

Jakten på kvalitetsindikatorerne

Sluttrapport fra prosjektet 'Ressursbruk og læringsresultater i grunnsopplæringen'

Vibeke Opheim, Cay Gjerustad og Jørgen Sjaastad

Rapport 23/2013

Jakten på kvalitetsindikatorerne

Sluttrapport fra prosjektet 'Ressursbruk og læringsresultater i grunnopplæringen'

Vibeke Opheim, Cay Gjerustad og Jørgen Sjaastad

Rapport 23/2013

Rapport 23/2013

Utgitt av Nordisk institutt for studier av innovasjon, forskning og utdanning
Adresse PB 5183 Majorstuen, NO-0302 Oslo. Besøksadresse: Wergelandsveien 7, 0167 Oslo

Oppdragsgiver Utdanningsdirektoratet
Adresse Boks 9359 Grønland, 0135 Oslo

Trykk Link Grafisk

ISBN 978-82-7218-925-8
ISSN 1892-2597 (online)

www.nifu.no

Forord

Dette er sluttrapporten fra prosjektet «Ressursbruk og læringsresultater i grunnopplæringen». Prosjektet er gjennomført på oppdrag fra Utdanningsdirektoratet og startet opp i begynnelsen av 2010. I løpet av de vel tre årene fra oppstart til begynnelse nav 2013, har prosjektet omfattet en rekke analyser presentert i tilsammen seks delrapporter. Formålet med sluttrapporten er å oppsummere og diskutere hovedfunn fra de ulike delrapportene sett under ett, og i lys av tidligere forskning. Formålet er også å peke på mulige kvalitetsindikatorer for grunnopplæringen, samt hvordan kvalitetsindikatorer kan videreutvikles med utgangspunkt i tilgjengelige datakilder på nasjonalt nivå.

Prosjektleder har vært Vibeke Opheim. Cay Gjerustad har hatt hovedansvar for kapittel 4 og 5, Jørgen Sjaastad har hatt hovedansvar for kapittel 3 og 6, Opheim har hatt hovedansvar for kapittel 1 og 7, mens kapittel 2 er utarbeidet av Gjerustad og Opheim i samarbeid. Idunn Seland og Jens Grøgaard har bidratt med tekst til teorigjennomgangen i kapittel 1. Vi takker professor Hans-Thore Hansen ved Universitetet i Bergen og Liv Anne Støren ved NIFU for nyttige kommentarer til tidligere utkast. Vi takker også elever, lærere og skoleledere ved skolene som har gitt av sin tid og levert viktige bidrag til prosjektet.

Oslo, 8. Mai

Sveinung Skule
Direktør

Vibeke Opheim
Forskningsleder

Innhold

Sammendrag	9
1 Innledning, bakgrunn og kunnskapsstatus	15
1.1 Innledning	15
1.2 Om prosjektet.....	15
1.2.1 Enkle og komplekse skoleressurser	18
1.3 Tidligere studier	19
1.4 Skoleeffektivitetsforskning	19
1.5 Klasseromsforskning.....	21
1.6 Ulikhetsforskning.....	22
1.7 Oppsummering	24
1.8 Rapportens videre gang	24
2 Hovedfunn fra tidligere delrapporter	26
2.1 Prosjektets delrapporter.....	26
2.1.1 Delrapport 1: Elevenes prestasjonsnivå på tre trinn i grunnskolen.....	26
2.1.2 Delrapport 2: Elevenes prestasjonsutvikling.....	27
2.1.3 Delrapport 3: Prestasjonsutvikling gjennom Vg1	28
2.1.4 Delrapport 4: Prestasjoner og læringsmiljø i to fag.....	29
2.1.5 Delrapport 5: Skoler med høyt prestasjonsnivå på nasjonale prøver	30
2.1.6 Delrapport 6: Den gode timen – casestudie av fire Oslo-skoler.....	30
2.2 Prosjektets hovedfunn: Bidrag til kunnskapsstatus	31
2.2.1 Sterk sammenheng mellom elevkjennetegn og læringsutbytte	31
2.2.2 Læringsmiljøet viktig for elevenes prestasjoner.....	32
2.2.3 Betydning av ulike undervisningsformer	32
2.2.4 Kan det bli for mye vurdering?	36
2.2.5 Små forskjeller mellom skoler	37
2.2.6 Betydning av skoleressurser	38
2.2.7 Stor variasjon mellom skoletimer.....	39
2.2.8 Sammenheng mellom skoleledelse og elevenes læringsutbytte?	40
2.3 Oppsummering	41
3 Ulike prestasjonsmål som indikatorer på kvalitet	42
3.1 Nasjonale prøver og karakterdata som mål på kvalitet.....	42
3.2 Kritikk mot ulike prestasjonsmål	42
3.3 Relative mål for prestasjonsutvikling	43
3.4 Personer som ikke er representert i datamaterialet.....	44
3.5 Snevert mål på skolekvalitet?	44
3.6 Hvor lett er det å måle bredden av skolens målsettinger?.....	45
3.7 Hva om andre prestasjonsmål hadde blitt anvendt?.....	46
3.8 Begrunnelsen for å benytte data fra nasjonale prøver og karakterdata som prestasjonsmål.....	46
4 Informasjon fra GSI som indikatorer på ressursbruk	48
4.1 Innhold i GSI: Ressurser og rammebetingelser	48
4.2 Informasjon fra GSI som bidrag til å forklare prestasjoner	49
4.3 Styrker og svakheter ved GSI	50
4.3.1 Mye informasjon.....	50
4.3.2 Uklart hva som måles – indirekte kvalitetsindikatorer?.....	50
4.4 Stabilitet og endring i timetall og antall PC-er per elev	52
4.5 Oppsummering	56
5 Elevundersøkelsen som indikator på elevenes læringsmiljø	57
5.1 Elevundersøkelsen: Bakgrunn og innhold	57
5.2 Funn fra Elevundersøkelsen i prosjektet	58
5.3 Styrker og svakheter ved Elevundersøkelsen som kvalitetsindikator	59
5.3.1 Innhold	60
5.3.2 Spørsmålsformuleringer og svarkategorier	62
5.3.3 Elevenes motivasjon for å svare	63
5.4 Elevundersøkelsen: Informasjon på ulike nivåer	65
5.4.1 Opplevelse av læringsmiljøet – betydningen av individ-, klasse- og skolenivå	65
5.4.2 Forklaringskraft til Elevundersøkelsen	67

5.5	Oppsummering	68
6	Spørreundersøkelsen til elever og lærere i to fag	69
6.1	Ett skritt nærmere kontekstnær informasjon	69
6.2	Data og design: Styrker og svakheter	70
6.2.1	Eksempel 1: Kobling av elevers og læreres responser	71
6.2.2	Eksempel 2: Spørsmål for to konkrete fag	73
6.3	Sammenligning med Elevundersøkelsen	75
6.4	En spørreundersøkelse med utviklingspotensial	79
6.4.1	En retning for fremtidige studier?	80
6.5	Oppsummering og konklusjon	82
7	Oppsummering: Veien videre mot gode kvalitetsindikatorer	84
7.1	Prosjektets hovedfunn: Bidrag til kunnskapsstatus	84
7.2	Anbefalinger for videre utvikling av kvalitetsindikatorer for grunnopplæringen	85
7.2.1	Anbefalinger for videre utvikling av nasjonale prøver og andre prestasjonsmål	85
7.2.2	Anbefalinger for videre utvikling av GSI	86
7.2.3	Anbefalinger for videre utvikling av Elevundersøkelsen	87
7.3	Tema for videre forskning	88
	Referanser	90

Sammendrag

Dette er sluttrapporten fra prosjektet «Ressursbruk og læringsresultater i grunnsopplæringen». Prosjektet er gjennomført på oppdrag fra Utdanningsdirektoratet og startet opp i begynnelsen av 2010. I løpet av de vel tre årene fra oppstart til begynnelsen av 2013, har prosjektet omfattet en rekke analyser presentert i tilsammen seks delrapporter. Formålet med sluttrapporten er både å oppsummere og diskutere hovedfunn fra de ulike delrapportene, og se resultatene i sammenheng, men også å peke på mulige kvalitetsindikatorer for grunnsopplæringen, og hvordan kvalitetsindikatorer kan videreutvikles med utgangspunkt i tilgjengelige datakilder.

Prosjektet har hatt som formål å utnytte bredden i de eksisterende datakilder på nasjonalt nivå. Gjennom analyser av et bredt sammensatt datamateriale har hensikten både vært å skape et nyansert bilde av sammenhenger mellom elevresultater og ulike innsatsfaktorer, samt legge grunnlag for utvikling av gode kvalitetsindikatorer for grunnsopplæringen. Dette gjøres i form av anbefalinger for videre utvikling av de sentrale nasjonale kildene til informasjon om skolen: De nasjonale prøvene og andre prestasjonsmål, Grunnskolens informasjonssystem (GSI) og Elevundersøkelsen.

Prosjektets bidrag til kunnskapsfeltet

Prosjektet har bidratt med ny kunnskap om elevenes prestasjonsnivå på 5., 8. og 10. trinn, samt prestasjonsnivået i Vg1. Prosjektet har undersøkt elevenes prestasjonsutvikling fra 5-8. trinn, 8-10 trinn, samt fra 10. trinn-Vg1. Dette har bidratt med ny kunnskap om hvilke faktorer ved skolens ressurser, elevenes læringsmiljø, elevsammensetningen ved skolen og kjennetegn ved elevene selv som bidrar til elevenes prestasjonsutvikling på tre ulike trinn i grunnsopplæringen.

Flere av analysene i prosjektet bekrefter funn fra tidligere forskning. Også tidligere studier har pekt på relativt små forskjeller mellom skoler i Norge, men store forskjeller innad ved den enkelte skole. Prestasjonsforskjellene mellom elever som går på samme skole og i samme klasse, er sterkt knyttet til kjennetegn ved elevene – deres sosiale bakgrunn, kjønn og minoritetsbakgrunn. Gjennom analyser av dette omfattende datamaterialet, har prosjektet bidratt med ny kunnskap om hvordan like rammefaktorer og kjennetegn ved elevenes læringsmiljø og undervisningssituasjon kan ha ulik betydning for ulike grupper av elever.

Et sentralt funn i prosjektet er betydning av elevenes læringsmiljø. Alle elever tjener på å ha et godt læringsmiljø, men for noen elevgrupper, særlig de som står i fare for å falle utenfor skolens fellesskap, er dette særlig viktig. Elever med minoritetsbakgrunn har større risiko for å være svakt integrert i skolens fellesskap, og dermed falle utenfor. Prosjektet peker på sammenhenger mellom integrasjon, elevenes oppfatning av sitt læringsmiljø og deres læringsutbytte.

Prosjektet har også bidratt med innsamling av nye empiriske data, både i form av en kvantitativ spørreundersøkelse til elever og lærere, samt en kvalitativ casestudie av fire skoler i Oslo. Spørreundersøkelsen til elever og lærere har gitt ny kunnskap om sammenhenger mellom elevenes og lærerens oppfatning om læringsmetoder i norsk og matematikk, og elevenes prestasjoner i disse fagene. Ved at både elever og faglærerne deres har fått likelydende spørsmål om undervisningssituasjonen og læringsmiljøet gir dette en kilde til direkte lærer-elev-informasjon. Analysene avdekker ulike oppfatninger om hva som skjer i klasserommet blant lærer og elever.

Et formål med casestudien var, med utgangspunkt i de kvantitative analysene i prosjektet, å undersøke nærmere ulike læringsprosesser i klasserommet. Casestudien har bidratt med nye data og gitt utdypende kunnskap om hva som skjer i klasserommet. Resultater fra casestudien har bidratt til diskusjonen av resultater og tolkninger av analysene fra registerdata og spørreundersøkelsen. Ved bruk av ulike datakilder og metodisk tilnærminger, har prosjektet samlet sett bidratt til å belyse komplekse sammenhenger mellom skoleressurser (input), læringsprosesser i klasserommet og elevenes prestasjoner (resultater/output). Skoleressurser er her forstått i vid forstand – både som skolens materielle og menneskelige ressurser, elevenes oppfatninger av eget læringsmiljø, samt kjennetegn ved elevene selv, deres foreldre, og skolens elevsammensetning.

Hovedfunn

Prosjektet har omfattet en stor mengde analyser av ulike datasett. En oppsummering av hovedfunn medfører rimeligvis en kraftig forenkling av resultatene fra de ulike analysene. Nedenfor har vi likevel forsøkt å sammenfatte hovedfunnene fra prosjektet.

Sterk sammenheng mellom elevkjennetegn og læringsutbytte: Resultatene viser gjennomgående sterke sammenhenger mellom individuelle kjennetegn som kjønn, sosioøkonomisk bakgrunn og innvandrerbakgrunn og elevenes prestasjoner. Også elevsammensetningen ved skolen har stor betydning for elevenes prestasjoner. Det betyr at ikke bare elevens egne kjennetegn, men også kjennetegn ved de andre elevene på skolen, har betydning for elevenes prestasjoner. Skolens elevsammensetning får dermed form av å være en ressursindikator på skolenivå.

Resultatene er i tråd med tidligere studier. Analysene i dette prosjektet går likevel videre enn mye av det som er gjort i tidligere studier, ved å undersøke nærmere hvordan indikatorer på skoleressurser og læringsmiljø kan ha ulik betydning for ulike grupper av elever. Dette utdypes i flere av hovedfunnene.

Elevkjennetegn har stor betydning for elevenes prestasjoner, men også for elevenes prestasjonsutvikling. Et mye omtalt tema i norsk skole er i hvilken grad skolen bidrar til å utjevne sosiale forskjeller i elevenes læringsutbytte eller om dette er forskjeller som forsterkes utover i skoleløpet. Resultatene tyder i liten grad på at sosiale forskjeller reduseres i løpet av skolegangen. Tvert i mot, synes de sosiale forskjellene mellom elevenes prestasjonsnivå å øke noe utover i skoleløpet. Dette kan tolkes som uttrykk for at skolen ikke kompenserer for forskjeller i elevenes bakgrunn. Samtidig kan vi ikke se bort i fra at skolen bidrar til at de sosiale forskjellene mellom elevenes læringsutbytte ikke er *enda* større.

Små forskjeller mellom skoler: Et gjentatt funn i de kvantitative analysene er at relativt lite av variasjonen i elevenes prestasjoner kan tilbakeføres til skolenivået. Omtrent 7 til 10 prosent av variasjonen i elevenes prestasjoner kan knyttes til forskjeller mellom skoler. Dette er lavere enn gjennomsnittet for OECD-land. Når relativt lite av variasjonen i elevens prestasjoner kan tilbakeføres til hvilken skole man går på, innebærer det at mye av variasjonen skyldes forskjeller mellom elever. Mesteparten av variasjonen i prestasjoner finnes mellom elever i samme klasse og på samme skole, og ikke mellom skoler eller klasser.

Selv om andelen variasjon som kan knyttes til skolenivået er relativt liten er det ikke grunnlag for at si at ulikheter mellom skoler ikke har betydning for elevenes prestasjoner. Forskjellene mellom ytterpunktene er betydelige. Å løfte skolene i den nederste delen av fordelingen opp på nivå med de

øverste skolene ville hatt betydelig effekt for de elevene det gjelder, og ville gitt utslag i høyere gjennomsnittlige prestasjoner på nasjonalt nivå.

Læringsmiljøet viktig for elevenes prestasjoner: Analysene viser gjennomgående at alle elevgrupper, uavhengig av kjønn, sosioøkonomisk bakgrunn og innvandrerbakgrunn, tjener på et godt læringsmiljø.

Et godt læringsmiljø, det vil si et miljø som er inkluderende, støttende, trivelig, punktlig og faglig fokusert, synes å være spesielt positivt for elever med minoritetsbakgrunn. På skoler med godt læringsmiljø er prestasjonsnivået blant minoritetselever høyere enn på skoler med svakere læringsmiljø.

Sammenheng mellom undervisningsform og elevprestasjoner: I analysene skilles det mellom to undervisningsformer: Lærerstyrt og elevaktiv undervisning. Lærerstyrt undervisning kjennetegnes, slik det er definert i dette prosjektet, av tavleundervisning, individuelt arbeid og diskusjon/samtale mellom læreren og elevene i klassen. Elevaktiv undervisning kjennetegnes av ulike former for samarbeid mellom elevene, som gruppearbeid, prosjektarbeid, praktiske øvelser og samarbeid to og to.

Lærerstyrt undervisning synes gjennomgående å være positivt knyttet til elevenes prestasjoner, både målt på individnivå og som gjennomsnitt på skolenivå. Elevaktiv undervisning er derimot i større grad negativt knyttet til elevenes prestasjoner. Sammenhengene er knyttet til elevenes læringsmiljø og omfang av uro i klasserommet: Det synes å være generelt mindre uro i timer dominert av lærerstyrt undervisning enn i timer med elevaktive undervisningsformer. En tolkning av resultatene kan derfor være at lærerstyrt undervisning i større grad er kjennetegnet av tydelig klasseledelse enn elevaktiv undervisning. Samtidig kan det ikke utelukkes at elevenes prestasjonsnivå kan ha betydning for læreres valg av undervisningsformer. Dette tema diskuteres nærmere i kapittel 2.

Enkelte analyser kan tyde på at den positive sammenhengen mellom lærerstyrt undervisning for elevenes prestasjoner, primært gjelder elever med majoritetsbakgrunn. Elever med minoritetsbakgrunn synes derimot å ha et bedre utbytte av elevaktive undervisningsformer. En mulig tolkning av dette kan være at elevaktive undervisningsformer har integrerende effekter ettersom de innebærer samarbeid mellom elevene i klassen, og at dette har en positiv betydning for læringsutbyttet blant elever med minoritetsbakgrunn.

Høyt omfang av vurdering negativt for elevenes prestasjonsnivå: Analysene finner gjennomgående negative sammenhenger mellom et høyt omfang av evaluering og elevenes prestasjonsnivå. Mye tidligere forskning har pekt på at dersom evaluering skal være positivt for prestasjonene er det viktig at den integreres i undervisningen og at den følges opp av gode og informative tilbakemeldinger til eleven. Evaluering i seg selv synes ikke å føre til økte prestasjoner, den må følges opp i etterkant gjennom tilbakemeldinger til eleven og justering av undervisningsopplegget for at den skal ha effekt. I tråd med dette, kan en tolkning av disse analysene være at mye tid på evaluering og vurdering av elevene ikke synes å følges opp på måter som forbedrer elevenes prestasjoner, men derimot kan synes å gå utover tid til undervisning og læring. Skillet mellom evaluering som en naturlig del av undervisningen (med positiv betydning for læringsutbytte) og evaluering som formalisert testing diskuteres nærmere i kapittel 2.

Betydning av skoleressurser: Analysene i prosjektet viser at skolenes ressurser, slik de er registrert i GSI, i liten grad har sammenheng med elevenes prestasjoner. Funnet kan gi inntrykk av at skolenes ressurser ikke er viktig for prestasjonene. Trolig er det mer riktig å se funnene som indikator på at skolene har relativt lik mengde med tilgjengelige ressurser og at spredningen i ressursbruk er liten. Selv om analysene i liten grad finner sammenheng med skolens materielle og økonomiske ressurser, slik det er registrert i GSI, og elevenes prestasjonsnivå, finnes det enkelte unntak.

Antall undervisningstimer er positivt relatert til elevenes prestasjoner på skolenivå. Sammenhengen er påvist i analyser av flere ulike klassetrinn. Dette er i tråd med tidligere studier som har vist en positiv relasjon mellom tid til læring og prestasjoner.

Antall pc-er per elev er en annen ressursindikator som synes å være positivt relatert til elevenes prestasjoner på skolenivå. Sammenhengen kan være uttrykk for at skoler med et høyt antall elevpc-er også har ressurser i form av skoleledelse som prioriterer slik ressursbruk, og lærere som kan utnytte disse teknologiske ressursene.

Stor variasjon mellom skoletimer: Casestudien finner stor variasjon i kvaliteten på undervisningstimene innad på samme skole, når det gjelder forhold med betydning for elevenes læring som elevenes konsentrasjon og arbeidsro. Forskjellene i skoletimene peker på betydningen av læreren og hvordan læreren organiserer undervisningen og håndterer forstyrrelser og uro. Casestudien viser imidlertid også eksempler på at den samme klasse oppførte seg klart forskjellig i to ulike timer, til tross for at de hadde den samme læreren ved begge anledningene. Selv om observasjoner i casestudien peker på betydningen av læreren, viser de også hvordan ulike faktorer sammen skaper gode eller mindre gode timer.

Sammenheng mellom skoleledelse og elevenes læringsutbytte? Sammenhenger mellom skolens ledelse og elevenes læringsutbytte er komplisert å studere og vanskelig å påvise. Casestudien peker likevel på noen mulige sammenhenger mellom skolens ledelse og den pedagogiske plattformen ved skolen og kvaliteten i de enkelte undervisningstimene. Samtidig er det vanskelig å skille skolens pedagogiske plattform og skoleleders faglige prioriteringer fra elevsammensetningen ved hver enkelt skole. Det kan være at skolens pedagogiske plattform først og fremst reflekterer elevgruppen ved skolen og de utfordringer/muligheter den innebærer. Resultatene kan imidlertid tyde på at enkelte skoler har et høyere prestasjonsnivå enn forventet, ut i fra skolens elevsammensetning, og at dette har sammenheng med at skolen har en tydelig pedagogisk plattform, en tydelig ledelse, og lærere som er tydelige ledere i klasserommet. Dette er forhold som synes å henge sammen.

Anbefalinger med hensyn til videreutvikling av de nasjonale prøvene og andre prestasjonsmål

I dette prosjektet blir nasjonale prøver og karakterdata benyttet som mål på elevenes prestasjoner. Nasjonale prøver er den eneste kilden til informasjon om elevenes prestasjoner på grunnskolenivå som dekker hele elevkullet. Dette er en forutsetning for analyser av sammenheng mellom skolens ressurser og elevenes læringsutbytte. Datamaterialet fra nasjonale prøver er svært solid, og elevenes resultater på denne type tester henger ofte svært tett sammen med andre målsettinger ved norsk skole som ikke måles direkte i testene. For å styrke anvendeligheten av nasjonale prøver som prestasjonsmål i norsk skole, gir vi her tre anbefalinger:

Større vektlegging av å ha en enhetlig fritakspraksis: Dette betyr ikke nødvendigvis en strengere fritakspraksis. Det frarådes å fremme en praksis som fremtvinger deltakelse av elever som ikke har utbytte av eller forutsetninger for å besvare oppgavene.

Utvikling mot et mål på reell prestasjonsutvikling: Denne utviklingen er allerede i gang og anbefales videreført. Deler av prøvene må holdes like og gjentas hvert år, noe man har god erfaring med blant annet i PISA- og TIMSS-undersøkelsene. Dette gir muligheter for å sammenligne elevers reelle prestasjonsnivå og -utvikling over tid, ikke bare det relative nivået.

Elevenes motivasjon: En forutsetning for høy datakvalitet er at elevene er motiverte til å gjøre sitt beste på nasjonale prøver. Det anbefales å vektlegge denne siden ved gjennomføring av nasjonale prøver i større grad enn i dag. Elevene kan for eksempel spørres om å vurdere egen motivasjon og innsats ved gjennomføring av prøvene. Dette gjøres blant annet i PISA-undersøkelsen.

Anbefalinger med hensyn til videreutvikling av GSI

Grunnskolens informasjonssystem (GSI) inneholder mye informasjon om skolene som brukes av ulike aktører og med ulikt formål. I arbeidet med utvikling av kvalitetsindikatorer, er det likevel et tankekors at informasjonen i GSI i liten grad synes å være relatert til elevenes prestasjonsnivå, målt som gjennomsnitt på skolenivå. Det betyr ikke at skoleressurser *ikke* har betydning for elevens læringsutbytte, resultatene må i stedet ses i sammenheng med den variasjon i elevenes prestasjoner som finnes innen den enkelte skole. Det er lite sammenheng mellom skolens *gjennomsnittlige* ressursituasjon og elevenes *gjennomsnittlige* prestasjoner. Innholdet i GSI og GSI som datakilde diskuteres i kapittel 4. I lys av analyser i ulike deler av prosjektet, foreslår vi nedenfor tre områder for en videreutvikling av GSI som kan øke mulighetene for å bruke opplysninger fra GSI som indikatorer på kvalitet i skolen.

Mer informasjon om lærernes kvalifikasjoner: GSI inneholder lite informasjon om lærernes kompetanse og kvalifikasjoner, utover andelen lærere på skolen med godkjent utdanning. Lærerens rolle som klasseleder og kunnskapsformidler fanges dermed i liten grad opp av GSI. På grunnlag av funn fra prosjektet anbefales en videreutvikling av denne siden ved GSI.

Opplysninger om timetall i ulike fag: GSI inneholder ingen informasjon om timetall i ulike fag. Det gjør at det ikke er mulig å undersøke om flere timer i et fag øker prestasjonene i det faget. Opplysninger om timetall i sentrale skolefag, som matematikk, norsk og engelsk ville gjort opplysningene i GSI mindre generelle og dermed mer egnet som kvalitetsindikatorer.

Mer kontekstnær informasjon: Informasjonen i GSI er primært gitt på skolenivå. I en del tilfeller kan ressursindikatorerne knyttes til bestemte klassetrinn. En utvikling av GSI i retning av å knytte deler av den informasjonen som registreres til bestemte klassetrinn, vil gi mer kontekstnær informasjon om de skoleressurser som er tilgjengelige for den enkelte elev.

Anbefalinger med hensyn til videreutvikling av Elevundersøkelsen

Som tidligere nevnt, finner prosjektet gjennomgående sammenhenger mellom elevenes læringsmiljø og undervisningssituasjon, slik det måles i Elevundersøkelsen, og deres skoleprestasjoner. At læringsmiljøet er trivelig, inkluderende, støttende, punktlig og faglig fokusert er positivt knyttet til elevenes prestasjonsnivå. Også kjennetegn ved undervisningen, som omfang av ulike undervisningsformer og omfang av vurdering, er funnet å være knyttet til elevenes prestasjoner på flere trinn i grunnopplæringen. Dette er kjennetegn ved elevenes læringsmiljø og undervisningssituasjon som kan fungere som kvalitetsindikatorer. Når vi likevel er forsiktige med å anbefale hvordan ulike sider ved elevenes læringsmiljø og undervisningssituasjon kan brukes som kvalitetsindikatorer for grunnopplæringen, har det sammenheng med usikkerheter og svakheter ved Elevundersøkelsen som datakilde. Dette er diskutert i kapittel 5. Elevundersøkelsen er en viktig kilde til informasjon om elevenes skolehverdag, men vi mener at potensialet ikke utnyttes fullt ut. Anbefalingene gir forslag til videre utvikling av Elevundersøkelsen, slik at den i større grad enn i dag kan gi informasjon om kvalitet i skolen.

Fagspesifikke spørsmål: I stedet for spørsmål om elevenes oppfatning av ulike sider ved undervisningssituasjonen generelt, som undervisning og evaluering, anbefales det at spørsmålene i større grad knyttes til spesifikke fag. Eksempelvis kan omfanget av ulike evaluerings- og undervisningsformer, variere fra fag til fag. Dette kan medføre at elevene forstår spørsmålene på ulike måter. Et alternativ til denne type generelle spørsmål, er mer fagspesifikke spørsmål som omhandler undervisningssituasjon og læringsmiljø i konkrete fag. I stedet for å formidle en generell oppfatning av undervisningen på skolen vil svarene på slike spørsmål gi innblikk i forskjeller mellom fag. Når spørsmål knyttes til bestemte fag, kan det samtidig åpnes for informasjon om elevenes oppfatning av læreren som kunnskapsformidler og klasseleder. Dersom slike opplysninger sammenholdes med

elevenes prestasjoner i de samme fagene, vil Elevundersøkelsen i større grad belyse sammenhenger mellom elevenes læringsmiljø og prestasjonsnivå.

Redusert omfang: Elevundersøkelsen dekker en rekke ulike tema knyttet til elevenes læringsmiljø. I tillegg til en fast del, bestående av spørsmål som alle får, kan den enkelte skole velge ett eller flere tilleggstema. Selv om antall tilleggsspørsmål/tema er begrenset, vil spørreundersøkelsen oppleves som lang og omfattende for mange elever. Trolig er dette en svakhet ved Elevundersøkelsen som har betydning for elevenes motivasjon og svar kvalitet. Det anbefales derfor å redusere omfanget av undersøkelsen. En mulighet er å innføre rotering av oppgavesett:

Rotering av oppgavesett: Samtidig som undersøkelsen ikke bør bli lengre, har vi her foreslått å stille spørsmål for spesifikke fag. En måte å imøtekomme begge disse behovene kan være å rotere hvilke oppgavesett som benyttes fra år til år, hvor en del består av faste spørsmål og en del består av spørsmål for utvalgte fag. Om man for eksempel velger ut seks sentrale undervisningsfag, kan man stille spørsmål for to fag hvert år. Da kan man følge utviklingen for de ulike fagene med kontrollpunkter hvert tredje år. En slik rotering av oppgavesett har man blant annet i PISA-undersøkelsen, hvor det veksles mellom å ha fokus på lesing, matematikk og naturfag. Videre kan Elevundersøkelsen bli kortere dersom man finner at det også er andre temaer hvor måling hvert tredje år er tilstrekkelig.

Mer nyanserte spørsmål om undervisningsformer: Slik undersøkelsen gjennomføres i dag er det mulig for elevene å svare både at mye av undervisningen er lærerstyrt og at mye av undervisningen er elevstyrt. Dersom spørsmålene konstrueres slik at elevene i større grad må angi hvilke undervisningsformer det er mest av og hva det er mindre av, vil Elevundersøkelsen i større grad få tak i kontraster i omfang og bruk av ulike undervisningsformer.

Vurdere svarkategoriene: Enkelte av spørsmålene har svarkategorier som synes å være for grove til å fange opp den variasjonen som finnes. Selv om det ikke er et mål i seg selv at svarene skal fordele seg jevnt på en skala, bør det være et mål å unngå at alle havner i samme svarkategori. Et eksempel er svarkategoriene i spørsmålene om undervisningsformer der svarfordelingen i noen tilfeller er svært skjev. I tilfeller der svært mange av elevene krysser av i øverste eller nederste svarkategori, bør disse vurderes endret.

Mer informasjon om bakgrunnsforhold: Elevundersøkelsen spør i liten grad om kjennetegn ved elevene eller deres foreldre, som kunne gi informasjon om elevenes sosioøkonomiske bakgrunn eller innvanderbakgrunn. Med utgangspunkt i den sterke sammenhengen mellom elevenes kjennetegn og deres prestasjoner, anbefales det å inkludere noe mer informasjon om individkjennetegn i Elevundersøkelsen.

Endre gjennomføringen: Med dagens ordning vil mange elever delta i Elevundersøkelsen en rekke ganger i løpet av grunnopplæringen. Hyppig deltakelse kan redusere elevenes motivasjon og respekt for undersøkelsen, noe som kan øke omfanget av useriøse svar. En mulig løsning på dette kan være å begrense skolens muligheter til å gjennomføre undersøkelsen for de samme elevene gjentatte ganger.

1 Innledning, bakgrunn og kunnskapsstatus

1.1 Innledning

Dette er sluttrapporten fra prosjektet «Ressursbruk og læringsresultater i grunnopplæringen» (kortnavn: Ressurser og resultater). Prosjektet er gjennomført på oppdrag fra Utdanningsdirektoratet og startet opp i begynnelsen av 2010. I løpet av de vel tre årene som prosjektet har vart, fra 2010 til vinteren 2013, har prosjektet omfattet en rekke analyser presentert i tilsammen seks delrapporter. Formålet med denne rapporten er først og fremst å oppsummere og diskutere hovedfunnene fra de ulike delrapportene, sett i sammenheng, og i lys av resultater fra tidligere norsk og internasjonal forskning.

Vi har kalt sluttrapporten 'Jakten på kvalitetsindikatorene'. I dette ligger et fokus på det som har vært en gjennomgående problemstilling i dette prosjektet: Å undersøke sammenhenger mellom ressursbruk i skolen og elevenes læringsresultater, for gjennom dette å identifisere indikatorer på kvalitet i grunnopplæringen. Prosjektet har hatt som formål å utnytte bredden i de eksisterende datakilder på nasjonalt nivå. Gjennom analyser av et bredt sammensatt datamateriale har hensikten både vært å skape et nyansert bilde av sammenhenger mellom elevresultater og ulike innsatsfaktorer, samt legge grunnlag for utvikling av gode kvalitetsindikatorer for grunnopplæringen.

I dette kapitlet gjennomgår vi først innholdet i prosjektet som helhet, hoveddelen av prosjektet og de to tilhørende delprosjektene – formål, datakilder og teoretisk grunnlag. Deretter tar vi et blick på tidligere studier. Dette er for å gi en oversikt over kunnskapsstatus, samt å sette prosjektet inn i en forskningsmessig sammenheng. Til slutt i kapitlet gjennomgås rapportens videre oppbygging.

1.2 Om prosjektet

Prosjektet har hatt som formål å undersøke sammenhenger mellom ressursbruk i skolen og elevenes læringsresultater. Dette er gjort ved bruk av ulike datakilder, og ved å koble sammen data fra ulike kilder for å kunne analysere dem i sammenheng. En hensikt med analysene har vært å identifisere indikatorer på kvalitet i grunnopplæringen. Oppdraget er besvart gjennom til sammen seks delrapporter. Delrapportene og resultater fra prosjektet presenteres nærmere i neste kapittel.

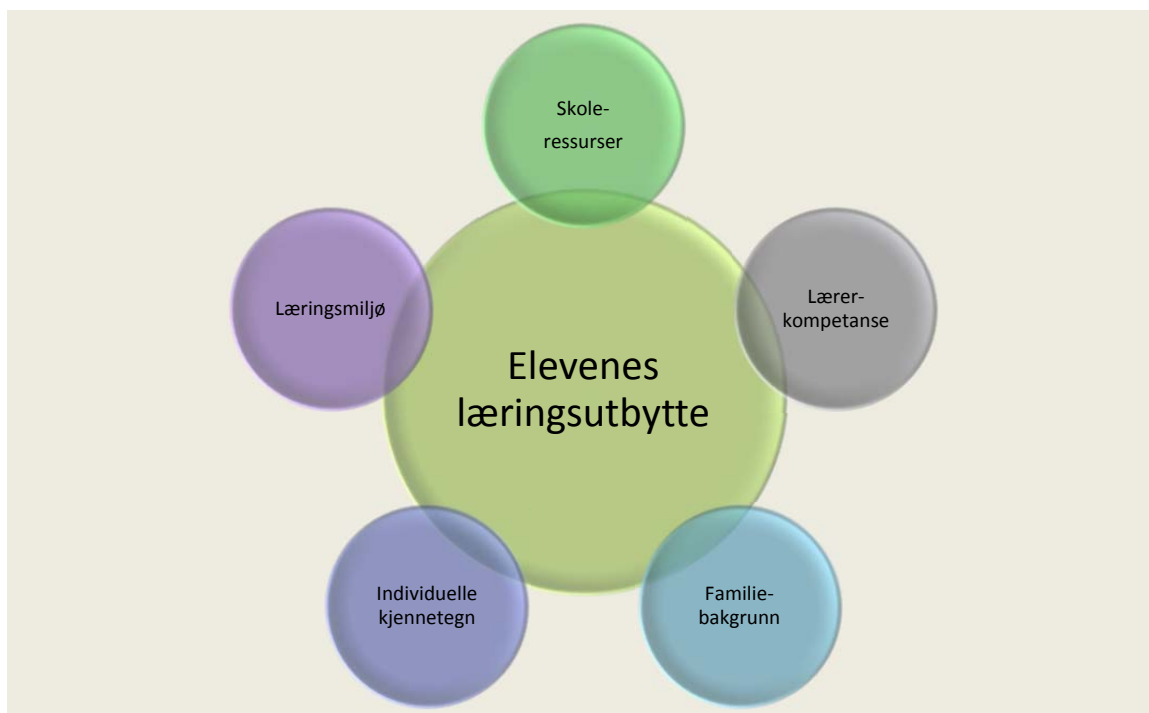
Prosjektet består av et hovedprosjekt og to delprosjekter. Hovedprosjektet innebærer å konstruere datasett bestående av elevkull på ulike trinn over flere år. Datasettet omfatter informasjon om elevene, skolen de går på, elevenes skoleprestasjoner, samt kjennetegn ved deres foreldre. Opplysningene er hentet fra ulike kilder og registre: Elevundersøkelsen, GSI, resultater fra nasjonale prøver, karakterer

ved utgangen av ungdomsskolen, samt befolkningsdata fra Statistisk sentralbyrå. I løpet av prosjektet har datasettet blitt utvidet med informasjon om elevenes prestasjoner på flere tidspunkt, etter hvert som nye data har blitt tilgjengelige. Dette har gjort det mulig å følge elevenes utvikling over tid og å undersøke skolens betydning for elevenes læringsutbytte og læringsutvikling (gjennom *value added*-analyser). Resultatene fra hovedprosjektet er formidlet i fire rapporter; delrapportene 1, 2, 3 og 5 (Opheim mfl 2010, Wiborg mfl 2011, Arnesen 2012, Grøgaard 2012).

Videre har det i prosjektet blitt gjennomført en spørreundersøkelse med elever på ungdomsskolen og videregående som informanter, samt en casestudie av fire skoler i Oslo. Spørreundersøkelsen og analysene av denne utgjør delprosjekt 1, mens casestudien utgjør delprosjekt 2 (se figur 1.2). Formålet med spørreundersøkelsen har vært å fremskaffe mer kontekstnær informasjon om sammenhenger mellom undervisningsformer og læringsutbytte i to konkrete fag. Casestudien har bidratt med kvalitative data fra intervjuer med rektorer og klasseromsobservasjoner. Resultatene fra de to delprosjektene er formidlet i egne delrapporter fra prosjektet, delrapport 4 og 6 (Opheim og Wiborg 2012, Markussen og Seland 2013).

Prosjektet har en bred tilnærming til ressursbegrepet. Ressurser er forstått både som skolens materielle, strukturelle og personmessige ressurser slik det registreres i GSI; som skolens ressurser i form av elevenes oppfatning av eget læringsmiljø slik det fremkommer gjennom den årlige Elevundersøkelsen; og som kjennetegn ved elevene og skolens elevsammensetning.

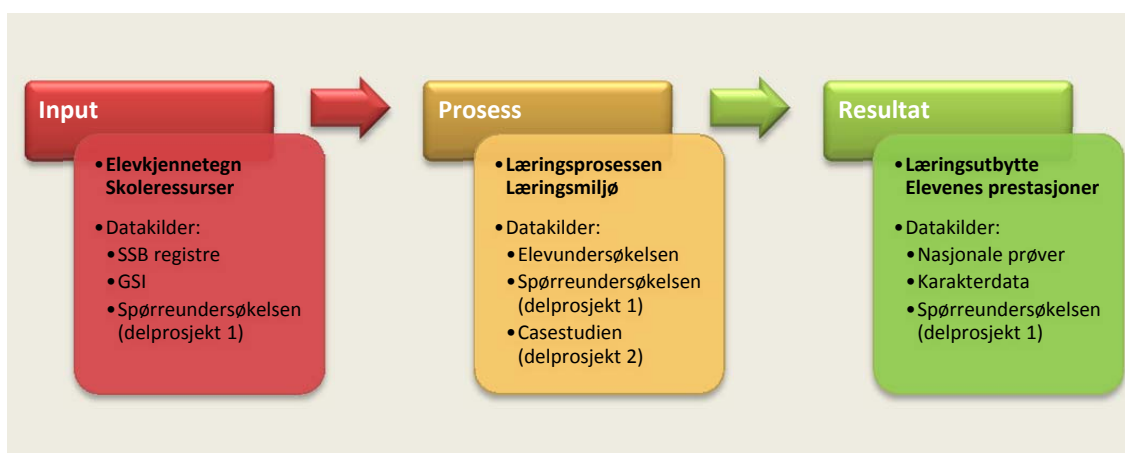
Figur 1.1 nedenfor illustrerer de ulike former for ressurser som analyseres i dette prosjektet. Elevenes læringsmiljø, skoleressurser, individuelle kjennetegn og familiebakgrunn er forhold som inngår i hoveddelen av prosjektet – analyser av registerdata. Kjennetegn ved lærerne og deres kompetanse inngår kun i svært begrenset grad i registerdataene. Dette er derimot blant temaene i spørreundersøkelsen til elever og lærere (delrapport 4) og i casestudiene (delrapport 6). Figuren er hentet fra delrapport 1.



Figur 1.1: Ulike former for ressurser med betydning for elevenes læringsutbytte.

Prosjektet har med andre ord som mål å analysere betydningen av ulike former for ressurser for elevenes læringsutbytte. Nedenfor kommer vi tilbake til differensiering mellom ulike former eller typer skoleressurser. Det finnes flere ulike tilnærminger til læringsutbytte som begrep og fenomen (Prøitz

2010). I denne sammenhengen har vi tatt utgangspunkt i en tradisjonell oppfatning av læringsutbytte. Dette innebærer at elevenes læringsutbytte betraktes som et resultat av en prosess. Både elevens forutsetninger for læring samt ulike typer ressurser/kjennetegn nevnt overfor, vil ha betydning for resultatet eller læringsutbyttet. Dette kan uttrykkes som en såkalt input-prosess-resultatmodell. Dette illustreres i figur 1.2. Input omfatter her alt eleven har med seg inn i læringsprosessen; hver enkelt elevs utgangspunkt og forutsetninger for læring. Her inngår faktorer knyttet både til arv og miljø. Læringsprosessen vil være påvirket av en mengde forhold; kjennetegn ved læringsmiljøet i klassen og på skolen, skolens ressursssituasjon, kjennetegn ved læreren, samspillet mellom læreren, eleven og klassen, og så videre. Når vi ønsker å undersøke hvordan man oppnår et best mulig læringsutbytte (resultat), trenger vi med andre ord mest mulig kunnskap om kjennetegn ved eleven; elevens utgangspunkt og forutsetninger for læring (input), samt informasjon om kjennetegn ved og hvordan læringsprosessen forgår (prosess). Skillet mellom input og prosess er ikke alltid like tydelig. Kjennetegn ved eleven kan for eksempel påvirke læringsprosessen. Det finnes en rekke teorier om ulike forståelser av læring og hvordan elevene lærer. Dette feltet går vi ikke nærmere inn på i denne rapporten (for en beskrivelse av ulike retninger innen læringsteori, se for eksempel Imsen 2006).



Figur 1.2: Input-prosess-resultat-modell (forenklet modell basert på Scheerens 1990).

Figur 1.2 er en forenklet modell basert på Scheerens (1990). I den opprinnelige modellen skilles det mellom input-faktorer og kontekstuelle forhold. Her regnes ulike ressursindikatorer på skolenivå, inkludert kompetanse og erfaring blant skolens lærere, samt støtte fra foreldre, som input-faktorer. Kjennetegn ved elevene, skolens størrelse og beliggenhet regnes som skolens demografiske faktorer og dermed kjennetegn ved skolens kontekst. Skillet mellom kontekst og input kan være utydelig, og begge forhold er antatt å ha en direkte effekt på prosessindikatorerne. Skolens kontekst regnes derfor her blant input-faktorene. Videre skilles det i den opprinnelige modellen mellom prosess-faktorer knyttet til ulike nivåer i skolen: skolenivå og klassenivå. Faktorer på skolenivå omhandler egenskaper ved skolens ledelse, mens faktorer på klassenivå omhandler kjennetegn ved læringsprosessene som forgår i klasserommet (Scheerens 1990). Dette er også forhold det er vanskelig å skille fra hverandre (handler forventninger til elevene bare om kjennetegn ved læreren og det som foregår i klasserommet, eller kan det også være uttrykk for en resultatorientert skoleledelse?). I casestudien diskuteres dette skillet, gjennom analysene av intervju med skolens leder og diskusjon av innholdet og kvaliteten på timene ved de ulike skolene.

Begrepet kvalitetsindikator er i norsk sammenheng mest brukt innenfor helsesektoren (se for eksempel Helsedirektoratet 2010). Selv om det er mange forskjeller mellom helsesektoren og skolesektoren, kan det likevel være nyttig å se hvordan dette begrepet diskuteres i en annen sammenheng enn knyttet til skolen og elevenes læringsutbytte. Selv om vi på ingen måte vil likestille elever med pasienter eller lærere og skoleledelse med helsepersonell, kan det likevel være nyttig å se hvordan det innen helsesektoren skilles mellom ulike grupper av kvalitetsindikatorer. I helsesektoren skilles det også mellom kvalitetsindikatorer innen struktur, prosess og resultat. Strukturindikatorer

beskriver rammer og ressurser, herunder de ansattes (helsepersonellens) kompetanse og tilgjengelighet til utstyr, teknologi og fasiliteter. Indikatorene beskriver med andre ord forutsetningene og rammene for det arbeidet som gjøres. Prosessindikatorer beskriver derimot konkrete aktiviteter og innholdet i det som gjøres i behandling av pasienter. Resultatindikatorer belyser pasientens gevinst – altså resultatene av behandling (Helsedirektoratet 2010:4). Denne inndelingen har med andre ord flere likheter med den vi har brukt i dette prosjektet.

1.2.1 Enkle og komplekse skoleressurser

En måte å skille mellom ulike former for ressurser i skolen er å ta utgangspunkt i hvor lett tilgjengelig og målbare de ulike ressursene er (Grubb 2009). Han skiller mellom fire former for ressurser i skolesektoren:

- Enkle ressurser: lærertetthet/klassestørrelse, timetallsnivå/fordeling og læreres erfaring og formelle utdanning og kompetanse
- Sammensatte (compound) ressurser: organisatoriske tiltak, læreres planleggingstid av undervisningen og lærere som underviser i fag som de er utdannet til
- Komplekse ressurser: pedagogiske tilnærminger og den faktiske/effektive undervisningstiden (motsetning til regulerte timetallet)
- Abstrakte ressurser: skolekultur og stabilitet i organisasjonen og institusjonelle omgivelser.

Årsaken til at så mange empiriske studier i beste fall finner små effekter av økonomiske ressurser, henger ifølge Grubb sammen med at penger bare kan kjøpe enkle ressurser. Men det er de sammensatte, komplekse og abstrakte ressurser som har klare og store effekter på elevprestasjoner, læringsmiljø og andre skolerestater. Hovedforklaringen er at enkle ressurser i liten grad påvirker undervisningen til lærerne (Grubb 2009 i Kunnskapsdepartementet 2011).

Statlige utdanningsmyndigheter står dermed overfor et styringsparadoks: Det de kan påvirke lettest, har liten effekt på skolerestater, mens det som har størst effekt på skolerestater, er det vanskelig å påvirke gjennom statlige virkemidler (Kunnskapsdepartementet 2011).

Denne forståelsen av sammenhengen mellom ressurser og skolerestater samsvarer for øvrig i stor grad med en annen studie som har fått mye oppmerksomhet i de senere årene. I rapporten '*Visible Learning*' sammenfatter og diskuterer John Hattie en mengde studier av ulike faktorer som kan påvirke elevens skoleprestasjoner (Hattie 2009). En hovedkonklusjon fra rapporten er betydning av lærerens innsats for å fremme elevenes prestasjoner, særlig ved å 'synliggjøre' læringsprosessen (jæmfør tittelen) samt å tydeliggjøre læringsmålene for sine elever. Dette omfatter blant annet at læreren har høye forventninger til alle elevgrupper, har tydelige læringsmål, og gir hyppige tilbakemeldinger underveis i opplæringen.

I dette prosjektet gir data fra GSI først og fremst indikatorer på skolens bruk/omfang av *enkle ressurser*. I følge Grubbs teorier, er det lite man kan forvente å finne av sammenheng mellom disse formene for ressurser og elevenes resultater (Grubb 2009). Ved å sammenligne skolens ressurssituasjon over tid, gir GSI imidlertid også informasjon om det Grubb definerer som *abstrakte ressurser*. Dette omfatter blant annet grad av *stabilitet* i skolens ressurssituasjon, og ettersom GSI inneholder informasjon om årlig ressursbruk og ressurssituasjon ved den enkelte skole, åpner dette for analyser av endring og stabilitet over tid. Ved å inkludere data om elevenes læringsmiljø fra Elevundersøkelsen i analysene, vil analysene også omfatte indikatorer som ifølge Grubbs inndeling må regnes som *abstrakte ressurser*. Derimot gir analysene av registerdataene lite informasjon om skolens sammensatte og komplekse ressurser. Gjennom spørreundersøkelsen til elever og lærere samt casestudien av utvalgte skoler, er hensikten å fange opp mer av dette.

1.3 Tidligere studier

Prosjektet omhandler flere sentrale tema innen utdanningsforskningen. Både sammenhenger mellom ressursbruk i skolen og elevenes læringsutbytte, betydning av læringsmiljø og hva som skjer i klasserommet for elevenes læring, samt sammenhenger mellom kjennetegn ved elevene og deres læringsutbytte er store forskningstema. Det første tema tilhører i høy grad det man gjerne kaller *skoleeffektivitetsforskningen*. Et sentralt spørsmål her er hvilke ressurser – både ressursnivå og ressursformer – som kjennetegner skoler der elevene oppnår gode resultater. Her finner vi mye norsk og internasjonal forskning, kanskje i hovedsak innenfor utdanningsøkonomi. Når det gjelder studier av læringsmiljø og hva som skjer i klasserommet, er dette et sentralt tema innen den pedagogiske *klasseromsforskningen*. I motsetning til det første tema, er dette feltet preget av kvalitative analysemetoder og casestudier av læringssituasjoner ved skoler og klasserom. Det tredje feltet – sammenhenger mellom kjennetegn ved elevene og deres læringsutbytte – er i likhet med det første tema dominert av kvantitative analysemetoder. Dette er et betydelig felt innenfor utdannings sosiologien hvor mye av forskningen har fokusert på *utjevning av forskjeller i læringsutbytte* mellom ulike grupper av elever. Prosjektet dekker med andre ord ulike tema som er forankret innenfor ulike fagfelt – både pedagogikk, økonomi og sosiologi.

I det følgende vil vi gjennomgå en del sentrale tidligere studier innenfor disse ulike temaene. Vi avgrensner oss til studier som særlig er relevante for problemstillinger og analyser i dette prosjektet. Dette vil primært være norske empiriske studier, men også sentrale internasjonale studier inngår i gjennomgangen. Vi begynner med en gjennomgang av skoleeffektivitetsforskningen. Deretter går vi over til klasseromsforskningen, for til slutt å se nærmere på studier av kjennetegn ved elevene og deres læringsutbytte.

1.4 Skoleeffektivitetsforskning

Skoleeffektivitetsforskningen er den retningen innen skoleforskningen som mest eksplisitt er knyttet til sammenhenger mellom skolens ressursbruk og elevenes resultater. Effektivitetsforskningen har studert hva som kjennetegner skoler og undervisning som bidrar til bedre enn forventet læringsframgang hos elever (Ogden 2005).

En retning i den internasjonale forskningen på skoleeffekter har primært vært opptatt av å vurdere om elevers læringsutbytte, slik dette måles ved standardiserte tester eller fagkarakterer og poengsummer, påvirkes av innsatsfaktorer i skolen som kan knyttes til bruk av økonomiske ressurser i skolen (bevilgninger/økte utgifter). Studier av sammenhenger mellom ressursbruk og skolerresultater har blitt sammenlignet med jakten på «den hellige gral» (Stiefel mfl 2005). Formålet har ikke bare vært å påvise sammenhenger, men å etablere *årsakssammenhenger* mellom ressursbruk og skolerresultater. Altså å kunne si noe sikkert om effekten av bestemte former for ressurser for elevenes prestasjoner. Dette er med andre ord ingen lett oppgave, men mange har forsøkt.

Helt siden de omfattende empiriske studiene av elevers læringsutbytte i USA på 1960- og 70-tallet (Coleman mfl 1966, Jencks mfl 1972) konstaterte at skolens betydning for elevenes utvikling var marginal, har dette vært et aktivt forskningsfelt. Studiene fant at skolen ikke bare reproduserte, men også forsterket tendensene til sosial ulikhet i det amerikanske samfunnet. Man fant liten sammenheng mellom undervisningens kvalitet og elevenes læringsutbytte, men man fant effekter av skolens elevsammensetning. Konklusjonene herfra om at «familien betyr alt» eller «skoler og lærere gjør ingen forskjell» var antagelig påvirket av at studien ikke tok utgangspunkt i direkte elev-lærer relasjoner, men sammenlignet elevenes gjennomsnittlige prestasjoner på skolenivå med skolens «gjennomsnittslærere». Likevel oppsto det en gryende optimisme for at skolen kunne bli mer egalitær og nå ut til de marginale elevene (Ogden 2005). Etter hvert rettet forskningen også oppmerksomheten mot kjennetegn ved effektive skoler og hvordan disse bidro til større enn forventet læringsframgang hos sine elever. Forskerne fant blant annet at læringsutbyttet varierte med hvor mye elevene ble eksponert for skolefagene og hvor mye tid de fikk til å lære (*time on task*) (Rutter mfl 1979, Brookover 1979).

Ved å forbedre designet fra Coleman-undersøkelsen påviste økonomene Summers og Wolfe (1975, 1977) effekter av lærerkvalitet på elevenes prestasjonsnivå. Med utgangspunkt i reelle lærer-elev-relasjoner påviste disse økonomene at kvaliteten på lærerutdanningsinstitusjonen var positivt assosiert med elevenes ferdighetsutvikling (*value added*) – alt annet likt. Studiene av Summers og Wolfe (1975, 1977) representerer samtidig et veiskille i amerikansk skoleforskning, hvor fokus gradvis ble flyttet nærmere primæraktivitetene i skolen; mot mer prosessnære studier av elevenes læringsutbytte.

Internasjonalt er det publisert en rekke «metastudier» som oppsummerer forskningen på dette feltet. Tidligere internasjonale studier av forskning på læringseffekter av økonomiske ressurser konkluderer gjerne med at ekstra ressursinnsats ikke synes å bedre elevenes læringsutbytte. En tolkning av disse empiriske resultatene er at skoler generelt, kanskje særlig de offentlige skolene, underutnytter ressursene sine (Hanushek 1986, 1997, Hanushek og Rivkin 2006, Hanushek og Vössmann 2007). Gustavsson (2003) gir en annen metaanalyse som diskuterer mye av den samme litteraturen, og her er konklusjonen at Hanushek og kollegers tolkning er for pessimistisk. Analyser av økt ressursinnsats i en rekke OECD-land dokumenterer positive læringseffekter av økte økonomiske ressurser i skolen. Særlig hvis man tar hensyn til at økonomiske ressurser benyttes selektivt, de settes ofte inn der behovene er størst, kan tolkningen av forskningsresultatene være at ekstra ressursinnsats reduserer muligheten for svært svake resultater i skolen. Et problem i de klassiske oversiktsstudiene er at man blander sammen direkte og indirekte bruk av ressurser.

Enkelte studier har likevel forsøkt å isolere effekten av mer ressurser til skolen ved å skille mellom ressurser knyttet til elevenes behov og andre typer – såkalte «frie» – ressurser. Dette er blant annet gjort ved å sammenligne tilstanden i skolekretser med betydelige ressurser (norske kraftkommuner) og tilstanden i landet for øvrig. Resultatene indikerer at en ekstra ressursinnsats har positive virkninger for elevene (Hægeland mfl 2008).

Et annet sentralt tema innen dette forskningsfeltet er betydning av *klassestørrelse eller antall elever per lærer* ved skolene for elevenes prestasjoner. Er det slik at mindre klasser gir bedre læringsutbytte? Internasjonal forskning viser ingen entydig sammenheng mellom skole- og gruppestørrelse og elevresultater, enten dette måles som gjennomsnittlig elevtall i skoler og elevgrupper eller som forholdstall mellom elever og lærere. En begrensning ved denne forskningen har vært at ressursituasjonen har variert lite fra skole til skole. Derfor har det vært vanskelig å undersøke hvilken betydning en omfattende økning eller reduksjon av ressurser ville ha (Ogden 2005). Effekten av økt ressursbruk er neppe universell, men vil blant annet avhenge av undervisningsorganisering og ledelse ved den enkelte skole, samt av rammebetingelser fastsatt av kommuner og nasjonale myndigheter (Borge og Rønning 2009).

Heller ikke norske studier viser noen entydig sammenheng mellom klassestørrelse og elevenes læringsutbytte (Borge og Rønning 2009). Funnene er sprikende, men enkelte studier har påvist en sammenheng mellom elevers sosioøkonomiske bakgrunn, klassestørrelse og læringsutbytte (Bonesrønning 2010, Bonesrønning og Iversen 2010). Resultatene tyder på at elever med lav sosioøkonomisk bakgrunn har størst utbytte av å gå i små klasser. Sammenhengen mellom elevers sosioøkonomiske bakgrunn, klassestørrelse og læringsutbytte er til stede både på 4. trinn (Bonesrønning 2010) og på 7. trinn (Bonesrønning og Iversen 2010).

Et annet omdiskutert tema er betydning av *lærerkvalitet/kjennetegn ved læreren* og elevenes prestasjoner. I stedet for å studere antallet lærere per elever ved den enkelte skole, ser disse studiene på (målbare) kjennetegn ved lærerne. Dette er i hovedsak lærernes formelle kompetanse, kjønn, alder og yrkeserfaring. En studie av sammenheng mellom lærernes kompetanse og elevenes prestasjoner, viste for eksempel at en høy andel lektorutdannede lærere på skolen på ungdomstrinnet, hadde positiv effekt på elevenes prestasjonsnivå i matematikk (Falch og Naper 2008). En slik effekt av

lærernes formelle utdanning kan indikere at bestemte opplæringsprogrammer som øker lærernes formelle (og reelle) kompetanse, gir positive læringseffekter på sikt.

En norsk studie av sammenhenger mellom læreres sannsynlighet for å slutte/bytte skole, tydet på at det foregår systematisk sortering av lærere mellom skoler. Lærere ved skoler med høy andel minoritets elever og/eller høy andel elever med spesialundervisning etter enkeltvedtak, hadde større sannsynlighet for å slutte enn andre lærere. Dette kan indikere at nivået på lærerressursene og lærerkompetansen varierer med elevsammensetning og skolestørrelse i den norske grunnskolesektoren. Analysene kan derimot ikke gi sikkert svar på om dette gir systematiske forskjeller i lærerkvalitet mellom skolene. Det er grunn til å tro at det er slik, men det påpekes et behov for mer forskning på dette området (Falck og Strøm 2005).

1.5 Klasseromsforskning

Klasseromsforskning er forskning på læring og undervisning knyttet til fag, læringsaktiviteter og læringsressurser i klasserommet. Denne forskningsretningen har tradisjonelt fokusert på prosesser i klasserommet og samspill mellom læreren og elevene. Klasseromsforskning bør ifølge Haug (2011:4) forstås vidt, som forskning på pedagogisk praksis eller prosessene i undervisning og læring der dette skjer – i eller utenfor klasserommet – så lenge det foregår i regi av skolen. I motsetning til effektivitetsforskningen, har klasseromsforskningen i all hovedsak en *kvalitativ* tilnærming til analyser av hva som foregår på skolen og i klasserommet, og sammenhenger mellom skolens ressurser og elevenes læringsutbytte. Tatt i betraktning hovedfokus i dette forskningsfeltet, er det ikke unaturlig at den mest utbredte metoden her består av *observasjon* i klasserom. Sentrale studier fra dette feltet ble gjennomgått i delrapport 4 (se Opheim og Wiborg 2012, kapittel 1). Deler av denne forskningen presenteres også i kapittel 3.

Også i dette feltet er man opptatt av hva som bidrar til god læring og prestasjonsutvikling blant elevene. Der stopper imidlertid de fleste likheter. I klasseromsforskningen undersøkes dette primært ved å gå inn i klasserommene og observere hva som foregår. Dette omfatter studier av elever, lærer og spillet mellom de ulike aktørene i klasserommet. Temaene dekker tidsbruk, arbeidsmåter, både hva som gjøres og hvordan det gjøres, aktivitetsnivå i ulike timer, læringsstrategier, læringsmiljø, rolleforståelse og -utøvelse blant elever og lærere. Flere studier er også opptatt av å analysere endringer over tid.

Studier av arbeidsformer og læringsaktiviteter viser mange stabile mønstre, men også enkelte endringer i hva som foregår i klasserommet. Blant annet viser enkelte studier at omfanget av «arbeid med oppgaver» på individuell basis har økt i løpet av de siste 10-20 årene. Tidsbruken anvendt til rene lærermonologer er redusert, og «spørsmål-svar-sekvensene» synes å være mer preget av dialog enn tidligere studier antyder. Når det gjelder innholds- og evalueringsaktiviteter peker så vel eldre som nyere studier mot høy grad av lærerstyring og med liten grad av elevinnflytelse, selv om omfanget av elevaktive undervisningsformer synes å ha blitt gradvis større:

«Tidligere studier viste hvordan klasserommet er lærerdominert og lærerstyrt og hvor lærer snakker og har ordet gjennomgående to tredeler av tiden. Slår vi sammen registrerte lærersentrerte aktiviteter og sammenligner dem med andelen individuelle og gruppeorienterte aktiviteter, dominerer fremdeles lærersentrerte aktiviteter tidsbruken i klasserommet. Elevaktive arbeidsformer har imidlertid fått et noe større gjennomslag enn tidligere studier og 2/3-regelen synes å antyde» (Klette 2003:60).

Tilsvarende finner Haug (2011) i sin oppsummering av klasseromsforskning, to hovedtendenser. Det ene er en tendens til gradvis endring av arbeidsformene i klasserommet. Den andre er en tendens til stor variasjon i arbeidsformene mellom ulike fag og i ulike klasserom. Haug peker på større samfunnsmessige endringer som påvirker skolens kontekst, som utviklingen av mediesamfunnet og omlegging av nasjonal og kommunal forvaltning. Skolens handlingsrom er blitt større, skriver han, noe

som også kan forklare den variasjonen han finner mellom skoler og klasser. En generell utvikling er likevel at omfanget av individuelle arbeidsoppgaver har økt, mens omfanget av kollektiv undervisning ser ut til å ha gått ned:

«Når vi ser undersøkningane over tid, stig omfanget av individuelt arbeid på kostnad av felles kateterundervisning. Det er ei stor endring. Fleire av dei nyare undersøkningane tyder på at det no i gjennomsnitt er ein tilnærma balanse mellom lærarstyrte og elevaktive arbeidsformer» (Haug 2011:12).

Klasserommets aktiviteter og lærerens undervisningsformer har med andre ord vært et sentralt tema over lengre tid innen dette forskningsfeltet (Klette 2003, Imsen 2004, Lyng 2004 og Haug 2011). Mye av denne forskningen handler om betydningen av den tydelige, tilstedeværende og ledelsesorienterte læreren som en motsats til en undervisningssituasjon som preges av elevenes egne valg og ansvar for læring. Elevpassivitet kan være en forklaring på svake skolefaglige prestasjoner, noe som kan ses i forbindelse med den tilbaketrukne lærerrollen. Samtidig har det ikke vært mulig å dokumentere gjennom den norske forskningen at en metode for undervisning generelt er bedre enn en annen (Haug 2011). Studiene har gitt mye informasjon om selv læringsprosessen – hva som foregår i klasserommet, lærerens og elevenes ulike roller, men resultatene har i begrenset grad blitt knyttet til elevenes læringsutbytte. Funn fra klasseromsforskningen er også diskutert i delrapport 4 og 6 (Opheim og Wiborg 2012, Markussen og Seland 2013).

1.6 Ulikhetsforskning

Ulikhetsforskning utgjør et betydelig felt innenfor utdannings sosiologien. Feltet er bredt og dekker studier av ulikhet i ulike deler av utdanningssystemet, samt ulikhet mellom ulike grupper. Den delen av feltet som er særlig relevant for dette prosjektet, omhandler sammenhenger mellom kjennetegn ved elevene og deres læringsutbytte. Feltet er i likhet med skoleeffektivitetsforskningen dominert av kvantitative analysemetoder, og har primært fokus på sammenhenger mellom *input* og *resultat*. Input er her mer knyttet til kjennetegn ved elevene og deres familiebakgrunn, og i liten grad til ressurser på skolenivå. Formålet har som regel ikke vært å isolere betydningen av spesifikke skoleressurser slik som i effektivitetsforskningen, men heller å analysere samspill mellom ulike ressurser og kjennetegn ved skolen og elevsammensetningen.

Et sentralt spørsmål innenfor utdannings sosiologien er om ulike grupper av elever oppnår like eller ulike skolerestater. Dette omfatter studier av kjønnsforskjeller i utdanningssystemet, prestasjonsforskjeller mellom elever med innvandrerbakgrunn og majoritets elever, samt sammenheng mellom foreldrenes utdanningsnivå og elevenes prestasjoner. I disse studiene er man også opptatt av forhold som kan bidra til å *jevne ut* prestasjonsforskjeller mellom ulike grupper av elever (og hvilke forhold som øker forskjellene). Mange av studiene omfatter å sammenligne prestasjoner blant ulike elevgrupper over tid, gjerne før og etter konkrete endringer i skolen, politiske tiltak, skolereformer, etc. Nedenfor oppsummeres hovedfunn og sentrale forklaringer i studier av sammenheng mellom foreldrenes utdanningsnivå og elevenes prestasjoner, kjønnsforskjeller i utdanningssystemet, og prestasjonsforskjeller mellom elever med innvandrerbakgrunn og majoritets elever.

Foreldres utdanningsnivå, familiesituasjon og engasjement

Studier av sosiale forskjeller i utdanningssystemet finner gjennomgående en sterk og stabil sammenheng mellom foreldrenes utdanningsnivå og elevenes prestasjoner. Elever med foreldre med høyt utdanningsnivå presterer bedre enn elever med foreldre med lav utdanning. Et sentralt spørsmål her er hvilke forhold som kan bidra til å redusere betydningen av elevenes sosiale bakgrunn.

I et nylig avsluttet prosjekt som del av evalueringen av Kunnskapsløftet ledet av NOVA, har målet vært å vurdere om reformen bidrar til å redusere sosiale forskjeller i læringsutbytte på ungdomstrinnet basert på foreldrenes utdanningsnivå, elevers kjønn og minoritetsstatus (prosjektet ble presentert i delrapport 1). Prosjektet har flere likhetstegn med dette prosjektet: Flere av de samme datakildene

brukes (karakterdata, GSI, Elevundersøkelsene), samt flere av de samme analysemetodene benyttes, inkludert flernivåanalyse og en kvalitativ casestudie. Prosjektet analyserer elevenes prestasjoner og prestasjonsutvikling før og etter innføringen av Kunnskapsløftet.

Analyser av elevenes prestasjonsutvikling, tyder på at de sosiale forskjellene synes å forsterke seg utover i skolesystemet. Det er med andre ord lite som tyder på at skolen klarer å utjevne de sosiale forskjellene mellom elevene, til tross for at dette er en politisk målsetting i Norge (Bakken og Elstad 2012). Resultatene viser at skolens læringsmiljø og læringskultur kan ha betydning både for utjevning av sosial ulikhet og for utjevning av kjønnsforskjeller i elevenes prestasjoner (Bakken 2010).

Kjønnsforskjeller i utdanningssystemet

Når det gjelder elevenes kjønn som forklaring på forskjeller i skoleprestasjoner, har det vært en utvikling i forskernes fokus og funn som tidligere handlet om at jenter fikk mindre oppmerksomhet i undervisningen enn gutter, til at gutter gjør det dårligere på skolen enn jenter (Bakken og Danielsen 2011). Naturligvis er det mange gutter som får gode skolerresultater, men som gruppe gjør jentene det bedre enn guttene. I gjennomsnitt har jentene bedre skolerresultater enn guttene (se for eksempel Steffensen og Ziade 2009; Bonesrønning og Iversen 2010, Bakken og Elstad 2012, Vibe mfl 2012). Dette mønsteret er gjennomgående i alle vestlige land (OECD 2012). Kjønnsforskjellene varierer imidlertid mellom fag og typer oppgaver. Analyser av norske resultater i PISA 2009 finner små kjønnsforskjeller i naturfag og matematikk, mens relativt store kjønnsforskjeller i lesing. Gutter i Norge leser betydelig dårligere enn jenter, og forskjellen er stor i en internasjonal sammenheng (Kjærnsli og Roe 2010).

Forklaringsmodeller på hvorfor gutter og jenter presterer ulikt på skolen og har ulik prestasjonsutvikling, er hentet både fra samfunnsvitenskapen og fra biologien – og samspillet mellom biologi og miljø. Flere forklaringsmodeller fra samfunnsvitenskapen tar utgangspunkt i skolens og lærerens rolle, og hvordan gutter og jenter kan respondere ulikt og ha behov for ulike undervisningsformer.

Sentralt i samfunnsdebatten står tesen om såkalt «feminisering» av skolen, og som dreier seg om at skolen representerer normer for språk og omgangsformer hvor jenter føler seg hjemme og gutter har vanskeligere for å innordne seg. Teorien har mange likheter med teorier om skolens rolle i å reprodusere sosiale forskjeller i elevenes prestasjoner – at elever fra høyere sosiale lag føler seg mer 'hjemme' på skolen enn elever fra lavere lag. Tilhengerne av slike oppfatninger har pekt på at en overvekt av *kvinnelige lærere*, slik vi særlig finner på de laveste årstrinnene i barneskolen, kan skape en marginalisering av gutter tidlig i utdanningsløpet. Enkelte studier har funnet en positiv sammenheng mellom andel kvinnelige lærere og elevenes prestasjoner (Falch og Naper 2008, Bakken 2009b, Bonesrønning og Iversen 2008, 2010). Ulike tolkninger av denne sammenhengen omtales i delrapport 1 (Opheim mfl 2010). En gjennomgang av forskningslitteraturen om skolens bidrag til kjønnsforskjeller i skoleprestasjoner, finner i liten grad studier som konkluderer med at skolen i seg selv *skaper* kjønnsforskjeller i skoleprestasjoner (Bakken mfl 2008). Tidligere studier peker derimot i retning av at skolen bidrar til å *reprodusere* kjønnsforskjeller som allerede eksisterer i klasserommet.

Minoritetsspråklige elever i utdanningssystemet

Elever med minoritetsspråklig bakgrunn oppnår i gjennomsnitt svakere skolerresultater enn elever med majoritetsbakgrunn. Samtidig er det store forskjeller i resultatene mellom ulike grupper av elever med minoritetsspråklig bakgrunn (Bakken 2010, Støren og Helland 2010). Elever fra enkelte innvandrergupper får jevnt over vel så gode karakterer som elevene uten innvandringsbakgrunn, men de fleste med innvandringsbakgrunn har et karakternivå under gjennomsnittet (Bakken og Elstad 2012:111).

En nærliggende forklaring handler om språkferdigheter og begrepsforståelse. Motivasjon for skolen synes imidlertid ikke å være lavere hos elevene med minoritetsspråklig bakgrunn enn hos majoritetselever. Utdanning har også blitt betegnet som en strategi for sosial mobilitet for elever og studenter med innvandrerbakgrunn. Ønsket om å skape seg en bedre tilværelse har derfor blitt foreslått som drivkraft bak de gode prestasjonene som man også kan finne innenfor denne befolkningsgruppen. Drivkraften kan handle om at foreldrene håper på en bedre fremtid for sine barn enn den bakgrunnen og den økonomiske situasjonen de selv kan tilby, og at foreldre derfor oppfordrer barna sine til å satse på skolen (Gibson og Ogbu 1991; Lauglo 1996; Modood 2004; Leirvik 2010; Fekjær og Leirvik 2011).

Bildet av minoritetselevenenes skoleprestasjoner synes dermed å være todelt. På den ene siden klarer mange minoritetselever seg bedre enn hva en skulle forvente ut fra deres sosiale bakgrunn (svært mange begynner for eksempel i høyere utdanning), på den annen side er minoritetselever klart overrepresentert blant dem som enten ikke fullfører videregående opplæring eller fullfører med svake karakterer (Støren 2010).

1.7 Oppsummering

Overfor har vi gitt et kortfattet bilde av tidligere forskning og kunnskapsstatus innen de områdene av utdanningsforskningen som prosjektet dekker. Prosjektet dekker tema som er forankret innenfor ulike fagfelt – både pedagogikk, økonomi og sosiologi. Det innebærer stor variasjon i de metodene, tilnærmingene, analyser og forklaringsmodeller som er brukt i de ulike studiene.

Prosjektets bidrag til den eksisterende forskningen er flere. Et bidrag er å presentere nye analyser av elevers prestasjoner på flere trinn i grunnopplæringen, samt analyser av sammenhenger mellom ulike former for skoleressurser og elevenes prestasjoner. Prosjektet bidrar også med nye analyser av elevenes prestasjonsutvikling, både gjennom grunnskolen og fra grunnskolen og over i videregående opplæring. Videre bidrar prosjektet med analyser av en spørreundersøkelse av elever og lærere på to trinn: 9. trinn og i Vg1. Spørreundersøkelsen inneholder informasjon om elevenes og lærernes oppfatninger av undervisningsformer og læringsmiljø i to konkrete fag: norsk og matematikk. Dette gir mer kontekstnær informasjon om læringsprosessen i klasserommet og de reelle elev-lærer relasjonene i klasserommet enn det som er tilgjengelig med bruk av registerdata. I tillegg til spørreundersøkelsen, bidrar prosjektet med nye data innsamlet fra en casestudie av fire ungdomsskoler i Oslo. Gjennom observasjoner av skoletimer samt intervju med skoleleder, gir casestudien data som bidrar til diskusjonen av de kvantitative analysene. Men casestudien bidrar også med nye analyser av sammenhenger mellom det som foregår i undervisningsøktene, hva som kjennetegner 'den gode timen', og kjennetegn ved skolen som helhet og skolens pedagogiske plattform. I denne sluttrapporten skal vi presentere og diskutere resultatene fra de ulike deler av dette prosjektet.

1.8 Rapportens videre gang

I neste kapittel oppsummeres hovedfunn fra de fem delrapportene fra prosjektet, samt tilleggsanalyser presentert i eget notat. Deretter diskuteres sentrale funn i lys av tidligere forskning på feltet.

I de påfølgende kapitler vil vi på grunnlag av resultatene, diskutere de sentrale datakildene som brukes i prosjektet. Utgangspunktet er å diskutere i hvilken grad informasjon fra disse datakildene kan brukes som kvalitetsindikatorer for grunnopplæringen. Indikatorer for kvalitet er her knyttet til elevenes prestasjoner på de nasjonale prøvene, samt elevenes karakternivå. Vi begynner derfor gjennomgangen av de sentrale datakildene med å se nærmere på hvordan elevenes prestasjoner måles i de nasjonale prøvene. Dette er tema i kapittel 3. I hvilken grad fungerer elevenes prestasjoner på de nasjonale prøvene som en kvalitetsindikator ved skolen?

I analysene i prosjektet fremkommer det tydelig at elevenes prestasjoner i høy grad er knyttet til kjennetegn ved elevene selv og deres familiebakgrunn. Et sentralt fokus i analysene har derfor vært å undersøke skolens bidrag til elevenes prestasjoner når man tar høyde for forskjeller mellom elevene, samt forskjeller i elevsammensetningen ved skolen. I kapittel 4 ser vi nærmere på data fra GSI og hvordan informasjon herfra kan brukes som kvalitetsindikatorer for grunnopplæringen. I kapittel 5 tar vi utgangspunkt i hovedfunn fra analyser av Elevundersøkelsen, samt diskuterer styrker og svakheter med å bruke informasjon fra Elevundersøkelsen som kvalitetsindikatorer for grunnopplæringen. I kapittel 6 ser vi nærmere på spørreundersøkelsen til elever og lærere, presentert i delrapport 4. Her diskuteres styrker og svakheter ved denne datakilden og hvordan spørreundersøkelsen har gitt mer kontekstnær informasjon om hva som bidrar til elevenes læringsutbytte. Til slutt i kapittel 7 oppsummeres prosjektets bidrag til kunnskapsstatus. Deretter oppsummeres anbefalingene for utvikling av kvalitetsindikatorer for grunnopplæringen. Avslutningsvis foreslår vi noen områder og temaer for videre forskning innen dette feltet.

2 Hovedfunn fra tidligere delrapporter

Som presentert i kapittel 1, består prosjektet «Ressurser og resultater» av et hovedprosjekt samt to delprosjekter, med resultater presentert i flere delrapporter. I alt er det publisert seks delrapporter fra prosjektet. Dette kapitlet består av to deler. Første del gir en kort presentasjon av hver av delrapportene. Presentasjonen legger vekt på rapportenes hovedtema, datamaterialet som er benyttet, framgangsmåte og analyser. Dette gjøres kronologisk – fra delrapport 1 til 6. Deretter presenteres prosjektets hovedresultater på tvers av delrapporter. Denne delen er tematisk inndelt, ettersom flere av delrapportene omhandler og belyser den samme tematikken: Hvilke forhold på skole- og individnivå som har betydning for elevers skoleprestasjoner. De seks delrapportene tilnærmer seg temaet på forskjellige måter, og ved hjelp av forskjellige typer data. Fire av rapportene er basert på registerdata, en er basert på en spørreundersøkelse blant elever og lærere på ungdoms- og videregående trinnet, og en presenterer resultater fra en casestudie av fire skoler. I tillegg til de seks delrapportene er det også utarbeidet et tilleggsnotat hvor noen av funnene i delrapport 1 analyseres nærmere. Dette tilleggsnotatet inngår i presentasjonen av delrapport 1.

2.1 Prosjektets delrapporter

Nedenfor presenteres hver av delrapportene fra prosjektet «Ressurser og resultater» i kronologisk rekkefølge.

2.1.1 *Delrapport 1: Elevenes prestasjonsnivå på tre trinn i grunnskolen*

Den første delrapporten fra prosjektet kom høsten 2010 med tittelen: *De gamle er eldst? Betydningen av skoleressurser, undervisningsformer og læringsmiljø for elevenes prestasjoner på 5. 8. og 10. trinn i grunnopplæringen* (Opheim mfl 2010). Tema for rapporten var sammenhenger mellom skolens ressurser – i utvidet forstand – og elevenes prestasjoner. Det utvidede ressursbegrepet omfatter en rekke kjennetegn ved skolen, fra strukturelle forhold til kjennetegn ved elevsammensetningen ved skolen, samt fordeling av lærerressurser, timetall og annet. Skolens ressurser omfatter også kjennetegn ved skolens læringsmiljø.

Sammenlignet med tidligere studier, gir rapporten i større grad en helhetlig analyse av hvordan ulike ressursindikatorer på individ- og skolenivå har betydning for elevenes prestasjoner. Til grunn for analysene ligger et omfattende datamateriale der ulike datakilder – på individnivå og skolenivå – er koblet sammen. Dette har muliggjort en rekke nye analyser av sammenhenger mellom skoleressurser og elevresultater både på 5., 8. og 10. trinn. I analysene av grunnskolens avgangselever (10.trinn) brukes data fra Elevundersøkelsene både på skolenivå og på individnivå. Dette er ikke tidligere gjort. Vi kommer tilbake til en nærmere beskrivelse av disse analysene senere i kapitlet.

Rapporten presenterer analyser av elevers resultater på nasjonale prøver på 5. og 8. trinn i grunnskolen årene 2007, 2008 og 2009, samt karakterdata for elevkullet på 10. trinn som gikk ut av grunnskolen våren 2009. Informasjon om elevenes prestasjoner er koblet til informasjon om demografiske og sosiale kjennetegn ved elevene og deres foreldre hentet fra Statistisk sentralbyrå. Dette omfatter informasjon om elevenes kjønn, alder, fødeland, ankomstår til Norge og søskenforhold. Opplysninger om foreldrene omfatter mors og fars utdanningsnivå, inntekt, yrkesstatus, fødeland og sivile status. Til dette datasettet, er det koblet til data på skolenivå hentet fra Grunnskolenes informasjonssystem (GSI) og data fra Elevundersøkelsene.

Siden variablene er på både individ- og skolenivå, benyttes flernivåanalyser. Denne metoden gjør det mulig å bruke variabler på ulike nivåer samtidig i analysene. Dermed viser resultatene både hvordan hver enkelt variabel er relatert til skoleprestasjoner, og hvor mye av variasjonen i skoleprestasjoner som kommer av forskjeller mellom skoler og forskjeller mellom individer.

Analysene viser gjennomgående at elevkjennetegn på individnivå, det vil si kjennetegn ved elevene og deres foreldre, har langt større betydning for elevenes prestasjoner på alle tre klassetrinn enn kjennetegn ved skolens ressurser. Analysene viser også tydelige sammenhenger mellom elevenes oppfatninger av sitt eget læringsmiljø, slik det måles gjennom Elevundersøkelsen, og elevenes prestasjonsnivå. Også enkelte av indikatorene hentet fra GSI synes å korrelere med elevenes prestasjoner. Både resultatene knyttet til kjennetegn ved elevene og deres foreldre, resultater knyttet til læringsmiljø og til andre kjennetegn ved skolene, vil bli nærmere presentert og diskutert i del to av dette kapitlet.

Tilleggsnotat: Analyser av segregering i skolen

Med utgangspunkt i de analysene som ble presentert i delrapport 1, ønsker Utdanningsdirektoratet en nærmere undersøkelse av enkelte funn herfra. Dette gjaldt særlig analyser som tydet på en positiv sammenheng mellom andelen ikke-vestlige innvandrerelever ved den enkelte skole og elevenes samlede prestasjoner. Resultatet ble i delrapport 1 tolket som at en viss etnisk uensartethet kan representere en prestasjonsressurs for elevene. Denne sammenhengen analyseres nærmere i arbeidsnotatet *Segregering, læringsmiljø og ikke-vestlige innvandrerelevers prestasjoner på nasjonale prøver* (Næss 2011). I tillegg undersøkes om et godt læringsmiljø er viktigere for innvandrerelever enn majoritetselever, og om det er forskjeller mellom skoler når det gjelder hvor vellykket integreringen av elever med innvandrerbakgrunn har vært. Siden Oslo og Akershus er de fylkene med klart høyest andel elever med innvandrerbakgrunn gjennomføres det separate analyser for disse fylkene og for resten av landet.

Resultatene i notatet tyder på at en moderat andel ikke-vestlige innvandrere på skolen er positivt knyttet til elevenes læringsresultater, så lenge andelen ikke overstiger 20 prosent. Imidlertid påpekes det at de statistiske modellene er komplekse, og at sammenhengen mellom innvandrerandel og prestasjoner kan skyldes forhold som ikke er undersøkt i analysene. Dermed kan analysene gi et skjevt bilde av betydningen av innvandrerandel. Temaet er også sentralt i delrapport 5 (Grøgaard 2012), og vi kommer tilbake til dette senere i kapitlet.

2.1.2 Delrapport 2: Elevenes prestasjonsutvikling

I delrapport 2 fra prosjektet, *Elevers prestasjonsutvikling – hvor mye betyr skolen og familien?* (Wiborg mfl 2011), var hensikten å bygge videre på analysene fra delrapport 1. I tillegg til å analysere elevenes prestasjonsnivå, var fokuset i denne rapporten å undersøke skolens bidrag til elevenes prestasjonsutvikling i løpet av grunnskolen – såkalte *value added*-analyser. Dette ble undersøkt ved å sammenligne prestasjonsutviklingen til to kull elever: Elevene som begynte på 5. trinn høsten 2007 og på 8. trinn høsten 2010, og elevene som begynte på 8. trinn høsten 2007 og fullførte 10. trinn våren 2010. Prestasjonsutviklingen fra 5. til 8. trinn er basert på resultater fra nasjonale prøver i lesing, engelsk og regning, og utviklingen fra 8. til 10. trinn er basert på nasjonale prøver fra 8. og karakterdata ved utgangen av 10. trinn. Siden det er utvikling over tid som undersøkes, er analysene basert på de elevene hvor det foreligger prestasjonsdata for på to tidspunkt.

Rapporten presenterer for første gang analyser av elevenes prestasjonsutvikling fra 5. til 8. trinn i grunnopplæringen. Med unntak av en tidligere studie av elevenes prestasjonsutvikling fra 8. til 10. trinn der en tidligere versjon av nasjonale prøver (fra 2004) ble benyttet (Bakken 2010), presenterer rapporten også nye analyser av elevenes prestasjonsutvikling gjennom ungdomstrinnet – fra 8. til 10. trinn.

Resultatene viser relativt høyt samsvar mellom elevenes prestasjoner på to tidspunkt i grunnopplæringen. Korrelasjonene (uttrykt ved Pearsons r) ligger mellom 0,6 og 0,8 både for sammenhengen mellom prestasjonene i 5. og 8. klasse, og for sammenhengen mellom prestasjonene i 8. og 10. trinn.

Mange av de forholdene som i første delrapport ble funnet å forklare prestasjonsnivå har også betydning for prestasjonsutviklingen. Det gjelder både betydningen av kjennetegn ved elevene selv og kjennetegn ved skolens ressurser og læringsmiljø. Et mye omtalt tema i norsk skole er i hvilken grad skolen bidrar til å utjevne sosiale forskjeller i elevenes læringsutbytte eller om dette er forskjeller som forsterkes utover i skoleløpet. Rapporten finner i liten grad at sosiale forskjeller reduseres i løpet av skolegangen: Skolen synes ikke å kompensere for forskjeller i elevenes bakgrunn. I likhet med resultatene fra første delrapport, er dette analyser og resultater som vil bli nærmere presentert og diskutert i del to av dette kapitlet.

2.1.3 Delrapport 3: Prestasjonsutvikling gjennom Vg1

Delrapport 3 bygger også videre på de foregående delrapportene: *Prestasjonsutvikling fra ungdomsskolen til første året i videregående opplæring* (Arnesen 2012). Mens andre delrapport undersøkte skolens bidrag til elevenes prestasjonsutvikling for to kull i den tiårige grunnskolen, undersøker delrapport 3 elevenes prestasjonsutvikling det første året etter fullført tiårig grunnskole og gjennom Vg1.

Rapporten er basert på registerdata for elever som fullførte 10. trinn i grunnskolen vår 2010 og som startet på videregående skole samme høst. Omtrent 95 % av alle som gikk ut av ungdomsskolen dette året fortsatte med videregående utdanning. Rapporten presenterer for første gang analyser av elevenes prestasjonsutvikling fra 10. trinn til Vg1 for en hel kohort elever. SSB har tidligere presentert analyser av prestasjonsutvikling fra 10. trinn til Vg1 blant elever i Oslo (Hægeland mfl 2010).

Rapportens hovedproblemstillinger berører to prosesser: For det første hvorvidt betydningen av familie og individkjennetegn endres når forhold ved skolene blir tatt hensyn til. For det andre om betydningen av forhold ved skolen endrer seg når kjennetegn ved elevmassen tas hensyn til.

I likhet med de foregående delrapportene, nyttes ulike statistiske metoder i rapporten fra enkle beskrivende analyser til mer avanserte flernivåanalyser. Som avhengig variabel benyttes to ulike mål på prestasjonsutvikling. Det ene målet er basert på forskjell mellom grunnskolepoeng og Vg1-poeng. Antall poeng er basert på karakterer i alle tellende fag. Det andre målet er basert på forskjeller i standpunktkarakterer i basisfagene norsk, engelsk og matematikk. De to karaktermålene er sterkt korrelert, men forskjeller i poeng er et bredere kompetansemål enn gjennomsnittet av de tre basisfagene.

I tillegg til de to målene på prestasjonsutvikling (smalt og bredt) blir elevenes studiegjennomføring også brukt som avhengig variabel. Et betydelig antall elever faller fra videregående utdanning, også i løpet av det første året. Disse elevene blir ikke fanget opp av analyser som kun ser på prestasjoner. Analysene av elevenes studiegjennomføring gir dermed et bilde av hvilke forhold ved skolen eller elevene selv som øker eller minsker sannsynligheten for at en elev gjennomfører eller dropper ut av skolen i løpet av Vg1.

I motsetning til grunnskolen kjennetegnes videregående skole av mange ulike utdanningsprogrammer og en relativt stor grad av heterogenitet. Denne heterogeniteten medfører blant annet at det foregår en seleksjon av bestemte elevgrupper til bestemte utdanningsprogrammer. Analysene viser for eksempel høyere prestasjonsnivå (i form av bedre gjennomsnittskarakterer fra 10. trinn) blant elever som startet på et studieforbereende utdanningsprogram enn elever som gikk til yrkesfaglige utdannelse. Slike forskjeller er velkjent, men medfører visse metodiske utfordringer i analyser av elevenes prestasjonsutvikling. Analysene skiller derfor mellom elever i studieforbereende og yrkesfaglige utdanningsprogrammer.

2.1.4 Delrapport 4: Prestasjoner og læringsmiljø i to fag

Den fjerde delrapporten fra prosjektet presenterer resultatene fra ett av de to delprosjektene innenfor prosjektet 'Ressurser og resultater': *Resultater fra en spørreundersøkelse blant elever og lærere på 9. trinn og i Vg1* (Opheim og Wiborg 2012). Grunnlaget for analysene i denne rapporten er en spørreundersøkelse gjennomført blant lærere og elever på 9. trinn i ungdomsskolen og i videregående trinn 1 (Vg1) våren 2011. I undersøkelsen har elevene fått spørsmål om undervisningssituasjon og læringsmiljø i to konkrete fag: norsk og matematikk. Dette gjør det mulig å analysere betydningen av læringsprosesser innenfor konkrete fag. Undersøkelsen omfatter også elevenes faglærere i norsk og matematikk. Faglærerne har fått spørsmål om undervisningssituasjon og læringsmiljø for en bestemt klasse. Spørsmålene til faglærerne omhandler klassen som helhet og ikke enkelte elever i klassen. Lærernes svar kan kobles til elevenes svar. Det betyr at samsvaret i elevenes og lærernes svar når det gjelder undervisningen i de to konkrete fagene kan studeres. Utvalget har en klyngestruktur: Ved deltakerskolene har elever i samtlige klasser på ett klassetrinn kunnet delta. Undersøkelsen gir dermed informasjon om tre nivåer: skole, klasse og elev.

Målsetningen med spørreundersøkelsen har vært å få tak i kontekstnær og generaliserbar kunnskap om hva som foregår i klasserommet - elevenes viktigste læringsarena. Etersom undersøkelsen omfatter utvalg av hele klasser samt opptil to av klassens faglærere, gir datastrukturen muligheter for å analysere sammenhenger mellom en gitt lærer og en bestemt gruppe elever – det vil si reelle lærer-elev-relasjoner. Dette står i motsetning til analyser av sammenhenger mellom elevenes prestasjoner og kjennetegn ved skolen basert på registerdata, slik det er gjort i de tidligere delrapportene i prosjektet.

Rapporten diskuterer metodologiske utfordringer ved utformingen av spørreundersøkelsen samt datamaterialets kvalitet og representativitet. Designet som er valgt i undersøkelsen er krevende og sårbart for frafall – både på skolenivå, klassenivå, blant enkelte elever og lærere. Frafall i ulike ledd i datainnsamlingen medfører at kun et begrenset utvalg klasser kan knyttes til faglærere. Et annet forhold som blir diskutert er omfanget av useriøse svar blant elevene som deltok i undersøkelsen. Nærmere 8 prosent av elevene ble fjernet fra analysene på grunn av useriøse svar. En sammenligning av resultatene med og uten denne elevgruppen (tøysekoppene), ga betydelige forskjeller. Å fjerne useriøse svar har derfor hatt stor betydning for analysene. Useriøse svar er også tidligere funnet å være et problem i undersøkelser av ungdommer (Elstad 2010, Lødding og Vibe 2010). Til tross for metodologiske utfordringer tyder frafallsanalyser på relativt god representativitet i datamaterialet av elevene. Utvalget av faglærere er derimot mindre og mer usikkert med hensyn til representativitet.

Analysene i rapporten omfatter betydning av undervisningsform, vurderingshyppighet, motivasjon, samt forhold knyttet til elevenes læringsmiljø – trivsel, motivasjon og uro i klasserommet. Til tross for at datamaterialet er forskjellig, peker flere av resultatene i samme retning som resultatene fra de foregående delrapportene i prosjektet. Dette diskuteres nærmere i del to i dette kapitlet.

2.1.5 Delrapport 5: Skoler med høyt prestasjonsnivå på nasjonale prøver

I femte delrapport fra prosjektet, *Hva kjennetegner barneskoler som oppnår høy skår på nasjonale prøver?* (Grøgaard 2012), undersøkes kjennetegn ved skoler der elevene i gjennomsnitt oppnår prestasjoner godt over gjennomsnittet på de nasjonale prøvene. Datamaterialet som brukes tilsvarer det som ble brukt i andre delrapport (Wiborg mfl 2011), rapporten er dermed en videreføring av analysene som her ble presentert. Rapporten konsentrerer seg om en gruppe elever: de som gjennomførte nasjonale prøver for 5. trinn i 2007 og for 8. trinn i 2010.

I likhet med de foregående rapportene, viser resultatene også her stor betydning av elevsammensetningen ved skolen. I motsetning til de foregående rapportene går analysene her imidlertid lenger i å undersøke i hvilken grad betydningen av skolens elevsammensetning er den samme for alle elevgrupper, eller om elevsammensetningen har ulike betydning for ulike grupper av elever. Rapporten inneholder i tillegg nye analyser av sammenheng mellom andelen elever med minoritetsbakgrunn ved skolen og elevenes prestasjonsnivå og prestasjonsutvikling.

Analysene i rapporten viser at det er en fordel for elever å gå på skoler der foreldrene i gjennomsnitt har høy utdanning, høy inntekt og der mange av foreldrene er gift. Derimot viser analysene at elever med bakgrunn fra ikke-vestlige land, i motsetning til majoritetselevne, ikke får positivt utbytte av å gå på skoler med slik elevsammensetning.

Resultatene viser at det er en fordel for elever å være tilknyttet en skole der det er et visst minoritetsinnslag blant elevene, trolig opp til en minoritetsandel tilsvarende 11–20 prosent av elevene. I Oslo kan det være prestasjonsgevinster også ved skoler med høyere minoritetsandel. Resultatene er i tråd med analyser fra andre deler av dette prosjektet (se tilleggsnotatet til Næss 2011).

Nærmere analyser viser imidlertid at det kun er majoritetselevne som 'får' et positivt utbytte av denne elevsammensetningen. Elever med innvandrerbakgrunn fra ikke-vestlige land har ikke høyere prestasjonsnivå ved skoler med en høy andel minoritetselever. Selv om minoritetselevne bidrar til en prestasjonsgevinst blant majoritetselevne, ser det med andre ord ikke ut til at de selv får det samme utbyttet av sin egen tilstedeværelse. Resultatene diskuteres som uttrykk for manglende integrasjon av elever med minoritetsbakgrunn ved en del norske skoler. Dette kommer vi tilbake til i del to av dette kapitlet.

2.1.6 Delrapport 6: Den gode timen – casestudie av fire Osloskoler

Sjette og siste delrapport fra prosjektet, *Den gode timen* (Markussen og Seland 2013), ble publisert vinteren 2013. Rapporten presenterer resultatene fra en casestudie av fire ungdomsskoler i Oslo. Dette utgjør et eget delprosjekt i «Ressurser og resultater». Casestudien omfatter observasjoner av ti forskjellige klasser og til sammen 40 timer/undervisningsøkter fordelt omtrent likt på niende og tiende klasse. I tillegg til klasseromsobservasjoner har casestudien omfattet intervjuer med rektorene på hver av skolene. Hensikten med casestudien har vært å undersøke nærmere noen av resultatene fra de kvantitative delene av prosjektet. Et sentralt spørsmål i har vært: Hva kjennetegner en god skoletime?

På grunnlag av observasjoner og intervju med rektor, diskuteres betydning av klasseledelse, pedagogisk plattform ved skolen, skolens elevsammensetning, ressursnivå og ressursbruk, samt bruk av ulike arbeidsformer og overganger mellom arbeidsformer i timene.

Resultatene fra casestudien bidrar til å utfylle funn fra de kvantitative analysene. Dette gjelder særlig på to områder. Det første er hvordan klasseromsobservasjonen tydeliggjør lærerens rolle i ulike situasjoner, og hvordan lærerens valg og handlinger har betydning for kvaliteten på undervisningen. Observasjonene viser at læreren har mye å si for oppstarten av timen og for ro og orden, at lærernes inngripen eller mangel på inngripen bestemmer hvor lenge forstyrrelser holder på og at lærerne gjennom valg av undervisningsopplegg påvirker hvordan timen blir.

Det andre området hvor casestudien bidrar til å utfylle funn fra de kvantitative analysene, er ved å tydeliggjøre sammenhenger mellom ulike elementer med betydning for undervisningskvaliteten. Klasseromsobservasjonene viser hvordan ro og orden i timen henger sammen med valg av undervisningsform, og hvordan en klasse kan oppføre seg svært forskjellig avhengig av hvilken lærer eller hvilket fag de har – samtidig som samme lærer kan ha både gode og dårlige timer med samme klasse. Observasjonene illustrerer hvordan forhold i klasserommet påvirker hverandre: elevenes arbeidsinnsats avhenger blant annet av hvordan undervisninger er lagt opp, elevenes oppførsel kan variere avhengig av lærer og fag, samtidig som også lærerens innsats og undervisningstimens kvalitet kan variere etter fag og undervisningsformer. Resultatene diskuteres i sammenheng med resultater fra andre deler av prosjektet i neste del av dette kapitlet.

2.2 Prosjektets hovedfunn: Bidrag til kunnskapsstatus

I første del av kapitlet har vi kort presentert de seks delrapportene fra prosjektet. I resten av kapitlet vil vi se nærmere på hovedresultatene fra de ulike delene av prosjektet og diskutere dem opp mot tidligere forskning. Hensikten er å tydeliggjøre prosjektets bidrag til kunnskapsstatus på dette feltet. Fremstillingen er inndelt tematisk. Eventuelle metodiske diskusjoner knyttet til funnene blir presentert i kapitlene 4 og 5.

2.2.1 Sterk sammenheng mellom elevkjennetegn og læringsutbytte

Prosjektet har som målsetting å legge til rette for et bedre kunnskapsgrunnlag og utvikling av kvalitetsindikatorer for vurdering av ressursinnsats og læringsresultater i grunnopplæringen. I dette ligger det et fokus på faktorer som er mulig å påvirke – enten i form av kjennetegn ved læringsmiljø, skolens ressursbruk eller andre faktorer. I utgangspunktet skulle prosjektet derfor omhandle sammenhenger mellom skoleressurser og elevenes prestasjoner, der kjennetegn ved elevene selv var tiltenkt en mindre rolle i analysene. Et gjennomgående funn er imidlertid de sterke effektene av elevkjennetegn og skolens elevsammensetning. Individuelle kjennetegn som kjønn, sosioøkonomisk bakgrunn og innvandrerbakgrunn er sterkt knyttet til elevenes prestasjoner. Dette er i tråd med tidligere studier.

Elevenes skoleprestasjoner kan langt på vei tilbakeføres til deres familiebakgrunn og innvandrerbakgrunn. Videre finner vi gjennomgående at elevsammensetningen ved skolen har betydning for elevenes prestasjoner. Det betyr at ikke bare elevens egne kjennetegn, men også kjennetegn ved de andre elevene på skolen, har betydning for elevenes prestasjoner. Skolens elevsammensetning får dermed form av å være en ressursindikator på skolenivå. Å diskutere sammenhenger mellom ressurser og resultater uten å ta inn over seg den sterke sammenhengen mellom elevkjennetegn og prestasjoner, vil være å utelate en sentral del av diskusjonen. Skolens elevsammensetning er en del av skolens kontekst som trolig langt på vei definerer skolens prioriteringer, handlingsrom og utfordringer. Analysene i dette prosjektet går likevel videre enn mye av det som er gjort i tidligere studier, ved å undersøke nærmere hvordan indikatorer på skoleressurser og læringsmiljø kan ha ulik betydning for ulike grupper av elever.

Et tema som særlig undersøkes, er hvilken utbytte ulike grupper av elever har av å gå på en skole med en gunstig elevsammensetning i betydningen at en høy andel av elevenes foreldre har et høyt utdanningsnivå. Resultatene viser at dette har positiv betydning for de fleste grupper av elever – på tvers av kjønn og sosial bakgrunn. Unntaket er elever med minoritetsbakgrunn. Resultatene tyder på at denne gruppen elever i liten grad får et positivt utbytte av å gå på en skole med en gunstig elevsammensetning (Grøgaard 2012). Dette paradokset kan være uttrykk for manglende integrasjon av elever med innvandrerbakgrunn i majoritetsgruppen. Analysene peker på sammenhenger mellom elevenes læringsutbytte og det å være godt integrert i læringsmiljøet på skolen.

Et nærliggende tema som også undersøkes, er hvilken betydning andelen minoritets elever ved skolen har for elevenes læringsutbytte. Flere analyser tyder på at en viss mengde elever med innvandrerbakgrunn på en skole synes å være positivt for elevenes prestasjoner (Opheim mfl 2010, Grøgaard 2012, Næss 2011). Imidlertid viser analysene også at når andelen elever med innvandrerbakgrunn overstiger et visst nivå, reduseres prestasjonsnivået. Andelen innvandrerelever ser ut til å være positivt opp til omtrent 20 prosent. Høyere andel enn dette synes å være negativt for prestasjonene. I Oslo-skolen er disse tallene noe høyere: Innvandrerandeler opp til omtrent 30 prosent synes å være positivt i hovedstaden. Paradoksalt nok tyder resultatene på at innvandrerelevne selv *ikke* får del i den positive effekten av en viss mengde elever med innvandrerbakgrunn. Det er med andre ord kun majoritets elevene som har høyere prestasjonsnivå ved skoler med en viss andel minoritets elever, ikke minoritets elevene selv (Grøgaard 2012).

2.2.2 Læringsmiljøet viktig for elevenes prestasjoner

Analysene viser gjennomgående at alle elevgrupper, uavhengig av kjønn, sosioøkonomisk bakgrunn og innvandrerbakgrunn, tjener på et godt læringsmiljø. Både data fra Elevundersøkelsen og fra spørreundersøkelsen til elever og lærere viser relativt sterke og positive sammenhenger mellom elevenes oppfatning av eget læringsmiljø og deres skoleprestasjoner. Også casestudien peker på sider ved læringsmiljøet i klasserommet som synes å ha betydning for kvaliteten på undervisningstimen. Et godt læringsmiljø vil si et miljø som er inkluderende, støttende, trivelig, punktlig og faglig fokusert. Sammenhengene mellom læringsmiljø og elevenes prestasjoner, slik de analyseres i de kvantitative analysene, består også når vi tar høyde for variasjon mellom elever og mellom ulike ressurser på skolene.

Funnene er i overensstemmelse med tidligere forskning. Både nasjonale (Bakken 2009a, 2010) og internasjonale studier (Stewart 2008), finner at elevers opplevelse av skolens læringsmiljø er relatert til deres prestasjoner. Det er stor grad av enighet om at læringsmiljøet på skolen har betydning for elevenes faglige utbytte (Nordahl, 2005; Skaalvik og Skaalvik, 2005; Wendelborg, Røe og Skaalvik 2011). Vanligvis sees sammenhengen mellom læringsmiljø og prestasjoner som en konsekvens av de positive effektene av et godt læringsmiljø. Imidlertid er det god grunn til å tro at også det motsatte er tilfelle: At elevenes prestasjoner på skolen kan påvirke deres trivsel. Elever som opplever de faglige utfordringene som passe vil trolig trives bedre på skolen enn elever som sliter med fag.

Et godt arbeidsmiljø, det vil si et miljø som er inkluderende, støttende, trivelig, punktlig og faglig fokusert, synes å være spesielt positivt for elever med innvandrerbakgrunn. På skoler med godt læringsmiljø er prestasjonsforskjellene mellom majoritets- og minoritets elever mindre enn på skoler med svakere læringsmiljø. Funnet, som er i tråd med tidligere norsk forskning (Bakken og Elstad 2012), tyder på at å satse på et godt miljø på skolen kan ha stor betydning for elevenes resultater. Både i delrapport 5 og i tilleggsnotatet gjøres det utdypende analyser av sammenhenger mellom elevenes oppfatning av sitt læringsmiljø og prestasjonsnivå blant elever med minoritets- og majoritetsbakgrunn (Grøgaard 2012, Næss 2011).

2.2.3 Betydning av ulike undervisningsformer

Mange av analysene i prosjektet har vist en sammenheng mellom elevenes prestasjoner og omfang av ulike undervisningsformer. Med utgangspunkt i data fra Elevundersøkelsen, skilles det mellom to undervisningsformer: Lærerstyrt og elevaktiv undervisning. Lærerstyrt undervisning, slik det er definert i prosjektet, kjennetegnes av tavleundervisning, individuelt arbeid under oppsyn av lærer og diskusjon/samtale mellom læreren og elevene i klassen. Elevaktiv undervisning kjennetegnes av ulike former for samarbeid mellom elevene, som gruppearbeid, prosjektarbeid, praktiske øvelser og samarbeid to og to.

Mens lærerstyrt undervisning gjennomgående synes å være positivt knyttet til skoleprestasjoner, både målt på individnivå og som gjennomsnitt på skolenivå, synes elevaktiv undervisning i større grad å ha en negativ betydning for elevenes prestasjoner. Sammenhengen er til stede i analyser av nasjonale prøver på 5. og 8. trinn, analyser av elevenes karakter på 10. trinn (Opheim mfl 2010), samt i analyser

av spørreundersøkelsen til elever på 9. trinn og i Vg1 (Opheim og Wiborg 2012). Sammenlignet med Elevundersøkelsen er spørsmålene om undervisning i spørreundersøkelsen knyttet til to konkrete fag: elevene ble spurt om ulike sider ved undervisningen i norsk og matematikk, og svarene ble vurdert opp mot deres prestasjoner i de samme fagene.

Sammenhengene mellom undervisningsformer og elevenes prestasjoner vi finner i prosjektet «Ressurser og resultater» er ikke veldig sterke. Det er flere andre forhold som synes å ha større betydning for elevenes læringsutbytte enn bruk og omfang av ulike undervisningsformer. Vi skal derfor være varsomme med å tillegge undervisningsformene alt for stor vekt. Det at sammenhengene mellom elevenes prestasjoner og omfang av ulike undervisningsformer gjenfinnes ved bruk av ulike datakilder og ulike prestasjonsmål, gir likevel grunn til å diskutere disse funnene nærmere.

Tidligere forskning

Som presentert i kapittel 1 har ulike forskningsretninger tilnærmet seg skole og utdanning på ulike måter. Skoleeffektivitetsforskningen i Norge har primært studert sammenhengen mellom innsatsfaktorer og resultater, og har i liten grad fokusert på betydningen av ulike undervisningsformer. Innenfor klasseromsforskningen har fokus vært på undervisning og hva som foregår i klasserommet, men hvordan undervisningsmetode henger sammen med elevenes resultater er i mindre grad undersøkt. Det innebærer at innenfor en norsk kontekst, er sammenhengen mellom undervisningsform og elevenes prestasjoner i liten grad studert.

Beveger vi oss til USA, har skoleeffektivitetsforskningen i større grad også inkludert prosessfaktorer, i tillegg til innsatsfaktorer. Allerede på 1970- og 80-tallet ble det her gjort studier av undervisningsformer og resultater – i form av sammenheng mellom lærerens bruk av ulike metoder for undervisning og instruksjon (*teacher behaviour*), og elevenes prestasjoner (*student achievement*). Et tema i disse studiene er betydningen av lærerens ledelse eller kontroll i klasserommet (*teacher management*). Kontroll handler her ikke bare om læreren har kontroll over elevenes oppførsel og holder ro i klasserommet, men også om kontroll over læringsoppgaver (hva skal læres og hvordan). Resultatene peker i positiv retning for lærere som opprettholder ro i klasserommet, snakker mye (bruker tid på instruksjon) og har kontroll over hvilke undervisningsmetoder som brukes (Brophy og Good 1986, Good og Brophy 2008). Dette synes å være i tråd med de positive effektene av lærerstyrt undervisning som fremheves i dette prosjektet. De amerikanske studiene tydet imidlertid på at sammenhengene mellom lærerens kontroll og elevenes prestasjoner varierte etter elevenes sosiale bakgrunn. Elever med lav sosial bakgrunn og/eller svake læringsforutsetninger, syntes å tjene mer på å ha mye strukturert undervisning og kontroll i klasserommet enn elever med høyere sosial bakgrunn. Interessant nok, peker resultatene i delrapport 4 i samme retning, selv om effektene er for svake til å gi en entydig tolkning (Opheim og Wiborg 2012).

I senere tid har det kommet til enkelte studier som, på tvers av de etablerte tradisjonene, forsøker å finne ut om det er systematiske sammenhenger mellom undervisningsmetode og prestasjoner. En slik studie finner at elever med lærere som oppgir at de bruker mye forelesende undervisning i gjennomsnitt har noe bedre skoleprestasjoner enn elever med lærere som i mindre grad bruker en slik type undervisning (Schwerdt og Wuppermann, 2011). Resultatet er i overensstemmelse med funn i dette prosjektet. En annen nyere studie fant ingen sammenheng mellom bruken av forelesende undervisning og skoleprestasjoner (Klaveren, 2011).

Studien til Schwerdt og Wuppermann (2011) viser videre en negativ sammenheng mellom bruk av oppgaveløsning i undervisningen; når læreren oppgir at de ofte bruker denne metoden presterer elevene dårligere enn når læreren sier at de ikke bruker den så ofte. Det framkommer ikke om oppgaveløsning innebærer at elevene samarbeider, så det er uklart hvordan dette funnet forholder seg til våre funn i dette prosjektet. Andre studier viser at samarbeide mellom elevene er positivt for deres prestasjoner (Slavin, Lake og Groff, 2009; Slavin, Lake, Davis og Madden, 2011). Disse funnene synes dermed ikke å samsvare med resultatene fra dette prosjektet, hvor elevaktiv undervisning i flere analyser synes å ha en svak negativ sammenheng med elevenes prestasjoner.

Den eksisterende forskningen gir ikke noe entydig bilde av sammenhengen mellom undervisningsformer og prestasjoner. Det er derfor vanskelig å fastslå hvorvidt funnene i «Ressurser og resultater» støttes av tidligere studier eller ikke. Studiene kommer fra forskjellige land, noe som kan være med å forklare de varierende funnene. Til forskjell fra de nevnte studier, er analysene i vårt prosjekt basert på elevenes vurdering av undervisningen. Det kompliserer sammenlikningen mellom dette prosjektet og tidligere studier. Analyser av surveyen i delrapport 4 (Opheim og Wiborg 2012) viser at elever og lærere opplever undervisningen forskjellig, og at det er elevenes oppfatning som har betydning for prestasjonene. Analysene viser ingen sammenhenger mellom lærernes oppfatninger av undervisningsform og elevenes prestasjoner. Det er derfor ikke urimelig at analyser av sammenhengen mellom elevenes opplevelse av undervisningen og deres skoleprestasjoner gir andre funn enn analyser av sammenhengen mellom lærerens opplevelse av undervisningen og elevenes prestasjoner.

Mulige forklaringer på sammenhengen mellom undervisningsform og resultater

Både spørreundersøkelsen til elever og lærere og casestudien tyder på at ro og orden i timen kan forklare noe av sammenhengen mellom undervisningsform og elevenes prestasjoner (Opheim og Wiborg 2012, Seland og Markussen 2013). Klasseromsobservasjonene viser at de ulike formene for samarbeid som inngår i elevaktiv undervisning ofte innebærer mye uro og forstyrrelser og, for noen, lite fokus på arbeidsoppgavene. Når elevene selv skal styre det de gjør er det mange som velger å snakke med andre om ting som ikke vedkommer fag, eller gjøre ting som reduserer muligheten til å lære. Når det gjelder lærerstyrt undervisning, tyder observasjonene derimot på stor grad av ro og orden. Observasjonsstudien peker altså på at elevstyrt undervisning fører til mer uro og mindre fokus på arbeidet enn lærerstyrt, noe som er i tråd med resultatene fra spørreundersøkelsen. Dette støttes av analyser i delrapport 4, som viser at ro og orden forklarer noen av sammenhengen mellom undervisningsform og prestasjoner, men ikke alt (Opheim og Wiborg 2012).

Observasjonene av timene viser stor variasjon i undervisningskvaliteten uavhengig av hvilken undervisningsform som dominerer i timen. Observasjonene illustrerer også hvordan det kan være vanskelig å skille ulike undervisningsformer fra hverandre. Et eksempel er observasjonene av individuelt arbeid som undervisningsform. I flere timer skulle elever arbeide individuelt med oppgaver, noe som vi har betegnet som en del av den lærerstyrte undervisningsformen. Men i hvilken grad elevene faktisk arbeidet på egen hånd eller samarbeidet, varierte betydelig. Grad av arbeidsro varierte også i stor grad i disse timene. Hvordan oppfattet elevene disse timene – som gruppearbeid eller individuelt arbeid? Det vet vi ikke. Observasjonene peker på noe av vanskelighetene ved å definere arbeidsformer i klasserommet som en bestemt undervisningsform. Dette ble for øvrig også diskutert i delrapport 4 (Opheim og Wiborg, 2012).

Dersom det er slik at lærerstyrt undervisning innebærer mer ro og mindre forstyrrelser, kunne man spørre seg hvorfor ikke lærerne bare har slike timer. Ved siden av det vi allerede har påpekt, at sammenhengen ikke trenger å være så direkte, kan et forhold også være hvor ressurskrevende de ulike undervisningsformene er. Lærerstyrt undervisning vil i mange tilfeller kreve grundig planlegging, og kan dermed være mer ressurskrevende enn elevaktive undervisningsformer. Dette synes for øvrig å være i tråd med tidligere forskning som har pekt på det krevende ved å undervise eller instruere hele klassen på likt i motsetning til mindre grupper (*whole class versus small group instruction*) (Brophy og Good 1986). Observasjoner fra casestudien fremhever også den lærerstyrte undervisningen, i betydning undervisning der læreren har en tydelig lederrolle, trolig krever mer forberedelse og lærerinnsetts, enn undervisningsformer der læreren inntar en mer tilbaketrukket rolle – noe som samtidig synes å redusere læringstrykket og kvaliteten på undervisningen (Markussen og Seland 2013). Det kan bety at omfang av lærerstyrt undervisning, blant annet kan avhenge av lærerens innsats, erfaring og ressurser/tid til planlegging av undervisningen. Nå kan det innvendes at gruppearbeid og annen elevaktiv undervisning også krever planlegging, noe som selvsagt er tilfelle. Resultatene kan likevel danne utgangspunkt for blant annet å diskutere hvilke forhold som kan bidra til lærerens bruk av ulike undervisningsformer og hvilke ressurser som kreves av læreren ved gjennomføring av de ulike undervisningsformene.

En annen forklaring på sammenhengen mellom lærerstyrt undervisning og prestasjoner kan være at en viss grad av lærerstyring er nødvendig for å få til god undervisning, uavhengig om elevene jobber sammen eller læreren leder undervisningen. I følge Schacter og Thum (2004) er lærernes *utførelse* av undervisningen av særlig stor betydning for elevenes resultater. Det de to forfatterne beskriver som gode måte å undervise på kjennetegnes av at læreren er tydelig og leder elevene gjennom de ulike oppgavene, uavhengig av om det er gruppearbeid eller instruksjon som brukes. Når vi ber elevene om å vurdere hva slags undervisning de har mye og lite av kan det være at de som har lærere som underviser i tråd med Schacter og Thums syn, i større grad enn andre oppgir at de har mye lærerstyrt undervisning. Slik kan elevenes svar tolkes som uttrykk for at læreren har kontroll i klasserommet og styrer undervisningen, mens den elevaktive undervisningen i større grad kan kjennetegnes ved at læreren har inntatt en mer fraværende og passiv rolle, der elevene i større grad har ansvar for egen læring.

Selv om analysene i prosjektet tyder på at det er undervisningsmetodene som fører til forskjeller i prestasjonene, er også andre forklaringer mulig. Årsaksrekkefølgen kan være den motsatte, at elevenes prestasjoner er årsaken til valg av undervisningsform. En slik forklaring innebærer at høyere elevprestasjoner fører til valg av lærerstyrte undervisningsformer, mens lavere elevprestasjoner fører til valg av elevaktive undervisningsformer. Det kan være grunn til å spørre hvor sannsynlig det er at læreren velger lærerstyrte undervisningsformer i større grad når elevene presterer godt enn når de ikke gjør det. Eller omvendt, at en lærer vil velge å bruke mer elevaktive undervisningsformer i en klasse med svakt prestasjonsnivå enn med et høyt prestasjonsnivå. Det intuitive er kanskje at lærere i klasser med svakt prestasjonsnivå i større grad vil velge undervisningsformer der de har høy grad av kontroll over klassen – altså lærerstyrte undervisningsformer. Samtidig kan det tenkes at det nettopp i slike klasser er særlig utfordrende for læreren å gjennomføre denne type undervisning, og at det derfor brukes mer elevaktive undervisningsformer i klasser med svakt prestasjonsnivå. Casestudien illustrerer kompleksiteten i samspillet mellom lærer og elev og de mange forhold som kan påvirke gjennomføringen av og innholdet i en undervisningstime. Det kan derfor ikke utelukkes at elevenes prestasjonsnivå kan ha betydning for læreres valg av undervisningsformer.

En annen mulig tolkning er at sammenhengen mellom undervisningsformer og læringsutbytte har sammenheng med bakenforliggende forhold, for eksempel kjennetegn ved lærerne. Det vil i så fall innebære at kjennetegn ved lærerne, som alder og ansiennitet/erfaring er bakenforliggende forhold som kan forklare lærernes bruk av ulike undervisningsformer. Dersom lærere med mye erfaring har en tendens til å bruke mer lærerstyrt undervisning enn andre, kan sammenhengen mellom undervisningsform og prestasjoner i virkeligheten handle om at lærere med mye erfaring har positiv betydning for elevenes læringsutbytte. Datamaterialet inneholder relativt lite informasjon om lærerne, noe som gjør det vanskelig å undersøke slike forhold. Enkelte kjennetegn ved lærerne inngår imidlertid i analysene, uten at dette har vist seg å ha særlig betydning for elevenes læringsutbytte. Dette er lærernes kjønn (kjønnsfordeling på skolenivå) og andel lærere med godkjent utdanning på skolen. For øvrig tyder også internasjonale studier på at egenskaper ved lærerne, som erfaring, har relativt lite å si for elevenes prestasjoner sett i forhold til undervisningsmetodene (Aslam og Kingdon 2011).

Selv om delrapportene gjennomgående viser positive effekter av lærerstyrt undervisning, finnes det i prosjektet også analyser som problematiserer dette. I delrapport 5 viser utdypende analyser av sammenhenger mellom elevenes minoritetsbakgrunn og læringsutbytte, at elever med innvandrerbakgrunn synes å ha et bedre utbytte av elevaktive undervisningsformer enn av lærerstyrt undervisning (Grøgaard 2012). Dette indikerer at den positive sammenhengen mellom lærerstyrt undervisning for elevenes prestasjoner, primært gjelder elever med majoritetsbakgrunn. En mulig tolkning av dette kan være at elevaktive undervisningsformer har integrerende effekter ettersom de innebærer samarbeid mellom elevene i klassen, og at dette er særlig positivt for læringsutbyttet blant elever med innvandrerbakgrunn.

2.2.4 Kan det bli for mye vurdering?

Analysene av sammenheng mellom vurdering/evaluering av skolearbeid og elevenes prestasjoner er i hovedsak basert på resultater fra Elevundersøkelsen. Her skiller det mellom til sammen seks former for vurdering av skolearbeid: Mappeevaluering, praktiske oppgaver, skriftlige prøver, muntlig høring i klassen, presentasjoner (for klassen) og hjemmearbeid (Opheim mfl 2010). I de fleste av analysene i prosjektet har de seks ulike formene blitt slått sammen til ett samlet mål for elevenes oppfatning av omfang av vurdering. Resultatene av analysene tyder på at et høyt omfang av vurdering av skolearbeid er negativt relatert til elevenes prestasjoner. Dette funnet bekreftes i analysene av spørreundersøkelsen i delrapport 4, som finner at høyere vurderingshyppighet er knyttet til lavere prestasjoner i norsk og matematikk (Opheim og Wiborg 2012). Sammenhengen består også etter at en rekke individuelle forhold er tatt hensyn til.

Innføringen av Kunnskapsløftet har medført et sterkere fokus på at elevene skal nå de faglige målene, og dermed også mer utstrakt bruk av vurdering og testing (Bakken og Elstad 2012). Dermed kan kunnskap om eventuelle negative konsekvenser av evaluering være særlig aktuelt for skolen i dag.

Tidligere forskning

Antallet studier som har undersøkt relasjonen mellom mengden evaluering og prestasjoner, slik det er gjort i dette prosjektet, er begrenset. En tidligere gjennomgang av internasjonal forskning på dette temaet konkluderte med en positiv sammenheng mellom evaluering og prestasjoner: elever som hadde hatt flere prøver i et fag gjorde det bedre på eksamen enn elever med få eller ingen prøver (Bangert-Drowns, Kulik og Kulik 1991). Dette står i klar kontrast til hva vi finner i dette prosjektet. Imidlertid viste gjennomgangen at effekten av vurdering sank betydelig etter hvert som antallet prøver økte. Elever med én prøve i løpet av et femten ukers kurs gjorde det nesten like godt på eksamen som de som hadde hatt mange prøver i løpet av kursperioden, og vesentlig bedre enn de som ikke hadde hatt noen prøver (ibid.).

Dette kan tyde på at vurdering av skolearbeidet er viktig for elevenes prestasjoner, men at det optimale omfanget av prøver ikke nødvendigvis er veldig stort. Denne tolkningen av resultatene er mer i tråd med funnene i dette prosjektet. Få eller ingen av elevene som deltar i Elevundersøkelsen oppgir at deres skolearbeid *ikke* blir evaluert. Mange av elevene rapporterer derimot om mye evaluering. De tre vanligste formene for evaluering er leksehøring, innlevering av lekse og skriftlig prøve. Mellom 40 og 45 prosent av elevene på 10. trinn oppgir at de har slike former for evaluering en eller flere ganger i uken. Bangert-Drowns og kolleger diskuterer ikke hva som bør eller kan være den optimale hyppigheten på evaluering, og det er vanskelig å sammenlikne våre resultater med funn fra deres gjennomgang. En mulig tolkning av våre resultater, sett i lys av tidligere studier, kan være at omfanget av vurdering blant norske elever kanskje er høyere enn det som er optimalt.

Det kan bety at den negative sammenhengen mellom evaluering og prestasjoner vi finner i dette prosjektet, kan skyldes at noen skoler bruker mye tid på evaluering, og at det går utover tid til undervisning og læring. Det betyr derimot ikke at all vurdering og evaluering av skolearbeid er negativt. Observasjonene i casestudien i delrapport 6 illustrerer dette (Markussen og Seland 2013). Her ble vurdering i mange tilfeller observert som en integrert del av undervisningen. Gjennom tilbakemeldinger som gjaldt eget og andres arbeid fikk elevene informasjon om hvordan de lå an og hva læringsmålene var. Denne formen for undervisning, som kanskje i større grad hadde preg av å være en type uformell muntlig høring av elevene, syntes å bidra positivt til elevenes læringsprosess. I casestudien ble det ikke gjort observasjoner av mer formaliserte og eksplisitte vurderingsformer, som skriftlige prøver. Vi kan derfor ikke utelukke at slike vurderingsformer er mindre egnet til å hjelpe elevens læring enn de mer implisitte vurderingsformene som foregår som en del av undervisningen.

Skillet mellom vurdering som en naturlig del av undervisningen og vurdering som formalisert testing kan bidra til å forstå hvorfor vår studie viser at vurdering er negativt knyttet til prestasjoner, mens andre studier viser en positiv sammenheng. Flere forskere hevder at dersom vurdering skal være positivt for prestasjonene er det viktig at den integreres i undervisningen og at den følges opp av gode

og informative tilbakemeldinger til eleven (Hattie og Timperley 2007, Black og Wiliam 2009). Også norske studer støtter dette: vurdering i seg selv fører ikke til økte prestasjoner, den må følges opp i etterkant gjennom tilbakemeldinger til eleven og justering av undervisningsopplegget for at den skal ha effekt (Bakken og Danielsen 2011, Ottesen og Møller 2010). Undersøkelser fra USA viser at det å innføre formaliserte test-regimer i enkelte tilfeller henger sammen med forbedrede elevprestasjoner, men at det ikke nødvendigvis er testingen i seg selv som fører til bedre resultater (Lee 2008, Lee og Reeves 2012). Derimot framstår oppfølging av skolene og satsing på lærerne som viktig for å forklare prestasjonsforbedringene.

Den eksisterende forskningen peker på at vurdering og testing som knyttes til undervisningen kan ha positiv effekt, men at økt vurdering uten oppfølging ikke gir bedre prestasjoner. Dersom mengden testing og vurdering har økt i norsk skole de siste årene, slik Bakken og Elstad (2012) hevder, kan det være grunn til å spørre hvilke former for vurdering og testing det har blitt mer av. Er det vurdering som en del av undervisningen? Eller er det former for vurdering som i liten grad fører til tilbakemelding til eleven og justering av undervisningen? I tilfelle det er den første formen for vurdering som har økt er det grunn til å spørre om lærerne faktisk klarer å følge opp gjennom å justere undervisningen og å lage tilpassede opplegg til elever som trenger det. Uavhengig av svarene på spørsmålene tyder funnene i «Ressurser og resultater» på at vurdering av skolearbeid, slik det gjennomføres i skolen i dag, ikke nødvendigvis fører til økt læring. Det kan derfor være grunn til å vurdere omfanget av vurdering og å undersøke i hvilken grad de ulike vurderingsformene er integrert i undervisningen.

Andre forklaringer

Diskusjonen så langt har tatt utgangspunkt i at det er forskjeller i mengden vurdering som forklarer skoleprestasjonene. Men årsakssammenhengene trenger ikke være slik. Det kan også tenkes at lærere med elever/klasser med svakt prestasjonsnivå, bruker vurdering som en måte å få alle elever med seg og dermed få framdrift i læringsarbeidet. Enkelte observasjoner i casestudien kan tyde på at vurdering kan ha ulike funksjoner for ulike grupper av elever og i ulike klasser. En grundigere undersøkelse av årsakssammenhenger, er derimot komplisert og vil kreve ytterligere datakilder.

Det kan også være at elever med ulikt prestasjonsnivå har ulik oppfatning av omfanget av vurdering. Elever med et høyt prestasjonsnivå kan oppleve det som mindre problematisk å bli vurdert, mens elever med et svakere prestasjonsnivå kan oppleve vurderinger av skolearbeidet, prøver og innleveringer som svært krevende. Dette kan igjen ha betydning for elevenes opplevelse av omfanget av vurdering. Subjektive variasjoner i elevenes oppfatninger synes imidlertid ikke å forklare all sammenheng mellom vurdering og prestasjoner. I delrapport 4 viste resultatene negative sammenhenger mellom omfang av vurdering og elevenes prestasjoner i norsk og matematikk, også når analysene tok høyde for hvilken klasse elevene gikk i og variasjon mellom klassene (Opheim og Wiborg 2012).

2.2.5 Små forskjeller mellom skoler

Et gjentatt funn i de kvantitative analysene er at relativt lite av variasjonen i elevenes prestasjoner kan tilbakeføres til skolenivået. Delrapport 1 (Opheim, Grøgaard og Næss, 2010) viser at omtrent 7 til 10 prosent av variasjonen skyldes forskjeller mellom skolene. Analyser av klassenivået i delrapport 4 (Opheim og Wiborg 2012) viser at mellom 13 til 22 prosent av variasjonen på 9. trinn kom av forskjeller mellom klasser. De tilsvarende tallene for Vg1 var 24 til 38 prosent. Det betyr at noe av variasjonen mellom elever innad den enkelte skole, har sammenheng med at elever er gruppert i ulike klasser. I grunnskolen utgjør dette kun en mindre andel av variasjonen i elevenes prestasjoner. I videregående skole er det derimot større forskjeller mellom klassene. Dette er ikke uventet, ettersom det kan være store forskjeller mellom klasser på Vg1 når det gjelder organisering, undervisning og elevenes avgangskarakterer fra ungdomsskolen.

Når relativt lite av variasjonen i elevens prestasjoner kan tilbakeføres til skoler og klasser så innebærer det at mye av variasjonen skyldes forskjeller mellom elever: mesteparten av variasjonen er mellom elever i samme klasse og på samme skole, og ikke mellom skoler eller klasser.

Selv om andelen variasjon som kan knyttes til skolenivået er relativt liten er det ikke grunnlag for at si at ulikheter mellom skoler ikke har betydning for elevenes prestasjoner. Forskjellene mellom ytterpunktene er betydelige. Å løfte skolene i den nederste delen av fordelingen opp på nivå med de øverste skolene ville hatt betydelig effekt for de elevene det gjelder, og ville gitt utslag i høyere gjennomsnittlige prestasjoner på nasjonalt nivå.

Tidligere forskning

En årsak til at skolenivået ikke har større betydning er trolig at skolene i Norge opererer innenfor det samme økonomiske og juridiske systemet. Ressurstilgangen er i stor grad regulert av offentlige bestemmelser som gjelder for alle skoler. Videre styres skolenes bruk av de økonomiske midlene i vesentlig grad av aktuelle lover, slik at elever sikres et minimum av ressurser uavhengig av for eksempel geografisk beliggenhet. Dermed blir skolenes rammebetingelser relativt like, og lite påvirket av kommunenes økonomi. Dette gjør det samtidig vanskelig å isolere betydningen av skolenes ressurser (i form av økonomisk ressursnivå) for elevenes læringsutbytte.

Også internasjonal forskning peker i samme retning. På bakgrunn av eksisterende studier konkluderer Hattie (2009) med at størstedelen av variasjonen i elevenes prestasjoner finnes innen skolene og ikke mellom skolene. Hattie mener at store forskjeller mellom lærerne er en viktig årsak til variasjonen i elevenes prestasjoner innad i skolen. En av studiene han trekker fram viser at mellom 0 og 20 prosent av variansen i elevenes prestasjoner kan tilbakeføres til skolenivået, mens 16 til 60 prosent kan tilbakeføres til lærer- eller klassenivået. I en annen gjennomgang av eksisterende forskning finner Luyten (2003) at i underkant av 10 prosent av variasjonen i prestasjonene kommer av forskjeller mellom skolene. Han finner videre at de få studiene som har undersøkt betydningen av både skole og klassenivået, vanligvis viser at klassenivået er viktigere enn skole.

Forskningen Hattie og Luyten baserer seg på er primært fra land som kan sammenlignes med Norge når det gjelder økonomiske ressurser. Hattie peker på at mindre ressurser trolig fører til større forskjeller mellom skoler, noe som kan være aktuelt i land med svak økonomi.

Selv om tidligere studier gir noe varierende estimat av hvor mye av variasjonen som kan tilbakeføres til forskjeller mellom skoler, skiller ikke våre funn seg vesentlig ut. Resultatene er i tråd med eksisterende forskning fra land som kan sammenlignes med Norge. Det at en relativt liten del av variasjonen i prestasjoner handler om forskjeller mellom skoler ser ikke ut til å være et særnorsk fenomen. Relativt sett, synes likevel forskjellene mellom skoler å være mindre i Norge enn i mange andre land. Dette understrekes i PISA-undersøkelsene, der de nordiske landene gjennomgående fremstår med relativt små forskjeller i prestasjoner mellom skoler sammenlignet med andre land som deltar i PISA-undersøkelsen (Kjærnsli og Roe 2010).

2.2.6 Betydning av skoleressurser

Analysene i prosjektet finner at elevenes prestasjoner bare i liten grad synes å være direkte knyttet til skolenes ressurser, slik de er registrert i GSI. Dette er i overensstemmelse med tidligere studier, både norske og internasjonale, som heller ikke finner at skolenes ressursbruk har vesentlig betydning for prestasjonsnivået (Hanushek 1986, Hanushek 1997, Hægeland mfl 2008, Bonesrønning mfl 2010). Funnet kan gi inntrykk av at skolenes ressurser ikke er viktig for prestasjonene. Imidlertid er det trolig mer riktig å se funnene som indikator på at skolene har relativt lik mengde med tilgjengelige ressurser og at spredningen i ressursbruk er liten. Studier som undersøker betydningen av spesielt mye ressurser finner at dette kan være positivt for prestasjonene (se kapittel 1). For eksempel viser en studie at skoler i spesielt rike kommuner i Norge har positive effekter av ekstra ressursinnsats (Hægeland mfl 2008). Formålet med studien var å si noe om betydningen av ressurser som går utover det ordinære, såkalte «frie ressurser». Dette er trolig avgjørende for hvorfor forfatterne finner en sammenheng. Først når skolene tildeles ressurser som går betydelig ut over det normale er det mulig å finne effekter. Variasjon i ressursene som ligger innenfor normalområdet vil ikke slå ut på prestasjonene fordi de i stor grad er bundet opp i lovpålagte oppgaver og derfor ikke gir skolene mulighet til å prioritere vesentlig annerledes enn andre skoler.

Blant de forholdene som *ikke* synes å ha betydning for elevenes prestasjoner, er antallet elever per lærer. Til tross for at det er rimelig å anta at dette forholdet påvirker elevenes læring, finner vi ikke at det har betydning for prestasjonsnivået. Funnet bekreftes av tidligere studier (Bonesrønning mfl 2010). Imidlertid er det grunn til å diskutere hvorvidt forholdstallet mellom elever og lærere er en god indikator på tilgjengelige lærerressurser. Hvordan funnet skal forstås blir derfor nærmere diskutert i kapittel 4.

Antall undervisningstimer

Selv om det på bakgrunn av funnene i «Ressurser og resultater» er rimelig å konkludere med at skolens ressurser har lite å si for elevenes prestasjoner, finner vi enkelte sammenhenger av betydning. For eksempel er antall pc-er per elev og andel lærere med godkjent utdanning relatert til elevenes prestasjoner. Sammenhengene heftes imidlertid av usikkerhet angående hvordan disse målene skal forstås. Dette diskuteres nærmere i kapittel 4.

En ressursindikator som gjennomgående ble funnet å være av betydning for elevenes prestasjoner, og som vi skal gå nærmere inn på her, er antall timer med undervisning i løpet av året (Opheim mfl 2010). Også tidligere studier har vist en positiv relasjon mellom tid til læring og prestasjoner (Rutter mfl 1979, Brookover 1979). Studiene fant at elever som hadde mer tid til rådighet også presterte bedre. Funn fra PISA-undersøkelsen støtter dette, men sammenhengen som kommer fram der er svak (OECD 2010). Heller ikke resultatene fra «Ressurser og resultater» tyder på en sterk sammenheng mellom antall timer og prestasjoner. Samlet peker dette på at tid til undervisning og læring har en viss, men ikke avgjørende, betydning for elevenes læring.

Mye av forskningen på tidsbruk i skolen har vært mer opptatt av hvordan tiden brukes enn hvor mye tid som er tilgjengelig. Studier viser at mye av undervisningstiden går bort til andre ting, og flere forskere mener at skolene kan effektivisere tidsbruken. Smith (2000) hevder at tilgjengelig tid til læring av basisfagene har blitt redusert og fragmentert i løpet av de siste årene. Grunnen til det er, ifølge Smith, for det første at det foregår mye i skolen som ikke handler om læring og undervisning, som testing og sosiale og kulturelle aktiviteter. For det andre mener Smith at skolene har feilaktige oppfatninger av hva som er effektivt når det gjelder læring. Dette stjeler verdifull tid som kunne vært brukt til undervisning. Alt i alt peker eksisterende forskning på at mange skoler kan tjene på å effektivisere undervisningen, noe som kanskje kan redusere behovet for å utvide antall skoletimer.

Hattie (2009) peker på at å øke tiden brukt på undervisning kun vil ha positive effekter dersom undervisningen har høy kvalitet. Dersom kvaliteten på undervisningen er lav, kan en utvidelse av antall skoletimer i verste fall være negativt for prestasjonene. Kvaliteten på undervisningen er viktig, og særlig dersom man skal ha mer av den.

2.2.7 Stor variasjon mellom skoletimer

Et sentralt spørsmål i casestudien er: Hva kjennetegner 'den gode timen'? Casestudien viser at skoletimene varierer betydelig når det gjelder konsentrasjon og arbeidsro, altså forhold som man kan forvente har betydning for elevenes læring. Også innad på de samme skolene viser observasjonene både det som kan betegnes som gode timer og timer med lav undervisningskvalitet (Markussen og Seland 2013). Ulik undervisning kan være en del av forklaringen på hvorfor det er så stor variasjon i elevenes prestasjoner innad på den enkelte skole.

Forskjellene i skoletimene peker på betydningen av læreren og hvordan læreren organiserer undervisningen og håndterer forstyrrelser og uro. Dette er i tråd med flere nyere studier som peker på lærernes atferd som avgjørende for elevenes læring (Schwerdt og Wuppermann 2010, Aslam and Kingdon 2010). Også norske studier har pekt på dette, flere av disse er omtalt i kapittel 1 (Klette 2003, Imsen 2004, Lyng 2004 og Haug 2011). Casestudien viser imidlertid eksempler på at den samme klasse oppførte seg klart forskjellig i to ulike timer, til tross for at de hadde den samme læreren ved begge anledningene. Mens den ene timen framsto som svært god var den andre timen i noen tilfeller preget av uorden og manglende struktur. Hvorfor ble timene så ulike når så mye var likt? Kom det av at det ble undervist i ulike fag? Eller var det andre forhold som var avgjørende? Observasjonene gir

ikke noe svar på hvorfor timene ble så forskjellige, men det er nærliggende å hevde at det ikke bare skyldes lærerens beslutninger og prioriteringer når det gjaldt undervisningen. Trolig er lærers valg av undervisningsopplegg og strategier for å håndtere bråk og uro et resultat av samspill mellom en rekke forhold ved skolen, læreren, elevene i klassen og faget. Også skoleledelsens vektlegging av prestasjoner og synspunkter når det gjelder undervisningsmetoder kan ha betydning for hvordan lærerne legger opp undervisningen. Selv om observasjoner i casestudien peker på betydningen av læreren, viser de også hvordan ulike faktorer sammen skaper gode eller mindre gode timer.

2.2.8 Sammenheng mellom skoleledelse og elevenes læringsutbytte?

Et tema som i liten grad fanges opp av de datakildene som har vært brukt i prosjektet, er sammenhengen mellom skoleledelsen (rektor) og elevenes resultater. Både i diskusjonen av funn fra analyser i dette prosjektet samt i tidligere studier, blir betydning av god skoleledelse trukket frem som en faktor som trolig har positiv betydning for elevenes resultater (se bl.a. Opheim mfl 2010, Bonesrønning mfl 2010). Med økende oppmerksomhet rettet mot elevenes resultater, etter innføringen av Kunnskapsløpet og de årlige nasjonale prøvene, er betydningen av god skoleledelse et stadig aktuelt tema. At god skoleledelse er viktig, er alle enige om. Hva som kjennetegner god skoleledelse og hvilken betydning (ulike kjennetegn ved) skolens ledelse har for elevenes resultater, er det derimot ingen enkle svar på (Scheerens 2012). Dette er heller ingen enkel oppgave å undersøke. Sammenhenger mellom skolens ledelse og elevenes læringsutbytte går via lærerne og må dermed studeres indirekte; ved å undersøke sammenhenger mellom skolens ledelse og lærernes undervisningspraksis og klasseledelse, samt å undersøke sammenhenger mellom det som skjer i klasserommet, lærerens rolle og elevenes læringsutbytte. Slike sammenhenger ble i dette prosjektet forsøkt fanget opp ved å gjøre en casestudie av et lite utvalg skoler i Oslo (Markussen og Seland 2013).

Casestudien besto av to deler; observasjoner av undervisningsøkter samt intervju med skolens rektor ved tre skoler i Oslo. Alle skoler med lik eier (offentlige), men variasjon i elevsammensetning og prestasjonsnivå på nasjonale prøver. På bakgrunn av dette datamaterialet synes casestudien å finne enkelte indikasjoner på sammenheng mellom skoleleders pedagogiske plattform og lærernes undervisningspraksis – noe som igjen syntes å være knyttet til kvaliteten på undervisningsøktene og elevenes læringsutbytte. Dette er resultater som det er vanskelig å løsrive fra kontekst – det er et tett samspill mellom skolens elevsammensetning og elevenes prestasjonsnivå. Skolens elevsammensetning kan ha betydning for skoleledelsens prioriteringer/innsatsfelt.

Skolene var valgt ut for nettopp å fange opp variasjon i elevprestasjoner, og skilte seg fra hverandre både med hensyn til gjennomsnittlig prestasjonsnivå på de nasjonale prøvene og med hensyn til skolens elevsammensetning. Målsettingen var å forsøke å finne faktorer som ikke er fanget opp i det kvantitative datamaterialet, og som kunne bidra til å forklare hvorfor enkelte skoler presterer godt mens andre skoler presterer svakere på de nasjonale prøvene.

Skolene ble valgt ut med utgangspunkt i prestasjoner og prestasjonsutvikling for kullet som gjennomførte ungdomsskolen 2007-2010. Dette innebærer at elevene som utgjorde grunnlaget for utvalget av skoler ikke var de samme som elevene som ble observert (se mer om dette i delrapport 6, Markussen og Seland 2013). Likevel er det verdt å merke seg at skolen som hadde hatt en positiv prestasjonsutvikling blant elever med et relativt lavt utgangspunkt, samtidig var den som viste tendenser til positive sammenhenger mellom skolens ledelse og undervisningskvaliteten i timene. Ved denne skolen var skolens ledelse kjennetegnet av et sterkt fokus på måloppnåelse, tett oppfølging av lærere og tydelig kommunikasjon om forventninger til skolens lærere. Dette er faktorer som også tidligere studier har pekt på som kjennetegn ved god skoleledelse (Kjærnsli mfl 2004, Scheerens 2012). Analyser av PISA 2003 viste blant annet at skoleledere som oppga å ha fokus på elevenes prestasjoner i gjennomsnitt hadde et høyere prestasjonsnivå blant elevene ved skolen enn skoleledere med mindre fokus på dette (Kjærnsli mfl 2004). Analysene kan imidlertid ikke fastslå årsaksretning for denne sammenhengen. Det kan være at skoleledernes fokus på elevenes læringsutbytte, medførte bedre undervisningskvalitet og dermed bedre læringsutbytte. Men det kan

også være at skoler med elever med et høyt prestasjonsnivå hadde skoleledere som fokuserte på dette – uten at skolelederne selv hadde betydning for elevenes læringsutbytte. Dette er en gjennomgående utfordring i slike analyser – noe som også diskuteres i dette prosjektet.

Casestudiene peker altså på noen mulige sammenhenger mellom skolens ledelse og den pedagogiske plattformen ved skolen og undervisningskvaliteten i timene. Samtidig kan dette både være uttrykk for en sammenheng med direkte betydning for elevenes prestasjoner, eller det kan være at skolens pedagogiske plattform først og fremst reflekterer elevgruppen ved skolen og de utfordringer/muligheter den innebærer, og altså i stor grad er knyttet til elevsammensetningen ved skolen. Dette er tema som krever mer omfattende undersøkelser og bør følges opp i videre forskning.

2.3 Oppsummering

I dette kapitlet har vi presentert og gjennomgått hovedinnholdet i de seks delrapportene fra dette prosjektet. Deretter har hovedresultater fra prosjektet blitt diskutert og knyttet til funn fra tidligere studier. I de neste kapitlene vil vi gå nærmere inn på de sentrale datakildene i dette kapitlet. Utgangspunktet er å diskutere i hvilken grad informasjon fra de ulike datakildene kan fungere som kvalitetsindikatorer for grunnopplæringen, basert på funn fra prosjektet og datakildenes form og oppbygging.

3 Ulike prestasjonsmål som indikatorer på kvalitet

3.1 Nasjonale prøver og karakterdata som mål på kvalitet

I dette prosjektet diskuteres hvilke faktorer ved læringsprosesser og læringsmiljø som fører til gode resultater, og som mål på slike resultater har vi valgt å legge vekt på nasjonale prøver og karakterdata. Det vi altså har gjort, er å identifisere faktorer i skolen som gir godt utbytte i form av karakterer og resultater på nasjonale prøver. Relevansen av disse kvalitetsindikatorerne vil dermed avhenge av hvorvidt man anser data fra nasjonale prøver og karakterdata som gode mål på det vi ønsker at skolen skal fremme.

I *Opplæringslova* (Kunnskapsdepartementet 2013) står det at formålet med opplæringen er at denne skal «opne dører mot verda og framtida og gi elevane og lærlingane historisk og kulturell innsikt og forankring». Opplæringen skal blant annet fremme demokrati, lære elever og lærlinger å tenke kritisk, handle etisk og utfolde utforskertrang. Elevene skal «utvikle kunnskap, dugleik og holdningar for å kunne meistre liva sine og for å kunne delta i arbeid og fellesskap i samfunnet». I vårt forsøk på å identifisere kvalitetsindikatorer i norsk skole, ser det ut til av ett av disse målene – kunnskap for å kunne delta i arbeid – har blitt vektlagt i særlig grad. Da må vi stille oss selv spørsmålet om vi har mistet skogen av syne fordi vi stirrer på ett enkelt tre. Det er ikke sikkert at disse prestasjonsmålene i tilstrekkelig grad fanger opp elevens læringsutbytte. I dette kapitlet skal vi drøfte hvorvidt nasjonale prøver og karakterdata fungerer som mål på kvalitet i norsk skole.

3.2 Kritikk mot ulike prestasjonsmål

Karakterdata og nasjonale prøver er prestasjonsmål som i stor grad er knyttet til standardiserte, ofte skriftlige, kunnskapstester. Kritikerne av slike mål på skolekvalitet er mange. Eksempelvis er det flere som hevder at TIMSS- og PISA-resultatene i for stor grad benyttes av politikere for å holde dom over norsk skole, og Sjøberg (2007) skriver at «faren med tall er at bare det som kan telles, teller.» Viktige forhold som naturglede og miljøbevissthet måles ikke i verken TIMSS eller PISA. Angående nasjonale prøver, er mye av kritikken mot disse samlet i boken *Kunnskapsbløffen* (Marsdal 2012). Her beskrives negative utslag testene har for ressursfordeling og fokus i norsk skole. Marsdal gir eksempler på forhold ved gjennomføringen av nasjonale prøver som gir grunn til å vurdere kritisk om resultatene herfra er velegnede mål på skolekvalitet. En grundigere gjennomgang av nasjonale prøver som system finnes for øvrig i en egen NIFU-rapport (Seland mfl 2013).

Det stilles også spørsmål til kvaliteten på karakterdata. En av hovedinnvendingene mot karakterer er knyttet til den store variasjonen i vurderingspraksis mellom skoler og mellom lærere (Allen 2005). Måles elevene relativt til kompetansemålene eller til medelevene? I hvor stor grad skal innsats og

forbedring avspeiles i karakterene? Eksamenskarakterene går fri av mange slike kritikker, ettersom eksterne sensorer er ment å skulle sikre objektive vurderinger av elevens prestasjoner. Eksamenskarakterer er i så måte mindre påvirket av personlige forhold og forutsetninger. Disse er likevel ikke uproblematisk. Flere studier viser at sammenhengen mellom gode eksamenskarakter og en dyp forståelse av fagenes sentrale begreper, er svakere enn antatt (Jakobsen mfl 1999). En indikasjon på problemene med karakterdata er synlig i forskjellen mellom standpunktkarakterer og eksamenskarakterer. Ved noen skoler har elevene et gjennomsnitt på eksamen som ligger på nivå med deres standpunktkarakter, mens det ved mange skoler er slik at elevene i snitt går ned en hel karakter fra standpunkt til eksamen (Utdanningsdirektoratet 2013).

Når man ser på summen av slike prestasjonsmål, som blant annet inkluderer TIMSS, PISA, de nasjonale prøvene, lokale prøver (f. eks Osloprøven), prøver for fastsetting av standpunktkarakterer og eksamener, bør man diskutere hvorvidt dette krever for mye av skolens og elevenes tid, ressurser og fokus. Møller, Prøitz og Aasen (2009) spør seg om det kompetansedrevne fokuset i norsk skole har ledet til et testregime. Resultatene i dette prosjektet kan tyde på at det er noe som heter *for mye testing*.

Det finnes altså velbegrunnet kritikk mot bruk av standardiserte kunnskapstester som eneste mål på skolekvalitet. Nå skal vi se nærmere på karakterdata og data fra nasjonale prøver og diskutere hvordan disse datakildene påvirket analysemulighetene i prosjektet.

3.3 Relative mål for prestasjonsutvikling

En strategi for å identifisere kvalitetsindikatorer i norsk skole, er å undersøke hvordan elevenes prestasjoner utvikler seg over en periode (Raudenbush 2004, Hægeland mfl 2011). Da vil resultatene i mindre grad avhenge av hvor dyktige elevene er idet de begynner på skolen, og i større grad avhenge av bidraget skolen gir i perioden mellom måletidspunktene. Dette gjør analysene mer robuste, i den forstand at de motvirker effekten av at noen skoler tar imot elever som har et bedre utgangspunkt enn elever ved andre skoler.

Slike «value added»-analyser stiller svært strenge krav til datamaterialet (Meyer 1997). Man må blant annet kunne argumentere for at endringen skyldes forhold i skolen, og ikke det eleven har lært på andre arenaer i det aktuelle tidsrommet. Videre vil et robust mål på prestasjonsutvikling kreve at man ved to måletidspunkter testes i oppgaver med sammenlignbare vanskelighetsgrader. I denne forbindelse har karakterdata og data fra nasjonale prøver en stor svakhet. De er ikke utviklet for å måle kompetanseutvikling over tid (Utdanningsdirektoratet 2012). Disse dataene benyttes kun som *relative* mål for kompetanseutvikling.

Karakterer er relatert til de kompetansemålene skolen har for elevene på bestemte tidspunkt. Med andre ord vil en karakterendring fra ett skoletrinn til neste ikke si noe om hvor mye *mer* en elev har lært, men hvor godt eleven presterer relativt til nye kompetansemål. Eksempelvis er det ikke slik at en elev som går fra karakteren 5 til karakteren 3 i matematikk fra ett år til et annet, kan mindre enn han gjorde året før. Eleven har derimot blitt stilt overfor andre og mer krevende kompetansemål.

Vedrørende nasjonale prøver, har det inntil gjennomføringen i 2012 vært slik at prøvene endres for hver nye gjennomføring (Utdanningsdirektoratet 2013). På individnivå betyr dette at en elev ikke kan sammenligne resultatet sitt med foregående resultater, og på skolenivå betyr dette at en skole ikke kan sammenligne resultatene for et trinn med resultater for samme trinn i foregående år. Det eneste man kan fastslå, er om man presterer bedre eller dårligere i forhold til andre skoler – igjen en *relativ* prestasjonsutvikling. Dette er viktig å ha i mente når man leser resultatene i delrapport 2 (Wiborg mfl 2011), delrapport 3 (Arnesen 2012) og delrapport 5 (Grøgaard 2012). Der beskrives blant annet kjennetegn ved skoler som har en positiv prestasjonsutvikling. Denne positive utviklingen dreier seg da om at skolens elever gjør det bedre i forhold til elever ved andre skoler enn de gjorde det ved forrige målepunkt. Fra 2013 vil det innføres noen oppgaver som skal være med på permanent basis. Disse vil gjøre det mulig å sammenligne resultater fra år til år.

3.4 Personer som ikke er representert i datamaterialet

En styrke ved å bruke karakterdata og data fra nasjonale prøver, er at disse omfatter en stor andel av elevene i norsk skole. De mange observasjonene medfører at statistiske resultater fremstår som svært robuste. Slik statistikk må likevel tolkes med en god porsjon skepsis. En årsak til dette er knyttet til hvem som *ikke* er representert i datamaterialet. For selv om vi har svært mange elever i datamaterialet, utgjør mindretallet som ikke er representert en vesentlig statistisk utfordring. Denne gruppen utgjør omtrent 10 % av elevkullet (Utdanningsdirektoratet 2012).

I delrapport 1 (Opheim mfl 2010) presenteres fritaksordningen knyttet til deltakelse i nasjonale prøver. Hovedregelen er at prøvene er obligatoriske for alle elever, og at muligheter for fritak er begrenset (Utdanningsdirektoratet 2012). Det er kun to kategorier av elever som skoleleder kan vurdere for fritak: elever med enkeltvedtak om spesialundervisning eller elever med enkeltvedtak om særskilt språkopplæring. Disse elevene får altså ikke automatisk fritak, men skal i hvert enkelt tilfelle vurderes av skoleleder og elevens lærer i fellesskap. Sentralt i vurderingen er om elevene vil ha nytte av prøvene i opplæringen. Utdanningsdirektoratet påpeker at praksis for fritak kan variere mellom skoler, kommuner og fylker (Utdanningsdirektoratet 2012).

Fritakspraksisen har vært gjenstand for gjentatt medieoppmerksomhet de senere årene. GSI-data vedrørende antall elever tyder på at omtrent 10 % av elevkullet uteblir fra nasjonale prøver. Til sammenligning hadde kun 3–4 % av elevkullet i 5. og 8. trinn i 2012 fritak fra den nasjonale leseprøven. Dette er et betydelig avvik, og enkelte skoleledere har blitt anklaget for bevisst å gi svake elever fritak fra deltakelse i håp om å øke skolens skår på slike prøver. Denne diskusjonen begir ikke vi oss inn på her. Poenget i vår diskusjon er ikke *hvorfor*, men *at* vi mangler mange svaktpresterende elever i datamaterialet. Dette medfører to konkrete problemer når kvalitetsindikatorer i norsk skole skal identifiseres.

En utfordring er knyttet til at fritakspraksisene varierer. Om det er slik at skoler med strenge fritakspraksiser har noen felles kjennetegn, vil disse kjennetegnene statistisk sett assosieres med svake resultater på skolenivå, ettersom flere svaktpresterende elever deltar fra slike skoler. Tilsvarende vil felles kjennetegn ved skoler som gir mange elever fritak, feilaktig kunne identifiseres som kvalitetsindikatorer. Å ta høyde for fritakspraksiser i analyser er ikke mulig med vårt datamateriale, ettersom antall elever med fritak ikke nødvendigvis er knyttet til fritakspraksis, men til faktiske forhold i elevmassen.

En annen utfordring er knyttet til hvilken informasjon som er tilgjengelig om gruppen av de svakeste elevene i norsk skole. Ettersom svært mange av disse ikke er representert i vårt datamateriale, minsker også muligheten man har til å gjøre analyser vedrørende denne spesifikke gruppen. Det kan gjerne tenkes at enkelte forhold ved læringsmiljøet eller bestemte undervisningspraksiser fungerer særlig godt for disse elevene. Muligheten til å gjøre slike analyser er redusert siden mange i nettopp denne gruppen ikke er representert i datamaterialet.

3.5 Snevert mål på skolekvalitet?

Hvilket læringsutbytte som bør måles, henger sammen med hvilke læringsmål skolen bør fremme. To ledende stemmer i denne diskusjonen er Gagné (1974) og Eisner (1979). Når de presenterer sine respektive syn på læring, identifiseres to motstridende syn på det læringsutbytte skolen bør fremme. Mens Gagné anlegger et resultatorientert perspektiv med et veldefinert og målbart læringsutbytte, fremmer Eisner et prosessorientert perspektiv hvor læringsutbyttet er mer fleksibelt og mindre målbart. Etter flere tiår med diskusjon har man ikke kommet nærmere en enighet om hvilket av disse perspektivene som bør være rådende i skolen (Prøitz 2010). Derfor er det også vanskelig å gi en entydig beskrivelse av hvordan læringsutbytte bør måles. Spørsmålet om karakterer er et snevert mål på skolekvalitet er en stor debatt. Er det virkelig slik at karakterer bare gir oss innblikk i en marginal del av utbyttet elevene har av skolegangen? Er det kanskje slik at det er de samme elevene som har gode

karakterer som også gjør det godt når andre prestasjonsmål tas i bruk, og at karakterene dermed er en god indikasjon på måloppnåelse innen et bredt spekter av skolens målsettinger?

I delrapport 5 (Grøgaard 2012) blir det diskutert hvilket perspektiv på skolekvalitet som anlegges i dette prosjektet. Grunnopplæringen er obligatorisk, og det faktum at elevene tilbringer store deler av hverdagen sin på skolen, gjør det blant annet viktig at elevene trives godt. Mange av skolens målsettinger er altså av «mykere» karakter enn kunnskapstestene som her benyttes som mål på skolekvalitet. Svært ulike ferdigheter, holdninger, egenskaper og opplevelser inngår i denne helheten. Opplæringen skal kunne anvendes på livets private, offentlige og profesjonelle arenaer og i møte med mennesker fra andre kulturer og i andre livssituasjoner.

Denne mangfoldigheten kontrasteres av standardiserte kunnskapsprøver som gjennomføres elektronisk eller med penn og papir, og hvor kontakten med andre mennesker ofte er svært begrenset. Det er av denne grunn at man i delrapport 5 (Grøgaard 2012) diskuterer hvorvidt nasjonale prøver og karakterdata er et tilstrekkelig mål på læringsutbytte. Andre ferdigheter, holdninger, egenskaper og opplevelser kan indikere kvalitet i norsk skole, og testing av disse kunne muligens supplert eller erstattet nasjonale prøver og karakterdata i vår forskning. Eksempler på noen slike er:

- Er eleven i stand til å samarbeide med andre mennesker?
- Har eleven respekt for mennesker fra andre kulturer og i andre livssituasjoner?
- Kan eleven tilegne seg ny kunnskap og gjøre nytte av dette i hverdagslivets utfordringer?
- Opplever eleven kulturell tilhørighet?
- Har eleven gode verdier?
- Er eleven i stand til å balansere tilpasningsdyktighet og selvstendighet?
- Vet eleven hvordan å utfolde sin utforskertrang?
- Handler eleven miljøbevisst?

Et eksempel på måling av slike perspektiver finner vi i ICCS-studien (Schulz mfl 2010), hvor ungdoms evne til demokratisk deltakelse undersøkes. Resultatene fra 2009 tyder på at norske elever er godt skikket til aktiv og reflektert deltakelse i et demokrati. Slike perspektiver bør inkluderes om man ønsker en helhetlig beskrivelse av skolens måloppnåelse. En kunnskapsprøve tegner alene ikke et tilstrekkelig bilde av kvaliteten ved norsk skole.

3.6 Hvor lett er det å måle bredden av skolens målsettinger?

Opplæringslovas mangfoldige siktemål gjør det naturlig nok til en omfattende jobb å vurdere hvorvidt man oppnår alt det man ønsker med norsk skole. Om man vil teste evner og holdninger som de nevnt over, støter man raskt på store utfordringer. Flere av egenskapene er både vanskelig å definere og enda vanskeligere å måle, som for eksempel etisk handling og holdninger som er viktige for å kunne mestre eget liv. Man kan riktignok tenke seg frem til kreative løsninger for hvordan elever kan vurderes etter slike dimensjoner, men løsningene er ofte meget ressurskrevende og ikke enkle å enes om. ICCS-studien vedrørende demokratideltakelse er et godt eksempel på hva som kan gjøres. Datamaterialet fra denne studien er det derimot ikke mulig å koble til kjennetegn ved elevenes ulike klasserom. Man kan derfor ikke bruke dette til å identifisere hvilke prosessfaktorer som i størst mulig grad fremmer slike viktige mål.

Videre er det et poeng at mange av skolens opplæringsmål er individuelle. Noen generelle egenskaper og holdninger skal formidles til og utvikles hos alle, enten de har en fremtid som tannlege eller truckfører. Samtidig er det slik at behovet for enkelte ferdigheter og opplevelser varierer fra person til person. Opplæringen både kan og bør gi ulike utslag avhengig av elevenes ulike forutsetninger og livssituasjoner. Det finnes altså dimensjoner det blir unaturlig å vurdere alle elever etter, ettersom mange viktige læringsmål er prosessorienterte og ikke forhåndsdefinerte (Eisner 1979).

Enkelte opplæringsmål kan også være kontroversielle. Å måle hvorvidt en person innehar gode verdier er ikke bare en testteoretisk utfordring, men det er også en utfordring med tanke på hva som er

«fasit». Det finnes riktignok noen verdier og holdninger som mange enes om, men variasjonen er fortsatt stor i hva ulike personer betrakter som «grunnleggjande verdier i kristen og humanistisk arv og tradisjon» (Kunnskapsdepartementet 2013).

3.7 Hva om andre prestasjonsmål hadde blitt anvendt?

I tillegg til praktiske, testteoretiske og verdimessige utfordringer med å måle mangfoldet som beskrives i Opplæringslova, bør det også formidles at nasjonale prøver og karakterer gjenspeiler mer enn bare prestasjoner i en gitt testsituasjon. PISA-undersøkelsen, som er et godt eksempel på en ferdighetsprøve elever gjennomfører individuelt med penn og papir, har vist seg å korrelere med flere viktige mål for opplæringen som ikke testes direkte i undersøkelsen. Blant annet er det en sammenheng mellom det å skåre godt på PISA-undersøkelsen og det å delta aktivt i høyere utdanning fire år senere (OECD 2007). Som PISA-oppgavene er oppgavene i nasjonale prøver utviklet av personer med ekspertise innen både bestemte fag og i testteori. Det er god grunn til å tro at også disse resultatene er korrelert med målsettinger som ikke testes direkte i prøvene.

Korrelasjonen mellom PISA-resultater og senere deltakelse i høyere utdanning peker på et alternativt prestasjonsmål som kunne blitt benyttet i dette prosjektet: Hvor mange er i utdanning eller jobb et visst antall år etter opplæringen? Dette kan anses som en indikasjon på at personen har oppnådd flere av skolens viktige målsettinger. Ytterligere et eksempel er inkludert i spørreundersøkelsen som ble presentert i delrapport 4 (Opheim og Wiborg 2012). Der ble elevenes selvvalgte kompetanseoppnåelse analysert. Inneholdt i et slikt mål kan både deres kompetanse og deres selvtilit og mestringsopplevelse være inkludert, hvor sistnevnte to punkter også er viktige formål med opplæringen. Hvor godt ville slike prestasjonsmål fungert? Og ville man da sannsynligvis ha kommet til samme konklusjoner som vi har kommet til ved bruk av karakterdata og data fra nasjonale prøver?

Vedrørende deltakelse i jobb og utdanning, er det mye som tyder på at det kan være krevende å finne entydige kategorier, tatt i betraktning de mange tilfellene av deltidsjobbing, deltidsstudier og ulike former for avbrekk i studier og jobb. Videre ville avstanden fra erfaringene i ulike klasserom frem til studie- og jobbsituasjonen flere år senere gjort det vanskelig å fastslå noe om bestemte prosessfaktorer. Hvilke prosessfaktorer i hvilke klasserom var det som gjorde at personen lyktes i jobbmarkedet? Selvvalgt kompetanseoppnåelse har i så måte større potensial som prestasjonsmål når prosessfaktorer skal identifiseres, ettersom det er større tematisk og tidsmessig nærhet fra prosess til prestasjonsmål.

Det er vanskelig å si om resultatene i dette prosjektet hadde blitt de samme om flere analyser hadde gjort nytte av slike datakilder. Analysene gjort i delrapport 4 (Opheim og Wiborg 2012) gjør det mest nærliggende å foreslå at de fleste resultatene hadde forblitt relativt uforandret: Den selvvalgte kompetanseoppnåelsen ser ut til å samsvare godt med de selvrapporterte karakterene, som igjen samsvarer godt med de faktiske karakterene. Andre studier viser sammenhenger mellom karakterer og blant annet lykke (Quinn og Duckworth 2007). En metastudie gjennomført i 2004 tyder på at skoleprestasjoner henger så tett sammen med både kreativitet og suksess på arbeidsmarkedet at forfatterne forsøker å utvikle et begrep som favner alle disse aspektene (Kuncel mfl 2004). Som nevnt viser også PISA-studien en sammenheng mellom resultater på PISA og deltakelse i utdanning og arbeidsliv. Disse resultatene inngår i argumentasjonen for at karakterer og resultater på nasjonale prøver er hensiktsmessige prestasjonsmål som indikerer måloppnåelse innen en større bredde av skolens målsettinger enn det som måles spesifikt i testene.

3.8 Begrunnelsen for å benytte data fra nasjonale prøver og karakterdata som prestasjonsmål

I det foregående har vi pekt på svakheter ved å benytte data fra nasjonale prøver og karakterdata som indikatorer på resultatkomponenten i input-prosess-resultat-modellen (Scheerens 1990). Disse datakildene spenner ikke over bredden av Opplæringslovas målsettinger (Kunnskapsdepartementet

2013), mange svaktpresterende elever er ikke representert i datamaterialet og muligheten for å studere prestasjonsutvikling er begrenset til en relativ, og ikke absolutt, utvikling.

Likevel vurderes nasjonale prøver og karakterdata til å være de datakildene hvor man får de best kvalitetssikrede og representative prestasjonsmålene med tanke på ressursene som kreves for å samle inn slikt datamateriale. Mange av Opplæringslovas målsettinger er vanskelige å måle. Solide vurderinger ligger stort sett bak når elevene får karakterer, og mangelen på karakterer i 5. trinn betyr at data fra nasjonale prøver er nødvendig. De nasjonale prøvene holder høy testteoretisk kvalitet, og data blir samlet inn fra en stor andel av elevene. Videre ble det argumentert for at en rekke av skolens målsettinger gjenspeiles i karakterdata og resultater fra nasjonale prøver. Prestasjonsmålene benyttet her, korrelerer med andre viktige mål med opplæringen.

Begrensningene i datamaterialet som er diskutert her, fører til at man likevel må trekke slutninger med forsiktighet. Enkelte forhold kunne økt relevansen, treffsikkerheten og analysemulighetene i dette prosjektet. Eksempelvis kunne man her benytte seg av data vedrørende flere av skolens «mykere» målsettinger hvor selvrapporteringer ser ut til å fungere. Elevenes opplevelse av tilhørighet, selvtillit og mestring er eksempler på viktige resultater av opplæringen som kan måles i blant annet Elevundersøkelsen. Noe av dette inngikk også i spørreundersøkelsen til elever på 9. trinn og i Vg1 (Opheim og Wiborg 2012). Analyser av spørreundersøkelsen, tyder imidlertid på relativt høy korrelasjon mellom elevenes opplevelse av mestring av ferdigheter og deres karakterer innen samme fag.

Den statistiske treffsikkerheten vil øke dersom man fortsetter arbeidet med å få en enhetlig fritakspraksis på nasjonale prøver. Videre ville fortjenesten vært stor av å ha datakilder som åpnet for muligheten til å se på reell prestasjonsutvikling. Her ser vi en positiv utvikling i at man nå skal la noen av oppgavene i nasjonale prøver gå igjen fra år til år. Forutsetningen for en slik praksis er at disse oppgavene blir holdt hemmelig, noe som har fungert godt blant annet i PISA- og TIMSS-undersøkelsene. Mulighetene for å identifisere kvalitetsindikatorer i norsk skole vil øke med økt kvalitet på ulike prestasjonsmål.

4 Informasjon fra GSI som indikatorer på ressursbruk

Opplysninger fra Grunnskolens informasjonssystem (GSI) har inngått i mange av analysene i prosjektet «Ressurser og resultater». I dette kapitlet skal vi diskutere GSI som indikator på kvalitet i skolen og i hvilken grad den tilgjengelige informasjonen forklarer prestasjonsforskjeller. Diskusjonene og forslagene er basert på analyser i prosjektets delrapporter, som er omtalt i denne rapportens kapittel 2.

GSI gir informasjon om de forhold som skolen jobber under. Sett opp mot den forenklede input – prosess – resultat modellen som ble presentert i kapittel 1, betyr det at opplysningene i GSI kun berører den første delen av modellen.

Kapitlet har som utgangspunkt å finne ut i hvilken grad data fra GSI kan brukes som kvalitetsindikatorer for skolesektoren. Dette er imidlertid ikke det primære formålet til informasjonssystemet. GSI er en sentral kilde til informasjon om skolen, uavhengig av vurderingene som gjøres her.

Innledningsvis i kapitlet presenterer vi kort GSI og GSI-indikatorer som er funnet å være relatert til elevenes prestasjoner i prosjektet "Ressurser og resultater". Deretter diskuteres styrker og svakheter ved GSI. Avslutningsvis skal vi foreslå enkelte endringer som kan gjøre opplysninger fra GSI mer relevante for å forstå sammenhengen mellom ressursbruk og skoleprestasjoner.

4.1 Innhold i GSI: Ressurser og rammebetingelser

GSI er et rapporteringsverktøy konstruert for å samle informasjon om ressursbruk og andre forhold ved landets barne- og ungdomsskoler. Skolene gir opplysninger om en rekke ulike forhold, som elevtall, timetall, ressurser, spesialundervisning, språklige minoriteter, målform, skyss, leirskole, tilvalgsfag, SFO, IKT og PPT. Til sammen registreres omtrent hundre forskjellige opplysninger for hver av landets barne- og ungdomsskoler, hvert år. Det betyr at GSI gir informasjon om forhold på skolenivå, altså forhold som varierer mellom skolene og som er felles for alle elevene ved en skole, og at GSI-dataene er objektive i den forstand at de uttrykker forhold ved skolen og arbeidet der som er målbart og konkret, og ikke grupper opplevelser eller synspunkter. Databasen inkluderer informasjon fra 1992, da GSI ble opprettet, og fram til i dag. En rekke ikke-sensitive opplysninger ligger tilgjengelig på internett, og kan hentes ut på skole-, kommune-, fylkes- og nasjonalt nivå.

En del av opplysningene i GSI er typiske indikatorer på skolens ressursbruk, som timetall, antall PC-er og lærertetthet. Selv om mye av skolens ressursbruk er lovregulert, som regel i form av minimumsinnsats innen ulike områder er dette samtidig ressursbruk som kan variere mellom skoler.

Andre mål i GSI er åpenbart ikke indikatorer på skolenes ressursbruk, som hvor skolen er plassert geografisk. Beliggenhet er en av de betingelsene skolene må forholde seg til og ikke en indikator på hvordan skolen bruker tilgjengelige ressurser. Samtidig er dette forhold som kan ha betydning for skolens ressursbruk – for eksempel vil utgifter til skoletransport per elev være høyere i mindre tettbygde strøk enn i byene.

Enkelte forhold kan på den ene siden sees på som indikatorer på skolens ressursbruk, og på den andre siden sees på mål på de forholdene skolene jobber under. Dette gjelder for eksempel andelen lærere med godkjent utdanning. På den ene siden kan andelen indikere i hvilken grad skolen er opptatt av å ansette lærere med godkjent utdanning. På den andre siden kan andelen være uttrykk for skolens rekrutteringsgrunnlag – i hvilken grad det er mulig for skolene å rekruttere lærere med godkjent utdanning.

Selv om ikke alle indikatorene handler om skolenes ressursbruk, er de likevel relevante for å forstå forskjeller i prestasjoner. Ytre forhold påvirker skolenes muligheter, og er derfor av betydning for prestasjoner. Opplysninger om dette er derfor viktig for å forstå prestasjoner. Det betyr at når vi studerer betydningen for prestasjoner av indikatorer registrert i GSI, må vi skille mellom hva som er ressursbruk og hva som måler rammebetingelser eller kjennetegn ved skolen. Videre er det viktig, slik det har blitt gjort i analysene i prosjektet «Ressurser og resultater», å ta hensyn til de forhold skolene jobber innenfor når betydningen av ressursinnsatsen vurderes.

4.2 Informasjon fra GSI som bidrag til å forklare prestasjoner

Vi skal her kort presentere indikatorer fra GSI som inngår i analysene i prosjektet «Ressurser og resultater», og som har blitt testet som mulige kvalitetsindikatorer:

Timetall: Høyere antall undervisningstimer på skolen er positivt relatert til elevenes prestasjoner.

Antall PC-er per elev: Både på barneskolen og på ungdomstrinnet finner vi at høyere antall PC-er per elev er positivt relatert til elevenes prestasjonsnivå.

Andel lærere med godkjent utdanning: Høyere andel lærere med godkjent utdanning er positivt knyttet til prestasjoner på 5. trinn, men negativt relatert til prestasjoner på 8. trinn.

Lærernes kjønnsfordeling: Andel kvinnelige lærere er positivt relatert til prestasjoner på barneskolen, men negativt relatert til prestasjonsutvikling fra 8. til 10. trinn.

Private skoler: Sammenlignet med elever på 10. trinn ved offentlige skoler har elever på samme trinn ved privatskoler noe bedre prestasjoner i norsk og engelsk, mens de presterer noe svakere i matematikk.

Geografisk beliggenhet: Skoler i Akershus og Oslo har høyere gjennomsnittlige prestasjoner enn skoler i resten av landet, også når kjennetegn ved eleven, som foreldrenes utdanning, er tatt hensyn til.

Klassestørrelse: Antall elever i klassen er i analysene ikke funnet å ha betydning for elevenes prestasjoner.

Lærertimer per elev til annet enn undervisning (som sosialpedagogisk tjeneste, kontaktlærertjeneste for elevråd, særskilt byrdefull undervisning, kontaktlærere, leseplikt og fjernundervisning eller annen undervisning): Hvor mange timer som ble brukt av lærerne til andre skolerelaterte forhold enn undervisning er ikke knyttet til elevenes prestasjoner på skolen.

Antall assistenttimer per elev: Analysene viser at dette ikke er av betydning for prestasjonene.

Skolestørrelse: Antall elever ved skolen er i prosjektet ikke funnet å være av betydning for prestasjonene.

To av variablene står igjen som mulige kvalitetsindikatorer: timetall og antall PC-er per elev. De to variablene var relatert til elevenes prestasjoner selv når en rekke andre forhold er kontrollert for. Selv om også andre forhold ble funnet å være knyttet til elevenes prestasjonsnivå er vår vurdering at de ikke kan sees på som kvalitetsindikatorer. Grunnen til det er at de enten beskriver forhold som ikke kan sees på som ressursbruk (som geografisk beliggenhet), eller at de påviste sammenhengene ikke er entydig positive eller negative (det er tilfelle for andel lærere med godkjent utdanning, lærernes kjønnsfordeling og private skoler).

4.3 Styrker og svakheter ved GSI

Sammenhengene mellom GSI-indikatorer og skoleprestasjoner som kommer fram i analysene, har blitt diskutert i og problematisert i mange av prosjektets delrapporter. På bakgrunn av erfaringer fra «Ressurser og resultater» skal vi diskutere sterke og svake sider ved GSI. Hensikten er å vurdere GSI som utgangspunkt for å utvikle indikator på kvalitet i skolen.

4.3.1 Mye informasjon

En styrke ved GSI er omfanget av tilgjengelig informasjon. For det første inkluderer datasettet alle landets grunnskoler. Det betyr at det er mulig å undersøke hele populasjonen av skoler, og opplysningene kan aggregeres til kommunalt-, fylkes- eller nasjonalt nivå og fremdeles være representativt. For det andre inneholder GSI mange og detaljerte opplysninger, noe som gjør det mulig å undersøke forskjeller mellom skoler innenfor en rekke områder. De aller fleste av opplysningene i GSI er ikke tilgjengelig i andre dataregistre. For å studere betydningen av forhold på skolenivå er slike data helt avgjørende. For det tredje er det registrert opplysninger i GSI siden 1992. Dermed er det mulig å følge og sammenlikne ressursbruk over tid på skolenivå. Videre gjør dataene det mulig å gå tilbake i tid og undersøke sammenhengen mellom ressursbruk og resultater i ulike perioder.

GSI gir mye informasjon om variabler på skolenivå. Imidlertid viser analysene i prosjektets delrapporter at en relativt liten andel av variansen i skoleprestasjoner kan forklares av dette nivået. Som omtalt i kapittel 2, er det kun om lag 7 til 10 prosent av variansen i elevenes prestasjonsnivå som kan tilbakeføres til forskjeller mellom skoler (Opheim mfl 2010). Dette innebærer at forhold på skolenivået, som GSI-indikatorene måler, i utgangspunktet ikke kan forklare mer enn en begrenset del av variasjonene i elevenes prestasjoner. Selv om opplysningene i GSI kunne forklare all varians på skolenivået, ville fremdeles om lag 90 til 93 prosent av prestasjonsforskjellene ikke være gjort rede for. For å forstå elevenes prestasjoner fullt ut er det nødvendig med supplerende opplysninger i tillegg til det som er tilgjengelig gjennom GSI.

At forhold på skolenivå har begrenset mulighet til å forklare den totale variansen i elevenes prestasjoner betyr ikke at det er uviktig å fokusere på dette. Kunnskap om hvordan forskjeller mellom skoler er knyttet til prestasjonsforskjeller er interessant fordi det kan si noe om hva som er god praksis for skolene. Å bedre praksis kan redusere forskjellene mellom skolene gjennom å løfte prestasjonene til de svakt presterende. Selv små endringer vil være positivt for dem det gjelder.

4.3.2 Uklart hva som måles – indirekte kvalitetsindikatorer?

I noen tilfeller er den tilgjengelige informasjonen i GSI svært begrenset. Dermed er det ikke alltid tydelig hva indikatorene egentlig måler. I tilfeller der dette er uklart blir det vanskelig å fastslå hva eventuelle sammenhenger mellom opplysninger i GSI og skoleprestasjoner betyr. Vi skal eksemplifisere dette nærmere ved å diskutere hvordan betydningen av fire GSI-indikatorer, som er funnet å være knyttet til prestasjoner, kan forstås.

Antall PC-er per elev

Analysene viser at antall PC-er per elev er knyttet til prestasjoner. Det kan forstås på flere måter. Den umiddelbare tolkningen er å se tilgangen til PC som en årsak til elevens prestasjoner: flere PC-er er positivt for skoleprestasjonene fordi det bedrer læringen. En slik tolkning er imidlertid ikke uproblematisk. Informasjonen i GSI viser ikke i hvilken grad PC-ene tas i bruk i undervisningen.

Dataene forteller bare at enkelte skoler har flere PC-er per elev enn andre, og at dette er positivt for skolens gjennomsnittlige prestasjonsnivå. Dermed vet vi ikke om sammenhengen skyldes at elever på skoler med mange PC-er faktisk bruker dem i skolearbeidet i større grad enn elever på andre skoler. Videre kan resultatet skyldes bakenforliggende variabler, at forhold ved skolen både påvirker elevenes prestasjoner og hvor mange PC-er skolen har. Et slikt forhold kan være skolens lærerressurser. Skoler med datakyndig faglig personale som særlig etterspør og har kunnskap om hvordan PCer kan utnyttes i undervisningen, kan velge å investere i elev-PCer i større grad enn andre skoler. Forholdet mellom antall elever og antall PC-er kan handle om i hvilken grad skolen prioriterer læringsutstyr til elevene. Skoler som mener dette er viktig, kan både ha mange PC-er og annet utstyr som kan brukes i undervisningen. Sammenhengen mellom andelen elev-PCer og elevenes prestasjoner, kan med andre ord involvere flere forhold ved skolen – og ikke nødvendigvis skyldes at flere elev-PCer har en direkte positiv betydning for elevenes læring og skoleprestasjoner. Det kan likevel være verdt å merke seg det gjennomgående positive utslaget av denne ressursindikatoren i analysene.

Andel lærere med godkjent utdanning

Andel lærere med godkjent utdanning er funnet å være relatert til prestasjonsnivået. På barneskolen er sammenhengen positiv: høyere andel lærere med godkjent utdanning er knyttet til bedre prestasjoner. Analyser i delrapport 2 (Wiborg mfl 2011) viser at barneskoler som har relativt mange lærere uten godkjent utdanning, også har mange førskolelærere i staben. Det indikerer at lærere uten godkjent utdanning i mange tilfeller kan være førskolelærere som underviser på trinn som utdanningen deres ikke er godkjent for. Tolkningen problematiserer bruken av førskolelærere på enkelte trinn i grunnskolen.

På ungdomsskolen er sammenhengen negativ: en høyere andel lærere med godkjent utdanning er knyttet til lavere prestasjoner. Også i dette tilfellet er det, på bakgrunn av opplysninger i GSI, vanskelig å slå fast hvorfor sammenhengen forekommer. En mulig forklaring, som diskuteres i delrapport 1 (Opheim mfl 2010), er at lærere kategorisert som «uten godkjent utdanning» har høy utdanning, men mangler den nødvendige pedagogiske kompetansen. Høyere utdanning fra universitet eller høyskole som ikke også inkluderer pedagogikk, regnes ikke som godkjent utdanning for å undervise på ungdomsskolen. Resultatene av analysene peker på at lærere med slik utdanning er positivt for prestasjonsnivået på skolen. Imidlertid kan vi ikke konkludere med sikkerhet på dette punktet, fordi vi ikke vet hva som ligger i kategorien «ikke godkjent utdanning». Selv om vi kan sannsynliggjøre at det er lærere med høyere utdanning i denne kategorien, kan vi ikke si det med sikkerhet. Hadde vi kunnet fastslå at dette handler om skoler med flere lærere med ikke-pedagogisk høyere utdanning, ville det vært et vesentlig innspill til debatten om kvalitet i skolen. Slik situasjonen er nå, blir dette forholdet bare delvis belyst.

Vi skylder å gjøre oppmerksom på at ikke alle analysene i prosjektet finner en sammenheng mellom andel lærere med godkjent utdanning ved skolen og elevenes prestasjonsnivå. Analysene i delrapport 5 (Grøgaard 2012), viser ingen sammenheng mellom andelen lærere med godkjent utdanning og om skolen oppnår høy skår på de nasjonale prøvene. Disse analysene omfatter imidlertid kun barneskoler.

Kjønnsfordeling blant lærerne

Når det gjelder betydningen av andelen kvinnelige og mannlige lærere, varierer resultatene i delrapportene. Høy andel kvinnelige lærere er funnet å være positivt for elevers prestasjoner på 5. trinn (Opheim mfl 2010), positivt relatert til prestasjonsutvikling fra 5. til 8. trinn (Wiborg mfl 2011) og negativt relatert til prestasjonsutvikling fra 8. til 10. trinn (Wiborg mfl 2011). Delrapport 5 (Grøgaard 2012) viser ingen sammenheng mellom kjønnsfordelingen blant lærerne og elevenes prestasjoner. Resultatene kan tolkes på flere måter. For det første er det mulig at kvinnelige lærere er positivt for læringsmiljøet og dermed for prestasjonene på barneskolen, mens mannlige lærere er positivt på ungdomsskolen. Imidlertid vet vi ikke om det er forskjeller mellom kvinnelige og mannlige lærere som velger å jobbe henholdsvis på barne- og ungdomsskolen. Er det for eksempel systematiske forskjeller i hva slags utdanning menn og kvinner på de ulike skoleslagene har? Slike forhold kan bidra til å

forklare hvorfor vi finner at kjønnsfordelingen på skolen er relatert til prestasjoner. Imidlertid er det ikke mulig ut fra opplysningene i GSI å analysere dette nærmere. Dermed blir det vanskelig å fastslå nøyaktig hvordan kjønnsfordelingen blant lærerne er relatert til elevenes prestasjoner.

Lærertetthet

Ingen av analysene i prosjektet har vist at antall lærere per elev er knyttet til elevenes prestasjoner: det er ikke vesentlige forskjeller i elevenes prestasjonsnivå på skoler avhengig av lærertettheten. Den umiddelbare tolkningen av mangelen på sammenheng mellom disse forholdene er at antall elever per lærer er uten betydning for elevenes læring. Imidlertid kan mangelen på sammenheng også skyldes at GSI-dataene om antall lærere per elev ikke er spesifikke nok. Riktignok gis det informasjon om hvor mange lærere det er i forhold til antall elever. Men det er ikke ensbetydende med at lærerressursene kommer den enkelte elev til gode. For eksempel er det mulig at lærertettheten er spesielt høy på skoler med store utfordringer, som for eksempel atferdsvansker blant elevene. På slike skoler kan de ekstra lærerressursene være mer rettet mot å opprettholde et visst nivå av ro og orden framfor å gjøre undervisningen bedre. Det er dermed ikke å forvente at de ekstra lærerressursene på disse skolene gir et generelt økt prestasjonsnivå. Videre kan det være at en liten gruppe elever med særlig behov for oppfølging, mottar mye av lærerressursene ved en skole, mens resten av elevene har få lærere. Slike skjevfordelinger innad den enkelte skole gir ikke GSI opplysninger om. Selv om det er opplysninger om antall elever med spesialundervisning, er informasjonen i GSI ikke detaljert nok til å skille skoler med høyt antall lærere per elev på grunn av spesielle utfordringer fra skoler som har høyt antall lærere per elev uten slike utfordringer. For å undersøke sammenhengen mellom lærertetthet og prestasjoner må vi vite ikke bare forholdstallet mellom elever og lærere ved en skole, men også hvilke lærerressurser den enkelte elev eller den enkelte klasse har hatt. Slik situasjonen er i dag, kan vi ikke klart konkludere når det gjelder betydningen av antall lærere per elev.

4.4 Stabilitet og endring i timetall og antall PC-er per elev

Timetall og antall PC-er per elev er i prosjektet «Ressurser og resultater» gjentatte ganger funnet å være positivt relatert til elevenes prestasjoner. For ytterligere å vurdere hvor egnet variablene er som indikatorer på kvalitet i skolen skal vi undersøke hvor stabile de er over tid, noe som i liten grad er gjort tidligere i prosjektet. Dersom timetall og antall PC-er er lite stabile over tid reduserer det nytten av å bruke variablene som kvalitetsindikatorer, siden innsamlet informasjon kun vil ha kort gyldighet. Opplysninger fra en måling kan være foreldet etter kort tid, og dermed lite egnet til å beskrive situasjonene i skolen. Dersom indikatorene er stabile over tid er det lettere å bruke dem som indikatorer på kvalitet, siden det er mer sannsynlig at den innsamlede informasjonen vil gjelde for en lengre periode.

I enkelte av analysene i dette prosjektet har GSI-indikatorene for flere år blitt slått sammen. Det betyr at det ikke er verdien på en GSI-indikator for ett enkelt år som er analysert, men gjennomsnittet av tre påfølgende år: 2007, 2008 og 2009. I andre tilfeller er det verdien på indikatorene for et enkelt år som er benyttet. Å bruke et gjennomsnitt av flere år reduserer sannsynligheten for at variasjon mellom skolene skyldes tilfeldige svingninger. Samtidig vil det å bruke en gjennomsnittsverdi for tre år bety at variabelen i mindre grad fanger forskjeller fra år til år.

Vi skal undersøke stabilitet og endring i variablene på to måter. For det første skal vi finne ut hvordan gjennomsnittet for alle skolene varierer for årene 2006 – 2009. For det andre skal vi undersøke de enkelte skolenes stabilitet i timetall og PC-er per elev i den samme perioden.

Timetall og PC-er per elev: skolen sett under ett

Analysene av hvordan gjennomsnittet endrer seg over tid tar utgangspunkt i timetall per år ved skolene på enkelte trinn i perioden 2006 – 2009. Resultatene vises i tabell 4.1.

Tabell 4.1: Timetall i årene 2006 – 2009. Gjennomsnitt for 2.-4. trinn, 5.-7. trinn og 8.-10. trinn.

	År	Gj.snitt	Std.avvik
2.-4. trinn (N = 2 491)	2006	667,47	36,48
	2007	668,54	39,79
	2008	712,82	27,82
	2009	716,48	29,70
5.-7. trinn (N = 2 440)	2006	777,69	29,95
	2007	777,86	31,37
	2008	778,75	27,08
	2009	779,49	26,50
8.-10. trinn (N = 1 234)	2006	859,99	33,81
	2007	860,49	31,03
	2008	860,27	26,65
	2009	859,54	22,62

Note: N angir minimumsantallet skoler som inngår i datamaterialet per år. Antallet skoler vil variere noe fra år til år.

Tabellen viser at timetallet på 2. til 4. trinn har økt fra 2006 til 2009; samlet antall timer i løpet av året er høyere mot slutten av perioden enn ved begynnelsen. Mye av økningen skjer fra 2007 til 2008. Trolig kommer dette av at det ble innført flere timer i barneskolen høsten 2008 som en del av Kunnskapsløftet. På de andre trinnene er endringene over tid minimale. Alt i alt tyder gjennomsnittstallene på stor grad av stabilitet i timetall når skolene vurderes under ett. Videre viser tabellen at standardavviket reduseres over tid. Det indikerer at forskjellene mellom skolene blir mindre. Dette gjelder for alle de trinnene som er undersøkt, og innebærer at flere skoler ligger tettere opptil skolene gjennomsnittlig timetall i 2009 enn i 2006.

Tabell 4.2: Antall PC-er per elev i årene 2006 - 2009: Gjennomsnitt for barneskoler og ungdomsskolen.

	År	Gj.snitt	Std.avvik
Barneskoler (N = 2 525)	2006	0,28	0,43
	2007	0,31	0,20
	2008	0,34	0,20
	2009	0,36	0,21
Ungdomsskoler (N = 1 234)	2006	0,30	0,22
	2007	0,35	0,23
	2008	0,39	0,23
	2009	0,41	0,26

Note: N angir minimumsantallet skoler som inngår i datamaterialet per år. Antallet skoler vil variere noe fra år til år.

Når det gjelder antall PC-er per elev, viser tabellen at gjennomsnittet øker i løpet av den undersøkte perioden, både på barneskolen og ungdomsskolen. Elevene har altså, i gjennomsnitt, flere PC-er til rådighet i 2009 enn i 2006. Tabellen viser også at standardavviket er relativt stort sammenlignet med

gjennomsnittsverdien. Det innebærer at variasjonene mellom skolene er stor: tilgjengeligheten til PC-er for elevene varierer betydelig avhengig av hvilken skole de går på. Videre finner vi ikke den samme tydelige tendensen til at standardavviket reduseres i løpet av perioden som vi fant for timetall. Det innebærer at forskjellene mellom skolene ikke blir mindre fra 2006 til 2009. Tvert imot tyder tallene på at det er en svak økning i standardavviket, altså at forskjellen mellom skolene blir større. Imidlertid er økningen liten, og den undersøkte perioden er relativt kort, så det er ikke mulig å fastslå hvorvidt dette er en tendens som varer over tid, eller om det skyldes tilfeldige svingninger.

Timetall og PC-er: Endring på den enkelte skole

Så langt har de presenterte tallene vist gjennomsnittet for alle skoler i Norge. Selv om størrelsen på standardavvik sier noe om variasjonen mellom skolene er en gjennomsnittsverdi i liten grad egnet til å fange opp utviklingen ved den enkelte skole. For å finne ut hvordan timetall og antall PC-er per elev varierer over tid på den enkelte skole skal vi foreta korrelasjonsanalyser. Gjennom å korrelere timetall ved et tidspunkt på en skole med timetall på andre tidspunkt ved den samme skolen får vi informasjon om stabiliteten til dette målet ved den enkelte skole. Det samme gjelder for antall PC-er per elev. Korrelasjonen viser i hvilken grad skolens relative plassering i antall timer og PC-er per elev opprettholdes over tid. En høy korrelasjon behøver ikke bety at skolen har det samme antall timer ved to anledninger, men viser at skolens relative plassering sett i forhold til de andre skolene når det gjelder antall timer, ikke varierer i særlig grad.

Tabell 4.3: Korrelasjon mellom timetall i årene 2006 – 2009.

		2006	2007	2008	2009
2.-4. trinn (N = 2 436)	2006	1			
	2007	0,78	1		
	2008	0,70	0,72	1	
	2009	0,63	0,67	0,78	1
5.-7. trinn (N = 2 375)	2006	1			
	2007	0,78	1		
	2008	0,72	0,75	1	
	2009	0,57	0,60	0,79	1
8.-10. trinn (N = 1 163)	2006	1			
	2007	0,64	1		
	2008	0,61	0,69	1	
	2009	0,59	0,61	0,74	1

Note: N angir minimumsantallet skoler som inngår i datamaterialet per år. Antallet skoler vil variere noe fra år til år.

En korrelasjon på 1 innebærer at alle skolene har den samme plasseringen ved de to tidspunktene, mens en korrelasjon på 0 hadde betydd at ingen av skolene hadde den samme plasseringen ved de to målingene. Når vi finner, som tabell 4.3 viser, at korrelasjonene ligger mellom 0,57 og 0,79, innebærer det relativt stor grad av stabilitet. Skoler som ved et tidspunkt har mange timer i forhold til de andre skolene, har også mange timer ved de andre tidspunktene.

Samtidig som korrelasjonene indikerer relativt stor grad av stabilitet er det ikke slik at forholdet mellom skolene er statisk. Det er visse endringer i hvilke skoler som har få og mange timer. Når tiden mellom målingene øker synker korrelasjonene. Det betyr at forandringene i den relative plasseringen er større over tre år enn over et år. Likevel er korrelasjonen mellom den første og den siste målingen relativt høy, uavhengig av hvilke trinn som undersøkes.

En tilsvarende korrelasjonsmatrise for antall PC-er per elev presenteres i tabell 4.4.

Tabell 4.4: Antall PC-er per elev. Korrelasjon mellom antall PC-er på skolen i 2006, 2007, 2008 og 2009.

		2006	2007	2008	2009
Barneskoler (N = 2 477)	2006	1			
	2007	0,80	1		
	2008	0,71	0,80	1	
	2009	0,63	0,71	0,80	1
		2006	2007	2008	2009
Ungdomsskoler (N = 1 163)	2006	1			
	2007	0,76	1		
	2008	0,64	0,77	1	
	2009	0,54	0,68	0,77	1

Note: N angir minimumsantallet skoler som inngår i datamaterialet per år. Antallet skoler vil variere noe fra år til år.

Også når det gjelder antall PC-er per elev er korrelasjonene relativt høye. Korrelasjonene ligger mellom 0,80 og 0,54, noe som er på nivå med det vi fant for timetall. Også her er det slik at korrelasjonene synker når avstanden i tid øker. Jo lenger tid det er mellom målingene, jo mer endres skolens plassering i forhold til de andre skolene. Men selv når perioden mellom målingene er tre år er korrelasjonene relativt høye.

Det innebærer at skoler som ved et tidspunkt har god PC-dekning også har det ved de andre tidspunktene. I tabell 4.2 kom det fram at antallet PC-er per elev øker i den perioden som undersøkes her. Det vil si at samtidig som dekningen av PC-er i norske skoler blir bedre og bedre, så er det i stor grad slik at de skolene som hadde best dekning i 2006 også er blant skolene med best dekning i 2007, 2008 og 2009.

Konklusjon

Samlet viser analysene at timetall og antall PC-er synes å være relativt stabile over tid. Gjennomsnittet for alle skolene i Norge endrer seg i liten grad i løpet av perioden som undersøkes, og korrelasjonene tyder på at skoler med høyt timetall og høyt antall PC-er ved et tidspunkt ofte vil ha tilsvarende plasseringer også ved de senere tidspunkt som inngår i tabellene. Samtidig finner vi visse svingninger innad ved den enkelte skole i de årene vi har studert. Analysene peker på at det å bruke gjennomsnitt for tre år i analysene er en rimelig framgangsmåte. Tilnærmingen reduserer betydningen av svingninger innad ved den enkelte skole fra år til år, samtidig som det er sannsynlig at forskjeller mellom skolene blir ivaretatt.

Analysene viser at timetall og antall PC-er per elev ikke varierer mye i de årene vi sammenligner. Det styrker de to variablene som mulige indikatorer på kvalitet i skolen: Timetall og antall PC-er per elev synes å være knyttet til elevenes prestasjoner, samtidig som de synes å være relativt stabile over tid.

4.5 Oppsummering

Kapitlet viser at GSI inneholder mye informasjon som kan brukes i analyser av skolens ressurstilgang og ressursbruk. Imidlertid er GSI ikke utformet med forskning som formål, og sammenhengen mellom indikatorer og skoleprestasjoner kan i mange tilfeller problematiseres.

Et av de mest fundamentale problemene med GSI er at vi ikke vet hvordan ressursbruk på skolenivå er relatert til den enkelte elev. Når vi ikke vet nøyaktig hvordan ressursene brukes er det vanskelig å fastslå hva som er effektene av forskjeller i ressurser. Dette ble diskutert i forbindelse med lærertetthet og antall PC-er per elev, men kan være en relevant innsigelse også for andre forhold som registreres i GSI.

Analysene i prosjektet viser at andelen lærere med godkjent utdanning er relatert til elevenes prestasjoner. Imidlertid inneholder GSI lite informasjon om lærerne uten godkjent utdanning, og vi kan ikke fastslå om de er førskolelærere som jobber på trinn hvor utdanningen deres ikke er godkjent, om de har høyere utdanning uten pedagogikk, eller om de ikke har høyere utdanning. Mer detaljerte opplysninger om lærernes utdanningsnivå og kvalifikasjoner vil bidra til å forstå sammenhengen mellom lærernes kompetanse og elevenes prestasjoner. Ytterligere opplysninger om lærerne, som alder, ansiennitet og tidligere arbeidserfaring, ville gitt større muligheter til å undersøke sammenhenger mellom elevenes prestasjoner og skolens lærerressurser. Videre ville det være positivt om informasjon fra andre dataregistre om læreres etter- og videreutdanning, kunne brukes som informasjon om den enkelte skole. Her ville det være viktig å se hvilke fag etterutdanningen ble tatt i, for eksempel i hvilken grad det gjelder skolens basisfag.

I prosjektet har antall undervisningstimer i flere analyser blitt funnet å være positivt knyttet til elevenes prestasjoner. Til tross for at skolens/klassestrinnets timetall ikke sier noe om innholdet eller kvaliteten på undervisningsøktene, synes likevel dette å fremstå som en ressursfaktor med betydning for elevenes prestasjoner, og dermed en mulig kvalitetsindikator. Timetallet sier heller ingenting om fordelingen av timer mellom ulike fag. Opplysninger om antall timer på ulike klassetrinn i enkelte sentrale fag, som matematikk, norsk og engelsk ville gjort det mulig å undersøke sammenhenger mellom innsats og resultater i spesifikke fag, noe som ytterligere vil kunne belyse betydningen av antall undervisningstimer for elevenes læringsutbytte.

Antall PC-er per elev synes gjennomgående å være positivt knyttet til elevenes prestasjoner. Vi finner at fordelingen er relativt stabil over flere år. Dette kan indikere at antall PC-er per elev kan fungere som en mulig kvalitetsindikator for skolen. Imidlertid er betydningen av denne variabelen uklar, og mindre intuitiv enn betydningen av antall undervisningstimer. Vi kan ikke vite sikkert om det er selve antallet PC-er per elev som har en positiv betydning for elevenes prestasjoner eller om denne sammenhengen skyldes andre forhold ved skolen. Det er med andre ord vanskelig å tolke sammenhengene mellom antall PC-er per elev ved skolen og elevenes gjennomsnittlige prestasjoner.

5 Elevundersøkelsen som indikator på elevenes læringsmiljø

Opplysninger fra Elevundersøkelsen har inngått i mange av analysene i prosjektet «Ressurser og resultater i grunnopplæringen». Som navnet tilsier, inneholder undersøkelsen elevenes egne vurderinger av sin skolehverdag og opplevelser på skolen. Det gjør at den skiller seg klart fra de andre datakildene på nasjonalt nivå som brukes i prosjektet.

I dette kapitlet skal vi se nærmere på Elevundersøkelsen som utgangspunkt for analyser av læringsmiljø og av sammenhengen mellom læringsmiljø og prestasjoner. Det innebærer både å diskutere hvilken informasjon om elevenes læringsmiljø som fanges opp av Elevundersøkelsen og å undersøke hvordan målene på læringsmiljø er relatert til skoleprestasjoner. Diskusjonen tar utgangspunkt i funn i de tidligere publiserte delrapportene i prosjektet (se kapittel 2). I dette kapitlet går vi nærmere inn på ulike sider ved Elevundersøkelsen som utgangspunkt for utvikling av indikatorer på skolekvalitet. Det innebærer å diskutere validiteten og kvaliteten på den informasjonen som fanges opp av Elevundersøkelsen. Vi vil understreke at vurderingene av Elevundersøkelsen som presenteres, er gjort fra vårt ståsted som forskere. Personer som skal bruke Elevundersøkelsen til andre formål enn forskning, kan ha en annen oppfatning av undersøkelsens styrker og svakheter enn det vi har.

Kapitlet starter med en kort gjennomgang av hva Elevundersøkelsen har bidratt med i analyser og resultater prosjektet «Ressurser og resultater i grunnopplæringen». Deretter vurderes sterke og svake sider ved både enkeltvariable/tema som inngår i undersøkelsen og ved Elevundersøkelsen generelt. Kapitlet inneholder også enkelte utdypende analyser av hvor mye av variansen i svarene i Elevundersøkelsen som kan tilbakeføres til individ-, klasse- og skolenivå.

5.1 Elevundersøkelsen: Bakgrunn og innhold

Elevundersøkelsen er en nettbasert spørreundersøkelse beregnet for elever fra og med 5. trinn i grunnskolen til og med 3. trinn i videregående opplæring. Hensikten med undersøkelsen er å få tak i hvordan elevene opplever og vurderer læringsmiljøet på sin skole. Skolene plikter å gjennomføre Elevundersøkelsen i 7. og 10. trinn og i videregående opplæring trinn 1. Selv om skolene må gjennomføre undersøkelsen er det frivillig for elevene å delta. Mange skoler gjennomfører undersøkelsen også for elever på andre trinn. Antallet deltakere er likevel klart høyest på de trinnene hvor gjennomføringen er obligatorisk. Det betyr at Elevundersøkelsen har mange deltakere på ett trinn på barneskolen, ett trinn på ungdomsskolen og ett trinn på videregående. I tillegg er det et betydelig antall elever som deltar også på de andre klassetrinnene.

Siden oppstarten i 2002 har flere hundre tusen elever deltatt i Elevundersøkelsen og svart på spørsmål om hvordan de har det på skolen, hvordan de opplever undervisningen og hvordan de har

det sammen med de andre elevene. Det er trolig flere grunner til at skolene velger å bruke tid og ressurser på å gjennomføre Elevundersøkelsen. For det første er det interessant å vite hvordan elever opplever skolen og det som foregår der. Elevundersøkelsen tar opp tema som relasjon til lærer, mobbing og trivsel. Det er i seg selv verdifullt at elever har gode relasjoner til lærere, at de ikke mobbes og at de trives på skolen. Informasjon om disse forholdene kan både brukes til å forstå elevers skolehverdag og som grunnlag for beslutninger som angår skolen. De senere årene har det blitt publisert en rekke rapporter på bakgrunn av data fra Elevundersøkelsen (Oxford Research 2009, 2010; Wendelborg, Røe og Skaalvik, 2011).

For det andre er det sannsynlig at forholdene som tas opp i Elevundersøkelsen har betydning også utover det å belyse elevenes opplevelser på skolen. Det er stor grad av enighet om at elevenes opplevelse av læringsmiljøet på skolen har betydning for deres faglige utbytte (Nordahl, 2005; Skaalvik og Skaalvik, 2005; Wendelborg, Røe og Skaalvik 2011). Trivsel, motivasjon og gode sosiale relasjoner til lærere og medelever regnes som positivt for læring og dermed for prestasjonene på skolen. Sammenhengen mellom læringsmiljø og skoleprestasjoner bekreftes både av analysene i tidligere delrapporter i dette prosjektet og av andre studier (Bakken 2009a, 2010). Det betyr at Elevundersøkelsen gir kunnskap om forskjeller i hvordan elever opplever læringsmiljøet, forskjeller som har betydning for deres læringsutbytte.

5.2 Funn fra Elevundersøkelsen i prosjektet

Data fra Elevundersøkelsen er brukt i en rekke ulike analyser. Analysene er forskjellige både når det gjelder hvilke variabler som inkluderes, hvilke klassetrinn som inngår i utvalget, hvilke indikatorer på læringsmiljø som brukes, hva som er avhengig variabel og om data fra Elevundersøkelsen brukes på individ- eller skolenivå. Variasjonen i hvordan analysene er bygd opp gjør det rimelig å forvente ulike resultater. Imidlertid viser gjennomgangen av delrapportene at funnene i stor grad peker i samme retning. Resultatene som er tatt med i oversikten er i all hovedsak funnet i flere typer analyser, og kan derfor betraktes som relativt robuste. Et av de mest betydningsfulle skillene mellom de ulike analysene er hvorvidt de inkluderer opplysninger kun fra Elevundersøkelsen, eller om informasjon også fra andre datakilder er med. Dette har konsekvenser både for tilgangen til kontrollvariabler og for om data kan brukes på individ- eller skolenivå.

I analyser basert kun på Elevundersøkelsen er sammenhengen mellom elevenes opplevelse av læringsmiljøet og deres prestasjoner undersøkt på individnivå. Det vil si at en elevs svar på spørsmålene i Elevundersøkelsen sees i forhold til denne elevens skoleprestasjoner. Imidlertid er det i analysene ikke kontrollert for elevens bakgrunnsforhold, som innvandrerbakgrunn og foreldrenes sosioøkonomiske status. Grunnen til det er at Elevundersøkelsen ikke kan kobles til slike data på individnivå. Disse analysene viser at mange av forholdene som studeres i Elevundersøkelsen er sterkt knyttet til skoleprestasjoner. Selv uten informasjon om bakgrunnsforhold forklarer opplysningene fra Elevundersøkelsen mellom 27 og 29 prosent av den totale variansen i elevenes prestasjoner. Faktoranalyse viser at de ulike indikatorene i undersøkelsen kan reduseres til fire dimensjoner: organisasjon og fysisk miljø, positivt læringsmiljø, læringsstrategier og evalueringsintensitet. Alle fire faktorene er knyttet til elevenes selvoppgitte karakterer på 10. trinn: en mer positiv opplevelse av læringsmiljøet, å ha uttalte læringsstrategier og mer formelle organisatoriske rammer er positivt for elevenes prestasjoner. Høy evalueringsintensitet er derimot negativt knyttet til elevenes skoleprestasjoner.

Også analyser som tar hensyn til forhold ved skolen, eleven og elevens bakgrunn, og hvor opplysningene fra Elevundersøkelsen brukes på skolenivå, viser at det er en sammenheng mellom elevenes prestasjoner og læringsmiljøet. At opplysningene brukes på skolenivå, betyr at alle elever ved en skole får den samme verdien på variablene i Elevundersøkelsen. Dermed er det ikke den direkte sammenhengen mellom hva den enkelte elev har svart på Elevundersøkelsen og deres skoleprestasjoner som undersøkes. Derimot er det sammenhengen mellom elevenes gjennomsnittlige svar på Elevundersøkelsen ved en skole og de enkelte elevenes prestasjoner på den samme skolen som undersøkes. Til tross for dette viser analysene i de ulike delrapportene at elevenes opplevelse av

skolen, slik det måles i Elevundersøkelsen, er av betydning for deres prestasjoner. Disse analysene peker på følgende forhold som knyttet til prestasjonsnivå:

Positivt skolemiljø

Analysene viser at et godt miljø på skolen er positivt knyttet til elevenes prestasjoner. Et godt skolemiljø er det forholdet som har størst betydning for skoleprestasjonene. Analysene i delrapport 5 (Grøgaard 2012) og i tilleggsnotatet (Næss 2011) viser at skolemiljøet kan ha spesielt stor betydning for elever med innvandrerbakgrunn.

Faglig støtte

I hvilken grad foreldre, lærere og medelever støtter elevene faglig er positivt relatert til prestasjonsutviklingen over tid.

Lærerstyrt undervisning

Lærerstyrt undervisning er i mange analyser funnet å være positivt knyttet til skoleprestasjoner. Den positive effekten av lærerstyrt undervisning er til stede både på 5., 8. og 10. trinn. Lærerstyrt undervisning kjennetegnes av tavleundervisning, individuelt arbeid og diskusjon/samtale mellom læreren og elevene i klassen.

Elevstyrt undervisning

Analysene viser gjennomgående at undervisning med stor grad av samarbeid mellom elevene, det som beskrives som elevstyrt undervisning, er negativt relatert til skoleprestasjoner. Dette gjelder å jobbe sammen to og to, gruppearbeid, prosjektarbeid og praktiske øvelser.

Evaluerings

For mye evaluering er knyttet til reduserte skoleprestasjoner. Skoler med utstrakt bruk av mappeevaluering i kombinasjon med evaluering gjennom praktiske oppgaver, har lavere prestasjoner enn skoler med mindre omfattende evaluering.

Læringsstrategier

Hvorvidt elevene har strategier for hvordan de skulle tilegne seg kunnskap, er av betydning for prestasjonene. Å ha flere strategier og å bruke dem i stor grad, henger sammen med bedre prestasjoner.

5.3 Styrker og svakheter ved Elevundersøkelsen som kvalitetsindikator

Alt i alt viser analysene i delrapportene at opplysninger fra Elevundersøkelsen bidrar til å forklare forskjeller i skoleprestasjoner. Hvordan elevene opplever skolen og undervisningen, er klart relatert til hvor godt de presterer. Dette er i tråd med teorier om betydningen av læringsmiljø (Skaalvik og Skaalvik, 2005) og med tidligere studier av Elevundersøkelsen (Bakken 2009a, 2010).

Selv om analyser bekrefter at opplysninger fra Elevundersøkelsen predikerer elevers skoleprestasjoner, er det likevel ikke gitt at undersøkelsens tilnærming er den som best fanger variasjon i læringsmiljøet og hvordan det er relatert til skoleprestasjoner. Gjennom de mange analysene i prosjektet «Ressurser og resultater i grunnopplæringen» har vi godt grunnlag for å vurdere hvor godt egnet Elevundersøkelsen er som indikator på kvalitet i skolen. På bakgrunn av den tilegnede kunnskapen skal vi diskutere sterke og svake sider ved Elevundersøkelsen. Hensikten er å peke på forhold ved undersøkelsens innhold og framgangsmåte som er godt egnet til å gi innblikk i læringsmiljø og hvordan det henger sammen med prestasjoner, og å peke på områder hvor det er rom for forbedring.

5.3.1 Innhold

I deres gjennomgang av betydningen av læringsmiljø skiller Skaalvik og Skaalvik (2005) mellom en begrenset og en utvidet forståelse av begrepet. En begrenset forståelse innebærer å se på læringsmiljø kun som de konkrete læringssituasjonene elevene deltar i, som for eksempel undervisning. En omfattende forståelse, derimot, innebærer å se på læringsmiljø som alle forhold som kan antas å ha betydning for elevenes læring. Dette synspunktet innebærer at læringsmiljø ikke er et klart definert begrep og at en rekke forhold ved skolen kan betraktes som indikatorer på læringsmiljø.

I Elevundersøkelsen er en omfattende tilnærming til læringsmiljø lagt til grunn. Undersøkelsen inneholder spørsmål om en rekke ulike forhold på skolen som kan være av betydning for elevenes læring, som fysiske forhold, relasjon til lærer, opplevd støtte hjemmefra og på skolen, trivsel, motivasjon, undervisning og evaluering. Bredden i temaene som undersøkes, gjør det mulig å undersøke betydningen av læringsmiljø i utvidet forstand. Samtidig gjør framgangsmåten det mulig å undersøke nærmere betydningen av spesifikke områder, som for eksempel undervisning og evaluering.

I utgangspunktet er det positivt at Elevundersøkelsen er bredt anlagt, i og med at det gir mulighet til allsidige undersøkelser av læringsmiljøet. Imidlertid har analysene i de ulike delrapportene vist at det er vesentlige forhold som ikke dekkes av undersøkelsen. Dette gjelder i særlig grad bakgrunnsforhold. Konsekvensene av disse manglene diskuteres under.

De siste årene har spørsmålene i Elevundersøkelsen i liten grad blitt endret. Det vil si at flere årskull med elever har svart på likelydende spørsmål om ulike forhold ved skolen. Det gjør det mulig å sammenlikne elevenes svar over tid. Imidlertid er det ikke mulig å undersøke utviklingen til enkeltindivider, siden opplysningene i Elevundersøkelsen ikke registreres på en slik måte at de kan kobles på individnivå. Det at undersøkelsen ikke er på individnivå, har også konsekvenser for muligheten til å koble opplysninger i Elevundersøkelsen til andre datasett. Dette diskuteres nærmere.

Begrenset informasjon om kjennetegn ved elevene: Input-faktorer

Forskning har gjentatte ganger påvist betydningen av elevenes bakgrunn for deres prestasjoner. Spesielt kjønn, sosioøkonomisk bakgrunn og innvandrerstatus er funnet å være viktig for skoleprestasjoner: jenter presterer i gjennomsnitt bedre enn gutter, elever fra høyere sosiale lag presterer i gjennomsnitt bedre enn elever fra lavere lag, og elever uten innvandrerbakgrunn presterer i gjennomsnitt bedre enn elever med innvandrerbakgrunn. Delrapportene i prosjektet bekrefter betydningen av disse forholdene for skoleprestasjoner.

Imidlertid er det få spørsmål om bakgrunnsforhold i Elevundersøkelsen. Riktignok blir elevene spurt om kjønn, men de blir ikke spurt om forhold som kan indikere foreldrenes sosioøkonomiske status, som for eksempel deres yrke, utdanning eller om de er arbeidsledige. Det spørres heller ikke om forhold er egnet til å indikere innvandrerbakgrunn, som for eksempel hvilket land elevene er født i eller hvilket språk som snakkes hjemme. Det betyr at undersøkelsen ikke har opplysninger om forhold som er av stor betydning for hvordan elever presterer på skolen, noe som har konsekvenser for hvilke slutninger det er mulig å trekke av analyser av Elevundersøkelsen.

For det første er det ikke mulig å undersøke hvor mye av variansen i svarene som skyldes bakgrunnsforhold og hvor mye som skyldes andre forhold. Dermed kan ikke analyser av Elevundersøkelsen gi svar på om hvor mye av variasjonen i elevers opplevelse av læringsmiljøet som skyldes ulik bakgrunn og hvor mye som handler om hva som skjer på skolen. Det vil si at det ikke er mulig å undersøke om det er slik at elever med høyere sosial bakgrunn trives bedre på skolen enn elever med lavere sosial bakgrunn, eller om elever med innvandrerbakgrunn trives bedre eller dårligere enn andre elever. Tidligere studier har vist at elevenes opplevelse av læringsmiljøet har sammenheng med kjennetegn ved elevene (Frøyland og Gjerustad 2012). For å forstå betydningen av læringsmiljøet for elevenes prestasjoner er det nødvendig å vite hvordan læringsmiljøet henger sammen med bakgrunnsforhold, noe som ikke er mulig med Elevundersøkelsen slik den er i dag.

For det andre betyr mangelen på bakgrunnsfaktorer at det ikke er mulig å kontrollere for effekten av viktige forhold når vi undersøker sammenhengen mellom svar avgitt i Elevundersøkelsen og skoleprestasjoner. Det fører mest sannsynlig til at betydningen av læringsmiljø for skoleprestasjoner overdrives. En del av variansen i både elevenes opplevelser av læringsmiljøet og deres prestasjoner kan trolig tilbakeføres til forhold som sosioøkonomisk bakgrunn og innvandrerbakgrunn. Når slike forhold ikke inkluderes blir det vanskelig å isolere effekten av læringsmiljøet: vi vet ikke nøyaktig hva som skyldes det elevene har med seg inn i skolen, og hva deres opplevelse av skolen bidrar med. Dette begrenser mulighetene til å forstå sammenhengen mellom læringsmiljø og skoleprestasjoner på bakgrunn av Elevundersøkelsen.

Begrenset informasjon om elevenes læringsutbytte: Output-faktorer

Elevundersøkelsen inneholder spørsmål til elevene om hvilke karakterer de hadde forrige termin i seks – sju sentrale fag. Dette gjelder på alle trinn der elevenes prestasjoner blir vurdert i form av karakterer, altså fra og med 8. trinn og ut videregående opplæring. På disse trinnene inkluderer dermed Elevundersøkelsen både opplysninger om elevenes opplevelse av læringsmiljøet på skolen (prosess) og deres karakterer (output), noe som gjør det mulig å se læringsmiljø og prestasjoner i sammenheng. Elevenes selvrapporterte karakterer brukes i analysen av elever på 10. trinn i delrapport 1 (Opheim mfl 2010). Resultatene indikerer høyt samsvar mellom elevenes selvrapporterte karakterer i Elevundersøkelsen og skolens karaktersnitt.

Blant elever på lavere trinn (5-7.trinn), inneholder Elevundersøkelsen ingen spørsmål om elevenes prestasjonsnivå. Dette er rimelig, ettersom elevene ikke får karakterer på disse trinnene. Samtidig kunne det vært mulig å inkludere spørsmål som på andre måter fanget opp elevenes prestasjonsnivå på disse trinnene. Når dette ikke gjøres, gir Elevundersøkelsen lite informasjon om elevenes læringsutbytte på disse trinnene.

Mulighet for å koble informasjon fra Elevundersøkelsen til andre datakilder

Som vist i forrige kapittel, omfatter prosjektet analyser av et omfattende datamateriale bestående av ulike datakilder koblet sammen. Elevundersøkelsen har blitt koblet sammen med registre som inneholder ulik type informasjon om skolene elevene går på (GSI, SSBs befolkningsregistre, data fra nasjonale prøver).

Imidlertid er det ikke mulig å koble til opplysninger fra Elevundersøkelsen på individnivå. Elevundersøkelsen inneholder ingen informasjon om eleven, verken personnummer, navn eller annet, som ville gjøre det mulig å identifisere den enkelte elev. Dette er viktig av personvern hensyn. Dersom undersøkelsen skulle inneholdt slik informasjon om elever under myndighetsalder, ville det vært nødvendig å innhente samtykke fra elevenes foreldre før undersøkelsen kunne gjennomføres. Det sier seg selv at dette hadde gjort Elevundersøkelsen langt mer komplisert å gjennomføre. Trolig ville også deltakelsen vært betydelig lavere med en slik fremgangsmåte. En lavere deltakelse ville samtidig øke risikoen for at enkelte grupper av elever ville vært underrepresentert, noe som ville gitt lav representativitet. I Elevundersøkelsen er det med andre ord lagt større vekt på gjennomførbarhet og høy deltakelse både på individnivå og på skolenivå, fremfor muligheter til å koble informasjon fra den enkelte elev til andre datakilder. Dette er en forståelig prioritering. Samtidig innebærer dette en begrensning ved dette datamaterialet.

Uten individkjennetegn er det umulig å hente inn informasjon fra andre kilder om for eksempel kjennetegn ved eleven (sosial bakgrunn, bosted, fødeland) eller om elevens læringsutbytte. Som diskutert tidligere i dette kapitlet, er dette informasjon som per i dag i liten grad innhentes i Elevundersøkelsen.

I stedet for å analysere sammenhenger mellom elevenes læringsmiljø og deres skoleprestasjoner på individnivå, kan dette analyseres på skolenivå. Dette gjøres ved å sammenligne elevenes gjennomsnittlige opplevelser av ulike sider ved sitt læringsmiljø med elevenes gjennomsnittlige prestasjoner på de nasjonale prøvene, noe som har blitt gjort i flere av delrapportene i prosjektet «Ressurser og resultater». Dette gir rimeligvis en svakere forståelse av sammenhengen mellom

elevenes læringsmiljø og prestasjoner. Senere i dette kapitlet vil vi se nærmere på hvor mye av variasjonen i elevenes oppfatning av læringsmiljøet som fanges opp ved å sammenligne gjennomsnitt på klassenivå og skolenivå – og hvor mye av variasjonen mellom elever som finnes innen den enkelte skole og enkelte klasse.

5.3.2 Spørsmålsformuleringer og svarkategorier

I spørreundersøkelser er det av stor betydning at spørsmålene er formulert slik at det er tydelig hva de måler. I tillegg er det også viktig at de tilgjengelige svarkategoriene er egnet til å fange variasjonen i elevenes svar. Gjennom arbeidet med prosjektet «Ressurser og resultater i grunnopplæringen» har det kommet fram at enkelte av spørsmålene i Elevundersøkelsen ikke er presise nok, og at svarkategoriene i enkelte tilfeller ikke fanger opp variasjonen i svarene. Dette påvirker muligheten til å bruke opplysninger fra undersøkelsen som indikator på kvalitet i skolen.

Manglende konkretisering og nyansering av spørsmål

I Elevundersøkelsen svarer elevene blant annet på spørsmål om undervisningen, hvordan de evalueres og hvilken relasjon de har til læreren. Disse spørsmålene er generelt formulert i den forstand at det ikke er spesifisert hvilke fag og hvilke lærere elevene skal ha i tankene når de svarer. Det medfører at den informasjonen som kommer fram er elevenes generelle oppfatninger av undervisning, evaluering og elev-lærerrelasjon. Imidlertid er det rimelig å anta at det er til dels store variasjoner i disse forholdene for den enkelte elev. Selv om en elev generelt opplever relasjonen til lærerne som midt på treet, kan det være at eleven har en spesielt god relasjon til en eller noen få lærere. Det samme gjelder for undervisning og evaluering. Trolig varierer elevenes opplevelse av disse forholdene fra fag til fag. Denne variasjonen fanges ikke opp av Elevundersøkelsen.

Det at spørsmålene er såpass generelt formulert kan gjøre det vanskelig å bruke dem til å forklare prestasjoner i konkrete fag. Dersom vi ønsker å forstå sammenhengen mellom elevenes opplevelse av undervisning og evaluering på den ene siden, og deres prestasjoner på den andre, er det nødvendig å vite mer om hvordan elevenes opplevelser varierer fra fag til fag. Dette diskuteres nærmere i neste kapittel.

Spørsmålene om undervisning i Elevundersøkelsen undersøker hvor ofte elevene har ulike former for undervisning. Elevene kan svare at de har alle undervisningsformene flere ganger i uken. De må ikke prioritere mellom ulike typer undervisningsformer eller anslå hvor mye de har av de ulike undervisningsformene. Det medfører at det i mange tilfeller er vanskelig å vite sikkert hvilke typer undervisning de har mest av, og hvilke de har minst av. Det kompliserer arbeidet med å analysere sammenhengen mellom undervisningsform og prestasjoner. For å forstå hvordan undervisningsformer er relatert til prestasjoner kreves mer konkret informasjon både om hvilke undervisningsformer som brukes i hvert fag, samt hvilke undervisningsformer som brukes mer enn andre. Dette kan oppnås ved bruk av mer nyanserte spørsmål.

Også spørsmålene om mobbing hadde tjent på ytterligere spesifisering. Forekomsten av mobbing dekkes i dag gjennom ett spørsmål. Før spørsmålet får elevene presentert en definisjon av begrepet mobbing, der mobbing som *gjentatt* negativ atferd kommer tydelig frem. Det kan være vanskelig å sette et klart skille mellom erting og mobbing, og mellom enkelttilfeller og mer systematisk eller repetert negativ atferd. Et slikt skille kan blant annet handle om hvem som utfører handlingene, i hvilke situasjoner, og hvordan/hvilke handlinger det er snakk om. Elevundersøkelsen inneholder spørsmål om hvem/hvilke grupper eleven opplever å bli mobbet av, men ellers ikke spørsmål som gir utdypende informasjon om hvilke former for mobbing eleven utsettes for. En annen innvending mot mobbspørsmålet, slik det er konstruert, har vært at det ikke skiller tilstrekkelig mellom mobbing og mindre alvorlige forhold som vanligvis ikke regnes som mobbing (Lødding og Vibe 2010). Det at Elevundersøkelsen ikke inneholder spørsmål om hvordan mobbingen foregår og/eller hvilke former for mobbing elevene har vært utsatt for, begrenser mulighetene til å bruke Elevundersøkelsen i skolens arbeid mot mobbing.

Grove svarkategorier i spørsmål om undervisning

Når elevene oppgir hvor ofte de har de ulike undervisningsformene kan de velge mellom følgende fem svarkategorier: «flere ganger i uken», «1 gang i uken», «1-3 ganger i måneden», «2-4 ganger i halvåret» og «sjeldnere». En nærmere gjennomgang av elevenes svarfordelinger viser at omtrent 90 prosent av elevene krysser av i kategorien «flere ganger i uken» på spørsmål om hvor ofte de har tavleundervisning, og omtrent 80 prosent svarer det samme på spørsmål om hvor ofte de jobber individuelt med oppgaver. Den betydelige andelen elever som krysser av i den høyeste svarkategorien tyder på at undervisningsformene som spørsmålene omhandler brukes svært mye i skolen, og at de anvendte svarkategoriene ikke i tilstrekkelig grad tar høyde for dette. Dermed oppstår det som kalles en takeffekt: svarkategoriene er ikke egnet til å fange variasjon i den øvre enden av skalaen. Takeffekten er særlig til stede på de to nevnte spørsmålene, men vi finner tendenser til det samme også for spørsmålene om omfang av andre undervisningsformer, som hvor ofte elevene jobber sammen to og to og hvor ofte undervisningen består av samtale/diskusjon mellom lærer og elever. Svarene er mer normalfordelt for tre gjenværende undervisningsmetodene – prosjektarbeid, gruppearbeid og praktiske øvelser - som studeres i Elevundersøkelsen.

Et viktig funn i prosjektet «Ressurser og resultater i grunnopplæringen» er at lærerstyrt undervisning er positivt relatert til prestasjoner. I samlemålet lærerstyrt undervisning inngår tavleundervisning og individuelt arbeid sammen med samtale/diskusjon mellom læreren og hele klassen. Det at vi ikke har informasjon om variasjonen i den store gruppen elever som har krysset av for at de har tavleundervisning og individuelt arbeid flere ganger i uken er problematisk. Det er overveiende sannsynlig at det er variasjon innad i gruppen, men dette kan vi ikke si med sikkerhet. Videre vet vi ikke hvordan variasjonen innad i denne store gruppen er relatert til skoleprestasjoner.

I utgangspunktet er det tre ulike muligheter for hvordan relasjonen kan være. For det første kan den være positiv, det vil si at mer lærerstyrt undervisning er knyttet til bedre karakterer. Dersom det er tilfelle vil en mer nyansert måling føre til en ytterligere styrking av sammenhengen mellom lærerstyrt undervisning og skoleprestasjoner. Den andre muligheten er at det ikke er noen sammenheng mellom hvor ofte elevene har lærerstyrt undervisning og skoleprestasjoner i den gruppen som har krysset av i kategorien «flere ganger i uken». Et slikt scenario innebærer at relasjonen mellom lærerstyrt undervisning og karakterer som er funnet i prosjektet svekkes. Den tredje muligheten er at det er en negativ sammenheng mellom lærerstyrt undervisning og prestasjoner innad i den store gruppen. I tilfelle betyr det at sammenheng mellom lærerstyrt undervisning og prestasjoner er kurvlineær: en viss mengde slik undervisning er positivt for prestasjonene, mens lite og mye er negativt.

Mangelen på informasjon om variasjon i den øverste delen av skalaen gjør at vi ikke kan slå fast hvilket av disse scenarioene som er riktig. Imidlertid viser analysene at mer lærerstyrt undervisning er positiv, selv om det er svakheter ved målingen. Dette indikerer at slik undervisning er positivt knyttet til prestasjonene, og det er ikke usannsynlig at sammenhengen er til stede også i gruppen som har krysset av i den øverste kategorien.

Mer nyansert informasjon om hvor ofte disse undervisningsformene brukes vil være positivt for forståelsen av sammenhengen mellom lærerstyrt undervisning og skoleprestasjoner. Dersom det også var inkludert kategorier med svaralternativer som indikerte høyere hyppighet, som for eksempel «hver dag» og «flere ganger om dagen», ville variablene fått tak i mer av variasjonen i hvor ofte undervisningsformene ble brukt. I tillegg kan mer fagspesifikke spørsmål, som foreslått under forrige punkt, være positivt. Dersom spørsmålene om undervisning avgrenses til å gjelde et fag vil behovet for å ha svarkategorier som måler høyere hyppighet trolig reduseres.

5.3.3 Elevenes motivasjon for å svare

I Elevundersøkelsen blir elevene bedt om å oppgi sine vurderinger av det som skjer på skolen og å fortelle om hvordan dette oppleves for dem. Framgangsmåten avhenger av at elevene er motivert for å svare ordentlig på spørsmålene. Dersom elevene ikke er motivert for å lese spørsmålene skikkelig og

å svare etter beste evne så reduseres kvaliteten på svarene. Det begrenser muligheten til å bruke Elevundersøkelsen for å forstå elevenes skolehverdag og hvilke forhold som bedrer prestasjonene.

Det er flere forhold ved Elevundersøkelsen som kan påvirke kvaliteten på elevenes svar. For det første er framgangsmåten sårbar for useriøse svar. Studier har vist at tulle svar er relativt vanlig i spørreundersøkelser med ungdom som respondenter (Elstad 2010, Lødding og Vibe 2010), noe som bekreftes i delrapport 4 i prosjektet «Ressurser og resultater» (Opheim og Wiborg 2012).

I Elevundersøkelsen skal alle svarene gis i forhåndsdefinerte kategorier og det er få spørsmål som tar opp sensitive tema. I utgangspunktet er dette positivt med tanke på å motvirke useriøse svar. Imidlertid trenger ikke det å bety at elevene ikke tuller. Derimot kan det innebære at slike svar kan være vanskelige å oppdage, fordi tøysing ikke blir like åpenbart. I en gjennomgang av mobbespørsmålet i Elevundersøkelsen fant Lødding og Vibe (2010) at en del respondenter tilsynelatende svarte useriøst på dette spørsmålet. Særlig gutter på 10. trinn så ut til å overrapportere hvor mye de ble mobbet. Lødding og Vibe konkluderte med at en del av respondentene tullet på dette spørsmålet og utelot dem fra analysene. Selv om det var en liten del av utvalget som ble identifisert som useriøse, omtrent en prosent, er det mulig at flere tullet på dette spørsmålet uten å bli oppdaget.

For å undersøke sannsynligheten for useriøse svar i dette prosjektet fant vi ut hvor mange som hadde oppgitt å ha spesielt dårlige karakterer. På 10. trinn var det i alt 1660 elever som oppga å ha 1 i alle fagene som undersøkes. Dette utgjør under en prosent av det samlede antall elever som hadde svart på spørsmålet og tallet kan være riktig. Imidlertid er det betydelig flere som sier at de har 1 i alle fag enn det er elever som sier at de har en karaktersum som tilsvarer 2 i alle fag. Denne overrepresentasjonen tilsier at dette kan være et spørsmål enkelte elever svarer useriøst på.

Når vi ser på sammenhengen mellom svar på karakterer og det å mobbe andre er det hele 65 prosent av de som sier at de har 1 i alle fag, som også oppgir å mobbe andre elever flere ganger i uken. Til sammenlikning svarer 3 prosent av alle elevene som deltok det samme. Tallene indikerer at det her er elever som svarer useriøst. Selv om det er relativt få som eventuelt tuller på disse spørsmålene, kan det være useriøse svar også på andre spørsmål. Samtidig er det vanskelig å konkludere med sikkerhet hvorvidt elever faktisk svarer useriøst, i og med at det er få spørsmål som er godt egnet til å avdekke dette. At elevene tuller med egne karakterer, er trolig ikke lett å oppdage ved hjelp av bare ett spørsmål. Det er svarene på flere spørsmål som sammen indikerer tulling, noe som også ble diskutert i delrapport 4 (Opheim og Wiborg 2012). Analyser av Elevundersøkelsen i delrapport 1, tyder imidlertid på at majoriteten av elevene svarer oppriktig på spørsmålene om egne karakterer (Opheim mfl 2010).

Det andre forholdet ved Elevundersøkelsen som kan påvirke kvaliteten på elevenes svar, er antallet spørsmål og spørsmålsformuleringer. I en undersøkelse kom det fram at mange skoleledere og -eiere ikke var fornøyde med spørsmålene i Elevundersøkelsen (Vibe og Hovdhaugen 2012). Rundt 10 prosent oppga at de var svært tilfreds med spørsmålene, mens et klart flertall svarte at de var litt tilfreds eller litt utilfreds med spørsmålene. Mange av skolelederne og -eierne benyttet et åpent kommentarfelt til å gi uttrykk for sitt syn på Elevundersøkelsen. De vanligste kommentarene handlet om at undersøkelsen var for lang, at spørsmålene var for kompliserte, og at dette reduserte elevenes motivasjon til å gi gjennomtenkte og seriøse svar. Lengden på Elevundersøkelsen er for øvrig noe skolene selv kan påvirke, gjennom valg av tilleggsspørsmål. En del skoleledere og -eiere oppfatter imidlertid Elevundersøkelsen for lang, selv uten tilleggsspørsmål.

Det tredje og siste forholdet av betydning for kvaliteten på svarene i Elevundersøkelsen er at mange elever vil delta i Elevundersøkelsen en rekke ganger. I tillegg til de obligatoriske undersøkelsene på 7. trinn, 10. trinn og på Vg1 er det mange elever som deltar også på andre trinn. Det vil si at en god del elever svarer på Elevundersøkelsen gjentatte ganger i løpet av grunntidningen. Det kan bidra til å svekke opplevelsen av at undersøkelsen er viktig, og dermed øke sjansen for useriøse svar.

5.4 Elevundersøkelsen: Informasjon på ulike nivåer

Elevundersøkelsen er konstruert for å undersøke elevenes subjektive oppfatning av skolen, som hvordan de trives og hvordan de opplever relasjonen til lærerne. Det taler for at Elevundersøkelsen først og fremst måler betydningen av elevenes individuelle opplevelse. Likevel kan det være slik at også forskjeller mellom skoler og klasser kan påvirke hvordan elevene svarer. Vi skal i det følgende undersøke dette nærmere, gjennom analyser av hvor mye av variasjonen som kan tilbakeføres til henholdsvis skole-, klasse- og individnivå. I prosjektet «Ressurser og resultater i grunnopplæringen» har ikke betydningen av ulike nivåer for svarene i Elevundersøkelsen vært undersøkt.

Videre skal vi undersøke hvor godt informasjonen i Elevundersøkelsen er til å forklare varians i skoleprestasjoner når de brukes på henholdsvis individ-, klasse- og skolenivå. Også disse analysene belyser betydningen av de ulike nivåene.

Analysene er basert på Elevundersøkelsen på 10. trinn og Vg1 for årene 2007 til 2010. I Elevundersøkelsen i 10. klasse og på videregående skole oppgir elevene karakterene sine i syv fag. Dermed er det mulig å se på sammenhengen mellom opplysninger gitt i Elevundersøkelsen og karakterer på disse trinnene. På disse trinnene er Elevundersøkelsen obligatorisk, slik at antallet elever som har deltatt er høyt. I alt inngikk 414 761 elever i datamaterialet. En del av disse ble ikke inkludert i analysene på grunn av manglende opplysninger på ett eller flere spørsmål.

Noen skoler opererer ikke med klasser i tradisjonell forstand, og innledende analyser viste at enkelte «klasser» var svært store, med over hundre elever. For å sørge for at det virkelig var klasseinndeling som ble målt ble derfor alle klasser med mer enn 30 elever tatt ut av analysene. Dette medførte at 9 290 elever ble tatt ut av analysene for ungdomstrinnet, mens 1 932 elever ble tatt ut av analysene for videregående skole. Alle analyser ble gjennomført separat for 10. trinn, Vg1 studieforberedende og Vg1 yrkesfag.

Tre av faktorene som ble funnet i delrapport 1 (Opheim mfl 2010) ble valgt ut som avhengige variabler: trivsel, lærerstyrt undervisning og læringsmiljø. Faktorene ble valgt fordi de måler ulike sentrale forhold ved skolen.

5.4.1 *Opplevelse av læringsmiljøet – betydningen av individ-, klasse- og skolenivå*

For å finne ut av hvor mye av variansen i elevenes opplevelse av læringsmiljøet som kan tilbakeføres til skole-, klasse- og individnivå, ble det gjennomført en rekke flernivåanalyser. De tre faktorene - trivsel, undervisningsform og læringsmiljø - inngikk som avhengige variabler i såkalte tomme modeller. Det vil si at modellene ikke har uavhengige variabler, utover at betydningen av skole- og klassenivå blir kontrollert for. Resultatene av analysene presenteres i tabell 5.1.

Tabell 5.1: Andel varians i tre indikatorer på læringsmiljø som kan tilbakeføres til skole-, klasse og individnivå på 10. trinn og Vg1.

		Trivsel	Lærerstyrt undervisning	Læringsmiljø
		N = 182 383	N = 141 091	N = 184 489
10. trinn	Skolenivå	2,0 %	1,4 %	4,7 %
	Klassenivå	6,1 %	3,9 %	11,3 %
	Individnivå	91,9 %	94,7 %	84,0 %
	Totalt	100,0 %	100,0 %	100,0 %
		N = 81 025	N = 69 661	N = 81 926
Vg1:				
Studieforberedende	Skolenivå	2,4 %	2,5 %	7,0 %
	Klassenivå	5,3 %	4,1 %	12,1 %
	Individnivå	92,3 %	93,4 %	80,9 %
	Totalt	100,0 %	100,0 %	100,0 %
		N = 120 566	N = 102 068	N = 121 964
Vg1: Yrkesfag	Skolenivå	1,7 %	1,4 %	2,4 %
	Klassenivå	8,8 %	6,5 %	19,5 %
	Individnivå	89,5 %	92,1 %	78,1 %
	Totalt	100,0 %	100,0 %	100,0 %

Tallene i tabell 5.1 viser med tydelighet at det individuelle forhold har stor betydning for elevenes opplevelse av skolen. Så mye som 78 til 93 prosent av variansen i elevenes oppfatninger av trivsel, undervisning og læringsmiljø handler om forskjeller mellom elevene. Andelen varians som kan tilbakeføres til skole- og klassenivået er relativt liten for alle tre faktorene. Videre viser tabellen at en høyere andel av variansen kan tilbakeføres til klassenivået enn til skolenivået. Det er til dels store forskjeller. Mellom 2 – 7 prosent av variansen skyldes forskjeller mellom skoler mens 4 – 20 prosent av variansen skyldes forskjeller mellom klasser. Det tyder på at forskjeller mellom klasser er av større betydning for elevenes opplevelse av skolen enn forskjeller på skolenivå.

Dette gjelder også for faktorene lærerstyrt undervisning og læringsmiljø, selv om utsagnene som inngår i faktorene er ment å måle det som skjer på skolen. Hva eleven har med seg inn i skolen, med andre ord ulike individkjenntegn, fremstår som viktigere for elevens opplevelse av skolen enn det som faktisk foregår på skolen og i klassen. Imidlertid forklares faktoren læringsmiljø i større grad av skole- og klassenivået. Omtrent 15 – til 20 prosent av variansen kan tilbakeføres til andre nivåer enn det individuelle når det gjelder læringsmiljø, noe som er klart høyere enn hva tilfellet er for de andre to faktorene (undervisningsform og trivsel). Dette peker på at elevenes opplevelse av læringsmiljøet på skolen i større grad kommer av forskjeller mellom skoler og klasser enn hva deres trivsel og opplevelse av undervisningen gjør.

5.4.2 Forklaringskraft til Elevundersøkelsen

Delrapport 1 viser at Elevundersøkelsen er godt egnet til å forklare forskjeller i elevers skoleprestasjoner (Opheim mfl 2010). Dette gjelder særlig når opplysningene brukes på individnivå, altså at den enkelte elevs svar i Elevundersøkelsen forklarer deres prestasjoner. Imidlertid har det vært nødvendig å aggregere data fra Elevundersøkelsen til skolenivå i mange av analysene i prosjektet «Ressurser og resultater i grunnopplæringen». Når dataene brukes på aggregert nivå vil det si at det er elevenes gjennomsnittlige svar på Elevundersøkelsen ved en skole som brukes til å forklare prestasjonene til de enkelte elevene på skolen. Forklaringskraften til aggregerte data er ofte svakere enn data som ikke er aggregert.

For å undersøke betydningen av å aggregere opplysninger fra Elevundersøkelsen nærmere skal vi gjennomføre analyser der data brukes på ulike nivåer. Hensikten er å finne ut hvordan det å bruke dataene på individ-, klasse- eller skolenivå påvirker deres evne til å forklare varians i elevenes prestasjoner. Regresjonsanalysene likner analyser som ble foretatt i delrapport 1 på data fra Elevundersøkelsen (Opheim mfl 2010). I analysene inngår i alt 16 indikatorer eller faktorer fra Elevundersøkelsen som uavhengige variabler: fysisk miljø, trivsel, arbeidsmiljø, fravær av mobbing, bruk av PC, evaluering med prøver, evaluering med presentasjon, evaluering med mappe, lærerstyrt undervisning, elevaktiv undervisning, faglig støtte, lærer-elev-relasjon, organisert leksehjelp, faglig tilbakemelding, medbestemmelse og motivasjon. Faktorene er tidligere benyttet i delrapport 1 for å predikere prestasjoner på skole, og er presentert mer detaljert der. Tre faktorer ble utelatt på grunn av høy andel med manglende svar. Analysene i seg selv presenteres ikke, kun hvor mye av variansen i skoleprestasjoner faktorene kunne forklare.

Resultatene av analysene presenteres i tabell 5.2. Kolonnene viser hvor mye av variasjonen i elevenes prestasjoner som forklares av Elevundersøkelsen når dataene brukes på individnivå, eller når de aggregeres til klasse- eller skolenivå. Analysene som ligger bak hver av kolonnene er separate.

Tabell 5.2: Forklart varians i prestasjonsnivå når variabler fra Elevundersøkelsen brukes på individnivå eller aggregeres til klasse- og skolenivå. 2007 – 2010.

	Individnivå	Klassenivå	Skolenivå
10. trinn (N = 117 413)	24,8 %	1,9 %	0,7 %
Vg1 studieforberedende (N = 50 233)	14,9 %	5,1 %	4,6 %
Vg1 yrkesfag (N = 60 760)	11,8 %	5,3 %	1,5 %

Tabellen illustrerer hvordan Elevundersøkelsens egnethet til å forklare varians i elevenes skoleprestasjoner svekkes når opplysningene brukes på aggregert nivå. Dataenes forklaringskraft reduseres betydelig når de aggregeres fra individnivå til klassenivå, og ytterligere når de aggregeres til skolenivå. Dette gjelder særlig for 10. trinn. Mens opplysninger om elevens opplevelse av skolen på individnivå forklarer omtrent 25 prosent av variansen i skoleprestasjoner på 10. trinn, reduseres denne andelen til omtrent 2 prosent når dataene blir aggregert til klassenivå og omtrent 1 prosent når dataene blir aggregert til skolenivå. Tabellen viser at å aggregere dataene reduserer forklaringskraften også på Vg1, selv om reduksjonen ikke er like sterk som for 10. trinn.

Tabellen viser videre at forklaringskraften til Elevundersøkelsen er lavere på individnivå for elever på Vg1 enn blant elever på 10. trinn. Samtidig er forklaringskraften på klassenivå og skolenivå høyere blant elever i Vg1 enn blant elever på 10. trinn.

Mange av analysene i prosjektet «Ressurser og resultater i grunnopplæringen» finner at variabler fra Elevundersøkelsen er relatert til elevenes prestasjoner, også når de er aggregert til skolenivå og når

en rekke faktorer ved eleven og skolen er kontrollert for. Det at sammenhengen er til stede på aggregert nivå, hvor forklaringskraften i utgangspunktet er svak, peker på at forholdene målt i Elevundersøkelsen er viktige for elevenes prestasjoner.

5.5 Oppsummering

I kapittel 1 ble det presentert en enkel analysemodell av elevers prestasjoner. Modellen viser hvordan elevenes skoleprestasjoner (output) avhenger både av det som skjer i skolen (prosess), og av hva de har med seg inn i skolen av individuelle kjennetegn og egenskaper (input). I utgangspunktet kan en spørreundersøkelse som Elevundersøkelsen bidra til å belyse alle delene i modellen. Imidlertid viser kapitlet at ikke alle områdene blir dekket like godt. Input-delen er svakt dekket. Til tross for at bakgrunnsforhold har stor betydning for elevers prestasjoner er det ikke spørsmål om foreldrenes yrkesstatus eller om innvandrerbakgrunn. Sammenliknet med input-faktorer er prosesser i skolen vesentlig bedre dekket. Elevundersøkelsen inneholder en rekke spørsmål om hva som foregår i skolen og hvordan dette oppleves av elevene. Men også her peker gjennomgangen på at det er rom for forbedring: i mange tilfeller er ikke spørsmålene tilstrekkelig konkrete og nyanserte. Dette gjør det vanskelig å bruke resultatene i arbeidet med å utvikle skolens læringsmiljø og forbedre elevenes læringsutbytte. Når det gjelder output-faktorer er dette relativt godt dekket, med spørsmål til elevene om deres karakterer i sentrale fag på alle trinn der elevenes prestasjoner blir vurdert i form av karakterer. Blant elever på lavere trinn, inneholder Elevundersøkelsen derimot ingen spørsmål som forsøker å fange opp elevenes prestasjoner eller mestringsnivå.

Selv om Elevundersøkelsen utvilsomt er en viktig kilde til informasjon om elevenes egen opplevelse av sitt læringsmiljø på skolen, viser gjennomgangen i dette kapitlet at undersøkelsen har sine begrensninger.

I kapittel 7 kommer vi tilbake til anbefalinger for en eventuell videre utvikling av Elevundersøkelsen med tanke på å gjøre den bedre egnet til å gi entydig og valid informasjon om viktige sider ved elevenes læringsmiljø. Anbefalingene er ment å styrke Elevundersøkelsen som kilde til kvalitetsindikatorer for grunnopplæringen.

6 Spørreundersøkelsen til elever og lærere i to fag

6.1 Ett skritt nærmere kontekstnær informasjon

Dette kapitlet har som formål å se nærmere på hva spørreundersøkelser kan bidra med i undersøkelser av sammenhenger mellom ressurser og elevresultater. Som nevnt tidligere i denne rapporten, har en sentral del av prosjektet vært å analysere data fra ulike registre for å belyse sammenhenger mellom ressursbruk i skolen og elevenes utbytte av opplæringen. Grunnskolenes informasjonssystem (GSI) bidrar med data om ulike ressurser (se kapittel 4). Data om elevenes standpunkt- og eksamenskarakterer benyttes for å vurdere elevers måloppnåelse med tanke på kompetansemålene, mens nasjonale prøver gir innblikk i elevenes ferdigheter i lesing, regning og engelsk, og data fra Elevundersøkelsene er brukt som indikatorer på elevenes læringsmiljø, først og fremst på skolenivå (se kapittel 5).

Disse datakildene gir mye informasjon om skolenes ressursituasjon og om elevenes læringsmiljø på skolenivå. En svakhet ved disse dataene er imidlertid mangelen på informasjon om hva som skjer i det enkelte klasserom og relasjonen mellom elever og lærer. Hvilke kjennetegn ved elev-lærer-relasjonen og læringsmiljøet i det enkelte klasserommet har betydning for elevenes læring? For å komme nærmere analyser av denne type problemstillinger, ble det våren 2011 gjennomført en spørreundersøkelse blant elever og lærere på 9. trinn og i videregående trinn 1 (Vg1); se nærmere omtale av delrapport 4 (Opheim og Wiborg 2012) i kapittel 2 i denne delrapporten.

I dette kapitlet diskuteres resultatene fra denne spørreundersøkelsen nærmere, samt at vi diskuterer hva denne type spørreundersøkelser kan bidra med mer generelt i undersøkelser av sammenhenger mellom ressurser og elevresultater. Diskusjonen underbygges med eksempler på forskjeller mellom elever og læreres svar på like spørsmål, sammenligning av elevenes svar på spørsmål knyttet til norsktimene og matematikktimene, samt sammenligning av svarfordelinger i spørreundersøkelsen og Elevundersøkelsen.

6.2 Data og design: Styrker og svakheter

Flere perspektiver og konkrete fag

Hovedtanken med å bruke et spørreskjema for å undersøke elevers erfaringer med skolen, er at man får tilgang til forhold ved skolen slik de oppleves av elevene selv. Det vil alltid være en styrke å innhente så konkret datamateriale som mulig, noe som i vårt tilfelle innebærer å spørre brukerne – elever – direkte om deres erfaringer. Dette er opplagt et viktig perspektiv når forhold i klasserommet skal belyses.

Videre vil det være en styrke å se hvordan andre aktører opplever de samme forholdene. Det kan gi verdifull innsikt å supplere elevens svar med responser fra deres lærere. På denne måten kan man se på klassen og aktivitetene med to ulike perspektiver. Videre kan samsvaret mellom responser fra disse to gruppene av aktører i seg selv si mye om elev-lærer-relasjonen, hva som er intensjonen med det som skjer i klasserommet og hva som faktisk skjer. Slike vurderinger var grunnlaget for valg av design av spørreundersøkelsen gjennomført i dette prosjektet. Ved å gi elever og lærere spørsmål om de samme forholdene, samt at elevenes besvarelser har fått unike koder for hver klasse, har man kunnet se på klasserommene hver for seg. Elevenes besvarelser har blitt koblet med besvarelsen til deres faglærer i norsk og/eller matematikk. Dette er en styrke ved denne spørreundersøkelsen sammenlignet med andre datakilder og har åpnet muligheter for flere verdifulle analyser av den direkte sammenhengen mellom elev og lærer, samt likheter og forskjeller mellom to skolefag. En slik analyse er vist i eksempel 1 nedenfor. Samtidig er dette et utfordrende og eksplorerende metodisk design som også er sårbart, noe som utdypes nedenfor.

Koblingen av responser fra ulike aktører er ikke uproblematisk. I spørreundersøkelsen viste det seg at antallet klasser hvor man hadde responser fra både læreren og fra lærerens elever, ble lavt. Grunnen til dette var både det ordinære frafallet man har i slike undersøkelser (elever, lærere, klasser og skoler som ikke besvarer spørreskjemaet) og et resultat av at lærerne kun skulle besvare spørsmålene med tanke på en bestemt klasse og ett bestemt fag. Dette er en svakhet ved datamaterialet som begrenser hvilke analyser man kan gjøre, og hvilke slutninger man kan trekke. Siden det alltid er en fordel med et stort antall respondenter og god svarprosent, må man ofte jobbe hardt med datainnsamlingen for å øke sannsynligheten for at mange personer svarer. Likevel skal det nevnes at spørreundersøkelser som metode er en tids- og kostnadseffektiv måte å innhente datamateriale fra mange respondenter på. Et stort antall responser gir muligheten til å undersøke hvor ofte hendelser inntreffer og hvor mange som deler ulike synspunkt. Med tilstrekkelig representativitet kan man generalisere fra utvalget til populasjonen som utvalget er trukket fra. Dette er svært viktig i vårt tilfelle. Resultatene fra slike undersøkelser kan potensielt påvirke nasjonal skolepolitikk, og da er det avgjørende at resultatene stammer fra et representativt utvalg av norske elever og lærere.

En generell svakhet ved bruk av spørreundersøkelser er knyttet til validiteten ved slutningene som trekkes basert på slikt datamateriale. Når et spørsmål blir stilt, er det ikke sikkert at elevene oppfatter dette spørsmålet på samme måte. De kan ha ulike forståelser av bestemte ord eller misforstå spørsmålet på andre måter. Det er heller ikke gitt at det er lett å svare presist. Spørsmål om «hvor ofte» ting skjer, kan ha stor feilmargin, enkeltepisoder kan være glemt eller huskes feil, og noen spørsmål opplever respondentene at de ikke har relevante erfaringer til å svare på.

Spørreundersøkelsen som er gjennomført i dette prosjektet, har en av sine styrker i det at spørsmålene er stilt for to konkrete fag – norsk og matematikk. Det å stille konkrete spørsmål bidrar til å forebygge feilslutninger knyttet til utfordringene nevnt i avsnittet over. Konkrete spørsmål er lettere å forstå, og de er lettere å besvare presist. Dette er særlig viktig når man skal belyse forhold i klasserommet.

Klasseromssituasjonen kan variere mye fra fag til fag, noen undervisningsmetoder brukes kun i enkelte fag og elevenes forhold til læreren avhenger, ikke overraskende, av hvilken lærer det er snakk

om. I eksempel 2 nedenfor presenteres en analyse hvor det skilles mellom de to fagene. Dette gjøres også i de påfølgende sammenligningene med Elevundersøkelsen.

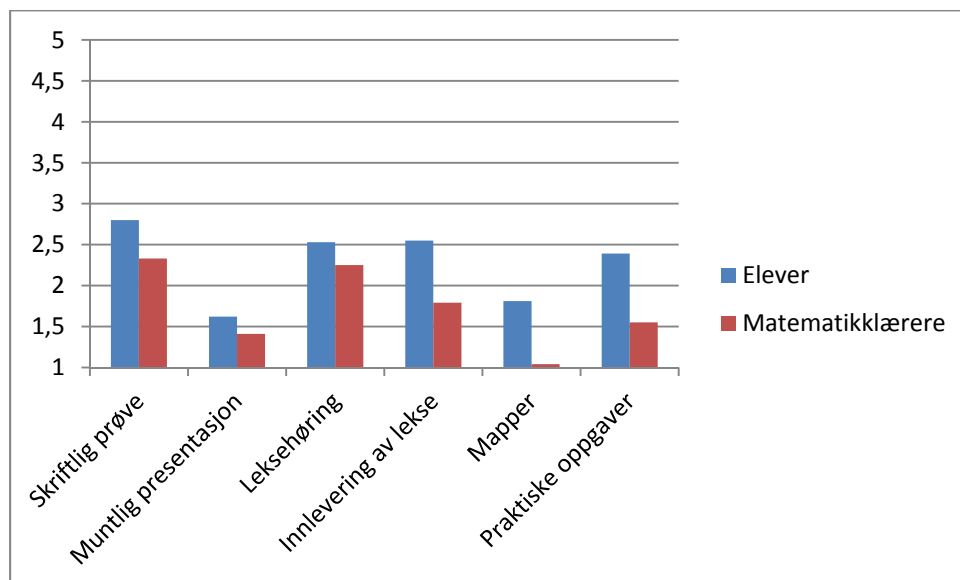
Et problem som oppstår når man ønsker å stille spørsmål om konkrete forhold i ulike fag, er knyttet til lengden på undersøkelsen. Den begrensede tiden og tålmodigheten respondentene har til rådighet betyr at man må prioritere hvilke temaer som inkluderes og hvor konkret man kan gå inn på hvert tema. Det er med andre ord begrenset hvor dypt og bredt man kan komme ved bruk av et rimelig antall spørsmål.

Spørreskjemalengden er en av flere faktorer som påvirker motivasjonen til de som svarer på slike undersøkelser. Hvis mange respondenter besvarer spørreskjemaet useriøst eller uten tilstrekkelig konsentrasjon, bidrar dette til å svekke validiteten. Slike responser kan være vanskelig å identifisere og dermed krevende å ta høyde for i analysene. Denne problemstillingen er særlig aktuell når et spørreskjema skal besvares av respondenter i skolealder, noe som viste seg tydelig i spørreundersøkelsen gjennomført i dette prosjektet. I delrapport 4 diskuteres de såkalte «tøysekoppene» blant respondentene. Hele 8 % av elevbesvarelsene ble ansett som useriøse og fjernet fra analysene. Dette er en relativt høy andel, og slike tøysekopper bidrar dermed til å senke antallet elever som inngår i analysene.

Vi har nå pekt på to viktige styrker ved spørreundersøkelsen gjennomført i dette prosjektet, nemlig at vi kan se på forhold i klasserommet fra to ulike perspektiver og at spørsmål er stilt for to konkrete fag. Dette kan føre oss ett skritt nærmere kontekstnær informasjon om hva som skjer i klasserommet. Her følger to eksempler hvor disse to egenskapene ved spørreundersøkelsen blir benyttet.

6.2.1 Eksempel 1: Kobling av elevers og læreres responser

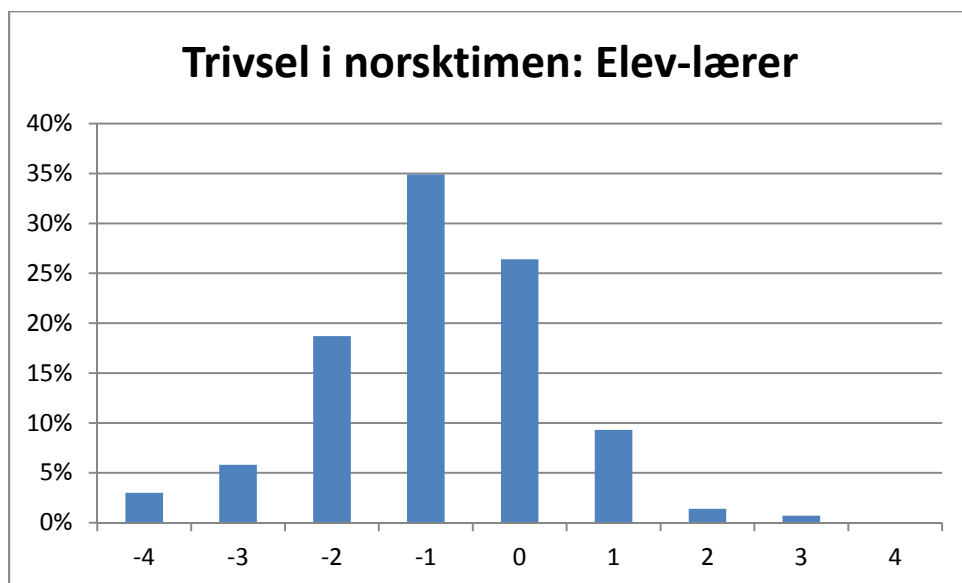
Til forskjell fra Elevundersøkelsen presentert i kapittel 5, har man i spørreundersøkelsen mulighet til å sammenligne responser fra elever med responser fra deres lærere. I figur 6.1 er alle lærere samlet i en gruppe og alle elever som har sin lærer representert i datamaterialet i den andre gruppen. Slik kan vi blant annet se hvordan lærere og elever besvarer de samme spørsmålene vedrørende vurderingshyppigheten i matematikk:



Figur 6.1: Spørsmålene «Hvor ofte blir arbeidet ditt i matematikk vurdert ved hjelp av ...» (elever) og «Hvor ofte vurderer du elevens arbeid ved hjelp av ...» (lærere), besvart på skalaen «Flere ganger i uken» (kodet 5), «En gang i uken», «1-3 ganger i mnd», «2-4 ganger i halvåret» og «Sjeldnere» (kodet 1). Gjennomsnittsverdier.

Figuren viser tydelig at elevene rapporterer en større hyppighet av ulike vurderingsformer i matematikk enn det læreren gjør. Særlig hevder elevene at de må levere inn lekser og mapper oftere enn det læreren selv hevder. Det samme svarmønsteret viser seg også for vurderingshyppighet i norsk.

Det er naturlig å tenke seg at trivselen i et klasserom oppleves ganske likt for lærere og elever. I Figur 6.2 vises følgende: Elevene har svart på spørsmålet «Hvor godt trives du i norsktimene» ved å krysse av på en skala fra «Ikke i det hele tatt» til «Svært godt». Deres norsklærer har svart på spørsmålet «Hvor godt trives du som lærer i denne klassen» med den samme skalaen. Differanse mellom elevens trivsel og lærerens trivsel er regnet ut. Dermed blir verdien 0 om eleven og læreren trives like godt. Verdien blir positiv om eleven trives bedre enn læreren, og verdien blir negativ om læreren trives bedre enn eleven. Merk at elevene er spurt om trivsel «i timen» mens læreren er spurt om trivsel «i klassen».



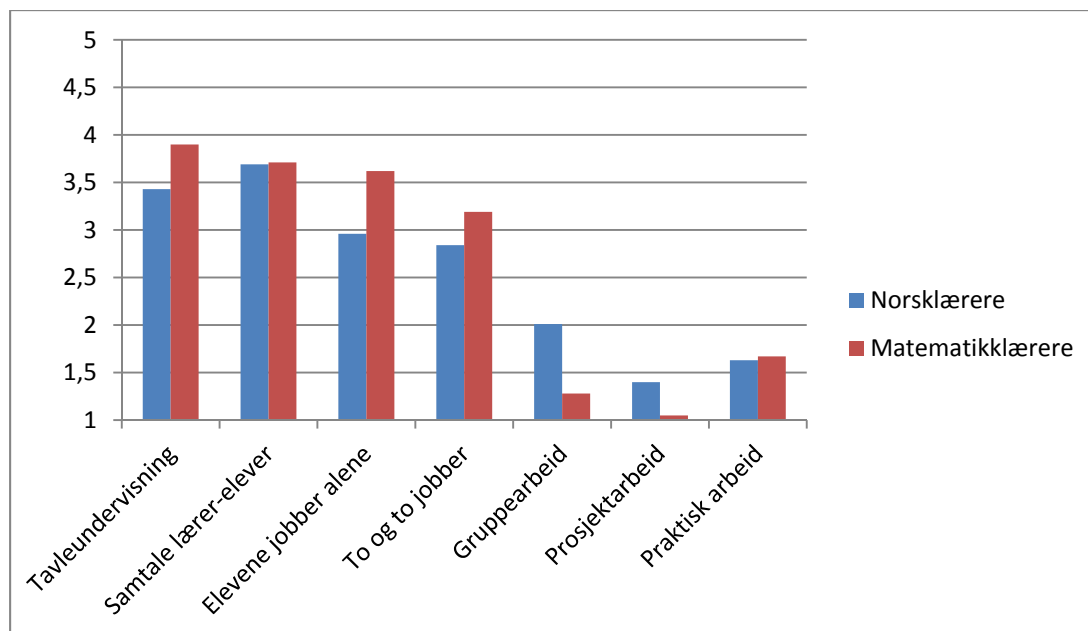
Figur 6.2: Differansen mellom elevers og deres lærers svar på spørsmålene «Hvor godt trives du i norsktimene» (elever) og «Hvor godt trives du som lærer i denne klassen» (lærere), besvart på en skala fra «Ikke i det hele tatt» (kodet 1) til «Svært godt» (kodet 5). Prosentandeler.

Mønsteret viser en normalfordeling rundt -1. Dette betyr at lærerne stort sett trives noe bedre enn elevene i disse klassene. Kun omtrent 1 av 10 elever rapporterer høyere trivsel enn sin lærer i norsk (stolper med positivt fortegn), mens det i 1 av 10 tilfeller er stort avvik i motsatt retning (-4 og -3), hvor lærerne trives «godt» eller «svært godt» og eleven trives «ikke noe særlig» eller «ikke i det hele tatt». Også her ser vi samme mønster for matematikk som for norsk.

Multivariate analyser, presentert i delrapport 4 (Opheim og Wiborg 2012), indikerer at lærernes syn på egne undervisningsmetoder alene ikke har signifikante utslag på elevens prestasjoner. Dersom det er et samsvar mellom lærerens og klassens syn på hva som skjer i klasserommet, har undervisningsmetodene derimot signifikante utslag på prestasjonene. Dette kan tolkes som at det ikke er lærerens ideer om undervisning som er avgjørende, men at det er lærerens ideer kombinert med hvorvidt elevene opplever at dette blir realisert. For eksempel viser analysene at det er en negativ sammenheng mellom elevenes resultater og tilfeller der læreren mener undervisningen er mer lærerstyrt enn det elevene oppfatter. En direkte sammenligning av elevenes og lærernes responser på spørsmål vedrørende bruk av ulike undervisningsmetoder er derimot vanskelig. Selv om spørsmålene er like og responsskalene har like mange kategorier, er benevnningen ulik. Elevene skulle besvare slike spørsmål ved å oppgi antall ganger dette ble gjort per uke, måned eller halvår, mens lærerne besvarte disse ved å oppgi hvor stor andel av undervisningstidene slike undervisningsformer ble brukt. Dette var for å få en bedre oversikt over omfang av ulike undervisningsmetoder blant lærerne enn det man

ville fått ved å bruke samme svaralternativ som til elevene. Ulempen er midlertid en begrenset mulighet til direkte å sammenligne elever og læreres svar på disse spørsmålene.

Senere i dette kapitlet skal vi se hva elevene svarer på spørsmål om bruken av undervisningsmetoder i norsk og matematikk hver for seg. Selv om deres besvarelser ikke kan sammenlignes direkte med lærernes besvarelser, har vi muligheten til å gjøre en tilsvarende sammenligning mellom norsk- og matematikklærere.

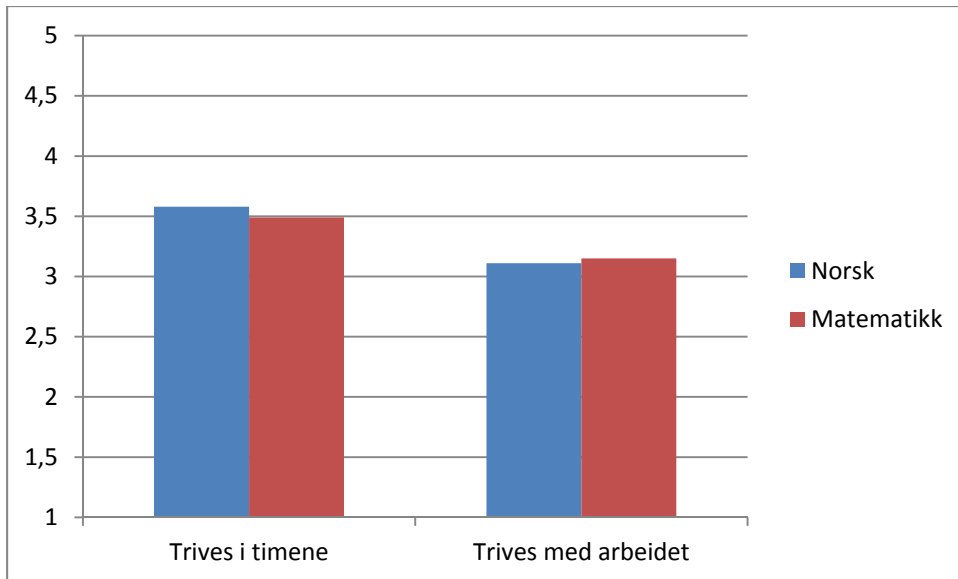


Figur 6.3: Spørsmålet «Hvor ofte bruker du følgende arbeidsmåter», besvart av lærere på en skala fra «Aldri eller nesten aldri» (kodet 1) til «I nesten alle timer» (kodet 5). Gjennomsnittsverdier.

I Figur 6.3 kommer det frem tydelige forskjeller mellom undervisningsmetodene norsk- og matematikklærere forteller at de benytter seg av. Matematikklærerne bruker signifikant mer tavleundervisning og lar elevene jobbe mer alene eller to og to. Norsklærerne bruker signifikant mer gruppe- og prosjektarbeid. Praktisk talt ingen matematikklærere forteller at prosjektarbeid er en del av deres undervisning.

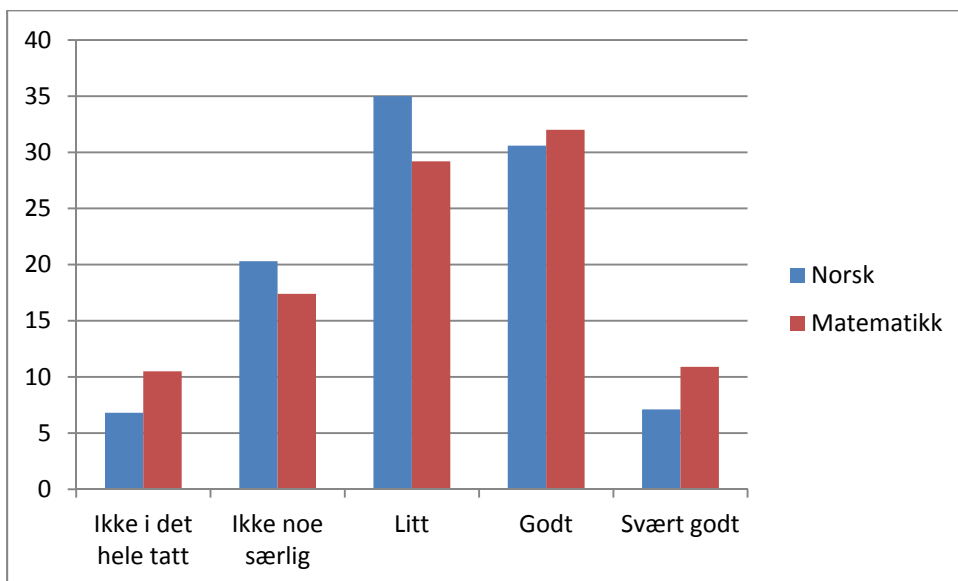
6.2.2 Eksempel 2: Spørsmål for to konkrete fag

I tidligere delrapporter er det blant annet fokusert på trivsel og arbeidsro når læringsmiljø har vært diskutert, og nå skal vi se på hva spørreundersøkelsen har å fortelle oss om forskjellen på trivsel i norsk og matematikk: Spørsmålene «Hvor godt trives du i timene» og «Hvor godt liker du skolearbeidet» er stilt for norsk og matematikk, med fem svaralternativer kodet fra 1 («Ikke i det hele tatt») til 5 («Svært godt»). Gjennomsnittsverdiene er vist i Figur 6.4:



Figur 6.4: Spørsmålene «Hvor godt trives du i timene» og «Hvor godt liker du skolearbeidet» gitt for norsk og matematikk, besvart av elever på en skala fra «Ikke i det hele tatt» (kodet 1) til «Svært godt» (kodet 5). Gjennomsnittsverdier.

Figuren viser at gjennomsnittverdiene til de to fagene er svært like. En liten forskjell mellom de to fagene finner man når man ser på fordelingen mellom de ulike svaralternativene for spørsmålet vedrørende trivsel med arbeidet (Figur 6.5). Her ser vi at matematikk, i større grad enn norsk, er litt mer «enten – eller». Det er både flere som ikke liker arbeidet i matematikk i det hele tatt, men samtidig også flere som liker dette svært godt.



Figur 6.5: Spørsmålet «Hvor godt liker du skolearbeidet» gitt for norsk og matematikk, besvart av elever på en skala fra «Ikke i det hele tatt» (kodet 1) til «Svært godt» (kodet 5). Prosentandeler.

De multivariate analysene av Elevundersøkelsens spørsmål knyttet til trivsel, viser at trivsel har en positiv effekt på resultatene i nasjonale prøver i regning, lesing og engelsk. Videre påpekes det at sammenhengen er sterkere med prestasjonsnivået enn med prestasjonsutviklingen. I spørreundersøkelsen viser det seg at denne positive effekten viser seg først og fremst i matematikk. Sammenhengen er svakere i norsk. Generelt ser det ut til at elever med svakt prestasjonsnivå trives bedre i norsktimene enn i matematikktimene, noe som kan belyse resultatene i Figur 6.5.

6.3 Sammenligning med Elevundersøkelsen

Noen generelle styrker og svakheter vedrørende spørreundersøkelsen er allerede nevnt. Flere av disse gjelder også Elevundersøkelsen, som er presentert i kapittel 5 og i anvendt i delrapport 1 og 2. I kapittel 5 ble det påpekt at man i Elevundersøkelsen har gitt de fleste spørsmålene for alle fagene samlet. Problemet med dette er at elevene kan oppleve ulike fag svært forskjellig. Variasjonen i undervisnings- og vurderingsformer kan være så stor mellom fagene at en samlet skår for alle fagene er vanskelig å tolke. Det er heller ikke sikkert at det som fremmer god læring i matematikk, er det samme som fremmer god læring i norsk, engelsk, historie eller kroppsøving. Videre kan det være krevende for respondentene å svare på enkelte spørsmål da disse omhandler alle fag samtidig. Når det for eksempel er spørsmål om hvor ofte noe inntreffer, vil noen prøve å regne ut et gjennomsnitt av alle fag, mens andre kan tenke på faget eller fagene hvor spørsmålet er mest relevant.

Vi skal i det følgende se på noen tilfeller hvor Elevundersøkelsens resultater utdypes ved hjelp informasjon innhentet i spørreundersøkelsen gjennomført i dette prosjektet. Denne sammenligningen er ment å skulle eksemplifisere verdien av kontekstnær informasjon generelt, og verdien av å stille spørsmål for fagene hver for seg spesielt.

Utvalget i denne sammenligningen

Spørreundersøkelsen ble gjennomført våren 2011 i 9. trinn og i Vg1. Datamaterialet fra Elevundersøkelsen som brukes her ble samlet inn våren 2010. For å fange opp respondenter fra det samme elevkullet, kunne man valgt å se på elever i Elevundersøkelsen som gikk i 8. og 10. trinn våren 2010. Dette var ikke ønskelig ettersom det kan være store forskjeller på opplevelsene i ungdomsskolen og i videregående skole. For å ha best mulig sammenlignbare tall, er derfor følgende respondenter inkludert i resultatene som gjengis her:

- Spørreundersøkelsen: Alle respondenter som gikk i 9. trinn og Vg1 våren 2011.
- Elevundersøkelsen: Alle respondenter som gikk i 9. trinn og Vg1 våren 2010.

Fordelingen av respondenter med tanke på klassetrinn og kjønn viser at de to utvalgene er tilfredsstillende like, til tross for en liten overvekt jenter blant respondentene i spørreundersøkelsen (Opheim og Wiborg 2012).

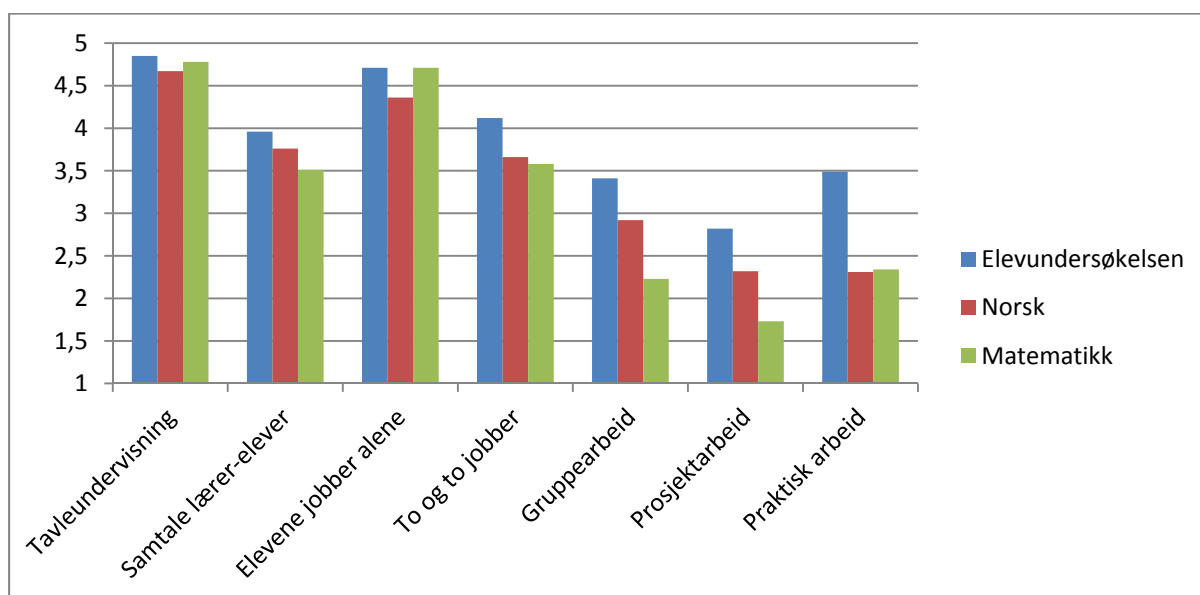
Tabell 6.1: Respondenter fra Elevundersøkelsen vår 2010 og spørreundersøkelsen vår 2011 som er inkludert i analysene i dette delkapitlet.

	Elevundersøkelsen vår 2010	Spørreundersøkelsen vår 2011
Klassetrinn		
9. trinn	31698 (36 %)	1025 (35 %)
Vg1	56241 (64 %)	1905 (65 %)
Kjønnsfordeling		
Jenter	42920 (49 %)	1489 (54 %)
Gutter	44504 (51 %)	1287 (46 %)

En forutsetning for å kunne gjøre en sammenligning av to påfølgende årskull, er at man kan generalisere fra ett årskull til det neste. Med andre ord er det forutsatt her at de som deltok i spørreundersøkelsen våren 2011, ville svart ganske likt på Elevundersøkelsen (2011) som de som deltok i Elevundersøkelsen året før (2010). Dette er en svakhet med sammenligningen som følger. Vi betrakter det likevel som sannsynlig at resultatene fra Elevundersøkelsen er tilstrekkelig stabile mellom to årskull til å peke på hovedtrekk i norsk skole.

Undervisningsformer

Ett av hovedtemaene i de foregående delrapportene i dette prosjektet er undervisningsformer. Særlig har det vært et fokus på bruken av lærerstyrte og elevaktive undervisningsformer. Lærerstyrte aktiviteter inkluderer tavleundervisning, plenumssamtaler i klasserommet og at elevene jobber alene, mens arbeid to og to, gruppearbeid, prosjektarbeid og praktisk arbeid klassifiseres som elevaktivt arbeid. I Elevundersøkelsen ble dette spørsmålet stilt: «Hvor ofte brukes følgende arbeidsmåter?» Deretter fulgte beskrivelsen av de sju arbeidsmåtene nevnt over, og respondentene skulle krysse av for ett av følgende fem alternativer: Flere ganger i uken, en gang i uken, 1-3 ganger i måneden, 2-4 ganger i halvåret eller sjeldnere. Disse ble kodet med verdier fra 1 (sjeldnere) til 5 (flere ganger i uken), og gjennomsnitt ble regnet ut for de ulike arbeidsmåtene. I spørreundersøkelsen ble samme spørsmål og svaralternativer gitt, men da ble spørsmålet stilt separat for norsk og matematikk. Gjennomsnittsverdiene fra Elevundersøkelsen og spørreundersøkelsen («norsk» og «matematikk») er vist i Figur 6.6:



Figur 6.6: Spørsmålet «Hvor ofte brukes følgende arbeidsmåter?» besvart av elever i Elevundersøkelsen (blå stolper) og av elever i spørreundersøkelsen for norsk (røde stolper) og matematikk (grønne stolper) separat, på skalaen «Flere ganger i uken» (kodet 5), «En gang i uken», «1-3 ganger i mnd», «2-4 ganger i halvåret» og «Sjeldnere» (kodet 1). Gjennomsnittsverdier.

De blå søylene er gjennomsnittsverdiene for de ulike arbeidsmåtene i Elevundersøkelsen. Ettersom det her spørres om erfaringer fra skolen generelt, noe som inkluderer alle fag, bør disse stolpene alltid være like høye eller høyere enn stolpene som gjelder spesifikt for norsk (røde) og matematikk (grønne). Ved å skille mellom ulike fag på denne måten, kan vi blant annet se at:

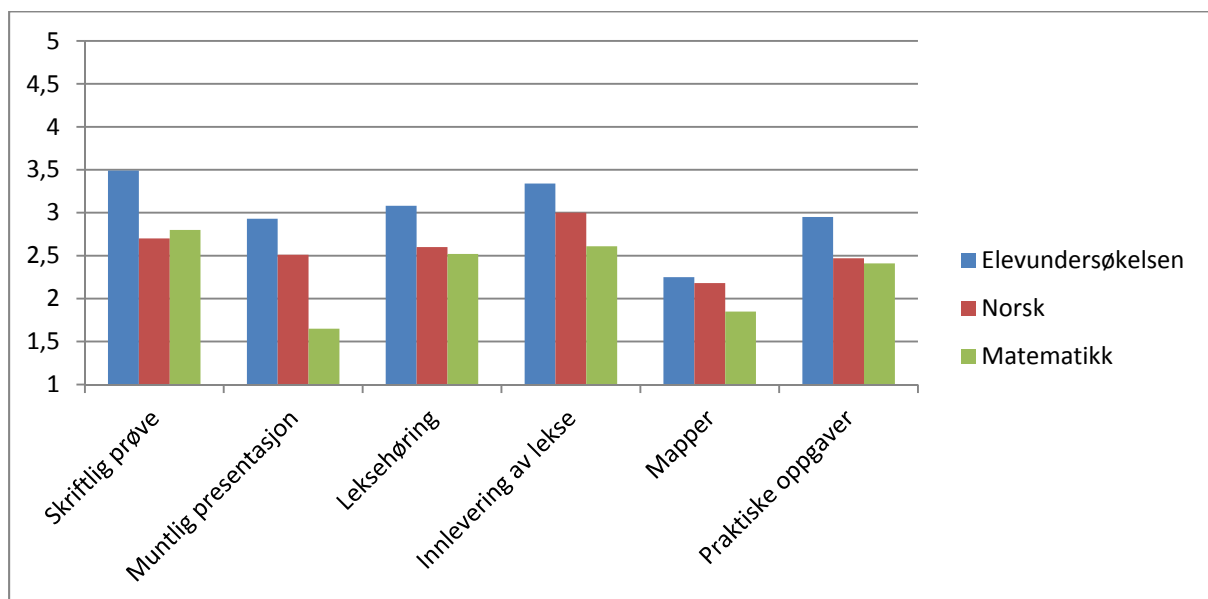
- Ut fra Elevundersøkelsen får man inntrykk av at praktisk arbeid skjer like ofte som gruppearbeid og noe sjeldnere enn arbeid to og to og klasseromsamtaler. I spørreundersøkelsen ser det ut til at dette i liten grad gjelder norsk og matematikk. Praktisk arbeid skjer primært i andre fag.
- Når det gjelder prosjektarbeid og gruppearbeid ser vi at det skjer sjeldnere i matematikk, signifikant oftere i norsk men sannsynligvis like ofte eller oftere i enkelte andre fag.
- For de resterende fire arbeidsformene fra venstre i figuren, skjer dette ofte både i norsk og matematikk. I disse to fagene ser det ut til å være utstrakt bruk av tavleundervisning og individuelt arbeid, men vi kan samtidig ikke utelukke at dette skjer like ofte i andre fag.
- Generelt ser det ut til at bruken av lærerstyrt undervisning er større enn elevaktiv undervisning, særlig i matematikkfaget.

De multivariate analysene fra Elevundersøkelsen viser at bruk av lærerstyrt undervisning korrelerer med elevaktiv undervisning, men at det er den lærerstyrte undervisningen som bidrar positivt til resultatene på nasjonale prøver i regning, lesing og engelsk. Analyser fra spørreundersøkelsen bidrar stort sett til å bekrefte disse funnene. Man finner dessuten en tendens til at lærerstyrt undervisning kan bidra til å svekke effekten av sosial bakgrunn i norskfaget i Vg1. Den tilsvarende tendensen er ikke signifikant i matematikk.

Selv om det tidligere i dette kapitlet ble nevnt at man ikke kunne sammenligne læreres og elevers besvarelser vedrørende undervisningsmetoder direkte, merker vi oss at Figur 6.3 og Figur 6.6 samsvarer godt på noen punkter. Både lærere og elever i matematikk rapporterer mindre bruk av gruppe- og prosjektarbeid enn lærere og elever i norsk. Hovedforskjellen mellom de to figurene, er at lærere i matematikk forteller at elevene jobber mer to og to enn lærere i norsk forteller. Denne forskjellen er ikke synlig når elevene rapporterer hva som skjer i klasserommet. Videre svarer ikke elevene like tydelig som lærerne at det er mer bruk av tavleundervisning i matematikk enn i norsk.

Vurderingshyppighet

Delrapport 1 og delrapport 4 tar blant annet for seg ulike former for vurdering. I Elevundersøkelsen fikk respondentene spørsmålet «Hvor ofte vurderes arbeidet ditt ved hjelp av...» etterfulgt av de seks vurderingsformene skriftlig prøve, muntlig presentasjon, leksehøring, innlevering av hjemmearbeid, mapper og praktiske oppgaver. Disse skulle besvares ved å markere ett av følgende alternativer: Flere ganger i uken, en gang i uken, 1-3 ganger i måneden, 2-4 ganger i halvåret eller sjeldnere. Ved å kode responsene med verdier fra 1 (sjeldnere) til 5 (flere ganger i uken) kunne gjennomsnittsverdier for vurderingshyppighet regnes ut for de ulike vurderingsformene. Samme spørsmål og svaralternativer ble gitt i spørreundersøkelsen for norsk og matematikk separat. De ulike gjennomsnittsverdiene for vurderingshyppighet kan sees i Figur 6.7:



Figur 6.7: Spørsmålet «Hvor ofte vurderes arbeidet ditt av ...» besvart av elever i Elevundersøkelsen (blå stolper) og av elever i spørreundersøkelsen for norsk (røde stolper) og matematikk (grønne stolper) separat, på skalaen «Flere ganger i uken» (kodet 5), «En gang i uken», «1-3 ganger i mnd», «2-4 ganger i halvåret» og «Sjeldnere» (kodet 1). Gjennomsnittsverdier.

Som i Figur 6.6 er det slik at de blå søylene viser gjennomsnittene fra Elevundersøkelsen, og siden denne omfatter alle fagene samlet, bør disse søylene være høyere enn spørreundersøkelsens tilsvarende søyler i norsk og matematikk. Figur 6.7 viser oss at spørreundersøkelsen gir oss følgende informasjon som Elevundersøkelsen ikke kan bidra med:

- Praktiske oppgaver, leksehøring og skriftlig prøve brukes omtrent like ofte i norsk og matematikk.
- Muntlig presentasjon og innlevering av lekse skjer signifikant sjeldnere i matematikk enn i norsk og andre fag. Særlig er det sjeldent med muntlige presentasjoner i matematikk.
- Mapper er lite brukt som vurderingsform i alle fag, og signifikant mindre i matematikk enn i norsk.
- Generelt er det mindre vurdering i matematikk enn i norsk, med unntak av omfanget av skriftlige prøver.

Multivariate analyser av Elevundersøkelsen viser at bruk av mapper og praktiske oppgaver er negativt assosiert med prestasjonsnivået, mens de fire andre vurderingsmetodene er positivt assosiert med dette. Generelt går hovedskillet mellom «de som praktiserer omfattende evaluering og de som ikke gjør det», og det ser ut til å være noe som heter «for mye»: Stor grad av evaluering er negativt assosiert med prestasjonene i 10. trinn. Analyser fra spørreundersøkelsen tyder på at den negative effekten av for mye vurdering er sterkere i matematikk enn i norsk. Kanskje er det mindre læring i vurderingsformene i matematikk enn i vurderingsformene i norsk?

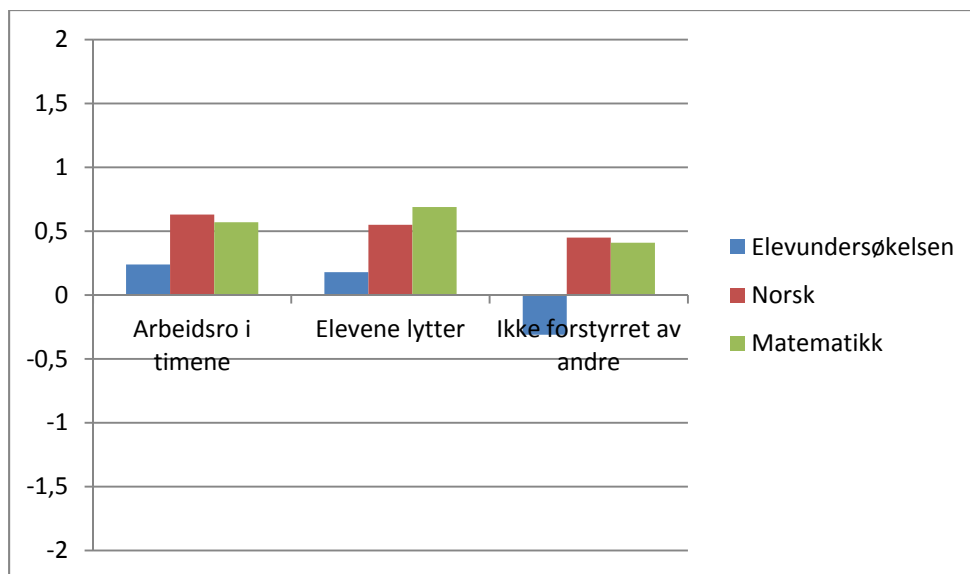
Arbeidsro

Vedrørende *arbeidsro* er disse uttalelsene gitt både i Elevundersøkelsen og i spørreundersøkelsen, med små variasjoner i formuleringen: «Det er god arbeidsro i timene», «Elevene er stille og hører etter når læreren snakker» og «Jeg blir ofte forstyrret av andre elever når jeg arbeider». I spørreundersøkelsen gis uttalelsene separat for norsk og matematikk. Hovedforskjellen mellom de to undersøkelsene er at det denne gangen ikke opereres med samme antall svarkategorier for spørsmålene knyttet til uro. I Elevundersøkelsen er det fem svaralternativer (helt uenig, litt uenig, verken eller, litt enig og helt enig) mens i spørreundersøkelsen er det fire svaralternativer (svært uenig, litt uenig, litt enig og svært enig)¹. Det finnes ingen fasit for hvordan et slikt sammenligningsproblem skal løses. Likevel, med tilstrekkelig normalfordelte responser og en omkodning av responsene, kan det argumenteres for at en sammenligning av gjennomsnittsverdier fortsatt kan være tilstrekkelig robust til vårt bruk i denne rapporten. Følgende steg er derfor tatt:

- Det ble bekreftet at alle spørsmålene har tilstrekkelig normalfordelte responser
- Skalaen fra 1 til 5 (Elevundersøkelsen) er omgjort til en skala fra -2 til 2, hvor 3 («verken eller») er omkodet til 0
- Skalaen fra 1 til 4 (spørreundersøkelsen) er også omgjort til en skala fra -2 til 2, men hvor ingen svarkategori har fått verdien 0.
- Spørsmålet «Jeg blir ofte forstyrret av andre elever når jeg arbeider» er reversert slik at alle tre spørsmål har samme fortegn (økning på skalaen betyr større i arbeidsro)
- Gjennomsnittsverdier er regnet ut

I Figur 6.8 kan vi derfor sammenligne elevenes responser på spørsmålene fra Elevundersøkelsen med spørsmålene fra spørreundersøkelsen. Merk igjen at disse resultatene bør tolkes med forsiktighet på grunn av de metodologiske utfordringene knyttet til bruken av ulike skalaer.

¹ Svarkategorien 'verken enig eller uenig' ble fjernet i spørreundersøkelsen ut fra begrunnelsen om at spørsmålene omhandler læringsmiljøet i konkrete fag, noe elevene kan forventes å ha en mening om – enten om de er enige eller uenige.



Figur 6.8: Spørsmålene «Det er god arbeidsro i timene», «Elevene er stille og hører etter når læreren snakker» og «Jeg blir ofte forstyrret av andre elever når jeg arbeider», besvart av elever i Elevundersøkelsen (blå stolper) og av elever i spørreundersøkelsen for norsk (røde stolper) og matematikk (grønne stolper) separat, på to ulike skalaer. Se diskusjon over figuren.

I de foregående figurene har det vært slik at stolpen som representerer Elevundersøkelsen logisk sett skal være minst like høy eller høyere enn stolpene for de spesifikke fagene. I dette tilfellet ser vi at de blå stolpene for alle fag under ett (Elevundersøkelsen) gjennomgående er lavere enn for de spesifikke fagene. En tolkning kan være at respondentene tolker spørsmålene som mål på «mengden uro», og at stolpen for alle fag under ett dermed gjenspeiler mengden uro akkumulert fra alle fag. Denne stolpen vil derfor trekke i motsatt retning siden positiv retning på skalaen betyr redusert uro.

I figuren ser det ut til at forskjellene mellom de to fagene er små. I frekvenstabellene for disse spørsmålene (ikke inkludert her) får man bekreftet at respondentenes svarfordeling er nesten identisk for norsk og matematikk. Det er heller ikke nevneverdig variasjon i forholdet mellom arbeidsro og prestasjoner i de to fagene.

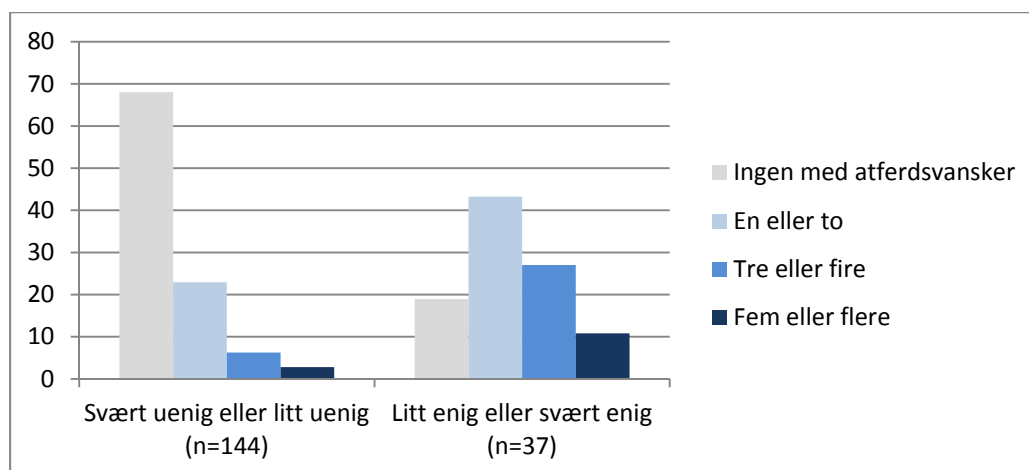
Merk hvordan konklusjonen vedrørende arbeidsro ville vært annerledes om man bare hadde hatt resultatene fra en av disse to undersøkelsene. Om «0» er det nivået for «akseptabelt» ville man med resultatene fra Elevundersøkelsen konkludert at for mange blir forstyrret av andre. Resultatene fra spørreundersøkelsen gjennomført i dette prosjektet, ville indikert at det ikke er et stort problem i de ulike fagene at medelever forstyrrer hverandre. Dette viser hvor avgjørende det er å stille spørsmål konkret for de ulike fagene. Det er vanskelig å tolke resultatene for et spørsmål som gjelder alle fagene er samlet, hvor man i dette eksemplet ser ut til å ha få resultater som er «mengden uro» akkumulert fra alle fag.

6.4 En spørreundersøkelse med utviklingspotensial

Før vi avslutningsvis i dette kapitlet diskuterer verdien av spørreskjema-data i debatten om ressurser og resultater, skal vi kort se på hvordan spørreundersøkelsen kan utvikles videre. Styrkene ved å belyse forhold i klasserommet fra ulike perspektiver og for konkrete fag, har kommet tydelig frem. Veien videre kan være å gå enda mer i dybden på de temaene man anser som viktige. Vi finner at blant annet at trivsel henger sammen med læring, og at lærerne rapporterer høyere trivsel enn elevene. Hvilke faktorer er det som skaper trivsel? Hva er det som skaper mistrivsel? Opplever elever og lærere at det er de samme forholdene som bidrar til et trivelig klasserom? Arbeidsro er trukket frem i eksemplet over. Her kan man gå mer i dybden og stille spørsmål om hvilke situasjoner som skaper uro og hva som skal til for at det skal bli arbeidsro. Dette er for øvrig også tema i casestudien, se kapittel 3 og 7. Videre er ulike utslag av vurderingshyppighet diskutert i dette prosjektet. På hvilke

måter er det vurdering fremmer læring? Hvordan forbereder elevene seg til skriftlige prøver, muntlige presentasjoner og til praktiske oppgaver? Endrer elevenes strategier seg med økt vurderingshyppighet, og varierer strategiene mellom de ulike fagene? På tilsvarende måte kan man velge å gå i dybden på andre områder man anser som viktige og dermed få mer kontekstnær kunnskap som kan knyttes til ulike former for prestasjonsmål.

I elev-versjonen av spørreskjemaet er flere spørsmål hentet fra Elevundersøkelsen. Denne er presentert i kapittel 5. Tilsvarende inneholder lærer-versjonen av spørreskjemaet spørsmål fra TALIS-prosjektet (Teaching and Learning International Survey). Dette er en studie ledet av OECD med fokus på lærere og skoleledere, og belyser forhold vedrørende kompetanseutvikling, vurdering og tilbakemelding, skoleklime, skoleledelse, undervisningssyn og pedagogisk praksis. På samme måte som i eksemplene for Elevundersøkelsen over, kan spørreundersøkelsen supplere TALIS med kontekstnær informasjon. Et eksempel på dette er knyttet til uro i klasserommet. I TALIS er det stilt spørsmål om «det er mye bråk i denne klassen». Det samme spørsmålet er stilt i spørreundersøkelsen, men her får lærerne også mulighet til å oppgi hvor mange elever det er med atferdsvansker i denne klassen. Hvis man grupperer lærerne etter hva de svarer på spørsmålet om det er mye bråk i denne klassen, ser man en klar sammenheng mellom bråk og antallet med atferdsvansker:



Figur 6.9: Prosentandel lærere med ulikt antall elever med atferdsvansker i klassen. Lærerne er delt i to grupper gitt av deres respons på utsagnet «Det er mye bråk i denne klassen» («Svært uenig» og «Litt uenig» i en gruppe, «Litt enig» og «Svært enig» i den andre gruppen).

Figur 6.9 gir et inntrykk av at en hovedforskjell går mellom de som har elever med atferdsvansker og de som ikke har noen slike elever. Fra gruppen som er uenige i at det er bråk i denne klassen, til gruppen som er enige i at det er bråk, synker andelen som svarer «ingen med atferdsvansker» fra nær 70 prosent til ca. 20 prosent, mens andelen lærere med «en eller to», «tre eller fire» og «fem eller flere» stiger. Nærmere analyser (ikke vist i figur) finner at i det klassen går fra å ha ingen til å ha én elev med atferdsvansker, nesten femdobles sannsynligheten for at lærere er enig i at det er en bråkete klasse.

6.4.1 En retning for fremtidige studier?

I spørreundersøkelsen gjennomført i dette prosjektet, har vi sett på to konkrete fag og gjort det mulig å koble elevenes svar fra deres lærere. Denne eksplorative undersøkelsen har tatt sikte på å belyse elevers og læreres erfaringer og øke innsikten i hva som faktisk skjer i klasserommet. Designet gjør det mulig å koble de ulike fenomenene som undersøkes med ulike prestasjonsmål. Spørsmålet er hvorvidt tillegget av informasjon som er eksemplifisert i dette kapitlet får implikasjoner som anses som tilstrekkelig verdifulle til at det er verdt å bruke tid og krefter på å videreutvikle, gjennomføre og analysere data fra slike spørreundersøkelser i årene som kommer.

Spørreundersøkelser for å samle kontekstnær informasjon – nyttig, men krevende

Innledningsvis i dette kapitlet ble det nevnt at kontekstnær informasjon potensielt kan løfte diskusjonen vedrørende ressurser og resultater i skolen. Dette kan blant annet omfatte de erfaringene elever og lærere har med skolehverdagen. Ved hjelp av spørreundersøkelser kan man fremskaffe slik kontekstnær kunnskap og samtidig knytte disse til elevens resultater på måter som vanskelig lar seg gjøre med kvalitative studier. Denne formen for kontekstnær kunnskap gir dermed muligheten til å vurdere hvorvidt ulike sider ved elevens læringsmiljø bidrar til ulike former for læringsutbytte. Resultatene fra Elevundersøkelsen gjengitt i kapittel 5 samt de supplerende resultatene fra dette prosjektets spørreundersøkelse, tyder på at slike forhold har en naturlig plass i en bred diskusjon om norsk skole.

Det kan være krevende å stille spørsmål som er tilstrekkelig konkrete. Sammenligningen av de to undersøkelsene i dette kapitlet viser verdien av å stille spørsmål konkret for ulike fag. En samlet skår for alle de ulike fagene gjorde det vanskelig – om ikke umulig – å trekke valide slutninger basert på responsene til enkelte spørsmål. I dette prosjektets spørreundersøkelse ble det sett nærmere på norsk- og matematikkfaget. Man kan altså være nødt til å vurdere på forhånd hvilke fag man ønsker å undersøke. Inkluderer man mange fag, betyr dette samtidig at man ikke kan stille like mange spørsmål om hvert fag.

I resultatene fra spørreundersøkelsen gjengitt over, viste det seg å være betydelige forskjeller mellom lærernes og elevenes opplevelser av samme situasjon. Dette er viktig lærdom for fremtidige studier og vil være av betydning når man velger ut respondenter til slike spørreundersøkelser. Her er det viktig å presisere at det kan være svært krevende å utføre analyser hvor man kobler elever, lærere og foreldre. Det ordinære frafallet i slike undersøkelser vil ofte, som i vårt tilfelle, kunne føre til statistisk usikkerhet i analysene.

Andre faktorer kan også svekke validiteten. I vårt tilfelle dukket det opp mange spørsmål: Hvor reliable er elevenes «opptellinger» av hva som skjer i klasserommet? I et tidvis kaotisk klasserom, oppfatter elevene undervisningsaktivitetene på samme måte? Skal det læreren forsøkte på eller det den enkelte eleven gjorde, telles? Regnes alt elevene gjør som læreren får tilgang på som vurdering, ettersom standpunkt karakteren er en helhetsvurdering, eller bare det elevene får karakter på? Respondentene har ganske sikkert svært ulike oppfatninger av disse tingene. Det er med andre ord krevende å utforme et spørreskjema angående kontekstuelle forhold som respondentene kan svare entydig på. Her må man gjøre en grundig jobb i forkant. Pilottester og fokusgrupper er nyttige hjelpemiddel i en slik prosess for å sikre god reliabilitet og dermed en styrket validitet.

Angående reliabiliteten av datamateriale fra spørreskjema, viste det seg å være ytterligere en utfordring i prosjektets spørreundersøkelse: De såkalte «tøysekoppene». Åtte prosent av respondentene hadde et svarmønster som indikerte at besvarelsen ikke var seriøs. Disse ble det besluttet å fjerne fra analysene. Å fjerne tøysekoppene hadde vesentlig betydning for resultatene, dette er nærmere beskrevet i delrapport 4 (Opheim og Wiborg 2012). Utfordringen med å unngå useriøse svar er ofte større med yngre respondenter, og vil sannsynligvis også være tilfelle for andre som ønsker å se på forhold i norsk skole ved hjelp av spørreskjemadata. Jobben med å få et tilstrekkelig antall elever til å levere seriøse besvarelser kan være stor.

Det å benytte spørreundersøkelser gir muligheten til å se på ulike prestasjonsmål. Som nevnt tidligere, er det ikke bare karakterer og testresultater som indikerer resultater av opplæringen. I undersøkelsene ble det blant annet stilt spørsmål om hvor godt elevene selv vurderer egen mestring av ulike oppgaver, og dette kan gjerne anses som et viktig mål for skolen. Om elevene opplever mestring, kan det blant annet tenkes at de vil jobbe mer målrettet i fremtiden og at de får større utholdenhet i utdanningsløpet. Informasjon om elevenes holdninger til fag, kunnskap og forskning er også et mål for utdanningen. Slike holdninger er et godt utgangspunkt for et arbeidsliv hvor evnen til omstilling og videreutdanning er stort. Generelt kan det sies at elevenes holdninger, selvtilitt og opplevelse av å mestre utfordringer

er viktige for mål for allmenndannelsen skolen skal bidra til. Spørreskjemaer ser ut til å være en hensiktsmessig måte å fremskaffe slik informasjon på.

Som eksemplifisert med de ulike prestasjonsmålene, kan man i et spørreskjema stille spørsmål om mange ulike forhold. Undersøkelsene i dette kapitlet fokuserer på blant annet undervisnings- og vurderingsmetoder, trivsel, arbeidsro og studieteknikk. Her finnes det flere temaer som kan være interessante å inkludere i et spørreskjema. Elevenes forhold til lekser er et slikt tema. Svært få spørsmål om dette er inkludert i Elevundersøkelsen. Debatten om bruk av lekser kommer jevnlig, hvor noen fremhever hvordan dette gir elevene mer tid med skolearbeidet, mens andre viser til at dette forsterker sosiale forskjeller ved at bestemte grupper av foreldre involverer seg forskjellig i barnas lekser. Hva tenker elevene selv om lekser? Hvordan er deres foreldre involvert i leksene? Opplever de selv at de lærer mer av dette, eller er leksene en kilde til frustrasjon, hvor de mislykkes og må gå videre uten å få nødvendig hjelp?

Et annet aktuelt tema er bruken av arbeidsplaner. Setter elevene pris på friheten til å velge tidspunkt for oppgaveløsning og vanskelighetsgrad på oppgavene? Tenker de selv at de bruker denne tiden fornuftig? Opplever de å mestre oppgaver, og samtidig bli utfordret i tilstrekkelig grad, når de velger oppgavenivå selv? Kommer de i utakt med de andre i klassen ved bruk av slike arbeidsplaner?

Videre er det flere viktige temaer hvor man i stor grad kun har data på skolenivå som utgangspunkt for diskusjon. Betydningen av klassestørrelse er et eksempel på dette. Her gir GSI-dataene kun informasjon om gjennomsnittlig klassestørrelse på skolen. Kontekstnær informasjon om elevenes og lærernes opplevelse av et klasserom med få elever og et klasserom med mange elever ville vært et verdifullt bidrag i denne debatten. Føler elevene i små klasser at de får mer hjelp og tilpasset undervisning enn elever i store klasser? Anser elevene det selv som problematisk at det er så mange andre som trenger hjelp og skaper uro? I hvor stor grad øker arbeidspresset på lærere med flere elever? Er det antallet elever eller antallet krevende elever som påvirker arbeidssituasjonen mest? Diskusjonen om klassestørrelse er på denne måten relatert til diskusjonen om elever som trenger ekstra oppfølging.

GSI-dataene om ressurser brukt til spesialundervisning er også på skolenivå. Begrenset informasjon finnes om spesialundervisningens plass i klasserommet og hvordan dette oppleves av elever og lærere. Hva tenker de andre elevene om de som får spesialundervisning? Hvordan påvirker dette klassemiljøet? Er det utelukkende positivt for læringsmiljøet at de som har spesielle utfordringer får andre tilbud? Med tanke på områder som spesialundervisning, klassestørrelse, arbeidsplaner, lekser og andre temaer som inngår i debatten om ressurser og resultater i skolen, vil det være en styrke å trekke inn kontekstnær informasjon i debatten.

6.5 Oppsummering og konklusjon

En rekke faktorer bør inngå i en diskusjon av ressurser og resultater i norsk skole. Kontekstnær kunnskap er nødvendig for å belyse ulike forhold, som blant annet inkluderer ulike aspekter ved læringsmiljøet og ulike prestasjonsmål. Spørreundersøkelser fremstår som en hensiktsmessig metode når slik informasjon skal innhentes fra et stort antall elever og lærere. Muligheten til å stille konkrete spørsmål for konkrete fag, å koble svar fra ulike aktører og å få informasjon om en rekke ulike temaer er utvilsomt av stor verdi. Ved hjelp av spørreskjemaer kan man få innsikt i hva som skjer i de enkelte klasserommene og knytte klasserommets ulike prosesser med elevenes ulike former for læringsutbytte. Spørreundersøkelser av elever utgjør dermed en kilde til informasjon om hele input-prosess-resultat-modellen, noe som verken er mulig med registerdata eller casestudier.

Samtidig er det viktig å påpeke at det er store utfordringer knyttet til bruk av spørreundersøkelser i slike kontekster. Analyser og diskusjoner av bruk av spørreundersøkelsen gjennomført i dette prosjektet fremhever sårbarheter og utfordringer ved denne type datakilde. Dette gjelder både utforming av spørreskjema, datainnsamling og kvaliteten på respondentenes svar. I kapittel 5 og i dette kapitlet diskuteres blant annet vanskeligheter ved å trekke valide slutninger fra

Elevundersøkelsen. Generelt kan det sies at det er krevende å velge hvilket detaljnivå en undersøkelse skal ha, å finne spørsmål som elevene har samme forståelse av og kan svare entydig på, å motivere elevene til å besvare spørreskjemaet seriøst og det å få nok responser til å gjøre robuste analyser hvor man kobler data fra ulike aktører.

Konklusjonen her blir derfor at spørreundersøkelser er viktig kilde til informasjon som rommer flere aspekter ved elevenes læringsprosess, samtidig som det gir mulighet til å skille mellom prestasjonsforskjeller på ulike nivå – på individnivå mellom elever, innad og mellom ulike klasser/basisgrupper og forskjeller innad og mellom ulike skoler. Slike data åpner for generaliserbar kunnskap om sammenhenger mellom det som skjer i klasserommet og elevenes læringsutbytte, noe som er nødvendig for å få en bred diskusjon om ressurser og resultater i norsk skole.

7 Oppsummering: Veien videre mot gode kvalitetsindikatorer

Prosjektet «Ressursbruk og læringsresultater i grunnopplæringen» startet opp i begynnelsen av 2010 og har nå pågått i vel tre år. Formålet med prosjektet har vært å undersøke sammenheng mellom ressursbruk i skolen og elevenes læringsresultater, og gjennom dette å finne gode indikatorer på kvalitet i grunnopplæringen.

Kapittel 1 begynte med en presentasjon av prosjektet i sin helhet. Det omfatter bruk av datakilder, analyser og metoder, teoretisk grunnlag, sentrale problemstillinger og hypoteser, samt innholdet i de to delprosjektene som sammen med hovedprosjektet utgjør prosjektet «Ressurser og resultater». Kapitlet ga en gjennomgang av sentrale tidligere studier og kunnskapsstatus innen de forskningsfelt som dekkes av analysene. Dette er i hovedsak tre felt: Skoleeffektivitetsforskningen (primært økonomisk forskning), klasseromsforskningen (primært pedagogisk forskning), og ulikhetsforskningen (primært sosiologisk forskning). Alle de tre feltene omhandler studier av elevenes læringsutbytte, men kjennetegnes av ulike problemstillinger, perspektiver, datakilder og metoder. I prosjektet har vi, så godt det har latt seg gjøre, kombinert de ulike tilnærmingene, for på den måten ha et bredt perspektiv på skolen.

I kapittel 2 ble de seks delrapportene fra prosjektet oppsummert. Deretter ble hovedresultater fra prosjektet, sett under ett, presentert og diskutert i lys av tidligere forskning. I de påfølgende kapitler har vi på grunnlag av resultatene, diskutert de sentrale datakildene som brukes i prosjektet.

Vi har kalt sluttrapporten 'Jakten på kvalitetsindikatorene'. I det ligger et fokus på det som har vært en gjennomgående problemstilling i dette prosjektet: Å undersøke sammenhenger mellom ressursbruk i skolen og elevenes læringsresultater, for gjennom dette å identifisere indikatorer på kvalitet i grunnopplæringen. I dette siste kapitlet vil vi oppsummere prosjektets bidrag til kunnskapsstatus, deretter presenterer vi flere anbefalinger for arbeidet med videre utvikling av kvalitetsindikatorer for grunnopplæringen. Avslutningsvis vil vi peke på sentrale tema for videre forskning på dette feltet.

7.1 Prosjektets hovedfunn: Bidrag til kunnskapsstatus

Prosjektet har bidratt med ny kunnskap om elevenes prestasjonsnivå på 5., 8. og 10. trinn, samt prestasjonsnivået i Vg1. Prosjektet har undersøkt elevenes prestasjonsutvikling fra 5-8. trinn, 8-10 trinn, samt fra 10. trinn-Vg1. Dette har bidratt med ny kunnskap om hvilke faktorer ved skolens ressurser, elevenes læringsmiljø, elevsammensetningen ved skolen og kjennetegn ved elevene selv som bidrar til elevenes prestasjonsutvikling på tre ulike trinn i grunnopplæringen.

Flere av analysene i prosjektet bekrefter funn fra tidligere forskning. Også tidligere studier har pekt på relativt små forskjeller mellom skoler i Norge, men store forskjeller innad ved den enkelte skole. Prestasjonsforskjellene mellom elever som går på samme skole og i samme klasse, er sterkt knyttet til kjennetegn ved elevene – deres sosiale bakgrunn, kjønn og minoritetsbakgrunn. Gjennom analyser av dette omfattende datamaterialet, har prosjektet bidratt med ny kunnskap om hvordan like rammefaktorer og kjennetegn ved elevenes læringsmiljø og undervisningssituasjon kan ha ulik betydning for ulike grupper av elever.

Et sentralt funn i prosjektet er betydning av elevenes læringsmiljø. Alle elever tjener på å ha et godt læringsmiljø, men for noen elevgrupper, særlig de som står i fare for å falle utenfor skolens fellesskap, er dette særlig viktig. Elever med minoritetsbakgrunn har større risiko for å være svakt integrert i skolens fellesskap, og dermed falle utenfor. Prosjektet peker på sammenhenger mellom integrasjon, elevenes oppfatning av sitt læringsmiljø og deres læringsutbytte.

Prosjektet har også bidratt med innsamling av nye empiriske data, både i form av en kvantitativ spørreundersøkelse til elever og lærere, samt en kvalitativ casestudie av fire skoler i Oslo. Spørreundersøkelsen til elever og lærere har gitt ny kunnskap om sammenhenger mellom elevenes og lærerens oppfatning om læringsmetoder i norsk og matematikk, og elevenes prestasjoner i disse fagene. Ved at både elever og faglærerne deres har fått likelydende spørsmål om undervisningssituasjonen og læringsmiljøet gir dette en kilde til direkte lærer-elev-informasjon. Analysene avdekker ulike oppfatninger om hva som skjer i klasserommet blant lærer og elever.

Et formål med casestudien var, med utgangspunkt i de kvantitative analysene i prosjektet, å undersøke nærmere ulike læringsprosesser i klasserommet. Casestudien har bidratt med nye data og gitt utdypende kunnskap om hva som skjer i klasserommet. Resultater fra casestudien har bidratt til diskusjonen av resultater og tolkninger av analysene fra registerdata og spørreundersøkelsen. Ved bruk av ulike datakilder og metodisk tilnærminger, har prosjektet samlet sett bidratt til å belyse komplekse sammenhenger mellom skoleressurser (input), læringsprosesser i klasserommet og elevenes prestasjoner (resultater/output). Skoleressurser er her forstått i vid forstand – både som skolens materielle og menneskelige ressurser, elevenes oppfatninger av eget læringsmiljø, samt kjennetegn ved elevene selv, deres foreldre, og skolens elevsammensetning.

7.2 Anbefalinger for videre utvikling av kvalitetsindikatorer for grunnopplæringen

I kapittel 3, 4 og 5 diskuteres styrker og svakheter ved henholdsvis nasjonale prøver og andre prestasjonsmål, GSI og Elevundersøkelsen som kilder til informasjon om sammenheng mellom skoleressurser og elevenes læringsresultater, samt mulig potensial som kvalitetsindikatorer for grunnskolen. I kapittel 6 diskuteres hvordan spørreundersøkelsen blant elever og lærere, men også spørreundersøkelser generelt, kan bidra til å gi informasjon om sammenhenger mellom ulike former for skoleressurser og elevresultater som ikke fanges opp av andre datakilder. Nedenfor oppsummeres våre anbefalinger for utvikling av nye kvalitetsindikatorer, samt anbefalinger for forbedring av eksisterende datakilder med dette som formål.

7.2.1 Anbefalinger for videre utvikling av nasjonale prøver og andre prestasjonsmål

I dette prosjektet blir nasjonale prøver og karakterdata benyttet som mål på elevenes prestasjoner. Til tross for at disse datakildene har vært gjenstand for kritikk, blant annet for å representere et snevert mål på skolekvalitet, fremstår disse kildene som naturlige valg for undersøkelser av typen som er gjort her. Datamaterialet fra nasjonale prøver er svært solid, og elevenes resultater på denne type tester henger ofte svært tett sammen med andre målsettinger ved norsk skole som ikke måles direkte i testene. For å styrke anvendeligheten av nasjonale prøver som prestasjonsmål i norsk skole, gis her tre anbefalinger. En grundigere gjennomgang av nasjonale prøver som system foreligger i en egen NIFU-rapport (Seland mfl 2013).

En enhetlig fritakspraksis: Retningslinjene for fritak fra de nasjonale prøvene fører til stor variasjon i fritakspraksis mellom skoler. Dette medfører statistisk usikkerhet, særlig ved analyser av elever på et lavt prestasjonsnivå. Å sikre større likhet i fritakspraksis mellom skoler vil øke den statistiske treffsikkerheten og gjøre det slik at man tydeligere kan kommunisere resultatene fra nasjonale prøver. *Tydeligere* retningslinjer behøver samtidig ikke å bety *strengere* retningslinjer. Det frarådes å fremme en praksis som fremtvinger deltakelse av elever som ikke har utbytte av eller forutsetninger for å besvare oppgavene.

Reell prestasjonsutvikling: Fortjenesten av å la noen oppgaver gå igjen fra år til år ser ut til å være betydelig. Dette gir muligheten til å analysere den reelle prestasjonsutviklingen mellom årskullene både nasjonalt og for den enkelte skole. Slike analyser er svært viktige om man skal kunne vurdere effekten av ulike innsatsordninger. Noen utvalgte oppgaver må altså holdes hemmelig, noe man har god erfaring med blant annet i PISA- og TIMSS-undersøkelsene.

Elevenes motivasjon: En forutsetning for høy datakvalitet er at elevene er motiverte til å gjøre sitt beste på nasjonale prøver. Dette kan ivaretas og styrkes ved å gjøre prøvene så kortfattede som mulig. Videre vil elevenes motivasjon avhenge av deres innsikt i verdien av nasjonale prøver. Om denne kommuniseres tydelig, med vekt på det potensielle utbyttet den enkelte elev kan ha å gjennomføre testene, kan dette redusere antallet useriøse og halvhjertede responser og styrke datakvaliteten. Informasjon om elevenes motivasjon og innsats kan også innhentes ved å spørre elevene om dette, som en del av gjennomføringen av prøvene. Slike spørsmål er blant annet inkludert i PISA-undersøkelsen.

7.2.2 Anbefalinger for videre utvikling av GSI

Prosjektet finner at få av variablene fra GSI framstår som potensielle indikatorer på kvalitet i skolen. To unntak er informasjon om antall undervisningstimer og antall pc'er per elev. I prosjektet er antall undervisningstimer og antall pc'er per elev gjennomgående funnet å være positivt relatert til elevenes prestasjoner. Analysene kan ikke belyse årsakssammenhenger mellom disse forholdene. Med dette som forbehold, framstår timetall antall pc'er per elev som mulige kvalitetsindikatorer for skolen.

GSI inneholder mye informasjon om skolene som brukes av ulike aktører og med ulikt formål. I arbeidet med utvikling av kvalitetsindikatorer, er det likevel et tankekors at informasjonen i GSI i liten grad synes å være relatert til elevenes prestasjonsnivå, målt som gjennomsnitt på skolenivå. Det betyr ikke at skoleressurser ikke har betydning for elevens læringsutbytte, men må ses i sammenheng med den variasjon i elevenes prestasjoner som finnes innen den enkelte skole. Det er lite sammenheng mellom skolens gjennomsnittlige ressursituasjon og elevenes gjennomsnittlige prestasjoner. I lys av analyser i ulike deler av prosjektet, foreslår vi nedenfor tre områder for en videreutvikling av GSI som kan øke mulighetene for å bruke opplysninger fra GSI som indikatorer på kvalitet i skolen.

Mer informasjon om lærernes kvalifikasjoner: GSI inneholder lite informasjon om lærernes kvalifikasjoner, utover andelen lærere på skolen med godkjent utdanning. Lærerens rolle som klasseleder og kunnskapsformidler fanges dermed i liten grad opp av GSI. På grunnlag av funn fra prosjektet anbefales en videreutvikling av denne siden ved GSI. Målet på hvor mange lærere som har godkjent utdanning gir i liten grad uttrykk for lærernes kompetansenivå, og fungerer dermed dårlig som ressursindikator på skolenivå. Lærernes utdanning er stadig oppe til debatt i norsk offentlighet. Nærmere opplysninger i GSI om hvor lang og hvilken type utdanning lærerne har, ville være gunstig for muligheten til å undersøke betydningen av lærernes kvalifikasjoner for elevens prestasjoner. Dette ville samtidig gitt flere opplysninger om skolens sammensetning av lærere. For eksempel gir informasjon om lærernes alder, ansiennitet og tidligere arbeidserfaring, større muligheter til å undersøke sammenhenger mellom elevenes prestasjoner og skolens lærerressurser.

Opplysninger om timetall i ulike fag: GSI inneholder ingen informasjon om timetall i ulike fag. Det gjør at det ikke er mulig å undersøke om flere timer i et bestemt fag øker prestasjonene i det faget. Sammenhengen kan kun undersøkes på et generelt plan: om flere timer totalt fører til bedre

prestasjoner. Opplysninger om timetall i sentrale skolefag, som matematikk, norsk og engelsk ville gjort opplysningene i GSI mindre generelle og dermed mer egnet som kvalitetsindikatorer.

Mer kontekstnær informasjon: Informasjonen i GSI er primært på skolenivå. I en del tilfeller kan ressursindikatorene knyttes til bestemte klassetrinn, eller grupper av klassetrinn (1. trinn, 2-4. trinn, 5-7. trinn eller 7-10. trinn). Fordi informasjonen i GSI ikke er knyttet til bestemte klasser eller elever, vet vi ikke hvordan de ulike ressursene som måles i GSI kommer den enkelte elev eller klasse til gode. Det gjør at variablene i GSI blir upresise som indikatorer på kvalitet. En utvikling av GSI i retning av å knytte deler av den informasjonen som registreres til bestemte klassetrinn, vil gi mer kontekstnær informasjon om de skoleressurser den enkelte elev møter. For eksempel vil informasjon om hvilke lærerressurser som er knyttet til den enkelte klasse, gjøre det mulig å undersøke betydningen av læreren for elevenes prestasjoner mer presist.

7.2.3 Anbefalinger for videre utvikling av Elevundersøkelsen

I prosjektet er elevenes læringsmiljø og undervisningssituasjon, slik det måles i Elevundersøkelsen, funnet å ha sammenheng med deres prestasjoner. At læringsmiljøet er trivelig, inkluderende, støttende, punktlig og faglig fokusert er positivt knyttet til elevenes prestasjonsnivå. Også kjennetegn ved undervisningen, som omfang av ulike undervisningsformer og omfang av vurdering, er funnet å være knyttet til elevenes prestasjoner på flere trinn i grunnopplæringen. Dette er kjennetegn ved elevenes læringsmiljø og undervisningssituasjon som kan fungere som kvalitetsindikatorer. Når vi likevel er forsiktige med å anbefale hvordan ulike sider ved elevenes læringsmiljø og undervisningssituasjon kan brukes som kvalitetsindikatorer for grunnopplæringen, har det sammenheng med usikkerheter og svakheter ved Elevundersøkelsen som datakilde. Dette er diskutert i kapittel 5. Elevundersøkelsen er en viktig kilde til informasjon om elevenes skolehverdag, men vi mener at potensialet ikke utnyttes fullt ut. Anbefalingene gir forslag til videre utvikling av Elevundersøkelsen, slik at den i større grad enn i dag kan gi informasjon om kvalitet i skolen.

Fagspesifikke spørsmål: I stedet for spørsmål om elevenes oppfatning av ulike sider ved undervisningssituasjonen generelt, som undervisning og evaluering, anbefales det at spørsmålene i større grad knyttes til spesifikke fag. Eksempelvis kan omfanget av ulike evaluerings- og undervisningsformer, variere fra fag til fag. Dette kan medføre at elevene forstår spørsmålene på ulike måter. Et alternativ til denne type generelle spørsmål, er mer fagspesifikke spørsmål som omhandler undervisningssituasjon og læringsmiljø i konkrete fag. I stedet for å formidle en generell oppfatning av undervisningen på skolen vil svarene på slike spørsmål gi innblikk i forskjeller mellom fag. Når spørsmål knyttes til bestemte fag, kan det samtidig åpnes for informasjon om elevenes oppfatning av læreren som kunnskapsformidler og klasseleder. Dersom slike opplysninger sammenholdes med elevenes prestasjoner i de samme fagene, vil Elevundersøkelsen i større grad belyse sammenhenger mellom elevenes læringsmiljø og prestasjonsnivå.

Redusert omfang: Elevundersøkelsen dekker en rekke ulike tema knyttet til elevenes læringsmiljø. I tillegg til en fast del, bestående av spørsmål som alle får, kan den enkelte skole velge ett eller flere tilleggstema. Selv om antall tilleggsspørsmål/tema er begrenset, vil spørreundersøkelsen oppleves som lang og omfattende for mange elever. Trolig er dette en svakhet ved Elevundersøkelsen som har betydning for elevenes motivasjon og svar kvalitet. Det anbefales derfor å redusere omfanget av undersøkelsen. En mulighet er å innføre rotering av oppgavesett (se neste punkt).

Rotering av oppgavesett: Samtidig som at undersøkelsen ikke bør bli lengre, har vi her foreslått å stille spørsmål for spesifikke fag. En måte å imøtekomme begge disse behovene kan være å rotere hvilke oppgavesett som benyttes fra år til år, hvor en del består av faste spørsmål og en del består av spørsmål for utvalgte fag. Om man for eksempel velger ut seks sentrale undervisningsfag, kan man stille spørsmål for to fag hvert år. Da kan man følge utviklingen for de ulike fagene med kontrollpunkter hvert tredje år. En slik rotering av oppgavesett har man blant annet i PISA-undersøkelsen, hvor det veksles mellom å ha fokus på lesing, matematikk og naturfag. Videre kan Elevundersøkelsen bli kortere dersom man finner at det også er andre temaer hvor måling hvert tredje år er tilstrekkelig.

Mer nyanserte spørsmål om undervisningsformer: Slik undersøkelsen gjennomføres i dag er det mulig for elevene å svare både at mye av undervisningen er lærerstyrt og at mye av undervisningen er elevstyrt. Dersom spørsmålene konstrueres slik at elevene i større grad må angi hvilke undervisningsformer det er mest av og hva det er mindre av, vil Elevundersøkelsen i større grad få tak i kontraster i omfang og bruk av ulike undervisningsformer.

Vurdere svarkategoriene: Enkelte av spørsmålene har svarkategorier som synes å være for grove til å fange opp den variasjonen som finnes. Selv om det ikke er et mål i seg selv at svarene skal fordele seg jevnt på en skala, bør det være et mål å unngå at alle havner i samme svarkategori. Et eksempel er svarkategoriene i spørsmålene om undervisningsformer der svarfordelingen i noen tilfeller er svært skjev. I tilfeller der svært mange av elevene krysser av i øverste eller nederste svarkategori, bør disse vurderes endret.

Mer informasjon om bakgrunnsforhold: Elevundersøkelsen spør i liten grad om kjennetegn ved elevene eller deres foreldre, som kunne gi informasjon om elevenes sosioøkonomiske bakgrunn eller innvandrerbakgrunn. Med utgangspunkt i den sterke sammenhengen mellom elevenes kjennetegn og deres prestasjoner, anbefales det å inkludere noe mer informasjon om individkjennetegn i Elevundersøkelsen.

Endre gjennomføringen: Med dagens ordning vil noen elever delta i Elevundersøkelsen en rekke ganger i løpet av grunnopplæringen. Hyppig deltakelse kan redusere elevenes motivasjon og respekt for undersøkelsen, noe som kan øke omfanget av useriøse svar. En mulig løsning på dette kan være å begrense skolens muligheter til å gjennomføre undersøkelsen for de samme elevene gjentatte ganger.

7.3 Tema for videre forskning

Prosjektet har dekket mange problemstillinger og spørsmål knyttet til hvilke forhold som har betydning for læringsutbytte for ulike grupper av elever. I tillegg til å besvare noen av disse spørsmålene, har prosjektet også åpnet opp for andre problemstillinger og pekt på nye spørsmål. Et generelt funn i analysene, er den relativt store forskjellen mellom elevenes læringsutbytte innen den enkelte skole og den enkelte klasse. Mye av denne forskjellen er knyttet til kjennetegn ved elevene. I lys av dette, er det naturlig å foreslå et større fokus på det som skjer ved den enkelte skole i videre forskning.

Helt avslutningsvis vil vi løfte frem noen av temaene som gjennom arbeidet med dette prosjektet, fremstår som særlig sentrale for å videreutvikle kunnskap om hvilke forhold som har betydning for elevenes læringsutbytte.

- Gjennomgangen og analyser av tilgjengelige datakilder og tidligere forskning avdekker et avvik mellom de tilgjengelige data og kvalitetsindikatorer og det tidligere forskning fremhever som viktig for elevenes læringsutbytte. Dette gjelder særlig betydning av læreren som kunnskapsformidler og tydelig klasseleder. Her finnes lite tilgjengelig informasjon i dagens registre og datakilder på nasjonalt nivå. Å utvikle kvalitetsindikatorer som har som mål å fange opp undervisningskvalitet og lærerens kompetanse i klasseledelse, bør derfor prioriteres. Mulige kvalitetsindikatorer vil kunne omfatte utvikling av sertifiseringsordninger for lærere, eller andre målbare former for informasjon om lærerens kompetanse i klasseledelse.
- Nærmere analyser av hvilke ulike sider ved læringsmiljøet som kan ha betydning for elevenes prestasjoner, samt analyser av sammenhenger mellom læringsmiljø og elevsammensetningen ved skolen, vil være en naturlig videreføring av resultatene i dette prosjektet. Resultater som peker i retning av betydning av integrasjon for prestasjoner blant elever med minoritetsbakgrunn er komplekse, men samtidig viktig å følge opp.
- Prosjektet peker mot mer detaljerte undersøkelser av sammenhenger mellom undervisningsformer og elevenes læringsutbytte i konkrete fag. En analytisk tilnærming for å

komme nærmere de reelle årsakssammenhengene mellom ressursindikatorer, hva som skjer i klasserommet (prosess) og elevenes prestasjoner, kan være å tilrettelegge for 'naturlige eksperimenter'. For eksempel ved å innføre endringer i skolen på ulikt tidspunkt, noe som muliggjør reelle effektstudier med reelle kontrollgrupper. Et eksempel på en slik form for implementering av tiltak i skolen er Ny GIV, som ble innført gradvis ved alle ungdomsskoler i perioden 2010-2011. Her er det imidlertid etiske forhold som må vurderes i hvert enkelt tilfelle. En annen metode for å undersøke nærmere de ulike elementer ved læringsprosessen som kan ha betydning for elevenes læringsutbytte, er å videreutvikle kvantitative datakilder som fanger opp dette. Spørreundersøkelsen til elever og lærere har dette som formål. Som diskutert i kapittel 6, kan spørreundersøkelser utgjøre en viktig kilde til informasjon om lærerens undervisning i et konkret fag, om den direkte interaksjon mellom lærer og elev, samt interaksjonen mellom lærer og klasse/gruppe.

- Studier av sammenheng mellom ulike former for ressurser i skolen og elevenes læringsutbytte bør i større grad enn i dag inkludere analyser av forskjeller mellom private og offentlig skoler. Her finnes lite empirisk forskning i Norge. Å inkludere skoler med ulike eieform og annet pedagogisk plattform i analysene, vil gi et bredere bilde av mulige sammenhenger mellom forholdet mellom skoleledelse, lærernes undervisningspraksis og elevenes læringsutbytte.
- Det bør tilrettelegges for forskning som i større grad forsøker å skille mellom kjennetegn ved elevene selv, inkludert elevens evner til læring, og kjennetegn ved deres foreldre. Dette for å få et riktigere bilde av hvilken betydning familiebakgrunn og miljøfaktorer har for elevenes prestasjonsutvikling, uten at dette blandes sammen med eventuelle medfødte egenskaper hos elevene. Forhold knyttet til familiebakgrunn og miljøfaktorer er ikke lett å skille fra hverandre, men forskningen innen dette feltet bør ha som ambisjon å utvikle analyser av sammenhenger og samspill mellom ulike faktorer med betydning for elevenes forutsetninger for læring.
- Casestudiens hypoteser om sammenheng mellom skolens pedagogiske plattform og kvaliteten på undervisningen bør følges opp i videre undersøkelser. Betydninger av faglig ledelse ved skolene og elevenes læringsutbytte, er et område med mye oppmerksomhet både politisk, fra media, foreldre og fra skoleeier/forvaltningen. Området preges imidlertid av mange antakelser, men lite sikker kunnskap.
- Det er behov for forskning som forsøker å isolere betydning av ulike undervisningsformer for elevenes læringsutbytte. Dette omfatter flere ulike tema. To aktuelle tema er betydningen av praktisk undervisning for elevenes læringsutbytte, og betydningen av nivådelt undervisning i ulike fag og blant ulike grupper av elever.
- Et tema i tilknytning til betydning av læreren som kunnskapsformidler og klasseleder, er behovet for mer forskning på effekter av insentiver for å øke lærernes kompetanse og å belønne gode lærere. Dette kan omfatte effekter av (ulike former for) resultatbasert lønn, etter- og videreutdanning, tilrettelegging i form av mindre tidsbruk på administrasjon og mer tid på undervisning, samt andre former for insentiver.

Referanser

- Alexander, R.J. (2000). *Culture and pedagogy: International Comparisons in Primary Education*. Oxford: Blackwell.
- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House* 78(5).
- Allison, Paul D. (2009). *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage Publications.
- Arnesen, C.Å. (2012). *Prestasjonsutvikling fra ungdomsskolen til første året i videregående opplæring. Delrapport 3 fra prosjektet 'Ressurser og resultater i grunnopplæringen'*. Oslo: NIFU.
- Aslam, M. og Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review*, 30, 559-574.
- Aukrust, V.G. (2003). «Å få ordet – å ta ordet. Jenter og gutters samtaledeltakelse i klasserommet.» Kapittel 5 i Klette, K. (red.) *Klasserommets praksisformer etter Reform 97*. Side 111-134. Oslo: Pedagogisk forskningsinstitutt.
- Bakken, A. (2010). *Prestasjonsforskjeller i Kunnskapsløftets første år. Kjønn, minoritetsstatus og foreldrenes utdanning*. Oslo: NOVA.
- Bakken, A. (2009a). *Ulikhet på tvers. Har foreldres utdanning, kjønn og minoritetsstatus like stor betydning for elevers karakterer på alle skoler?* NOVA-rapport 8/2009. Oslo: Norsk institutt for studier om oppvekst, velferd og aldring.
- Bakken, A. (2009b). «Er mannlige lærere viktige for gutters skoleprestasjoner?» I M. Raabe (red.): *Utdanning 2009: Læringsutbytte og kompetanse*. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Bakken, A., E. Borg, K. Hegna og E. Backe-Hansen (2008). *Er det skolens skyld? En kunnskapsoversikt om skolens bidrag til kjønnsforskjeller i skoleprestasjoner*. NOVA-rapport 9/2008. Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.
- Bakken, A. og K. Danielsen (2011). *Gode skoler – gode for alle? En casestudie av prestasjonsforskjeller på seks ungdomsskoler*. NOVA-rapport 10/2011. Oslo: NOVA.
- Bakken, A. og J.I. Elstad (2012). *For store forventninger? Kunnskapsløftet og ulikhetene i grunnskolekarakterer*. NOVA-rapport 7/2012. Oslo: Norsk institutt for studier om oppvekst, velferd og aldring.
- Bangert-Drowns, R.L. mfl (1991). Effects of frequent classroom testing. *Journal of Educational Research* 85(2): 89-99.
- Black, P. og D. Wiliam (2009). Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability* 21(1): 5-31.
- Bonesrønning, H. og J.M.V. Iversen (2008). *Suksessfaktorer i grunnskolen: Analyse av nasjonale prøver 2007*. Trondheim: Senter for økonomisk forskning.
- Bonesrønning, H. og J.M.V. Iversen (2010). *Prestasjonsforskjeller mellom skoler og kommuner: Analyse av nasjonale prøver 2008*. SØF-rapport nr. 01/10. Trondheim: Senter for økonomisk forskning.
- Borge, L.E. og M. Rønning (2009). «Bidrar økt ressursbruk i grunnskolen til bedre elevprestasjoner?» I Raabe (red). *Utdanning 2009 – læringsutbytte og kompetanse*. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Boudon, R. (1974). *Education, Opportunity and Social Inequality*. New York/London: John Wiley.
- Bourdieu, P. (1977). Cultural Reproduction and Social Reproduction. In Karabel, J. og Halsey, A.H (Eds.), *Power and Ideology in Education*. New York: Oxford University Press.
- Brophy, J.E. og T.L. Good (1986). *Teacher behavior and student achievement*. In M. Wittrock (red.), *Handbook of research on teaching* (3. utg. s. 328–375). New York: Macmillan.
- Brookover, W. mfl (1979). *Schools, Social Systems and Student Achievement: Schools can make a difference*. New York: Praeger.

- Byrhagen, K., Falch, T. og Strøm, B. (2006). *Frafall i videregående opplæring: Betydningen av grunnskolekarakterer, studieretning og fylke*. Rapport 8. Trondheim: SØF.
- Coleman, J.S. mfl (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.
- Eisner, E. W. (1979). *The education imagination. On the design and evaluation of school programs*. London, Collier Macmillan Publishers.
- Elstad, J.I. (2010). *Datakvalitet i Ung i Oslo 2006*. Notat 1/10. Oslo: NOVA.
- Falch, T. og B. Strøm (2005). Fordeling av lærerressurser mellom norske grunnskoler. I *Utdanning 2005*. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Falch, T. og L.R. Naper (2008). *Lærerkompetanse og elevresultater i ungdomsskolen*. SØF-rapport 1/08. Trondheim: Senter for økonomisk forskning.
- Fekjær, S. og M.S. Leirvik (2011). Silent gratitude. Education among second-generation Vietnamese in Norway. *Journal of Ethnic and Migration Studies* 37(1): 117-134.
- Gagné, R.M. (1974). *Learning for instruction*. Illinois: The Dryden Press Hinsdale.
- Gibson, M.A. og J.U. Ogbu (1991). *Minority Status and Schooling. A Comparative Study of Immigrant and Involuntary Minorities*. New York: Garland Publishing, Inc.
- Good, T.L. og J.E. Brophy (2008). *Looking in Classrooms* (10. utg.). New York: Pearson.
- Gravaas, B. C., T. Hægeland mfl (2008). *Skoleresultater 2007: en kartlegging av karakterer fra grunn- og videregående skoler i Norge*. Oslo: Statistisk sentralbyrå.
- Grubb, N.W. (2009). *The Money Myth – School resources, outcomes and equity*. New York: Russell Sage Foundation.
- Grøgaard, J.B. (2012). *Hva kjennetegner barneskoler som oppnår høy skår på nasjonale prøver? Delrapport 5 fra prosjektet 'Ressurser og resultater i grunnopplæringen'*. Oslo: NIFU.
- Grøgaard, J.B. (1997). *Skolekontroversen. Belyst ved to norske utvalgsundersøkelser*. Avhandling til dr.philos. Fafo-rapport 222. Oslo: Fafo.
- Grøgaard, J.B., H. Helland og J. Lauglo (2008). *Elevenes læringsutbytte. Hvor stor betydning har skolen? En analyse av ulikhet i elevers prestasjonsnivå i fjerde, syvende og tiende trinn i grunnskolen og i grunnkurset i videregående*. Oslo: NIFU STEP.
- Gustafsson, J.E. (2003). What do we know about effects of school resources on educational results? Paper presentert på konferansen: What educational policies achieve? *Swedish Economic Policy Review* 10 (2003): 77-110.
- Hacking, I. (1999). *The Social Construction of What?* Cambridge, Mass.: Harvard University Press.
- Halpern, D.F. (1997). Sex differences in intelligence. Implications for education. *American Psychologist* 52: 1091-1102.
- Hanushek, E.A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24:1141-1177.
- Hanushek, E.A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19:2: 141-164.
- Hanushek, E.A. og L. Wössmann (2007). *The Role of School Improvement in Economic Development*. Cambridge, Mass.: National Bureau of Economic Research. NBER Working Paper Series.
- Hanushek, E.A. og S.G. Rivkin (2006). Teacher quality. I: E.A. Hanushek og F. Welch (red.), *Handbook on the economics of education*. Amsterdam: Elsevier.
<http://edpro.stanford.edu/hanushek/admin/pages/files/uploads/HESEDU2018.pdf>
- Hattie, J. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J. og H. Timperley (2007). The power of feedback *Review of Educational Research* 77(1): 81-112.

- Haug, P. (2011). *Klasseromsforskning. Kunnskapsstatus og konsekvensar for lærarrolla og lærarutdanninga*. Rapport 21. Volda: Høgskulen i Volda.
- Haug, P. (red.) (2004). *Resultat frå evalueringa av Reform 97*. Oslo: Norges forskningsråd.
- Helsedirektoratet (2010). *Rammeverk for et kvalitetsindikatorsystem i helsetjenesten. Primær- og spesialisthelsetjenesten*. Oslo: Helsedirektoratet.
- Hernes, G. (1974). Om ulikhetens reproduksjon. Hvilken rolle spiller skolen? I Mortensen, M. S. (Ed.), *I Forskningens Lys: 32 artikler om norsk forskning i går, i dag og i morgen, NAVF 1949-1974*. Oslo: NAVF.
- Hernes, G. og K. Knudsen (1976). *Utdanning og ulikhet*. NOU 1976:46. Oslo: Universitetsforlaget.
- Hægeland, T., Kirkebøen, L. J., B. Bratsberg og O. Raaum (2011). *Value added-indikatorer. Et nyttig verktøy i kvalitetsvurdering av skoler?* Rapport 42/2011. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Hægeland, T., Kirkebøen, L. J. og O. Raaum (2010). *Skolebidragsindikatorer for videregående skoler i Oslo*. Rapport 36/2010. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Hægeland, T., L.J. Kirkebøen og O. Raaum (2008). *Ressurser i grunnskole og videregående opplæring i Norge 2003-2007*. Rapport 2008/3. Oslo-Kongsvinger: Statistisk sentralbyrå.
- Imsen, G. (2003). *Skolemiljø, læringsmiljø og elevutbytte. En empirisk studie av grunnskolenes 4., 7. og 10. trinn*. Trondheim: Tapir akademiske forlag.
- Imsen, G. (red.) (2004). *Det ustyrlige klasserommet. Om styring, samarbeid og læringsmiljø i grunnskolen*. Oslo: Universitetsforlaget.
- Jackson, C. (2006). *Lads and Ladettes in School*. London: Open University Press.
- Jakobsen, A, C. Rump, T. Clemmensen, og M. May (1999): *Kvalitetsudviklingsprosjektet «Faglig Sammenheng»*. Hovedrapport. CDM's skriftserie nr. 1. København: Danmarks Tekniske Universitet.
- Jonsson, J.O. og F. Rudolphi (2011). Weak Performance – Strong Determination: School Achievement and Educational Choice among Children of Immigrants in Sweden. *European Sociological Review* 27(41): 487-508.
- Kjærnsli, M. mfl (2004). *Retts spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003*. Oslo: Universitetsforlaget.
- Kjærnsli, M. og A. Roe (red.) (2010). *På retts spor. Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009*. Oslo: Universitetsforlaget.
- Klaveren, C.V. (2011). Lecturing style teaching and student performance. *Economics of Education Review*, 30, 729-739.
- Klette, K. (red.) (2003). *Klasserommets praksisformer etter Reform 97*. Oslo: Pedagogisk forskningsinstitutt.
- Klyve, A. (2009). *Sinte unge menn i kunnskapssamfunnet*. Nestun: Folilo.
- Kuncel, N. R. mfl (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology* 48(1): 148-161.
- Kunnskapsdepartementet (2011). *Myten om pengenes makt. Analyser om grunnopplæring*. Publisert på nett 6. juni 2011. Oslo: Kunnskapsdepartementet.
http://www.regjeringen.no/nb/dep/kd/dok/rapporter_planer/aktuelle-analyser/aktuelle-analyser-om-andre-tema/myten-om-pengenes-makt.html?id=646135
- Lauglo, J. (1996). *Motbakke, men mer driv?* Oslo: Ungforsk.
- Lee, J. (2008). Is test-driven external accountability effective? Synthesizing the evidence from cross-state causal-comparative and correlational studies. *Review of Educational Research* 78(3): 608-644.
- Lee, J. og T. Reeves (2012). Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990-2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis* 34(2): 209-231.
- Leirvik, M.S. (2010). For mors skyld. Utdanning, takknemlighet og status blant unge med pakistansk og indisk bakgrunn. *Tidsskrift for ungdomsforskning* 10(1): 23-47.

- Luyten, H. (2003). The size of school effects compared to teacher effects: An overview of the research literature. *School Effectiveness and School Improvement* 14(1): 31-51.
- Lyng, S.T. (2004). *Være eller lære? Om elevroller, identitet og læring i ungdomsskolen*. Oslo: Universitetsforlaget.
- Lødding, B. (2009). *Sluttere, slitere og sertifiserte: Bortvalg, gjennomføring og kompetanseoppnåelse blant minoritetsspråklige ungdommer i videregående opplæring*. Rapport 41/2008. Oslo: NIFU STEP.
- Lødding, B og N. Vibe (2010). "Hvis noen forteller om mobbing ... ". *Utdypende undersøkelse av funn i Elevundersøkelsen om mobbing, urettferdig behandling og diskriminering*. Rapport 48/2010. Oslo: NIFU.
- Løken, G. (2012). *Levd observasjon. En vitenskapsteoretisk kommentar til observasjon som forskningsmetode*. Oslo: Cappelen Damm akademisk.
- Markussen, E. og I. Seland (2013). *Den gode timen*. Oslo: NIFU.
- Marsdal, M. (2012). *Kunnskapsbløffen*. Oslo: Forlaget manifest.
- Meyer, R.H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review* 16(3): 283-301.
- Modood, T. (2004). Capitals, ethnic identity and educational qualifications. *Cultural Trends* 13(2): 87-105.
- Møller, J., T. S. Prøitz og P. Aasen (2009). *Kunnskapsløftet—tung bør å bære?* Oslo: NIFU STEP.
- Nordahl, T. (2005). *Læringsmiljø og pedagogisk analyse. En beskrivelse og evaluering av LP-modellen*. NOVA-rapport 19/2005. Oslo: Norsk institutt for forskning oppvekst, velferd og aldring.
- Nordenbo, S.E. mfl (2008). *Lærerkompetanse og elevers læring i barnehage og skole. Et systematisk review utført for Kunnskapsdepartementet, Oslo*. København: Danmarks Pædagogiske Universitetsforlag og Dansk Clearinghouse for Uddannelsesforskning.
- Norsk Offentlig Utredning (NOU) 1995: 18. *Ny lovgivning om opplæring «...og for øvrig kan man gjøre som man vil*. Oslo: Kirke-, utdannings- og forskningsdepartementet.
- Norsk Offentlig Utredning (NOU) (2002). *Førsteklasses fra første klasse. Forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem av norsk grunnopplæring*. NOU 2002:10, Utdannings- og forskningsdepartementet.
- OECD (2007). *PISA 2006 – Science competencies for tomorrow's world. Volume 1: Analysis*. Paris: OECD.
- OECD (2010). *PISA 2009 at a Glance*, Paris: OECD.
- Ogden, T. (2005). *Skolens mål og muligheter*. Stockholm: Statens folkhälsoinstitut.
- Opheim, V., J.B. Grøgaard og T. Næss (2010). *De gamle er eldst? Betydning av skoleressurser, undervisningsformer og læringsmiljø for elevenes prestasjoner på 5., 8. og 10. trinn i grunnopplæringen*. Oslo: NIFU STEP.
- Opheim, V. og Ø. Wiborg (2012). *Resultater fra en spørreundersøkelse blant elever og lærere på 9. trinn og i videregående trinn 1. Delrapport 4 fra prosjektet 'Ressurser og resultater i grunnopplæringen'*. Oslo: NIFU.
- Opplæringslova. *Lov om grunnskolen og den videregående opplæringa*, 17. juli 1998 nr 61.
- Ottesen, E. og J. Møller (red.) (2010). *Underveis, men i svært ulikt tempo. Et blick inn i ti skoler etter tre år med Kunnskapsløftet*. Oslo: NIFU STEP.
- Oxford Research (2009). *Elevene svarer! Analyse av Elevundersøkelsen 2009*. Kristiansand: Oxford Research.
- Pihl, J. (1998). *Minoriteter og den videregående skolen*. Avhandling for dr. polit., Pedagogisk forskningsinstitutt. Oslo: Universitetet i Oslo.

- Prøitz, T.S. (2010). Learning outcomes: What are they? Who defines them? When and where are they defined? *Educational Assessment, Evaluation and Accountability* 22(2): 119-137.
- Quinn, P.D. og A.L. Duckworth (2007). Happiness and academic achievement: Evidence for reciprocal causality. *Annual meeting of the American Psychological Society*. Washington, D.C.
- Rabe-Hesketh, S. og A. Skrondal (2012). *Multilevel and Longitudinal Modeling Using Stata*, Third Edition, two volumes. Stata Press.
- Raudenbush, S.W. (2004). *Schooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Educational Testing Service.
- Reynolds, D. og C. Teddlie (2000). The Processes of School Effectiveness. I: Teddlie, C. og Reynolds, D.: *The international handbook of school effectiveness research*. London: Falmer Press.
- Rumberger, R. (2011). *Dropping out. Why students drop out of high school and what can be done about it*. Harvard University Press: Cambridge.
- Rutter, M., B. Maughan, P. Mortimore, J. Ouston og A. Smith (1979). *Fifteen thousand hours: secondary schools and their effects in children*. London: Open Books.
- Sadker, D. og E. Silber (2007). *Gender in the classroom. Foundations, skills, methods, and strategies across the curriculum*. Mahwah, N.J.: L. Erlbaum.
- Sammons, P., J. Hillman og P. Mortimore (1995). *Key Characteristics of Effective Schools: A Review of School Effectiveness Research*. London: Office for Standards in Education and Institute of Education.
- Schacter, J. og Y.M. Thum (2004). Paying for high- and low-quality teaching. *Economics of Education Review*, 23, 400-430.
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1(1), 61-80.
- Schulz W. mfl (2010). *ICCS 2009 International report: Civic knowledge, attitudes and engagement among lower secondary students in 38 countries*. Amsterdam: IEA.
- Schwerdt, G. og A.C. Wuppermann (2011). Is traditional teaching really all that bad? A within-student between-student approach. *Economics of Education Review*, 30, 365-379.
- Seland, I., E. Hovdhaugen og N. Vibe (2013). *Evaluering av nasjonale prøver som system*. Rapport 4/2013. Oslo: NIFU.
- Sjøberg, S. (2007). *Naturfag som allmenndannelse - en kritisk fagdidaktikk*. Oslo: Gyldendal Norsk Forlag AS.
- Skaalvik, E.M. og S. Skaalvik (2005). *Skolen som læringsarena. Selvoppfatning, motivasjon og læring*. Oslo: Universitetsforlaget.
- Slavin, R.E., C. Lake og C. Groff (2009). Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis. *Review of Educational Research* 79(2): 839-911.
- Slavin, R.E. mfl (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review* 6(1): 1-26.
- Smith, B. (2000). Quantity matters: Annual instructional time in an urban school system. *Educational Administration Quarterly* 36(5): 652-682.
- Snijders, T. og R. Bosker (2002). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modelling*. London: Sage.
- Steffensen, K. og S.E. Ziade (2009). *Skoleresultater 2008. En kartlegging av karakterer fra grunnskolen og videregående skoler i Norge*. Oslo: Statistisk sentralbyrå.
- Stiefel, L., A.E. Schwartz, R. Rubenstein og J. Zabel (2005). Measuring School Efficiency: What Have We Learned? I Stiefel, L. mfl, *Measuring School Performance and Efficiency: Implications for Practice and Research*, 2005 Yearbook of American Education Finance Association, Eye on Education, New York City, New York: 1-16.

Stiefel, L. mfl (2005). *Measuring School Performance and Efficiency: Implications for Practice and Research*, 2005 Yearbook of American Education Finance Association, Eye on Education, New York City, New York.

Støren, L.A. (2010). *Unge innvandrere i utdanning og overgang til arbeid*. Oslo: NIFU.

Støren, L.A. og H. Helland (2010). Ethnicity differences in the completion rates of upper secondary education: how do the effects of gender and social background variables interplay? *European Sociological Review*, vol. 26, no. 5, s. 585-601.

Telhaug, A. O. (1997). Forty Years of Norwegian Research in the History of Education. *Scandinavian Journal of Educational Research*, 41(3-4).

Utdanningsdirektoratet (2012). *Utdanningsspeilet 2012 – Tall og analyse av grunnsopplæringen i Norge*: http://www.udir.no/Upload/Rapporter/Utdanningsspeilet_2012.pdf?epslanguage=no.

Utdanningsdirektoratet (2013). Skoleporten: <http://skoleporten.udir.no>.

Utdanningsdirektoratet (2013). "Uttalelse publisert elektronisk 17.01.2013: <http://www.udir.no/Stottemeny/Presse/Uttalelser/2013/Fritak-ved-nasjonale-prover-/>".

Vibe, N. mfl (2012). *Strukturer og konjunkturer. Evaluering av Kunnskapsløftet. Sluttrapport fra prosjektet «Tilbudsstruktur, gjennomføring og kompetanseoppnåelse i videregående opplæring»*. Oslo: NIFU.

Wagner, Å.K.H. (2004). *Hvordan leser minoritetsspråklige elever i Norge? En studie av minoritetsspråklige og majoritetsspråklige 10-åringers leseresultater og bakgrunnsfaktorer i den norske delen av PIRLS 2001*. Stavanger: Norsk senter for leseopplæring og leseforskning.

Warrington, M. og M. Younger (2006). *Raising boys' achievement in primary schools. Towards a holistic approach*. Maidenhead: Open University Press.

Wendelborg, C. mfl (2011). *Elevundersøkelsen 2011*. Trondheim: NTNU Samfunnsforskning.

Wiborg, Ø., C.Å. Arnesen, J.B. Grøgaard, L.A. Støren og V. Opheim (2011). *Elevers prestasjonsutvikling - hvor mye betyr skolen og familien? Andre delrapport fra prosjekter "Ressurser og resultater"*. Oslo: NIFU.

Nordisk institutt for studier av
innovasjon, forskning og utdanning

Nordic Institute for Studies in
Innovation, Research and Education

www.nifu.no