

Med omleggingen fra regel- til målstyring fikk den enkelte utdanningsinstitusjon ansvar for lokalt å nå de sentralt formulerte målene, og det ble innført systemer som skulle brukes til å kontrollere lokal

måloppnåelse. Mens politikktutformingen på 1990-tallet i stor grad handlet om formulering av mål og utforming av læreplaner basert på mål, har perioden siden årtusenskiftet dreid seg om resultatmåling og resultatoppfølging. Nå dominerer interessen for hvordan pedagogiske prosesser som undervisning, vurdering og tilbakemeldinger kan bedre elevenes læringsutbytte. Vi ser med andre ord omrisset av et reformarbeid som har strukket seg over mer enn 20 år.

Alt dette gjør at store deler av K06 inngår som del av en større reformbevegelse. Den er et «bevegelig mål» og det er derfor ikke hensiktsmessig å tenke på den som en reform som skjedde på et bestemt tidspunkt. Det bidrar til større forståelse for reformen å se på endringer over tid uten å låse dette fast til et tidsskille før og etter 2006

Syntese: Seks organiserende tema

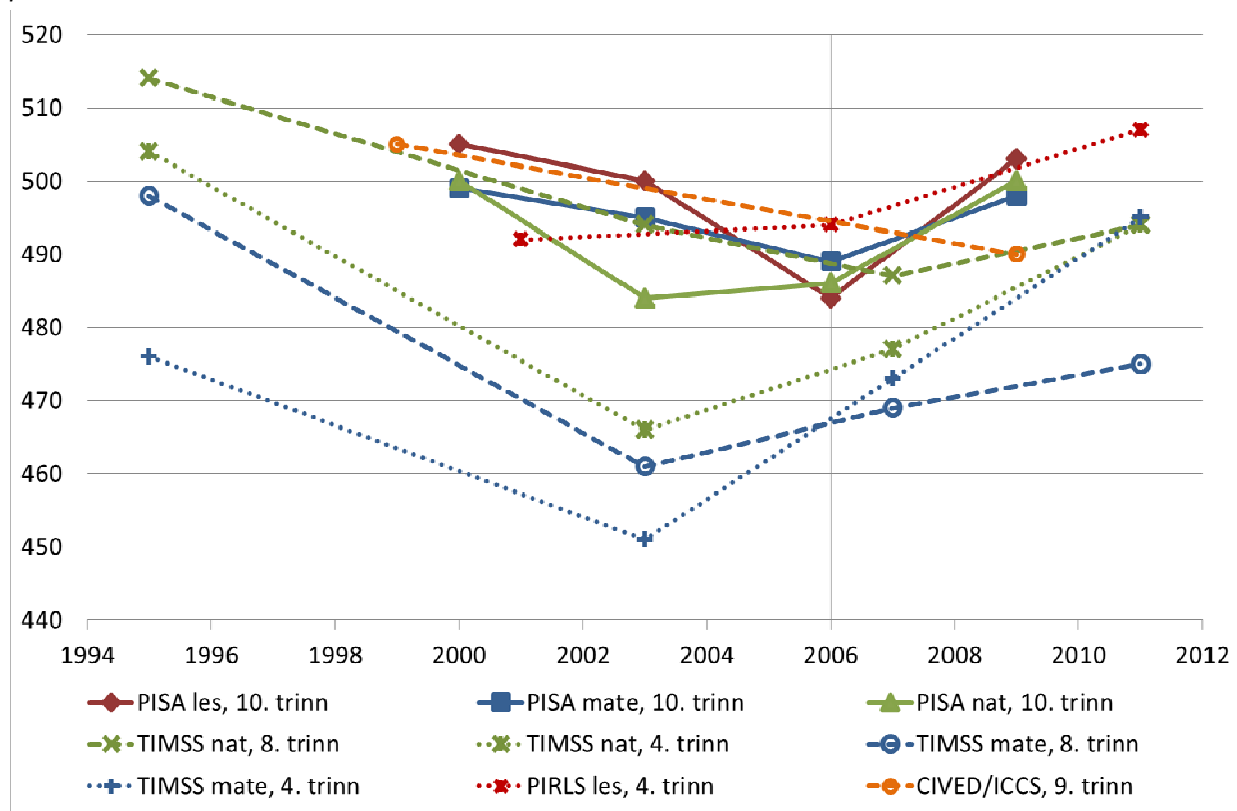
Syntesen er strukturert gjennom seks organiserende temaer. Noen springer direkte ut fra problemstillingene for syntesen, mens andre kom til underveis i arbeidet med å presisere noen av de mer vage begrepene og underveis i lesingen av litteraturen

TEMA 1: Læringsresultater i grunnskolen

De internasjonale undersøkelsene har design som gjør det mulig å studere utvikling over tid for elevenes læringsutbytte i noen sentrale fagområder. Grunnen er at studiene har prøver hvor mange av de samme oppgavene blir brukt i testene i påfølgende år. Denne synteserapporten gir ikke rom for å beskrive disse studiene, og vi viser derfor til de siste nasjonale rapportene som tallmaterialet i figur 1 er hentet fra (Grønmo mfl. 2012, Kjærnsli og Roe 2010, Mikkelsen mfl. 2011, Van Daal mfl. 2012).

Figur 1 på neste side gir en samlet framstilling av tidsserier fra disse studiene, men det er viktig å påpeke at figuren må leses med forsiktighet. Det er en tradisjon for at disse studiene rapporterer resultater på en skala der et internasjonalt gjennomsnitt er satt til på 500 og med et standardavvik på 100. Resultatene kan likevel ikke sammenliknes på tvers av studiene, fordi de måler ulike typer dyktigheter hos elevene, alder/klassetrinn varierer på tvers av studiene, og skalaene er forankret i ulike utvalg av land

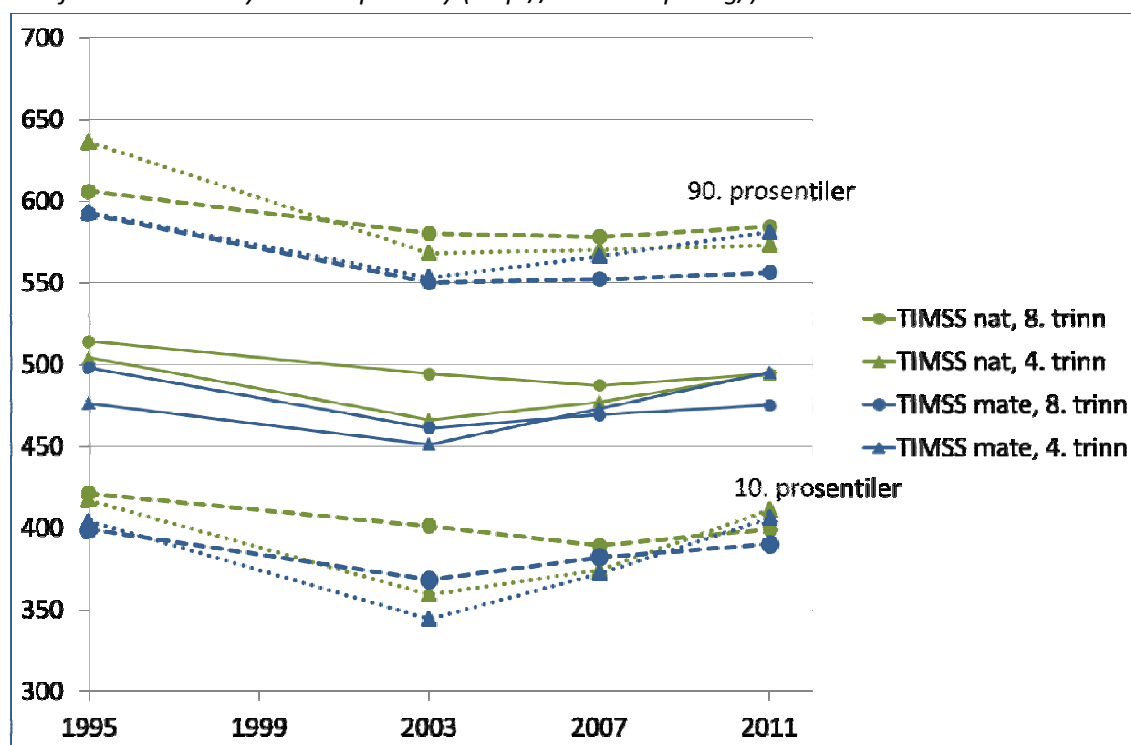
Figur 1: Resultater over tid for PISA, TIMSS, ICCS og PIRLS. Merk at resultater ikke kan sammenliknes på tvers av studiene.



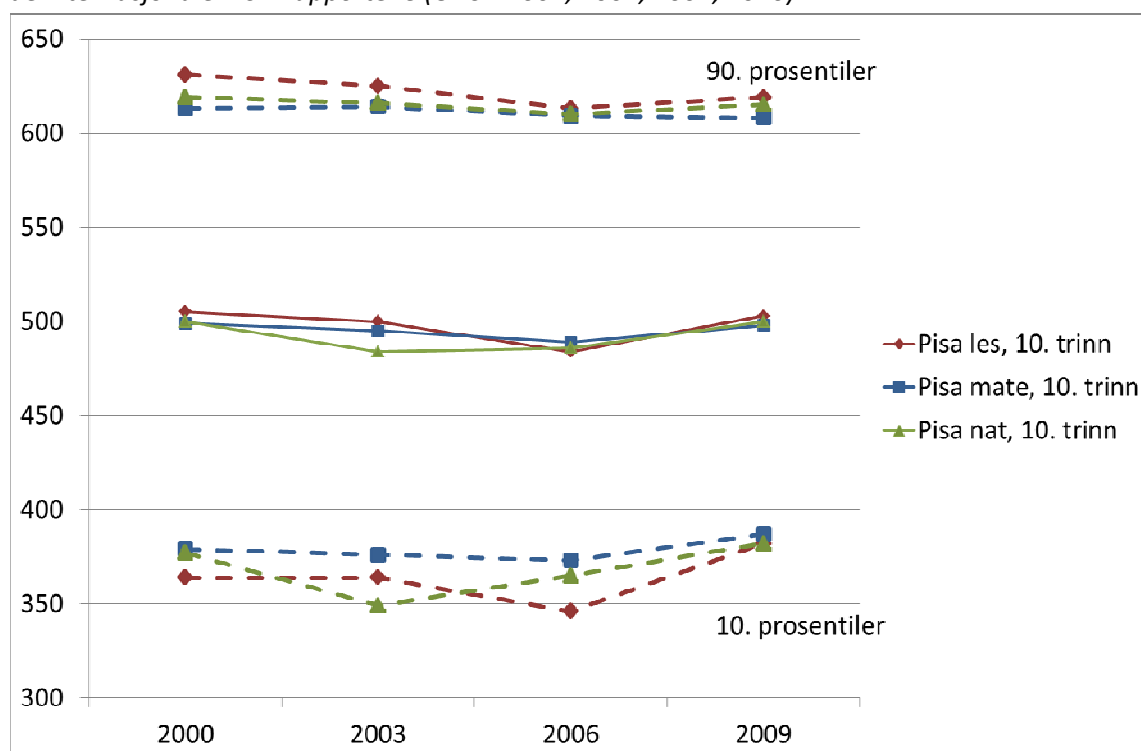
Når vi likevel har valgt å gi en samlet framstilling, er det fordi det gir oss mulighet til å få fram et interessant sammenfall i utviklingen av de norske elevenes prestasjoner som synes å være til dels uavhengige av fagområde og de ulike måtene disse fagområdene er definert på i de ulike studiene¹. En populær, men noe upresis og forenklet framstilling av disse skalaene, er at 30-40 poeng svarer til omtrent ett års skolegang. Fra de første målepunktene som vi har tilgjengelig tilbake til 1995 og fram til omtrent tidsrommet 2003-2006, ser vi en klar tendens til det som må kalles en betydelig tilbakegang. Vi snakker her om en nedgang som svarer til ett års skolegang, og det skjer til tross for at elevene som var med i de senere studiene har gått ett år lenger på skolen enn elevene som var med i 1995. At perioden før K06 var preget av betydelig tilbakegang var en bekymring som i stor grad motiverte flere av grepene i reformen, blant annet satsingen på grunnleggende ferdigheter. I tillegg ser vi en tydelig tendens til at denne utviklingen bremser opp og etterfølges av en like markant framgang i perioden som har gått siden K06 ble innført. Denne dataserien vil få et verdifullt tilskudd når PISA 2012 rapporteres i desember 2013.

¹ Et unntak her er utviklingen for CIVED/ICCS som er en undersøkelse relatert til samfunnsfaglige temaer, men denne trendlinjen er spesielt lite robust fordi den til dels er basert på svært få oppgaver. I tillegg er det ingen framgang å se i naturfag for 8. trinn.

Figur 2: Endringer i 10. og 90. prosentil, samt gjennomsnittet, for TIMSS-undersøkelsen. Data lastet ned fra the IEA Study Data Repository (<http://rms.iea-dpc.org/>)



Figur 3: Endringer i 10. og 90. prosentil, samt gjennomsnittet, for PISA-undersøkelsen. Data hentet fra de internasjonale PISA-rapportene (OECD 2001, 2004, 2007, 2010).



Figur 2 og 3 går litt bak utviklingen for TIMSS- og PISA-undersøkelsene. Dette er datapunkter som representerer det faglige nivået for henholdsvis en typisk svaktpresterende elev (10. prosentil) og en typisk faglig sterk elev (90. prosentil). Den store tilbakegangen fra 1995 til 2003 i TIMSS-undersøkelsen som vises i fig. 2, gjenspeiles både i toppen og bunnen av fordelingen. På 2000-tallet ser vi imidlertid at det er relativt små endringer i toppen av fordelingene for både PISA og TIMSS-undersøkelsene, med matematikk 4. trinn i TIMSS som et unntak. I motsatt ende av fordelingen ser vi at endringene er langt større, og de følger i stor grad endringene i gjennomsnittet. Særlig markant er utviklingen for de aller svakeste elevene i lesing i løpet av bare en treårs-periode fra 2006 til 2009 på omlag 40 poeng (fig. 3). Gjennomsnittene i PISA-undersøkelsen i 2009 er tilbake til omtrent samme nivå som de var i 2000, men det er tydelig at de svakeste elevene i lesing i samme periode har hatt en relativt stor og signifikant framgang. Vi ser også i TIMSS at den 10. prosentilen har endret seg betydelig i perioden 2003-2011. Selv om man tar høyde for at usikkerheten knyttet til endring over tid er størst i bunnen av fordelingen (blant annet vil dette påvirkes mer av eventuelle endringer i frafall fra studiene) er den eneste rimelige tolkningen av disse to figurene at norsk skole har hatt et betydelig løft fra bunnen. En oppsummering for dette temaet blir derfor:

- I perioden rundt innføringen av K06 viser de internasjonale undersøkelsene at nedgangen i faglige prestasjoner stopper opp. Vi ser deretter en positiv utvikling i det gjennomsnittlige prestasjonsnivået blant norske grunnskoleelever. Utviklingen er felles for de tre fagområdene naturfag, matematikk og lesing.
- Spredningen i de norske prestasjonene blir mindre, altså blir forskjellene mellom elevene mindre i perioden.
- Både framgangen i gjennomsnittet og den reduserte spredningen er i all hovedsak drevet av at de svakeste elevene har fått et betydelig løft - fremganger som i noen tilfeller svarer til effekten av omtrent ett år på. I denne perioden med framgang i gjennomsnittet, endrer imidlertid ikke prestasjonene seg for de sterkeste elevene.
- Resultatene fra de internasjonale undersøkelsene viser at i perioden rundt innføringen av K06, har det blitt gjort grep som slår ulikt ut for elever med ulike forutsetninger. Det er spesielt de faglig svakeste elevene som drar fordel av virkemidler som er gjennomført i perioden. Hvilke virkemidler det rent konkret er snakk om, kan vi imidlertid ikke si noe om.

TEMA 2: Overganger og gjennomføring i videregående skole

I denne delen beskriver vi utviklingen i gjennomføring av videregående skole (vgs). Vi inkluderer her kun noen overordnede trekk og går ikke inn på yrkesutdanningen da dette er tema for en egen synteserapport. Status for gjennomføring i videregående skole er velkjent, og vår beskrivelse av dette materialet tilfører derfor lite eller ingenting til det som allerede er godt kjent gjennom ulike publikasjoner og nettsteder. Den vanligste indikatoren for resultat kvalitet i vgs er andelen elever som fem år etter påbegynt utdanning har fullført vgs med formell kompetanse. I tillegg rapporteres andelen av elever med ordinær progresjon fra et årstrinn til det neste. Fordelen med disse indikatorene er at de kan gi et bilde av utvikling over et ganske langt tidsrom. Disse tallene er allment tilgjengelig gjennom både Skoleporten.no og Statistisk sentralbyrå sine nettsider, og dataene publiseres jevnlig både i Utdanningsspeilet og SSBs publikasjoner med nøkkeltall fra utdanningssektoren. I tillegg er det nå en tettere oppfølging og en større forpliktelse når det gjelder, å følge opp disse indikatorene gjennom Ny Giv som har satt klare resultatmål for perioden. Disse målene vil, sammen med flere nye indikatorer, kontinuerlig rapporteres i et eget

Gjennomføringsbarometer på Utdanningsdirektoratets hjemmesider (Kunnskapsdepartementet 2012).

Innledningsvis er det også grunn til å peke på at dersom vi ser for oss at K06 skal ha en effekt på gjennomføringen, så er dette indikatorer som i altfor liten grad allerede kan ha vært påvirket av virkemidler i reformen. Ett virkemiddel er for eksempel satsingen på grunnleggende ferdigheter i tidligere skoleår ut fra blant annet en antakelse om at dette vil gi en mer solid plattform for videre utdanning, og vi har allerede sett at de svakeste elevene har hatt en tydelig positiv utvikling på barne- og ungdomstrinnene. Man vil imidlertid måtte vente enda noen år før det kan måles effekter på gjennomføringsandelen av en slik satsing. Men tiltak som f.eks. strukturendringer som skal sikre at elever ikke velger seg bort i kryssløp med tap av tid, kan virke raskt.

NIFU har i sin evaluering «Tilbudsstruktur, gjennomføring og kompetanseoppnåelse i videregående opplæring» fulgt de kullene som begynte i vgs i 2004, 2005, 2006 og 2007, altså to kohorter før og to kohorter etter innføringen av K06² (Vibe mfl. 2012). Dersom de hadde brukt femårs-målet for gjennomføring, kunne de ikke så kort tid etter implementeringen av reformen ha gitt en før-etter beskrivelse av situasjonen. De har derfor i all hovedsak brukt fullføring på normert tid som utfallsvariabel i sine analyser. I tillegg har det kommet et nytt datapunkt (gjennomføring etter skoleåret 2010/11) etter at NIFU satte sluttstrek i sitt datasett. Det er også slik at de to hovedkildene for denne delen av syntesen, NIFUs studie og Gjennomføringsbarometeret, har ulike mål på det som begge kaller for ordinær progresjon.

Hovedbildet som avtegner seg er altså godt kjent: Fullføring av vgs til en formell kompetanse er veldig stabil over et relativt langt tidsrom. Andelen som fullfører på normert tid er 57 % for det første kullet etter K06³, og denne er svært lik for kullene i de to årene før K06 ble innført (Vibe mfl. 2012). Tilsvarende andeler har også blitt rapportert i en tidligere NIFU studie som fulgte kohortene som begynte i vgs i 1999-2001 (Støren mfl. 2007). Når det gjelder andelen som har fullført med kompetanse etter fem år, har den variert fra 68 % til 72 % for kohortene som begynte i vgs i perioden 1994-2006. Endringer i denne størrelsesorden er i seg selv store nok til å være av interesse, men når man studerer endringene over et lengre tidsrom, blir det tydelig at dette er fluktuasjoner rundt et gjennomsnitt på 70 % uten en tydelig trend i noen retning.

Når det gjelder overganger mellom ulike år i vgs, gir ikke de to hovedkildene et entydig bilde. Overgangen fra grunnskole til vgs ligger stabilt på 96-97 % i perioden 2007-2011. Andelen med ordinær progresjon fra vg1 har imidlertid økt med 3,6 prosentpoeng i perioden 2007-2011 (fra 80,9 % til 83,5 %). For begge disse to overgangene er NIFUs og Gjennomføringsbarometerets definisjoner sammenfallende. Utvikling er imidlertid ikke like entydig for overgangene fra vg2 og vg3, og her er ulike definisjoner lagt til grunn i de to indikatorsettene⁴. Ut fra de ulike definisjonene er det naturlig at NIFU rapporterer lavere andeler for progresjon fra vg2 og vg3, men det er vanskeligere å forstå hvorfor Gjennomføringsbarometeret rapporterer om til dels store fluktuasjoner i disse andelene, mens NIFU gir et bilde av stabilitet for de fem kohortene som de har inkludert.. Så store fluktuasjoner

² I tillegg har de kunnet inkludere flere årskull i oversikten for overganger mellom de lavere trinnene.

³ Dette er det eneste kullet som NIFU har kunnet følge gjennom et normert løp siden utdanningene som leder fram mot yrkeskompetanse er basert på en 2+2 modell.

⁴ NIFU følger kohorter fra de kommer inn, og beregner andelene av disse kohortene som er i et ordinært løp, mens gjennomføringsbarometeret regner andeler av de som begynner et år av de som avsluttet foregående skoleår på det lavere trinnet

virker urimelige og vi legger her til grunn også at Gjennomføringsbarometeret påpeker noe usikkerhet i datagrunnlaget for de tidlige årene i perioden. Den overordnede tendensen synes å være at selv om den totale gjennomføringen til slutt er stabil, så har det vært endringer over tid for *når* frafallet og forsinkelsene skjer. Frafall og forsinkelser i progresjonen synes å skje noe senere i løpet.

Et viktig utviklingstrekk som ligger delvis skjult bak disse overordnede gjennomførings- og progresjonsindikatorne, er det som NIFU betegner som en akademisering av vgs. Denne akademiseringen viser seg i flere steg. For det første viser søkemønstrene i tiden før og like etter innføringen av K06 en økning i søkningen til studieforberedende, spesielt blant jentene. Det er imidlertid litt forhastet å konkludere med at skjevheten mot valg av allmennfag har blitt større fordi reformen også medførte endringer i struktur med hensyn til hva som regnes som henholdsvis yrkesfaglige eller studiespesialiserende program. Videre er det en stor økning i andelen elever som i utgangspunktet starter på et yrkesfaglig tilbud og ender opp med studiespesialisering. Dette skyldes at én studieretning kan se ut til å være feilplassert i systemet, nemlig Medier og kommunikasjon. Her ender det overveiende flertallet av de som fullfører med studiespesialisering, selv om programmet er yrkesfaglig. Dessuten er det en stor økning i valg av Påbygging til studiespesialisering, spesielt i det nye Helsearbeiderfaget. Til sammen gjør dette at fire av fem som fullfører vgs etter K06 ender med studiespesialisering, mens det i utgangspunktet er en omtrent lik fordeling av elever som begynner i hvert av de to hovedløpene. Dette er ikke en ny utvikling med innføringen av K06, men en trend som startet etter R94. Øia (2011) finner for øvrig i sin studie knyttet til ungdomskoleelevers motivasjon, at det var flere i 2010 som ønsket seg allmennfaglig utdanning eller utdanning på universitet eller høyskoler enn i en tilsvarende studie i 2002. Spesielt gjelder dette for jentene.

Et spesifikt tiltak i Kunnskapsløftet er faget Utdanningsvalg. Dette er et tiltak hvor elever på ungdomstrinnet kan velge å følge fag i vgs eller de kan hospitere i en bedrift. Formålet er å sette ungdommene bedre i stand til å velge riktig videregående utdanningsløp. Evalueringen av dette tiltaket er foreløpig litt uklar fordi man ikke har gode tall på omfanget, og vi mangler studier som kan si noe om tiltakets effekt. Det er en økning i andelen rektorer som rapporterer at de har etablert en slik ordning (om lag 30 %), men dette er basert på en spørreundersøkelse blant skoleledere med usikkerhet knyttet til utvalgsprosedyrene (Borgen og Lødding 2009). Elever og lærere uttrykker usikkerhet om hvorvidt faget har noen effekt på elevenes utdanningsvalg (Lødding og Holen 2012).

I tillegg gir NIFU i sine evalueringsrapporter en grundig og robust multivariat analyse av en hel rekke faktorer som er relatert til elevers valg og frafall fra videregående opplæring. De samme faktorene behandles i en rekke rapporter fra Senter for økonomisk forskning og i statistikkrapporter hos SSB (Borgen og Lødding 2009, Falch mfl. 2010, Falch mfl. 2009, Falch og Nyhus 2010, 2011, Falch mfl. 2011, Frøseth mfl. 2008, 2010, Høst og Evensen 2009, Lødding og Borgen 2008, Nyen mfl. 2011, Statistisk sentralbyrå 2011, Vibe mfl. 2011, Vibe mfl. 2012). Analysene gir imidlertid ikke innblikk i hvordan disse sammenhengene eventuelt har endret seg med innføringen av K06

Som NIFU selv flere ganger påpeker, vil fluktuasjoner i arbeidsmarkedet ha en direkte påvirkning på fullføringsandeler i mange av yrkesfagene, og det er derfor ikke enkelt å slutte om eventuelle endringer som observeres er knyttet til endringer i skolen eller endringer i ytre forhold. Vi vil likevel anbefale at man om noen år gjentar noen av disse multivariate analysene, men at man endrer analysemodellene for eksplisitt å se på hvordan disse sammenhengene endrer seg over tid. Dette er også en type analyser som gjerne kan knyttes til den pågående evalueringen av Ny Giv. Det er for

tidlig å si noe om dette er et tiltak som virker, og den ene rapporten som foreløpig er publisert fra NOVAs følgeevaluering av tiltaket, har ikke et datagrunnlag som gjør det naturlig å trekke konklusjoner i noen retning (Sletten mfl. 2012).

Hovedkonklusjonene når det gjelder eventuelle endringer etter innføringen av KL er nokså tydelige:

- Det er god evidens for å hevde at andelen elever som gjennomfører videregående opplæring med fagbrev eller studiekompetanse, ikke har endret seg betydelig i perioden rundt innføringen av K06, men dette kunne man heller ikke forvente fordi de elevkullene som det har mulig å studere, ikke har vært eksponert for reformen i særlig grad.
- Det er god evidens for å konkludere at det er en betydelig økning i andelen elever som har normal progresjon fra vg1 til vg2, og dette er en utvikling som har vart over tid. Også Støren (2007) viste at denne andelen økte betydelig etter innføringen av R94. Bedre progresjon mellom vg1 og vg2 er en nødvendig, men tydeligvis ikke tilstrekkelig, forutsetning for økt gjennomføring.
- Det er også god evidens for å slå fast at andelen av de som fullfører og kommer ut med studiekompetanse, har økt, men dette er også en trend som var tydelig før innføringen av K06. På den ene siden gir dette flere unge muligheten til å ta høyere utdanning, men utviklingen er likevel bekymringsfull, og vi gjentar NIFUs egen oppsummering av dette utviklingstrekket: «Hvis man er bekymret for rekrutteringen av faglært arbeidskraft, så framstår dette som en direkte årelating av fag- og yrkesopplæringen» (Vibe mfl. 2012, s. 12).
- Et funn med svakere evidens er at det kan være grunn til optimisme for de neste årene. Vi har sett en framgang i læringsresultater i ungdomsskolen. Dette tolker vi som at kommende kull i vgs vil ha færre elever med svake grunnleggende ferdigheter i lesing og regning. Dette er et av virkemidlene som Kunnskapsløftet og den senere satsingen Ny Giv har framhevet som viktig for å få flere til å fullføre videregående opplæring, og det er rimelig å håpe at dette kan gi økt progresjon og fullføring i årene som kommer.

TEMA 3: Vurdering av og for læring

Et område som har fått stor oppmerksomhet i K06, er feltet vurdering. Programmet *Vurdering for læring* er en nasjonal satsing (2010-2014), hvor Utdanningsdirektoratet på oppdrag fra Kunnskapsdepartementet, har igangsatt skolering av ca. 170 skoleeier, lærere og skoleledere i alle landets fylker⁵. Satsingen er en oppfølging av prosjektet *Bedre Vurderingspraksis* som ble gjennomført ved 77 skoler i perioden 2007 – 2009. Det er verdt å merke seg at disse satsingene har rettet seg mot å styrke læreres kunnskaper om og ferdigheter i vurdering for læring, noen ganger også kalt formativ vurdering. Flere internasjonale kunnskapsoppsummeringer har dokumentert at gode vurderingspraksiser kan fremme elevenes læringsutbytte (Black og Wiliam 1998, Hattie 2009). Det finnes imidlertid få norske studier som har undersøkt sammenhengen mellom tiltak som vurdering for læring i klasserommet, og effekten tiltaket har på elevenes læringsutbytte. Rapportene som denne synteserapporten bygger på, har heller ikke forskningsdesign som kan gi oss svar på om vurderingspraksisen har påvirket elevenes læring. Det er også vanskelig å trekke noen konklusjon om hvorvidt, hvordan og i hvilken grad lærernes vurderingspraksis har endret seg i tiden etter innføringen av Kunnskapsløftet, med utgangspunkt i de foreliggende studiene.

⁵ <http://www.utdanningsdirektoratet.no/Vurdering-for-laring/VFL-skoler/>

Det vi kan si, er at kommuner selv mener at vurdering er satt på dagsorden og har fått gjennomslag som en sentral og viktig satsing lokalt. I Utdanningsdirektoratets spøringer til skoler og skoleeiere (Vibe 2012) ser en at 81 % av skolene oppfatter skoleeier som en pådriver i vurderingsarbeidet, en økning fra 72 % i samme spørsmål stilt i 2010 (Vibe 2010). Når det gjelder spørsmål knyttet til innhold og vurdering, er derimot forskerne noe overrasket over at de ikke finner større forskjeller mellom skoler som deltar i satsningen Vurdering for læring, og skoler som ikke deltar. Dette funnet sammenfaller med evalueringen av Bedre vurderingspraksis, hvor en heller ikke fant signifikante forskjeller mellom elevene ved deltakerskolene og kontrollskolene på spørsmål tilknyttet elevmedvirkning i vurdering (Hopfenbeck mfl. 2012, Throndsen mfl. 2009).

Dersom Kunnskapsløftet har ført til endringer innen feltet vurdering, finner vi disse først og fremst blant læreres holdninger til vurdering. I sluttrapporten fra Nordlandsforskning evalueringstudie (Hodgson mfl. 2012), kommer det fram at 7 av 10 lærere mener at reformen har ført til bedre vurderingspraksis. Samtidig rapporteres det fra lærere at K06 har medført større oppmerksomhet om testing, og at satsingen på vurdering kan ha sine negative sider. Noen er bekymret for at økte krav om dokumentasjon kan føre til at det blir for mye vurdering for vurderingens skyld. Forskerne advarer imidlertid mot å tolke intervjudata for langt, grunnet skjevt utvalg.

I en delrapport fra 2010 har Nordlandsforskning rapportert om intervjuer med 59 lærere våren 2008 og 2009, der temaet vurdering var det sentrale (Hodgson mfl. 2010b). Fra lærerintervjuene kommer også de klareste meldingene om endret praksis for underveisvurderingen. Lærerne mener at de selv har blitt flinkere til å vise elevene hva som skal vurderes, de bruker ny terminologi som "framovermeldinger" og de er mer konkrete i sine tilbakemeldinger. Forskerne selv oppsummerer med at over halvparten av lærerne uttrykker en forståelse for vurdering som er i tråd med den nye vurderingsforskriften. 75 % av lærerne rapporterer om en endret vurderingspraksis, og det er barnetrinnets lærere som i størst grad viser til praktiske eksempler. Lærere på de høyere trinnene rapporterer i mindre grad om økt bruk av underveisvurdering i det daglige (ibid 2010).

Det kan også være grunn til å nevne at sluttrapporten fra evalueringstudien om K06 som styringsform i samarbeid mellom UiO og NIFU (FIRE-prosjektet) hevder det har blitt større interesse for resultater, altså summativ vurdering, etter K06 (Aasen mfl. 2012). Dette skyldes særlig innføringen av det nasjonale systemet for kvalitetsvurdering (NKVS) som har bidratt til at rektorene har utviklet nye rutiner for resultatoppfølging. Endringer i skolens planer og satsingsområder ser særlig ut til å gjelde grunnskolen. Samtidig påpeker forskerne at det er store forskjeller mellom skoler når det gjelder hvordan informasjon om resultater brukes i skoleutvikling eller på klassenivå. Dette knytter seg særlig til i hvilken grad man har kapasitet og kompetanse til å nyttiggjøre seg resultatene i skolens kvalitetsarbeid (Aasen mfl. 2012).

Disse resultatene kan for øvrig også leses i lys av trendmålingene fra PISA. Her blir rektorer blant annet spurt om elevenes eksamensresultater og resultater fra nasjonale prøver offentliggjøres og brukes i evalueringen av rektors og læreres arbeid. Andelen rektorer som svarer «Ja» på disse spørsmålene har økt med mellom 10 og 15 prosentpoeng fra 2006 til 2009 (Hopfenbeck mfl. 2010). Til sammen gir dette grunnlag for å hevde at det har skjedd en endring i skolens praksis med å bruke elevresultater for ulike formål.

Når det gjelder summativ vurdering i form av karaktersetting, finnes det lite informasjon. Opplæringsloven og Vurderingsforskriften endret seg i perioden, og krav om hva som kunne eller ikke

kunne inngå i grunnlaget for sluttvurdering, ble tydeliggjort. Det eneste lille kunnskapsgrunnlaget vi har om eventuelle endringer i måten karakterer settes på, finner vi i NOVAs evalueringsstudie som rapporterer om svakt økende gjennomsnittskarakter i alle fag på ungdomstrinnet gjennom 2000-tallet, og dette knytter de spesielt til en økt bruk av karakteren seks, både til standpunkt og eksamen (Bakken og Elstad 2012). Det er flere mulige årsaker som kan ligge bak en slik observasjon: Det kan ha blitt flere flinke elever i perioden, det kan ha blitt innført andre kriterier for toppkarakteren, slik at flere «rettmessig» fortjener en slik vurdering, eller lærerne kan ha blitt snillere i perioden. Vi kan ikke skille mellom disse mulige forklaringene, men det er relevant å minne om at de internasjonale undersøkelsene viser at prestasjonene til de 10 % flinkeste elevene har vært stabil i perioden.

Ingen av de studiene vi har vist til ovenfor, representerer noe overbevisende kunnskapsgrunnlag for å generalisere, men sett i sammenheng framstår likevel noen funn med middels grad av evidens:

- Alt i alt viser den tilgjengelige forskningen om vurdering i norsk skole at perioden etter innføringen av K06 er preget av økt oppmerksomhet om vurdering, og at det kan se ut som om satsingen knyttet til underveisvurdering eller formativ vurdering er godt mottatt.
- Når det gjelder en nærmere beskrivelse av hvordan dette konkret praktiseres, og hvilket omfang vurdering for læring har i sektoren, må vi avvente evalueringene av prosjektet Vurdering for læring.
- Data fra flere kilder rapporterer om økt bruk av elevenes læringsresultater for ulike formål, og det ser ut til å være en generelt forsterket positiv holdning til de ulike prøvene som inngår i NKVS, spesielt blant skoleledere og rektorer, og særlig på barnetrinnet.

TEMA 4: Elevenes læringsmiljø

Med læringsmiljø kan man inkludere elevenes subjektive opplevelser knyttet til læringssituasjonen, og forhold som ligger utenfor eleven selv. Forhold utenfor eleven kan for eksempel være læreres generelle holdninger til læring, deres forventninger om elevenes læring, eksterne faktorer som for eksempel hvordan samfunnet generelt verdsetter skole og læring, det fysiske skolemiljøet og læreplaner (Skaalvik og Skaalvik 2009). Vi vil her hovedsakelig forholde oss til de elevopplevde sidene ved læringsmiljø som er best dekt i primærstudiene. Utover det som behandles her, er ytre forhold ved elevenes læringsmiljø knyttet til beskrivelser av undervisningen, vurderingspraksis og tilpasset opplæring omhandlet i egne organiserende temaer i syntesen.

Gjennom Elevundersøkelsen kan man studere utviklingen fra 2007 og framover for en del sider ved elevenes læringsmiljø. Videre er det spørsmål i spørreskjemaene til elever og skoleledere i PISA-undersøkelsen som kan belyse utviklingstrekk fra 2000 til 2009. Det er også gjennomført en såkalt «Lærerundersøkelse» og en «Foreldreundersøkelse», men dette er data som foreløpig kun rapporteres på lokalt nivå. Det kan nevnes at en Lærings- og Instruktørundersøkelse også er på trappene, men den første rapporteringen fra denne er mer å betrakte som en mulighetsstudie og beskrivelse av den første implementeringsfasen (Nyen mfl. 2011).

Elevundersøkelsen gjennomføres hvert år, og er obligatorisk for elever på 7., 10. trinn og vg1. Ved å kombinere enkeltspørsmålene fra undersøkelsen i samlemål er det utviklet flere indikatorer knyttet til elevenes oppfatninger og erfaringer fra skolen. Det er noe fallende svarprosent med alder, slik at undersøkelsene for 7. trinn framstår som mest representative, mens det knytter seg noe mer usikkerhet til tallene for vg1. Generelt viser både Elevundersøkelsen og PISA at elevene har et

positivt forhold til sin skolegang. Når det gjelder utvikling i elevenes læringsmiljø i tiden etter innføring av K06, er det overordnede budskapet fra Elevundersøkelsen at endringene er så små at de ikke gir grunnlag for å slutte at det er en utvikling i en bestemt retning siden 2007 (Wendelborg mfl. 2011).

I PISA-undersøkelsene var det både i 2000 og i 2009 et sett med spørsmål som fanget inn lærer-elev-relasjonen. Utviklingstrenden for disse enkeltspørsmålene er ikke tidligere publisert i de nasjonale rapportene. Tre av spørsmålene er uendret i perioden, og dreier seg henholdsvis om lærerne hjelper når man har behov for det, om elevene føler seg rettferdig behandlet av lærerne sine og om lærerne er lyttende. Elevene oppfatter situasjonen i 2009 som mer eller mindre den samme som i 2000. Det er også noen spørsmål til skolelederne i PISA som er relevante for dette temaet. Gjennom et sett med spørsmål har de blitt bedt om å vurdere i hvilken grad en rekke forhold som kan kalles problematisk elevatferd bidrar til å hemme elevenes læring. Disse viser en overordnet positiv utvikling, men forskjellene over tid er likevel for små til at man med sikkerhet kan slå fast at det har vært en endring. Øia (2011) har tidsserier fra 1992, 2002 og 2010 for liknende forhold og konkluderer blant annet med at det er betydelig færre skolekonflikter i 2010 sammenliknet med 2002. I sin studie finner han også at det er færre som skulker skolen.

Avslutningsvis er det verdt å nevne at de begrepene som måles i Elevundersøkelsen er positivt relatert til læringsresultater, noe som for så vidt ikke er overraskende, men enda viktigere er at Wiborg mfl. (2011) har vist at elevenes læringsmiljø, slik dette måles i Elevundersøkelsen, gir signifikante og unike bidrag i multivariate modeller som sikter mot å forklare elevens og skolers prestasjonsutvikling. Elevundersøkelsen gir dermed tilgang til kvalitetsindikatorer som synes å være viktige komponenter i en læringskultur.

Både Elevundersøkelsen og spørsmålene til skoleledere og elever i PISA, er typiske eksempler på såkalte brukerundersøkelser hvor aktører blir bedt om å beskrive sine opplevelser ved å ta stilling til en rekke beskrivende utsagn. Det er vanskelig å tolke resultater fra slike undersøkelser fordi opplevelser ikke kun reflekterer den faktiske opplevde virkeligheten, men også er influert av forventninger som man har til de ulike forholdene som det spørres om. Dersom disse målene endrer seg over tid, er det derfor ikke mulig å slå fast om det er det objektive fenomenet eller respondentenes forventninger som har endret seg. Vi velger derfor å oppsummere dette temaet slik:

- Det finnes middels evidens for å hevde at elevenes læringsmiljø, slik dette har blitt fanget inn av PISA og Elevundersøkelsen, ikke synes å ha endret seg med K06.
- Det er noe svakere evidens for å hevde at spesielt problematisk adferd forekommer i mindre grad enn tidligere.

TEMA 5: Tilpasset opplæring

Tilpasset opplæring har en snart førti år lang historie i norsk skolepolitikk, men defineres forskjellig av ulike regjeringer (Jenssen og Lillejord 2009). Vi har derfor valgt å oppfatte begrepet «tilpasset opplæring» ganske vidt i vår problemstilling brukes begrepet «tilpasset opplæring *for alle*», og vi har derfor valgt å ikke inkludere spesialundervisning i vår syntese. Det er imidlertid god dekning for å si at det har vært en markant økning av spesialundervisning i perioden fra 2006. Analyser av kommunenes tilstandsrapporter indikerer at i mange kommuner kan tidlig innsats ha blitt det samme som spesialundervisning (Rambøll Management 2012, s. 36). Fordelingen av ressursene på de ulike

årstrinnene og effektene på læringsresultater av spesialundervisningen, er heller ikke entydige (Bonesrønning mfl. 2010, 2011).

Tabell 2 Prosentandel rektorer som rapporterer om ulike grupperinger av elevene i alle fag eller i noen fag i PISA 2006 og 2009

I noen eller alle fag	2006		2009	
	I alle fag	I noen fag	Alle fag	I noen fag
Elevene gruppert etter evner/forutsetninger på tvers av trinn	0	9	1	24
Elevene er gruppert etter evner/forutsetninger innenfor samme trinn	4	36	5	64

Tilpasset opplæring blir gjerne forbundet med undervisning som tar hensyn til den enkelte elevens forutsetninger og interesser. Ingen av evalueringsstudiene sier noe om hvorvidt eller i hvilken grad dette kjennetegner skolenes praksis. Vi tyr derfor nok en gang til PISA-undersøkelsen der skolelederne i 2006 (like før K06 ble innført) og i 2009 fikk følgende spørsmål: "Noen skoler organiserer undervisningen forskjellig for elever med ulike forutsetninger. Hvilken løsning velger eventuelt din skole for elevene på 10. trinn?" I svaret skulle det oppgis hvorvidt dette skjedde på tvers av og innen samme trinn. Svaralternativene var: "I alle fag", "I noen fag" og "Ikke i noen fag". Tabell 2 viser hvor stor prosentandel av rektorene som svarte "I alle fag" og "I noen fag" i 2006 og 2009.

Det er svært få rektorer som oppgir at elevene deles inn etter ulike forutsetninger i alle fag, uavhengig av om det er innen eller på tvers av trinn, og her har det ikke skjedd noen endring fra 2006 til 2009. Når det gjelder enkelte fag, ser imidlertid undervisning i nivådelte grupper ut til å forekomme i økende grad. I 2006 oppga 40 prosent av rektorene at nivådeling forekom ved deres skole på samme trinn, mens i 2009 er denne andelen økt til 69 prosent. Vi kan ikke fra dette gi et anslag av omfanget av slik organisatorisk differensiering, men tallene viser at det utvilsomt har vært en tiltakende tendens til å velge en eller annen variant av elevgruppering for å møte kravet om tilpasset opplæring. Dette bildet i 2009 fra PISA-undersøkelsen, bekreftes også av den årlige surveyen «Spørsmål til Skole-Norge» som ble gjennomført høsten 2010, og der 62 % av skolene (på tvers av skoleslag) bekrefter at de praktiserer en eller annen form for nivådifferensiering i tråd med det Opplæringsloven åpner for (Vibe 2010).

En annen side ved tilpasset opplæring er skolens bidrag for utjevning på tvers av kjønn, innvandringsstatus og hjemmets sosioøkonomiske status (SØS)⁶. Det er godt kjent fra flere studier at det er en sterk sammenheng mellom SØS og skolefaglige prestasjoner. Denne sammenhengen er imidlertid relativt sett svakere i Norge enn i land det er naturlig å sammenlikne seg med (Olsen og Turmo 2010). Videre er det godt kjent at elever som er født i et annet land, heretter kalt minoritets elever, presterer betydelig svakere enn elever som er født i Norge, men at disse ulikhetene er ubetydelige når man kontrollerer for foreldrenes utdanningsnivå (se for eksempel Birkeland mfl.

⁶ Sosioøkonomisk status er et svært sammensatt begrep, og det finnes mange ulike definisjoner og operasjonaliseringer. Størrelsen på slike effekter kan derfor vanskelig sammenliknes på tvers av studier.

2010). Vi vet også at jenter presterer bedre enn gutter i nesten alle skolefag (se for eksempel Utdanningsdirektoratet 2012 for en kortfattet oversikt). Dette overordnede bildet bekreftes også i rapportene fra NOVAs evalueringsstudie «Kunnskapsløftet – også et løft for utjevning av sosial ulikhet i læringsutbytte» (Bakken 2009, 2010, Bakken og Danielsen 2011, Bakken og Elstad 2012). Men det som gjør NOVAs studie spesielt relevant for vår syntese, er at den gjør flere analytiske grep for å undersøke om innføringen av K06 ser ut til å ha moderert disse effektene.

Den første tilnærmingen i NOVAs studie, er å beskrive hvordan sammenhengen mellom elevenes hjemmebakgrunn (her er registerdata om foreldrenes utdanning og inntekt brukt) og karakterer (både standpunkt og avgangskarakterer) har endret seg for avgangselevne i ungdomsskolen i perioden 2002-2010. Denne effekten uttrykkes i det følgende som SØS-gradienten. I korte trekk viser analysene at endringene i de ulikhetene som studeres er små, men der det er endringer, er det i de fleste tilfeller mot større forskjeller. Dette gjelder spesielt forskjellene mellom elever fra ulike sosioøkonomiske hjemmeforhold hvor man ser et tydelig knekkpunkt der SØS-gradienten, som for så vidt er jevnt og svakt stigende i hele perioden, gjør et ekstra lite byks opp for det første kullet i K06. Det samme gjelder for kjønnsforskjellene: Før K06 er de faktisk litt på vei nedover, men i tiden etter implementeringen av reformen er forskjellene svakt stigende. Når det gjelder elever med minoritetsstatus, er bildet litt mer broket: Totalt sett øker også denne ulikheten marginalt etter K06, men for noen grupper av minoritetselever er forskjellen minkende. Nærmere undersøkelser av endringen i SØS-gradienten viser at effekten er noe sterkere for jenter og at familiens inntekt har en økende betydning. I studien har NOVA i tillegg kontrollert for en rekke alternative strukturelle effekter, men dette påvirker ikke konklusjonen om at SØS-gradienten har økt.

Det er imidlertid andre forhold som også kan bidra til å forklare økningene i ulikheter. NOVAs studie viser til at andelen elever som får karakteren seks har økt i den samme perioden. De sier selv at dette er en av hovedkildene til at forskjellene har økt. Denne effekten forsterkes også av endret praksis i motsatt ende av karakterskalaen ved at andelen med manglende karakterer i fag øker etter innføringen av K06 at (2012, s. 63-68). En alternativ hypotese er derfor at læreplanene i K06, og presiseringer i vurderingsforskriften, har ført til tydeligere vurderingskriterier. Dette kan bety at de latente sammenhengene egentlig ikke har endret seg, men at reliabiliteten i karaktermålet har økt, og at sammenhengen dermed også framstår som sterkere. Denne alternative forklaringen er ikke undersøkt eksplisitt, men forfatterne berører den selv i sin rapport (Bakken og Elstad 2012, s. 246-7). Det finnes også andre funn i denne studien som er konsistente med en slik hypotese. Det er kjent fra før (se f.eks. Galloway mfl. 2011), og bekreftes av Bakken og Elstad (2012, s 223) at standpunktkarakterer har en tendens til å bli normert i forhold til gjennomsnittsnivået på skoler, med andre ord at elever som går på skoler med mange flinke elever får relativt svakere standpunktkarakterer enn de ville fått om de hadde gått på en skole med relativt svakere elever. Bakken og Elstad (ibid) viser at denne effekten av skole på standpunktkarakterer svekkes i perioden, noe som kan tolkes som en utvikling mot større grad av konsistens i lærernes karaktersetting. Dette funnet bidrar til å forsterke en hypotese om at kriteriene for karaktersetting har blitt mer entydige, noe som fører til mer reliabel karaktersetting på tvers av lærer og skole. Dette vil i seg selv kunne føre til at SØS-gradienten øker, selv om den underliggende sammenhengen skulle være uendret.

I motsatt retning er det en tendens til at skolens gjennomsnittlige SØS har en økende betydning for elevenes prestasjoner til eksamen, altså har det for eksamen mer å si hvilken skole du går på etter innføringen av K06.

Videre gjøres det to interessante grep for å nærme seg det kausale spørsmålet om K06 har hatt en effekt på disse sammenhengene. I den første tilnærmingen antas det at skoler har implementert reformens virkemidler i ulik hastighet – noen skoler var allerede før innføringen av K06 i gang med tiltak som var i tråd med reformens intensjoner, mens andre skoler hang litt etter. Dersom virkemidlene i K06 kan bidra til å dempe eller øke ulikhetene i karakterprestasjoner mellom ulike elevgrupper, bør skoler som i større grad har implementert virkemidlene, ha en nedgang i ulikhetene over tid. Dette kaller Bakken og Elstad (2012) for *implementeringshypotesen*. Vi viser til deres rapport for detaljer i hvordan denne hypotesen har blitt undersøkt. Resultatet av analysen er at graden av implementering, slik dette blir målt i noen av de indikatorene som utvikles, har en effekt på SØS-gradienten, men de påvirker ikke endringer over tid for de rapporterte sammenhengene i nevneverdig grad. Det er ikke rimelig å tro at indikatorene for implementeringshastighet er svært reliable mål, bl.a. er det lite variasjon mellom skolene for disse målene, og det er derfor ikke overraskende at analysen konkluderer med at implementeringshypotesen ikke kan bekreftes.

Den andre hypotesen som undersøkes er den såkalte *kunnskapshypotesen* som tar utgangspunkt i at K06 sin kjerneambisjon er å øke kunnskapsnivået hos elevene. Dersom dette også er en ambisjon som er forenlig med utjevning, bør skoler som viser en bedring i elevresultater i perioden (kontrollert for ulikhet i skolens elevsammensetning), også ha en utvikling der SØS-gradienten blir mindre i samme periode. Denne hypotesen antar altså at gode skoler er gode på kompensatorisk vis, dvs at de svake elevene vil ha en relativt større fordel av å gå på en god skole enn de relativt sterkere elevene. Analysene som presenteres, er konsistente med en slik hypotese: For de skolene som har hatt en slik positiv utvikling i elevenes justerte læringsresultater i perioden, er SØS-gradienten minkende, mens for de skolene som har en motsatt utvikling i justerte læringsresultater, er SØS-gradienten økende. En liknende effekt ser man imidlertid ikke for forskjellen mellom minoritets- og majoritets elever eller mellom kjønn. Men nok en gang: Dette er svært små interaksjoner for allerede små effekter. Likevel viser denne analysen at ambisjonen i reformen om å løfte kunnskapsnivået ikke står i motstrid til en eventuell målsetting om å redusere ulikheter mellom ulike elevgrupper.

Disse endringene over tid kan også studeres med data fra PISA-undersøkelsene, som viser en svak tendens til at effekten av elevenes hjemmebakgrunn øker over det samme tidsrommet, men endringene er for små til å framstå som statistisk signifikante (Olsen og Turmo 2010). Dette står ikke nødvendigvis i motstrid til funnene i NOVA sin studie. Sosioøkonomisk status er definert på en annen måte i PISA enn i NOVA-studien. Målet i PISA er basert på relativt usikre elevrapporterte mål, og det er en utvalgsundersøkelse, slik at effekter i den størrelsesorden som NOVA rapporterer, ikke framstår som signifikante. Når det gjelder de såkalte kjønns- og minoritetsgapene viser de internasjonale undersøkelsene at kjønnsgapet primært økte i jentenes favør før innføringen av K06, og fra PISA ser vi at minoritetsgapet ikke har endret seg vesentlig i den perioden undersøkelsen har blitt gjennomført (Grønmo og Onstad 2009, Kjærnsli og Roe 2010).

Det som fortsatt er litt vanskelig å forklare, er at vi i Tema 1 har dokumentert at de gjennomsnittlige prestasjonene for norske elever i den samme perioden har bedret seg og spredningen har blitt mindre, og i all hovedsak er begge disse trendene knyttet til at de svakeste elevene presterer bedre. Dette er ikke lett å forene med funnet om en forsterking av SØS-gradienten. Man ville snarere tro at når forskjellene i elevprestasjoner minker fordi de svakeste elevene forbedrer seg, så vil også SØS-gradienten svekkes.

Det er også relevant for temaet tilpasset opplæring å nevne at leksehjelp ble innført som et obligatorisk tilbud for 1.-4. trinn skoleåret 2010/11. Dette tiltaket skal evalueres av NOVA, men evalueringen har så vidt kommet i gang, og det er for tidlig å trekke noen konklusjoner om hvordan tiltaket er gjennomført i skolene, og hvordan det eventuelt kan bidra til å gi en mer tilpasset opplæring som også kan ha en utjevnende effekt, men det er klart fra beskrivelsene av tiltaket at nettopp dette er et av hovedformålene (Seeberg mfl. 2012).

For å oppsummere dette temaet:

- Et hovedfunn er at det finnes middels evidens for å hevde at tilpasset opplæring etter K06 i større grad enn før skjer gjennom å organisere elever i ulike grupper etter forutsetninger.
- Vi har ikke vært i stand til å identifisere annen kunnskap om hvordan skoler og enkeltlærere møter utfordringen med å tilpasse opplæringen for ulike elevgrupper.
- Et annet hovedfunn er at det finnes isolert god evidens for å hevde at effekten av elevenes kjønn, foreldrenes utdanning og inntekt, og elevenes minoritetsstatus på læringsresultater (dvs karakterer i 10. trinn) har økt marginalt etter innføringen av K06.
- Det er utfordrende å forstå hvordan den norske framgangen i de internasjonale undersøkelsene, som kan knyttes til et løft fra bunnen av fordelingen, er forenlig med en forsterking av SØS-gradienten.
- Det er middels evidens for at skoler med faglig framgang i større grad lykkes med å jevne ut forskjeller i læringsresultater for elever med ulik hjemmebakgrunn. Det er derfor mulig å bedre de faglige prestasjonene samtidig som forskjeller reduseres.
- Det er ellers verdt å merke seg at det ikke finnes studier som kan si noe utviklingen for barnetrinnet eller videregående skole.

TEMA 6: Mangler i evalueringen av K06

I lesingen av studier som berører de spørsmålene vi har fått i oppdrag å inkludere i vår syntese, er det noen påfallende mangler. Vi vil her se på fire temaer som vi mener burde ha vært inkludert i evalueringen: Satsingen på realfag, satsingen på grunnleggende ferdigheter, kvaliteten i videregående opplæring og kunnskap om hvordan bredden av tiltak i reformen har påvirket det som skjer i klasserommet.

Satsingen på realfag: Realfagsatsingen er en del av Soria Moria-erklæringen og nedfelt i et eget strategidokument (Kunnskapsdepartementet 2010). En tidligere evaluering av strategien viste at de fleste av målene ikke var nådd (Rambøll Management 2007). Målene er grovt sett fortsatt de samme og mange av tiltakene fortsetter. Med K06 kom nye og gjennomgripende tiltak, og i den siste versjonen av strategiplanen er det nå sammenfall mellom reformtiltak og realfagsatsingen. I tillegg til at det kom nye læreplaner også i realfagene, ble matematikkfaget styrket gjennom økte timetall og ekstra ressurser for å øke lærertettheten i faget på 1.-4. trinn. Men den mest markante endringen var innføringen av et nytt obligatorisk fag på vg2 for elever som velger studiespesialisering. I tillegg er matematikkfaget forsterket gjennom etableringen av regning som grunnleggende ferdighet og innføring av tilhørende nasjonale prøver i regning og kartleggingsprøver på lavere årstrinn i barneskolen og i oppstarten av vg1. Disse tiltakene er først og fremst rettet mot alle elevene, uavhengig av hvilken spesialisering de ender opp med til slutt. Effekten av tiltakene bør derfor være en generell økning i kunnskapsnivået i matematikk for alle elever som går ut av grunnskolen, og ikke minst, en økning i nivået for elevene som fullfører vgs med studiekompetanse. Nivået i grunnskolen kan følges ganske godt gjennom de internasjonale studiene, og som vi har sett har det vært en

framgang i alle fagområder som er inkludert her, også i matematikk. Men det er bemerkelsesverdig at effekter av det nye programfaget i vgs ikke er evaluert. I tillegg ble programområdet for realfag endret ved at de nye programfagene Teknologi og forskningslære og Geofag ble opprettet. Dette er tilbud som skulle gjøre det mer attraktivt å velge en realfaglig fordypning. Så langt er imidlertid utviklingen nedslående: Etter en periode med positiv utvikling i søkningen til programområdet, har det vært en nedgang de siste årene og kjønnskjevheten i valg av de ulike fagene består (Bjørkeng 2011).

Satsing på grunnleggende ferdigheter: Det er noe, men mangelfull, evidens knyttet til implementeringen av grunnleggende ferdigheter i de foreliggende evalueringsstudiene. Vi har identifisert en liten komponent i evalueringsstudien FIRE (ILS/NIFU) som gir beskrivelser og analyser av utsagn fra rektorer og lærere ved tre skoler som antyder en noe større bevissthet hos disse lærerne om grunnleggende ferdighet som en tverrfaglig komponent (Hertzberg 2010). I sluttrapporten fra denne evalueringsstudien gis det også resultater for noen av spørsmålene fra den felles spørringen for evalueringsstudiene som angår implementeringen av de grunnleggende ferdighetene. Skoleeiere og rektorer mener i stor grad at KLO6 har bidratt til å styrke de fem grunnleggende ferdighetene, mens lærerne er noe mer ambivalente. Videre viser svar fra lærerne at det er skriving og lesing som i størst grad er implementert i fagene, mens regning også nevnes av barneskolelærere som noe de jobber med på tvers av fag. I vgs er det spesielt digital kompetanse som nevnes som sentralt på tvers av fag. Men ingen av disse kunnskapsfragmentene kan egentlig opplyse oss om hvor langt vi har kommet i implementeringen av de grunnleggende ferdighetene.

Når det gjelder de enkeltstående ferdighetene, finnes det andre kilder som kan trekkes inn. Vi ser for eksempel at satsingen på lesing og regning som grunnleggende ferdighet har hatt tydelige effekter i form av tiltak, blant annet ved innføring av nasjonale prøver og kartleggingsprøver. Vi ser også som tidligere nevnt at elevene presterer bedre på disse to områdene i de internasjonale undersøkelsene. I tillegg finner evalueringsstudien til Nordlandsforskning gjennom en videoanalyse av 259 undervisningsøkter over tre år, fra 2007 til 2010, at den mest markerte økningen i undervisningsaktiviteter er stillelesing (Hodgson mfl. 2012). For grunnleggende digitale ferdigheter gir ITU Monitor noen enkeltstående funn som kan trekkes fram (Egeberg mfl. 2012). Et viktig utviklingstrekk som rapporteres, er at det er en dreining i skoleledernes prioriteringer fra grunnleggende IKT ferdigheter som det å kunne beherske regneark og tekstbehandling, til økt innsats for å integrere IKT i ulike undervisningsmetoder, og sammenfallende med dette rapporteres også lærerne i langt større grad at de nå bruker IKT ressurser som en del av sin undervisning. Dette gjelder spesielt i vgs, men økningen gjelder for alle skoleslag. Men hvordan det går med den grunnleggende digitale ferdigheten fra et elevperspektiv, vet vi dessverre lite om.

Resultatkvalitet i videregående opplæring: Vi har relativt god oversikt over gjennomstrømmingen av elevene i videregående opplæring. Registre gjør det mulig å følge elevenes løp i detalj. Men det er påfallende at det ikke finnes nasjonale indikatorer som mer direkte reflekterer læringsresultater i vgs. Det kan kanskje synes paradoksalt at det skolenivået hvor det samles mest informasjon om elevers prestasjoner, er det nivået hvor det kan sies minst om utviklingen i elevenes faglige prestasjoner. Videregående skole lever fortsatt i en tradisjon med relativt hyppige lærergitte prøver og eksamener som gir ytterligere presiseringer av standarder for hva som til slutt forventes av elevene. Men gjennom lærernes egne prøver kan man ikke få informasjon som er sammenliknbar på tvers av skoler, og verken prøver eller eksamener er utformet for å kunne studere utvikling over tid.

Endringer i klasseromspraksis: Hva slags endringer man kan observere når det gjelder klasseromspraksis er kanskje ikke et opplagt tema å etterlyse i evalueringen av en reform hvor nettopp sentrale føringer om bestemte undervisningsmetoder og aktiviteter bevisst er utelatt fra styringsdokumentene. Men som nevnt vet vi svært lite om hvordan lærerne i praksis møter utfordringene knyttet til grunnleggende ferdigheter i fag og kravet om å gi tilpasset opplæring innenfor rammene av ordinær helklasseundervisning. Kun studier basert på observasjoner av klasserommene og læreres undervisningspraksis kan gi et sikkert kunnskapsgrunnlag om disse aspektene ved reformimplementeringen. At det mangler et representativt og generaliserbart kunnskapsgrunnlag for å studere endringer i klasseromspraksis i K06-evalueringene og andre studier av norsk skole, henger naturlig nok sammen med at observasjon av undervisning i utgangspunktet er komplekst, omfattende og ressurskrevende. Utfordringen blir enda større dersom slike observasjoner skal gjøres i en skala som kan gi et representativt og generaliserbart bilde av situasjonen for hele landet, og ytterligere kompliserende blir det å skulle ha et longitudinelt eller tidsseriedesign som gjør det mulig å si noe om hvordan undervisningen endrer seg. Med ny teknologi er det imidlertid mulig å forsere noen av de største hindringene her. Vi er enn så lenge overlatt til tre typer kilder som i lys av denne syntesens problemstillinger må sies å representere evidens på et relativt lavt nivå: Kvalitative studier av noen utvalgte klasserom, intervjustudier og spørreskjemaundersøkelser.

I Nordlandsforsknings evalueringsstudie inngår en videostudie av mer enn 250 undervisningstimer over en periode på tre år etter innføringen av reformen, og det finnes intervju- og spørreskjemadata fra lærere og elever som i noen grad er rettet inn mot beskrivelser av undervisningen (Hodgson mfl. 2012). Med unntak av økt stillelesing konkluderer forfatterne med at det er få tegn til endring av undervisningsaktiviteter fra 2007 til 2010. Vi har tidligere presentert funn som tilsier at aktiviteter knyttet til underveisvurdering har økt, men det er uklart hvordan dette i praksis gjennomføres i klasserommene. Videre har vi god evidens for å hevde at elevene mer enn før deles inn i grupper etter forutsetninger. Hovedpoenget er imidlertid at dette er små drypp av relativt usikker kunnskap som ikke gir grunnlag for å trekke noen overordnede konklusjoner om eventuelle endringer i klasserommets praksisformer med innføringen av K06.

Mot et system for kontinuerlig evaluering av kvalitet i opplæringen

Denne syntesen har vist at selv om det finnes noen godt dokumenterte endringer i elevenes læringssituasjon så er det lite evidens for at K06 er årsaken til disse endringene. Vi har vært inne på flere årsaker til dette tidligere: Kausale effekter er generelt vanskelig å dokumentere, både fordi det er metodisk komplekst, men også fordi reformer er omfattende, langsiktige og implementeres i en kontekst. Det er mange andre skoleinterne og -eksterne forhold som påvirker det som skjer. Internasjonal forskning dokumenterer at reformer trenger tid for å bli implementert. De fleste studiene som vi har inkludert her, har kun sett på endringer noen få år etter innføringen av reformen, og man kunne derfor i utgangspunktet ikke forvente å finne store endringer.

«Kunnskapsløftet» eller «reform» er dessuten uklare størrelser. Det kan kanskje la seg gjøre å måle effekter av særskilte reformtiltak (for eksempel innføring av grunnleggende ferdigheter), men ikke av reformintensjoner og normative idealer (som for eksempel "et godt læringsmiljø"). Et særlig problem oppstår når man skal finne effekter mellom to eller flere uklare størrelser. Ett av våre oppdrag var å svare på hvordan reformen bidrar til utvikling av en bedre kultur for læring og bedre tilpasset opplæring for alle. Her finner vi tre uklare størrelser: *reform*, *kultur for læring* og *tilpasset opplæring*. Ikke bare skal vi si noe om effekten mellom disse tre, men vi skal avgrense det til *bedre* kultur for

læring og *bedre* tilpasset opplæring. Her blir det innført en normativitet i problemstillingen som vanskeliggjør empirisk arbeid.

Med disse forbeholdene om begrensninger i oppdraget og grunnlagsmaterialet for syntesen (primærstudiene), vil vi understreke at vi i dag vet mye mer om forhold ved opplæringen enn vi gjorde bare for noen få år siden. Kravet om kunnskapsbasert politikkutvikling har ført til at vi i dag har mer systematiske registerdata og at det samles inn mye god informasjon, og til sammen utgjør dette en base for å utvikle kvalitetsindikatorer. Ved hjelp av disse indikatorene har det vært mulig for oss å evaluere utviklingen for en rekke forhold i skolen. Selv om OECD sin ferske review av det norske vurderingssystemet identifiserer noen mangler og forbedringsområder (Nusche mfl. 2011), ser vi på sikt for oss at evalueringsstudiene delvis vil kunne bli erstattet av løpende analyser med jevnlige oppdateringer.

Referanser

- Aarkrog, V. og Bang, S. W. (2012). *Resultater af evalueringer af Kunnskapsløftet. En syntese*. Aarhus: Aarhus Universitet.
- Aasen, P., Møller, J., Rye, E., Eli, O., Prøitz, T. S. og Hertzberg, F. (2012). *Kunnskapsløftet som styringsreform - et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen*. (Rapport 20/2012). Oslo: NIFU.
- Bakken, A. (2009). *Ulikhet på tvers. Har foreldres utdanning, kjønn og minoritetsstatus like stor betydning for elevers karakterer på alle skoler?* (Rapport 8/2009). Oslo: NOVA.
- Bakken, A. (2010). *Prestasjonsforskjeller i Kunnskapsløftets første år - kjønn, minoritetsstatus og foreldres utdanning*. (Rapport 9/2010). Oslo: NOVA.
- Bakken, A. og Danielsen, K. (2011). *Gode skoler – gode for alle? En casestudie av prestasjonsforskjeller på seks ungdomsskoler*. (Rapport 10/2011). Oslo: NOVA.
- Bakken, A. og Elstad, J. I. (2012). *For store forventninger? Kunnskapsløftet og ulikhetene i grunnskolekarakterer* (Rapport nr 7/12). Oslo: NOVA.
- Birkeland, G. E., Hermansen, A. S. og Evensen, Ø. (2010). *Skolesegregering – et problem? Elevsammensetning, frafall og karakterer i Oslo-skolen*. Oslo: Institutt for sosiologi og samfunnsgeografi, Universitetet i Oslo.
- Bjørkeng, B. (2011). *Jenter og realfag i videregående opplæring* (Rapport 3/2011). Oslo-Kongsvinger: Statistisk sentralbyrå.
- Black, P. og Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Bonesrønning, H., Iversen, J. M. V. og Pettersen, I. (2010). *Kommunal skolepolitikk etter Kunnskapsløftet: Med spesielt fokus på økt bruk av spesialundervisning* (Rapport nr. 07/10). Trondheim: Senter for økonomisk forskning AS.
- Bonesrønning, H., Iversen, J. M. V. og Pettersen, I. (2011). Antall enkeltvedtak i grunnskolen øker – til fordel for elever som ikke mottar spesialundervisning? I N. P. Sandbu (red.), *Utdanning 2011 – veien til arbeidslivet* (s. 101-120). Oslo/Kongsvinger: Statistisk sentralbyrå.
- Borgen, J. S. og Lødding, B. (2009). *Implementering av faget utdanningsvalg på ungdomstrinnet. Delrapport II fra prosjektet Karriereveiledning i overgangen mellom ungdomsskole og videregående opplæring. Evaluering av Kunnskapsløftet*. (Rapport 39/2009). Oslo: NIFU Step.
- Dansk Clearinghouse for Uddannelsesforskning. (2007). Konseptnotat. Lastet ned 7. sept., 2011, fra http://www.dpu.dk/fileadmin/www.dpu.dk/danskcldclearinghouseforuddannelsesforskning/konceptnotat/Konceptnotat_om_Clearinghouse_version_1.2_110607.pdf
- Egeberg, G., Guðmundsdóttir, G. B., Hatlevik, O. E., Ottestad, G., Skaug, J. H. og Tømte, K. (2012). *Monitor 2011. Skolens digitale tilstand*. Oslo: Senter for IKT i utdanningen.
- Falch, T., Borge, L.-E., Lujala, P., Nyhus, O. H. og Strøm, B. (2010). *Årsaker til og konsekvenser av manglende fullføring av videregående opplæring* (Rapport nr. 03/10). Trondheim: Senter for økonomisk forskning AS.

- Falch, T., Johannesen, A. B. og Strøm, B. (2009). *Kostnader av frafall i videregående opplæring* (Rapport nr. 08/09). Trondheim: Senter for økonomisk forskning AS.
- Falch, T. og Nyhus, O. H. (2010). *Videregående opplæring og arbeidsmarkedstilknytning for unge voksne innvandrere* (Rapport nr. 04/10). Trondheim: Senter for økonomisk forskning AS.
- Falch, T. og Nyhus, O. H. (2011). *Betydningen av fullført videregående opplæring for sysselsetting blant unge voksne* (Rapport nr. 01/11). Trondheim: Senter for økonomisk forskning AS.
- Falch, T., Nyhus, O. H. og Strøm, B. (2011). *Grunnskolekarakterer og fullføring av videregående opplæring* (Rapport nr. 03/11). Trondheim: Senter for økonomisk forskning AS.
- Frøseth, M. W., Hovdhaugen, E., Høst, H. og Vibe, N. (2008). *Tilbudsstruktur og gjennomføring i videregående opplæring. Delrapport I. Evaluering av Kunnskapsløftet*. (Rapport 40/2008). Oslo: NIFU Step.
- Frøseth, M. W., Hovdhaugen, E., Høst, H. og Vibe, N. (2010). *En, to...tre? Den vanskelige overgangen. Evaluering av Kunnskapsløftet. Fra andre til tredje år i videregående opplæring*. (Rapport 21/2010). Oslo: NIFU Step.
- Fullan, M. (2007). *The new meaning of educational change*. London New York: Routledge Teachers College Press.
- Galloway, T. A., Kirkebøen, L. J. og Rønning, M. (2011). *Karakterpraksis i grunnskoler. Sammenheng mellom standpunkt- og eksamenskarakter* (4/2011). Oslo/Kongsvinger: Statistisk sentralbyrå.
- Gough, D., Oliver, S. og Thomas, J. (2012). *An introduction to systematic reviews*. Los Angeles: Sage.
- Grønmo, L. S. og Onstad, T. (2009). *Tegn til bedring : norske elevers prestasjoner i matematikk og naturfag i TIMSS 2007*. Oslo: Unipub.
- Grønmo, L. S., Onstad, T., Nilsen, T., Hole, A., Aslaksen, H. og Borge, I. C. (2012). *Framgang, men langt fram. Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2011*. Oslo: Akademika forlag.
- Hattie, J. (2009). *Visible learning : a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Haug, P. (2003). *Evaluering av Reform 97. Sluttrapport frå styret for Program for evaluering av Reform 97*. Oslo: Norges Forskningsråd.
- Hertzberg, F. (2010). Arbeid med grunnleggende ferdigheter. I E. Ottesen og J. Møller (red.), *Underveis, men i svært ulikt tempo. Et blikk inn i ti skoler etter tre år med Kunnskapsløftet. Delrapport 3 Underveisanalyse av Kunnskapsløftet som styringsform* (s. 77-89). Oslo: ILS/UiO og NIFU STEP.
- Hodgson, J., Rønning, W., Skogvold, A. S. og Tomlinson, P. (2010a). *På vei fra læreplan til klasserom. Om læreres fortolkning, planlegging og syn på LK06*. (Rapport nr. 3/2010). Bodø: Nordlandsforskning.
- Hodgson, J., Rønning, W., Skogvold, A. S. og Tomlinson, P. (2010b). *Vurdering under Kunnskapsløftet. Læreres begrepsforståelse og deres rapporterte og faktiske vurderingspraksis*. (Rapport nr 17/2010). Bodø: Nordlandsforskning.
- Hodgson, J., Rønning, W. og Tomlinson, P. (2012). *Sammenhengen mellom undervisning og læring. En studie av læreres praksis og deres tenkning under Kunnskapsløftet* (Rapport nr. 4/2012). Bodø: Nordlandsforskning.
- Hopfenbeck, T. N., Kjærnsli, M. og Throndsen, I. (2010). Rektorenes svar på spørsmål i skolelederspørreskjemaet. I M. Kjærnsli og A. Roe (red.), *På rett spor - Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009* (s. 178-192): Universitetsforlaget.
- Hopfenbeck, T. N., Throndsen, I., Lie, S. og Dale, E. L. (2012). Assessment with distinctly defined criteria. A research study of a National project. *Policy Futures in Education*, 10 (5).
- Høst, H. og Evensen, M. (2009). *Ny struktur – tradisjonelle mønstre? Om Kunnskapsløftets strukturendringer i det videregående opplæringssystemet, og utforming av tilbud og dimensjonering i fylkeskommunene*. (Rapport 28/2009). Oslo: NIFU Step.
- Jenssen, E. og Lillejord, S. (2009). Tilpasset opplæring: Politisk dragkamp om pedagogisk praksis. *Acta Didactica*, 3(1), 1-15.

- Kjærnsli, M. og Roe, A. (red.). (2010). *På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009*. Oslo: Universitetsforlaget.
- Kunnskapsdepartementet. (2010). *Realfag for framtida. Strategi for styrking av realfag og teknologi 2010–2014*. Oslo: Kunnskapsdepartementet.
- Kunnskapsdepartementet. (2012). Midtveisrapport Ny GIV. Lastet ned fra http://www.regjeringen.no/upload/KD/Kampanjer/NyGiv/Midtveisrapport_endelig.pdf
- Kunnskapssenteret. (2009). *Slik oppsummerer vi forskning. Håndbok for Nasjonalt kunnskapssenter for helsetjenesten*. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.
- Lillejord, S. (1997). Målstyring, skoleutvikling og skoleledelse. I O. L. Fuglestad og S. Lillejord (red.), *Pedagogisk ledelse - et relasjonelt perspektiv*. Bergen: Fagbokforlaget.
- Lødding, B. og Borgen, J. S. (2008). *Karriereveiledning i overgangen mellom ungdomsskole og videregående opplæring. Delrapport I. Evaluering av Kunnskapsløftet*. (Rapport 41/2008). Oslo: NIFU Step.
- Lødding, B. og Holen, S. (2012). *Utdanningsvalg som fag og utfordring på ungdomstrinnet. Sluttrapport fra prosjektet Karriereveiledning i overgangen mellom ungdomsskole og videregående opplæring. Evaluering av Kunnskapsløftet* (Rapport 28/2012). Oslo: NIFU.
- Mikkelsen, R., Fjelstad, D. og Lauglo, J. (2011). *Morgendagens samfunnsborgere: Norske ungdomselevers prestasjoner og svar på spørsmål i den internasjonale demokratiundersøkelsen ICCS*. Oslo: Institutt for lærerutdanning og skoleforskning.
- Noblit, G. W. og Hare, R. D. (1988). *Meta-ethnography : synthesizing qualitative studies*. Newbury Park [CA]. Sage.
- Nordenbo, S. E. (2012). *Kunnskapsløftet som reformprosess*. Aarhus: Institut for uddannelsesforskning og pædagogik.
- NOU. (2002:10). *Førsteklasses fra første klasse. Forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem av norsk grunnopplæring*. Oslo: Utdannings- og forskningsdepartementet.
- Nusche, D., Earl, L., Maxwell, W. og Shewbridge, C. (2011). *OECD Reviews of Evaluation and Assessment in Education: NORWAY*. Paris: OECD.
- Nyen, T., Næss, T., Skålholt, A. og Tønder, A. H. (2011). *På veien til fagbrev: Analyser av Lærlingundersøkelsen* (Rapport 2011:28). Oslo: FAFO.
- OECD. (1989). *Ekspertvurdering fra OECD. OECD-vurdering av norsk utdanningspolitikk*. (A.-M. Smith, oversettelse.). Oslo: Aschehoug.
- OECD. (2001). *Knowledge and Skills for Life. First results from PISA 2000*. Paris: OECD Publications.
- OECD. (2004). *Learning for Tomorrow's World. First Results From PISA 2003*. Paris: OECD Publications.
- OECD. (2007). *PISA 2006: Science competencies for tomorrow's world*. Paris: OECD Publishing.
- OECD. (2010). *PISA 2009 Results: What Students Know and Can Do. Student Performance in Reading, Mathematics and Science* (Vol. I). Paris: OECD Publishing.
- Olsen, R. V. og Turmo, A. (2010). Et likeverdig skoletilbud? I M. Kjærnsli og A. Roe (red.), *På rett spor - Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009* (s. 207-225). Oslo: Universitetsforlaget.
- Rambøll Management. (2007). *Evaluering av strategiplanen 'Realfag, naturligvis'*. Oslo: Rambøll Management.
- Rambøll Management. (2012). *Evaluering av årlige tilstandsrapporter. Delrapport 1*. Oslo: Rambøll Management.
- Rønning, W., Fiva, T., Henriksen, E., Krogtoft, M., Nilsen, N. O., Skogvold, A. S. mfl. (2008). *Læreplan, læreverk og tilrettelegging for læring. Analyse av læreplan og et utvalg læreverk i naturfag, norsk og samfunnsfag*. (Rapport 2/2008). Bodø: Nordlandsforskning.
- Sarason, S. B. (1990). *The predictable failure of educational reform : can we change course before it's too late?* San Francisco: Jossey-Bass.
- Seeberg, M. L., Seland, I. og Hassan, S. C. (2012). «Litt vanskelig at alle skal med!». *Rapport 1: Evaluering av leksehjelpstilbudet 1.–4. trinn* (Rapport 3/2012). Oslo: NOVA.

- Sivesind, K. (2012). *Kunnskapsløftet: implementering av nye læreplaner i reformen*. Oslo: Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Skaalvik, E. M. og Skaalvik, S. (2009). *Skolen som læringsarena. Selvoppfatning, motivasjon og læring*. Oslo: Universitetsforlaget.
- Sletten, M. A., Bakken, A. og Haakestad, H. (2012). *Ny start med Ny GIV? Kartlegging av intensivopplæringen i regi av Ny GIV-prosjektet skoleåret 2010/11* (Rapport 23/11). Oslo: NOVA.
- St. mld. nr. 28. (1998-99). *Mot rikare mål. Om einskapsskolen, det likeverdige opplæringstilbudet og ein nasjonal strategi for vurdering og kvalitetsutvikling i grunnskolen og den vidaregåande opplæringa*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- St. mld. nr. 30. (2003-04). *Kultur for læring*. Oslo: Utdannings- og forskningsdepartementet.
- St. mld. nr. 33. (1991-92). *Kunnskap og kyndighet. Om visse sider ved vidaregåande opplæring*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- St. mld. nr. 37. (1990-91). *Om organisering og styring i utdanningssektoren*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- St. mld. nr. 47. (1995-96). *Elevvurdering, skolebasert vurdering og nasjonalt vurderingssystem*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- St. prp. nr. 1. (2002-03). *Tillegg nr. 3*. Oslo: Utdannings- og forskningsdepartementet.
- Statistisk sentralbyrå. (2011). *Utdanning 2011 – veien til arbeidslivet*. Oslo/Kongsvinger: Statistisk sentralbyrå.
- Støren, L. A., Helland, H. og Grøgaard, J. B. (2007). *Og hvem stod igjen...? Sluttrapport fra prosjektet Gjennomstrømning i vidaregåande opplæring blant elever som startet vidaregåande opplæring i årene 1999-2001* (Rapport 14/2007). Oslo: NIFU STEP.
- Thronsdén, I. S., Hopfenbeck, T. N., Lie, S. og Dale, E. L. (2009). *Bedre vurdering for læring. Rapport fra "Evaluerer av modeller for kjennetegn på måloppnåelse i fag"*. Oslo: Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.
- Tyack, D. B. og Cuban, L. (1995). *Tinkering toward utopia : a century of public school reform*. Cambridge, Mass.: Harvard University Press.
- Utdanningsdirektoratet. (2012). *Utdanningsspeilet 2012*. Oslo: Utdanningsdirektoratet.
- Van Daal, V., Solheim, R. G. og Gabrielsen, N. N. (2012). *Godt nok? Norske elevers leseferdighet på 4. og 5. trinn*. Stavanger: Lesesenteret, Universitetet i Stavanger.
- Vibe, N. (2010). *Spørsmål til Skole-Norge høsten 2010: Resultater og analyser fra utdanningsdirektoratets spørreundersøkelse blant skoler og skoleeiere* (Rapport 40/2010). Oslo.
- Vibe, N. (2012). *Spørsmål til Skole-Norge 2011: Resultater og analyser fra Utdanningsdirektoratets spørreundersøkelse blant skoler og skoleeiere* (Rapport 5/2012). Oslo: NIFU.
- Vibe, N., Brandt, S. S. og Hovdhaugen, E. (2011). *Underveis i vidaregåande opplæring. Evaluering av Kunnskapsløftet. Underveissrapport fra prosjektet "Struktur, gjennomføring og kompetanseoppnåelse"*. (Rapport 11/2011). Oslo: NIFU.
- Vibe, N., Frøseth, M. W., Hovdhaugen, E. og Markussen, E. (2012). *Strukturer og konjunkturer. Evaluering av Kunnskapsløftet. Sluttrapport fra prosjektet "Tilbudsstruktur, gjennomføring og kompetanseoppnåelse i vidaregåande opplæring"*. (Rapport 26/2012). Oslo: NIFU.
- Wendelborg, C., Røe, M. og Skaalvik, E. (2011). *Elevundersøkelsen 2011: Analyse av Elevundersøkelsen 2011*. Trondheim: NTNU Samfunnsforskning.
- Wiborg, Ø., Arnesen, C. Å., Grøgaard, J. B., Støren, L. A. og Opheim, V. (2011). *Elevers prestasjonsutvikling – hvor mye betyr skolen og familien? Andre delrapport fra prosjektet «Ressurser og resultater»* (Rapport 35/2011). Oslo: NIFU.
- Øia, T. (2011). *Ungdomsskoleelever : motivasjon, mestring og resultater* (Rapport 9/2011). Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.



Institutt for lærerutdanning og skoleforskning

Det utdanningsvitenskapelige fakultet

Universitetet i Oslo

Postboks 1099 Blindern

0317 OSLO

Dept. of Teacher Education and School Research

Faculty of Education

University of Oslo

P.O.Box 1099 Blindern

0317 Oslo

Norway

www.ils.uio.no

ISSN 1503-8076

ISBN 978-82-90904-98-7