



## Skolekvalitet i de videregående skolene i Østfold

En analyse og drøfting av skolebidragsindikatorer og mål på skolekvalitet for skolene i Østfold fylkeskommune i SØF-rapport nr. 1-2016

Eifred Markussen  
Rune Borgan Reiling

Arbeidsnotat 2016:7



# Skolekvalitet i de videregående skolene i Østfold

En analyse og drøfting av skolebidragsindikatorer og mål på skolekvalitet for skolene i Østfold fylkeskommune i SØF-rapport nr. 1-2016

Eifred Markussen  
Rune Borgan Reiling

Arbeidsnotat 2016:7

Arbeidsnotat 2016:7

Utgitt av Nordisk institutt for studier av innovasjon, forskning og utdanning  
Adresse Postboks 2815 Tøyen, 0608 Oslo. Besøksadresse: Økernveien 9, 0653 Oslo.

Prosjektnr. 12820702

Oppdragsgiver Østfold fylkeskommune  
Adresse Postboks 220, 1702 Sarpsborg

Trykk Link Grafisk

Bilddesign Cathrine Årving  
Foto Shutterstock

ISBN 978-82-327-0197-1  
ISSN 1894-8200 (online)

[www.nifu.no](http://www.nifu.no)

---

# Forord

Dette arbeidsnotatet er utarbeidet på oppdrag fra Østfold fylkeskommune. Bakgrunnen for oppdraget er en rapport om skolekvalitet i videregående opplæring som Senter for økonomisk forskning (SØF) presenterte i april 2016. I dette arbeidsnotatet analyseres og drøftes tallene for Østfold og de videregående skole i fylket. Vi drøfter også om målsettingene med rapporten til SØF er nådd. Notatet er utarbeidet av Eifred Markussen og Rune Borgan Reiling. Markussen har hatt hovedansvaret for kapitlene 1, 3 og 4, mens Reiling har hatt hovedansvaret for kapittel 2. Forskningsleder Roger Andre Federici og assisterende direktør Vibeke Opheim har stått for intern kvalitetssikring av notatet.

Oslo, 29.juni 2016

Sveinung Skule  
Direktør

Roger Andre Federici  
Forskningsleder



# Innhold

<b>Skolekvalitet i Østfold: Et sammendrag .....</b>	<b>7</b>
<b>1 En kort introduksjon til SØF-rapporten og arbeidsnotatet .....</b>	<b>11</b>
<b>2 Mål på skolebidrag og skolekvalitet – hvordan kan tallene forstås.....</b>	<b>13</b>
2.1 Beregning av skolebidragsindikatorer og skolekvalitet.....	13
2.2 Forskjeller mellom skoler.....	17
2.2.1 Hva skjer hvis vi setter alle skolebidrag som ikke skiller seg fra gjennomsnittet lik 0?.....	18
2.2.2 Hva skjer hvis vi utelater standpunkt karakterer?.....	19
2.3 Forskjeller mellom fylker.....	20
<b>3 Målsettinger og konsekvenser .....</b>	<b>22</b>
3.1 Målsetting nr. 1: Å etablere et dekkende og gyldig mål på <i>Skolekvalitet</i> .....	22
3.2 Målsetting nr. 2: Å etablere et grunnlag for forbedringsarbeid i videregående skole.....	24
3.2.1 Data om elever som begynte i videregående skole for 7-8 år siden .....	24
3.2.2 Mange skolebidragsindikatorer og mål på skolekvalitet for mange skoler er ikke signifikant forskjellig fra gjennomsnittet.....	25
3.2.3 Yrkesfag og studieforbereidende måles på samme måte .....	26
3.2.4 Elevresultater kan tillegges en skole elevene knapt har gått på .....	26
3.2.5 Alle indikatorene teller like mye .....	26
3.2.6 Formidlingen .....	27
3.2.7 Grunnlag for forbedring: oppsummering .....	27
3.3 Konsekvenser .....	28
<b>4 Anbefalinger .....</b>	<b>30</b>
<b>Vedlegg .....</b>	<b>31</b>
<b>Referanser .....</b>	<b>32</b>





# Skolekvalitet i Østfold: Et sammendrag

I dette notatet har vi sett på tallene for Østfold i rapporten om skolekvalitet i videregående skole fra SØF (Senter for økonomisk forskning). Vi har vist hvordan tallene er beregnet så langt det er mulig på grunnlag av den informasjonen som foreligger i selve rapporten.

Vår analyse av tallene viser at man ikke kan utelukke at Østfold egentlig skårer gjennomsnittlig på skolekvalitet fordi svært mange av skolebidragsindikatorerne og målene på skolekvalitet ikke er signifikant forskjellige fra landsgjennomsnittet.

Vi har foretatt noen alternative beregninger. I SØFs beregninger inngår alle skolebidragsindikatorerne i grunnlaget for målet på skolekvalitet med den beregnede skåren for skolebidrag, også når de ikke er signifikant forskjellig fra landsgjennomsnittet. Men hva hadde skjedd dersom man hadde sagt at de skolebidragsindikatorerne som ikke skiller seg signifikant fra landsgjennomsnittet på indikatoren, ikke skal vektlegges like mye som signifikante skolebidrag i skolekvalitetsmålet? For å teste dette foretok vi en alternativ beregning hvor vi satte alle skolebidragsindikatorer som ikke skiller seg signifikant fra gjennomsnittet lik gjennomsnittet på 0.

En annen mulig endring av SØFs beregningsmåte kunne være å ta ut standpunkt karakterer som en indikator. Et argument for dette er at det er usikkerhet rundt standpunkt karakterer; de er subjektive og det eksisterer ulike karakterregimer (jf. avsnitt 3.2.3). Vi har gjort en slik alternativ beregning.

I begge disse alternative beregningene endrer målene på skolekvalitet seg for de aller fleste skolene, det blir mindre spredning mellom skolene, i de fleste tilfellene blir skåren høyere og rangeringen mellom skolene endres.

Vi har også sammenlignet Østfold med Vestfold som har høyere skåre på målet for skolekvalitet enn Østfold. Også i Vestfold er mange skolebidragsindikatorer og mål på skolekvalitet ikke signifikant forskjellig fra landsgjennomsnittet. Når Vestfold i snitt skårer høyere enn Østfold er én forklaring at alle skolebidragsindikatorer – også de ikke-signifikante – inngår i målet på skolekvalitet med den beregnede skåren for skolebidrag, og disse har høyere verdi i Vestfold enn Østfold. Når alle de ikke-signifikante indikatorene i Vestfold gis gjennomsnittsverdien null slik vi gjorde for Østfold, endres skåren for de fleste skolene og gjennomsnittet for fylket reduseres og blir likt landsgjennomsnittet på 3,1 og litt høyere enn Østfolds 3,0 når ikke-signifikante indikatorer var gitt gjennomsnittsverdien null. Vi ser altså at når de ikke-signifikante skolebidragsindikatorerne gis gjennomsnittsverdien null både i Vestfold og Østfold, er det så godt som ingen forskjell mellom disse to fylkene.

Både ved å ta ut standpunkt karakterer ut av beregningene og ved å sette alle de ikke-signifikante skolebidragsindikatorerne til null (både i Østfold og Vestfold), ser vi at skårene på skolekvalitet endres. Dette viser at det som kan fremstå som objektive mål, viser seg å være basert på subjektive valg. Disse valgene kunne vært andre, og det ville gitt andre resultater. Dette illustrerer at skårene i SØF-rapporten er beheftet med usikkerhet, og at de ikke kan tolkes som den absolutte sannhet om skolekvalitet i videregående skole i Norge.

Slik vi leser rapporten er de to viktigste målsettingene med rapporten a) Å etablere et dekkende og gyldig mål på *Skolekvalitet*, og b) at informasjon om skolekvalitet til skoler og fylkeskommuner kan utgjøre et grunnlag for forbedringsarbeid. Vi har i dette kapitlet drøftet om målsettingene nås.

Når det gjelder den første målsettingen om å etablere et dekkende og gyldig mål på skolekvalitet, er vi helt enige i at mål på gjennomstrømming og faglige prestasjoner er helt sentrale indikatorer. I vår drøfting peker vi imidlertid på at disse indikatorerne ikke er tilstrekkelige når man skal måle skolekvalitet. En rekke andre indikatorer fortjener en plass når skolekvalitet skal måles. Vi peker på en skoles evne til å gi god opplæring til skolefaglig svake elever, omfang av mobbing, grad av trivsel, kvalitet på læringsmiljø samt graden av opplæring i ferdigheter for det 21. århundre. Flere forhold kunne vært nevnt. Fordi mange sentrale kjennetegn ved kvaliteten ved en skole ikke er inkludert, er det vår vurdering at det etablerte målet ikke er et mål på *Skolekvalitet*, men et mål på *Gjennomstrømming og prestasjoner*. Vi argumenterer for å bruke denne korrekte og presise betegnelsen fordi *Skolekvalitet* er noe langt mer enn det som er målt. Og dette er ikke flisespikkeri. Det handler om realiteter, at språk er makt og at å definere akkurat dette som *Skolekvalitet* kan få noen konsekvenser og legge noen føringer på arbeidet i de videregående skole.

Vi har også drøftet målsettingen om at informasjon om skolekvalitet til skoler og fylkeskommuner kan utgjøre et grunnlag for forbedringsarbeid. Vi har drøftet en rekke forhold i relasjon til dette: dataenes alder, spørsmålet om signifikante funn, lik beregningsmåte for yrkesfag og studieforbereidende, spørsmålet om skolesammenslåinger og elevmobilitet, valget om ikke å vekte indikatorerne, at 27 prosent av skolene bare har seks indikatorer, samt at rapporten vurderes som vanskelig tilgjengelig for potensielle brukere i sektoren.

Vår vurdering er at disse forholdene vi har drøftet, bidrar til at rapporten ikke vil fungere som et så godt grunnlag for forbedringsarbeid som målsettingen var. Særlig mener vi at gamle data som ikke tegner et bilde av dagens situasjon, at målet for skolekvalitet ikke er signifikant forskjellig fra gjennomsnittet for halvparten av skolene og at rapporten er vanskelig tilgjengelig for skolefolk flest, svekker rapporten som grunnlag for forbedringsarbeid.

Vi har også drøftet mulige konsekvenser av offentliggjøring av denne rapporten og måten det har vært gjort på. Bl.a. har man fått en unyansert debatt basert på et tall, en *karakter* til skolene. Denne debatten har i mange tilfeller bygd på gale premisser ved at mange ikke har forholdt seg til at halvparten av skolene ikke skiller seg signifikant fra gjennomsnittet, og ved at mange har debattert som om den situasjonen tallene beskriver gjelder dagens situasjon, mens dette handler om situasjonen på skolene for 4-6 år siden. Dette har, slik vi har erfart det, vært uheldig og skapt unødvendig uro og bekymring på skoler og i fylkeskommuner.

En annen sannsynlig konsekvens av rapporten kan være at dette vil bli en rapportering som gjentas med jevne mellomrom. Dersom så skjer, og det dermed slås fast at disse sju indikatorerne defineres som det som til sammen er et mål på kvaliteten ved de videregående skolene i Norge, kan konsekvensen av det være en dreining av skolenes virksomhet og en tilpasning av aktiviteten for å oppnå høyest mulig skåre på det etablerte målet på skolekvalitet. Dette kan være uheldig, og også av den grunn ville det, som vi har poengtert foran, vært bedre om dette samlemålet het *Gjennomføring og prestasjoner* enn *Skolekvalitet*, fordi skolekvalitet er så mye mer enn det som er målt av SØF i denne rapporten.

For di Østfold fremstår som et gjennomsnittsfylke, fordi datagrunnlaget er gammelt, og fordi tallene endrer seg ved noen alternative beregninger, er det vår vurdering at det ikke er grunn til noen strakstiltak innenfor videregående opplæring i Østfold på grunnlag av denne rapporten. Det er viktig å være bevisst forbedrings- og utviklingsarbeidet som har vært satt i verk de siste årene og å holde fast ved dette og ikke stoppe pågående tiltak på grunnlag av SØF-rapporten. Man bør også være særlig bevisst på de resultatene man vet at man har oppnådd de siste årene, samt eksisterende kunnskap om egen skole og fylkeskommune og vurdere resultatene fra SØF-rapporten opp mot dette. På denne måten kan SØF-rapporten inngå som et av mange grunnlag for forbedringsarbeid på skolene i Østfold, men den alene bør ikke danne grunnlag for brå endringer og strakstiltak.



# 1 En kort introduksjon til SØF-rapporten og arbeidsnotatet

I april 2016 offentliggjorde Senter for økonomisk forskning (SØF) rapporten *Skolekvalitet i videregående opplæring. Utarbeidelse av skolebidragsindikatorer og mål på skolekvalitet* (Falch, Bensnes & Strøm, 2016). Rapporten var utarbeidet på oppdrag for Kunnskapsdepartementet.

Rapporten presenterte sju resultatmål som det ble beregnet skolebidragsindikatorer for (tabell 1).

**Tabell 1. Definisjon av resultatmålene som det beregnes skolebidragsindikatorer for**

Resultatmål	Definisjon
<b>Fullføring</b>	Fullført videregående opplæring innen fem år etter avsluttet grunnskole.
<b>Deltakelse</b>	Deltatt i videregående opplæring alle semestre etter avsluttet grunnskole til og med 6. semester. Det inkluderer alle som fullfører på normert tid og alle med normert progresjon.
<b>Normert progresjon</b>	Er i VG3 seks semestre etter avsluttet grunnskole.
<b>Normert fullføring</b>	Fullført videregående opplæring på normert tid. Normert tid er satt til tre år for elever som starter studieforbereende utdanningsprogram, med unntak av idrettsfag hvor normert fullføring er satt til fire år fordi enkelte skoler har fire år som normert tid. For elever som starter på yrkesfaglig utdanningsprogram settes normert fullføring til fire år, med unntak av program som har mer enn fire år som normert tid og medier og kommunikasjon som er satt til tre år.
<b>Standpunktkarakterer</b>	Gjennomsnittlig standpunktkarakter.
<b>Norskeksamen</b>	Gjennomsnittskarakter på avsluttende skriftlig eksamen i norsk hovedmål.
<b>Skriftlig eksamen</b>	Gjennomsnittskarakter på alle skriftlige eksamener.

Kilde: SØF-rapport nr.01/16.

De fire første indikatorene handler om fullføring og gjennomføring av videregående opplæring, og de tre siste handler om faglige prestasjoner målt med karakterer. Det ble beregnet et gjennomsnitt av de sju indikatorene for den enkelte skole, som igjen ble regnet om til et samlemål for skolen med skalaen 1-6. Dette samlemålet ble presentert som et mål på skolekvalitet.

I rapporten fra SØF presenteres:

- Ujusterte resultatmål for hver skole (side 123-132). Dette er skolens faktiske skåre på det enkelte resultatmål. De fire første resultatmålene angir andelen som tilfredsstillende de definerte kriteriene, mens de tre siste målene er på karakterskalaen 1-6. Resultatmålene er på to ulike skalaer, og kan derfor ikke sammenlignes med hverandre.

- Standardiserte skolebidragsindikatorer for hver skole (side 115-122). Her er tallene standardisert slik at tallene for alle skolebidragsindikatorerne er på samme skala, og kan sammenlignes. I beregningen av disse standardiserte skolebidragsindikatorerne har man tatt hensyn til elevenes karakterer fra grunnskolen.
- Standardiserte beregnede mål for skolekvalitet for hver skole (side 143-150). Dette er beregnet på grunnlag av gjennomsnittet av de standardiserte skolebidragsindikatorerne. Målet på skolekvalitet angis på skalaen 1-6.
- Standardiserte beregnede mål for skolekvalitet for hver fylkeskommune (side 37). Dette er beregnet som et gjennomsnitt av målet på skolekvalitet for alle skolene i fylket.

Indikatorerne og målene vil bli grundig presentert i kapittel 2 i dette notatet.

Gjennomsnitt for landet på målet for skolekvalitet var 3,1. Gjennomsnitt for alle skolene i Østfold fylkeskommune var 2,74. En av skolene i Østfold, Halden videregående skole, skåret likt med landsgjennomsnittet. To av skolene, Glemmen og Borg videregående skoler, skåret over landssnittet, mens alle de øvrige skolene skåret under landsgjennomsnittet.

Dette ble omtalt både i Østfoldpressen og i landsdekkende media (Aftenposten). Østfold fylkeskommune og de videregående skolene i fylket fikk negativ omtale og ble fremstilt som et fylke med lav skolekvalitet i videregående (bare de tre nordligste fylkene presterte dårligere). Situasjonen, slik den ble formidlet og oppfattet, vakte bekymring i opplæringsavdelingen i Østfold fylkeskommune. Man ønsket derfor å få en bedre forståelse av rapporten og hva tallene for Østfold og de videregående skolene i fylket faktisk betyr. Opplæringsavdelingen i Østfold henvendte seg til NIFU for å få bistand til dette, og dette arbeidsnotatet er NIFUs svar på oppdraget fra Østfold fylkeskommune.

I notatets kapittel 2 skal vi se nærmere på tallene for Østfold, og beregningene som ligger bak skårene på skolebidragsindikatorerne og samlemålet for skolekvalitet på den enkelte skole ved å:

- vise hvordan alle de sju standardiserte indikatorerne på hver enkelt skole er beregnet på grunnlag av regresjonsanalyser der det kontrolleres for at elevgrunnlaget varierer mellom skoler
- vise hvordan gjennomsnitt for alle skolebidragsindikatorerne er regnet om til en skolekarakter på skalaen 1-6
- kommentere forholdet mellom de sju skolebidragsindikatorerne og samlemålet for skolekvalitet for hver enkelt skole
- sammenligne skolene i Østfold og presentere en tolkning av hva forskjellene betyr, både når det gjelder de sju skolebidragsindikatorerne og samlemålet på skolekvalitet
- sammenligne Østfold med Vestfold fylkeskommune og vise hva som skiller fylkeskommunene fra hverandre, samt å undersøke hvordan de enkelte skolebidragsindikatorerne påvirker skolekarakteren i de to fylkeskommunene

I kapittel 3 drøfter vi om det etablerte samlemålet er et dekkende og gyldig mål på skolekvalitet, vi drøfter valg av indikatorer og om evt. andre indikatorer kunne/burde vært inkludert, samt hvilken betydning valg av indikatorer har for skolekvalitetsmålet. Vi drøfter også andre forhold ved rapporten som kan være verdt å vie oppmerksomhet når funnene skal tolkes og kanskje legges til grunn for tiltak. Til slutt, i kapittel 4 vil vi, i den grad analyser og funn gir grunnlag for det, presentere forslag til tiltak eller strategier for oppfølging av funnene.

## 2 Mål på skolebidrag og skolekvalitet – hvordan kan tallene forstås

Hensikten med rapporten til Falch mfl. (2016) er å bidra til en bedre forståelse av kvalitet i og kvalitetsforskjeller mellom norske videregående skoler. Spørsmålet de ønsker å besvare er om elevprestasjoner avhenger av hvilken videregående skole elevene går på. Dette spørsmålet besvares gjennom å utvikle indikatorer som måler hver enkelt videregående skoles bidrag til syv ulike resultatmål som skal fange opp elevenes læringsutbytte og gjennomstrømning. Videre benyttes indikatorene til å beregne et samlet mål på det de omtaler som kvaliteten på hver enkelt skole.<sup>1</sup>

I denne delen av notatet beskriver vi først hvordan skolebidragsindikatorerne i rapporten til Falch mfl. beregnes og hvordan indikatorene deretter regnes om til en samlet karakter på skalaen 1-6. Videre vil vi, med fokus på Østfold, sammenligne bidragene mellom ulike skoler og presentere en tolkning av hva forskjellene betyr. Til slutt vil vi også sammenligne Østfold med Vestfold og vise hva som skiller de to fylkene fra hverandre, samt undersøke hvordan de enkelte skolebidragsindikatorerne påvirker målet på skolekvalitet på fylkesnivå.

### 2.1 Beregning av skolebidragsindikatorer og skolekvalitet

Vi begynner med å beskrive hvordan de enkelte skolebidragsindikatorerne er beregnet og hvordan disse deretter brukes til å komme frem til et samlet mål på skolekvalitet. Det kan være nyttig å ta utgangspunkt i den enkleste sammenligningen av skoler man kan gjøre, nemlig å sammenligne oppnådde resultatmål (f.eks. fullføringsandelen) ved den enkelte skole.<sup>2</sup> Dette gjøres for skolene i Østfold i tabell 2.

I tabell 2 er tilsynelatende Fredrik II og St. Olav de to skolene i Østfold med høyest kvalitet (andelen elever som fullfører videregående innen fem år er lik 0,83<sup>3</sup>), mens Borg og Kalnes kommer dårligst ut (fullføringsandeler lik henholdsvis 0,55 og 0,50). Utfordringen med denne sammenligningen er at den ikke tar hensyn til at elevgrunnet på Fredrik II og St. Olav er forskjellig fra elevgrunnet på Borg og Kalnes (se siste kolonne i tabell 2, som rapporterer gjennomsnittlig standpunktkarakterer fra grunnskolen). Falch mfl. argumenterer for at faktisk fullføringsandel (og andre resultatmål) er et resultat av både skolens bidrag til fullføring og elevrekrutteringen til skole. For å kunne si noe om skolens bidrag til ulike resultatmål, er det helt nødvendig å skille mellom hva som er et resultat av skolens bidrag og

---

<sup>1</sup> I kapittel 3 drøfter vi om det etablerte samlemålet er et dekkende og gyldig mål på skolekvalitet. Der konkluderer vi at det er for mange sider ved kvalitet i skolen som ikke inngår i skolekvalitetsmålet. I dette kapitlet bruker vi likevel begrepet *Skolekvalitet* slik Falch mfl. har definert det.

<sup>2</sup> Se tabell 1 i kapittel 1 for en oversikt over resultatmålene som det beregnes skolebidragsindikatorer for.

<sup>3</sup> Dette betyr at 83 prosent fullførte etter fem år.

hva som er et resultat av elevrekrutteringen. For å finne skolens bidrag til resultatmålene, må man justere resultatmålene for sistnevnte.

**Tabell 2: Fullføring ved de videregående skolene i Østfold.**

Skole	Andelen elever som fullfører videregående innen fem år etter oppstart (0,83=83%)	Avvik fra landsgjennomsnittet	Gjennomsnittlig standpunktkarakterer fra grunnskolen
Fredrik II	0,83	0,09	4,35
St. Olav, Sarpsborg	0,83	0,09	4,29
Kirkeparken	0,81	0,07	4,42
Askim	0,71	-0,03	4,00
Halden	0,69	-0,05	3,86
Greåker	0,66	-0,08	3,91
Mysen	0,65	-0,09	3,93
Glemmen	0,64	-0,10	3,73
Malakoff	0,58	-0,16	3,62
Borg	0,55	-0,19	3,30
Kalnes	0,50	-0,24	3,35
Landsgjennomsnittet	0,74	0,00	4,25

Kilde: SØF-rapport nr.01/16.

Den ideelle måten å gjøre dette på er å gjennomføre et kontrollert eksperiment der man fordeler elever på en tilfeldig måte på videregående skoler, og deretter observerer hva som skjer med resultatmålene for disse elevene. Forskjeller i oppnådde resultater vil da reflektere forskjeller i det «korrekte» skolebidraget i den forstand at man har tatt hensyn til forskjeller i elevgrunnet. Dette eksperimentet er *ikke* praktisk gjennomførbart i Norge.<sup>4</sup>

Derimot mener Falch mfl. (2016, s. 7) at norske registerdata inneholder så presis informasjon om elevenes kunnskapsnivå, målt med karakterer på ungdomskolen, at man kan beregne relativt «korrekte» skolebidrag ved å bruke andre tilnærminger enn en rent eksperimentell tilnærming. Mer spesifikt benytter Falch mfl. *multippel regresjonsanalyse* for å estimere skolens bidrag til elevers resultatoppnåelse på videregående. En *multippel regresjonsanalyse* betyr at man beregner forholdet mellom en *avhengig variabel* (f.eks. fullføring, deltakelse eller skriftlig eksamenskarakter) og en *interessevariabel* (i dette tilfellet: skolen en elev går på) ved å benytte en beregningsmodell der man tar hensyn til potensielle sammenhenger mellom *andre variabler* (f.eks. elevgrunnet ved skolen) og den *avhengige variabelen*.<sup>5</sup> På denne måten tar man hensyn til at elevgrunnet varierer mellom skoler. Det beregnede forholdet mellom *avhengig variabel* (f.eks. fullføring, deltakelse eller skriftlig eksamen) og *interessevariabelen* (skolen en elev går på) i modellene til Falch mfl. (2016) kan tolkes som skolens bidrag til de ulike resultatmålene, og vil uttrykke forskjellen i resultat mellom videregående skoler for elever med samme karakternivå fra grunnskolen.

Resultatene av regresjonsanalysene til Falch mfl. er presentert i tabell A5.1.2. i SØF-rapporten. Resultatene fra analysene brukes videre til å beregne skolebidragsindikatorerne for hver enkelt skole, og disse indikatorerne blir presentert på *standardisert form* for hver enkelt skole (se tabell A11.2 i rapporten.). *Standardisering* er nødvendig for å kunne sette sammen et samlemål på skolekvalitet (se s. 53 i rapporten),<sup>6</sup> og innebærer at hver av de syv indikatorerne gjøres om slik at de får samme

<sup>4</sup> Angrist mfl. (2015) kommer nærme dette eksperiment ved å utnytte at noen elever blir plassert på skoler etter loddrekning i enkelte delstater i USA.

<sup>5</sup> I prinsippet skal man kontrollere for alle variabler som potensielt kan påvirke den avhengige variabelen. Falch mfl. (2016) beregner også modeller der de tar hensyn til en rekke slike forhold, men konkluderer med at modellen der de kun justerer for elevgrunnet for skolen fungerer best.

<sup>6</sup> Grunnen til dette er at de ulike resultatmålene i utgangspunktet følger ulike skalaer. For eksempel kan fullføring kun ha verdiene 0 og 1 (0=ikke fullført, 1=fullført), mens standpunkt- og eksamenskarakter kan ha verdier fra og med 1 til og med 6.



gjennomsnitt og standardavvik<sup>7</sup>. Dette gjøres ved å benytte en standard formel (som vi kommer tilbake til). I tillegg til at standardisering er nødvendig for å kunne sette sammen et samlemål på skolekvalitet, medfører det også at det blir enklere å sammenligne på tvers av de enkelte resultatmålene. Dette kan illustreres ved hjelp av følgende eksempel.

**Eksempel:** Vi kan tenke oss en ung friidrettsutøver som ønsker å satse seriøst på friidrett, og som må bestemme seg for hvilken friidrettsøvelse hun skal spesialisere seg i. Hun er ganske god i både 400 meter løping (personlig rekord 61,20 sekunder) og høydehopp (personlig rekord 1,35 meter), men vet ikke hvor god hun er sammenlignet med de beste nasjonale utøverne i disse øvelsene. Hun ønsker å spesialisere seg i den øvelsen hun er nærmest de beste i. Gjennomsnittstiden for de beste nasjonale utøverne på 400 meter er 60 sekunder, mens gjennomsnittshøyden for de beste i høyde er 1,50 meter. Hun er altså 1,2 sekunder bak de beste på 400 meter, mens hun er 15 cm unna de beste i høyde. Utfordringen er å finne ut om 1,2 sekunder bak de beste på 400 meter er bedre sammenlignet med å være 15 cm i unna de beste i høyde.

Det er nødvendig å gjøre friidrettsutøverens avvik fra gjennomsnittet sammenlignbare, altså at 1,2 sekunder gjøres sammenlignbart med 15 cm. Dette gjøres ved å standardisere og kan enkelt gjøres ved å benytte følgende matematiske formel:

$$(1) Z = \frac{X - \text{gjennomsnittet}}{\text{Standardavvik}}$$

X er i dette tilfellet friidrettsutøverens prestasjoner på friidrettsbanen, henholdsvis 61,2 sekunder på 400 meter og 1,35 meter i høyde. Mens gjennomsnittet henholdsvis er gjennomsnittstiden for de beste nasjonale utøverne på 400 meter og gjennomsnittshøyden for de beste nasjonale høydehopperne. La oss si at standardavviket for 400-meterløperne er 3, mens standardavviket for høydehopperne er 0,15 (spredningen blant 400-meterløpere er altså større enn blant høydehopperne). Dermed har vi all informasjon vi trenger for å standardisere, og vi kan sette inn i formelen:

$$Z_{400 \text{ meter}} = \frac{61,2 - 60}{3} = \frac{1,2}{3} = 0,4$$

$$Z_{\text{Høyde}} = \frac{1,35 - 1,50}{0,15} = \frac{-0,15}{0,15} = -1$$

Dette betyr at de standardiserte prestasjonene til friidrettsutøveren er lik 0,4 på 400 meter og -1 i høydehopp. Disse tallene kan sammenlignes. Hun er ikke like god som de beste utøverne i noen av øvelsene. Men hun er kun 0,4 standardavvik unna gjennomsnittet på 400 meter, men 1 standardavvik under gjennomsnittet i høyde. Hun er altså nærmere de beste i løping enn i høyde.

Dette eksempelet beskriver nøyaktig det samme som Falch mfl. har benyttet for å beregne de standardiserte skolebidragsindikatorerne som presenteres i tabell A11.2 i deres rapport. De standardiserte skolebidragsindikatorerne kan tolkes som avvik fra et landsgjennomsnitt målt i standardavvik. Vi presenterer de standardiserte skolebidragsindikatorerne for skolene i Østfold i slik de er presentert av Falch mfl. i vår tabell 3 (kolonne 1 til 7).

---

<sup>7</sup> Standardavvik er et mål på hvordan hver enkelt videregående skole bidrag til ulike resultatmål plasserer seg rundt det gjennomsnittlige skolebidrag, altså et mål på spredning. Desto større spredning, desto større standardavvik. **For å illustrere dette:** Tenk dere at vi ser på gjennomsnittsalderen for hele i Norges befolkning, og er interessert i å se på hvordan befolkningen fordeler seg aldersmessig rundt dette gjennomsnittet. Her vil vi få stor spredning (og dermed et stort standardavvik) fordi befolkningen vil fordele seg over en stor skala mellom 0 og 110 år. Tenk dere deretter at vi ser på ser på gjennomsnittsalderen for videregående elever i Norge, og er interessert i å se hvordan elevene fordeler seg aldersmessig rundt dette gjennomsnittet. Her vil vi få liten spredning (og dermed et lite standardavvik) fordi elevene fordeler seg på en mindre skala mellom 16 og 20 år. For å kunne sammenligne avvik fra gjennomsnittet for de to fordelingene (befolkningens alder og elevens alder) er det nødvendig å standardisere fordelingene, slik at de blir like.

**Tabell 3: Standardiserte skolebidragsindikatorer. Skolene i Østfold.**

Skole	(1) Full føring	(2) Del- tag- else	(3) Nor- mert progre- sjon	(4) Normert full- føring	(5) Stand- punkt- karak- terer	(6) Norsk- eks- amen	(7) Skriftlig eks- amen	(8) Gjennom- snitt	(9) Standard gjennom- snitt	(10) Skole- kvalitet
Fredrik II	0	0.9	0.32	-0.21	-1.83	-0.27	-0.62	-0.24	-0.35	2,7
St. Olav	0.37	0.52	0.68	0.68	-1.91	-0.12	-0.35	-0.02	-0.03	3,0
Kirkeparken	-0.48	-0.24	-0.09	-0.54	-1.29	0.22	-0.54	-0.42	-0.60	2,5
Askim	-0.26	-1.3	-0.73	-0.05	-1.46	-0.58	-0.34	-0.67	-0.95	2,1
Halden	0.21	0.54	1.12	0.49	-0.41	-0.57	-1.05	0.05	0.06	3,1
Greåker	-0.56	-0.54	0.11	-0.67	-0.61	-0.56	-0.64	-0.50	-0.70	2,4
Mysen	-0.86	-1.03	-0.37	-0.52	-0.88	-0.02	-0.63	-0.62	-0.87	2,2
Glemmen	-0.29	-0.32	-0.08	0.12	0.2	-0.43	1.47	0.10	0.13	3,2
Malakoff	-0.6	-0.47	-0.09	-0.47	0.04	-0.17	-0.34	-0.30	-0.43	2,6
Borg	0.28	0.42	0.59	0.51	0.43		-0.78	0.24	0.33	3,4
Kalnes	-0.6	-0.92	0.26	-0.22	0.90		-2.04	-0.44	-0.65	2,4
Gjennomsnitt Østfold	-	-	-	-	-	-	-	-	-	2,7
Lands- gjennomsnitt	0	0	0	0	0	0	0	0	0	3,1
Standardavvik	1	1	1	1	1	1	1	0,71	1	0,97

Kilde: SØF-rapport nr.01/16.

I kolonne (10) i tabell 3 er det beregnede målet på skolekvalitet rapportert. Dette målet beregnes i tre steg. Først beregnes gjennomsnittet av alle bidragsindikatorerne for alle skoler i hele landet. Gjennomsnittet for hver enkelt skole i Østfold rapporteres i kolonne (8). Deretter standardiseres også dette gjennomsnittet ved å benytte formel (1) over. Vi illustrerer hvordan dette gjøres i praksis i eksempelet for Fredrik II under:

**Eksempel:** For å beregne det standardiserte gjennomsnittet av skolebidragene til Fredrik II setter vi inn -0,24 for X, 0 for gjennomsnittet og 0,71 for standardavviket i formel (1).<sup>8</sup> Altså:

$$(1) Z = \frac{X - \text{gjennomsnittet}}{\text{Standardavvik}} = \frac{-0,24 - 0}{0,71} = \frac{-0,24}{0,71} = -0,35$$

Til slutt gjøres disse standardiserte gjennomsnittene om til en skala fra 1 til 6 (altså samme karakterskala som benyttes i videregående skole) der gjennomsnittet er lik 3,1 og standardavviket er lik 0,97 (altså det samme gjennomsnittet og standardavviket som for skriftlig eksamen).<sup>9</sup> Dette er ikke et nødvendig grep av Falch mfl., men snarere et grep for å fremstille målet på skolekvalitet på en pedagogisk måte. Siden de standardiserte gjennomsnittene også kan tolkes som avvik fra landsgjennomsnittet målt i standardavvik kan vi regne ut målet på skolekvalitet ved å legge den standardiserte summen (multiplisert med 0,97) til 3,1. Se eksempelet for St. Olav under. Dermed får vi målet på skolekvalitet som er rapportert i kolonne (10).

**Eksempel:** Hvis vi for eksempel skal beregne målet på skolekvalitet for St. Olav, gjør vi følgende beregning:<sup>10</sup>

$$(1) \text{ Skolekvalitet (St. Olav)} = 3,1 + (\text{standardisert sum (St. Olav)} \times 0,97) = 3,1 + (-0,03 \times 0,97) = 3,1 - 0,03 = 3,08$$

Dette eksempelet kan gjentas for alle skolene i Østfold, for å finne målet på skolekvalitet for hver enkelt skole. Etter at vi nå har vist hvordan målet på skolekvalitet for den enkelte skole beregnes, skal vi fortsette med å se på forskjeller mellom skolene i Østfold.

<sup>8</sup> Disse tallene er gjennomsnittet og standardavviket til gjennomsnittet som er rapportert i kolonne (8). Disse er beregnet basert på alle videregående skoler i landet.

<sup>9</sup> Gjennomsnittet og standardavviket til skriftlig eksamen er hentet fra tabell A3.2.1 i rapporten til Falch mfl.

<sup>10</sup> Tallene som rapporteres i dette notatet avviker noe fra tallene til Falch mfl. på grunn av ulik avrunding på et eller annet nivå.

## 2.2 Forskjeller mellom skoler

Det er ikke helt enkelt å sammenligne skoler basert på de standardiserte skolebidragene, og forklare hva tallene faktisk betyr, uten at vi har tilgang på rådataene som ligger bak tallene. Vi kan imidlertid forklare hvordan man skal forstå tallene som rapporteres i tabell 3. Tall som er større enn null kan tolkes som at skolens bidrag til et gitt resultatmål (f.eks. fullføring, deltakelse eller skriftlig eksamen) ved denne skolen er større enn landsgjennomsnittet. Desto større tallet er, desto større vil skolens bidrag til resultatmålet (f.eks. fullføring, deltakelse eller skriftlig eksamen) være. På samme måte kan tall som er mindre enn null tolkes som at skolebidraget til et gitt resultatmål (f.eks. fullføring, deltakelse eller skriftlig eksamen) er mindre enn landsgjennomsnittet. Desto lenger unna null dette tallet er, desto lengre unna landsgjennomsnittet vil bidraget være. Igjen illustrerer vi dette ved å benytte et eksempel.

**Eksempel:** Se på kolonne (1) i tabell 3. De beregnede skolebidragene for St. Olav, Halden og Borg til fullføring er henholdsvis 0,37, 0,21 og 0,28, mens bidragene til fullføring for de andre skolene i Østfold er mindre enn null. En elev som begynner på St. Olav, Halden eller Borg vil altså ha større sannsynlighet for å fullføre videregående opplæring sammenlignet med om denne hadde begynt på gjennomsnittsskolen i Norge, mens eleven ville hatt lavere sannsynlighet for å fullføre videregående opplæring dersom den hadde begynt ved en av de andre skolene i Østfold sammenlignet med om eleven hadde begynt på en gjennomsnittsskole. Videre indikerer tallene at det positive skolebidraget til fullføring er størst ved St. Olav og minst ved Halden. Og på samme måte er det negative skolebidraget til fullføring størst ved Mysen og minst ved Askim.

Tallene i kolonne (2) - (7) kan tolkes på akkurat samme måte. Tall som er større enn null, indikerer at skolen løfter en gitt elev sammenlignet med om eleven hadde gått på gjennomsnittsskolen. Tall som er mindre enn null, indikerer at en gitt elev gjør det dårligere på denne skolen sammenlignet med om eleven hadde gått på gjennomsnittsskolen.

Det er viktig å merke seg at skolebidragene som presenteres i rapporten til Falch mfl. er *beregnet* (se side 12 i dette notatet), og at det derfor vil være usikkerhet knyttet til om man har beregnet «korrekte» skolebidrag. Usikkerheten skyldes at enkeltelever med samme karakternivå fra grunnskolen som går på samme videregående, oppnår noe ulike resultater. I rapporten til Falch mfl. er man opptatt av om de *beregnete* skolebidragene med tilstrekkelig sannsynlighet kan sies å skille seg fra det gjennomsnittlige skolebidraget, eller, mer teknisk, om bidragene skiller seg statistisk signifikant fra gjennomsnittet. Rapporten til Falch mfl. viser at de aller fleste skolene *ikke* skiller seg signifikant fra gjennomsnittet når det gjelder skolebidragsindikatorene. Dette betyr altså at man *ikke* med tilstrekkelig sannsynlighet kan si at de *estimerte* skolebidragene skiller seg fra det gjennomsnittlige skolebidraget. Dette gjelder for mellom 53 og 88 prosent av skolene, avhengig av hvilken indikator man ser på.

Gul celle i tabell 2 indikerer at skolebidraget i denne cellen ikke kan sies å være signifikant forskjellig fra landsgjennomsnittet. Gul celle betyr altså at man *ikke* (med stor nok sannsynlighet) kan si at disse skolebidragene er forskjellige fra bidraget til den gjennomsnittlige norske videregående skolen. Ingen av skolene i Østfold skiller seg fra gjennomsnittet når det gjelder fullføring, deltagelse, normert fullføring og norsk eksamen, altså fire av de syv resultatmålene som Falch mfl. benytter for å beregne det samlede skolekvalitetsmålet.<sup>11</sup> Halden er den eneste skolen i Østfold som kan sies å være forskjellig fra gjennomsnittet når det gjelder normert progresjon, mens Halden og Glemmen er de to eneste skolene der bidraget til skriftlig eksamen skiller seg fra landsgjennomsnittet. For standpunkt karakterer er det fire skoler som skiller seg fra gjennomsnittet (Fredrik II, St. Olav, Kirkeparken og Askim). Ved Greåker, Mysen og Malakoff kan ingen av skolebidragene sies å være forskjellige fra landsgjennomsnittet. Oppsummert kan vi si at skolene i Østfold i liten grad skiller seg fra en gjennomsnittlig norsk videregående skole når det gjelder skolebidragsindikatorene. Bare åtte av 75 indikatorer skiller seg signifikant fra landsgjennomsnittet på null. Dette er viktig å ha i bakhodet når man tolker resultatene for skolene i Østfold. Videre i denne delen av notatet, gjør vi to alternative beregninger som vi håper gir en sterkere forståelse av hvordan man kan lese resultatene i rapporten til Falch mfl. (2016). I tillegg gjør vi

<sup>11</sup> Merk at man på to av skolene hadde man ikke elever som hadde norskeksamen på Vg3 studieforbereende.

disse alternative beregningene for å tydeliggjøre at målene på skolekvalitet er beregnet, og derfor ikke må tolkes som den absolutte sannheten om kvalitet i norsk videregående skole.

### **2.2.1 Hva skjer hvis vi setter alle skolebidrag som ikke skiller seg fra gjennomsnittet lik 0?**

I beregningen av målet på skolekvalitet blir alle skolebidragsindikatorerne tillagt like stor betydning, uavhengig av om bidraget er signifikant forskjellig fra gjennomsnittet eller ikke. For eksempel betyr dette at målet på skolekvalitet ved Askim videregående påvirkes like mye av det skolebidraget til deltagelse som skolebidraget til standpunkt karakterer, selv om det bidraget til deltagelse ikke kan sies å skille seg fra gjennomsnittet. Samtidig ender Greåker, Malakoff og Mysen opp med mål på skolekvalitet som ligger signifikant under landsgjennomsnittet selv om ingen av deres skolebidragsindikatorer kan sies å være signifikant forskjellig fra landsgjennomsnittet. Dette kan skje fordi usikkerheten i skolekvalitetsmålet er beregnet på en annen måte enn usikkerheten i skolebidragsindikatorerne. Enkelt forklart har Falch mfl. anslått usikkerheten i skolekvalitetsmålet ved å variere betydningen av de syv underliggende skolebidragsindikatorerne, og deretter beregnet skolekvalitetsmålet på 1000 ulike måter. Variasjonen som da oppstår brukes for å beskrive usikkerheten i skolekvalitetsmålet, denne variasjonen kalles konfidensintervallet. Et eksempel: Når konfidensintervallet for Borg er 3,0-3,7 (se tabell A12-1 i SØF-rapporten), betyr det at skolekvalitetsmålet med en viss sikkerhet vil ligge innenfor dette intervallet. 3,4 er det beste anslaget, men hele intervallet er sannsynlig. Konfidensintervallet for skolekvalitetsmålet ved hver enkelt skole er presenter i tabell A12-1 i rapporten til Falch mfl. Dersom landsgjennomsnittet for skolekvalitet faller inn i konfidensintervallet sier man at skolekvalitetsmålet ikke kan sies å være forskjellig fra landsgjennomsnittet. I det nettopp nevnte eksemplet faller landsgjennomsnittet på 3,1 innenfor konfidensintervallet, og skolekvalitetsmålet for Borg skiller seg dermed ikke signifikant fra landsgjennomsnittet.

Men hva hadde skjedd dersom man hadde sagt at de skolebidragsindikatorerne som ikke skiller seg signifikant fra landsgjennomsnittet på indikatoren, ikke skal vektlegges like mye som signifikante skolebidrag i skolekvalitetsmålet? For å teste dette har vi i tabell 4 har satt alle skolebidragsindikatorer som ikke skiller seg signifikant fra gjennomsnittet lik gjennomsnittet på 0. Sammenlignet med de målene som rapporteres i tabell 3 og i rapporten til Falch mfl., fører dette til at målet på skolekvalitet går ned for St. Olav og Borg mens målet for de øvrige skolene går opp. Det blir vesentlig mindre spredning mellom skolene, og skolekvaliteten for fylket samlet ligger nå rett i underkant av landsgjennomsnittet.

Glemmen er nå den skolen som kommer best ut, mens Fredrik II og St. Olav er de skolene som kommer dårligst ut. Dette skyldes i stor grad at Glemmen kommer godt ut på skriftlig eksamen, mens Fredrik II og St. Olav kommer dårlig ut på standpunkt karakterer. Askim og Kirkeparken gjør det dårligere enn landsgjennomsnittet i tabell 4, men kommer faktisk bedre ut her i forhold til i tabell 2. Dette skyldes at mange av skolebidragsindikatorerne for disse skolene er relativt store, uten å være statistisk signifikante. Når disse settes til null gir det store utslag på målet for skolekvalitet.

Halden ligger rett over landsgjennomsnittet på skolekvalitet. Dette kan forklares fordi de gjøre det bra på normert progresjon, men tilsvarende dårlig på skriftlig eksamen. Kalnes ligger under landsgjennomsnittet. De gjør det bedre enn gjennomsnittet på standpunkt karakterer, men vesentlig dårligere enn landsgjennomsnittet på skriftlig eksamen.

Resultatene som er presentert i tabell 4 indikerer at det å la skolebidragsindikatorerne bidra like mye uavhengig av om man med tilstrekkelig grad av sikkerhet kan si at de skiller seg fra gjennomsnittet, har stor betydning for det samlede skolekvalitetsmålet (i hvert fall for enkelte skoler).

**Tabell 4: Skolekvalitet når alle skolebidrag som ikke er signifikant forskjellig fra gjennomsnittet settes lik null. Skolene i Østfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunkt-karakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet	Avvik fra Skolekvalitet i tabell 3
Fredrik II	0	0	0	0	-1.83	0	0	2,7	0
St. Olav	0	0	0	0	-1.91	0	0	2,7	-0,3
Kirkeparken	0	0	0	0	-1.29	0	0	2,8	+0,3
Askim	0	0	0	0	-1.46	0	0	2,8	+0,7
Halden	0	0	1.12	0	0	0	-1.05	3,2	+0,1
Greåker	0	0	0	0	0	0	0	3,1	+0,7
Mysen	0	0	0	0	0	0	0	3,1	+0,9
Glemmen	0	0	0	0	0	0	1.47	3,5	+0,3
Malakoff	0	0	0	0	0	0	0	3,1	+0,5
Borg	0	0	0	0	0		0	3,1	-0,3
Kalnes	0	0	0	0	0,9		-2,04	2,9	+0,5
Gjennomsnitt Østfold	-	-	-	-	-	-	-	3,0	
Landsgjennomsnitt	0	0	0	0	0	0	0	3,1	
Standardavvik	1	1	1	1	1	1	1	0,97	

Kilde: SØF-rapport nr. 01/16

### 2.2.2 Hva skjer hvis vi utelater standpunkt-karakterer?

Videre er det viktig å ha i mente at det beregnede skolekvalitetsmålet også er et resultat av hva slags indikatorer Falch mfl. har valgt å ta med (jf. også drøftingene i avsnitt 3.1). Det er derfor interessant å se på hva som skjer med skolekvalitetsmålene dersom vi utelater et av de syv resultatmålene som det beregnes skolebidragsindikatorer for. Basert på tabell 3, ser det ut til at skolen bidrag til standpunkt-karakterer er den viktigste årsaken til at skolekvalitetsmålet blir lavere enn gjennomsnittet for Fredrik II, St. Olav, Kirkeparken og Askim. Hvis vi ser på tabell 2, er det disse skolene som kommer best ut på fullføringsmålet når man ikke justerer for ulikheter i elevgrunnlaget. Vi velger derfor å undersøke hva som skjer med skolekvalitetsmålene når vi utelater standpunkt-karakterer fra analysen. Et argument for å utelate standpunkt-karakterene som mål på skolekvalitet kan være at dette målet kan sies å være relativt subjektiv sammenlignet med de andre resultatmålene, både fordi de sette av ulike lærere på ulike skoler og fordi karakterregimene varierer mellom linjer på videregående (jf. avsnitt 3.2.3).<sup>12</sup>

I tabell 5 har vi fjernet standpunkt-karakter som skolebidragsindikator. Det fører til at målet på skolekvalitet øker betydelig for Fredrik II og St. Olav og mer moderat for Kirkeparken, Askim og Halden. For de øvrige skolene er målet på skolekvalitet uforandret eller marginalt mindre. Å utelate standpunkt-karakterer som skolebidragsindikator, innebærer også at St. Olav nå er den skolen som kommer best ut i Østfold, men Askim, Mysen og Kalnes er de skolene som kommer dårligst ut. Fredrik II får en skolekvalitet tilsvarende landsgjennomsnittet, mens Halden, Glemmen og Borg skårer over landsgjennomsnittet. Kirkeparken, Greåker og Malakoff kommer dårligere ut enn landsgjennomsnittet. Kirkeparken har marginalt høyere kvalitet enn Malakoff.

Dette indikerer at valg av hvilke resultatmål det skal lages skolebidragsindikatorer for, kan få stor betydning for hvordan enkelt skoler kommer ut med tanke på skolekvalitetsmålet. Dette er det viktig å være oppmerksom på, når man tolker og drøfter tallene, og særlig om resultatene i rapporten skal ligge til grunn for tiltak eller utviklingsarbeid i fylkeskommuner eller på enkelt-skoler.

<sup>12</sup> Vi har også gjort to alternative analyser hvor vi har utelatt norskeksamen og skriftlig eksamen fra analysen. De beregnede målene på skolekvalitet man da får er rapportert i henholdsvis tabell V1 og V2 i vedlegg. Å ta ut norskeksamen fra analyse har liten betydning for det samlede målet på skolekvalitet. Å ta ut skriftlig eksamen fører til at Halden og Kalnes kommer bedre ut med tanke på skolekvalitet sammenlignet med tabell 3, mens de andre skolene i Østfold oppnår omtrent samme resultat.

**Tabell 5: Skolekvalitet når standpunktkarakterer er utelatt. Skolene i Østfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunktkarakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet	Avvik fra skolekvalitet i tabell 3
Fredrik II	0	0.9	0.32	-0.21		-0.27	-0.62	3,1	+0,4
St. Olav	0.37	0.52	0.68	0.68		-0.12	-0.35	3,5	+0,5
Kirkeparken	-0.48	-0.24	-0.09	-0.54		0.22	-0.54	2,7	+0,2
Askim	-0.26	-1.3	-0.73	-0.05		-0.58	-0.34	2,4	+0,3
Halden	0.21	0.54	1.12	0.49		-0.57	-1.05	3,3	+0,2
Greåker	-0.56	-0.54	0.11	-0.67		-0.56	-0.64	2,5	+0,1
Mysen	-0.86	-1.03	-0.37	-0.52		-0.02	-0.63	2,4	+0,2
Glemmen	-0.29	-0.32	-0.08	0.12		-0.43	1.47	3,2	0
Malakoff	-0.6	-0.47	-0.09	-0.47		-0.17	-0.34	2,6	0
Borg	0.28	0.42	0.59	0.51			-0.78	3,3	-0,1
Kalnes	-0.6	-0.92	0.26	-0.22			-2.04	2,2	-0,2
Gjennomsnitt Østfold	-	-	-	-	-	-	-	2,8	
Landsgjennomsnitt	0	0	0	0	0	0	0	3,1	
Standardavvik	1	1	1	1	1	1	1	0,97	

Kilde: SØF-rapport nr. 01/16

## 2.3 Forskjeller mellom fylker

Videre er det interessant å undersøke hva som gjør at fylkene skiller seg fra hverandre, samt å undersøke hvordan de enkelte skolebidragsindikatorerne påvirker målet for skolekvalitet i de ulike fylkene. For enkelhets skyld, har vi her bare valgt å sammenligne Østfold og Vestfold. Begrunnelsen for å sammenligne med Vestfold er at dette fylket har relativt like andeler fullført og bestått etter fem år som Østfold for kullene som begynte i videregående skole i 2008 og 2009<sup>13</sup>, mens de på skolekvalitetsmålet skårer 3,1, som er vesentlig bedre enn Østfold. Hvorfor er det slik?

I tabell 6 presenterer vi de standardiserte skolebidragsindikatorerne samt målet på skolekvalitet for skolene i Vestfold. Vi ser at også for Vestfold er det stor andel av skolebidragsindikatorerne (86 prosent) og skolekvalitetsmålene (8 av t0) som ikke skiller seg signifikant fra landsgjennomsnittet. Vestfold fremstår som Østfold som et gjennomsnittsfylke. Likevel kommer de ut med en betydelig bedre skåre på SØFs mål for Skolekvalitet. Med utgangspunkt i beregningene som er gjort, gjør vi i det følgende et forsøk på å forklare forskjellen mellom de to fylkene.

**Tabell 6: Standardiserte skolebidragsindikatorer. Skolene i Vestfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunktkarakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet
Færder	-0.27	-0.41	-0.21	-0.26	0.61	0.74	0.25	3,1
Greveskogen	-0.35	0.11	0.36	-0.33	-0.09	1.22	0.51	3,3
Holmestrand	-0.2	0.46	0.73	-0.22	0.13	-0.53	-0.89	3,0
Horten	0.33	0.62	0.79	0.61	0.04	-0.40	-0.24	3,4
Melsom	-1.69	-0.86	-0.12	-1.33	0.42		-0.22	2,2
Nøtterøy	0.79	1.38	1.39	0.45	0.67	1.49	0.43	4,3
Re	1.36	1.21	1.65	0.28	0.82	-0.92	-0.61	3,8
Sande	0.04	-0.44	0.13	0.41	-0.62	-0.31	-0.28	2,8
Sandefjord	-0.22	0.33	0.15	-0.23	-0.46	0.05	0.03	3,0
Thor Heyerdahl	-0.24	-0.08	0.55	0.1	0.18	-0.57	-0.83	2,9
Gjennomsnitt Vestfold	-	-	-	-	-	-	-	3,3
Landsgjennomsnitt	0	0	0	0	0	0	0	3,1
Standardavvik	1	1	1	1	1	1	1	0,97

Kilde: SØF-rapport nr. 01/16

I tabell 4 så vi at det å la skolebidragsindikatorerne bidra like mye uavhengig av om man kan si med sikkerhet at de skiller seg fra gjennomsnittet, har stor betydning for det samlede skolekvalitetsmålet for

<sup>13</sup> Bestått etter fem år. Østfold 2008: 66%, 2009: 66%. Vestfold 2008: 69%, 2009: 68%.

hver enkelt skole i Østfold. Vi så også at skolene nærmet seg hverandre, og landsgjennomsnittet når vi satte alle bidrag som ikke kunne sies å skille seg fra landsgjennomsnittet lik null. Det er interessant å undersøke om dette også er tilfellet i andre fylker, og vi sammenligner altså med Vestfold her.

**Tabell 7: Skolekvalitet når alle skolebidrag som ikke er signifikant forskjellig fra gjennomsnittet settes lik null. Skolene i Vestfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunkt-karakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet
Færder	0	0	0	0	0	0	0	3,1
Greveskogen	0	0	0	0	0	0	0	3,1
Holmestrand	0	0	0	0	0	0	-0,89	2,9
Horten	0	0	0	0	0	-0,4	0	3,0
Melsom	-1,69	0	0	-1,33	0		0	2,4
Nøtterøy	0	0	1,39	0	0	1,49	0	3,8
Re	1,36	0	1,65	0	0,82	0	0	4,1
Sande	0	0	0	0	0	0	0	3,1
Sandefjord	0	0	0	0	0	0	0	3,1
Thor Heyerdahl	0	0	0	0	0	0	0	3,1
Gjennomsnitt Vestfold								3,1
Landsgjennomsnitt	0	0	0	0	0	0	0	3,1

Kilde: SØF-rapport nr. 01/16

I tabell 7 har vi derfor satt alle skolebidragsindikatorer som ikke skiller seg signifikant fra gjennomsnittet i Vestfold lik 0. Som for Østfold beveger alle skolene seg mot landsgjennomsnittet, det blir altså mindre spredning. Fylkets gjennomsnittlige skolekvalitet blir helt lik landsgjennomsnittet. Dette indikerer at selv om skolene i Vestfold skårer bedre enn skolene i Østfold når vi fokuserer på samlet skolekvalitet i henholdsvis tabell 6 og 3, så er det de mange ikke-signifikante bidrag i de to fylkene som bidrar til at Vestfold skårer bedre enn Østfold. Når vi setter alle disse bidragene lik null, kommer de to fylkene likt ut. Igjen får vi illustrert at når ikke-signifikante skolebidragsindikatorer teller med i skolekvalitetsmålet, har det stor betydning for den målte skolekvaliteten.

De alternative beregningene viser at de endelige målene på skolekvalitet er beheftet med usikkerhet, og påvirkes av valg av indikatorer. Bruk av andre kvalitetsindikatorer ville gitt kanskje gitt helt andre resultater enn de som er presentert i rapporten til Falch mfl. Det betyr at SØF-rapporten ikke forteller den absolutte *sannhet* om Skolekvalitet i videregående skole i Norge. Når man ser dette i sammenheng med diskusjonen i neste kapittel, blir denne konklusjonen enda tydeligere.

## 3 Målsettinger og konsekvenser

I dette kapitlet skal vi drøfte om det etablerte samlemålet er et dekkende og gyldig mål på skolekvalitet, og også drøfte hvor godt rapporten er egnet som grunnlag for et forbedringsarbeid i videregående skole. Til slutt vil vi drøfte noen mulige konsekvenser av rapporten.

### 3.1 Målsetting nr. 1: Å etablere et dekkende og gyldig mål på *Skolekvalitet*

Det heter innledningsvis i rapporten at «*Hensikten med dette prosjektet er å framskaffe kunnskap om kvalitet og kvalitetsforskjeller mellom videregående skoler*» (Falch mfl. 2016: 1). For å kunne gjøre dette bør man ha et mål som faktisk måler skolekvalitet. Vi tolker derfor dette til at en målsetting med prosjektet og rapporten er å etablere et dekkende og gyldig mål på *Skolekvalitet*.

Sju skolebidragsindikatorer inngår i målet på skolekvalitet. Fire av disse handler om gjennomføring av videregående opplæring og tre handler om faglige prestasjoner målt med karakterer. At disse forholdene inkluderes i et mål på skolekvalitet er riktig. Det er en sentral målsetting for videregående opplæring i Norge at flest mulig skal fullføre og bestå, og skolen jobber for at flest mulig skal prestere best mulig, dvs. oppnå best mulig karakterer. Sentrale aspekter ved kvaliteten på en videregående skole vil være hvor stor andel av elevene som fullfører og hvilke karakterer de oppnår.

Det man imidlertid kan stille spørsmål ved, er om fire indikatorer på gjennomføring og tre prestasjonsmål er tilstrekkelig for å etablere et samlemål på *Skolekvalitet*.

Det kan argumenteres for at om man skal etablere et mål på skolekvalitet, så bør et slikt mål inneholde langt flere aspekter ved det som foregår på en skole. Man vet bl.a. at mobbing og dårlig trivsel påvirker elevenes læringsarbeid negativt. Et godt læringsmiljø påvirker læringsutbytte positivt. Man kunne tenke seg at graden av mobbing, hvordan trivselen er blant elevene og hvor godt læringsmiljøet er, kunne inngå som indikatorer i et mål på skolekvalitet. Det kan hevdes at alle disse forholdene er forutsetninger for god gjennomføring og gode prestasjoner, slik at disse indikatorene er indirekte representert i de sju valgte skolebidragsindikatorerne. Det er grunn til å tro at det er slik, men det er også grunn til å tro at disse indikatorene også ville kunne yte selvstendige bidrag til skolekvalitet, og derfor argumenterer vi for at de kunne vært inkludert som skolebidragsindikatorer.

I Norge begynner så godt som hele årskullene i videregående skole. Det betyr at det er meget stor variasjon i elevenes faglige forutsetninger når de begynner. Det er tidligere påvist at om lag ti prosent av årskullene har karakteren 1 eller mangler karakter i minst ett fag når de går ut av grunnskolen (Frøseth m.fl. 2008). Stortingsmeldingen *På rett vei* (Meld.St.20, 2012-2013) anslår at «(...) rundt 20 prosent av et elevkull har svært lav sannsynlighet for å gjennomføre videregående opplæring med



studie- eller yrkeskompetanse (...)» (s 100). Det er en målsetting for videregående opplæring at flest mulig også blant disse elevene skal nå lengst mulig. I SØFs undersøkelse er elever med færre enn ti karakterer fra grunnskolen utelatt. Det samme gjelder spesialskoler. Dette betyr at de skolefaglig svakeste elevene ikke inngår i datamaterialet. En tolkning av dette er at SØF-undersøkelsen måler skolekvalitet for elever over et visst prestasjonsnivå. Det å lykkes med skolefaglig svake elever inngår ikke som en indikator på skolekvalitet. Alternativt kunne man tenke seg at det å gi gode tilbud til elever med skolefaglig svakt grunnlag fra grunnskolen, f.eks. elever med behov for spesialundervisning og elever som sikter mot planlagt grunnkompetanse som lærekandidat eller praksisbrevkandidat, kunne inngå i et mål på skolekvalitet.

Sommeren 2015 leverte det såkalte Ludvigsen-utvalget sin innstilling *Fremtidens skole* (NOU 2015:8). Utvalget fremhevet betydningen av et bredt kompetansebegrep, og mener at «skolen mer systematisk enn i dag skal støtte elevenes sosiale og emosjonelle læring og utvikling i fagene» (s 80), fordi dette «har betydning for faglig læring og for elevenes senere liv» (s 21-22). Begrepet sosiale og emosjonelle kompetanser defineres av utvalget som personers holdninger, oppførsel, emosjoner og sosiale ferdigheter og relasjoner, og man eksemplifiserer med ferdigheter som bl.a. etisk vurderingsevne, engasjement, utholdenhet, forventninger til egen mestring, å kunne ytre seg og bidra, ta hensyn til fellesskapet ved å regulere egne tanker, følelser og handlinger, respektere og se verdien av andres synspunkter, nysgjerrighet, utholdenhet og åpenhet (s 22). Dette er ferdigheter som også omtales som *non-cognitive skills* eller *21<sup>st</sup> Century skills*. F.eks. nevner Lamb et al. (2015) bl.a. motivasjon, utholdenhet, selv-kontroll, sosial kompetanse, kreativitet, kritisk tenkning, fleksibilitet, ansvarlighet, samarbeidsevne som eksempler på *21<sup>st</sup> Century skills*. Man kunne tenke seg at indikatorer på sosiale og emosjonelle ferdigheter, kunne inngå i et mål på skolekvalitet.

Vi har pekt på noen forhold som kunne inngått som indikatorer på skolekvalitet. Flere kunne vært nevnt, men hovedpoenget har vært å vise at mange sentrale aspekter ved videregående skoles virksomhet og målsettinger mangler i det etablerte målet på skolekvalitet.

Også forfatterne av SØF-rapporten er inne på at noen indikatorer mangler når de skriver at «Vi kunne for eksempel tenke oss å inkludere informasjon om hvor dyktig skolene er til å skaffe lærlingeplasser, samt informasjon om ikke faglige ferdigheter slik som for eksempel sosiale ferdigheter og samarbeidsevner» (s 2-3). Men, skriver de også: «Det er imidlertid ikke mulig å inkludere alle ønskelige resultatmål i en databasert kvantitativ studie som denne».

Dette kan forstås slik at forfatterne ser at et mål på skolekvalitet burde bygge på et utvalg av indikatorer som representerer en større bredde av det som kjennetegner skolers virksomhet og av det som er skolens målsettinger. Men fordi dette ikke er informasjon som fins tilgjengelig for kvantitative analyser, velger man å bruke det som fins. Dette er forståelig.

Det kan imidlertid stilles spørsmålsteget ved bruken av selve begrepet *Skolekvalitet*. Bruk av dette begrepet kan signalisere at man har inkludert flere aspekter ved skolens oppgaver og målsettinger enn kun fullføring/gjennomføring og faglige prestasjoner. Bruk av dette begrepet kan love mer enn man kan holde.

Forfatterne av rapporten er også inne på dette når de skriver: «Det er viktig å understreke at skolekvalitet slik den måles her ikke fanger opp alle aspekter ved skolens oppdrag» (s 46), og videre at «Indikatoren for skolekvalitet må derfor betraktes som en relevant, men ikke en komplett indikator på den «sanne» skolekvaliteten» (s 46).

Det er grunn til å spørre, når forfatterne ser disse svakhetene og betenkelighetene ved å bruke begrepet *Skolekvalitet*, hvorfor de likevel velger å bruke begrepet? Kanskje burde man heller kalt det etablerte samlemålet for *Gjennomføring og prestasjoner*. Etter vår vurdering ville det vært mere korrekt og presist å bruke en slik betegnelse. Ved å bruke begrepet *Skolekvalitet* signaliserer man et mer omfattende innhold i dette samlemålet enn det er grunnlag for.

Noen vil helt sikkert mene at dette er flisespikkeri, og at med sju sentrale indikatorer for skolekvalitet inkludert, er det forsvarlig å kalle dette for *Skolekvalitet*. Noen vil kanskje mene at det er ikke så viktig hva man kaller det man har målt, og at *Skolekvalitet* er akseptabelt. Vi er ikke enige. Vi vet at språk er makt, og dersom dette etableres som et anerkjent mål på *Skolekvalitet*, er det sannsynlig at dette vil få noen konsekvenser for virksomheten i videregående skole. Vi kommer tilbake til dette i avsnitt 3.3.

I forlengelsen av dette kan man også reise spørsmål om man i stedet for å etablere dette ene målet på skolekvalitet, ville vært bedre tjent med å etablere et mål på fullføring/gjennomføring basert på de fire fullførings-/gjennomføringsindikatorerne og et mål på faglige prestasjoner basert på de tre karakterindikatorerne man har inkludert i det etablerte målet på skolekvalitet.

Konklusjonen på drøftingen ovenfor er at det målet på skolekvalitet som brukes i SØF-rapporten utelater flere indikatorer som kunne inngått i et mål på skolekvalitet. Målet som er etablert er relevant og anvendbart, men med kunnskap om hvilke indikatorer som inngår og at mange sentrale indikatorer mangler, er det etter vår vurdering ikke etablert et dekkende og gyldig mål på *Skolekvalitet*. Det ville vært mere korrekt og presist å kalle dette for et mål på *Gjennomføring og prestasjoner*.

## 3.2 Målsetting nr. 2: Å etablere et grunnlag for forbedringsarbeid i videregående skole

Den andre målsettingen med arbeidet er at informasjon om skolekvalitet til skoler og fylkeskommuner kan utgjøre et grunnlag for forbedringsarbeid. I rapporten formuleres dette slik:

*Dersom det er store kvalitetsforskjeller mellom videregående skoler, kan det gi et godt grunnlag for at skoler kan forbedre seg ved å lære av andre skoler. Videre kan påvisning av kvalitetsforskjeller mellom skoleeiere muliggjøre forbedring på fylkeskommunenivå. Aktiv kvalitetsutvikling rettet mot enkeltskoler i regi av skoleeierne kan bli mer målrettet og effektiv hvis pålitelig informasjon om kvaliteten på skolene foreligger (side 1).*

Et viktig spørsmål å stille er om rapporten faktisk kan fungere slik. Dette ser vi nærmere på nedenfor, gjennom å drøfte noen ulike sider ved rapporten.

### 3.2.1 Data om elever som begynte i videregående skole for 7-8 år siden

Elevene som inngår i datamaterialet begynte i videregående opplæring i 2008 og 2009, altså nesten sju og åtte år før rapporten fra SØF ble presentert. Det betyr at tallene for skolebidrag og skolekvalitet for skolene som ble presentert i april 2016 er beregnet på grunnlag av fullføring/gjennomføring og faglige prestasjoner blant elever hvor svært mange var ferdige allerede i juni 2011 og hvor majoriteten var ferdige i 2012, altså 4-5 år før denne rapporten ble offentliggjort.

Rapporten er tydelig på dette, men det blir uheldig når den offentlige debatten føres som om resultatene gjelder dagens elever, skoler, lærere og ledere. Det finnes skoler og fylkeskommuner som var oppmerksomme på at situasjonen var kritisk, og som har satt i verk tiltak for å bedre situasjonen. Mange har lyktes, og vet at situasjonen på egen skole eller i egen fylkeskommune er bedre enn det skolebidragsindikatorerne og målet på skolekvalitet viser. Når man vet at man har snudd en negativ trend og at forholdene er bedret, kan det oppleves som urimelig at man dømmes og kritiseres for en situasjon som var, og ikke for den situasjonen som er. Når man får kritiske spørsmål fra media og fylkestingspolitikere samtidig som man vet at beskrivelsen i rapporten ikke er gyldig for dagens virkelighet, kan dette oppleves urimelig.

Forskerne er også oppmerksomme på dette forholdet, men argumenterer for å bruke så gamle data. Hovedargumentet er at dersom man skal inkludere fullført og bestått etter fem år som en av indikatorerne, så må data være for kull som begynt i videregående opplæring så langt tilbake i tid. Dette er et gyldig argument, gitt premisset. Vi mener imidlertid at det er god grunn til å vurdere om man bør bruke data for nyere kull for å undersøke og etablere mål på skolekvalitet. Da ville resultatene

sannsynligvis oppleves som mer gyldige, relevante og aktuelle som grunnlag for en debatt om skolekvalitet. Nå er det røster i debatten som, kanskje med rette, argumenterer for at dette trenger man ikke ta hensyn til ettersom dette gjelder en situasjon som har vært, og at bildet er et helt annet i dag, slik assisterende rektor på Nannestad videregående skole sa til Eidsvoll Ullensaker Blad i april 2016:

*Situasjonen er snudd siden disse elevene gikk ut i 2011 og 2012. Skoleledelsen tok tak i utfordringene da, og resultatene vi kan vise til i dag er vi svært godt fornøyde med. Rapporten sier hvordan det var, ikke hvordan det er i dag. (...) Vi hadde ønsket at bildet viste skolen slik den framstår i dag. Da kunne vi hatt nytte av informasjonen til videreutvikling av skolen.*

På grunnlag av drøftingen over mener vi at der er grunn til å vurdere å bruke data for kull som begynte i videregående skole nærmere i tid, selv når det innebærer at fullført og bestått etter fem år ikke kan inkluderes. Etter vår vurdering er det fullt mulig å bruke andre mål på gjennomføring og prestasjoner i videregående skole. For gjennomføring kunne man f.eks. bruke fullført og bestått Vg1 og Vg2 som to slike mål. Et tredje gjennomføringsmål kunne være andel blant elevene som starter i Vg2 i august et gitt skoleår som fortsatt er i videregående opplæring per 1. november skoleåret etter. Dette er relevant fordi vi vet at svært mange ungdommer avslutter sin videregående opplæring etter det andre året på yrkesfag (bl.a. Markussen mfl. 2008, Markussen 2014). Derfor ville et gjennomføringsmål som inkluderer overgangen mellom VG2 og Vg3/lære være en god indikator på skolekvalitet.

Som mål på faglige prestasjoner kunne man bruke gjennomsnittlig standpunkt karakter fra Vg1 og Vg2. Ved å samarbeide med fylkeskommunene direkte om datainnsamling ville man kunne ha data tilgjengelig for et kull som begynte i videregående skole et gitt skoleår, i januar når ungdommene er i sitt tredje år i videregående opplæring, og resultater av analyser ville kunne presenteres høsten etter at de som har brukt tre år på videregående opplæring er ferdige.

### **3.2.2 Mange skolebidragsindikatorer og mål på skolekvalitet for mange skoler er ikke signifikant forskjellig fra gjennomsnittet**

Vi har foran i kapittel 2 sett på tallene for Østfold, og bl.a. drøftet signifikans og usikkerhet. For 48 prosent av alle skolene som inngår i denne undersøkelsen skiller ikke skåren for skolekvalitet seg signifikant fra landsgjennomsnittet på 3,1. Det betyr at man ikke (med tilstrekkelig sannsynlighet) kan utelukke at disse skolene faktisk skårer gjennomsnittlig på kvalitet.

Hver skole har inntil sju skolebidragsindikatorer, med unntak av de skolene som ikke har elever med eksamen i norsk studieforberedende Vg3. I alt er det 2410 skolebidragsindikatorer fordelt på de 359 skolene. I alt 73 prosent av disse er ikke signifikant forskjellig fra gjennomsnittet på 0 (null). For Østfold del gjelder dette 89 prosent av alle skolebidragsindikatorerne. Bare åtte av 75 indikatorer her er signifikant forskjellige fra null.

Slik vi har fulgt offentliggjøringen av rapporten, presentasjonen av resultatene og debatten som har fulgt, er det vår vurdering at dette har vært så godt som usynlig i det offentlige ordskiftet om resultatene. Mange skoler og fylkeskommuner har sett på skårene, og forholdt seg til disse som om de er signifikant forskjellige fra gjennomsnittet på 3,1.

Vi har bl.a. observert at skoler som har en skåre lavere enn 3,1 har oppfattet at de «gjør det dårlig», selv om de ikke skiller seg signifikant fra landsgjennomsnittet, og at det dermed ikke kan utelukkes at de er gjennomsnittsskoler. På den andre siden har skoler som har skåret bedre enn 3,1, oppfattet at de «gjør det godt», selv om de ikke skiller seg signifikant fra landsgjennomsnittet. Vi bruker to skoler fra Østfold for å illustrere dette: Kalnes og Borg videregående skoler. Kalnes skårer 2,4 og Borg skårer 3,4, altså en forskjell på 1 «karakter». Men som tabell 3 foran i denne rapporten viser, skiller ingen av disse to skolene seg signifikant fra landsgjennomsnittet på 3,1. Den ene har fremstått som en «vinner», den andre som en «taper», men i realiteten kan vi ikke utelukke at begge er

gjennomsnittsskoler fordi analysene viser at ingen av dem skiller seg signifikant fra landssnittet, dvs. at begge teoretisk kan ha en skåre på 3,1.

Fordi signifikans og usikkerhet har blitt underkommunisert og vært usynlig i debatten, har mange skoler opplevd å «ha gjort det dårlig» eller å «ha gjort det bra», selv om de ikke skiller seg signifikant fra landsgjennomsnittet. Dette har bidratt til å tegne et feilaktig bilde i offentligheten, og mange skoler har unødvendig følt seg som tapere i «konkurransen».

### **3.2.3 Yrkesfag og studieforberevende måles på samme måte**

Elever på studieforberevende og yrkesfaglige utdanningsprogrammer måles og beregnes likt i undersøkelsen. Dette er problematisk at to grunner. For det første går yrkesfagelevne i videregående skole bare i to år før de går over i lære. Det betyr at det som skjer i lærebedriften i læretida tillegges den skolen elevene begynte på. F.eks. når en elev fullfører og består fagprøven, så vurderes det som en del av kvaliteten ved den skolen der eleven begynte i Vg1.

For det andre er det grunn til å regne med at det er ulike karakterregimer studieforberevende og på yrkesfag, og at det er et tredje karakterregime i ungdomsskolen. Det er tidligere vist (bl.a. Markussen & Sandberg 2005) at elever som begynner på yrkesfag i snitt forbedrer karakterene sine, mens elever som begynner på studieforberevende i snitt oppnår dårligere karakterer enn på ungdomsskolen. Dette indikerer at en 5'er ikke er en 5'er, men at en 5' er i det minste er tre ulike ting: En 5'er betyr en ting i ungdomsskolen, noe annet på studieforberevende og en tredje ting på yrkesfag. Vår vurdering er at når man beregner skolekvalitet bl.a. på grunnlag av karakterer og også tar hensyn til karakterene fra grunnskolen og forutsetter at karakterene på yrkesfag og studieforberevende er uttrykk for det samme, så er det en feilkilde. Også dette kan bidra til å svekke rapporten som et grunnlag for forbedringsarbeid. På grunnlag av dette er det grunn til å spørre om man hadde vært tjent med at skolekvalitet for yrkesfagelever og for studieforberevende elever hadde vært presentert som to ulike mål.

### **3.2.4 Elevresultater kan tillegges en skole elevene knapt har gått på**

Det fremgår av rapporten at 15 prosent av de videregående skolene og 20 prosent av elevene som inngår i datamaterialet har vært berørt av skolesammenslåinger (s 10). Resultatene relateres til skoler som eksisterte våren 2014. Det betyr at for en ikke ubetydelig andel av elevene i undersøkelsen er resultatene deres tillagt en skole som ikke eksisterte i samme form og med samme struktur i 2008/2009 som i 2014. Forfatterne skriver at resultatene for disse skolene må tolkes med varsomhet. For disse 15 prosent av de videregående skolene som ikke eksisterte i samme form og med samme struktur i 2008/2009 som i 2014, er det vår vurdering at resultatene i SØF-rapporten vanskelig kan tjene som grunnlag for forbedringsarbeid. Et eksempel i rapporten er Kuben videregående skole i Oslo, som er en nybygd skole som åpnet i 2013, og dermed ikke fantes i 2008/2009, men som likevel står med mål på skolebidragsindikatorer og skolekvalitet i rapporten.

En god del elever bytter skole i løpet av videregående opplæring. I utvalget av elever i SØFs undersøkelse var 19 % av elevene registrert siste gang i videregående opplæring ved en annen skole enn skolen de startet på etter grunnskolen (s 12). I undersøkelsen tillegges elevenes resultater (med unntak av de to indikatorene som bygger på eksamensresultater) den skolen elevene begynte på. Det betyr at for hver femte elev kan resultatene deres tillegges en skole hvor de ikke har gått hele tiden.

### **3.2.5 Alle indikatorene teller like mye**

De sju indikatorene teller alle like mye i målet på skolekvalitet. Forskerne skriver at man har vurdert vektning, dvs. å tillegge skolebidragsindikatorer ulik betydning, men etter en del testing valgt å ikke gjøre det. Det betyr at den ene eksamenskarakteren i norsk betyr like mye inn i skolekvalitetsmålet som gjennomsnittet av alle standpunkt-karakterene. Man kan stille spørsmål ved om det er rimelig at alle indikatorene teller likt, eller om man burde tillagt skolebidragsindikatorer ulik betydning.

Et annet interessant forhold ved skolebidragsindikatoren *norskeksamen i Vg3* bare gjelder for elever på studieforbereidende og ikke på yrkesfag. På skoler uten elever i Vg3 studieforbereidende inngår derfor ikke denne skolebidragsindikatoren. Det betyr at på 95 av skolene er det seks og ikke sju indikatorer som inngår i målet på skolekvalitet. Forskerne har testet dette og konkluderer med at det ikke har vesentlig betydning for skolenes resultat.

Dette betyr at det kanskje særlig er grunn til å være oppmerksom på tolkning av indikatoren *norskeksamen i Vg3*, både fordi den bare representerer ett fag i videregående skole og fordi den ikke inngår i målet på skolekvalitet på 27 prosent av skolene.

### **3.2.6 Formidlingen**

SØF-rapporten er laget av samfunnsøkonomer, og de har selvsagt brukt sitt fags metoder og terminologi. Det er forståelig og rimelig. Konsekvensen har blitt at for mange som arbeider innenfor videregående-sektoren; lærere, skoleledere, fylkesutdanningsbyråkrater og fylkesutdanningspolitikere, er rapporten vanskelig å lese og forstå. NIFU har fått en rekke henvendelser som underbygger dette. Formler, figurer, omfattende tabeller av ulikt slag sammen med bruken av metodebegreper som f.eks. signifikans, konfidensintervall, usikkerhet, ujustert og standardisert uten at disse begrepene er forklart med tanke på at personer uten metodekompetanse skal ha mulighet for å forstå, bidrar til at rapporten blir vanskelig tilgjengelig.

Når rapporten blir så vanskelig tilgjengelig, kan man tenke seg at det som i hovedsak sitter igjen, og som skoler og fylkeskommuner forholder seg til, er karakteren man har fått. Men hva betyr det å ha 2,3, 3,1 eller 4,8 i skolekvalitet? Selv om de aller fleste er fortrolig med karakterskalaen 1-6, er det sannsynlig at disse tallene oppfattes svært ulikt. Og dersom det som i hovedsak oppfattes er tallet på skolekvalitet, og at tallet i seg selv kan være vanskelig å virkelig forstå hva innebærer, er det grunn til å spørre hvor godt grunnlag dette kan være for utviklingsarbeid på skolene.

### **3.2.7 Grunnlag for forbedring: oppsummering**

Vi har foran vist at

- Dataene som analysene av skolekvalitet bygger på gjelder elever som begynte i videregående skole i 2008 og 2009. Dette betyr at man ikke kan forutsette at de funnene man gjør gjelder skolene slik de er i dag.
- Videre har vi vist målet på skolekvalitet ikke er signifikant forskjellig fra gjennomsnittet på 3,1 for 48 prosent av skolene i landet og for seks av elleve skoler i Østfold.
- Skolekvalitet måles likt for yrkesfag og studieforbereidende til tross for store forskjeller mellom de to utdanningsretningene.
- Elevers gjennomføring og karakterer kan bli tillaget skolekvaliteten på skoler der de knapt har gått pga. skolesammenslåinger og at elever bytter skole.
- Alle indikatorene teller like mye, f.eks. er det ikke forskjell på betydningen av den ene eksamenskarakteren i norsk og gjennomsnitt av alle standpunkt-karakterene.
- 27 prosent av skolene ikke har eksamen i norsk på Vg3 studieforbereidende, slik at for disse skolene er det bare seks skolebidragsindikatorer som gjelder.
- Rapporten og presentasjonen av resultatene er vanskelig tilgjengelig for personer uten samfunnsfaglig metodekompetanse.

Vår vurdering er at disse forholdene vi har drøftet, bidrar til at rapporten ikke vil fungere som et så godt grunnlag for forbedringsarbeid som målsettingen var. Særlig mener vi at gamle data som ikke tegner et bilde av dagens situasjon, at målet for skolekvalitet ikke er signifikant forskjellig fra gjennomsnittet

for halvparten av skolene og at rapporten er vanskelig tilgjengelig for skolefolk flest, svekker rapporten som grunnlag for forbedringsarbeid.

### 3.3 Konsekvenser

Dette er første gang en rapport som måler gjennomføring og prestasjoner og offentliggjør resultatene for identifiserte skoler, har blitt presentert i Norge. Etter offentliggjøring av en slik rapport, er det grunn til å spørre om hvilke konsekvenser offentliggjøringen av og innholdet i rapporten kan komme til å få.

Noen konsekvenser kan vi se allerede nå, men det meste vil vi først få kjennskap til senere. Vi vil peke på noen konsekvenser som allerede er synlige, men også trekke frem noen mulige konsekvenser.

Det vi har sett allerede, er at rapporten har lagt premisser for en debatt om kvaliteten på de videregående skolene i Norge. Nå diskuterer man skolekvalitet med utgangspunkt i denne rapporten. Et interessant fenomen er at debatten i stor grad har et innhold som om resultatene gir et bilde av situasjonen på skolene i dag. Man har glemt at disse resultatene er basert på data om elever som for lengst er ferdige i videregående skole. Dette betyr at debatten ikke er reell: Man diskuterer noe som var på en måte som om det er. Dette kan være uheldig, særlig for skoler med lav skåre som har opplevd en forbedring de siste årene.

Vi har også sett at offentliggjøringen av rapporten og særlig «karakteren» på den enkelte skole har skapt stor uro blant lærere, rektorer, utdanningsbyråkrater og utdanningspolitikere på fylkesnivå, særlig når man har skåret lavt. I noen tilfeller uten grunn, fordi man har bekymret seg over skoler med lav skåre uten å ha oppfattet at resultatene ikke skiller seg signifikant fra landsgjennomsnittet.

At analysemetodene er avanserte og at formidlingen av dem har vært vanskelig å forstå, kan i flere tilfeller ha ført til at skoler og fylkeskommuner har oppfattet skåren de har fått, og ikke alt som ligger bak, blant annet alle forbeholdene og anbefalinger i rapporten om forsiktighet ved tolkning av resultatene. Dermed har man fått en unyansert debatt på grunnlag av én karakter.

En konsekvens vi kan komme til å se - det vil tiden vise - er at det tas beslutninger på et feilaktig/misforstått/ukjent grunnlag. Et eksempel på dette er at skoler som skårer lavt setter i verk tiltak basert på rapporten, uten at de har oppfattet at skolens resultat ikke skiller seg signifikant fra gjennomsnittet, og at det derfor ikke kan utelukkes at de er en gjennomsnittsskole. Et annet eksempel er at man glemmer at man har analysert data for elever som begynte i videregående skole i 2008 og 2009, og at man kan komme til å ta beslutninger for i dag og i morgen basert på analyser av historiske data.

En sannsynlig konsekvens av denne rapportens offentliggjøring og forskernes anbefaling av at dette bør gjentas, er at analysene og rapporteringen faktisk blir gjentatt. I debatten har det kommet frem synspunkter knyttet til valg av indikatorer og forslag om andre indikatorer som kan måle forhold som ligger nærmere i tid. Motargumentet har vært at da kan man ikke sammenligne, derfor bør man bruke de samme indikatorene om man skal gjenta dette. Dette er et argument som kan vinne gehør, selv om det kan argumenteres med at det i denne fasen av etableringen av et slikt samlemål på *Skolekvalitet* (eller *Gjennomføring og prestasjoner*) bør være viktigere å finne de «riktige» indikatorene enn å kunne sammenligne over tid. Men det er mulig at sneballen har begynt å rulle. Nå har man gjort dette en gang, og da gjør man det en gang til med de samme indikatorene slik at man kan sammenligne over tid. Og har man gjort det to ganger er det ikke vanskelig å spå fortsettelsen.

En konsekvens av en gjentakelse med de samme indikatorene og med fortsatt bruk av begrepet *Skolekvalitet*, kan være endring av fokus hos skoleeier, hos rektorene og hos lærerne. Når man vet hvilke indikatorer som gjelder kan dette legge føringer for prioriteringer. Når det er indikatorer som handler om gjennomføring og prestasjoner (karakterer) som teller, kan det medføre mindre fokus på andre sider ved skolens aktivitet som også er indikatorer på skolekvalitet (jf. avsnitt 3.1).

En mulig konsekvens av endring av fokus kan være at man legger om kursen for det utviklings- og forbedringsarbeidet man driver på skolene og i fylkeskommunene. Man tar nye initiativ som følge av rapporten, og konsekvensen kan være at tidligere prioriteringer gjøres om og at tidligere satsinger nedprioriteres.

En siste mulig konsekvens vi vil peke på er at det kan bli mindre viktig for skolene å gi gode tilbud til skolefaglig svake elever. I undersøkelsen til SØF er elever som har karakter fra grunnskolen i færre enn ti fag ikke inkludert. Heller ikke spesialskoler. Det betyr at mange skolefaglig svake elever ikke er inkludert i datamaterialet. Når disse ikke blir målt og ikke inngår i målet på *Skolekvalitet*, kan en konsekvens bli at det ikke blir like viktig å prioritere disse elevene.

## 4 anbefalinger

På grunnlag av våre analyser av tallene for Østfold i SØF-rapporten er det mulig å komme med noen forsiktige anbefalinger til Østfold fylkeskommune.

Det viktigste vi har identifisert er at Østfold fylkeskommune er et gjennomsnittsfylke og at de videregående skolene i fylket i stor grad er gjennomsnittsskoler. Med utgangspunkt i at så mange av skolebidragsindikatorerne og skolekvalitets-indikatorerne ikke er signifikant forskjellig fra landsgjennomsnittet, bør man være svært varsomme i sin tolkning av tallene. Når en skåre ikke er signifikant forskjellig fra gjennomsnittet, er det god grunn til – enten skåren ligger over eller under landsgjennomsnittet – å behandle skåren som om den ligger på landsgjennomsnitt. Det er også viktig å huske at dette gjelder elever som begynte for 7-8 år siden, og ikke reflekterer situasjon i fylket og ved skolene i dag.

De er også viktig å være klar over at målet på *Skolekvalitet* ikke er «sannheten». Vi har både argumentert for at mange sentrale indikatorer på skolekvalitet mangler, og at det ligger noen subjektive valg bak valg av indikatorer og beregningsmåter. Også disse aspektene taler for en varsom tolkning av tallene.

Fordi Østfold fremstår som et gjennomsnittsfylke og fordi datagrunnlaget er gammelt, er det vår vurdering at det ikke er grunn til noen strakstiltak på grunnlag av SØF-rapporten. Vi mener at det er viktig å være bevisst det forbedrings- og utviklingsarbeidet som har vært satt i verk i Østfold de siste årene, å holde fast ved dette og ikke stoppe pågående tiltak på grunnlag av SØF-rapporten. Det er også viktig å være særlig bevisst på de resultatene man vet man har oppnådd de siste årene, og å være bevisst eksisterende kunnskap om egen skole og egen fylkeskommune, og vurdere resultatene fra SØF-rapporten opp mot dette. På denne måten kan SØF-rapporten inngå som et av mange grunnlag for forbedringsarbeid på skolene i Østfold, men den alene bør ikke danne grunnlag for brå endringer og strakstiltak.



## Vedlegg

**Tabell V1: Skolekvalitet når standpunktkarakterer er utelatt. Skolene i Østfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunktkarakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet	Avvik fra skolekvalitet i tabell 3
Fredrik II	0	0.9	0.32	-0.21	-1.83		-0.62	2.8	+0.1
St. Olav	0.37	0.52	0.68	0.68	-1.91		-0.35	3.1	+0.1
Kirkeparken	-0.48	-0.24	-0.09	-0.54	-1.29		-0.54	2.4	-0.1
Askim	-0.26	-1.3	-0.73	-0.05	-1.46		-0.34	2.2	+0.1
Halden	0.21	0.54	1.12	0.49	-0.41		-1.05	3.3	+0.2
Greåker	-0.56	-0.54	0.11	-0.67	-0.61		-0.64	2.5	+0.1
Mysen	-0.86	-1.03	-0.37	-0.52	-0.88		-0.63	2.1	-0.1
Glemmen	-0.29	-0.32	-0.08	0.12	0.2		1.47	3.3	+0.1
Malakoff	-0.6	-0.47	-0.09	-0.47	0.04		-0.34	2.7	+0.1
Borg	0.28	0.42	0.59	0.51	0.43		-0.78	3.4	0
Kalnes	-0.6	-0.92	0.26	-0.22	0.9		-2.04	2.5	+0.1
Gjennomsnitt Østfold	0	0	0	0	0		0	2.8	
Landsgjennomsnitt	1	1	1	1	1		1	3.1	
Standardavvik								0.97	

**Tabell V2: Skolekvalitet når standpunktkarakterer er utelatt. Skolene i Østfold.**

Skole	Fullføring	Deltagelse	Normert progresjon	Normert fullføring	Standpunktkarakterer	Norsk-eksamen	Skriftlig eksamen	Skolekvalitet	Avvik fra skolekvalitet i tabell 3
Fredrik II	0	0.9	0.32	-0.21	-1.83	-0.27		2.9	+0.2
St. Olav	0.37	0.52	0.68	0.68	-1.91	-0.12		3.1	+0.1
Kirkeparken	-0.48	-0.24	-0.09	-0.54	-1.29	0.22		2.6	+0.1
Askim	-0.26	-1.3	-0.73	-0.05	-1.46	-0.58		2.1	0
Halden	0.21	0.54	1.12	0.49	-0.41	-0.57		3.4	+0.3
Greåker	-0.56	-0.54	0.11	-0.67	-0.61	-0.56		2.5	-0.1
Mysen	-0.86	-1.03	-0.37	-0.52	-0.88	-0.02		2.3	-0.1
Glemmen	-0.29	-0.32	-0.08	0.12	0.2	-0.43		2.9	-0.3
Malakoff	-0.6	-0.47	-0.09	-0.47	0.04	-0.17		2.7	+0.1
Borg	0.28	0.42	0.59	0.51	0.43			3.6	+0.2
Kalnes	-0.6	-0.92	0.26	-0.22	0.9			2.9	+0.5
Gjennomsnitt Østfold	0	0	0	0	0	0			
Landsgjennomsnitt	1	1	1	1	1	1		3.1	
Standardavvik								0.97	

## Referanser

- Angrist, J. et al. (2015). *Leveraging lotteries for school value-added: Testing and estimation*. No. w21748. National Bureau of Economic Research, 2015.
- Falch, T., Bensnes, S. & Strøm, B. (2016). *Skolekvalitet i videregående opplæring. Utarbeidelse av skolebidragsindikatorer og mål på skolekvalitet*. Rapport nr 1. Trondheim: SØF
- Frøseth, M.W., Hovdhaugen, E., Høst, H. & Vibe, N. (2008). *Tilbudsstruktur og gjennomføring i videregående opplæring. Delrapport I. Evaluering av Kunnskapsløftet*. Rapport 40. Oslo: NIFU STEP
- Lamb, S., Jackson, J. & Rumberger, R. (2015). *ISCY Technical Paper: Measuring 21st Century Skills in ISCY*. Melbourne: Victoria University
- Markussen, E. (2014). *Utdanning lønner seg. Om kompetanse fra videregående og overgang til utdanning og arbeid etter avsluttet grunnskole 2002*. Rapport 1. Oslo: NIFU
- Markussen E. & Sandberg, N. (2005). *Stayere, sluttet og returnerte. Om 9756 ungdommer på Østlandet, deres vei gjennom, ut av, eller ut og inn av videregående opplæring, og om deres prestasjoner to år etter avsluttet grunnskole*. Skriftserie nr. 6. Oslo: NIFU STEP
- Markussen, E., Frøseth, M.W., Lødding, B. & Sandberg, N. (2008). *Bortvalg og kompetanse. Gjennomføring, bortvalg og kompetanseoppnåelse i videregående opplæring blant 9749 ungdommer som gikk ut av grunnskolen på Østlandet våren 2002: hovedfunn, konklusjoner og implikasjoner fem år etter*. Rapport 13. Oslo: NIFU STEP.
- Meld.St. 20 (2012–2013). *På rett vei. Kvalitet og mangfold i fellesskolen*. Oslo: Kunnskapsdepartementet.
- NOU 2015:8. *Fremtidens skole*. Oslo: Kunnskapsdepartementet.



Nordisk institutt for studier av  
innovasjon, forskning og utdanning

Nordic Institute for Studies in  
Innovation, Research and Education

[www.nifu.no](http://www.nifu.no)