



Karakterer i offentlige og private videregående skoler

En analyse av eksamens- og standpunktkarakter i norsk og matematikk og rutiner for standpunktvurdering i offentlige og private videregående skoler

Elisabeth Hovdhaugen, Idunn Seland, Berit Lødding, Tine S. Prøitz og Nils Vibe

Rapport 24/2014

Karakterer i offentlige og private videregående skoler

En analyse av eksamens- og standpunktkarakter i norsk og matematikk og rutiner for standpunktvurdering i offentlige og private videregående skoler

Elisabeth Hovdhaugen, Idunn Seland, Berit Lødding,
Tine S. Prøitz og Nils Vibe

Rapport 24/2014

Rapport	24/2014
Utgitt av Adresse	Nordisk institutt for studier av innovasjon, forskning og utdanning PB 5183 Majorstuen, NO-0302 Oslo. Besøksadresse: Wergelandsveien 7, 0167 Oslo
Oppdragsgiver Adresse	Utdanningsdirektoratet Postboks 9359 Grønland, 0135 Oslo
Trykk	Link Grafisk
ISBN	978-82-327-0011-0
ISSN	1892-2597 (online)

www.nifu.no

Forord

Denne rapporten dokumenterer prosjektet *Karakterpraksis i offentlige og private videregående skoler*, som NIFU har gjennomført på oppdrag av Utdanningsdirektoratet. Prosjektleder har vært Elisabeth Hovdhaugen, og Marit Solvoll har vært kontaktperson i Utdanningsdirektoratet.

Prosjektet har tatt utgangspunkt i både kvantitative og kvalitative data. Forskerne Nils Vibe og Elisabeth Hovdhaugen har hatt ansvar for den kvantitative delen av prosjektet, mens den kvalitative delen er gjennomført av forskerne Idunn Seland, Berit Lødding og Tine Sophie Prøitz. Idunn Seland har hatt hovedansvar for de kvalitative analysene av lærerintervjuene, mens Berit Lødding har analysert elevmaterialet. Alle i forskergruppen har bidratt med kommentarer og innspill til kollegaenes arbeid. Rapporten er gjennomgått og kvalitetssikret av Jorunn Spord Borgen, pedagogikkprofessor ved Idrettshøgskolen og Jens B. Grøgaard, forsker ved NIFU og sosiologiprofessor ved Høgskolen i Buskerud og Vestfold.

Vi vil rette en spesiell takk til de åtte skolene vi fikk lov til å besøke i forbindelse med datainnsamlingen, og da særlig til lærerne og elevene som stilte opp som informanter i gruppeintervjuene om karakterpraksis og standpunktkaraktersetting.

Oslo, juli 2014

Sveinung Skule
Direktør

Vibeke Opheim
Forskningsleder

Innhold

Sammendrag.....	7
1 Introduksjon.....	11
1.1 Avviksproblematikk og karakterer	12
1.2 Karakterer og regelverk	12
1.3 Karakterer og fag	13
1.4 Offentlige og private skoler	13
1.5 Karakterer og troverdighet.....	14
1.6 Problemstillinger.....	15
1.7 Teoretisk rammeverk.....	15
2 Metodisk tilnærming	19
2.1 Statistiske analyser av karakterer på skolenivå	19
2.1.1 Utvalg av skoler til kvalitative intervjuer	20
2.2 Kvalitative datakilder: Intervju og dokumentanalyse.....	20
2.2.1 Dokumentbaserte analyser.....	21
2.2.2 Epost-intervju	21
2.2.3 Gruppeintervju.....	21
3 Kvantitative analyser, avstand mellom standpunkt- og eksamenskarakter	23
3.1 Forutsetninger som ligger i problemstillingen.....	24
3.2 Forskjeller mellom offentlige og private skoler	26
3.3 Oppsummering.....	30
4 I spenningsfeltet mellom pålegg og praksis	31
4.1 Rutiner for elevvurdering	31
4.1.1 Designede rutiner kontra rutiner i daglig bruk	31
4.1.2 Aritmetisk og analytisk vurdering	32
4.1.3 Universell og differensiell behandling	32
4.2 Bruk av vurderingsforskriften	33
4.3 Bruk av eksamensforberedende prøver.....	34
4.3.1 Gamle eksamensoppgaver som forberedelse til heldagsprøver	35
4.3.2 Eksamensveiledninger som forberedelse til heldagsprøver	35
4.3.3 Hva heldagsprøven kan måle: forskjeller mellom norsk og matematikk	36
4.3.4 Når lærere skal sette karakter på heldagsprøven	36
4.3.5 Betydningen av fagsamarbeid i vurdering av heldagsprøven	38
4.3.6 Oppsamlingsprøver, vippeprøver og ekstra vurdering av strykkandidater	39
4.4 Sluttvurdering på bakgrunn av undervisvurderinger og samtaler med elevene	40
4.4.1 Tilbakemeldinger til elever etter prøver.....	41
4.4.2 Formelle og uformelle elevsamtaler	42
4.5 Sensorerfaring og skoleinitierte prosjekter for vurdering	43
4.5.1 Betydningen av sensorerfaring.....	44
4.6 Forskjeller mellom standpunktkarakter og eksamenskarakter.....	46
4.7 Oppsummering: rutiner for standpunktvurdering.....	48
5 «Vi er ungdommer, og vi gjør så godt vi kan ...».....	53
5.1 Intervjuenes innhold og noen tolkningsrammer.....	53
5.2 Valg av skole.....	54
5.3 Antakelser og erfaringer med forskjeller i belønning	54
5.4 Betydningen av karakterer.....	55
5.5 Standpunktvurdering blir ofte bedre enn eksamen.....	57
5.6 Forskjeller og likheter mellom fag i vurdering.....	59
5.6.1 Matematikk – et «fasitsvarfag»	59
5.6.2 Norsk og kunsten å knekke koden.....	61
5.7 Juks – hva er det?	63
5.8 Så hva er en rettferdig karakter?	66
5.9 Oppsummering og drøfting.....	67

6	Avslutning	71
6.1	Rapportens hovedfunn	71
6.2	Hvorfor er det avvik mellom eksamen og standpunktkarakterer mellom offentlige og private skoler?	74
6.2.1	Betydningen av tolkningsfellesskap	75
6.2.2	Betydningen av sensorerfaring og andre muligheter for kalibrering av vurderingsrutiner	76
6.2.3	Betydningen av forhandlingsrom i karaktersetting	77
6.2.4	Betydningen av formelle tilpasninger til designede rutiner	78
6.2.5	Oppsummering: tema for videre forskning	79
	Referanser	81
	Nettbaserte referanser	83
	Vedlegg 1 Guide fokusgruppe lærere	85
	Vedlegg 2 Guide fokusgruppe elever	89
	Vedlegg 3 Karakterforskjeller – delrapport levert til Udir i desember 2013	93

Sammendrag

Denne rapporten kartlegger praksiser og rutiner for hvordan standpunktkarakterer settes ved offentlige og private skoler. Bakgrunnen for oppdraget er en avisdebatt i 2013 der det ble påvist større avvik mellom standpunkt- og eksamenskarakter ved private enn ved offentlige videregående skoler, og der det ble reist spørsmål om hva dette kan skyldes.

Rapporten har tre problemstillinger:

1. Er det større avvik mellom standpunktkarakter og eksamenskarakter når private og offentlige skoler sammenlignes?
2. Hvilke rutiner og retningslinjer ligger til grunn for vurderingen ved de ulike skolene?
3. Er det forskjeller i rutiner og retningslinjer ved de ulike skolene?

Vi har brukt statistiske data for å undersøke den første problemstillingen, mens vi har brukt gruppeintervjuer med lærere og elever, samt e-postintervjuer med rektorer ved fire private og fire offentlige skoler for å kartlegge hvilke rutiner og retningslinjer skolene legger til grunn for vurderingen. Lærerne som har deltatt i gruppeintervjuene underviser i matematikk og norsk.

De kvalitative analysene har tatt utgangspunkt i et teoretisk rammeverk som er ment å fange opp og belyse interaksjonen mellom rutiner og praksis i vurderingssituasjonen. Det teoretiske rammeverket er beskrevet i kapittel 1.

Større avvik mellom standpunkt- og eksamenskarakter ved noen av de private skolene

De statistiske analysene tar utgangspunkt i individdata, der vi har informasjon om både eksamens- og standpunktkarakter i enkeltfag for samme person. I analysene sammenligner vi gjennomsnittlige avvik for grupper av elever ved fire ulike typer skoler. Analysene viser at det er forskjeller innad mellom ulike typer private skoler. Avviket mellom standpunktkarakter og eksamenskarakter er forholdsvis likt ved offentlige videregående skoler, private kristne skoler og private spesialiserte skoler, mens det er et større avvik mellom standpunktkarakter og eksamenskarakter ved private allmenne videregående skoler. Dette gjelder også når vi kontrollerer for andre forhold som har betydning for karakterene. Det er altså større avvik mellom standpunktkarakter og eksamenskarakter når vi sammenligner private og offentlige skoler, men dette gjelder ikke for *alle* private skoler, bare for de private allmenne videregående skolene. Imidlertid finner vi at det som har størst påvirkning på hvor stort avviket mellom de to karakterene er, er elevens standpunktkarakter. I praksis innebærer dette at elever med gode karakterer har større sannsynlighet for å gå tilbake til eksamen enn elever med svakere karakterer. Dette henger delvis sammen med at standpunktkarakterer i gjennomsnitt ligger høyere enn eksamenskarakterene i alle typer skoler, både offentlige og private.

Rutiner og retningslinjer som ligger til grunn for vurdering

Den andre problemstillingen dreier seg om hvilke rutiner, praksiser og retningslinjer som lærerne legger til grunn for standpunktvurderingen. Gruppeintervjuene viser at lærerne legger et mangfold av rutiner til grunn når de arbeider med vurdering og karaktersetting, og disse framtrer både som formelle pålegg og som rutiner i daglig bruk. Eksempler på formelle pålegg er læreres forståelse og bruk av vurderingsforskriften, av eksamensforberedende prøver (heldagsprøver) inkludert eksamensveiledningene, erfaringer med de lovpålagte elevsamtalene for vurdering og erfaringer med sensorarbeid samt skoleinterne prosjekter rettet mot vurderingspraksis. Rutiner i daglig bruk kan forstås som praksiser som enkeltlærere eller grupper av kolleger utvikler for å utfylle eller supplere de formelle påleggene, og som dekker et behov som lærerne opplever i det løpende arbeidet med vurdering. Viktige daglige rutiner er kollegasamarbeid for retting av besvarelser og tolkning av vurderingskriterier, uformelle samtaler med elever og praksiser for å vurdere helhet i en elevs prestasjon eller skoleresultater hvis eleven gjør det uventet dårlig på heldagsprøven.

Gjennomgangen av rutiner og praksiser for vurdering og karaktersetting viser klart at lærerne på alle de åtte skolene vi har besøkt, forholder seg nokså likt til de formelle påleggene. Tolkningene deres av disse påleggene er stort sett like på tvers av skoler. Lærernes fortellinger om praksiser i daglig bruk danner også et gjenkjennelig mønster på tvers av skolene. Dermed betyr dette at vi ikke observerer noen systematiske forskjeller mellom offentlige og private skoler i hvordan pålegg og praksiser brukes i arbeidet med vurdering og karaktersetting.

Derimot finner vi forskjeller med hensyn til rutiner og praksiser mellom faggruppene matematikklærere og norsklærere. Hovedtendensen i materialet er at matematikklærerne ofte følger en aritmetisk tilnærming til vurdering, som innebærer at de gir poeng for riktig svar og lar karakteren fremkomme på basis av en prosentsats av totalt antall oppnåelige poeng for prøven. Norsklærerne forholder seg derimot analytisk til vurderingen av elevbesvarelser. Dette innebærer en kvalitativ bedømming av elevens måloppnåelse med grunnlag i kompetansemålene for faget. Samtidig bruker også matematikklærerne en analytisk vurderingsmåte når de skal se på helheten i en elevbesvarelse og vurdere elevens *forståelse* av faget. Disse to tilnærmingene til vurdering er også styrende for hvordan norsklærere og matematikklærere omtaler vurderingsforskriften når de snakker om muligheten for å gi en sluttvurdering i de respektive fagene. Norsklærerne har utviklet mer eller mindre formaliserte praksiser som gjør at de foretar vurdering av elevenes sluttkompetanse over et lengre tidsrom og med en generelt bredere innfallsvinkel til vurderingssituasjonen enn hva matematikklærerne gjør.

Når det gjelder bruk av vurderingskriterier, fremhever norsklærerne dette som en gjennomgående og fortolkende praksis hvor formelle pålegg og daglige rutiner til en viss grad flyter over i hverandre. Matematikklærerne uttrykker et langt mindre behov for å tolke vurderingskriteriene i faget sitt. Dette påvirker også kollegasamarbeidet som kjennetegner de to faggruppene av lærere. Mens norsklærere har et løpende behov for å samtale og samarbeide om fortolkning og anvendelse av kriteriene, avgrenses matematikklærernes behov for kollegasamarbeid til forberedelse av prøven og i noen grad til tvilssituasjoner som oppstår under rettingen. Denne fagforskjellen gjenspeiles også i noen grad i elevintervjuene, ved at elevene gir uttrykk for at det er vanskeligere å vite hva som kreves for å oppnå en god karakter i norsk enn i matematikk. Elevene gir uttrykk for at de oppfatter kravene når matematikklæreren forklarer hva som skal til for å oppnå en viss karakter, mens det er vanskeligere å få grep om det samme i norsk. Med andre ord ser det ut til at det er noe vanskeligere for lærerne i norsk å formidle hva som ligger til grunn for vurderingen enn for matematikklærerne. Likevel, ved en av de offentlige skolene mente elevene at de fikk god veiledning og klare beskjeder om hva som kreves for å få god karakter i norskfaget. Norsklærerne ved denne skolen var også ganske eksplisitte på hva de ser etter når de vurderer besvarelser. Dette er en skole som har drevet Vurdering for læring over flere år. Uansett hvor avgjørende Vurdering for læring har vært i dette, viser eksemplet at kriteriene for vurdering i norskfaget ikke med nødvendighet må være diffuse eller ugjennomtrengelige for elevene. Spørsmålet er hvilken trening og erfaring lærere har når det gjelder å forklare vurderingskriteriene og gjøre dem eksplisitte for elevene.

Ingen forskjeller i rutiner og retningslinjer mellom offentlige og private skoler

Generelt ser vi i de åtte skolene vi har besøkt, ingen klare forskjeller mellom offentlige og private videregående skoler i hvordan det arbeides med vurdering og karaktersetting. Det synes som lignende rutiner, praksiser og retningslinjer er å finne både ved offentlige og private skoler, og forskjeller i slik praksis kan dermed ikke forklare hvorfor vi observerer et større statistisk avvik mellom standpunkt- og eksamenskarakterer ved private allmenne skoler. Både ved offentlige og private skoler jobber lærerne aktivt med prosjekter eller satsninger for å forbedre vurderingsarbeidet, men heller ikke her finner vi noen klare mønstre som kan bidra til å forklare forskjellen mellom offentlige og private skoler når det gjelder avvik mellom eksamen og standpunktkarakterer.

Andre forhold som kan bidra til å forklare avviket

Spørsmålet reiser seg da om hva som kan bidra til å forklare avviket vi observerer. Basert på våre data kan vi bare fremsette antakelser, som må undersøkes videre dersom man ønsker å finne ut av hva avviket skyldes. Vi vil påpeke at dette er temaer vi *antar* kan ha innvirkning, men som vi ikke kan konkludere om basert på denne studien. Med andre ord er dette antakelser som det med fordel kan forskes videre på.

Det første spørsmålet som kommer opp er om lærere på private skoler har tilgang til et tilstrekkelig stort tolkningsfellesskap til at de får anledning til å utvikle vurderingskriterier som er i samsvar med en gjengs oppfatning om slike kriterier. Det er mulig at noen av de private skolene har litt for små eller for lite formaliserte tolkningsfellesskap for å utvikle slike kriterier og anvende dem aktivt og på en ensartet måte i det daglige arbeidet. Dersom dette er tilfelle, kan en mulig løsning være at skoleeier eller Utdanningsdirektoratet, som har samordnings- og koordineringsansvar, legger til rette for å gi lærere muligheter for å møtes og å diskutere vurdering på tvers av skoletilhørighet. Vi vil anta at små skoler generelt, og ikke bare små private skoler, kan ha nytte av slike tiltak.

Det andre spørsmålet er hvordan det legges til rette for at flest mulig, og da kanskje særlig lærere ved private videregående skoler, får mulighet til sensorskolering. Mange av lærerne som har sensorerfaring, anbefaler at sensorskolering blir gjort obligatorisk og at så mange som mulig blir gitt muligheten til å høste erfaringer som sensor. Det er flere av lærerne ved de private skolene som er sensorer, men likevel vil det å inkludere flere yngre lærere kunne virke positivt for å skape et tolkningsfellesskap.

Det tredje spørsmålet vi vil stille er om elever ved private skoler har noe større rom for forhandling i standpunktkaraktersettingen, ved at omfanget og bruken av ekstraprøver som virker karakterjusterende kan være større på private skoler sammenlignet med offentlige skoler. Dersom et slikt forhandlingsrom er til stede, kan det virke drivende for høyere karakterer.

Til sist reiser vi et forskningsspørsmål som fremtrer i forlengelse av analysen uten å være knyttet direkte til mulige forskjeller mellom offentlige og private videregående skoler. I materialet har vi sett enkelte tendenser til at opplevde krav om at eksamens- og standpunktkarakter skal være så like som mulig, kan resultere i at lærere tilpasser standpunktkarakteren etter hva de tror eleven kommer til å få på eksamen. Da kan resultatet bli at avviket mellom standpunktkarakter og eksamenskarakter blir mindre, men samtidig faller i realiteten hensikten med å gi eleven standpunktkarakter bort. Forskningsspørsmålet kan følgelig formuleres slik: Dersom avvik mellom standpunktkarakter og eksamenskarakter blir mindre over tid, skyldes dette tilpasninger til resultatmålingsindikatorer eller reelle skoleutviklingsprosesser? Vi ser at dette også kan aktualiseres for skoler som er i en konkurransesituasjon og hvor små avvik mellom standpunktkarakter og eksamenskarakter kan ha positiv betydning for skolens omdømme.

Eksamenskarakter som kalibrering av standpunktkarakteren?

Et siste, overordnet tema som har vært til stede i prosjektet helt fra starten, er spørsmålet om i hvilken grad eksamenskarakteren faktisk kan og skal brukes for å kalibrere standpunktkarakteren. Spørsmålene som var formulert i utlysningen av dette prosjektet tyder på en slik tankegang. En

forutsetning for at eksamenskarakter skal kunne virke kalibrerende er at den er stabil over tid. Vi viste imidlertid i analysene av karakterer at eksamenskarakterene slett ikke er stabile over tid, og dermed er det heller ikke gitt at de kan brukes til kalibrering. I intervjuene med lærerne kommer også denne problemstillingen opp. Mange gir uttrykk for at det er forståelig at det er et avvik mellom de to typene karakterer, og at dette henger sammen med vurderingssituasjonen. Eksamen oppfattes som en langt smalere prøving av elevens kompetanse enn det standpunktkarakteren ideelt sett skal være, selv om begge vurderingsformer skal vurdere eleven ut fra kompetansemålene ifølge vurderingsforskriften.

Lærere som har sensorerfaring gir dessuten uttrykk for at de to vurderingssituasjonene ikke er like. Når de er sensorer, vurderer de den anonyme elevbesvarelsen kun utfra vurderingskriteriene, men når de setter standpunktkarakter, vil kjennskap til eleven og dennes faglige utvikling kunne påvirke læreren i vurderingssituasjonen, og oftest på en overveiende positiv måte. Med andre ord blir situasjonene forskjellige, og dette kan bidra til å forklare hvorfor det oppstår avvik mellom eksamens- og standpunktkarakterer. Imidlertid gir både lærere og elever uttrykk for at de ikke synes at avviket bør være for stort.

1 Introduksjon

Debatten om private skoler er ikke ny i Norge, og den nye regjeringens varslede endringer i privatskoleloven har gjort at debatten har blitt aktualisert igjen (KD 2014). Debatten dreier seg ofte om hvilken rolle de private skolene skal spille i det norske utdanningssystemet (Berge & Hyggen 2011). Tema for denne rapporten er ikke private skoler generelt eller debatten kring dette, heller ikke i hvilken grad private skoler virker sosialt segregerende, som Helland & Lauglo (2007) har forsket på, men snarere det forholdsvis smale temaet standpunkt-karaktersetting i offentlige og private videregående skoler. Prosjektet er et oppdrag for Utdanningsdirektoratet og tar utgangspunkt i direktoratets utlysning i konkurransegrunnlag kunngjort i mai 2013. Prosjektet har to overordnede hovedformål:

1. Fremskaffe sikker kunnskap om hva som ligger til grunn for medienes debatt våren 2013 om påståtte forskjeller i fastsettelse av standpunkt-karakterer mellom offentlige og private skoler.
2. Generere kunnskap om ulike praksiser og rutiner som ligger til grunn for standpunkt-karaktersetting ved ulike videregående skoler.

Hoveddelen av ressursene i prosjektet har vært fokusert på den andre delen av formålet, å kartlegge praksiser og rutiner for standpunkt-karaktersetting ved offentlige og private skoler. Resultatet av undersøkelsen skal utgjøre et grunnlag for videre arbeid med karakterpraksis og bidra med forslag til tiltak for å sikre kvalitet og tillit til standpunkt-karaktersetting ved private og offentlige skoler.

Bakgrunnen for forskningsoppdraget er at Utdanningsdirektoratet da oppdraget ble utlyst hadde lite dokumentasjon på om det foreligger konkrete forskjeller i fastsettelse av standpunkt-karakterer mellom offentlige og private skoler, og eventuelt hvordan og hvorfor dette forekommer. Dette gjelder særlig tall som har blitt presentert i offentlige medier om mulig avvik i karaktersetting mellom standpunkt og eksamens-karakterer på individnivå og mellom ulike skoler og i ulike fag. I desember 2013 publiserte Utdanningsdirektoratet et statistikknotat som konkluderer med at halvparten av elevene får lavere karakter til skriftlig eksamen enn de fikk til standpunkt, og at det er liten forskjell mellom offentlige og private skoler, når alle private skoler ses under ett (Utdanningsdirektoratet 2013a).

Dette innledningskapittelet presenterer de tre problemstillingene som rapporten skal besvare, og setter opp et teoretisk rammeverk som siden brukes til å fortolke de praksiser og rutiner vi finner ved skolene. I tillegg har vi også en innledende drøfting av karakterer, både i forhold til diskusjonen om avvik mellom standpunkt-karakter og eksamens-karakter, men også i forhold til regelverket, til troverdighet og til forskjeller mellom ulike fag. Dessuten avgrenser og definerer kapittelet hvilke typer av private skoler vi ser på i denne studien.

1.1 Avviksproblematikk og karakterer

Bakgrunnen for prosjektet er en mediedebatt om påståtte forskjeller mellom fastsettelse av standpunktkarakterer i offentlige og private skole, som i første rekke var en debatt drevet av Aftenposten. Denne debatten ble det henvist til i beskrivelsen av formålet med forskningsoppdraget fra Utdanningsdirektoratet. Aftenposten (05.02.2013) viste til at det var avvik mellom karakterer norske elever i videregående opplæring fikk til standpunkt og til eksamen i norsk hovedmål på program for studiespesialisering. Aftenposten poengterte særlig at avviket kunne knyttes til hvorvidt elevene går ved en offentlig eller privat videregående skole og fulgte dermed opp en debatt Dagens Nyheter i Sverige har ført om systematiske forskjeller i karaktersetting mellom svenske friskoler og offentlige skoler (Aftenposten 15.02.2013). Debatten i Sverige er imidlertid ikke helt lik den norske debatten, siden man der har vært mer fokusert på *betygsinflation*, som er målt ved at karakterene går opp selv om elevenes kunnskapsnivå er lavere enn det var et tiår tidligere (Vlachos 2010). I Sverige har man heller ikke separat eksamenskarakter, og dermed blir det å måle avvik mellom standpunktkarakter og eksamenskarakter i samme fag noe særnorsk.

Generelt sett er avvik mellom karakterer gitt til standpunkt og eksamen ikke en ny problemstilling i norsk grunnopplæring (Prøitz & Borgen 2010, Galloway et al 2011, Opheim et al 2010). Problemstillingen handler dels om ulike praksiser for standpunktkaraktersetting, men det handler også om i hvilken grad det er hensiktsmessig å sammenlikne disse karakterene ettersom de er basert på ulikt vurderingsgrunnlag og per definisjon skal måle ulike sider ved elevens oppnådde kompetanse. Samtidig er det et faktum at både skoleeiere, skoleledere og forskere sammenlikner standpunkt og eksamen for ulike formål. Hos skoleeiere og blant skoleledere blir avvik mellom karakterer brukt som kvalitetsindikator på regionalt nivå, der oppfatninger om at for store avvik mellom standpunkt og eksamen kan tyde på at lærere og skoler enten vurderer elever for «snilt eller for strengt» sammenliknet med karakterer gitt til eksamen. Eksamen blir brukt som rettesnor og har på denne måten en kalibrerende funksjon (Prøitz & Borgen 2010). Praksis er imidlertid ulik og oppfatningene om hvorvidt karakterene for standpunkt og eksamen ideelt sett skal være like eller hvor store avvik som er akseptable, varierer. Spørsmålet om hva god eller riktig praksis er, er sterkt knyttet til hvilken status og funksjon standpunkt og eksamen er gitt i det norske opplæringssystemet formelt og uformelt, samtidig som det er gitt få klare retningslinjer om dette fra myndighetenes side.

1.2 Karakterer og regelverk

Standpunktvurdering skal gi informasjon om den kompetansen eleven har oppnådd i faget ved avslutningen av opplæringen og i fag i videregående opplæring. Dette er fastsatt i forskrift til opplæringsloven, og med § 3-3 understrekes det at: Grunnlaget for vurdering i fag er dei samla kompetansemåla i læreplanane for fag slik dei er fastsette i læreplanverket, jf. § 1-1, § 1-3.

Videre heter det at «I og med at standpunktkarakteren skal vise elevens samlede kompetanse i faget skal elevens kompetanse vurderes på et så bredt grunnlag som mulig. Kompetansemålene skal ses i sammenheng, og det er ikke anledning til å vurdere kompetansen bare på grunnlag av noen utvalgte kompetansemål» (www.utdanningsdirektoratet.no 28.05.13). Vurderingen skal også gi muligheter for forbedring og ses i lys av den underveisvurdering som eleven har gjennomgått i løpet av opplæringen før endelig standpunktkarakter settes ved slutten av opplæringen. Det presiseres i rundskriv til forskrift til Opplæringsloven at: «I underveisvurderingen handler det om å bidra til utvikling og faglig progresjon, og i standpunktvurderingen handler det om hvor langt eleven har nådd» (Merknad til paragraf 3-18).

I forskrift til Opplæringsloven § 3-25 heter det om eksamen at «Eksamen skal vere i samsvar med læreplanverket» og at «Eksamen skal organiserast slik at eleven eller privatisten kan få vist kompetansen sin i faget. Eksamenskarakteren skal fastsetjast på individuelt grunnlag og gi uttrykk for kompetansen til eleven eller privatisten slik denne kjem fram på eksamen.» Gjennomføring av eksamen er også fastsatt i regelverket med opptil til fem timer skriftlig eksamen og opptil 30 minutter

for muntlig eksamen. Samlet sett så angir dette at eksamen er en langt smalere prøving av elevens kompetanse enn standpunkt ideelt sett skal være.

Regelverket angir som vist at det er klare forskjeller i det som ligger til grunn for fastsettelse av standpunktkarakterer og eksamenskarakterer selv om begge vurderingsformer skal vurdere eleven ut fra kompetansemålene for fag (§3-3).

1.3 Karakterer og fag

På den annen side er det praksiser og tradisjoner knyttet til standpunkt og eksamen i ulike fag som gjør at det kanskje kan forventes en større sammenheng mellom standpunkt og eksamen i enkelte fag, som for eksempel i norsk.

Lange tradisjoner for utvikling av sentralt gitt eksamen i eksamensnemnder med tette koblinger mellom fageksperter, lærere i norsk, kursing av lærere for ekstern sensurering og ikke minst omfattende bruk av eksempeloppgaver med tilhørende sensorveiledninger i undervisning, underveisvurdering og som grunnlag for standpunktkaraktersetting gjør at avvik eller samsvar mellom karakterer i faget norsk kan handle om hvor tett på ulike skoler er til disse tradisjonene og praksisene. Det er dermed et spørsmål om i hvilken grad dette er noe som kan slå ulikt ut i læreres prosedyrer for karaktersetting i ulike skoler.

Spørringene som NIFU gjennomfører på oppdrag fra Utdanningsdirektoratet bekrefter dessuten at det er ulike oppfatninger blant skoleledere om hvordan vurdering av kompetanse i fag skal foretas og om innsats og aktivitet skal være en del av vurderingsgrunnlaget (Vibe og Evensen 2009; Vibe 2012). I spørringene fra våren 2014 fremgår det at 42 prosent av skolelederne sier seg helt eller litt uenige i at elevenes innsats og aktivitet i klasserommet bør vurderes for seg og ikke være en del av grunnlaget for vurdering av elevenes kompetanse i fag (Vibe & Lødding 2014).

I invitasjonen til forskningsoppdraget fra Utdanningsdirektoratet vises det til tidligere dokumentasjon på at standpunkt ofte settes på et mer sammensatt grunnlag enn bare fag (Prøitz og Borgen 2010, Galloway et al 2011). Dette er en velkjent problemstilling i internasjonal forskning om vurdering og læring (assessment), som ofte forklares med et manglende samsvar mellom læreres og skolemyndigheters oppfatninger om hva som utgjør et godt grunnlag for vurdering av prestasjoner (Angelo 1998, Barnes 1985, Brookhart 1991, 1993, 1999, Manke & Loyd 1990, Stiggins et al 1989, Eggen 2004, Wiliam 1996, Resh 2010, Melograno 2007).

1.4 Offentlige og private skoler

Mesteparten av de videregående skolene i Norge er offentlige. I følge Utdanningsspeilet 2013 var det i høsten 2012 81 private skoler av totalt omtrent 430 videregående skoler i Norge. Antallet skoler varierer fra år til år, fra 2011/12 til 2012/13 var det 15 skoler som hadde blitt lagt ned, samtidig som 8 nye skoler har blitt opprettet. Det er størst endringene ser vi blant private skoler: 11 av skolene som hadde blitt lagt ned, og 7 av de nyopprettede skolene var private (Utdanningsdirektoratet 2013b). Dermed er det krevende å få en fullgod og dekkende oversikt over private videregående skoler. I tillegg er det forholdsvis få elever som går på private skoler, i perioden vi undersøker (2010-2013) er det 8,3 prosent av elevene som går på en privat videregående skole. Andelen elever som går på en privat skole har vært stabil i de fire årene vi ser på.

Det finnes forskjellige typer private skoler, og disse kan kategoriseres på ulike måter. Vi har valgt å kategorisere de private skolene i Norge i seks ulike typer:

- 1) Kristne videregående skoler
- 2) Internasjonale videregående skoler (herunder sorterer også Den tyske skolen og Den franske skolen)
- 3) Videregående skoler med alternativ pedagogikk (f.eks. Steiner og Montessori)
- 4) Yrkesskoler
- 5) Skoler som tilbyr toppidrett, estetiske fag eller naturbruk med sikte på studiekompetanse.
- 6) Allmenne private videregående skoler.

Allmenne private videregående skoler er skoler som tilbyr studieforberedende programmer. Noen av disse skolene vil dels tilby eksamensrettet privatistopplæring i enkeltfag og dels ordinære tilbud med eksamensrett, med alle tre klassetrinn og organisert på samme måte som offentlige skoler. I vår studie er det den siste typen tilbud som vil være aktuelt å se på, fordi det her normalt vil foreligge både standpunkt karakterer og eksamens karakterer, noe som er en forutsetning for våre analyser. Flere av disse skolene ble godkjent på grunnlag av Friskoleloven før omlegging til ny privatskolelov i 2005. Det er også i denne gruppen vi gjenfinder skoler som har reklamert med «karaktergaranti», men da gjerne knyttet til den delen av tilbudet som gjelder privatisteksamen i enkeltfag og som ikke inngår i våre analyser.

I analysene i denne rapporten sammenligner vi offentlige skoler med tre ulike typer private skoler, som alle gir tilbud rettet mot studiekompetanse. Kristne videregående skoler, spesialiserte skoler (som tilbyr toppidrett, musikk eller naturbruk) og allmenne private skoler, det vil si skoletype 1, 5 og 6 i oversikten ovenfor. Grunnen til at vi tar utgangspunkt i disse tre typene skoler er at de er sammenlignbare med de offentlige skolene, mens det ikke gjelder for de internasjonale skolene og skoler med alternativ pedagogikk. Yrkesskolene har ikke eksamen i fag som er sammenlignbare med skoler som er rettet mot studiekompetanse og dermed ikke aktuelle for vårt formål. For skolene med alternativ pedagogikk har vi ikke tilgang til standpunkt karakterer og dermed har vi ikke noe grunnlag å gjøre sammenlignende analyser. De internasjonale skolene følger en annen læreplan i tillegg til den norske.

1.5 Karakterer og troverdighet

Et lands karaktersystem er avhengig av troverdighet for å ivareta sin funksjon som verktøy for seleksjon. Sentralt i dette står det at folk har tillit til systemet. Tillit er knyttet til om folk oppfatter retningslinjer, rutiner, prosedyrer og praksis for karaktersettingen som rettferdig og likeverdig. I et system der prosedyrer og/eller praksis for karaktersetting oppfattes som påvirket av tilfeldige, variable eller systematisk urettferdige faktorer, vil et system miste sin troverdighet, og folks tillit til systemet blir svekket.

Med tanke på at standpunkt karakterer er det som er mest avgjørende for elevers muligheter til videre utdannings- og yrkeskarriere, er det helt sentralt at prinsippet om en likeverdig og rettferdig vurdering i det norske opplæringssystemet holdes høyt uavhengig av om skolen er offentlig eller privat. Samtidig er det mye som tyder på at selve prosessen med å sette standpunkt karakter trenger veiledning og regulering fra sentralt hold.

De siste årene har det blitt satt i gang flere sentrale tiltak rundt vurdering og læring i Norge, og det har skjedd endringer i hvordan lærere, skoleledere og skoleeier arbeider med vurdering i Kunnskapsløftet (Aasen et al. 2012, Hodgson, Rønning & Tomlinson 2012). Over lengre tid er det gjennomført flere tiltak for å øke kompetansen på vurdering i norsk skole, som den nasjonale satsingen Vurdering for læring (2010-2014) der både offentlige og private skoleeiere deltar samt utforming av veiledninger med kjennetegn for måloppnåelse til støtte i arbeid med standpunkt karaktersetting og eksamen. Tiltakene er i første rekke rettet mot læreres arbeid med underveisvurdering og utvikling av

tolkningsfellesskap ved den enkelte skole, og er i begrenset grad tiltak rettet direkte mot læreres arbeid med standpunktkaraktersetting. Det er dokumentert at innsats og aktivitet tas med i karaktervurderingen (Thronsdén et al 2009, Prøitz & Borgen 2010), og at elevgruppers sosiale bakgrunn vurderes inn ved karaktersetting (Hægeland et al 2005). Wiborg et al. (2011) fant at foreldres utdanning har større innvirkning på standpunktkarakterer enn på eksamenskarakterer. Mye av forskningslitteraturen har lenge pekt på behovet for å øke vurderingskompetansen hos lærere for å bøte på praksis som ikke er i tråd med anbefalinger og regelverk (Brookhart 2013). Prøitz og Borgen (2010) fant imidlertid at lærere i stor grad forsøkte å etterleve nye reguleringer for standpunktvurdering og ny læreplan som følge av implementeringen av Kunnskapsløftet, men at flere strevde med å få det til innenfor fagenes ulike rammer. Dette er i tråd med nyere forskning som viser til at lærere i ulike fag og på ulike trinn anvender en miks av ulike typer informasjon på en logisk og fornuftig måte (Brookhart 2013). Denne forskningen på karaktersetting viser også til at ulike kontekster krever ulike tilnærminger til karaktersetting og antyder at det ikke nødvendigvis er behov for mer kompetanse om vurdering lærere og skoler trenger, men at det er et behov for mer støtte og veiledning i arbeid med karaktersetting (Brookhart 2013, Wyatt-Smith & Klenowski 2013)

Dette indikerer at mange faktorer virker inn på hvordan lærere setter karakterer. Det er et spørsmål om i hvilken grad avvik i karakterstatistikk handler om forskjeller mellom typer av skoler (offentlig/privat hvor privat står for ulike typer skoler), eller om avvik kan være knyttet til andre faktorer.

1.6 Problemstillinger

Utdanningsdirektoratet ønsker å få belyst følgende problemstillinger:

1. Er det større avvik mellom standpunktkarakter og eksamenskarakter når private og offentlige skoler sammenlignes?
2. Hvilke rutiner og retningslinjer ligger til grunn for vurderingen ved de ulike skolene?
3. Er det forskjeller i rutiner og retningslinjer ved de ulike skolene?

Oppdragets tematikk handler om to hovedelementer slik vi har tolket det. For det første handler det om avvik mellom standpunktkarakterer og eksamenskarakterer, og dette temaet innbefatter også praksis og rutiner for standpunktkaraktersetting. For det andre dreier det seg om hvorvidt grad av avvik systematisk kan knyttes til om skolen er offentlig eller privat.

I Utdanningsdirektoratets konkurransegrunnlag heter det at prosjektet skal være en evaluering av karakterpraksis i offentlig og private videregående skoler, og at det er ønskelig å få gjennomført en evaluering av fremsatte påstander i mediene om forskjeller i standpunktkarakterer mellom offentlige og private skoler. En evaluering forutsetter at det finnes et regelverk eller fastsatte retningslinjer for hvordan forholdet mellom karakterer til standpunkt og eksamen skal være. Slik vi ser det finnes det et omfattende regelverk og tolkninger av dette tilgjengelig for hvordan standpunktkarakterer skal settes og hvordan eksamen skal gjennomføres, det er imidlertid uklart hvordan disse skal ses i forhold til hverandre utover implisitte antydninger og omtalt praksis.

Dette innebærer at dette prosjektet i første rekke kan karakteriseres som en kartlegging og undersøkende studie av forholdene rundt standpunktkaraktersetting i offentlige og private skoler i tråd med de problemstillinger og temaer som er nevnt i konkurransegrunnlaget og som er videre utviklet og fortolket i dette prosjektet.

1.7 Teoretisk rammeverk

Dette prosjektet omfatter flere elementer og det er nødvendig å anlegge et rammeverk som kan strukturere og fokusere arbeidet med datainnsamling og analyse. I og med at en sentral bakgrunnsvariabel er offentlig vs. privat skole blir det viktig å inkludere verktøy som kan identifisere eventuelle

særegenheter i praksiser og rutiner ved karaktersetting knyttet til skoletype, dette kommer i tillegg til modeller for mer generelle praksiser og rutiner knyttet til læreres standpunkt-karaktersetting.

For å fange opp eventuelle skoletypiske karaktertrekk er prosjektets teoretiske rammeverk inspirert av ideer utviklet av Spillane (2012). Han har studert bruk av informasjon/data i skoler (for eksempel prøveresultater og karakterdata) og identifisert ulike typer rutiner som helt sentrale for hvordan arbeid foregår og utvikler seg i skole som organisasjon. Spillane definerer rutiner som gjentatte og gjenkjennelige handlingsmønstre som involverer aktørene i skolen; de kjennetegnes av å strukturere skolehverdagens praksiser ved å ramme inn og fokusere samhandlingen mellom de ansatte, koordinerer arbeidet, reduserer konflikt og usikkerhet omkring hvordan ting skal gjøres, og de kan være en mekanisme for både endring men også konservering av status quo (Spillane 2012). Selv om hans teorier om "data use" primært er utformet for studier av praksis, for eksempel ved observasjon av hva som skjer i interaksjon lærere mellom, mener vi at hans begreper kan være formålstjenlige også i dette studiet basert på karakterstatistikk og intervjudata.

Forenklet beskrevet skiller Spillane (2012) mellom to former for organisasjonsrutiner; rutiner som design og rutiner i bruk. Den første omfatter nedtegnede og formaliserte rutiner, for eksempel i form av regelverk, retningslinjer eller veiledninger, beskrevet som intensjoner, i abstrakt, generell og skjematisk form. Rutiner i bruk handler om hvordan aktørene i skole etablerer bestemte praksiser for det som skal gjøres. Disse rutineene er ikke nedtegnede og kan være basert på designede rutiner eller helt andre rutiner vokst frem gjennom praksis. Spillane definerer dette som: Bestemte handlinger, gjennomført av bestemte personer, på bestemt sted og til bestemt tid – for å poengtere at det dreier seg om lokal praksis utformet av de som er involvert innenfor de kontekstuelle rammene de befinner seg i (2012:117 egen oversettelse).

Først og fremst vil vi benytte oss av Spillanes begreper om designede rutiner og rutiner i bruk for innrammingen av prosjektet. Dette lar oss fokusere på arbeidet med standpunkt-karaktersetting i form av lokale rutiner spesielt og passer således godt til å studere hva som skjer mellom designede rutiner (for eksempel formaliserte nasjonale og lokale retningslinjer for vurdering og andre verktøy for standpunkt-karaktersetting som foreligger skriftlig) og rutiner i bruk i den aktuelle skolen.

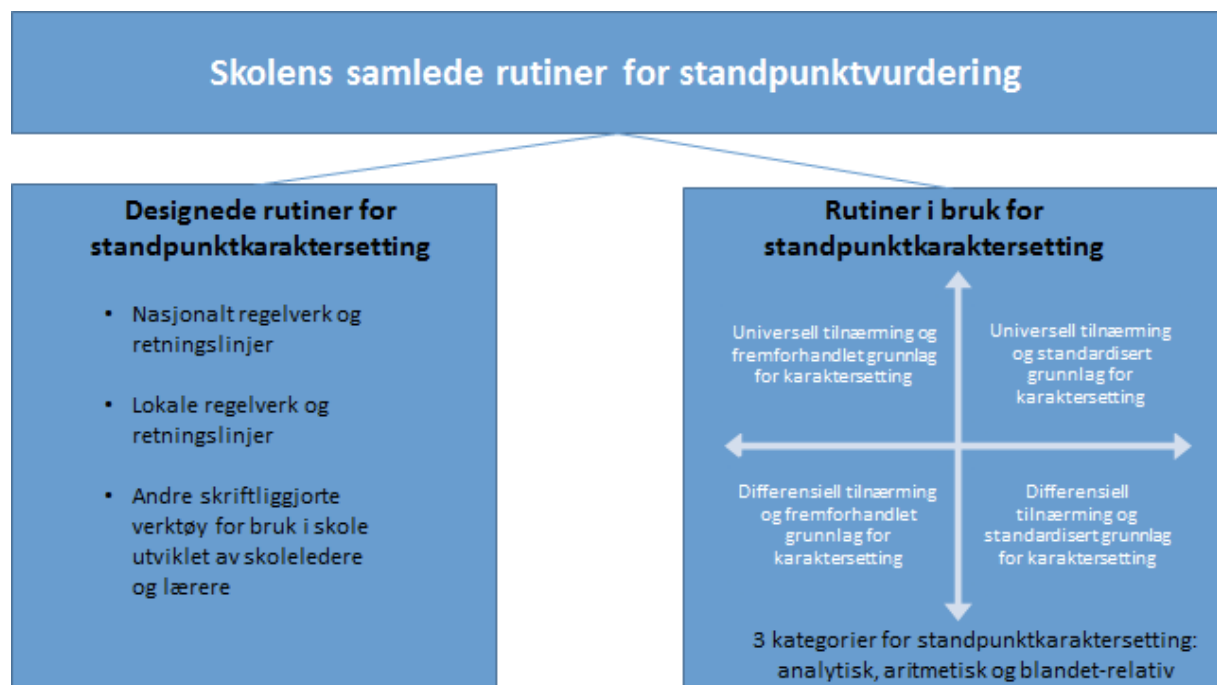
For å identifisere rutiner i bruk vil vi anvende en modell for læreres karaktersetting utviklet i Prøitz & Borgen (2010) og Prøitz (2013). I Prøitz & Borgen (2010) identifiserte vi ulike typer av grunnlag og verktøy lærere oppgir å bruke når de setter standpunkt-karakterer. Blant annet stadfestet vi at en tilnærming er å bruke poeng på prøver som avgir et bestemt karakternivå der tolkningsrommet oppfattes som lite, standardisert og gitt (standardisert grunnlag for vurdering). En annen tilnærming var å anvende ulike verktøy basert på definerte kjennetegn på måloppnåelse, vurderingsmatriser eller nedbrutte målformuleringer der tolkningsrommet ble beskrevet som større og åpent for stadig forhandling (fremforhandlet grunnlag for vurdering). Prøitz og Borgen (2010) identifiserte også to tilnærminger til rettferdighet i karaktersetting basert på internasjonal forskning og informantenes beskrivelser av praksis (universell og differensiell vurdering). Universell karaktergivning handler om å behandle alle elever likt, mens differensiell karaktergivning handler om å behandle elever ulikt etter hvorvidt de er oppfattet som svake eller sterke elever. Modellen har flere fellestrekk med tre kategorier av karaktersetting identifisert i en svensk studie av læreres karaktersetting og forholdet til deres nasjonale prov.

- Analytisk: Elevenes kunnskapsprofil analyseres i forhold til målene i kursplanen og kriterier for karaktersetting. De nasjonale prøvene brukes som en blant flere kilder for informasjon om elevens måloppnåelse.
- Aritmetisk: Elevenes karakterer beregnes på basis av prøveresultater og karakterer på ulike oppgaver gitt i løpet av opplæringen. Prøven og oppgavene er utformet for å vise elevenes kunnskaper i forhold til målene i kursplanen.

- Blandet-relativ: Elevenes karakter bestemmes ut fra en blanding av vurderinger, prøveresultater, lærerens allmenne inntrykk av elevens kunnskapsnivå, tilstedeværelse, holdninger og aktivitet i undervisningssammenheng. Fokus er på lærerens vurdering av elevens allmenne kunnskapsnivå i faget. Referanserammen er lærerens samlede undervisnings- og karaktersettingserfaring med ulike grupper i faget. Denne praksisen er basert på lærerens skjønn og følelse av hvilken karakter eleven skal ha. De nasjonale prøveresultatene brukes for å bekrefte lærerens vurdering, og kan være utslagsgivende når læreren anser eleven for å være i vippeposisjon mellom to karakterer. I andre tilfeller er det prøveresultatene som er utgangspunktet for karaktersettingen i faget innenfor denne modellen. (Skolverket 2005: 124, egen oversettelse)

Vi vil bygge videre på funnene fra Prøitz og Borgen (2010) i kombinasjon med de tre presenterte svenske kategoriene i denne studien for å få et så dekkende rammeverk som mulig.

Samlet sett kan det teoretiske rammeverket for prosjektets analyse og datainnsamling fremstilles i en figur som viser at skolens samlede rutiner for standpunktvurdering omfatter rutiner som er skriftliggjorte og formaliserte (designede rutiner) og rutiner anvendt i praksis (rutiner i bruk):



Figur 1.1: Illustrasjon av teoretisk rammeverk for prosjektet (inspirert av Spillane 2012)

I dette prosjektet har vi operasjonalisert det teoretiske rammeverket i utviklingen av verktøy for datainnsamling og i analysen. Dels er dette gjort ved å utforme intervjuguider tematisk og med konkrete spørsmål om praksis i forhold til kjente designede rutiner som: regelverk, veiledninger, eksamensoppgaver og sensorveiledninger. Vi har også stilt spørsmål om deltakelse på sensorskolering, annen kursvirksomhet knyttet til karaktersetting og lokalt utviklet materiale eller pålegg om bestemte praksiser for karaktersetting. Videre har vi stilt mer åpne spørsmål om hva lærere gjør når de setter karakter og hva de tenker om dette. Dette er også spørsmål som er fulgt opp i intervjuene med elever. Både de mer åpne spørsmålene og i forlengelse av de mer konkrete spørsmålene har informantene gitt informasjon om praksiser knyttet til både designede rutiner og praksisrutiner.

I analysen har vi søkt å identifisere beskrivelser av kjennskap til, bruk av eller henvisninger til regelverk, formaliserte retningslinjer eller veiledninger, eller også formelle beskjeder og pålegg fra rektor i møte med lærerkollegiet. I vår analyse regner vi også kurs og opplæring, herunder

sensoropplæring eller den nasjonale satsningen *Vurdering for læring* som eksempler på slike designede rutiner. I analysen henviser praksisrutiner eller rutiner i bruk til praksiser som kan springe ut av de designede rutinene, og være en videreutvikling av disse som enkeltlærere eller grupper av lærere (for eksempel norsklærere som arbeider på samme trinn) har skapt over tid som en støtte for arbeidet. Dagligrutinene kan imidlertid også ha oppstått mer eller mindre av seg selv, det vil si som løsning på et behov eller et problem som en eller flere lærere har erfart. Vi har også søkt å identifisere beskrivelser av typer av praksiser som kan karakteriseres som mer aritmetisk eller analytisk og strategier for rettferdig vurdering i form av universell eller differensiell behandling. I analyse av intervjudata handler det alltid om informanternes beskrivelser av praksiser og både informant og intervjuers fortolkning av disse beskrivelsene. Dette innebærer at det må tas forbehold om at disse kan avvike fra det som faktisk skjer. Likevel mener vi at våre informanternes beskrivelser kan gi et inntrykk av hvilke praksiser som råder. Valg av gruppeintervjuer i dette prosjektet gir også et godt grunnlag for dialog, diskusjon og korreksjon både mellom informanter og mellom informanter og intervjuer som sikrer et solid datagrunnlag.

En mer maktkritisk innfallsvinkel til vurderingspraksis kan utledes av Bernsteins kodeteori og hans begreper innramming og klassifisering i pedagogisk praksis (Bernstein 1990). Riksaasen (1994) anvender Bernsteins begrepsapparat i skillet mellom synlig versus usynlig pedagogikk. Førstnevnte karakteriseres ved overføring av spesifikke ferdigheter og faste kriterier for evaluering av resultater, i motsetning til sistnevnte hvor det ikke er spesifikke ferdigheter og klare vurderingskriterier som gjelder, men heller selve læringsprosessen (å lære å lære) som står i sentrum. Nøkkelbegrepet for førstnevnte er *disiplin* mens det for sistnevnte er *lek*. I usynlig pedagogikk vil lærerens kontroll over eleven være implisitt, læreren tilrettelegger en kontekst som eleven forventes å utforske, og eleven forventes dessuten å selv styre sine aktiviteter. Poenget her er at pedagogikken er usynlig for den som er under opplæring.

Denne dikotomien synlig versus usynlig pedagogikk er ikke sammenfallende med de dimensjonene vi har skissert i det foregående, og den er også langt mer omfattende (med forskjeller i grensesetting mellom disipliner, maktstrukturer, språkkoder og kommunikasjonskoder) enn vi gjengir her. Kanskje kan en også stille spørsmål ved om lek er et meningsfullt begrep om pedagogikk i Vg3 studiespesialisering. Det som likevel synes relevant er Bernsteins vektlegging av at hovedforskjellen mellom synlig og usynlig pedagogikk ligger i måten evalueringen foregår på samt i hvilken grad en spesifiserer kriteriene for å tilegne seg ferdigheter. Riksaasen konkretiserer Bernsteins teoretiske bidrag slik: «Jo mer diffuse kriterier, jo mer usynlig pedagogikk. Jo mer eksplisitte overføringskriterier, jo mer synlig pedagogikk. Det er viktig å merke seg at ved en usynlig pedagogikk er det større deler av elevens personlighet som evalueres enn ved en synlig pedagogikk» (Ibid: 50).

Denne grensedragningen mellom ferdigheter på den ene siden og personlighet på den andre siden som ulike omdreiningspunkt for vurdering, har verdi i tolkningen av elevenes erfaringer, som presentert i kapittel 5. Den gir også dybde til lærernes etos slik det formidles i kapittel 4 om at det er elevenes kompetanse de bestreber seg på å vurdere og nettopp ikke elevenes personligheter.

2 Metodisk tilnærming

I dette prosjektet har vi lagt opp til et forskningsdesign der vi kombinerer kvantitative og kvalitative analyser. Et sentralt element i designet er at de kvantitative analysene av avvik mellom standpunkt-karakter og eksamenskarakter ved offentlige og private videregående skoler er gjennomført i forkant av de kvalitative undersøkelsene. Dermed er utvalget av skoler for den kvalitative delen av studien gjort på grunnlag av de kvantitative analysene. Hensikten har vært å inkludere både offentlige og private videregående skoler med relativt sett store og relativt sett små avvik mellom standpunkt- og eksamenskarakterer. I dette kapitlet redegjør vi for utvalg, metodisk fremgangsmåte og gjennomføring av de datainnsamlinger som inngår i de ulike delene av prosjektet.

2.1 Statistiske analyser av karakterer på skolenivå

Den første delen av dette prosjektet besto i å gjøre kvantitative analyser av karakterforskjeller mellom offentlige og private videregående skoler, der vi også i noen analyser skiller mellom ulike typer private skoler. Datasettet vi har tilgang til er fra Utdanningsdirektoratet, og disse dataene inneholder forholdsvis lite informasjon om eleven, kun kjønn og utdanningsprogram er variabler som finnes på individnivå, i tillegg til informasjon om standpunkt- og eksamenskarakter i ulike fag. Vi har dermed ingen mulighet til å kontrollere for sosial bakgrunn, etnisk opphav eller tidligere skoleprestasjoner i våre analyser. Delrapporten som omfatter analysene av de kvantitative dataene ble presentert for oppdragsgiver i desember 2013, og finnes i sin helhet som vedlegg 3 til rapporten. Hovedfunnene i delrapporten er oppsummert i kapittel 3, og der vises også noen bivariate sammenhenger grafisk.

Det er flere former for statistiske analyser som kunne anvendes for sammenlikning av karakterpraksis i private og offentlige videregående skoler. En mulighet er å standardisere karakterene, men vi har valgt i våre analyser å heller se på standpunktkarakter og eksamenskarakter og sammenligne gjennomsnittlig avstand mellom standpunktkarakter og eksamenskarakter for de elever som har begge karakterene i et fag. Oversikt over hvor mange personer dette omfatter finnes i vedlegg 3 og i tabell 3.1 i kapittel 3.

En utfordring i dataene er at det kun er 8,3 prosent av elevene i materialet som har gått på en privat skole. Ettersom antallet private skoler er lavt, og mange av dem er små, vil det dermed være relativt stor statistisk usikkerhet knyttet til resultater på skolenivå. Dette vil vi vise i analysene ved ikke bare å vise gjennomsnittet men også vise konfidensintervall i alle de bivariate analysene. Alle figurer som vises i kapittel 3 er basert på bivariate observerte forskjeller, men i vedlegg 3 finnes også regresjonsanalyser som kontrollerer for flere variabler samtidig.

Siden det er mange forhold som kan påvirke karakterene, har vi brukt lineær regresjon for å kontrollere for flere variabler samtidig. I regresjonsanalysene vil de private skolene deles inn i tre ulike

kategorier, slik det er beskrevet i kapittel 1.4. Det vi først og fremst ville se etter i en slik analyse, er om private skoler, eller bestemte typer private skoler, har en annen karakterutvikling fra standpunkt til eksamen enn offentlige skoler, også når vi kontrollerer for sammensetningen av elevmassen.

Prosjektet tar utgangspunkt i karakterdata for fire avgangskull, nemlig årene 2010 til 2013. Å analysere endringer fra år til år for ulike typer skoler vil i seg selv være en viktig oppgave fordi dette kan gi en indikasjon om hva som er normal variasjon, for eksempel i forskjellen mellom standpunkt- og eksamenskarakterer, og hva som må kunne sies å være mer systematiske forskjeller, for eksempel mellom private skoler og offentlige skoler. En særlig viktig problemstilling vil være om det fokuset som ble rettet mot resultatene fra 2012 har ført til endringer i karakterpraksis i 2013. Nettopp derfor er det viktig å kunne skille mellom tilfeldig og systematisk variasjon.

2.1.1 Utvalg av skoler til kvalitative intervjuer

Siden det ikke finnes så mange private skoler tok vi utgangspunkt i disse da vi gjorde utvalget av caseskoler. Et utvalgsriterium var at alle de tre typene private skoler (kristen, spesialisert og allmenn) skulle være representert blant de fire private skolene vi skulle besøke. I tillegg la prosjektets økonomiske rammer begrensninger som gjorde at alle skolene vi har besøkt er lokalisert på Østlandet, for å minimere reisekostnader. Det tredje kriteriet vi brukte var direkte relatert til den kvantitative delen av prosjektet, ved at vi brukte data til å se på forskjeller i avvik mellom enkeltskoler.

Vi ønsket både å ha skoler som har stort og lite avvik mellom standpunkt- og eksamenskarakter representert blant skolene vi besøkte. Det er få av de offentlige skolene som har stort avvik, men det er noen av de offentlige skolene som har et noe høyere avvik enn andre offentlige skoler og som også har hatt det over tid.

Stort avvik ble definert som et avvik på mer enn -1,0, mens lite avvik var de skolene som hadde et avvik som var mindre enn -0,4. Ved gjennomgangen av avvik per skole så vi at det i praksis ikke er noen skoler som har et positivt avvik, det vil si der eksamenskarakter er høyere enn standpunkt karakter, men det finnes skoler som har svært små avvik. I tillegg så vi på om avviket var stabilt over tid, for de fire årene vi har tilgang på data om.

Tabell 2.1 viser hvordan avviket mellom standpunkt karakter og eksamenskarakter så ut ved de fire private og de fire offentlige skolene som har vært caseskoler. For begge typer skoler har vi en skole med høyt eller relativt avvik, en skole med endring over tid (fra høyt til gjennomsnittlig eller lavt) og to skoler med stabilt lavt eller gjennomsnittlig avvik. Med andre ord matcher caseskolene ved private og offentlige skoler hverandre forholdsvis godt.

Tabell 2.1: Oversikt over caseskolenes eierform samt avvik og utvikling i forholdet mellom standpunkt karakterer og eksamenskarakterer

Skole 1	Skole 8	Skole 2	Skole 5	Skole 6	Skole 4	Skole 3	Skole 7
Privat	Privat	Privat	Privat	Offentlig	Offentlig	Offentlig	Offentlig
Stabilt høyt avvik	Endring fra høyt til gjennomsnittlig avvik	Relativt stabilt lavt avvik	Relativt stabilt lavt avvik	Relativt høyt stabilt avvik	Endring fra høyt til lavt avvik	Stabilt gjennomsnittlig avvik	Stabilt lavt avvik

2.2 Kvalitative datakilder: Intervju og dokumentanalyse

I dette prosjektet er det benyttet flere typer kvalitative metoder. Vi har valgt en kombinasjon av dokumentbasert analyse, kvalitative epost-intervjuer med skoleleder/rektor, gruppeintervju med lærere og gruppeintervju med elever.

2.2.1 Dokumentbaserte analyser

Utgangspunktet for å gjennomføre standpunktvurdering ligger i regelverket for grunnopplæringen og vi ser det som viktig å inkludere sentralt fastsatt regelverk, retningslinjer og veiledninger i datamaterialet i denne undersøkelsen i en dokumentbasert analyse. Videre har vi etterspurt og samlet inn materiell som skal støtte opp under arbeidet med standpunktvurdering både på skoleleder- og lærernivå av ulik type. Denne typen datamateriale inngår i det teoretiske rammeverket som såkalt designede rutiner (jfr. modell). Dokumentene bidrar til å skape en ramme for tolkningen av lærer- og elevintervjuene, og styrker muligheten for å gjøre sammenligninger på tvers av skoler.

2.2.2 Epost-intervju

I epost-intervjuer (Meho 2006, Opdenakker 2006, Burns 2010) med skoleleder/rektor har vi stilt spørsmål om rutiner og retningslinjer for standpunktvurdering og hvordan de forholder seg til eksamen ved den enkelte skole. E-postintervjuer har den fordel at man kan få svar på konkrete spørsmål uten å forårsake nevneverdige forstyrrelser av arbeidsdagen til den som skal svare på spørsmålene, og det at svarene er skriftlige, sparer tid for forskeren. Det gjør det også lett for avsender og mottaker å dokumentere det som kommer frem i intervjuet. På den annen side gjør formatet at man må begrense antall spørsmål i et e-postintervju, og i motsetning til en samtale på telefon eller ansikt til ansikt, må man forvente at svarene blir forholdsvis korte.

Derfor har vi brukt e-postintervju med skoleledere for å få kjennskap til skoleledelsens arbeid med vurderingsrutiner, herunder også igangsatte prosjekt for endring av vurderingsrutiner, og vi har fått skoleledernes synspunkter på dette arbeidet. Epost-intervjuene hadde form av tre åpne spørsmål som skolelederne svarer skriftlig på.

Da vi mottok svarene fra rektorene, så vi imidlertid at enkelte av dem hadde svart utfyllende, mens andre var svært knappe. Noen av rektorene har utførlig vist til hva slags satsninger eller prosjekter skolen deltar for å bedre arbeidet med vurdering, mens andre beskriver rutinemessig arbeidet med vurdering knyttet til forskrift og behandling av elever som står i fare for å stryke eller å få karakteren «ikke vurdert». Vi ser i ettertid at selv om vi gikk ut fra at spørsmålene til rektorene var enkle å besvare og ville gi grunnlag for sammenligning på tvers av skoler, ble dette i mindre grad resultatet. I denne rapporten har vi derfor bare brukt svarene på e-post-intervjuer med rektor for å registrere hvorvidt skolen deltar i en satsning eller har et prosjekt for vurdering, men feilkilder i denne registreringen kan ha oppstått. Feilkildene består i at selv om mange rektorer har fortalt om slike satsninger eller programmer, kan vi ikke utelukke at rektorene som ikke har nevnt slike forhold i e-posten, likevel jobber på en skole hvor det foregår slikt arbeid knyttet til en satsning eller prosjekt. Vi innser at våre spørsmål har ikke vært godt nok formulert til å fange opp dette.

2.2.3 Gruppeintervju

Intervjuene ble gjennomført i løpet av våren 2014 på åtte utvalgte skoler (fire offentlige og fire private). Alle de utvalgte skolene befinner seg på Østlandsområdet og er valgt ut etter kriterier som eierform, type skole, grad av avvik mellom standpunkt karakterer og eksamens karakterer og geografisk lokalisering.

Det er gjennomført åtte gruppeintervjuer med lærere og åtte gruppeintervjuer med elever. Vi har snakket med til sammen 32 lærere, hvorav halvparten av lærerne hadde erfaring med å sette standpunkt karakter i norsk og den andre halvparten hadde erfaring med å sette standpunkt karakter i matematikk. Lærerne dekker et bredt spekter hva angår utdanning, fagbakgrunn, ansiennitet og erfaring. Kjønnsfordelingen blant lærerne har vært godt balansert. Videre har vi snakket med 31 elever, 17 jenter og 14 gutter på de til sammen åtte skolene. Gruppene bestod av 3-5 elever og alltid av både jenter og gutter. Elevene var rekruttert til undersøkelsen av skolen, etter vårt ønske om at de skulle gå i tredje trinn og være fylt 18 år. I flere av gruppene ga elevene inntrykk av at de var høytpresterende elever gjennom sine uttalelser om karakterer. I to av gruppene var rekrutteringen mer tilfeldig. I det ene tilfellet var de som først var invitert, blitt opptatt på annet hold, og som erstatning for

disse ble andre elever rekruttert direkte fra gangen. I det andre tilfellet var enkelte forespurt, og andre kom til for å fylle opp for de som ikke var på skolen den aktuelle dagen sent i terminen. I noen av intervjuene kom det for en dag at elevene kjente hverandre fra tidligere trinn, inkludert fra ungdomsskolen, mens dette ikke var tilfelle i andre intervjuer.

Hvert gruppeintervju varte cirka 45 minutter. Ettersom fokusgruppeintervjuer er en krevende intervjuform, har samtlige intervjuer vært gjennomført med to erfarne forskere. Dette er en intervju praksis vi har gode erfaringer med fra tidligere prosjekter (bl.a. Prøitz & Borgen 2010; Kårstein & Carlsten 2013) ved at det sikrer en effektiv og etterrettelig datainnsamling og databehandling. Intervjuene ble gjennomført ved at en forsker hadde ansvar for å stille spørsmålene, drive samtalen, fange opp og få utdypet forskjellige poeng underveis, mens en annen forsker hadde ansvar for å skrive ned det som ble sagt. Samtalene ble også tatt opp på bånd og er senere transkribert.

Intervjuene har vært organisert rundt en såkalt semistrukturert intervjuguide inndelt i temaer og pre-formulerte spørsmålsstillinger. En slik intervjuguide er en samling spørsmål ordnet i en rekkefølge som vil være et utgangspunkt for samtalen, men der noen spørsmål kan vise seg viktigere enn andre samtidig som enkelte spørsmål ikke vil bli stilt, og forskeren har mulighet for å hoppe mellom spørsmål og tema for å sikre flyt i samtalen. Dette utgangspunktet sikrer at de samme temaene blir berørt i hvert intervju, men tillater også en dialogbasert intervjuform der temaene blir berørt naturlig gjennom samtalsgang (Kvale & Brinkmann 2009). Intervjuguidene følger denne rapporten som vedlegg (1 og 2).

I utgangspunktet ble deltakerne orientert om at det ikke finnes noen riktige svar, samt at det ikke var noe formål at de skulle nå enighet om hvordan spørsmålene burde besvares. I ethvert fokusgruppeintervju vil det finnes mer og mindre aktive deltakere. Det hender at en deltaker tar mye ansvar for å holde samtalen i gang, mens andre ser ut til å vurdere det som lite hensiktsmessig å gjenta det som allerede har vært sagt når de ikke har annet å tilføye. Andre deltakere kan demonstrere iver etter å se et forhold fra en annen vinkel og tilføre nye momenter til samtalen. Utfordringen for moderatoren i intervjuet er å få alle i tale gjennom å stille direkte spørsmål, oppsummere det som har vært sagt og fritte ut alternative synspunkter. Dynamikken vil aldri kunne bli fullt ut forutsigbar. I en skolekontekst vil deltakerne kjenne hverandre, og de kan ha egne felles referanser. De vil dessuten ha ulike forventninger til hverandre som det er vanskelig å overskue.

Selv om deltakerne oppfordres til å gi uttrykk for det når de ikke er enig med siste taler, kan et gruppeintervju ofte vise seg å dreie mot en enighet, eventuelt en enighet om hvordan erfaringene varierer. Synspunkter som hevdes tidlig og med styrke, vil kunne prege samtalen. En kan også observere hvordan erfaringer og argumenter prøves ut og møter motstand eller får tilslutning fra andre deltakere. Alt dette tilsier at de synspunktene som fremmes, til en viss grad kan være situasjonsbestemte og dessuten preget av den enkelte bidragsyters omdømme, formuleringsevne eller forventning om tilslutning fra de andre. Vi mener likevel at intervjuene har gitt innsikt nettopp gjennom deltakernes eksplisering av synspunkter og argumenter samt graden av (u)enighet fra andre deltakere. Når det gjennom en serie på åtte intervjuer ved forskjellige skoler etter hvert begynner å komme frem betraktninger som forskerne har hørt før, kan vi si at datatilfanget er i ferd med å nå en metningspunkt og vi kan anta at undersøkelsen har fanget opp de aller fleste relevante synspunktene på et spørsmål.

Skoleleder/rektor har vært prosjektets kontaktpunkt og hjulpet til med å finne informanter til gruppeintervjuene. Et generelt inntrykk fra intervjuene er at både lærere og elever har møtt temaene med interesse og engasjement. Flere lærere forteller om at tema for intervjuet berører utviklingsarbeid de selv har hatt i fagseksjonen over tid. Dette styrker antagelsen om at lærerne kan ha meldt seg selv som interesserte i å delta som informanter og at de ikke har vært spesielt utvalgte av skolens ledelse. I første rekke synes det som om informantene vi har møtt, i stor grad har stilt seg til rådighet fordi de fyller kriteriene (lærere med erfaring fra karaktersetting i norsk og matematikk) elever (i hovedsak 3. trinns elever) og fordi de har hatt mulighet til å møte til avtalt tidspunkt.

3 Kvantitative analyser, avstand mellom standpunkt- og eksamenskarakter

Dette kapitlet tar for seg prosjektets første problemstilling: *Om det er større avvik mellom standpunktkarakter og eksamenskarakter når private og offentlige skoler sammenlignes*. Kapitlet er en oppsummering av en delrapport som ble presentert for oppdragsgiver underveis i prosjektet. Delrapporten finnes i sin helhet som vedlegg 3 i denne sluttrapporten.

Som nevnt i kapitlet om metodisk tilnærming undersøker vi avvik mellom standpunktkarakter og eksamenskarakter ved å sammenligne gjennomsnittsavstanden mellom standpunktkarakter og eksamenskarakter for de elever som har vurdering fra både standpunkt og eksamen i samme fag. Det er flere metodiske utfordringer i å sammenligne karakterer, og det er særlig viktig at analysene baseres på tilstrekkelig store utvalg, både i forhold til antallet elever på ulike skoler og antallet elever som har standpunktkarakter og eksamenskarakter i et gitt fag.

Den første utfordringen er at det bare er en liten andel av elevene som går på en privat skole. I materialet vi bruker går 8,3 prosent av elevene på en privat skole, og dermed sammenligner vi en relativt liten gruppe av elever, i alt ca 3 300 hvert år, med det store flertallet som går i offentlige skoler. Vi skiller dessuten i analysene mellom tre ulike typer private videregående skoler: Kristne skoler, spesialiserte skoler og allmenne private skoler. Nærmere redegjørelse for hva som betegner de ulike typene skoler finnes i kapittel 1.

Den andre utfordringen er at det bare er relevant å gjøre analyser der vi har både standpunktkarakter og eksamenskarakter for de samme elevene. Med unntak av norsk hovedmål, der alle elever i videregående opplæring har eksamen, praktiseres det uttrekk til eksamen i alle andre fag, hvilket gjør at ikke alle elever som har hatt et fag også har en eksamenskarakter i det faget. I de fullstendige analysene (vedlegg 3) har vi tatt utgangspunkt i ni fag, fem skriftlige og fire muntlige. De ni fagene er norsk skriftlig hovedmål, norsk skriftlig sidemål, norsk muntlig, realfagsmatematikk, matematikk påbygg, rettslære, religion og etikk, historie studiespesialisering og historie påbygg. Tabell 3.1 viser en oversikt over hvor mange elever analysene er basert på, hvor stor andel av dem som går på en privat skole samt gjennomsnittlig standpunkt- og eksamenskarakter i de aktuelle fagene. Vi vil i dette kapitlet fokusere på analysene av norsk skriftlig hovedmål. Argumentet for dette er at de analysene utgjør hoveddelen av delrapporten og at det særlig er forskjellene i norsk skriftlig hovedmål det har vært fokus på i media. I tillegg finner vi samme mønster for de andre fagene med skriftlig eksamen, samt at det gjennomgående er slik at avstanden mellom standpunkt og eksamenskarakter er større i fag med skriftlig eksamen enn i fag med muntlig eksamen.

Tabell 3.1 Karakterdata for ni utvalgte fag i fire skoleår. Avsluttende vurdering. Antall elever

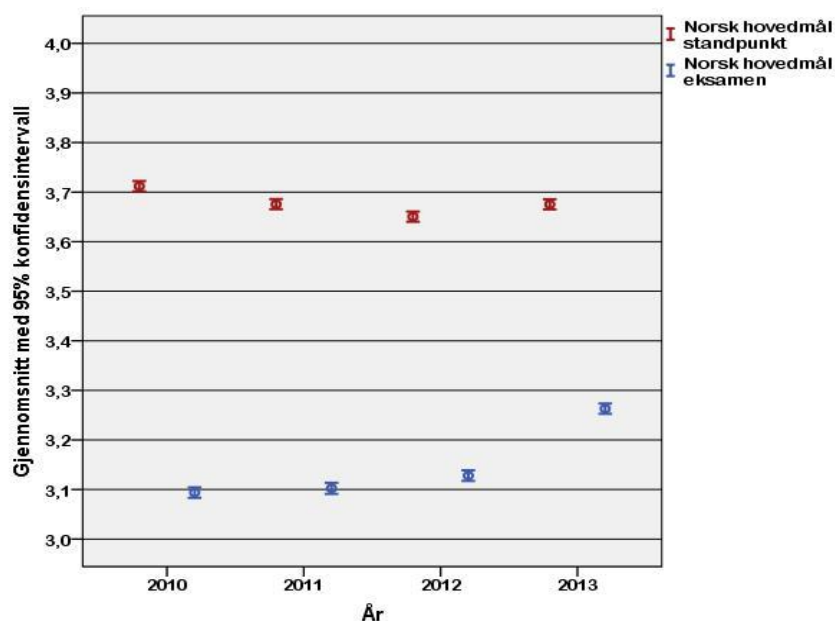
	Jenter	Alder	Privat	Standpunkt	Eksamen	Antall
NOR1211 + 1231 Norsk hovedmål	57 %	19,4	9 %	3,68	3,15	157 625
NOR1212 + 1232 Norsk sidemål	59 %	19,3	8 %	3,42	3,03	71 468
REA3024 Realfagsmatematikk	39 %	19,1	9 %	3,94	3,58	13 839
MAT1005 Matematikk påbygg	59 %	20,0	10 %	3,31	2,62	18 109
SAM3023 Rettslære	63 %	19,1	14 %	4,01	3,51	9 413
NOR1213 + 1233 Norsk muntlig	57 %	19,4	9 %	4,08	4,11	29 621
REL1001 Religion og etikk	56 %	19,1	10 %	4,13	4,34	15 918
HIS1002 Historie studiespesialisering	56 %	19,1	8 %	4,12	4,24	18 581
HIS1003 Historie påbygg	60 %	20,1	8 %	3,61	3,69	8 682
Alle ni fag	57 %	19,5	8 %	3,76	3,26	172 852

NB: tabellen er identisk med tabell 2.4 i vedlegg 3.

Med andre ord vil vi i dette kapittelet oppsummere analysene av norsk skriftlig hovedmål, og sammenligne avvik mellom standpunktkarakter og eksamenskarakter i offentlige skoler med tre typer private skoler: kristne skoler, spesialiserte skoler og allmenne private skoler. For de øvrige fagene viser vi til vedlegg 3.

3.1 Forutsetninger som ligger i problemstillingen

Før vi starter sammenligningen av standpunkt- og eksamenskarakterer, vil vi se nærmere på noen av forutsetningene for hele problemstillingen. Problemstillingen spør om det er avvik mellom standpunkt- og eksamenskarakterer, og den underliggende forutsetningen er dermed at det avviket skal være så lite som mulig. Her ligger med andre ord en forventning om at eksamen kan bli brukt som rettesnor og på denne måten ha en kalibrerende funksjon i forhold til standpunktkaraktersettingen (Prøitz & Borgen 2010, Prøitz 2013). I våre analyser av data for avgangskullene i 2010-2013 finner vi at standpunktkarakterene i norsk hovedmål er forholdsvis konstant over tid, rundt 3,7, mens eksamenskarakterene varierer mer. Dette gjelder særlig for det siste året. I perioden 2010 til 2012 ligger eksamenskarakterene på 3,1, men i 2013 øker gjennomsnittlig eksamenskarakter til mellom 3,2 og 3,3. Dette er illustrert i figur 3.1.

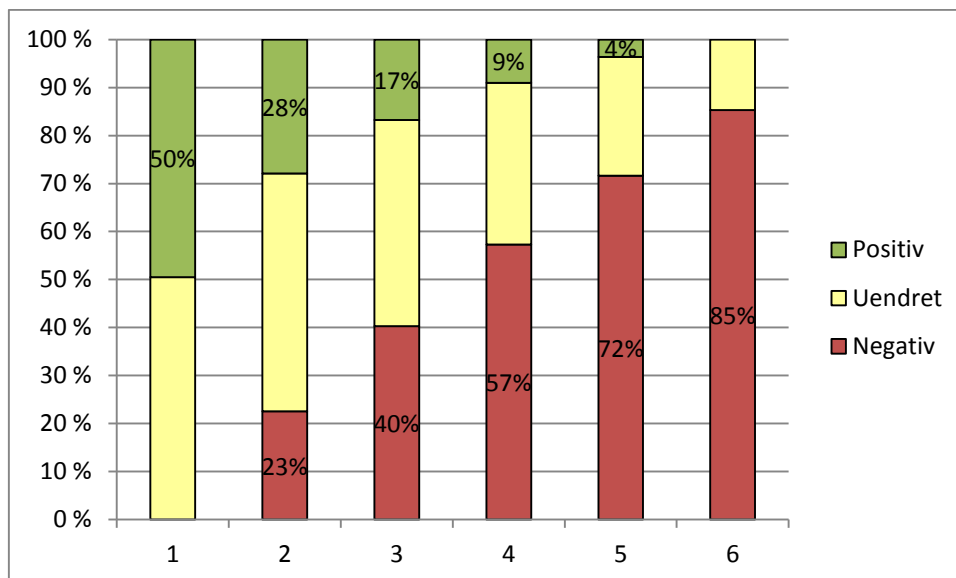


Figur 3.1: Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk skriftlig hovedmål i fire skoleår.

NB: figuren er identisk med figur 3.1 i vedlegg 3.

Dermed bryter mønsteret vi observerer i figur 3.1 med forutsetningen som ligger for problemstillingen. Siden vi bare har tilgang på data for 4 år, kan vi ikke si om dette er et enkelt utslag eller ikke, men det viser likevel at eksamenskarakterene ikke er stabile over tid, og at de kan være mindre stabile enn standpunktkarakterene. Figuren viser karakterer i norsk skriftlig hovedmål, men det samme mønsteret gjelder for de andre fagene med skriftlig eksamen. I alle tilfeller varierer begge former for karakterer noe, men variasjonene er større i eksamenskarakter enn i standpunktkarakter i alle fag, og dette gjelder særlig i 2013. Dermed ser vi at oppfatningen av eksamens funksjon som objektivt målepunkt eller utgangspunkt for kalibrering ikke er uproblematisk, men at det faktisk er noe det kan stilles spørsmål ved. Dette funnet er fulgt opp i de kvalitative intervjuene med lærere på offentlige og private videregående skoler.

En annen utfordring i sammenligninger mellom standpunktkarakter og eksamenskarakter er i hvilken retning vi skal forvente at avvikene går. Karakterskalaen går fra 1 til 6, og dersom en elev har en av de to karakterene, så kan ikke karakteren blir dårligere respektive bedre. Figur 3.2 viser observerte gjennomsnittlige forskjeller i standpunkt og eksamenskarakter i norsk skriftlig hovedmål og viser akkurat dette. For de som hadde karakteren 1 gikk halvparten opp i karakter til eksamen. For de som hadde 6 i karakter var det 85 prosent som gikk ned til eksamen. Figuren viser klart og tydelig at jo bedre standpunktkarakter en elev har desto større er risikoen for å gå ned til eksamen, og at det er vanligere å gå ned til eksamen enn å gå frem for de som har standpunktkarakter 3 eller bedre.

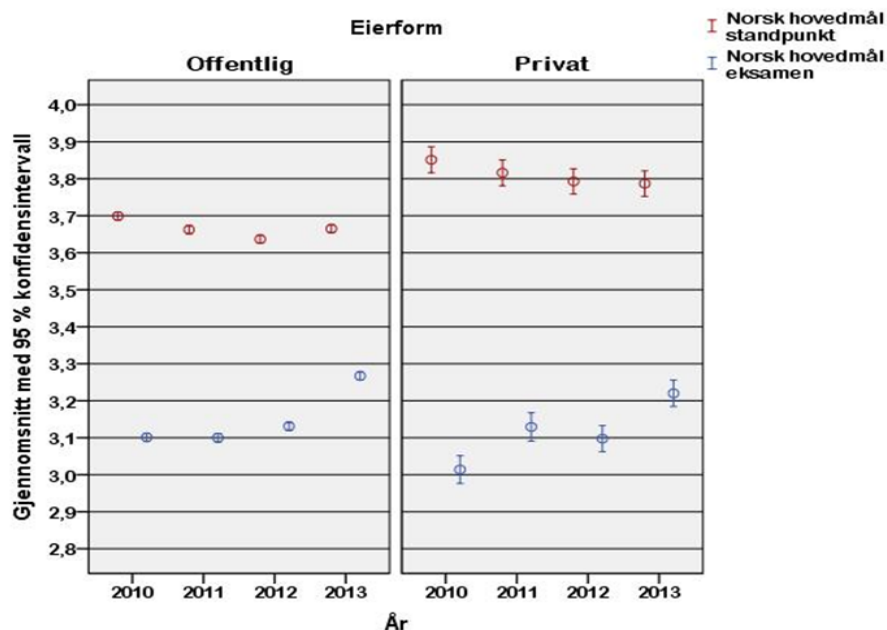


Figur 3.2: Forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål. Gjennomsnitt basert på data for fire år (2010 – 2013).

Fenomenet som figur 3.2 viser, kalles regresjon mot gjennomsnittet (se for eksempel Galton 1886, Grøgaard 2012). Når det er en statistisk samvariasjon mellom to forhold, slik det er mellom standpunktkarakter og eksamenskarakterer i norsk skriftlig hovedmål (Pearsons $r = 0,591$), og en av disse karakterene avviker fra gjennomsnittet, så vil det være størst sannsynlighet for at den andre karakteren ligger nærmere gjennomsnittet enn den første. Når gjennomsnittlig standpunktkarakter er 3,7 og gjennomsnittlig eksamenskarakter er 3,2, vil en elev med standpunktkarakteren 5 ha større sannsynlighet for å få en firer enn en femmer til eksamen. Dette fenomenet gjør at det i analysene er viktig å kontrollere for hvilken standpunktkarakter eleven hadde, fordi det sier noe om risikoen for å gå opp, ned eller bli stående til eksamen. Regresjonsanalyser av hvilke forhold som påvirker gjennomsnittsforskjellen mellom standpunkt- og eksamenskarakter for et individ viser at den viktigste forklaringsfaktoren er nivået på standpunktkarakteren. Selv om effekten varierer noe fra fag til fag, finner vi hele tiden en tydelig tendens til regresjon mot gjennomsnittet, ved at de med best karakter går mest tilbake, mens de med dårligst karakter går fram. Dette funnet er noe det alltid må tas hensyn til når grupper skal sammenlignes.

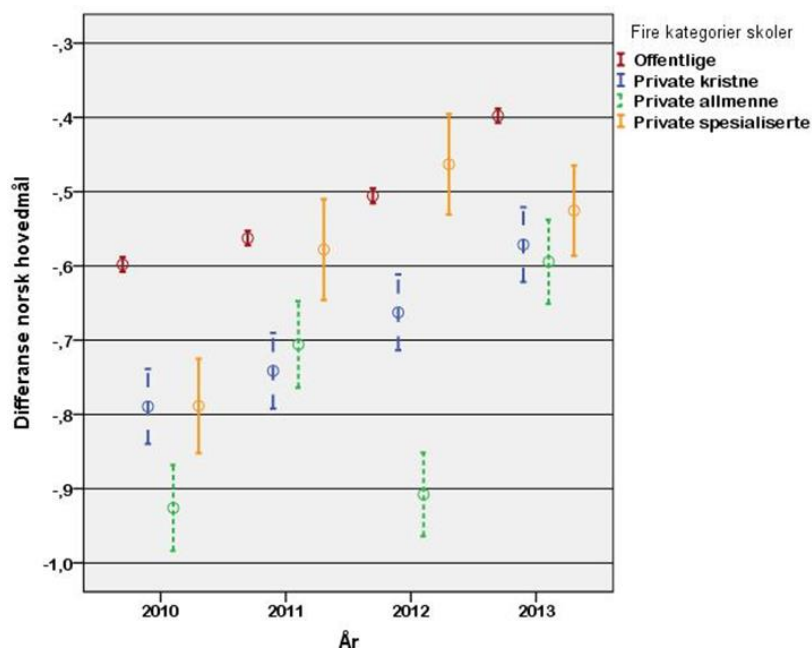
3.2 Forskjeller mellom offentlige og private skoler

Hovedproblemstillingen i dette kapittelet er å se på om det er forskjeller i avvik mellom standpunkt- og eksamenskarakter for offentlige og private videregående skoler. Figur 3.3 viser de observerte gjennomsnittlige standpunkt og eksamenskarakterene for offentlige og private skoler sett under ett. Av figuren fremgår det at det er større avvik på private skoler enn på offentlige skoler. For begge typer skoleslag er mønsteret det samme, nemlig at avviket er mindre i 2013 enn de andre årene og dette skyldes at eksamenskarakterene det året er i gjennomsnitt noe høyere.



Figur 3.3: Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk skriftlig hovedmål i fire skoleår for offentlige og private skoler. Figuren er identisk med figur 3.4 i vedlegg 3.

Figur 3.3 skiller ikke mellom ulike typer private skoler, men i figur 3.4 viser vi forskjellene i karakteravvik mellom offentlige skoler og tre typer private skoler: Kristne skoler, spesialiserte skoler og allmenne private skoler. Figuren indikerer at forskjellene i avvik har variert over tid og mellom ulike typer private skoler. For de offentlige skolene har avviket minsket over tid, fra 0,6 i 2010 til et avvik på 0,4 i 2013. De private kristne skolene følger samme tendens, avviket har minsket fra 0,8 i 2010 til 0,6 i 2013. De private spesialiserte skolene har kun et avvik som er signifikant forskjellig fra de offentlige skolene i 2010 og 2013, og i begge tilfeller er dette avviket forholdsvis lite. For de private allmenne skolene er derimot mønsteret mer sammensatt. Her ble forskjellen kraftig redusert fra 2010 til 2011, for så å øke nesten like mye fra 2011 til 2012. Fra 2012 til 2013 foregår det så igjen en kraftig reduksjon i forskjellen mellom standpunktkarakter og eksamenskarakter ved de private allmenne skolene. Figuren viser med andre ord at variasjonene i avvik mellom standpunktkarakter og eksamenskarakter er noe større ved private videregående skoler enn ved offentlige, men at avvikene i begge tilfeller er redusert over tid.

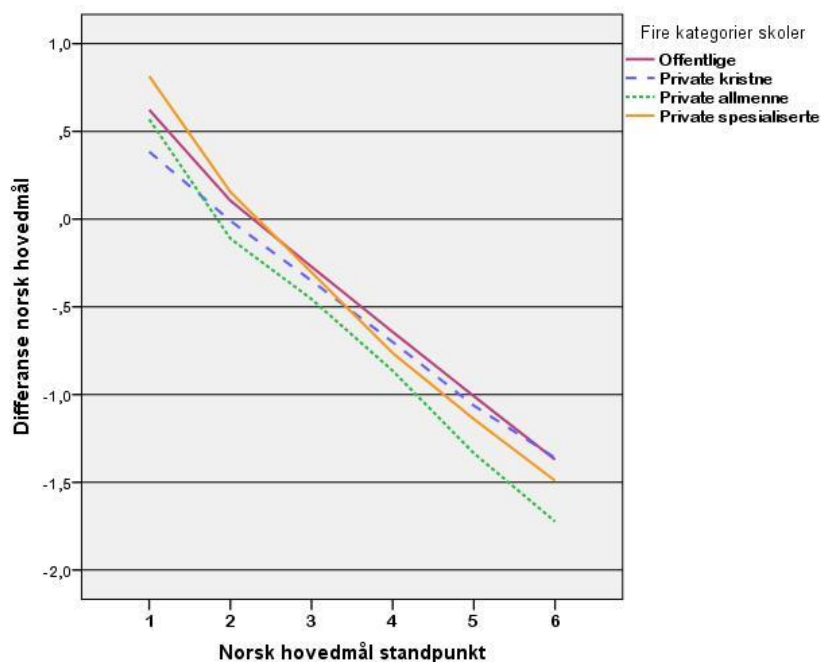


Figur 3.4 Gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål i fire skoleår for offentlige og tre typer private skoler.

Figuren er identisk med figur 3.5 i vedlegg 3.

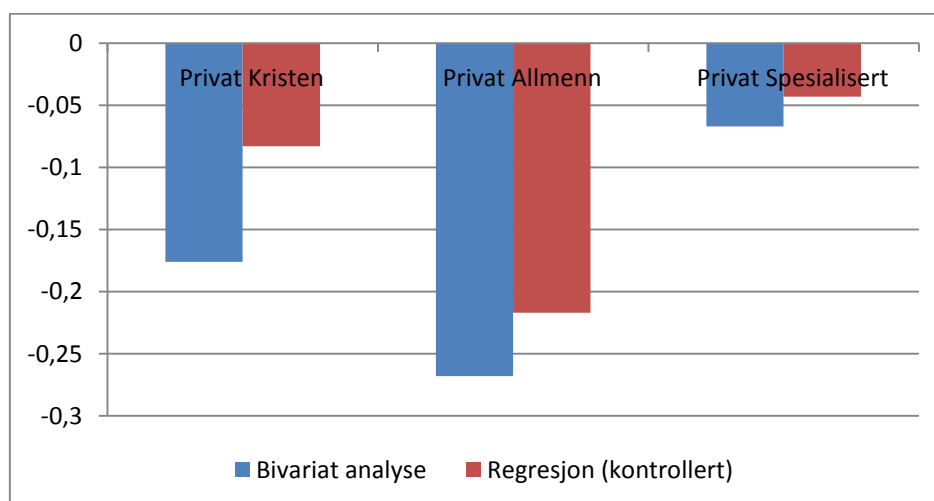
Siden det var 2012-tallene som fikk offentlig oppmerksomhet og utløste dette prosjektet, kan det være grunn til å knytte en ekstra kommentar til disse. Vi ser av figur 3.5 at det dette året ikke var noen forskjell av betydning mellom offentlige skoler og de private spesialiserte, mens det var noe større forskjeller mellom offentlige skoler og de to øvrige private skoleslagene. Særlig gjaldt dette de allmenne private skolene. Året etterpå derimot, i 2013 er avvikene mellom de offentlige skolene og de private skolene mindre, og det er ikke lenger signifikante forskjeller mellom de tre typene privatskoler. Siden vi bare kan observere fire år, kan vi ikke si om dette er en trend, men vi merker oss at avvikene er mindre året etter at det var stor debatt i media om avvikene mellom standpunkt- og eksamenskarakterer.

Dersom vi ser alle årene under ett, kan vi se nærmere på om det er forskjeller i hvordan avvikene mellom standpunkt- og eksamenskarakter fordeler seg avhengig av hvilken karakter eleven hadde i utgangspunktet. I denne bivariate sammenligningen får vi vist at standpunktkarakteren påvirker hvor stor risikoen er for å gå ned til eksamen (regresjon mot gjennomsnittet). Figur 3.5 indikerer at det er små forskjeller mellom offentlige skoler og de tre typene private skoler i avvik, men at avviket for alle karakterer bedre enn 1 er noe høyere for de private allmenne skolene.



Figur 3.5 Avvik mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål i for offentlige og tre typer private skoler, etter standpunktkarakter.

Dette funnet bekreftes også dersom vi tar utgangspunkt i regresjonsanalysen (se vedlegg 3). Også her fremkommer det at private allmenne skoler har et større avvik enn de andre typene private skoler, og at dette ikke blir noe særlig redusert når vi kontrollerer for andre forhold. Figur 3.6 viser hvor stort avviket er i de tre typene private skoler. Det er kun små forskjeller i avvik mellom offentlige skoler og private spesialiserte skoler. Private kristne skoler har et observert avvik på -0,17 i de bivariate analysene, men denne forskjellen halveres når vi kontrollerer for kjønn, alder, programområde, skolestørrelse og om skolen har høy inntakskvalitet. Med andre ord rekrutterer private kristne skoler i noen grad andre typer elever enn offentlige skoler, som bidrar til å forklare avviket vi finner i de bivariate analysene. For de private allmenne skolene er derimot avviket på -0,27 i de bivariate analysene og -0,22 i regresjonsanalysen, og dermed reduseres avviket lite når vi har kontrollert for andre forhold.



Figur 3.6: Gjennomsnittlig avvik i bivariate analyser og i regresjonsanalyser, etter type privatskole.

Regresjonsanalysen viser dessuten at det ikke bare er skillelinjen offentlig - privat som påvirker avvik mellom standpunkt- og eksamenskarakter. Vi finner også at jenter går tilbake noe mindre enn gutter, at avviket er noe større naturbruk og påbygg enn på andre programområder, samt at de som går på en skole med høye opptakskrav har mindre negativt avvik og de som går på små skoler har et noe større negativt avvik, uavhengig av om dette er en offentlig eller en privat skole. Med andre ord er det ikke bare skillet offentlig – privat skole som har påvirkning på avviket mellom standpunkt- og eksamenskarakter.

3.3 Oppsummering

Vi har i dette kapittelet sett at private allmenne skoler har noe større avvik mellom standpunkt- og eksamenskarakter enn hva som er tilfelle for offentlige skoler, private kristne skoler og private spesialiserte skoler. Samtidig viser regresjonsanalyser at det også er andre variabler som påvirker avviket, blant annet kjønn, type program, om skolen har høye opptakskrav og om det er en liten eller stor skole. Små skoler har noe større avvik enn større skoler.

Det som har størst påvirkning på hvor stort avviket er, er elevens standpunktkarakter; elever med gode karakterer har større risiko for å gå tilbake til eksamen enn elever med svakere karakterer. For en elev med 6 i standpunktkarakter er risikoen for å gå tilbake 85 prosent og for alle elever med karakter 3 eller bedre er det større risiko for å gå tilbake til eksamen enn å gå frem.

Vi ser også at det er forskjeller i avvik mellom de fire årene vi ser på (2010-2013), ved at avviket er mindre i 2013, og dette kommer av at gjennomsnittlig eksamenskarakter er høyere det året. Siden vi ikke har tilgang til informasjon om mer enn de fire årene, kan vi ikke si om dette er en generell trend, men det er likevel interessant å merke seg at det er eksamenskarakterene som varierer mest over tid. Dermed kan det diskuteres om eksamenskarakterer faktisk har den kalibrerende funksjonen som de antas å ha. Det forutsetter i så fall at karakterene ikke varierer mye fra år til år. Samtidig kan det være noe ved eksamen det enkelte året som forklarer avviket, men det har vi ikke informasjon om her. Siden eksamen forutsettes å måle det samme fra år til år, bør ikke vanskelighetsgraden variere for mye. Det kan den heller ikke gjøre dersom den skal sies å ha en kalibrerende funksjon.

4 I spenningsfeltet mellom pålegg og praksis

I dette kapittelet skal vi ta for oss intervjuer med lærere som har ansvar for undervisning og standpunktvurdering i norsk og matematikk ved til sammen åtte videregående skoler, som vi fikk besøke i løpet av våren 2014. På hver skole har vi intervjuet en gruppe på mellom tre og fem lærere. Utvalget av skoler har vi gjort rede for i kapittel 2. Blant de åtte skolene er det fire private og fire offentlige videregående skoler, og ett av spørsmålene denne rapporten søker å finne svar på, er hvorvidt det er forskjeller i rutiner for karaktersetting blant lærere i de to ulike skoleslagene. I kapittel 2 beskriver vi også hvordan intervjuene foregikk. Intervjuguiden som vi brukte i samtalene med lærerne, følger denne rapporten som vedlegg.

Kapittelet er bygget opp på følgende måte: etter et kort tilbakeblikk på vårt grunnlag for analyse av rutiner for elevvurdering, tar vi for oss hovedelementene i slike rutiner og beskrivelser av praksis, belyst med eksempler fra intervjumaterialet. Kapittelet avsluttes med en diskusjon av funnene i datamaterialet.

4.1 Rutiner for elevvurdering

Slik vi har gjort rede for i teorigjennomgangen i kapittel 1 i denne rapporten, har vi gjennom intervjuene søkt å bringe frem lærernes beskrivelser av rutiner for å sette standpunktkarakter i norsk og matematikk for elever på Vg3 studiespesialiserende program. Rutiner forstår vi her som gjentatte og gjenkjennelige handlingsmønstre som strukturerer praksis, det vil si at handlingsmønstrene legger rammer for lærernes arbeid med vurdering, at de ligger til grunn for koordinering av dette arbeidet mellom lærerne og kan bidra til å redusere eventuell konflikt og usikkerhet om hvordan vurdering skal foregå. Sammenfattet kan vi si at vi i dette kapittelet søker å beskrive og forstå hvordan skolen og lærerne overfører, fortolker og utvikler de pålegg de får om vurdering – og først og fremst standpunktvurdering – gjennom enkeltlærerers praksis og samhandling mellom lærere i det daglige arbeidet.

4.1.1 *Designede rutiner kontra rutiner i daglig bruk*

Slik vi viste til i kapittel 1, kan slike rutiner være nedfelt i et regelverk, i formaliserte retningslinjer eller veiledninger, eller de kan være kunngjort som formelle beskjeder og pålegg fra for eksempel rektor eller representanter for den øvrige skoleledelsen i møte med lærerkollegiet. Slike skriftlige dokumenter eller offentlig kunngjorte budskap kan være utviklet av nasjonale myndigheter, av skoleeier eller av den enkelte skole for å regulere, styre og understøtte vurderingsarbeidet på skolene. Ideelt sett skal slike dokumenter bidra til at praksis blir mest mulig lik på tvers av skoler, eller i det minste at læreres

vurderingspraksis innenfor den enkelte skolen blir mest mulig ensartet. Dette kalles i denne rapporten for *designede rutiner*. I vår analyse regner vi også kurs og opplæring, herunder sensoropplæring eller den nasjonale satsningen *Vurdering for læring* som eksempler på slike designede rutiner.

I tillegg til de formelle, designede rutinene vil det også i det daglige arbeidet finnes hva vi i denne rapporten kaller *rutiner i bruk*. Disse dagligrutinene kan springe ut av de designede rutinene, og være en videreutvikling av disse som enkeltlærere eller grupper av lærere (for eksempel norsklærere som arbeider på samme trinn) har skapt over tid som en støtte for arbeidet. Dagligrutinene kan imidlertid også ha oppstått mer eller mindre av seg selv, det vil si som løsning på et behov eller et problem som en eller flere lærere har erfart. Kjennetegn på slike dagligrutiner er at de er ikke-formelle, det vil si ikke skriftliggjorte, men at de har blitt til en gjenkjennelig arbeidsmåte for lærerne. Et eksempel på dette kan være norsklærerne som har gjort en vane av å rette noen elevoppgaver i fellesskap for å bli enige om konkret vurderingspraksis, før de setter seg med hver sin bunke elevbesvarelser og retter disse hver for seg. Et annet eksempel kan være matematikklærerne som har laget seg et nettverk med kolleger på andre skoler for å utveksle kvalitetssikrede oppgaver som kan brukes i heldagsprøver.

Slik det går fram av disse beskrivelsene, kan det være glidende overganger mellom designede og daglige rutiner, i den grad slikt samarbeid er initiert av rektor eller seksjonsleder (designet rutine) eller har oppstått «nedenfra» som svar på et behov som lærerne har løst på egen hånd (dagligrutiner). De daglige rutinene kjennetegnes likevel som nettopp «rutiner» ved at handlingene blir utført på noen lunde samme vis over tid, at de blir gjennomført av bestemte personer med en bestemt hensikt.

4.1.2 Aritmetisk og analytisk vurdering

I kapittel 1 i denne rapporten har vi også gjort rede for to tilnærminger til vurdering og karaktersetting. Vi skal kort repetere disse tilnærmingene her. Den første måten å vurdere elever på, refereres i forskningslitteraturen som *aritmetisk vurdering*, og kan omsettes til poengberegning. Ifølge denne tilnærmingen er et svar enten riktig eller galt, og svarene kan tilordnes en verdi som igjen kan summeres og fremstilles som en prosentandel av det totale antallet riktige svar som det er mulig å oppnå på oppgavene i prøven. Karakterskalaen vil basere seg på prosentsatser, der 20 prosent riktige svar vil gi en bestemt karakter i den lave enden av skalaen, mens 75 prosent riktige svar vil gi en på forhånd bestemt karakter i den høyere enden av skalaen. En motsats til denne standardiserte måten å vurdere elevbesvarelser er en *analytisk* tilnærming til vurdering. Den analytiske tilnærmingen kjennetegnes ved en individuell vurdering av elevens oppnåelse av bestemte mål for opplæringen og hvordan disse målene kommer til uttrykk i elevens arbeid. Her er tolkningsrommet betydelig større enn i den første, aritmetiske beregningen av elevens resultater. En analytisk vurdering av elevarbeidet gir også større rom for forhandling, det være seg mellom kolleger eller mellom lærer og elev.

Som vi skal se i den følgende empiriske gjennomgangen, forbindes ofte den aritmetiske fremgangsmåten med realfag, mens humanistiske og samfunnsvitenskapelige fag er preget av en analytisk tilnærming til vurdering. I vårt materiale vil dette tilsvare skillet mellom matematikkfaget og norskfaget. Som en norsklærer på en privat skole sier i intervjuet: «I matematikken er det én fasit, men i norsken er det kanskje én fasit per elev.» Likevel må ikke aritmetiske og analytiske tilnærminger til vurdering oppfattes som fullstendig utelukkende kategorier som følger disse faggrensene. En matematikklærer kan oppgi å la aritmetisk vurdering veie tyngst, men deretter bruke et analytisk blikk for å vurdere helheten i prøven og ikke minst hvorvidt eleven har forstått oppgaven, men kommet i skade for å bruke feil tallverdi i utregningen. Denne feilen er kjent for matematikklærerne som følgefeil, og er et typisk eksempel på at matematikklæreren må kjenne til og kunne anvende begge vurderingsmåter.

4.1.3 Universell og differensiell behandling

Til sist når det gjelder rutiner for elevvurdering, viste vi i kapittel 1 til at lærerne kan etterstrebe å behandle alle elevene likt (universell behandling) men likevel ende med å behandle elever ulikt (differensiell behandling tilpasset eleven). Forskningslitteraturen som underbygger dette, viser til

undervisningssituasjoner i klasserommet og empiriske studier der lærerne under gitte omstendigheter behandler svake og sterke elever ulikt. I vurderingspraksis fremstår slik eventuell forskjellsbehandling i utgangspunktet i strid med formelle, designede rutiner, som tilsier at alle elevene skal behandles på samme måte slik at karakteren blir rettferdig. Likevel vil det, både innenfor de designede og de ikke-formaliserte, daglige vurderingsrutinene, oppstå rom for tvil, forhandling eller også dilemma som kan gjøre at en lærer ser det som nødvendig å gi en elev særskilt behandling. Dette kan for eksempel ta form av en ekstra prøve for en elev som står i fare for å stryke. Et annet spørsmål er hvordan aritmetiske og analytiske vurderingsmåter (se over) i realiteten legger til rette for universell eller differensiert behandling av elever, intendert eller uintentert, og hvilke utfordringer lærerne mener knytter seg til dette.

De klareste eksemplene på designede rutiner er vurderingsforskriften, bruk av eksamensforberedende prøver inkludert eksamensveiledninger, deltakelse på formelle kurs eller programmer for vurderingsarbeid – også i forbindelse med sensorarbeid – samt uttalte prosjekter for skoleinternt utviklingsarbeid i forbindelse med lærernes vurderingspraksis. Disse rutinene suppleres av en rekke formaliserte praksiser for vurdering knyttet til den løpende undervisningen og samarbeid i bestemte faggrupper eller seksjoner i lærerkollegiet. Her finnes også glidende overganger til arbeidsmåter som har oppstått og blitt videreført uten å ha blitt formalisert – mange av dem i forlengelse av eller som støtte til de formaliserte prosedyrene. Vi skal ta for oss hver av disse designede rutinene etter tur, og vise hvordan de utfylles av formalisert og ikke-formalisert praksis. Fremstillingen i det følgende er belyst med eksempler fra intervjumaterialet.

4.2 Bruk av vurderingsforskriften

Alle lærerne vi har intervjuet i denne studien, viser til vurderingsforskriften eller viser indirekte til denne når de sier at de følger nasjonale myndigheters pålegg om vurdering. Vi har imidlertid ikke inntrykk av at vurderingsforskriften er det vi kan kalle et «aktivt dokument» i lærernes daglige arbeid. Med det mener vi at vi ikke har inntrykk av at lærerne konsulterer forskriften jevnlig eller hver gang de ulike vurderingssituasjonene oppstår, men at de er kjent med at denne forskriften legger rammene for arbeidet deres og styrer rutinene for vurdering på skolen og i arbeidshverdagen.

Når lærerne i intervju viser oss at de er kjent med forskriften, gjør de i særlig grad dette ved å understreke betydningen av sluttvurdering. Dette er likt for alle lærere vi har snakket med, og det er ingen forskjeller på lærernes uttalelser avhengig av om de arbeider på en privat eller en offentlig skolen. Alle lærerne fastslår i intervju at standpunktkarakter skal bygge på sluttvurdering, forstått som at det er elevens måloppnåelse definert av kompetansemålene i læreplanen som skal gjenspeiles i lærerens vurderingspraksis i avslutningen av opplæringen. På en av de private skolene kontrasterer lærerne dette mot å sette standpunktkarakter ved å regne ut et gjennomsnitt av delkarakterer gitt underveis i skoleåret eller i opplæringsløpet, og understreker at slik praksis ikke skal gjelde. Likevel er det ikke noen fullstendig enighet blant lærerne om hvordan «sluttvurdering» faktisk skal tolkes i det daglige arbeidet. En matematikklærer på en offentlig skole uttrykker en prinsipiell tolkning på vegne av skolen og fagseksjonen sin slik: «Vi har i dag en forskrift som sier at elevene skal ha sluttvurdering, og den har vi valgt å tolke veldig bokstavelig. Det betyr at det eleven gjør gjennom året, egentlig ikke påvirker standpunktkarakteren.» En norsklærer på samme skole, som ble intervjuet samtidig som denne matematikklæreren, nyanserer dette for norskfagets del:

Vi har en sluttvurdering hvor vi legger vekt på utviklingen i elevens arbeid, og vi prøver å få sirklet dette inn så godt som mulig i løpet av året slik at karakteren vi setter, blir riktig. Vi avslutter ganske mange temaer underveis, vi har ikke den måten å gjøre det på som i matematikk. Norskfaget er stort, og man blir nødt til å «lukke» kompetansemål og si at i dette emnet nådde eleven så og så langt på det nivået. Så det blir en balanse mellom å ta hensyn til sluttkompetanse og utvikling.

Forskjellen som kommer til syne mellom norskfaget og matematikkfaget i disse sitatene, er et gjennomgående funn i de kvalitative lærerintervjuene, og det tydeligste funnet på tvers av skoler i vårt materiale. For norsklærerne generelt får derfor «sluttvurdering» et noe annerledes innhold og betydning enn for matematikklærerne. Dette kan beskrives som en gradsforskjell i vekten av undervisvurderingens betydning når sluttkompetansen til eleven skal stadfestes. En matematikklærer på en offentlig skole sier for eksempel: «Jeg synes ikke en 2er i matematikk i september skal bety noe som helst når jeg setter karakteren i mai.» En norsklærer på en privat skole sier derimot: «Vi vil legge mest vekt på det som skjer i 2. termin mot slutten, men det er jo vurderingssituasjoner ellers i året og i 1. termin, som man må se på.» I denne uttalelsen, som er representativ for norsklærerne i materialet vårt, blir overgangene mellom undervisvurdering og sluttvurdering i norskfaget mer glidende. En norsklærer på en privat skole som reflekterer over det samme, mener likevel at det i norsk muntlig er mulig å måle alle kompetansemål på en avsluttende prøve, mens dette ikke lar seg gjøre i de to fagene med skriftlig eksamen: norsk hovedmål og norsk sidemål. Dette er det imidlertid ikke enighet om blant alle norsklærerne. Vi har også intervjuet en norsklærer på en offentlig skole som mener at det lar seg gjøre å teste de fleste kompetansemål i norsk skriftlig ved avslutningen av Vg3 «Likevel kan ikke standpunktkarakter i norsk være så avhengig av siste terminprøve som det matematikken er,» sier en kollega til denne norsklæreren i det samme gruppeintervjuet. «Man har jo et ganske godt inntrykk av eleven gjennom tredje klasse, og kanskje karakteren frem til første termin, så man justerer litt,» utdyper hun.

Selv om matematikklærerne er mer prinsipielle i sin omtale av sluttvurdering, sier imidlertid også disse lærerne i intervjuene at det kan bli nødvendig å supplere vurderingen av en elevs eventuelle overraskende dårlige resultat på en avsluttende heldagsprøve med å se tilbake på hva eleven har prestert gjennom året. Derfor kan «sluttvurdering» også i matematikk inneholde et element av hva eleven har prestert underveis i siste termin eller siste år på videregående skole. En matematikklærer på en offentlig skole omtaler disse utvidede vurderingene som er en mer pragmatisk tolkning av «sluttvurdering» i matematikk enn det vi så eksempel på innledningsvis:

I matematikk har vi en årsprøve, og den omfatter jo hele pensum fra året, så den blir det lagt stor vekt på. Og så bruker man kanskje karakterene fra tidligere i året som en korreksjon, hvis man ligger og vipper. Likevel legger jeg mest vekt på det som blir gjort gjennom våren, for det skal være mulig å vise at man har forbedret seg.

I dette sitatet og i foregående henvisning til norsklæreres utsagn om utfordringer ved vurdering, er det tydelig at de overordnede designede rutinene (vekt på sluttvurdering pålagt i vurderingsforskriften) får glidende overganger til rutiner i daglig bruk (læreren må se på hva eleven presterer jevnt gjennom året).¹ Dette hovedfunnet i materialet er tydeligst i skillet mellom matematikklærere og norsklærere uavhengig av om lærerne arbeider på en offentlig eller en privat skole, og skillet kommer til uttrykk gjennom utsagn der lærerne forteller om behovet for å gjøre en analytisk tilnærming til vurdering og en viss differensiering eller i det minste individualisert behandling av elever. Imidlertid har også vurdering i matematikk innslag av analytisk vurdering. Vi skal nå se hvordan lærerne beskriver og balanserer disse overgangene mellom formalia og utøvelse av profesjonelt skjønn gjennom de øvrige rutinene de forteller om i intervjuene.

4.3 Bruk av eksamensforberedende prøver

Alle lærerne vi har intervjuet omtaler eksamensforberedende prøver, eller heldagsprøver i norsk og matematikk på våren i Vg3, som det aller viktigste elementet som inngår i standpunktvurderingen av

¹ Dette er også i tråd med Utdanningsdirektoratets egen fortolkning av forskriften i rundskriv der det brede og helhetlige blikket på elevprestasjoner understekes .

<http://www.udir.no/Regelverk/Rundskriv/20101/Udir-1-2010-Individuell-vurdering/>

elevene. Heldagsprøvene avholdes følgelig også på alle skolene vi har besøkt, med noen variasjoner. De fleste skolene i materialet vårt avholder kun én ordinær, skriftlig heldagsprøve i de respektive fagene norsk hovedmål, norsk sidemål og matematikk.² En av de offentlige skolene som vi har besøkt, deler imidlertid matematikkprøven i to og holder den over to dager med om lag en ukes mellomrom. Dette er ifølge matematikklærerne på denne skolen en fornuftig ordning for å teste alle kompetansemål i faget og på den måten sikre en reell sluttvurdering, hvor begge prøvene teller like mye. En av de private skolene lar elever få søke om å ta en ny heldagsprøve dersom resultatet av den ordinære prøven er svært dårlig. Den første heldagsprøven gir da grunnlag for en midlertidig standpunkt karakter, som eleven kan forsøke å forbedre ved å ta en ny heldagsprøve. Hvis det blir aktuelt å revidere standpunkt karakteren på grunnlag av ny heldagsprøve, settes denne karakteren før den formelle fristen for å sette endelig standpunkt karakter utløper.

4.3.1 Gamle eksamensoppgaver som forberedelse til heldagsprøver

At heldagsprøvene er eksamensforberedende, betyr for lærerne at prøvene skal ligne så mye på eksamen som mulig. Lærerne lar også elevene øve seg til disse prøvene ved å løse gamle eksamensoppgaver. Heldagsprøvene lages av faglæreren eller i et samarbeid mellom lærere som har samme fag eller kurs. I dette arbeidet tar lærerne ofte utgangspunkt i gamle eksamensoppgaver. «Jeg bruker veldig ofte gamle eksamensoppgaver,» sier en matematikklærer på en offentlig skole, «jeg liker det best, for da får elevene et inntrykk av hva de skal testes i til slutt.» Lærere som oppgir at de lager heldagsprøvene selv på grunnlag av gamle eksamensoppgaver, understreker at de også bruker de tilhørende eksamensveiledningene. En matematikklærer på en privat skole viser imidlertid til at mange gamle eksamensoppgaver ligger tilgjengelige på nettet med ferdige forslag til løsninger, og at han derfor ikke vil lene seg for tungt på slike gamle oppgaver i utformingen av heldagsprøven. Denne læreren lager egne varianter av oppgavene. Lærere på offentlige så vel som de private skolene forteller at de kan også benytte seg av forslag til heldagsprøver utarbeidet av forlag på bakgrunn av eksisterende læreverk, eller kombinere disse kildene for å lage heldagsprøver. Her er vi over i rutiner i lærernes daglige arbeid som ikke er nedfelt noe sted, men som likevel finner sin form i forlengelse av hvordan heldagsprøvene (en designet rutine) brukes av skolen og av lærerne for å sette standpunkt karakter for elevene.

4.3.2 Eksamensveiledninger som forberedelse til heldagsprøver

Alle lærerne vi har intervjuet, oppgir at de kjenner til eksamensveiledningene i sine respektive fag. Det varierer imidlertid hvordan de sier at de bruker dem i undervisningen eller som forberedelse til heldagsprøvene, og i hvilken grad de formidler veiledningene til elevene. Disse variasjonene synes å være individuelle, og kan ikke knyttes til skillet mellom offentlige og private skoler. De kan heller ikke knyttes til forskjeller mellom fagene norsk og matematikk. «Jeg har lest dem. Men da mest i forbindelse med en situasjon hvor jeg skal vurdere til eksamen, jeg har ikke brukt dem mye i min egen vurdering,» sier en norsklærer på en offentlig skole. «Jeg ser på dem i forbindelse med for eksempel karakterskala, det står jo hvilken poengsum som gir de ulike karakterene, så det er mest for å se at jeg ligger i nærheten av det samme» sier en matematikklærer på en annen offentlig skole. Denne læreren gir samtidig et tydelig eksempel på hvordan en designet, formell rutine kan bestrides og endog avvises fra skolenivå, uten å ha blitt omformet til praksis på skolene. Eksempelet denne læreren gir, gjelder matematikk, der ny eksamensveiledning anga at karakterkravet for karakteren 2 skulle settes ned fra 25 prosent til 20 prosent måloppnåelse. Dette ble av matematikklærerne på denne skolen oppfattet som et signal om hva myndighetene ønsker, men samtidig avvist av disse lærerne, som sier at de ikke kommer til å endre sin vurdering av karakteren 2 med mindre de får ytterligere pålegg. Samme passus om senkning av karakterkrav i eksamensveiledningen blir omtalt med undring av andre matematikklærere vi har intervjuet, fra offentlig så vel som privat sektor.

² Skolene holder imidlertid også oppsamlingsprøver for elever som ikke var til stede under den ordinære heldagsprøven. Vi skriver mer om dette senere i kapittelet. Her vil vi først vise forskjellen mellom skoler i organiseringen av de ordinære prøvene.

En del lærere gir eksamensveiledningene videre til elevene sine, og gjennomgår sentrale deler av disse dokumentene i timene. En grunn til at ikke alle lærere videreformidler eksamensveiledningene direkte til elevene, er at de erfarer at språket i veiledningene er for abstrakt for elevene. I noen utstrekning arbeider derfor enkeltlærere med å «oversette» sentrale punkt fra eksamensveiledningene for elevene og gi dem forenklede kjennetegn på måloppnåelse. Lærerne som oppgir at de bruker eksamensveiledningene aktivt, sier imidlertid at de selv etterstreber å gi elevene vurderinger i tråd med disse veiledningene også underveis i skoleåret, slik at elevene skal bli vant til den strukturen og de kjennetegn på gode besvarelser som det vil bli lagt vekt på til eksamen. På flere skoler har elevene tilgang til eksamensveiledningene gjennom elektroniske læringsplattformer, uavhengig av om lærerne viser til veiledningsdokumentene i undervisningen eller ikke. Også her, i beskrivelsen av hvordan eksamensveiledningene brukes og formidles til elevene, ser vi derfor at designede rutiner (selve veiledningsdokumentene) avfører rutiniserte, men ikke-formaliserte praksiser på den enkelte skole og hos den enkelte lærer. Disse praksisene synes å være individuelle, det vil si at de har funnet sin form hos den enkelte lærer etter hva som passer best for hans eller hennes arbeidsmetoder. Vi finner ikke skiller mellom lærere i offentlige og private skoler, eller skiller mellom lærere i norsk og matematikk i dette spørsmålet.

4.3.3 Hva heldagsprøven kan måle: forskjeller mellom norsk og matematikk

I forlengelse av diskusjonen om hvordan begrepet «sluttvurdering» må tolkes og hvordan dette kan komme til uttrykk i de to fagene norsk og matematikk, aktualiseres igjen lærernes ulike syn på hvordan heldagsprøven faktisk kan brukes til å teste elevenes måloppnåelse i alle kompetansemål. I hovedsak mener alle de intervjuede lærerne (både i norsk og matematikk) at heldagsprøver i matematikk egner seg til dette formålet. «Matematikken er jo ettårig, det er mye mer begrenset enn norsk,» sier en matematikklærer på en privat skole, og fortsetter: «i matematikk kan du faktisk måle alle hovedmålene. Du får ikke målt alle delmålene, det er jo umulig, for pensum er stort. Men du får stikkprøver på hovedmålene.» De aller fleste norsklærerne anser det derimot som ikke gjennomførbart å la en heldagsprøve måle alle kompetansemål ved avslutningen av Vg3, og i dette utsagnet møter de forståelse hos matematikklærerne som deltar i samme intervju. En norsklærer på en privat skole sier likevel at selv om han må «lukke» kompetansemål i norsk underveis gjennom året, kan eleven velge en oppgave til avsluttende heldagsprøve som tester det samme kompetansemålet og som i prinsippet kan vise elevens utvikling og fremgang. Slik forblir etter denne lærerens mening ikke kompetansemålene i norsk fullstendig «lukket», selv om de må avsluttes som del av undervisningen tidligere på året. Enkelte lærere, i vårt materiale fordelt på en privat og en offentlig skole, viser også til fagsamtaler de har med elever som del av vurderingspraksisen, og at eleven kan benytte slike fagsamtaler på slutten av året til å vise forbedret måloppnåelse i et kompetansemål som ble «lukket» i undervisningen tidligere i skoleåret.

4.3.4 Når lærere skal sette karakter på heldagsprøven

Vi spurte lærerne om hvordan de konkret arbeider med elevbesvarelsene fra heldagsprøvene for å sette standpunkt karakteren. Svarene vi fikk, belyser igjen avstanden eller kontrastene mellom norsk og matematikk som fag, samtidig som bredden i hvert av disse fagene blir utdypet. Matematikklærerne starter svarene sine på dette spørsmålet ved å vise til bruk av poeng og prosent når de vurderer elevbesvarelser, det vil si en typisk aritmetisk tilnærming til vurdering av kompetansemål. «Hvis du har 50 prosent skåre, så skal det mye til at det ikke er en 3er, og får du 95 prosent rett så skal det mye til at det ikke er en 6er,» sier en matematikklærer på en offentlig skole. «Matematikk er et veldig ryddig fag,» sier en matematikklærer på en annen offentlig skole, «enten så behersker du det, eller så behersker du det ikke. Det gjør det lett å nivådele og sette karakter.» En tredje matematikklærer, også han fra en offentlig skole, kontrasterer matematikkfaget og norskfaget når det gjelder klarheten i nivådeling på denne måten:

I matematikk er det lav, middels og høy måloppnåelse, og så kan vi gå inn etterpå og se hvor i det nivået elevbesvarelsen ligger. Jeg har aldri opplevd at det er vanskelig å sette inn lav, middels eller høy målforståelse, men det kan være vanskelig innenfor disse nivåene. Men i

biologi, der har jeg opplevd at det kan være mer sprik. Det er et skriftlig fag som kanskje kan relateres mer til norskfaget, hvor det er mer synsing, litt mer om begrepene er brukt korrekt eller på hvilket nivå de har brukt begrepene. Da må man bruke ganske mye tid på å se hvor eleven ligger.

En matematikklærer fra en privat skole følger opp dette med å si at hun oppfatter hovedmålene i matematikk som svært tydelige, hvor lærer først skal sette poeng og deretter foreta en helhetsvurdering. «Og så kommer karakteren ut nærmest av seg selv,» sier denne læreren, «grensene er veldig tydelige og det er klart hvordan man som lærer skal gjøre vurderingene». En fagfelle på en av de offentlige skolene betegner selve faget som «mekanisk,» og utdyper dette med å si at i matematikken er det "veldig rett frem, som lærer kan du enkelt se hvor mange oppgaver eleven klarer, du kan se hvor det stopper opp og hvor dypt kunnskapen strekker seg." Enda et sitat fra en matematikklærer i en offentlig skole kan oppsummere hvordan disse lærerne retter:

I realfagene er det forholdsvis enkelt, ofte er et svar riktig eller feil, og muligens er man kommet litt halvveis til svaret, Da kan man få delpoeng. Vi opererer med poeng og prosent. Basert på dette setter vi en karakter basert på en karaktergrense som vi har blitt enige om enten på skolen eller i fylket. Noen ganger er prosentene i et grenseområde sammenlignet med måloppnåelsen. Vi retter først riktig og feil i oppgaven, og etterpå ser vi gjennom resultatet en gang til.

Hvis matematikkfaget kan beskrives som «mekanisk» eller aritmetisk, betoner norsklærerne faget sitt på en helt annerledes måte. Også i dette faget vil det være elevenes oppnåelse av hovedmålene for faget som teller, men hvordan denne oppnåelsen skal måles, er mindre entydig og klart enn i matematikkfaget. «Norsk er et veldig komplekst fag,» sier en av norsklærerne på en offentlig skole, og fortsetter:

Elevene skal ha tre standpunkt karakterer i norsk. Vi opererer jo ikke med poeng, og bare det å sette karakter på heldagsprøver er et helt prosjekt. Der har vi stor nytte av fagmøter, hvor vi plukker ut oppgaver og diskuterer, alle lærerne på trinnet. Vi skal se på innhold, struktur, språk, hvorvidt de har gjort hensiktsmessige valg.

Tredelingen mellom språk, struktur og innhold går igjen i norsklærernes svar på hvordan de går frem for å vurdere elevenes arbeid når de skal sette karakter på heldagsprøven. Helheten i teksten vurderes gjennom disse enkeltdelene. Lærerne viser i dette til vurderingskriteriene utarbeidet av Utdanningsdirektoratet. To norsklærere på en privat skole forklarer også at disse tre nodene i norskfaget kan utfylle hverandre når de vurderes, på den måten at såfremt oppgaven er relevant besvart, kan en god struktur eksempelvis veies opp mot et svakt innhold. På en annen privat skole nevner norsklærerne vurderingskriteriene kommunikasjon og skriftlighet, som etter vårt skjønn kan operasjonaliseres i samme retning som språk, struktur og innhold.

På en av de offentlige skolene fremhever norsklærerne at de ser på bruken av fagbegreper, der eleven må vise at hun kan reflektere. I analysen av fagbegrepene har norskseksjonen på skolen laget seg en egen liste over kriterier. «Hvis elevene klarer å komme innom de der, i alle fall tre av dem, så kan de gjøre det ganske bra,» sier en norsklærer på denne skolen, og legger til at oppgaven naturligvis også må besvares, og språkføringen må være god. Listen over fagbegreps-kriteriene fremstår her som et godt eksempel på en formalisert rutine utviklet i norskseksjonen under ledelse av seksjonsleder, et hjelpemiddel som mellomnivå mellom de formelle påleggene som kommer fra sentralt hold i form av vurderingskriterier og uuttalt individuelt skjønn i lærernes daglige vurderingsarbeid.

Alle norsklærerne på skolene vi har besøkt, beretter om bruk av vurderingskriterier i faget når de snakker om retting av heldagsprøven. I intervjuene har det ikke blitt helt klart for oss hvorvidt dette er kriteriene utarbeidet av Utdanningsdirektoratet eller lærerne i større eller mindre fagfellesskap, og om det i tilfelle har vært seksjonsleder e.l. som har initiert utarbeidelsen av slike kriterier. Dette er heller

ikke så viktig for vår analyse. Mer sentralt i lys av våre problemstillinger er det at norsklærerne her oppgir at de bruker slike kriterier i arbeidet sitt, for å motvirke det de omtaler som «synsing», samtidig som de understreker at det å sette karakter i norsk aldri kan bli en eksakt vitenskap. Ved å bruke vurderingskriterier som det er faglig enighet om, og gjennom samarbeid med fagfeller søker norsklærerne etter bredest mulig grunnlag for vurdering som de vil etterstrebe å tolke etter så objektive kriterier som mulig. En norsklærer beskriver dette som at fagfellesskapet utøver og utvikler «det profesjonelle skjønnnet.»

4.3.5 Betydningen av fagsamarbeid i vurdering av heldagsprøven

Det er et gjennomgående funn blant norsklærerne vi har intervjuet at disse lærerne vektlegger fagsamarbeidets betydning for vurdering av heldagsprøven i særdeleshet og vurdering i sin alminnelighet. Selv om disse lærerne sier at de oppfatter vurderingskriteriene de bruker som svært gode, gir kriteriene samtidig rom for tolkning som må utvikles og kvalitetssikres i faglig samarbeid med kolleger. Der norsklærerne vektlegger faglig samarbeid i etterkant av heldagsprøven, fremhever matematikklærerne særlig betydningen av fagsamarbeid i det å utvikle heldagsprøven. «En god vurdering krever en god prøve. Det gjør vurderingen enklere,» sier en matematikklærer på en offentlig skole. Han utdyper dette med å si at en prøve er god når den inneholder oppgaver som svarer til prøvens tema, og som lar elevene vise kompetanse på ulike måter. En matematikklærer på en annen offentlig skole sier at utfordringen er å konstruere prøven slik at den kan gjenspeile kompetansen til eleven, og på den måten kan brukes til å innplassere eleven på mestringsnivå og hjelpe læreren i arbeidet med å nyansere karakterene. «Hvis vi er flere lærere som har samme kurs, lager vi prøven sammen og forbereder hvordan vi skal vurdere svarene deres,» sier en tredje matematikklærer, som arbeider på en privat skole.

Vi har ikke holdepunkt i materialet vårt for å si at matematikklærere samarbeider mindre enn hva norsklærerne gjør i etterkant av heldagsprøven. Likeledes kan vi ikke si, med utgangspunkt i de skolene vi har besøkt, at det er mer eller mindre lærersamarbeid på private skoler sammenlignet med offentlige skoler. Samtidig har vi gode holdepunkter i materialet for at måten alle lærerne omtaler matematikkfaget på, legger en hovedsakelig aritmetisk forståelse av vurderingssituasjonen til grunn. Vi tror at dette kan legge til rette for mer individuelt rettellesarbeid blant matematikklærerne enn det som er tilfelle blant norsklærerne. En matematikklærer sier for eksempel at han tror de fleste (matematikk)lærere vil bedømme en elevbesvarelse i dette faget ganske likt. En norsklærer sier det slik: «Det er veldig mye subjektive ting i en tekst. Derfor er det veldig viktig at vi kan samarbeide om ting vi er usikre på, mens i matematikk er det kanskje mer 'to streker under svaret'.»

Selv om denne påstanden, at det er «to streker under svaret» i matematikk, er gjengs for alle lærerne i utvalget vårt, forteller også matematikklærerne som tidligere nevnt om behov for analytiske tilnærminger til vurdering. Da kan det bli nødvendig å konsultere kolleger, og særlig hvis man er i tvil om karakter. For matematikklærerne er det viktig å ha konsensus blant kolleger om hvordan man skal vekte elevenes forståelse for faget versus slurve- eller regnefeil som kan oppstå i løsningen av oppgaven. Videre ønsker matematikklærerne lik praksis med kolleger i bedømming av formale feil. Formalia i matematikk defineres av en lærer på en offentlig skole som å kunne bruke faglige termer, korrekte benevnelser og definere alt man skal gjøre i en besvarelse før man utfører regneoperasjonen. Hvis man som lærer kun forholder seg til fasit i matematikk, vurderer man ikke forskjeller på følgefeil og reell forståelse, sier flere matematikklærere i intervjuene. «Da går det veldig fort å rette en oppgave, hvis du retter på den måten der, altså,» sier en av dem, som arbeider på en privat skole.

Alle matematikklærerne vi har snakket med, sier at eleven skal få uttelling i poeng for å ha brukt riktig metode selv om tallverdien er feil. «Da har eleven fått litt trekk, men det kan fort være jeg som helhetsvurdering tenker at eleven ikke skal trekkes så mye i karakter som eleven er trukket i poeng,» sier en av disse lærerne, ansatt i offentlig skole. «Jeg ser etter hvordan eleven har løst oppgaven. Er det på den tungvinte måten, er det småfeil, eller er det store og alvorlige feil?» sier en matematikklærer på en privat skole, og fortsetter: «Det er den type ting du ser etter for å bestemme om du skal vippe eleven opp eller ned.»

Selve lærersamarbeidet i matematikk kan foregå ved at en lærer lar en kollega se på en elevbesvarelse, ved at flere lærere ser på besvarelsen i fellesskap, eller at kolleger presenterer enkeltproblemer eller utfordringer i elevbesvarelsene for hverandre for å få råd og lik vurderingspraksis. Vi har også fått høre eksempler på lærere som kommuniserer på e-post eller Skype med kolleger i retteprosessen. På alle skolene forteller matematikklærerne om rutiner for kollegasamarbeid i rettingen, men vårt inntrykk er at dette samarbeidet kanskje i noe mindre grad er institusjonalisert enn hva som er tilfelle for fag- og rettefellesskapet blant norsklærerne når vi ser alle skolene under ett. På noen skoler arbeider alle matematikklærerne systematisk sammen om problemløsning i rettingen, mens på andre skoler vil kanskje to matematikklærere som har samme kurs, ha laget seg rutiner for å jobbe sammen. Av vårt materiale kan vi ikke se systematiske forskjeller i hvordan slikt kollegasamarbeid foregår på offentlige kontra private skoler. Mer sannsynlig er det at slike rutiner henger sammen med størrelsen på kollegiet, erfaringer med institusjonalisert samarbeid innad i kollegiet og hvilke pålegg kollegiet eventuelt har fått fra skoleledelsen.

4.3.6 Oppsamlingsprøver, vippeprøver og ekstra vurdering av strykkandidater

Oppsamlingsprøver, det vil si ekstraprøver for elever som ikke var til stede på heldagsprøven, praktiseres både på de offentlige og de private skolene i utvalget. Oppsamlingsprøvene krever lærersamarbeid i forberedelse og retting tilsvarende de ordinære heldagsprøvene, og innebærer også ekstraarbeid for lærerne. Oppgavesett til oppsamlingsprøve må endres fra den opprinnelige prøven samtidig som begge prøver skal teste de samme kompetansemål og bidra til sammenlignbare nivåer i elevens resultater. Lærerne på en av de private skolene forteller oss at de ikke har plikt til å holde oppsamlingsprøver, men at slike prøver likevel er rutine. «Jeg føler ofte at vi strekker oss langt overfor elever som har høyt fravær og mangler vurderingssituasjoner,» sier en norsklærer på denne skolen. «Jeg tenker at vi har blitt så snille nå at noen kanskje til og med spekulerer i å ikke komme til heldagsprøven, fordi de vet at det kommer en utsatt prøve,» legger han til.

Det er flere av skolene i materialet vårt som har institusjonalisert rutiner for å justere informasjonen som skal ligge til grunn for standpunktvurdering også etter heldagsprøven. Vi har allerede fortalt om den private skolen som tillater elever å søke om å få ta ny heldagsprøve med mulighet for justert standpunktarakter, mens lærere på to andre skoler (en offentlig og en privat) beretter om fagsamtaler mellom faglærere og elever som kan gi eleven mulighet til å vise mer presis eller forbedret kompetanse i de målene som lå til grunn for oppgavene i heldagsprøven. Her oppfatter vi det som om justering kan gjelde for alle elever på alle karakternivåer.

I forlengelse av disse rutinene finnes praksiser som skoler og lærere har utviklet for å gjøre ytterligere vurderinger av elever. Dette kan inntre for elever som enten har strøket på heldagsprøven, der hvor lærer mener at det ikke finnes tilstrekkelig grunnlag for å skille mellom stryk- og ståkarakter i faget, eller bli aktuelt der elevene (uansett karakternivå) befinner seg i en vippeposisjon mellom to karakterer. I vårt materiale praktiseres ikke dette med noe påviselig skille mellom offentlige og private skoler. Vi legger likevel til at vi her ikke har godt nok grunnlagsmateriale for å si at det slett ikke eksisterer noe skille mellom offentlig og private skoler med hensyn til hvor mange og hva slags vippeprøver lærerne tilbyr elevene. Tema om vippeprøver kom opp i intervjusituasjonene uten at dette var særlig forberedt eller planlagt fra vår side på forhånd, slik at spørsmålene om dette emnet kunne vært planlagt bedre og datainnsamlingen på dette punktet kunne vært mer systematisk – og helst på et større utvalg av skoler. Samtidig kan vi med utgangspunkt i materialet vårt slå fast at der lærerne forteller om slike rutiner, praktiseres disse rutinene likevel ikke nødvendigvis likt innenfor samme skole, eller av lærerne innenfor samme fag.

På en offentlig skole forteller for eksempel matematikklærerne om rutiner som gir mulighet til skriftlig eller muntlig vippeprøve for elever på alle karakternivåer hvis resultatet av heldagsprøven er uklart. Norsklærerne på samme skole praktiserer ikke slike vippeprøver, men vil legge andre vurderingssituasjoner i løpet av skoleåret til grunn for standpunktarakteren dersom resultatet av heldagsprøven er uklart. På to andre offentlige skoler forteller matematikklærerne om at det tidligere har vært gitt mulighet til vippeprøve for elever som lå an til stryk. Etter at disse begge lærerne fikk

klage fra foreldre som mente at også elever med sterkere prestasjoner burde blitt gitt mulighet til å forbedre karakteren, er disse lærerne imidlertid i tvil om de vil arrangere slike prøver igjen. En av disse lærerne legger til at hun etter klagen har sørget for å ha flere vurderingssituasjoner gjennom skoleåret, slik at hun sikrer vurderingsgrunnlag for standpunkt uten å tilby elever vippeprøve.

Hvis eleven vipper mellom stryk- og ståkarakter, men ikke benytter seg av disse formelle sjansene, settes det som regel strykkarakter. Noen lærere vil imidlertid gjøre et siste forsøk på å legge til rette for en ny vurderingssituasjon. En norsklærer på en offentlig skole forteller at hun i slike tilfeller vil forsøke å få eleven til å levere et arbeid hvor det er mulig å sette ståkarakter. En matematikklærer på en annen offentlig skole sier at han gir strykkandidater en siste mulighet til å ta en prøve for å vise minimumskompetanse for å sette ståkarakter. Som forberedelse til prøven gir han dem en liste over hva de må kunne. Enkelte lærere nevner også at de kan bruke muntlige vurderingssituasjoner på tampen av året for å redde en elev fra å stryke. Dette omtales som en uformell vippeprøve. «Hvis resultatene har vært dårlige gjennom hele året og han stryker på heldagsprøven, så stryker han,» sier en matematikklærer på en privat skole. «Men har det vært litt variasjon, så kaller jeg ham inn til samtale for å gi ham en 2er.»

I avveining mellom stryk- og ståkarakter vil lærerne også søke råd hos kolleger. Dette gjelder både for norsklærere og for matematikklærere, som for eksempel vil be kollegaen om hjelp til å finne minimumsgrunnlag i elevens tidligere prøver og innleveringer slik at vedkommende kan få berettiget ståkarakter. Her understreker lærerne, både på offentlige og private skoler, at det vil være forskjell på tilrettelegging for elever som har hatt god grunn til fravær eller svake skoleprestasjoner (sykdom eller personlige problemer) og elever som rett og slett ikke møter frem. På flere av skolene, primært de offentlige, sier lærerne at de opplever et visst press på gjennomstrømning, det at elever skal gjennom videregående opplæring uten forsinkelser. «Jeg synes det skal mye til for å stryke når det er standpunkt. Spesielt når det er 3. klasse og du ikke kan gjøre noe [for eleven]. Altså, går du i 2. klasse og får stryk i standpunkt, så har du et år på å gjøre noe med det og likevel få et førstegangsvitnemål,» sier en norsklærer på en offentlig skole. Dette kan i noen grad forklare noen av de rutinene som utvikler seg der lærere leter etter begrunnelser for å la en elev stå, men faglærere forteller også om at de vil det beste for eleven. I praksis fører dette til at faglærerne snevrer inn bunnen av karakterskalaen, slik at også de aller svakeste elevene kan gå ut av skolen med en ståkarakter.

Alt i alt ser vi særlig i dette feltet et mangfold av praksiser som beveger seg i feltet mellom formaliserte rutiner (oppsamlingsprøver, fagsamtaler) og rutiner i daglig bruk (variasjoner av vippeprøver, muntlige vurderingssituasjoner). Fra et forskerståsted er det vanskelig å skille mellom de ulike nivåene av rutiner og praksiser i dette feltet. Vi antar likevel at en del av de ikke-formaliserte praksisene har utviklet seg nettopp som en støtte eller et supplement til de ordinære rutinene for å dekke behovene som lærerne har, men vi ser også muligheten for at lærerne her kan føle seg presset av elever til å tilby nye vurderingssituasjoner. Som et resultat av praksisene ser vi at særlig svake elever kan unngå stryk og vippes opp på karakteren 2. En av årsakene til at det blir slik, skal vi komme nærmere inn på senere i kapittelet, når vi tar for oss lærernes erfaringer og opplevelser av det å følge elever et helt år som faglærer og hvordan dette kan få betydning for avvik mellom standpunkt-karakter og eksamens-karakter.

4.4 Sluttvurdering på bakgrunn av underveisvurderinger og samtaler med elevene

Slik vi har gjort rede for så langt i kapittelet, sier norsk- og matematikklærerne at heldagsprøven vil gi det aller viktigste grunnlaget for standpunkt-vurdering, men at det også kan bli aktuelt å se på vurderingssituasjoner fra tidligere i skoleåret. Dette gjelder særlig hvis heldagsprøven gir et uklart resultat med hensyn til karakternivå, eller læreren mener å se at eleven har misforstått en eller flere oppgaver på heldagsprøven slik at elevens reelle forståelse og kompetanseoppnåelse i faget ikke avspeiles. «Den siste prøven blir den viktigste, for den viser jo totalkompetansen, men så er det lov til å ha en dårlig dag. Du kan feile på den siste prøven og likevel få en god karakter hvis det viser seg at

du har hatt et jevnt og høyt nivå det siste året,» sier en matematikklærer på en offentlig skole. Derfor skal vi også se på hva lærerne sier om vurderingssituasjoner gjennom året, inkludert samtaler for undervisningsvurderinger. Dette gir etter vår mening et ytterligere innblikk i hvordan lærerne arbeider frem mot den endelige standpunktkarakteren.

4.4.1 Tilbakemeldinger til elever etter prøver

I intervjuene kommer norsk- og matematikklærerne selv inn på hvordan de arbeider med å gi elevene tilbakemeldinger etter prøver underveis i skoleåret. Mange av lærerne reflekterer i intervjuet over hvordan elevene tar imot tilbakemeldingene de får etter prøver, og i hvilken grad disse tilbakemeldingene hjelper elevene til å bli bedre. «Elever er generelt opptatt av karakterer», sier en norsklærer på en privat skole, «men det er ikke alltid de forstår hva en god oppgave egentlig er.» «Mange elever har jo gjerne et ønske om å få den og den karakteren, og de ser på en måte ikke at det ikke er realistisk, for de ser ikke hvor de selv er hen,» sier en kollega av denne norsklæreren. Lærerne vil søke å diagnostisere elevens problemer. En matematikklærer som ble intervjuet sammen med disse to norsklærerne, sier:

Når jeg går rundt til elevene, så sier jeg jo at 'grunnen til at du gjorde feil var fordi du ikke leste oppgaveteksten skikkelig, eller her har du skrevet av feil fra oppgaven. Eller, nå ser jeg at du har pugget en metode, og så prøver du å bruke den uansett, men her fungerer ikke det.' Men så er det noen som du ser, som jobber og jobber og jobber, de gjør lekser, de møter til timen og er flinke, men så får de kanskje ikke resultatene de vil likevel. Du kan jo ikke si til dem at de må lese en time ekstra til neste prøve for å få en bedre karakter, det er noe annet som må til, og som lærer må du prøve å diagnostisere det.

Mange av lærerne forteller at de bruker de formelle vurderingskriteriene for å gi elevene tilbakemelding etter prøver gjennom skoleåret, men at elevene enten ikke er interessert i å sette seg inn i det som står der eller at de ikke forstår det. «Tilbakemelding er vanskelig, for du må ikke si for mye og heller ikke for lite,» sier en av norsklærerne på den private skolen som vi har sitert over. «Man må finne ut hva man skal legge vekt på, og så plukke ut to eller tre punkter som eleven må jobbe med til neste gang.» En norsklærer på en offentlig skole forteller at han bruker eksempeltekster for å demonstrere gode besvarelser for elevene:

Jeg bruker mye eksempeltekster som enten er vurdert ved eksamen eller i rettelleskap på skolen for å vise hva som skal til. Når elevene får lest disse tekstene og vi snakker om dem, så er de egentlig like gode til å vurdere som det lærerne er, de er veldig flinke til å bedømme nivået på tekstene. Det er jo også veldig nyttig for å vise dem noen grep de kan gjøre, samtidig som det er enkelt å snakke om dette, men vanskeligere å gjøre det. Men de ser i hvert fall hva som forventes.

Det er flere lærere, både på offentlige og private skoler, som forteller at de lar elevene få rette og gi forslag til vurdering av egne besvarelser for å gi dem innsikt i vurderingskriterier. Elevene kan også få i oppgave å sette sammen vurderingene han eller hun har fått i løpet av året, med tanke på å bli oppmerksom på ting som går igjen og som de må øve på frem mot eksamen. Når slike øvelser blir gjort i matematikk, bemerker disse lærerne at elevene gir inntrykk av å ha god forståelse for vurderingskriteriene, og at de også vurderer egne besvarelser strengere enn hva læreren deres ville ha gjort. Akkurat dette kan bero på den analytiske helhetsforståelsen som også matematikklærerne sier ligger til grunn for vurdering og som disse lærerne behersker, mens elevene kanskje vil vurdere besvarelsen sin utelukkende på aritmetisk grunnlag. «I matematikk er det jo enkelt,» sier en matematikklærer på en offentlig skole. «Elevene ser jo hva de har gjort feil og hvilke oppgaver de sliter med, så det er enkelt å rettlede dem frem mot en årsprøve.» En matematikklærer på en annen offentlig skole sier med et lite smil: «Det er derfor jeg er så glad jeg er mattelærer og ikke norsklærer, fordi alle forstår jo at det er noe helt annet å skulle grunngi en karakter i norsk i forhold til matematikk.» Norsklærerne kan da også fortelle om en betydelig svakere forståelse blant elevene for vurderingskriteriene de bruker. Kompetansemålene med tilhørende vurderingskriterier i norsk er

kvalitative, og vår tolkning som forskere er at elevene ikke har like lett for å forstå den analytiske vurderingsmåten som de tilsynelatende forstår den aritmetiske måten å vurdere på. «Norskfaget er jo mer enn å kunne skrive og snakke,» som en norsklærer på en privat skole uttrykker det.

De fleste av lærerne vi snakket med, har systemer for å loggføre karakter, vurderinger eller tilbakemeldinger, og dessuten hvorvidt det har vært noen kommentarer fra kollegaer eller samtaler med eleven om den aktuelle prøven eller innleveringen. Noen skoler har dette formalisert i den digitale læringsplattformen, andre lærere har et eget arkiv- og loggføringssystem for dette. En norsklærer på en offentlig skole som bruker den elektroniske læringsplattformen It's learning til dette formålet, sier: «Der har du leksene og vurderingene, alt som er av kommentarer som vi har sendt til eleven. Det er ganske oversiktlig, det viser hvor eleven ligger.»

Kollegasamarbeid er også nødvendig for lærerne i den løpende vurderingen av elevene. Mange av lærerne forteller om praksiser for blindretting – at en kollega ser på en elevbesvarelse uten å kjenne elevens identitet – men dette kommer som hovedregel bare til anvendelse hvis den første faglæreren er i tvil. Enkelte lærere forteller om elever som til tross for gjentatte vurderingssituasjoner med skuffende resultat, insisterer på å få besvarelser vurdert på nytt av andre lærere i håp om at karakteren skal gå opp. Da kan det være en styrke for læreren at kolleger har rettet prøven blindt. For læreren styrker dette reliabiliteten i vurderingsarbeidet. På noen skoler, blant de offentlige så vel som de private, har lærerkollegiet diskutert anonym retting av prøver, men dette er ikke satt i verk som et fullskalaprojekt på noen av skolene som vi har besøkt.

Lærere som i intervjuet reflekterer over tvilssituasjoner de har stått i med hensyn til vurdering, understreker alle at de lar slik tvil komme eleven til gode og vipper ham eller henne opp hvis det står mellom to karakterer. Flere av dem sier at hvis det skal brukes ord som «trynefaktor», så innebærer det at de lar tvil komme eleven til gode. Ellers understreker både norsklærerne og matematikklærerne at felles vurderingspraksis er det beste middelet mot såkalt «trynefaktor»-vurdering av elevene. Noen av lærerne innrømmer imidlertid at det kan være elever man liker mindre enn andre. En lærer sier at han vil anstrenge seg i særlig grad for å være hyggelig mot disse elevene, og ikke behandle dem annerledes enn andre. Blant lærerne som snakker om slike forhold, er det ellers gjengs oppfatning at antipatier mot elevene må motarbeides.

4.4.2 Formelle og uformelle elevsamtaler

Alle skolene gjennomfører de lovpålagte elevsamtalene (også kalt fagsamtaler og vurderingssamtaler) underveis i skoleåret, en samtale pr. semester. Vi har ikke så mye data fra lærernivå i materialet vårt om hva disse samtalene inneholder, bortsett fra at standpunkt-karakteren og arbeidet fram mot denne er naturlige emner for samtalene. «Typisk innhold for en vurderingssamtale er at vi snakker om grunnlaget for standpunkt-karakteren og de målene en elev har,» sier en matematikklærer på en offentlig skole. «Jeg spør gjerne hva målet er i karakter, og formulerer da de tingene som eleven må rette på eller forbedre for å nå dit. Egentlig er det så enkelt,» sier en norsklærer på samme skole. Videre er det slik at elevene skal kvittere for fremmøtet, slik at man i ettertid kan si at eleven har fått en formell vurderingssamtale.

Noen lærere vi har snakket med, gir uttrykk for at de ser på de lovpålagte elevsamtalene som en rutine de er pliktige til å gjennomføre, uten at innholdet i samtalene nødvendigvis har stor betydning. Derfor var det interessant å erfare gjennom intervjuene våre hvor mye vekt lærere legger på de uformelle samtalene med elevene. «Det er jo minst like viktig, det som ikke står i kalenderen,» sier en lærer på en privat skole. Disse samtalene kan skje i etterkant av en undervisningstime, i gangen, i friminuttet eller ved at eleven oppsøker læreren på arbeidsrommet utenfor undervisningen. E-post og kommunikasjon mellom lærer og elev på elektroniske læringsplattformer kan også regnes inn blant disse uformelle samtalene. Lærerne som nevner slike uformelle møter med elevene, tillegger også disse samtalene stor betydning for å gi eleven tilbakemelding og eventuell oppmuntring. Hvis eleven oppsøker læreren etter undervisningen, kan det være for å få hjelp til et bestemt problem. Da kan læreren gi råd om hva eleven kan gjøre for å få større forståelse for faget, samtidig som samtalene

også gir læreren informasjon om elevenes utvikling. Også den løpende samtalen i klasserommet fremheves av lærerne i intervjuene for at lærer kan få slik informasjon. På denne måten mener lærerne at de kan få god oversikt over elevens jevne forståelse og nivå i faget også *utenfor* de formelle vurderingssituasjonene. En lærer som jobber på en privat skole setter dette spøkefullt på spisen ved å si at hvis man som lærer holder god kontakt med elevene og lar muligheten for samtale være åpen underveis i skoleåret, kan man spare ti minutter av den formelle vurderingssamtalen hvert semester.

Vektleggingen av uformelle møter og samtaler med elevene underveis i skoleåret aktualiserer pålegget om å gi eleven sluttvurdering og elevenes innsats og fremmøte. Alle lærerne vi har snakket med, problematiserer disse tre forholdene når det gjelder deres mulighet til å sette det de oppfatter som en riktig standpunkt karakter. I dette understreker lærerne at de er klar over at det er sluttvurdering som skal ligge til grunn for standpunkt, men at hvis resultatet av siste eksamensforberedende prøve er uklart, må lærerens inntrykk av eleven gjennom året nødvendigvis komme i betraktning. Da er det en klar fordel at eleven har vært til stede i undervisningen, og eleven må også ha vært til stede i flere vurderingssituasjoner gjennom året. I intervjuene gjør lærerne grundig rede for rutiner for bruk av «ikke vurdert» når en elev uteblir fra en vurderingssituasjon, men problematiserer samtidig dette i lys av behovet for å ha et bredere vurderingsgrunnlag for standpunkt karakter enn hva heldagsprøven kan gi dem. Vurdering av elevers innsats og elevers fremmøte/fraværsbestemmelser i videregående skole er ikke tema for denne rapporten, men de aller fleste lærerne nevner i intervjuet at dette regelverket skaper problemer for dem med hensyn til å gjøre gode og rettferdige vurderinger av elever. Årsaken er at disse lærerne mener elevens innsats og oppmøte gir lærerne gode og valide holdepunkter for å vurdere elevenes faglige utvikling over tid.

4.5 Sensorerfaring og skoleinitierte prosjekter for vurdering

Flere av skolene vi har besøkt, har hatt egne prosjekter for å bedre lærernes vurderingspraksis. I noen grad er dette gjenkjennelig i den nasjonale satsningen på *Vurdering for læring* (Utdanningsdirektoratet), men i tilknytning til denne satsningen eller istedenfor denne har skoler også utviklet egne prosjekter. I tabell 4.1 har vi brukt svarene fra e-postintervju med rektorene for å indikere hvor mange skoler i utvalget vårt som har satsninger eller prosjekter knyttet til vurdering, og vi viser fordelingen av disse skolene på offentlig og privat sektor.

Tabell 4.1 Oversikt over skolenes eierform og arbeid med vurdering

Skole 1	Privat	Skoleinternt prosjekt for vurdering
Skole 2	Privat	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering
Skole 5	Privat	Skoleinternt prosjekt for vurdering, også samarbeid med UH-miljø.
Skole 8	Privat	Ikke eget prosjekt, men uttalt oppmerksomhet rettet mot eventuelle avvik mellom eksamens- og standpunkt karakter.
Skole 3	Offentlig	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering
Skole 4	Offentlig	Workshops knyttet til satsningen Vurdering for læring.
Skole 6	Offentlig	Satsning på vurdering i samarbeid med UH-miljø
Skole 7	Offentlig	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering

Lærerne sier som et generelt inntrykk av prosjektene at de opplever at det har gitt kollegiet større bevissthet om vurdering og en mer ensartet vurderingspraksis på skolen, men særlig på en av de offentlige skolene i materialet vårt kunne det likevel virke som om lærerne trakk et skille mellom den løpende vurderingen som skjer i lærernes daglige arbeid og gjennom kollegasamarbeidet, og et prosjekt initiert ovenfra. På en av de private skolene forteller lærerne (både i norsk og matematikk) om et lokalt prosjekt for vurdering som aktivt grep inn i kollegasamarbeidet og forandret rutinene for slikt samarbeid på en måte som lærerne opplevde som positivt. Her ble lærerne blant annet satt sammen med en diskusjonspartner som de ikke hadde valgt selv, mens praksis i mange fagseksjoner vil være

at lærere med ansvar for samme kurs eller samme trinn automatisk jobber sammen. På den aktuelle skolen tok også rektor aktivt rollen som lærernes diskusjonspartner i prinsipielle vurderinger. En annen privat skole har hatt en serie med foredrag og workshops gjennom et skoleår, hvor lærerne har fått egne oppgaver som de skulle løse, og lærerne har gitt hverandre tips og råd.

Alle disse prosjektene har vært ledet av rektor, men disse fortellingene kan gi en pekepinn om at det ikke alltid vil være samsvar mellom satsninger initiert fra et ledernivå og det lærerne legger vekt på i sin arbeidsdag. Vi vil tro at det kan bero på satsningens innhold og organisering, og om den aktivt griper inn i gjeldende vurderingspraksis med oppgaver og refleksjon, eller forblir på foredragsnivå.

På noen skoler, og først og fremst på de private skolene vi har besøkt, har vi fått vite at man i lærerkollegiet har hatt diskusjoner om vurderingspraksis og avisoppslagene som viser at standpunkt og eksamenskarakter har større avvik på disse skolene sett under ett. Dette spørsmålet er imidlertid ikke noe som bare opptar lærere og ledere på private skoler, da vi også har offentlige skoler i vårt materiale hvor lærerne kan berette om slike diskusjoner. Vi skal komme tilbake til lærernes synspunkter på slike avvik senere i dette kapitlet.

4.5.1 Betydningen av sensorerfaring

Over halvparten av de intervjuede lærerne på de åtte videregående skolene har sensorerfaring fra sentralt gitt skriftlig eksamen i enten norsk eller matematikk. Blant de øvrige lærerne er det flere som har erfaring fra muntlig eksamen, eller lokalt gitt eksamen i norsk, matematikk eller andre fag. Blant lærerne som har erfaring som sensor på skriftlig eksamen, har flesteparten gjennomført en form for sensorskolering eller møter og samlinger for sensorene hvor felles vurderingspraksis har vært tema. Gjennom intervjuene har det ikke vært mulig for oss å avdekke nøyaktig hvorvidt lærerne snakker om formelle sensorskoleringer eller formelle sensormøter, og det er også ofte uklart i intervjuene hvem som har vært arrangør av slike møter eller tilbudt kurs for sensorer. Likevel velger vi i det følgende å behandle disse erfaringene som lærerne har fra slike møteplasser under ett, fordi de i alle tilfeller peker mot at lærerne synes tilbudene er svært nyttige og sier de har stor utbytte av dem for sin egen vurderingspraksis i det løpende arbeidet. Materialet avdekker ingen forskjeller mellom lærere på private eller offentlige skoler i dette spørsmålet. «Med tanke på det å sette standpunktkarakter, vil jeg si at hvis man har sensorerfaring, så er man mye, mye tryggere i det jevne vurderingsarbeidet,» sier en norsklærer på en privat skole. «Du får mer faglig tyngde, du kan veilede og rette bedre,» sier en norsklærer på en annen privat skole. En annen norsklærer med lang erfaring fra både privat og offentlig skole (nå ansatt på privatskole) viser til hvordan sensorerfaring kan komme ikke bare den enkelte lærer til gode, men hele fagfellesskapet på skolen:

Vi bruker jo mye tid på å få til dette tolkningsfellesskapet i faget innad på skolen, det at vi vurderer på samme måte og diskuterer karakterer. Og da tenker jeg at den erfaringen de har, de som har vært sensor til sentralt gitt eksamen, de tar jo med seg den kompetansen inn i kollegafellesskapet. Så dette er viktig, ikke bare i egen vurderingspraksis, men også fordi kunnskap og kompetanse kan tilflyte andre. Mange hos oss har lang erfaring, men noen er ferske og skal sette standpunktkarakterer for første gang. Det er viktig for oss å følge dem opp.

Denne interessen og anerkjennelsen av sensorerfaringen bekreftes også av lærere som ikke har vært sensor til skriftlig og sentralt gitt eksamen, og noen av dem sier i intervju at de gjerne kunne tenke seg slik erfaring eller ser fram til å få den gjennom å delta i sensur våren 2014. En av disse sier at han tror sensorerfaring vil gjøre ham tryggere i egen karaktersetning. Lærere omtaler også sensorerfaringen ikke bare som et redskap til bedre vurdering, men også som en kalibrering av hva man bør legge vekt på i egen undervisning. «Jeg synes det er en utrolig grei måte å se nivået på, altså hva som kreves på eksamen,» sier en norsklærer på en privat skole. En matematikklærer på en offentlig skole følger opp dette når han sier at han opplever slik erfaring som en nyttig korreksjon til ham selv, som også kommer kollegene hans til nytte. En annen matematikklærer på en privat skole sier at han bedre kan rettlede elevene sine frem mot eksamen når han gjennom sensorarbeid får stadfestet hvordan

oppgaver skal føres for å få maks antall poeng. Imidlertid har vi også lærere i materialet vårt som sier at de ikke ønsker sensorarbeid på grunn av arbeidsmengden det medfører. Denne arbeidsmengden erkjennes og beklages også av lærerne som sier de har stor nytte av slik erfaring, men disse lærerne vurderer altså denne erfaringen som så verdifull at de likevel vil ta imot 200 besvarelser til retting. Som en av disse lærerne, ansatt på en privatskole, sier: «Det er jo fellessensuren og skoleringen du har utbytte av og som du har noe igjen for, mens alle de timene med rettearbeid, det er jo ikke så veldig stimulerende i seg selv.»

Det er ikke alle lærerne med erfaring fra skriftlig, sentralt gitt eksamen som har formell sensorskolering, men alle med sensorerfaring kan berette om møter med andre sensorer. Lærerne som sier at de har deltatt i sensorskolering, sier at dette har vært svært nyttig, og foreslår at slik skolering kunne være obligatorisk både fordi det gir så stor utbytte i egen vurderingspraksis, men også fordi de tror det vil gi mer ensartet vurdering på tvers av skoler. I norsk kan man være opptatt av forskjellige ting, sier en erfaren norsklærer med mange sensorskoleringer bak seg, og derfor er fagsamarbeidet, også gjennom sensorskolering, så viktig. Denne læreren arbeider i offentlig skole:

Du kan være opptatt av refleksjon og jeg kan være opptatt av struktur, men sammen utgjør det en komplett kompetanse. Det er også dette som blir presisert når vi er på sensorskoleringene, at vi ikke skal ta med oss våre «hangups» og vektlegge det i vurderingen, vi må forholde oss til flere kriterier. Og det er ganske utfordrende når man er sensor, og man gjerne har sin «greie». Uansett tror jeg det er veldig nyttig at vi jobber i team og at vi hjelper hverandre i vurderinger og bevisstgjør hverandre på hva som er viktig.

Vi har allerede skrevet at det er vanskelig for oss å skille formell sensorskolering og formelle sensurmøter fra hverandre i materialet vårt, men dette mener vi heller ikke betyr noe for vår analyse fordi lærerne legger uansett vekt på det å møtes og å diskutere oppgaver for å enes om felles praksis. «Jeg er veldig glad i denne dagen,» sier en norsklærer på en offentlig skole, «jeg kan sitte sammen med en hel haug andre norsklærere og diskutere tekster. Det er veldig nyttig og lærerikt for meg, og det bygger opp under et fellesskap.» Denne læreren sier at han søker et tolkningsfellesskap også utenfor skolen, og dette behovet får han delvis oppfylt gjennom ekstern sensorskolering/møter for fellessensur. Med bakgrunn i materialet vårt kan vi se disse uttalelsene i kontrast til at ingen av lærerne vi har snakket med, sier de har fått noen annen formell opplæring i det å sette karakterer – verken gjennom utdanning eller som nyansatte lærere. «Det er i teamet og på skolen jeg har lært alt om vurdering,» sier en norsklærer på en offentlig skole. Flere av lærerne stusser over at det er slik, men viser samtidig til den uformelle opplæringen de har fått gjennom deltakelse i et fagfellesskap. Mange av dem har fått hjelp av eldre lærere det første året til å bli trygge i egen vurderingspraksis. Her understreker samtidig lærerne at det utvikler seg «skolekultur» for vurdering som overføres fra eldre til yngre kolleger. Vi som forskere antar at denne skoleinterne vurderingskulturen, som kan utvikle seg på godt og vondt, brynes, justeres og bekreftes i sensormøtene og gjennom sensorskoleringen.

Gjennom flere av disse eksemplene vi nå nettopp har gjengitt, ser vi at sensur kan skape møteplasser og fagfellesskap som mange lærere sier er en styrke for profesjonalitet i læreryrket generelt og vurdering spesielt. Samtidig er det en matematikklærer på en privat skole i materialet vårt med lang sensorerfaring som sier at han opplever nytten av slike fagfellesskap som avtagende når han har deltatt på slike møter gjennom flere år:

Om vi kommer så mye nærmere en felles forståelse, det vet jeg ikke helt. Jeg kunne godt tenke meg mer sentral styring av hvordan vi skal rette, hva vi skal legge vekt på. Disse møtene blir bare en diskusjon med forskjellige personer hver gang, som fremmer sine oppfatninger.

Det er interessant, også som en overgang til avslutningen av denne diskusjonen om rutiner, når en del av lærere med sensorerfaring i intervjuene diskuterer det de kaller «sensorbriller». Dette beskriver det blikket en sensor har når han eller hun sitter med oppgavebunken foran seg, uten å kjenne elevenes identitet. Disse lærerne kontrasterer det å sitte med en anonym bunke mot det å være faglærer og

skulle gi elevene standpunktkarakter, og dette tjener også som overgang i dette kapittelet til å se nærmere på hva lærerne mener om eventuelle avvik mellom eksamens- og standpunktkarakter. «Jeg kaller det sensorbriller,» sier en norsklærer på en privat skole:

Jeg er jo strengere som sensor, tror jeg, enn jeg er som faglærer. Det menneskelige aspektet, det er vanskelig, altså å ikke skulle ta det med. Som sensor vurderer du jo akkurat den oppgaven eleven har gjort på disse fem timene. La si at de har skrevet en sammenligning av to dikt fra modernistisk og realistisk tradisjon. Og bare det at de har gjort det på fem timer, er jo helt vanvittig godt, spør du meg. Så egentlig burde man være mindre streng som sensor, men det er jo selvfølgelig lettere å være pirkete og streng når du sitter der med en helt upersonlig besvarelse.

En kollega av denne læreren følger opp i samme intervju med å beskrive hvordan en faglærer blir kjent med elevene gjennom de daglige møtene i undervisning og underveisvurdering:

Du blir jo etter hvert vant til stilen elevene skriver i, du vet at de har en bestemt skrivemåte, men du vet også at hvis du klarer å legge godviljen til, hvis du setter deg ned og leter etter den, så finner du den røde tråden i besvarelsen. Som sensor får du en tjukk bunke, og hvis den røde tråden eller sammenhengen ikke er der sånn umiddelbart, så ser du den ofte ikke heller.

En privat ansatt norsklærer med lang erfaring også fra sensorarbeid, sier at man alltid skal være så objektiv som mulig, og sette karakter på prestasjon og ikke på person. «Likevel er jo dette enklere når du ikke har et navn å forholde deg til,» sier hun.

Alle lærerne vi har snakket med i intervjuene, oppgir at de benytter seg av hele karakterskalaen til standpunktvurdering. Lærerne synes det er lettest å vurdere besvarelser som ligger i ytterpunktene av skalaen, mens det krever mer arbeid å grunngi karakteren 3 mot 4 enn for eksempel karakteren 5 mot 6. Tidligere i dette kapittelet har vi likevel vist hvordan bunnen av skalaen kan snevres inn til standpunktvurdering, der elever som står i fare for å stryke får stadig nye sjanser. Med «sensorbrillene» på, kan imidlertid toppen av karakterskalaen snevres litt inn. En erfaren norsklærer på en offentlig skole sier:

Når jeg selv sensurerer, så ser jeg at det er litt lenger mellom 6erne enn det kanskje er i enkeltvurderinger i løpet av året. I en bunke på 200 besvarelser er det ikke mange 6ere. Jeg er faktisk litt usikker på om 6 til eksamen ligger litt høyere, eller hva som er grunnen til det. Men helt ærlig, du ønsker virkelig at eleven din skal få den 4eren, du vil gjerne se den utviklingen. Og så tenker man gjerne at man gir den høyeste av to karakterer mot slutten.

Denne læreren, som her gjør en ærlig refleksjon over sin vurderingspraksis, viser både til at kollegasamarbeid er viktig for å gi rett vurdering i slike tilfeller av standpunktvurdering, men også til et gjennomgående funn i vårt materiale: alle lærerne mener at det er forskjell på å sette standpunktkarakter og eksamenskarakter. Dette skal vi ta for oss nedenfor.

4.6 Forskjeller mellom standpunktkarakter og eksamenskarakter

Ovenfor har vi gjengitt tre sitater fra faglærere som også har vært eksamenssensor, som viser hvordan faglærere blir kjent med elevene sine og deres faglige utvikling, og ønsker dem de beste resultatene. Lærerne vi har snakket med, omtaler eksamen som en stikkprøve på elevens kompetanseoppnåelse, hvor eleven kan være heldig eller uheldig med oppgavene, ha en god eller dårlig dag. «Alt skal funke den dagen,» sier en norsklærer på en privat skole, «da skal hodet være tipptopp, du skal ha hatt en god natts søvn, du skal ha spist frokost, du skal ikke ha kjærlighetssorg.» En matematikklærer sier i samme intervju: «Eksamen er jo en spesiell form å teste elever på. En elev kan jo være god i matte selv om han ikke er flink til å sitte og jobbe i timer i strekk med matematikkoppgaver.» Dette er forhold som for flere av lærerne begrunner påstanden når de sier at

de opplever standpunktkarakteren som en «riktigere» karakter enn eksamenskarakteren når det gjelder vurdering av elevens kompetanse. En matematikklærer på en offentlig skole uttrykker dette som at eksamen er et sjansespill for eleven, mens standpunkt bygger på lærerens vurdering og gir et mer helhetlig inntrykk av eleven. «Dette kan ikke bli identisk, det sier seg selv,» sier en norsklærer på en annen offentlig skole.

Selv om alle lærerne vi har snakket med, mener at vurderingen bak standpunktkarakteren skiller seg fra vurderingen bak eksamenskarakteren, er det svært ulike meninger i dette utvalget av lærere når det gjelder hvorvidt disse to karakterene bør være sammenfallende. Lærere på et par av skolene sier at avvik opp mot 0.5 mellom de to karakterene er alt for mye, mens lærere på andre skoler setter dette til et avvik på 1.0. Mellom dette finnes det lærere som ikke ønsker å tallfeste hvor stort et eventuelt avvik kan være for at det regnes som akseptabelt, men de sier de registrerer et ønske eller en føring som tilsier at avviket bør være så lite som mulig. Ved noen av skolene er det klart at denne føringen kommer fra ledelsen, mens det i andre tilfeller kan dreie seg om mediedebatt om emnet. En del av lærerne, uavhengig av disse posisjonene om størrelse på avvik, synes dette er en bra og riktig føring som kan gi bedre vurderingspraksis på skolen, mens andre holder fast ved at eksamen og standpunkt er to ulike størrelser som riktignok kan ses i sammenheng, men hvor man ikke skal forvente fullstendig samsvar. I materialet vårt er det følgelig også lærere som omtaler dato for eksamenssensur som «årets høydepunkt», fordi de da får grunnlag for kalibrering av egen undervisnings- og vurderingspraksis. Likeens inneholder materialet vårt uttalelser fra lærere som avviser at eksamen kan oppleves som en «fasit» på deres eget undervisnings- og vurderingsarbeid. «Man må jo måle seg opp mot noe,» sier en matematikklærer på en privat skole: «For meg er det veldig nyttig å se, jeg blir veldig glad hvis elevene mine gjør det omtrent like bra på eksamen som det jeg har gitt dem. Jeg tenker at ok, da er jeg i hvert fall sånn høvelig på rett spor.» De ulike synspunktene som vi her har gjort rede for, fremtrer ellers i materialet vårt uavhengig av om lærerne er ansatt på en offentlig eller privat skole.

Videre er det lærere i materialet vårt som skiller mellom avvik for enkeltelever eller en liten gruppe elever og avvik som gjelder hele klasser eller trinn, og lærere som skiller mellom avvik i et år eller over lang tid. Det første, avvik for små utvalg elever eller i et bestemt år, bekymrer ingen av lærerne vi har intervjuet. Her kan det være tilfeldigheter som slår inn. Hvis avviket derimot gjelder store grupper elever eller fremtrer systematisk over flere år, vil de fleste av lærerne ta dette til etterretning og se nærmere på egen praksis. Dette avhenger igjen av hvordan lærerne ser på forholdet mellom eksamen- og standpunktkarakter, og hvor stort avvik de vil akseptere. «Standpunkt måler jo mye bredere, det måler jo ideelt sett alle kompetansemål, men eksamen måler jo bare den ene dagen på de oppgavene som ble gitt den dagen. Samtidig bør ikke spriket være for stort, i hvert fall ikke over tid,» sier en norsklærer på en privat skole hvor lærerne har det strengeste synet på avvik mellom de to karakterene. Her aktualiseres standpunkt som sluttvurdering på nytt. Matematikklæreren på en offentlig skole som vi siterte helt i begynnelsen av dette kapittelet, og som la en streng tolkning av vurderingsforskriften til grunn, sier dette om eventuelle avvik mellom eksamen og standpunkt:

Eksamen og standpunkt bør være nokså like. Hvis ikke, så mener jeg at det enten er noe galt med eksamensformen eller hvordan vi gjør det. Hvis de sier at det skal være en sluttvurdering i matematikk, så skal det ikke være veldig stor forskjell på den sluttvurderingen og den eksamenen de har noen uker senere. Så lenge det er sluttvurdering og ikke innsats og sånne ting som skal telle, så skal det være ganske likt.

Lærere på en annen privatskole, som sier de har oppfattet at det er sluttvurdering som skal gjelde, setter det på spissen ved å si at dersom standpunkt skal sammenfalle med eksamen, mener de at standpunktkarakteren like gjerne kan sløyfes. En matematikklærer på en offentlig skole, hvor både denne læreren og kolleger reflekterer over menneskelige faktorer som kan inntre i vurderingssituasjoner, står for motsatt posisjon. Han sier:

Det står jo i disse veiledningene at man skal legge vekt på elevens kompetanse på slutten av vurderingssituasjonen. Sånn sett kan man tenke at karakterene burde være ganske like. Men samtidig er det slik at når du har hatt en elev i et helt år, så får du et helt annet forhold til den

eleven enn om du bare hadde fått en eksamensbesvarelse. Jeg tror at sånn vil det bare være. Det er ikke så uventet, kanskje, at standpunkt er bedre enn eksamen, sett ut fra at du kjenner eleven. Og når du retter og setter standpunkt, så er det kanskje også et bilde av hvordan du sjøl har vært som lærer. Det kan være at det virker inn, selv om det ikke skal gjøre det. Jeg kan ikke si at jeg ikke legger vekt på det, selv om jeg prøver å ikke gjøre det. Så jeg mener at det bør være forskjell på standpunkt- og eksamenskarakter, selv om retningslinjene legger opp til at det ikke skal være det. Når du har gått på skole i et helt år, er det naturlig at standpunkt tar hensyn til hele det året, og ikke bare det du gjorde på en prøve rett før sommeren.

På en annen privat skole reflekterer to lærere, en norsklærer og en matematikklærer, over ledelsens føring i retning av at standpunkt- og eksamenskarakter skal vise større grad av sammenfall. Disse lærerne sier at de tar seg selv i å tenke på hva en aktuell elev ville fått hvis det var eksamen i dag, og la denne vurderingen bli grunnlag for standpunkt karakteren. De mener at dette er en feilaktig fremgangsmåte for å sette standpunkt karakter, som ikke nødvendigvis kommer eleven til gode. På samme skole sier en norsklærer at man som sensor og som faglærer er underlagt helt ulike dokumentasjons- og saksbehandlingsrutiner for vurderingen man gjør. Som sensor er man ferdig med jobben når man har kommet til bunnen av oppgavebunken, og ved eventuell klage vil det bli oppnevnt en ny komité. Altså forholder man seg som sensor ikke til besvarelsen eller eleven på nytt. Som faglærer må man på den annen side være forberedt på å møte klager på standpunkt karakterer, og da må dokumentasjonen av vurderingen være i orden. Dette forholdet, mener denne læreren, kan også skape ulike vurderingspraksiser for faglærer kontra sensor, og presse faglærerens standpunkt vurdering oppover.

4.7 Oppsummering: rutiner for standpunkt vurdering

Vi har i dette kapittelet sett at det er tydelige forskjeller i tilnærming til vurdering blant norsklærere kontra matematikklærere. Dette skillet følger faggrensene, og gjelder uavhengig av om lærerne jobber i offentlig eller privat videregående skole. Som en hovedtendens i materialet oppgir matematikklærerne å følge en aritmetisk tilnærming til vurdering, hvor de gir poeng for riktig svar og der karakter fremkommer på basis av prosent av totalt antall oppnåelige poeng for prøven. Norsklærerne forholder seg derimot analytisk til vurderingen av elevbesvarelser. Poengberegning er ikke i bruk i denne tilnærmingen til vurdering, som tar sikte på å gi en kvalitativ bedømming av elevens måloppnåelse med grunnlag i kompetansemålene for faget. Samtidig understreker alle matematikklærerne vi har snakket med, at en analytisk vurdering kommer til anvendelse når de skal se på helheten i en elevbesvarelse og vurdere elevens *forståelse* av matematikk opp mot slurve- eller regnefeil som kan oppstå i elevens oppgaveløsning.

Disse to ulike tilnærmingene til vurdering vil etter vår oppfatning være styrende for hvordan henholdsvis norsk- og matematikklærere tolker og anvender de designede rutineene, og hvordan de har utviklet ikke-formaliserte dagligrutiner i arbeidet som støtte eller supplement til de påleggene de har fått. På mest grunnleggende vis kan vi se dette i lærenes omtale av vurderingsforskriften, hvor matematikklærerne og de aller fleste av norsklærerne deler seg i to blokker når de snakker om muligheten av å gi en sluttvurdering i de respektive fagene. Mens matematikklærerne mener at en sluttvurdering er mulig fordi faget matematikk på Vg3 studiespesialiserende utdanningsprogram er ettårig, viser norsklærerne til at norskfaget på videregående avsluttes først etter tre år og at en sluttvurdering ikke kan dekke alle fagets kompetansemål. Norsklærerne ser følgelig behovet for å «lukke» kompetansemål i løpet av Vg3, mens matematikklærerne i prinsippet ser det som mulig å kunne holde samtlige kompetansemål i faget «åpne» frem til og med siste vurderingssituasjon i skoleåret.

Med henvisning til det teoretiske rammeverket som vi har lagt til grunn for denne analysen, kan vi si at lærerne omsetter pålegget om sluttvurdering til praksis med utgangspunkt i fagets og læreplanens særtrekk. Norsklærerne især har utviklet mer eller mindre formaliserte praksiser som gjør at de foretar

vurdering av elevenes sluttkompetanse over et lengre tidsrom og med en generelt bredere innfallsvinkel enn hva matematikklærerne oppgir at de gjør. Forenklet kan vi si at norsklærerne sier de er avhengig av flere vurderingssituasjoner over tid for å kunne sette det de mener er en riktig standpunktarakter, mens matematikklærerne i prinsippet mener at årets siste vurderingssituasjon vil være tilstrekkelig for å sette standpunktarakter. Samtidig, og som vi allerede har vist til, kan det også bli aktuelt for matematikklærere å se på vurderingssituasjoner tidligere i skoleåret, hvis læreren trenger mer informasjon om helheten i elevens utvikling og forståelse for faget.

Med bakgrunn i den aritmetiske og den analytiske tilnærmingen til vurdering kan vi også oppsummere lærernes bruk av vurderingskriterier. Her skiller norsklærerne seg ut, ved at de fremhever en gjennomgående og fortolkende bruk av vurderingskriterier i arbeidet sitt. Også matematikklærerne viser til vurderingskriterier, men disse kriteriene vil i større grad gi seg selv gjennom lærernes forarbeid med å lage prøven, når de bestemmer hvor mange poeng svarene skal gi. Norsklærernes bruk av vurderingskriterier fremstår dermed som mer gjennomgripende og omfattende, men deres ønske om faglig enighet om vurderingskriterier kan sammenlignes med matematikklærernes ønske om faglig enighet om poengberegning. Satt på spissen kan vi si at godt utviklede vurderingskriterier for norsklærerne tilsvarer matematikklærernes poengsystemer, dog med den viktige forskjellen at norsklærernes vurderingskriterier beholder kontakten med et kvalitativt, analytisk utgangspunkt for vurdering og derfor forblir åpne for bearbeiding og fortolkning. Likevel vil norsklærerne tilstrebe at vurderingskriterier skal kunne brukes på et ensartet vis av kolleger som vurderer den samme elevbesvarelsen, på samme måte som matematikklærernes poengberegningssystemer skal gi høy grad av reliabilitet i hver enkelt lærers vurdering av en elevbesvarelse.

Både systemer for poengberegning og vurderingskriterier har opphav i designede rutiner og søker å etablere allmenn, faglig enighet om verdien av et svar på en oppgave, men vårt inntrykk er at norsklærere generelt må legge mer arbeid i å utvikle og alminneliggjøre bruken av vurderingskriteriene. Da kan de formaliserte rutinene få tilsnitt av en «slik gjør vi det her»-måte å tolke, utfylle og supplere vurderingskriterier på. Vi som forskere har imidlertid ingen grunn til å tro at slik utvikling av vurderingskriterier i norsklæreres daglige arbeid forringer kvaliteten av de opprinnelige, designede kriteriene. Vi konstaterer bare at norsklæreres behov for å bruke disse kriteriene resulterer i at de må holdes ved like i den løpende samtalen og samarbeidet i fagfellesskapet, og at dette behovet synes å være mindre blant matematikklærerne.

Aritmetisk kontra analytisk tilnærming til vurdering legger etter vår forståelse også premisser for kollegasamarbeid om vurdering for henholdsvis norsklærere og matematikklærere. Her legger vi igjen til grunn at norsklærerne må gå bredere til verks enn hva matematikklærerne må gjøre for å oppfylle pålegg om å gi elevene sluttvurdering, og at de formaliserte rutinene blir bygd ut med et stort «nett» av praksiser som danner glidende overganger mellom formaliserte vurderingskriterier og vurderingsrutiner i daglig bruk. På denne bakgrunnen fremtrer også behovet for kollegasamarbeid i vurdering med et annet omfang og med større tyngde for norsklærerne enn hva som kan sies å gjelde for matematikklærerne. Vi har antydnet med grunnlag i datamaterialet vårt at kollegasamarbeidet for vurdering mellom matematikklærere kan ha hovedtyngden i utformingen av prøven, for å skape enighet om prøvens kvalitet og kriteriene for poengberegning. Hovedtyngden av kollegasamarbeid blant norsklærere ser derimot ut til å ligge i etterkant av prøven, hvor rettellesarbeidet og det å skrive kommentarer til elever kan kreve lang tid. Her har norsklærerne, på bakgrunn av den analytiske tilnærmingen til hver enkelt elevbesvarelse, et utstrakt behov for å diskutere og sammenligne med kolleger. Vi fant ikke forskjeller mellom private og offentlige skoler i materialet vårt når det gjelder kollegasamarbeid, men vi har antydnet i dette kapittelet at omfanget av og kvaliteten på samarbeidet kan henge sammen med størrelsen på fagseksjonen, kollegiets erfaring og i hvilken grad ledelsesnivået har pålagt lærerne rutiner for slikt samarbeid.

Både private og offentlige skoler praktiserer «ekstraprøver» for elever som ikke møtte til eksamensforberedende prøve (eller enkelte andre vurderingssituasjoner i løpet av året). Skoler praktiserer også en rekke ulike justeringsmuligheter for elever hvor det enten er uklarhet om karakter,

svakt vurderingsgrunnlag eller eleven står i fare for å stryke. Her ser vi ingen klare forskjeller mellom fagene norsk og matematikk i materialet vårt. Dette feltet av vurderingspraksiser er svært sammensatt og har variasjoner både innenfor fag på den enkelte skole og mellom kolleger som har samme fag på samme skole. Materialet vårt er ikke tilstrekkelig til å si om det er klare forskjeller mellom private og offentlige skoler i hvor mange «ekstrasjanser» elever kan få i en vurderingssituasjon, eller om justering av karakterer bare gjelder strykkandidater eller også elever som vipper mellom to karakterer i den høyere delen av skalaen. Det vi vet, er at alt dette forekommer i kombinasjoner på de private så vel som de offentlige skolene i materialet vårt. Vårt utvalg er for lite til å trekke bastante konklusjoner om omfanget av disse prøveformene eller bruken av dem fordelt på de to skoleslagene, men i kapittel 6 nevner vi dette som et av flere aktuelle forskningsspørsmål for eventuelle fremtidige undersøkelser.

På generell basis kan vi likevel slå fast at elever ser ut til å få mange sjanser til å bli vurdert, både i norsk og matematikk. Her ser vi for første gang i materialet at lærerne gir uttrykk for avveininger mellom universell og differensiell vurdering av elever, alt etter hva årsaken til fravær fra prøver består i. Elever som strever med personlige problemer, som har vært mye syke og som derfor har gode grunner for å ha forsømt skolearbeidet, får mer tilrettelegging for ekstraprøver enn elever som taust uteblir fra vurderingssituasjonen. Samtidig ser vi at vurdering av strykkarakter, uansett årsak, tilsynelatende svært ofte fremkaller behov for ekstraprøve med håp om at eleven kan prestere til ståkarakter. På dette grunnlaget mener vi at de mange ekstra vurderingssituasjonene i praksis bidrar til å snevre inn den nederste delen av karakterskalaen. Som forskere oppfatter vi det også slik at det er de dårligste eller de svakeste elevene som blir gitt flest sjanser til vurdering sammenlignet med flinkere eller mer ressurssterke elever, og vi ser at dette i tilfelle peker mot en differensiell vurdering av elever til standpunkt-karakter basert på elevens evner, forutsetninger og/eller tidligere prestasjoner.

Formelle elevsamtaler eller fagsamtaler blir gjennomført på alle skolene vi har besøkt, men det er interessant å se i materialet hvordan denne designede rutinen tones ned til fordel for verdien læreren legger i alle de uformelle samtalene han eller hun har med eleven tilknyttet skolesituasjonen. Her er det ingen forskjeller mellom fagene norsk og matematikk, og de variasjoner som måtte finnes i hvordan lærere praktiserer slike «småsamtaler», ligger på individnivå.

I den grad man kan bruke slike ord i en nøktern forskningsrapport, vil vi si at lærerne omtaler sensorerfaringer og ikke minst sensorskolering med begeistring. Det er ikke tvil om at lærerne vi har intervjuet, setter verdien og utbyttet av slik erfaring høyt, og at de mener de trekker veksler på slike erfaringer både i sitt eget vurderingsarbeid og som input til kollegasamarbeidet om vurdering. Sensorerfaring og sensorskolering er kanskje også den eneste reelle muligheten lærerne har til å kalibrere individuell praksis eller skolers felles praksis i vurdering med kolleger fra andre skoler. Her omtales sensorgjerningen som noe annet eller noe langt mer enn retdebunker, og kan beskrives som en møteplass eller en kanal der man som lærer kan få tilgang til et større tolkningsfellesskap som presumptivt er skikket til å sette standard for designede så vel som daglige vurderingsrutiner innenfor et fag. Ellers ser vi at de fleste skolene i materialet vårt har ulike skoleinterne prosjekter for vurdering. Lærerne rapporterer at bevisstheten om vurdering er høy i de ulike kollegiene, og de setter dette delvis i forbindelse med de skoleinterne prosjektene.

Når lærerne tar oppgaver som sensor, inntreffer også det de kaller for «sensorbriller». Dette betegner et blikk for vurdering som ikke tar hensyn til at man kjenner eleven og vet hva han eller hun vanligvis er i stand til å prestere. Her har det i vårt materiale ingen betydning hvorvidt sensor til vanlig arbeider på en offentlig eller en privat skole, og de tidligere påviste forskjellene mellom norsk- og matematikklærere synes også å være underordnet. Derimot skiller lærerne her mellom rollen som faglærer og sensor. Som faglærer følger de eleven over tid, de er i stand til å finne «den røde tråden» i elevens måte å uttrykke seg på i norsk, eller vet at eleven kan ha utmerket forståelse av matematikk selv om den skriftlige fremstillingen ikke gjenspeiler dette. Dette vil komme eleven til gode i vurderinger. Vi mener å ha holdepunkter i materialet for å si at en faglærer vil se muligheten for å vippe en elev opp, dersom hun er i tvil mellom to karakterer, basert på kjennskap til eleven og inntrykk av elevens tidligere prestasjoner og eventuelle fremgang.

Sensor mangler imidlertid alle disse opplysningene og erfaringene som tegner bildet av eleven som et enkeltindivid, og har ingen forutsetninger eller incentiver for å vippe vedkommende elev opp dersom oppgaven ikke svarer godt til vurderingskriteriene. På denne måten kan vi si at faglærer har et utgangspunkt for å behandle elever differensielt, mens sensor sitter med en anonym besvarelse som han bare kan tilstrebe å behandle likt med de øvrige besvarelsene i bunken, og i samsvar med kriterier diskutert på sensormøte eller i sensorskolering.

Dette gjør eksamen til en designet rutine i ekstrem grad, hvor det riktignok må skilles mellom aritmetisk og analytisk tilnærming til vurdering og følgelig åpning for diskusjon og forhandling med andre sensorer, men hvor differensiell vurdering i praksis skal være umulig. Standpunktvurdering bygger på den andre siden også på en rekke designede rutiner, men jo lenger ned i hierarkiet av slike designede rutiner vi kommer, jo større innslag av variasjon og ikke-formaliserte dagligrutiner finner vi. Her inntreffer et tilnærmet uendelig antall situasjoner som innebærer formell eller uformell kommunikasjon mellom lærer og elev, og som uttrykker krav til faglig forståelse og elevens respons (eller mangel på sådan). Vi har sett hvordan dette tilrettelegger blant annet for en differensiell behandling av elever som står og vipper mellom to karakterer, og i særlig grad elever som ligger helt i bunnen av karakterskalaen og står i fare for å stryke. Sensor, på den annen side, vil ikke vippe elever oppover med mindre besvarelsen i seg selv kan forsvare det. Resultatet er at faglærere kan gi færre 1ere til elever enn man kan forvente av en sensor, mens sensor på sin side kan gi færre 6ere enn man kanskje ser hos faglærere. Her, i gradsforskjeller av stringens i designede rutiner mellom standpunkt og eksamen, ligger derfor muligens også en forklaring på at eksamens- og standpunktkarakter kan avvike. Lærerne uttrykker imidlertid, og med forklaringer som er beslektet med de argumentene vi nettopp har gjengitt, at de synes det er naturlig at eksamens- og standpunktkarakterene avviker noe fra hverandre. Avvikene bør imidlertid ikke være for store når man ser på hele trinn eller kull av elever og over flere år.

Avslutningsvis kan vi si at det arbeides og reflekteres svært mye over karaktersetting både på offentlige og private skoler. Likevel har vi ikke funnet holdepunkter i materialet for å si at rutinene fortolkes eller anvendes forskjellig på skolene, inndelt etter deres eierform. De største forskjellene i hvordan designede rutiner og ikke-formaliserte praksiser utvikles og anvendes kommer derimot til syne mellom de to fagene norsk og matematikk. Likevel er det grunn til å tro, også på grunnlag av materialet i denne rapporten, at det finnes forskjeller mellom enkeltskoler når det gjelder anvendelsen av rutiner for vurdering. Dette gjelder særlig graden av formalisert kollegasamarbeid om vurderingssituasjoner. Slike forskjeller kan bero på flere ting, blant annet skoleledelse, men også egenskaper ved lærerstaben eller stort innslag av sensorerfaring blant lærerne tror vi kan ha innvirkning på kollegasamarbeid. Her har vi ikke funnet holdepunkter i materialet for å beskrive noen forskjeller mellom offentlige og private skoler. Videre antar vi at de sjansene elever kan få for å forbedre standpunktkarakteren sin, også kan påvirke utfallet av denne karakteren. Her har vi heller ikke data som sier at private skoler skiller seg fra offentlige skoler i antall ekstra vurderingssituasjoner de tilbyr elevene sine. Vi ser at det kunne vært satt et sterkere søkelys på nettopp dette i selve datainnsamlingen, og vi ser dette som et mulig forskningsspørsmål for en fremtidig undersøkelse.

5 «Vi er ungdommer, og vi gjør så godt vi kan ...»

Hensikten med dette kapitlet er å formidle elevenes erfaringer med standpunktvurdering i fagene norsk og matematikk. Elevenes erfaringer er ikke noe speilbilde av hva lærerne forteller, deres synspunkter divergerer av og til på interessante måter. Elevene uttaler seg ofte mer direkte og tydeligere enn lærerne, kanskje noen ganger med en viss overdrivelse, som en vil se. Vi ser ikke noen grunn til å tone dette ned eller stille spørsmål ved det deres forståelse. Tvert imot mener vi at verdien av disse intervjuene nettopp ligger i at det er elevers forståelser av karaktervurderingen som kommer til uttrykk.

Vi tar for oss elevenes omtaler av hva som lå til grunn for valget av skole. Deretter refererer vi elevers synspunkter på hvorvidt skolevalget har betydning for standpunktvurderingen, før vi konsentrerer oss om deres forståelse av karakterer og hvilken plass de har i deres skolehverdag. Synspunkter på forholdet mellom standpunktvurdering og eksamensvurdering tematiseres, og vi går deretter inn på elevenes erfaringer fra matematikkfaget og norskfaget når det gjelder hva de mener lærerne vektlegger i vurderingen og hva som ligger til grunn for den karakteren de får. Vi er også innom temaet juks og hva som faller innenfor og utenfor en slik betegnelse, før vi konsentrerer oss om det store spørsmålet om hva elevene forstår med en rettferdig karakter.

5.1 Intervjuenes innhold og noen tolkningsrammer

De viktigste spørsmålene til elevene handlet om hva de oppfatter lærerne vektlegger når de setter standpunktkarakterer samt hvilke krav som ligger til grunn for den karakteren de får.

Som en innledning stilte vi spørsmål om hvilke fag de hadde valgt og hvorfor de hadde valgt den aktuelle skolen. Temaet standpunktkarakterer ble raskt introdusert i samtalen gjennom ulike spørsmål om hva de synes om å få standpunktvurdering i norsk og matematikk og om deres erfaringer med hvordan disse karakterene blir satt, og om likheter og forskjeller mellom standpunktvurdering og eksamensvurdering. Elevene ble også bedt om å beskrive hva som er en rettferdig karakter. Intervjueguiden for samtalen med elevene følger som vedlegg i denne rapporten.

Et inntak til tolkning av elevenes uttalelser er dikotomien synlig versus usynlig pedagogikk, fra Bernsteins begrepsutvikling, formidlet av Riksaasen (1994), som vi var inne på i kapittel 1. Det er ikke vår hensikt her å gå dypt inn i Bernsteins relativt komplekse kritiske teori, men simpelthen anvende hans innsikt i en nokså forenklet form: Det han kaller usynlig pedagogikk er kjennetegnet ved diffuse vurderingskriterier, stort rom for at eleven selv skal oppdage og utforske sammenhenger samt, som en følge av dette, et relativt stort rom for at vurderingen kan komme til å handle om elevens personlighet.

Synlig pedagogikk karakteriseres ved spesifiserte kriterier for overføring av kunnskap og eksplisitte vurderingskriterier og desto mindre rom for at det er elevens personlighet, men heller hennes kompetanse som vurderes.

Andre tolkningsrammer for forståelse av intervjumaterialet, er aritmetisk versus analytisk vurdering som gjennomgått i kapittel 1 og reflektert i lærernes forståelse av vurderingsoppgavene i kapittel 4. Matematikk og norsk kan i utgangspunktet beskrives som kjennetegnet ved henholdsvis aritmetisk vurdering (poeng- og prosentbasert) og analytisk vurdering (basert på profesjonelt skjønn), selv om dette også må nyanseres. Universell versus differensiell vurdering, og standardisert versus forhandlingsbasert vurdering er andre analytiske perspektiver som er gjennomgått i kapittel 1, og som vi mener gir økt forståelse for meningsinnholdet i elevenes uttalelser om erfaringene de har gjort med standpunkt- og eksamensvurdering.

5.2 Valg av skole

Spørsmålet om hvorfor eleven hadde valgt skolen de gikk på, ble besvart mer eller mindre utførlig. For enkelte hadde det ikke vært noen problemstilling i det hele tatt, gitt at skolen var den som lå nærmest hjemmet, eventuelt at kollektivtilbudet i området pekte mot den aktuelle skolen. Andre hadde vurdert og eventuelt også søkt andre skoler. En gutt ønsket seg et nytt miljø som var forskjellig fra det som preget ungdomsskolen og var fornøyd med å ha fått en ny start. To jenter ved en annen skole ga uttrykk for det samme. Som respons til dette snakket elevene om at det koster mer å etablere seg i et nytt miljø, men at det også var lettere å være seg selv og få aksept for det enn det hadde vært på ungdomsskolen, hvor en stadig hadde måttet anstrenge seg for å passe inn. Ved et par andre skoler ble det å fortsette sammen med venner vektlagt som en grunn for skolevalget, og dette kom til uttrykk ved en offentlig skole, men også ved en privat skole hvor det så ut til å være tradisjon i enkelte av elevenes nabolag for å velge den spesifikke skolen.

Skolens omdømme ble også trukket frem i mange av samtalene om valget av skole. I begrunnelsen for valget kunne også skolens fasiliteter eller det faktum at bygningen var ny, bli trukket frem. Ved en offentlig skole la elevene vekt på et stort tilbud av valgmuligheter med hensyn til fag. Dessuten er skolen attraktiv med høye inntakskrav, ble det poengtert. Nettopp det høye faglige nivået og skolens gode lærere ble oppgitt som motivet for skolevalget av en gutt i elevintervjuet ved denne skolen. På en privat skole ble læreres kompetanse kontrastert av en elev som hadde erfaring fra en offentlig skole tidligere. Der hadde lærerne vært presset på tid til alle elevene som ba om hjelp, mens ved den private skolen opplevde eleven at de hadde fått «litt kremen av det som finnes av lærere. Her tar de ting veldig seriøst, og det betyr veldig mye for meg, fordi jeg mener at hver elev fortjener det beste.» Hun fremhevet dessuten små forhold og lave elevtall som en stor fordel, til forskjell fra den offentlige skolen: «De lærte ikke navnet mitt på et halvt år», påstod hun. Ved en annen privat skole ble det fremhevet at mange av elevene anstrengte seg litt ekstra, ikke bare med skolearbeid, men også for å gjøre verden til et bedre sted, og dessuten: «det er nok mennesker som er så interessert i så mye spesielt at det finnes en gjeng for det du vil drive med, som regel.»

Vi kan oppsummere at høye inntakskrav og høyt faglig nivå var fremhevet både ved private og ved offentlige skoler. Et godt sosialt miljø uten klikker eller sosialt hierarki ble også beskrevet som gunstig ved begge de to typene skoler. Likevel var valget av skole ofte mer eksplisitt ved de private enn ved de offentlige, men det å følge strømmen av medelever fra ungdomsskolen og i vennekretsen også ble poengtert ved en av de private skolene. Ved et par av de offentlige skolene syntes valget av skole imidlertid ikke å være noe elevene hadde lurt særlig mye på, gitt geografi og infrastruktur.

5.3 Antakelser og erfaringer med forskjeller i belønning

Ingen av elevene ga uttrykk for at deres skolevalg var basert på noen orientering om hvor det ville bli lett å få gode standpunktkarakterer. På spørsmål til elever ved en offentlig skole om de noen gang hadde vurdert å søke en privat skole, fortalte en gutt om venners erfaringer som hadde valgt et privat

alternativ og som hadde fortalt ham at «de så vidt jobba og fikk ganske bra karakterer». Dette syntes gutten var «litt urettferdig», og at det var derfor han hadde stilt til intervju. Han utdypet dette med et annet eksempel: «jeg hadde en kompis som faktisk lå på sånn lav middels, og når han kom på [navn på annen offentlig videregående skole] så slo han meg i alle fag, mens jeg [tidligere] pleide å ta ham skikkelig.» En annen gutt tilføyde at det var vanskelig å tenke seg at disse elevene plutselig hadde begynt å jobbe: «Det er ikke typer som pleier å jobbe hardt på skolen, og så plutselig har de 5 i snitt».

Antakelsen om at vurderingsregimet varierer mellom skoler kan, som vi ser her, bli underbygget av elevers kjennskap til andre elever som nå går på andre skoler. De sammenligner også andre skolars opptakspoeng og medelevers og egen karakterutvikling. Det er interessant nok erfaringene de har med tidligere medelevers innsats, som løftes frem til støtte for at andre skoler setter høyere karakterer enn deres egen skole. En undertekst i dette er at deres egen innsats ikke belønnes. Når disse elevene også har erfart at karakterene gikk ned etter at de startet i den videregående skolen, brukes dette også til støtte for en mistanke om at det skolevalget de gjorde ikke var det lureste: «Så skal man liksom bli straffet for at man har vært litt smart». En elev resonnerer om hvordan han kanskje burde ha vært mer taktisk i valget av skole, men tilføyde at han kanskje vil sitte igjen med mer etter å ha gått på den skolen han faktisk valgte. Det aller mest interessante med disse uttalelsene er ikke om det elevene sier stemmer med de faktiske forholdene, men at mistanken finnes om at andre elever ved andre skoler kommer lettere til de gode karakterene. Elevene er seg stadig bevisst at de vil konkurrere mot de andre og heldigere elevene senere og viser til at det er karakterer som teller og ikke hvilken skole en har gått på.

Elever kan dessuten sette spørsmålstegn ved en lærers dyktighet i å vurdere når alle elevene i en klasse får omtrent de samme karakterene i et språk fag samtidig som det er åpenbart for den som betrakter dette, at elevenes ferdigheter er svært forskjellige. Et annet tema var klasseforskjeller og ulempen ved å gå i en klasse med mange meget høyt presterende elever fordi disse ser ut til å sette standarden for hva som er en meget god prestasjon. Om en kan alt som forventes vil en likevel ikke få høyeste uttelling i standpunkt karakter, fordi det finnes andre i klassen som kan enda mye mer. Igjen kommer det en mistanke til uttrykk for at en selv kan ha kommet i en ufordelaktig posisjon. Slike forhold ser ut til å inngå i elevenes beskrivelser av rammer de er underlagt og som bidrar til at de får de karakterene de gjør. På denne måten kan en hevde at elevene begrenser gyldigheten av den bærende ideen om at enhver belønnes for kompetanse og ferdigheter de har tilegnet seg. Begrensningen anskueliggjøres gjennom påvisningen av at det eksisterer ulikhet i vilkår. På denne måten blir også verdien av egen innsats relativ.

5.4 Betydningen av karakterer

Før vi går nærmere inn på elevenes synspunkter på den spesifikke vurderingen i fagene matematikk og norsk, vil vi referere fra elevenes svar på spørsmålet om hva de synes om å få standpunkt karakterer i disse og andre fag. Mer enn én gang vakte dette spørsmålet en smule forundring, det ble gjerne stille, før en eller annen mumlet: Skulle vi ikke det?

Det kan synes som det å få sine prestasjoner og kompetanse vurdert er noe elevene ser som nesten ensbetydende med det å gå på skole. Gjennom siste del av ungdomsskolen og videregående kan det se ut til at de nærmest er blitt to sider av samme sak. På spørsmål om de er opptatt av karakterer, svarte elevene alltid bekreftende, og på flere av skolene ble det også påpekt at på den aktuelle skolen er elevene særlig konkurranseorienterte. Elevene fortalte om at de sammenligner resultatene seg imellom, det går sport i å få bedre resultater enn en annen som en har gått i klasse med lenge, og særlig stas er det å vinne over de elevene som regnes som særlig gode. «Det er litt morsomt, og noe vi tar med et smil», forklarte en gutt. I et annet intervju reflekterte en elev over hvordan karakteren dominerer:

Men altså det er på en måte tallet som betyr noe, ikke hva du har lært. For du vet at til slutt så er det gjennomsnittet av de karakterene du har fått som betyr noen ting. Selvfølgelig betyr

kunnskap noen ting, og det avspeiler jo karakteren, men til syvende og sist så tenker du at karakteren er det viktigste, det tallet du får på arket, som egentlig ikke burde ha vært relevant.

Dette kan tolkes som en vektlegging av indre motivasjon og orientering om å lære fremfor ytre motivasjon orientert mot belønning, men eleven erkjenner at karakteren fremstår som viktigere og av større varig betydning enn kunnskapen en har tilegnet seg. Andre elever hadde synspunkter på nyanseforskjellene i karakterene, eller det en elev kalte irriterende grenser mellom karakterene, før hun fortsatte:

For 4/5 er jo på slutten av året en 4er. Like mye som en 4/3. Ikke sant? Sånn at det er jo vanvittig kjipt der du havner på 4/5 og ikke 5/4, for en 5/4 er på en måte – feil å si like mye verdt, men det blir jo på en måte det, da. For det er en 5er og ikke en 4er. Så den er like mye verdt som en 5+ på slutten av året. For det blir jo også en 5er.

Som regel svarte elevene benektende på spørsmålet om de noen gang hadde klaget på karakterer, og flere fremhevet at de da kan risikere å gå ned. I ett av intervjuene ga to av guttene uttrykk for at de angret på at de ikke hadde klaget på en karakter i Vg1. På den tiden tenkte de at det ikke spilte så stor rolle, og at det ville jevne seg ut, mens de etter hvert og særlig i Vg3 var blitt klar over hvor mye hver karakter teller.

En diskusjon dreide seg om hvorvidt det er viktigere å få god karakter i noen fag fremfor andre. En jente var kommet til at alle de mindre fagene hadde stor betydning for det samlede sluttresultatet, og hennes strategi, fortalte hun, gikk ut på å legge ned det lille ekstra som trengtes for å få 5 og 6 i lesefagene, for å veie opp for karakterene i realfag. I et annet intervju hevdet en jente at selv om hun likte kroppsøving, ville hun i praksis nedprioritere kroppsøving til fordel for religion eller historie når hun måtte velge hvilke hun skulle lese til. De andre fagene ble viktigere. «For», sa hun: «selv om du ikke er så god i ishockey, så kan det hende du er god til andre ting, ikke sant. Jeg syns det blir rart med standpunktskarakter i gym.» En gutt sluttet seg til dette:

Jeg er veldig enig. For jeg mener at hvis du har et fag, så skal du kunne oppnå toppkarakter uansett hvilket utgangspunkt du er i. Og det er faktisk ikke mulig i gym. Hvis du for eksempel er litt overvektig eller veldig tynn, så er det vanskelig å få toppkarakter i styrke eller i utholdenhet. Så jeg synes det stilles for høye krav, i hvert fall i gym. Da burde man tatt hensyn til den enkelte. Sånn at alle kunne oppnådd toppkarakter som i andre fag.

En annen gutt i gruppeintervjuet erklærte seg uenig. Han mente at det var for lett å oppnå gode karakterer i kroppsøving: «Jeg føler at det bare er å gjøre en innsats, så får du god karakter.» Selv hadde han drevet med idrett helt siden han var liten og dette burde han belønnes for, hevdet han, før han gjentok: «det er litt for lett å få god karakter i gym.»

Bare en elev i hele utvalget stilte seg kritisk til karakterregimet som sådan, og dette var etter at spørsmålet var blitt stilt på en litt annet måte: Hva synes dere om det å få karakterer? Hun sa:

Jeg syns den tanken om at alle får inn litt sånn samme generelle kompetanse, allmennkompetanse om norsk, om historie, om religion og alt sammen, det er med på å utvikle deg som menneske og sånn og fungere i samfunnet, men jeg syns det er for mye fokus på selve karakteren: At den karakteren skal ta deg med (sic) videre uansett. For alle kommer inn med forskjellige utgangspunkt, og har samme vurderingskriterier. Og det er som å be en fisk og en apekatt klatre opp i et tre og si at den som, ja, begge to må gjøre det for å stå på en eksamen. Jeg synes ikke det er helt, helt kult, men jeg skjønner at det er enklere å vurdere på den måten.

Denne jenta vendte også senere tilbake til sitt standpunkt om at hun var imot karakterer. Andre deltakere i samme gruppeintervju innvendte at tanken på karakterene får en til å strekke seg og arbeide for å nå et bestemt mål, men jakten på desimalene for å komme inn på for eksempel

medisinstudiet, ble omtalt som snever sammenlignet med et mer ideelt regime med intervju av søkere, hvor også engasjement for studiet burde vært vurdert.

De siterte elevene fremstår som høyt presterende elever gjennom de ambisjonene de signaliserer både med hensyn til karakterer og med hensyn til orientering mot medisinstudiet. Når elevene uttaler seg om karakterer og hvilke strategier de legger an for å komme ut med gode resultater, har vi sett at en elev argumenterer for at det bør være slik at en skal kunne øve seg til å bli god på alle felt uansett utgangspunkt. En annen elev vektlegger elevers ulike utgangspunkt langt sterkere og mener en felles vurdering er urimelig. De to guttene som diskuterer vurderingspraksis i kroppsøvfaget er grunnleggende enige i at en skal få belønning for å øve, men også at forutsetningene hos elevene er forskjellige, i alle fall i dette spesifikke faget. Den ene synes å ha resignert ettersom han ikke har utsikter til å få toppkarakter i gym gjennom å anstrenge seg, slik han tydeligvis ellers er vant til. Den andre synes han får for liten uttelling relativt til andre elever, tatt i betraktning at han har øvd, det vil si drevet med idrett, svært lenge.

Selv om denne diskusjonen kan synes på siden av hva fokusgruppeintervjuene i hovedsak handlet om, nemlig standpunkt-karakterer i norsk og matematikk, kan vi ane noen felles dimensjoner som vil bli tydeligere etterhvert. Betydningen av å øve og forventninger om at øving belønnes karaktermessig står i kontrast til spørsmålet om hvilken betydning ens utgangspunkt eller forutsetninger har. De to dimensjonene kommer ikke i konflikt dersom øving alltid gir uttelling og med alle tenkelige utgangspunkt. Det synes imidlertid å oppstå en konflikt eller brist i logikken når elever mener å erfare at øving ikke hjelper; mulighetene for å forbedre seg forsvinner når kjennetegnene ved en god prestasjon er uforståelige eller ugjennomtrengelige.

5.5 Standpunkt-vurdering blir ofte bedre enn eksamen

Blant fordelene med standpunkt-karakterer ble det fremhevet at det forbereder for eksamen, både på den måten at en vet omtrent hvor på karakterskalaen en befinner seg og på den måten at en over tid merker seg hva en lykkes med, og dette bidrar til et hensiktsmessig valg av oppgave, når det er aktuelt. En annen klar fordel med standpunkt-vurderingen er at den gjøres over tid, med mange mindre prøver og innleveringer, til forskjell fra eksamen hvor det bare er det en får til den bestemte dagen som vurderes. Standpunkt-karakteren er således en «riktigere vurdering» enn eksamens-karakteren, resonnerer en elev. Med mindre prøver og innleveringer får en dessuten insentiv til å øve kontinuerlig, det forebygger skippertaksstrategi, resonnerer en annen elev.

I en elevgruppe var diskusjonsdeltakerne innom ordningen med å bli trukket ut til eksamen. En viss oppgitthet kom til uttrykk fra en elev som viste til at hun hadde fått en dårlig karakter å bære på til forskjell fra medelevene som ikke var blitt trukket ut. Under dette ligger en uuttalt antakelse om at eksamen nærmest logisk sett vil være strengere enn en standpunkt-vurdering. Begrensning i hvilke kompetansemål en blir prøvd i til eksamen, dårlig dagsform og generelt større spillerom for tilfeldigheter, er blant årsakene til at eksamens-karakteren i elevenes øyne nærmest logisk sett vil ligge lavere enn standpunkt-karakteren.³ En innvending om at en vil bli bedre mot slutten av skoleåret på grunn av all øvingen og tilegnelse av stadig mer kompetanse, hvilket skulle kunne tilsi at eksamens-karakteren kan være den høyeste, syntes ikke å endre elevenes resonnement. I ett av intervjuene ble det hevdet med styrke av en elev at en lærer vil ha gjort en dårlig jobb hvis standpunkt-karakteren er lavere enn eksamens-karakteren:

Det er jo veldig mye fokus på at det er elever som får veldig bra standpunkt og så dårligere til eksamen. Men jeg tror ofte så er det også problemer andre vei. Læreren vår har jo blant annet skrytt av at alle gikk opp [i karakterer] til eksamen. Da tenker læreren wow, jeg har gjort en god jobb. Men det er feil. Da du har gjort en dårlig jobb for du har gitt eleven feil karakter.

³ Intervjuene med lærere har avdekket andre strukturelle årsaker til at standpunkt-karakteren gjerne ligger høyere enn eksamens-karakteren.

Denne gutten synes å mene at eksamen har avdekket hva en elev faktisk er god for og feilen består i at læreren ikke har identifisert elevens «faktiske» kompetanse. På et individplan og for en elev som er orientert om å skaffe seg et best mulig utgangspunkt for konkurranse om en attraktiv studieplass, vil lærerens standpunktbedømmelse kunne tolkes som en forsømmelse. Andre elever var mer åpne for at en forskjell kan gå begge veier uten at det nødvendigvis er kritikkverdige, men i likhet med lærerne ga elevene uttrykk for at forskjellen ikke bør være veldig stor, kanskje en, men ikke to karakterer.

Det varierer i hvilken grad elevene har oppfattet at det er kompetansen mot slutten av skoleåret som vurderes til standpunkt fremfor prestasjonene underveis i skoleåret. Som vi har sett ovenfor oppfatter elever at det er mange mindre prøver og innleveringer som danner grunnlaget for standpunktbedømmelsen, noe som også bygger opp under oppfatningen om at standpunkt utgjør et «riktigere» mål på elevens kompetanse enn en eksamen vil være. Samtidig vet svært mange av elevene vi snakket med, at prestasjonene mot slutten av året skal telle mest. Enkelte elever ser ut til å være særlig oppmerksomme på en mulig selvmotsigelse i dette. Det gjaldt for eksempel en gutt ved en privat skole, som fortalte at han hadde en liten feide med en lærer med utgangspunkt i hvor mye karakteren til jul skulle bety:

Jeg skjønner ikke hvorfor han har satt den karakteren, jeg får ikke noen forklaring utenom at før jul hadde jeg en dårligere karakter enn det jeg mener jeg vil ha nå. Jeg mener jeg har hatt progresjon i faget, og da bør jeg få 6 til sommeren. Men han læreren mener at siden jeg fikk 5 før jul, så er det grunn nok til å ikke gi meg 6. Og han mener at alle lærerne gjør som ham, men det er ikke slik. Vi får inntrykk av at det er sluttbedømmelsen som teller, og det er veldig irriterende at en lærer sier at det ikke er sånn.

En annen elev var enig i dette og fortalte at hun hadde gjort en mye bedre innsats til tentamen, og likevel: «Læreren sa at hun ikke kunne gi meg en god karakter fordi karakteren min før jul ikke var bra nok.» I et annet intervju, denne gangen ved en offentlig skole, hevdet en gutt at det har vært mange skifter i skolepolitikken, og at det er relativt nytt at kompetansen mot slutten av året er avgjørende. Han mente å observere at de eldre lærerne ikke følger dette, men står fast på at alt teller. Han mente det tar tid før endringer når lærerne og enda lengre tid før de praktiserer endringene slik at endringene når elevene.

Som det kom frem i samtaler med lærerne, er det kjent for lærerne at det er sluttkompetansen som skal måles i standpunktbedømmelsen, men lærerne vet likevel at elevene bør arbeide kontinuerlig fremfor å satse på skippertak. I intervjuene med elevene så vi det ikke som vår oppgave å orientere dem om hvilket regelverk lærerne skal følge, men vi konstaterte likevel at underveis- og sluttbedømmelse er to vurderingsformer som ser ut til å eksistere side ved side for elevene så vel som blant lærerne: elevene vurderes underveis og elevenes sluttkompetanse teller mest. De to logikkene ser ut til å leve side ved side uten at de nødvendigvis kommer i konflikt, inntil elever mener seg berettiget til karaktermessig premiering for at de har hatt progresjon i et fag mot slutten av skoleåret.

Den siste eleven som er sitert, tilkjennegir en iver etter å argumentere seg frem til en bedre karakter med henvisning til at det er sluttkompetansen som skal telle for standpunkt karakteren. Han synes overbevist om at det er han og ikke læreren som har rett. Denne og lignende situasjoner (for eksempel det vi kaller strategisk fravær nedenfor) får oss til å stille spørsmålet: Kan det være slik at elever ved de private skolene er noe sterkere orientert om sine rettigheter enn elevene i de offentlige skolene?

Vippeprøve er et annet moment som kan komme inn ved slutten av skoleåret før standpunkt karakteren settes. Ved en av de private allmenne skolene uttalte en av elevene om vippeprøve: «Stort sett er det opp til læreren, og maser du, så kan du nok få det», en annen elev tilføyde: «Noen lærere vil deg oppriktig vel, vil hjelpe deg til å få en bedre karakter. De kan sette opp en vippeprøve». Lærere ved den samme skolen bekreftet at de gir vippeprøver når de er i tvil, men særlig for elever som ligger an til stryk. Uansett synes de elevene som vi snakket med ved denne skolen, å ha en bevissthet om et visst forhandlingsrom. Vi kan ikke konkludere bastant om forskjeller i

bruk av vippeprøver mellom offentlige og private skoler, men vi kan heller ikke utelukke at vippeprøver brukes i forskjellig utstrekning.

5.6 Forskjeller og likheter mellom fag i vurdering

I flere av intervjuene kom elevene frem til at sammenfallet mellom standpunkt og eksamen bør være større i matematikk enn det kanskje noen gang kan bli i et fag som norsk. En gutt så norskfaget i sammenheng med flere andre fag som har det til felles at store deler av vurderingen dreier seg om refleksjon:

... det vi kaller å drøfte og sånn. Og jeg tror også mye av forskjellen i karakter også mellom eksamen og standpunkt på prøver og personer er pga. det kompetansemålet – om å reflektere, det er veldig subjektivt. Og det er veldig vanskelig å vurdere en persons refleksjon. For eksempel i matte så er det lettere å få en rettferdig karakter enn i historie, hvor du bare må reflektere nesten.

Denne forskjellen mellom fag er gjennomgående i elevintervjuene, i overensstemmelse med hvordan lærerne omtalte vurderingsarbeidet. Norsk assosieres med stikkord som refleksjon og subjektivitet, skjønn og personavhengighet, mens matematikk har vært betegnet som et «fasitfag» i elevintervjuene. Det mest interessante med dette er de kontrastene som fremheves av elevene, men vi skal etter hvert ta for oss de to fagene etter tur. Først kan det være på sin plass å gjengi elevers forståelse av hva som teller for å få en god karakter for begge fagene. Følgende uttalelser kom i et intervju som respons på spørsmålet Hva må til for å få gode karakterer i norsk og matematikk?

Elev1: Bare være i timen og vise en innsats og på en måte vise at man jobber og er villig til å investere tid i faget.

Elev2: Så må du vise en viss kompetanse. Ja, du må jo selvfølgelig ha en viss kompetanse. Du kan ikke få en 5er og en 6er hvis du ikke kan noen ting.

Elev3: Det er bare å nå kompetansemålene som dere har satt. Ganske enkelt. Det er satt visse krav for å nå en høy oppnåelse eller middels oppnåelse eller noe sånt.

En kan forstå det slik at poengene disse elevene fremhever, bygger på hverandre, men det er også tydelig at elevene assosierer nokså forskjellig på spørsmålet. Det er betydelig forskjell i oppmerksomhet fra å vise læreren at en jobber til bevisstheten om høy og middels måloppnåelse. Kanskje kan en også si at det er en perspektivforskyvning fra ytre kontroll – at det læreren tenker om en er avgjørende, til indre kontroll – at ens kompetanse blir målt. Å vise interesse eller innsats er også et tema som ble utførlig drøftet i flere elevintervjuer, ofte i sammenheng med det å komme godt overens med læreren. Dette kommer vi tilbake til.

Når vi i det følgende tar for oss henholdsvis matematikk og norsk, er målet å finne ut hvordan elevene oppfatter hva som vektlegges når standpunkt-karakterer skal settes, og deres oppfatning av hvilke krav som ligger til grunn for de karakterene de får. Som vi skal se, er dette nokså forskjellig mellom de to fagene.

5.6.1 Matematikk – et «fasitsvarfag»

Elevene hadde ulike former for matematikk. Enkelte av de som deltok, hadde ikke lenger matematikk som Vg3-elever, dette gjaldt 11 elever. Ved to av skolene traff vi ingen elever med matematikk inneværende skoleår. 12 elever hadde R2 eller realfagsmatematikk. Ved to av skolene gjaldt dette alle eller så godt som alle i gruppene vi intervjuet. 8 elever hadde S2, spredt på skoler hvor andre i elevintervjuet hadde R2 eller ikke hadde matematikk. Til grunn for elevenes uttalelser om standpunkt-vurdering i matematikk, finnes det med andre ord vesentlig variasjon i hvor krevende den matematikken er som de har erfaring med. Det synes likevel å være betydelig konsistens i elevenes

betraktninger om vurdering i matematikkfaget. Uttalelsene fra ett av intervjuene illustrerer dette, og disse kom som svar på spørsmålet: Hva skal til for å få god karakter i norsk og matematikk:

Elev1: I matte er det veldig lett, da. Du skal løse oppgavene. Riktig.

Elev2: Men det er nok, det går. Det kommer en del an på prøven. Hvis det er en lærer som gir veldig vanskelige prøver, så kan jo det også påvirke karakteren veldig.

E21: Du har gode lærere og dårlige lærere, men prøven er som regel ganske lik. Det som er kjekt i matte, er at du kan ha en forferdelig lærer, så lenge du kan bare sitte med boka og gjøre alle oppgavene, så får du de samme oppgavene på prøven, altså de samme type oppgavene.

Vi ser at det er en liten uenighet om hvilken betydning læreren har. Den første eleven som er sitert, hadde R2, og klarte seg tydeligvis bra i faget, hvilket også kom frem i andre deler av intervjuet. Også i andre intervjuer ble lærerens betydning for hvor godt en gjør det i matematikk diskutert, men også andre elever syntes å vektlegge viktigheten av å jobbe med faget:

Elev2: I matte, så er det jo ganske enkelt, der må du bare jobbe, du må forstå matten, og du må få riktige svar på prøven.

Elev3: Det er noen som ikke klarer å forstå, da, det er det som er så trist. De ser ikke sammenhengen i tallene og sånn.

Elev2: Men det er sånn, det føler jeg, det er sånn det må være. Matte er et sånt liksom fasitsvarfag. Norsk er mye mer diffust, ikke sant, du kan diskutere mye mer.

I et annet intervju viste en elev til at 96 prosent riktige svar, gir karakteren 6 i matematikk, mens det er langt vanskeligere å forstå hvordan prestasjonene vurderes i norsk. Det ble også hevdet at en såkalt «trynefaktor» har mindre spillerom i matematikk enn i norsk. Samtidig er det lettere å juks i matematikk enn det er i norsk, ble det påstått. Både trynefaktor og juks er stikkord vi kommer tilbake til.

I et av intervjuene omtalte elevene kravene i norsk som nærmest ugjennomtrengelige. Dette gjorde det nærliggende å utfordre elevene litt på hvor åpenbart det er hva som skal til for å få en god karakter i matematikk:

Intervjuer: Man kunne jo tenke seg det helt motsatte, at matematikk er sånn at enten har du talent, eller så sliter du uforholdsmessig mye. Men dere snur dette på hodet når dere sier at matematikk er enkelt å forholde seg til, men norsk er et spørsmål om talent?

Elev1: Ja, men i matte er det lettere å jobbe seg opp til god karakter, det er et puggefag. Du kan gjøre en haug med oppgaver, og det er lettere å få veiledning. Så det er vanskeligere med norsk. Men noen forstår matte bedre enn andre, og sånn vil det alltid være.

Elev2: Har man talent i matte, kan man bare gjøre mindre. Trenger ikke så mye mengdetrening. I norsk funker ikke mengdetrening.

Med dette har vi allerede forespeilet noen av elevenes erfaringer med vurdering og krav i norskfaget. At matematikk er enklere å forholde seg til, slik vi ser elevene hevder, betyr ikke at det oppfattes som et fag hvor det nødvendigvis er lett å få gode karakterer. I alle fall uttrykte enkelte elever at det å få en god karakter i matematikk eller fysikk, krever stor arbeidsinnsats. Dette har vi også vært innom i det foregående, i omtalen av hvordan en jente var kommet til at det var taktisk lurt å legge inn ekstra innsats i «lesefagene» som tilsvarer mange standpunktkarakterer enn å legge alt arbeidet i de mest krevende realfagene. Det er med andre ord ikke noe entydig sammenfall mellom hvor godt en gjør det i et fag som matematikk og det å forstå hvordan matematikkprestasjonene vurderes. Prestasjoner og forståelse av vurderingskriteriene fremstår som relativt uavhengige av hverandre.

Sammenholdt med de analytiske kategoriene som er gjennomgått i kapittel 1 i denne rapporten, kan en si at elevenes erfaringer er forenlig med at prestasjonene i matematikk bedømmes på et *standardisert grunnlag* samt at vektingen er *universell* (fremfor differensiell). Elevene kjenner også til poengberegning av prøver med fastsettelse av prosentandel korrekte svar, såkalt aritmetisk vurdering. Dette betyr ikke at innsats eller holdninger er uten betydning i elevenes bevissthet om matematikk, snarere tvert imot, de erkjenner at matematikk kan være arbeidskrevende. Vi har sett at enkelte ser egen innsats som viktigere enn hvilken matematikklærer en har. De elevene som vektlegger innsats, synes likevel å være innforstått med at det er resultatet av arbeidsinnsatsen og mengdetreningen i form av forståelse og mestring som avgjør.

5.6.2 Norsk og kunsten å knekke koden

En elev kom med friske betraktninger om norsk i sammenheng med fenomenet hun kalte karakterinflasjon, og hun mente å vite at norskkarakteren har vært stabil, mens andre karakterer har krøpet oppover i gjennomsnitt:

Det gjør jo på en måte at den norsk-6eren skal henge utrolig høyt, da, og lærerne kvier seg for å gi ut den. Mens i de andre fagene, så skal det liksom bare være at du skal kunne dette her, ikke sant. Mens i norsk så skal den liksom ligge så utrolig mye høyere enn alt annet, da.

At norskkarakteren er avhengig av læreren var mange elever enige om:

Norsk er veldig diffust, det kommer an på hvilke lærere du har hatt tidligere, om du f.eks. fra ungdomsskolen har bra grunnlag der, og så går det på hvilke lærere du hadde på videregående og om de legger vekt på det samme. Og det er jo sånn at hvis du skriver et kåseri, da. Broren min hadde det problemet. Læreren tok ikke humoren hans. Han byttet skole, så får han 6er, og 6er på eksamen og alt mulig, men forrige læreren tok det ikke, så han fikk 4 på alt han leverte inn der. Så der går det jo kjempemye på skjønn.

Det kan nevnes at svært mange historier fra elevene har dreid seg om hvordan vurderingene i norsk endret seg med skifte av lærer, og her trekker de også på erfaringer fra ungdomsskolen. At lærere liker ulike sjangere, og at det lønner seg å imøtekomme deres preferanser, synes å være en udiskutabel sannhet for en del elever. «Hvis du skriver det læreren liker, kan du gå opp en hel karakter», fremholdt en elev. Man kan også spørre om erfaring med norskvurdering varierer etter elevers erfaring med vurdering i matematikkfaget. I intervjuet som det siste lange sitatet er hentet fra, hadde alle elevene R2. De var enige om hvor vanskelig det var å nå målene i norskfaget:

Elev2: *Når man leser igjennom de kravene, så kreves det jo sinnsykt mye i norsk.*

Elev1: *Egentlig mer der enn i matte.*

Elev2: *Ja. Det kreves mye mer der enn det gjør i realfagene, egentlig. På en måte, for det er sånn. For der kan du si at det og det er rett og det er galt. Det er et fasitsvar, men i norsk så er det så, du må skrive som Ibsen, liksom for å få 6. Det er helt sjukt.*

Elev1: *Og han [dvs. Ibsen] skriver bra!*

Elev2: *Han skriver bra! Det er jo søren meg ... Det føles i hvert fall veldig sånn, også når vi sitter og ser igjennom [kompetansemålene]. Vi er ungdommer og vi prøver så godt vi kan, liksom, men det er tydeligvis veldig, veldig mye som skal til. Syns jeg.*

En kan tenke seg at denne karakteristikken av hva som kreves i norsk, kan være særegen for elever som har valgt den mest krevende matematikken, men vanskeligheter med å forstå vurderingskriteriene i norskfaget kom også for dagen i en elevgruppe hvor flertallet hadde S2. Disse pekte på læreren som ganske avgjørende:

I matte er det lett, for der er det en fasit. Men i norsk så er det jo opp til læreren å bestemme hva du fortjener, og det kan oppleves urettferdig. Men det er jo ikke noe eleven kan gjøre med, det. Det er hva læreren synes du er verdt. Hvis man bruker tilbakemeldingen og leverer et nytt arbeid, men får samme resultat, så oppleves det urettferdig og demotiverende.

Uttrykkene denne eleven bruker «hva læreren synes du fortjener» og «hva læreren synes du er verdt» er nokså usedvanlige sammenholdt med resten av intervjumaterialet fra elevsamtalene. Poenget for denne eleven var at hun hadde forsøkt å forholde seg til tilbakemeldingene fra norsklæreren på konkrete besvarelser hun hadde levert om hva som skal til for å få en bedre karakter. Når dette ikke gjorde noen forskjell synes hun å ha konkludert med at det er ens personlighet som bedømmes i et fag som norsk, og der tilsa karakteren at hun stadig falt igjennom vis-a-vis denne læreren.

En annen elev ga uttrykk for at han oftere ble overrasket i norskfaget. Han kunne gå ut fra tentamen med en god følelse, men så oppdage at det ikke gikk så bra som han hadde trodd. Han la til at det også er vanskeligere å forholde seg til vurderingen i norsk enn det er i matematikk. En jente brukte uttrykket: «Det er en kode som må knekkes» og la til at den koden klarte hun ikke å knekke. En lignende erfaring ble beskrevet av en elev ved en annen skole:

Ja, det er litt sånn når du skriver, så enten er du en 5er i norsk, eller så får du det bare ikke til. Og det er ikke noe sånn som du kan forbedre veldig fort, sånn som i matte, OK, regn mer av de regnestykkene. I norsk blir det mer sånn på helhetsforståelse og sånt noe. Og det er noen som bare ikke får det til.

Uttrykket «å være en 5er» i norsk, ble brukt av elever, men også av lærere. Denne taleformen ligger tett opp til uttrykk som «hva du fortjener» og endog «hva du er verdt» på den måten at karakteren kan synes nærmest å utgjøre en del av en persons identitet, som uttrykt av en gutt: «jeg mener jeg er en 5er i nynorsk». Med slike formuleringer synes det som elevene antyder at det personen eller personligheten fremfor kompetansen som settes i sentrum. Vi er klar over at disse formuleringene er sjargong og mer eller mindre bevisste forenklinger, men dette gjør dem ikke mindre interessante for analyse. Med disse taleformene som bakteppe, vil vi undersøke nærmere hva elevene sa om erfaring med vurdering i norskfaget ved de skolene hvor norsklærerne var aller mest eksplisitte på hva de leter etter og bygger på i vurderingen i norsk.

Intervjuer: *Vet dere hva som skal til for å få god karakter i norsk og matematikk, føler dere at dere har forståelse av det?*

Elev2: *I hvert fall i matte, det har jeg veldig oversikt over. Der har du boka hvor det står veldig konkret hva du skal øve på. Alle kompetansemålene, mens i norsk ... I år har jeg fått mer kontroll på hva jeg bør kunne, men tidligere har jeg følt at norsk bare er sånn hvor du får tilfeldig karakter, egentlig. Men i år har det blitt litt bedre, det er kanskje på grunn av lærere, jeg vet ikke.*

Elev1: *Ja det syns jeg også, læreren min er veldig strukturert, og klar på hva vi må gjøre for å få en god karakter, og hun gir gode tilbakemeldinger, så jeg har en veldig god oversikt over hva man må gjøre bedre for å få god karakter i norsk.*

En kunne kanskje hevde at dette er elever som nettopp har «knekt koden». Det er interessant at de begge viser til lærere som avgjørende for den oversikten de nå mener å ha fått. Ved en annen skole har elevene også begynt å knekke koden:

Han lærer oss veldig mye formelle ting som er viktig nå, med kildehenvisning. Og det er så viktig hvis man ikke har med det, da blir man trukket en karakter med en gang. Det er jo sånn det er.

Ved en tredje skole antydte også elevene at de fikk pekepinn om hvordan de skulle arbeide med norskfaget:

Elev1: *I norsk er det flere ting å jobbe med, f.eks. innhold osv. Det er litt mer sånn konkrete ting man kan jobbe mot, da. Men i maten er det bare sånn at du må forstå det, du må kunne det.*

Elev2: *Men jeg ser ikke hva læreren skulle gjort så veldig mye annerledes for å kunne gi bedre vurderinger tilbake.*

Elev1: *Det er vanskelig, selvfølgelig.*

Vi ser at erfaringene er svært ulike mellom elever og elevgrupper med hensyn til beskjeder de får om hva som kreves og hvordan de kan arbeide for å få gode vurderinger i norskfaget. Blant mange av de elevene vi har snakket med, synes det som norsk fortøner seg som et enten eller, man lykkes eller man lykkes ikke, og hvis man ikke lykkes, er det veldig vanskelig å gjøre noe med det. En lignende logikk kom til uttrykk i et lærerintervju, hvor en norsklærer hevdet at det er ikke alle gitt å få 5 i norsk.

For matematikkfaget ser det ut til at det går an å forstå vurderingskriteriene uten å være god i matematikk, slik elevene ser det. I norskfaget forholder dette seg litt annerledes når vi legger uttrykket «å knekke koden» til grunn. Å forstå hva som kreves fremstår i større grad som sammenfallende med å prestere godt, det vil si å nå kompetansemålene.

På andre skoler har elevene gitt uttrykk for at lærere formidler hva som kreves og hva som gir god uttelling i norskvurderingen, og ikke minst tilkjennegir disse elevene at de forstår hva læreren sier. Skillelinjene på dette området går ikke mellom private og offentlige skoler i vårt intervjumateriale. De to skolene hvor uttalelsene fra elevene er mest ulike med hensyn til hvordan de oppfatter mulighetene for forbedring, er begge offentlige skoler.

Det er kanskje litt pussig at norskfaget vurderes relativt forskjellig av elevene, tatt i betraktning at alle elevene har det samme norskfaget, mens omtalene av matematikk er nokså entydige uavhengig av hvilken matematikk elevene holder på med eller tidligere har hatt. Vi har sett at betydningen av egeninnsats er sterkt fremme i bevisstheten hos elever når de vurderer mulighetene for å gjøre det bra i matematikk, mens betydningen av læreren vektlegges sterkere i norskfaget. Dette er konsistent med oppfatningen om at vurderingen i norskfaget er mer subjektiv eller personavhengig slik det formuleres av elevene, mens lærere vektlegger det profesjonelle skjønnet. Flere elever har dessuten omtalt vurderingskriteriene i norskfaget som diffuse. Sammenhold med de analytiske begrepene som er gjennomgått i kapittel 1, kan vi se at elevenes erfaringer er forenlige med at det foregår en analytisk fremfor en aritmetisk vurdering, den er mer åpen for stadige forhandlinger og avhengig av læreres tolkningsfellesskap. Vi kan merke oss at enkelte elever hevder at de ikke får uttelling for anstrengelsene de mener at de gjør for å nå kompetansemålene. Kanskje kan vi si at noen av elevene synes å oppfatte vurderingen som standardisert samtidig som de opplever at standarden ligger urimelig høyt. Elever har også vektlagt at veiledning er lettere i matematikk enn i norskfaget.

Erkjennelsen blant pedagoger av at tolkningsfellesskapet er en kritisk faktor for vurdering ikke minst i norskfaget er ikke av ny dato. Alle anstrengelsene for å styrke dette fellesskapet og nå frem til enhetlige vurderingskriterier, er ikke tydelig reflektert i elevintervjuene, men vi kan også legge merke til at enkelte elever mener at de får god veiledning i hvordan de kan oppøve de ferdighetene og demonstrere den kompetansen som vurderes som høy måloppnåelse.

5.7 Juks – hva er det?

På spørsmål om det er mulig å jukse seg til en god karakter i standpunkt eller til eksamen, hevdet mange elever at de ikke så hvordan det skulle kunne gå til, når internett og telefoner er lagt vekk.

Å låne andres besvarelser er noe som nokså utvetydig faller i kategorien juks. Det forhindrer ikke at elever eksperimenterer med dette. Vi har ikke informasjon om at elever har lånt av andre, men en elev

fortalte at han sendte sin besvarelse til en kamerat, og da han senere leverte den samme besvarelsen som han selv hadde laget, var det han som fikk beskyldningen om plagiat.

I noen av gruppene har elevene vært svært påståelige på at det går an å jukse. Eksempler på dette er bruk av ulovlige hjelpemidler, som for eksempel, et automatisk oversetterprogram til nynorsk, som elevene er sikre på at læreren ikke kan oppdage. Fordømmelse av dette kom også til uttrykk fra elevene i dette og andre gruppeintervjuer. Juksemakere ble identifisert på følgende vis av en gutt:

Trangen til juksing kommer jo av mangel på kunnskap. Du kan ikke spørsmålene og sånn. Så det er ofte, hva skal jeg si, nedre sjiktet som jukser. Som regel, da.

Stigmatiseringen ved å jukse var også fremme for andre elever som viste til at hvis du blir oppdaget kan læreren bli sur på deg, og «da er løpet kjørt». Et ganske annet tema var smisking med læreren, og noe diskusjon oppsto da dette ble omtalt som juks. I et intervju ved en offentlig skole diskuterte elevene med utgangspunkt i det som fremstår som tolkninger av andre deltakers atferd og meninger om denne fra tidligere. På spørsmålet om smisking er juks, falt følgende replikkveksling:

Elev2: *Nei, jeg syns ikke det.*

Elev1: *Nei, du syns ikke det, nei! (Mye latter.)*

Elev2: *Ikke det. Hva er smisking? Hvis man kjøper ting til læreren ...*

Elev1: *Det er ikke smisking. Det er bestikking.*

Elev2: *I alle fall, det er ikke noe greit, men hvis man har en god tone med læreren og ja, hva skal jeg si? Det er lov å være hyggelig og, kanskje.*

Elev3: *Jeg tror det hjelper på karakteren å ha en god tone med læreren.*

Elev2: *Ja det gjør det helt klart.*

Elev1: *Det gjør det absolutt, men det er bare litt forskjellige måter å gjøre det på. Liksom, det er noen som ... Nei, jeg vet ikke hva jeg skal si. Det er vanskelig å forklare. Jeg vet ikke om det å liksom ta læreren ut etter hver eneste vurdering og snakke med dem, om det er smisking. Det er noe som jeg ikke tør å gjøre. Fordi, det er på en måte å vise læreren at du jobber med faget, det er å vise interesse for hva du kan gjøre bedre til neste gang. Og det er jo på en måten en form for smisk. Jeg ser på kommentarene og så prøver jeg å forbedre meg, uten å ...*

Elev2: *Det må være lov å snakke med læreren og spørre om hva du skal gjøre. Skal man ikke gjøre det da?*

Elev3: *Jeg er veldig enig i det dere sa om å engasjere seg faglig og kanskje tenke litt mer strategisk, tenke på de prøvene som faktisk blir vektlagt, at du gjør det bra på de, og at du er flink med lekser og gjør egentlig alt du skal, alt læreren vil. Jeg synes egentlig ikke det kan oppfattes som smisking, men du får en bedre tone med læreren.*

Et undertema i denne diskusjonen synes å være spørsmålet om hvor grensene for rettfærdig vurdering går, og når blir elever forskjellsbehandlet. At det er dumt å bli uvenner med læreren, var det ingen som protesterte på, men en mistanke om at andre lettere får gode karakterer på grunn av vennlighet overfor læreren, synes å utfordre forestillingen om at det er kompetansen som vurderes. En lignende diskusjon løp i et annet elevintervju, da en gutt brøt inn:

Jeg tror at det å gi folk belønning for å ha en god holdning og være snill med andre, også læreren, altså å ta med det inn i arbeidslivet, det tror jeg er viktig. Å gi elever den holdningen at vær hyggelig mot medelever og vær hyggelig mot læreren, og ta det med dere inn i

arbeidslivet, for det får dere belønning for, det tror jeg er veldig viktig. Da synes jeg god holdning og godt engasjement skal gi resultater på standpunktkarakter.

Denne gutten mente altså at vennlighet aktivt burde belønnes med en god standpunktkarakter, og begrunnelsen for dette var at slik atferd er en god trening for arbeidslivet. Begge de refererte diskusjonene fant sted i offentlige skoler. I en av de private skolene kom vi imidlertid på sporet av en annen type atferd som elever kan forsøke for å høyne sjansene for en best mulig standpunktkarakter. Temaet var først brakt på bane i lærerintervjuet, hvor det ble uttrykt mild frustrasjon over taktikken, derfor fikk elevene spørsmålet i det påfølgende elevgruppeintervjuet:

Intervjuer: Kan man manipulere litt for å sikre en god standpunktkarakter? Hvis dere skal ha prøve i noe dere vet at dere ikke behersker så godt, går det an å være syk på prøven?

Elev1: Det er et tema, det skjer ofte.

Elev2: Jeg gjør det aldri.

Elev3: Jeg har gjort det.

Elev1: Jeg også.

Intervjuer: Får dere den karakteren dere da gambler på at dere skal få?

Elev3: Ja, for det er bedre å ikke møte opp med legeerklæring. Så ser læreren karakterene dine, og ser bort fra den prøven. Kun de andre prøvene blir lagt vekt på, og hvis de er bra, så har du flaks.

Elev2: Men hvis du er syk, kan du skaffe deg bedre tid og utsatt prøve. Ja, det er jo en annen prøve, men du får de samme kompetansemålene. Og du får lenger tid. Men du må jo igjennom det likevel.

Intervjuer: Hva slags reaksjoner får dere fra lærere da?

Elev1: De ser jo et mønster på noen, men hvis du gjør det en eller to ganger i løpet av året tror jeg ikke at lærerne ser på det som manipulasjon. Men hvis du gjør det fire fem ganger i løpet av et år, da ville jeg antydde at læreren ser et mønster.

Elev3: Jeg har møtt opp til alle prøver unntatt en, og det var med vilje. Læreren ble veldig sur, og da måtte jeg ta den på nytt.

Ved å måtte «ta prøven på nytt», kan en si at eleven har oppnådd å få utsatt vurderingen. Også ved en annen privat allmenn skole ble dette tematisert, og der fremhevet elevene at lærerne ikke tillot elevene å ta prøven dersom de hadde vært borte, med mindre de hadde legeerklæring. De siterte elevene er tydeligvis innforstått med hvilket grunnlag læreren kan bygge på ved standpunktvurdering og hva som ikke skal legges til grunn, selv om de også erfarer at lærere gjennomskuer strategien, og dette er en risiko. Som fremhevet av andre er «løpet kjørt» hvis læreren blir sur på en. Det er interessant å se elevenes uttalelser i sammenheng med læreres omtaler av hvordan de prøver å være seg bevisst egne sympatier og antipatier med elever. Trynefaktor er et uttrykk som var fremme i samtlige elevintervjuer, og nesten uten unntak var det elever som introduserte det i samtalen. Lærere hadde også formeningene om hva elever tror om dette, og hvor feil de kan ta i mistanken om hvordan elever belønnes.

5.8 Så hva er en rettferdig karakter?

Mange av diskusjonene vi har gjengitt kan sies å dreie seg om spørsmålet: hva er en rettferdig karakter. En elev forsøkte å beskrive kjennetegn ved en rettferdig karakter, og ble etter hvert supplert av medelever:

Elev1: Det er en karakter du har oppnådd, og ikke fått. Det er liksom, jeg tenker vi snakker så sykt mye om hva slags karakter fikk du og hva slags karakter fikk du, men det handler mer om hva slags karakter man har oppnådd. Vi jobber jo med dette her, det er noe vi legger sykt mye arbeid i, og hvis du legger mye arbeid i en ting og oppnår en god karakter, så er jo det en rettferdig karakter. Men hvis du ikke legger noe særlig arbeid i det og, som vi snakket om, smisker deg til ting eller jukser deg til ting, og får en karakter som du ikke kan egentlig stå for, da er det en urettferdig karakter.

Elev2: En rettferdig karakter er noe du har oppnådd uten juks og at det er ordentlig grunnlag for læreren, da, hvorfor han satte den karakteren.

Elev3: Ja, jeg er veldig enig. Jeg synes også at karakteren skal kunne grunngis ut fra kompetansemålene. Jeg synes mye av læring og vurdering går utenfor kompetansemålene, eller de kommer ikke synlig frem. Og det er jo det vi skal lære, kompetansemålene, det er ikke noe annet. Så jeg syns mer av vurderingen og læringen burde vært basert på kompetansemål, egentlig.

At det er ordentlig grunnlag, slik den andre eleven er inne på, kan bety at eleven ikke har unndratt seg prøver gjennom strategisk fravær, som tematisert ved en av skolene. Det synes ellers å være mange betingelser som skal være oppfylt for at en skal anse en karakter som rettferdig, kan vi se av elevenes uttalelser. En elev stilte spørsmål ved om karakterer noen gang kan bli rettferdige, gitt at én elev ikke behøver å lese mer enn en gang mens en annen elev «pugger hele uka» og likevel får dårligere karakter. Med dette trekkes elevens forutsetninger inn i forståelsen av hva rettferdighet er og om det er mulig. Denne eleven konkluderte fort med at det må være «om man har fått det med seg» som testes, uavhengig av om innsatsen har vært stor eller liten.

Anonym retting ble fremholdt som et tiltak mot trynefaktoren ved svært mange av skolene. At lærere bytter besvarelser og retter for hverandre, ble også fremholdt som noe som kan dempe trynefaktoren. Utover dette mente en elev at det må være en viss avstand mellom lærer og elev for at læreren skal klare å være objektiv i karaktersettingen. Ved en skole poengterte en gutt:

Elev1: Jeg synes ikke at det at noen er flinkere enn andre til å argumentere for seg, skal gi grunnlag for en bedre karakter.

Elev2: Det kommer jo an på faget.

Elev3: Jo, men noen kan være sjenerte og ikke tørre å si det man mener. Og da blir jo ikke karakteren rettferdig.

Denne samtalen fra en privat skole, minner om en annen elevsamtale, gjengitt ovenfor, som foregikk ved en offentlig skole, hvor elevene diskuterte hvorvidt smisk er å forstå som juks eller ikke. Sjenanse kan oppfattes som et personlighetstrekk som vil gi dårligere betingelser for en god karakter. Som vi har sett, mente en elev at vennlighet, også overfor læreren, aktivt burde belønnes i standpunkt-karakteren, fordi det er noe en bør trenes i for arbeidslivet. En annen gutt som deltok i et annet intervju, ved en privat skole, uttalte at kampen mot trynefaktoren er meningsløs:

Men det er jo sånn i arbeidslivet også, på en måte, så det kan jo ikke, å bruke så mye energi på å få bort den trynefaktoren. Det er jo kanskje noen steder det er altfor ekstremt. Men det er på en måte, hvorfor skal du slåss mot det umulige – det uunngåelige? [...] Smisk slash muntlig aktivitet, er det innenfor det faglige eller er det utenfor? Det at du kan gå til læreren i

friminuttet, for eksempel, og fortelle om – ja. Ting, da. Jeg sa det til og med rett før prøven i dag, jeg fortalte en historie til læreren. Det viser at du er faglig engasjert og har entusiasme og er villig til å lære ting utenfor skolen, ikke sant. Så de kompetansemålene og det tullet der, det er de du viser med smisk, da.

Denne eleven synes å ville demonstrere en viss kynisme i bruken av egne fortrinn til beste for seg selv og for å oppnå gode resultater. Det er vanskelig å finne uttalelser som ligner hos andre elever. Samtidig som mange tilkjennegir at de vurderer ulike strategier for å oppnå størst mulig konkurransedyktighet ved avslutningen av tretten års skolegang, har de fleste elevene også lagt for dagen et dypt engasjement i spørsmålet om rettferdighet. Dette kan kalles en bærende ide ved standpunktvurderingen – at den skal være rettferdig. Når elever mistenker at den ikke er det, eller når kravene som stilles er uforståelige, er det nærliggende for dem å resignere – å forbedre norsk karakteren fremstår som veldig vanskelig. Det er ingen opplagte forskjeller mellom uttalelsene fra elevene ved de private skolene sammenlignet med elevene ved de offentlige skolene. En kan kanskje fornemme at elevene ved de private skolene muligens er noe mer orienterte om sine rettigheter enn elevene ved de offentlige skolene, men det er vanskelig å trekke noen annen konklusjon enn at de alle har vist betydelig engasjement i spørsmålet om rettferdighet.

Ingen av elevene har på noe tidspunkt tematisert fordeler ved å ha foreldre med høyere utdanning eller at det kunne være forskjeller i vilkår for å oppnå gode resultater avhengig av foreldres utdanning og yrker, slik utdanningssosiologien så klart viser. Temaet ble heller ikke brakt på bane av intervjuerne. Elevene insisterer ofte på at gode karakterer er og bør være et resultat av egne anstrengelser, og noen legger for dagen en skepsis til at gode manerer, vennlighet eller smiger skal belønnes. Andre ser ut til å mene at det ikke er illegitimt, mens et par av elevene fremholder at det aktivt bør belønnes.

Det synes å være stor enighet om at i matematikk gjelder entydige og rettferdige vurderingskriterier, mens i norsk er dette langt vanskeligere. Taleformer som «hva du er verdt» eller «å være en 5-er» peker mot identifisering av personen med prestasjonen, og en mistanke om at det er personligheten som vurderes. Det synes å være et stort uutnyttet potensiale i klargjøring for elevene av hvilke krav som stilles i norskfaget og hva som er henholdsvis god og mindre og god prestasjon. Vi har sett at enkelte elever mener de får vite dette av strukturerte og dyktige lærere. En kan spørre om det er slik at når elevene forstår det, er det fordi de også behersker de aktuelle kravene. Det vil imidlertid kunne oppstå en tillitskrise hos de elevene som mener å ha erfaring for at det er personligheten deres som vurderes og at det er derfor de stadig faller igjennom.

5.9 Oppsummering og drøfting

Vi har gjennom dette kapitlet sett nærmere på elevers forståelse av standpunktvurderingen og på hvilket grunnlag den settes. Assosiasjonene som elevene har gitt uttrykk for, spenner fra å demonstrere engasjement og innsats for at læreren skal få et godt inntrykk av en, til orientering om kompetansemål, og vi har pekt på at det her synes å ligge en forskyvning fra ytre kontroll til indre kontroll i strategiene elevene legger for dagen når de vil oppnå gode karakterer.

Alle elever er enige om at det er veldig uheldig om man blir uvenner eller kommer dårlig overens med læreren. Vi har også sett at elevene har nokså ulike synspunkter på hvorvidt anstrengelser for å gjøre seg til venner med læreren, forstått som smisk, er å betrakte som juks eller ikke. Vi har pekt på at til grunn for disse diskusjonene synes det å ligge et dypt engasjement hos elevene i at vurderingsprosessen må kunne oppfattes som rettferdig. Sammenholdt med læreres omtale av prosessen ser vi noen interessante forskjeller. I flere lærerintervjuer har vi hørt at tvilen kommer eleven til gode, og vi har også hørt at dersom en lærer har vanskelig for å like en elev, vil læreren legge ekstra vekt på ikke å forskjellsbehandle, og som resultat kan vedkommende elev kanskje komme til å få en mer velvillig vurdering enn en annen elev ville ha fått. Lærere har også fremhevet at de ser på innsats eleven har gjort over tid, og at de belønner slik innsats. Disse uttalelsene fra lærerne

kan forstås som et etos i deres yrkesutøvelse, vurderingen skal være rettferdig og elevene skal ha de samme mulighetene. Når elever har så sterk tro på anonym retting av skriftlige prøver mot slutten av skoleåret, kan det være at de overser den velviljen lærere legger inn i vurderingen av dem.

Fra lærerintervjuene vet vi mye om deres forståelse for hvorfor standpunktkarakteren tenderer til å ligge høyere enn eksamenskarakteren. De fleste elevene vi har snakket med, synes å forstå at forskjellen kan gå begge veier, men vi har også hørt en elev mene at dersom standpunktkarakteren ligger under eksamenskarakteren, da har læreren satt «feil karakter». Beslektet med dette er ambivalensen i læreres forståelse av om innsatsen og eventuelt progresjonen mot slutten av skoleåret skal tas med i vurderingen, eller om det skal være kompetansen mot slutten av året som står i fokus. Mange av elevene vi har snakket med, synes å leve med og forholde seg til begge de to vurderingsprinsippene. De to logikkene ser ut til å leve side ved side uten at de nødvendigvis kommer i konflikt, inntil elever mener seg berettiget til karaktømmessig premiering for at de har hatt progresjon i et fag mot slutten av skoleåret og læreren etter sigende legger prestasjonene gjennom året til grunn.

I likhet med lærere understreker elever forskjellene mellom fagene norsk og matematikk når de omtaler vurderingskriterier og hvorvidt elevene får vite hvordan de kan gjøre det bedre. Litt overraskende for oss er det at en del elever omtaler gode resultater i matematikk som et spørsmål om mengdetrening mens mengdetrening oppfattes som umulig i norsk. Diffust og personavhengig er betegnelser flere av elevene bruker om vurdering i norskfaget. Et unntak er de elevene vi intervjuet ved en offentlig skole som har drevet Vurdering for læring i flere år. Ved denne skolen var norsklærerne meget eksplisitte på hva de ser etter i vurderingen av elevenes prestasjoner. Samtidig ga elevene ved denne skolen uttrykk for at de hadde en ganske god forståelse for hva som kreves for å få en god karakter i norsk. Ved en annen skole var elevene sterkt overbevist om at det er ens personlighet det kommer an på, og at forbedring av prestasjoner ikke er mulig. Begge disse skolene er offentlige skoler.

Det er ingen av elevene vi har snakket med, som har gitt uttrykk for at de valgte skole ut fra antakelser om hvor det vil være lett å få gode karakterer. Samtidig har flere av elevene tilkjennegitt at de følger med på karakterutviklingen hos tidligere medelever som nå går på andre skoler. Med en anelse om at det kan være lettere å få gode karakterer på andre skoler enn den en selv går på, peker elevene både på offentlige og private skoler.

Sammenholdt med de analytiske kategoriene som er gjennomgått i kapittel 1 i denne rapporten, kan en si at elevenes erfaringer er forenlig med at prestasjonene i matematikk bedømmes på et *standardisert grunnlag* samt at vektningen er *universell* (fremfor differensiell). Elevers erfaringer med norskfaget er ganske forenlige med at vurderingen arter seg som stadige forhandlinger og er avhengig av et tolkningsfellesskap. Vi har gjennom intervjuene med norsklærere sett hvor viktig og krevende deres bestrebelser på å nå et tolkningsfellesskap er for dem, samtidig som nettopp dette tolkningsfellesskapet er ganske avgjørende når de ønsker å foreta rettferdig standpunktvurdering.

I lys av dette er en del av elevenes påstander om at vurderingen er personavhengig og subjektiv bemerkelsesverdig. Enkelte uttalelser om hva som vurderes, som «hva du er verdt», har vi tolket dit hen at de mistenker at det er deres identitet eller personlighet som er gjenstand for vurdering. Erfaring for at anstrengelser (eller mengdetrening) ikke nytter, betegner elever selv som demotiverende. Vi tror ikke lærere har intensjoner om å mystifisere vurderingskriterier for elever, tvert imot, lærere har også vært inne på at de må forenkle og konkretisere kompetansemål for at elevene skal forstå dem. Det som er bekymringsfullt er at det finnes elever som tror at de er forhåndsvurdert som mindre gode. Dette betyr at lærere ikke har lyktes i å formidle vurderingskriteriene til elevene slik at de forstår hvordan vurderingen foregår og hva som ligger til grunn for den karakteren de får. Selv om det ikke er intendert fra lærerens side, fremstår pedagogikken som usynlig for elevene, og usynlig pedagogikk kjennetegnes først og fremst ved diffuse vurderingskriterier (Bernstein i Riksaasen 1994). Et helt avgjørende poeng fra elevintervjuene er likevel at formidling av hvilke vurderingskriterier som gjelder, viser seg å ha lyktes godt ved en av skolene. Det betyr at diffuse eller mystiske vurderingskriterier ikke er noen nødvendighet i norskfaget.

På akse standardisert versus forhandlingsorientert vurdering, kan norskfaget klart plasseres nærmest sistnevnte. Blant elever som strever, synes det likevel å eksistere en anelse om at det eksisterer en standard, men denne er urimelig høy. En må skrive som Ibsen, er en slutning som trekkes. Andre uttrykker at det finnes en kode de ikke klarer å knekke. Vi mener enighet og eksplisittgjøring av vurderingskriterier som lærerne bruker i norskfaget ville kunne redusere slike misforståelser blant elevene.

Vi har sett flere eksempler på at elever diskuterer hva som er en rettferdig karakter også med utgangspunkt i spørsmål om like vurderingskriterier og i hvilken grad elevers ulike forutsetninger bør trekkes inn. Om det er for vanskelig eller for lett å få god karakter i kroppsøving, om en fisk og en apekatt kan måles på samme oppgave og forskjeller mellom elever i hvor mye tid de må bruke for å lære det samme, er alle eksempler på dette. Vi kan si at elevene drøfter universell versus differensiell vurdering uten at de bruker disse begrepene.

Elevene ved de private og offentlige skolene synes å ha en ganske samstemmig oppfatning av hvordan vurderingen foregår, og vi finner ingen klare holdepunkter for at dette varierer avhengig av eierforhold. Vi har likevel antydning med grunnlag i vårt intervjumateriale fra åtte skoler at det kan synes som elevene ved de private skolene er mer orienterte om sine rettigheter enn elevene ved de offentlige skolene er. Det vi kan kalle strategisk fravær, at en elev velger å utebli fra en prøve som skal måle kompetanse på et felt hvor eleven føler seg usikker, har vært et fremtredende tema på en av de private skolene. Det hadde vært ønskelig at vi ble oppmerksomme på dette fenomenet tidligere i intervjuprosessen for å kunne utforske det ved de øvrige skolene. Uansett er det ikke tvil om at elever er orientert om ulike strategier for å oppnå et best mulig vitnemål. Noen tilkjenner at deres strategi er hardt arbeid, målrettet innsats og at de tar til seg signaler fra lærerne og prøver å lære av det, mens andre i større grad ser muligheter i å samtale og eventuelt også forhandle med lærere om sluttresultatet.

6 Avslutning

I dette kapittelet oppsummerer vi hovedfunnene fra rapporten og drøfter de tre problemstillingene som har ligget til grunn for arbeidet vårt. Mot avslutningen av kapittelet løfter vi frem noen nye problemstillinger som kommer til syne i rapporten, og vi søker å belyse og klargjøre disse problemstillingene ved å reise noen nye spørsmål som kan egne seg til videre forskning.

6.1 Rapportens hovedfunn

De statistiske analysene av avvik mellom standpunkt- og eksamenskarakter viser at de private allmenne skolene har et noe større avvik enn hva som er tilfelle for offentlige skoler, private kristne skoler og private spesialiserte skoler, også når vi kontrollerer for andre variabler. Samtidig viser regresjonsanalyser at det også er andre variabler som påvirker avviket, blant annet kjønn, type program, om skolen har høye opptakskrav og om det er en liten eller stor skole. Små skoler har generelt noe større avvik enn større skoler.

Det som har størst påvirkning på hvor stort avviket mellom de to karakterene er, er elevens standpunktkarakter. Elever med gode karakterer har med andre ord større risiko for å gå tilbake til eksamen enn elever med svakere karakterer. For en elev med 6 i standpunktkarakter er risikoen for å gå tilbake 85 prosent og risikoen for å gå tilbake til eksamen er større enn sjansen for å gå frem for alle elever med karakter 3 eller bedre. Dette fenomenet kalles regresjon mot gjennomsnittet.

Vi ser også at det er forskjeller i avvik mellom de fire årene vi ser på (2010–2013), ved at avviket er mindre i 2013, og dette henger sammen med at gjennomsnittlig eksamenskarakter er høyere det året. Vi kan ikke si om dette er en generell trend siden vi ikke har tilgang til informasjon om mer enn de fire årene, men det er likevel interessant å merke seg at det er eksamenskarakterene som varierer mest over tid. Dermed kan det diskuteres om eksamenskarakterer faktisk har den kalibrerende funksjonen som de antas å ha. Det forutsetter i så fall at karakterene ikke varierer mye fra år til år. Samtidig kan det være noe ved eksamen det enkelte året som forklarer avviket, men det har vi ikke informasjon om her. Siden eksamen forutsettes å måle det samme fra år til år, bør ikke vanskelighetsgraden variere for mye. Det kan den heller ikke gjøre dersom den skal sies å ha en kalibrerende funksjon. Diskusjonen om den kalibrerende funksjonen ved eksamen tas opp både i lærerintervjuene og i elevintervjuene.

De kvalitative analysene viser at lærerne legger et mangfold av rutiner til grunn når de arbeider med vurdering og karaktersetting. Rutinene kan i mange tilfeller betegnes som *designede rutiner*, noe som betyr at de ofte er nedskrevet eller i det minste formelt kunngjort, og som viser til standardiserte handlingsrepertoar for lærernes arbeid med vurdering. Av slike designede rutiner har vi i denne rapporten tatt for oss læreres forståelse og bruk av vurderingsforskriften, av eksamensforberedende

prøver (heldagsprøver) inkludert eksamensveiledningene, erfaringer med de lovpålagte elevsamtalene for vurdering og erfaringer med sensorarbeid samt skoleinterne prosjekter som skal fremme god vurderingspraksis. I kjølvannet av slike designede rutiner har vi imidlertid sett at det oppstår et mangfold av rutiner i daglig bruk. Dette kan forstås som praksiser som enkeltlærere eller grupper av kolleger utvikler for å utfylle eller supplere de designede rutinene, og som dekker et behov som lærerne har erfart i det løpende arbeidet med vurdering. Dagligrutinene eller praksisene er ikke nedskrevet, og kanskje heller ikke uttalt som regler eller normer for arbeidet, men de kan like fullt eksistere og være gjenkjennbare som noe som gjentar seg og som former lærernes handlinger over tid. I kapittel 4 har vi identifisert og beskrevet et vell av slike dagligrutiner som eksisterer i forlengelse av eller parallelt med de designede rutinene. Viktige dagligrutiner er kollegasamarbeid for retting av besvarelser og tolkning av vurderingskriterier, uformelle samtaler med elever og praksiser for å vurdere helhet i en elevs prestasjon eller skoleresultater hvis eleven gjør det uventet dårlig på heldagsprøven.

Av gjennomgangen av rutiner og praksiser i kapittel 4 er det klart at lærerne på alle de åtte skolene vi har besøkt, forholder seg noen lunde ens til de designede rutinene. Tolkningene deres av disse rutinene er stort sett like på tvers av skoler, skjønt vi ser variasjoner i hvordan begrepet «sluttvurdering» tolkes og anvendes respektivt i fagene norsk og matematikk. Lærernes fortellinger om praksiser i daglig bruk danner også et gjenkjennelig mønster på tvers av skolene. Alle lærerne anvender et mangfold av slike praksiser, og i noen grad er disse praksisene institusjonalisert på seksjons- eller faggruppenivå. Da blir det et åpent spørsmål hvorvidt de glir over til å bli designede rutiner, initiert og styrt av skoleleder eller seksjonsleder. I hvilken grad disse praksisene skal kalles for designede eller daglige rutiner, kan i tilfelle avgjøres ved å se på hvordan rutinene har blitt innført. Hvis de har oppstått «nedenfra» som et svar på behov som lærerne erfarer i det løpende arbeidet med vurdering, kan vi fremdeles kalle dem for institusjonaliserte praksiser. Hvis de har blitt innført ved en bestemmelse ovenfra som er ment å gripe aktivt inn i lærernes vurderingspraksis, kan vi kalle dem for designede rutiner. Poenget er imidlertid at lærerne behersker alle disse rutinene og kombinasjonene av dem, på måter som kan variere litt mellom faggrupper og skoler. Imidlertid har vi ikke klart å se at disse forskjellene følger skillet mellom offentlige og private skoler, siden det ikke er klare forskjeller mellom skoleslagene i hvordan lærerne jobber. Det er dermed ikke noe i selve intervju materialet som tilsier at det statistiske funnet at noen private skoler har et litt høyere avvik mellom eksamen og standpunkt karakter, skyldes ulik praksis.

De klareste forskjellene i materialet vårt med hensyn til rutiner og praksiser inntreffer mellom faggruppene matematikklærere og norsklærere. Hovedtendensen i materialet er at matematikklærerne følger en aritmetisk tilnærming til vurdering, som innebærer at de gir poeng for riktig svar og lar karakteren fremkomme på basis av en prosentsats av totalt antall oppnåelige poeng for prøven. Norsklærerne forholder seg derimot analytisk til vurderingen av elevbesvarelser. Dette innebærer en kvalitativ bedømming av elevens måloppnåelse med grunnlag i kompetansemålene for faget. Samtidig bruker også matematikklærerne en analytisk vurderingsmåte når de skal se på helheten i en elevbesvarelse og vurdere elevens *forståelse* av faget.

Disse to tilnærmingene til vurdering vil etter vår oppfatning være styrende for hvordan henholdsvis norsklærere og matematikklærere tolker og anvender de designede rutinene, og hvordan de har utviklet praksiser som støtte eller supplement til påleggene de har fått. På mest grunnleggende vis kan vi se dette i lærernes omtale av vurderingsforskriften, hvor alle matematikklærerne og de aller fleste norsklærerne deler seg i to blokker når de snakker om muligheten for å gi en sluttvurdering i de respektive fagene. Norsklærerne har utviklet mer eller mindre formaliserte praksiser som gjør at de foretar vurdering av elevenes sluttkompetanse over et lenger tidsrom og med en generelt bredere innfallsvinkel til vurderingssituasjoner enn hva matematikklærerne gjør.

Når det gjelder bruk av vurderingskriterier, fremhever norsklærerne dette som en gjennomgående og fortolkende praksis hvor designede og daglige rutiner til en viss grad flyter over i hverandre. Matematikklærerne har et mindre behov for å tolke vurderingskriteriene i faget sitt. Dette påvirker også

kollegasamarbeidet som kjennetegner de to faggruppene av lærere. Mens norsklærere har et løpende behov for å samtale og samarbeide om fortolkning og anvendelse av kriteriene, avgrenses matematikklærernes behov for kollegasamarbeid til forberedelse av prøven og i noen grad til tvilssituasjoner som oppstår under rettingen. Lærere fra begge faggrupper omtaler det å ha erfaring fra sensorarbeid som svært viktig for å justere egen og kollegers vurderingspraksis.

Et viktig, men ikke godt nok utforsket funn i materialet vårt, er det mangfold av ekstra vurderingssituasjoner som skolen som institusjon eller enkeltlærere setter inn for elever som enten ikke møtte til de ordinære vurderingssituasjonene, som står i fare for å få stryk i standpunktkarakter eller som står og vipper mellom to karakterer når standpunkt skal settes mot slutten av året. Også her beveger vi oss i grenselandet mellom designede rutiner og praksiser som er i daglig eller i det minste i årlig bruk. Vippeprøvene eller ekstraprøvene mot slutten av året er dels innrettet mot å ivareta elevers formelle rett til sluttvurdering, dels skal de på mer uformelt vis sørge for at elever ikke stryker eller at elevene har mulighet til å gi sin beste prestasjon i en sluttvurderingssituasjon. Det kan se ut som om en del lærere strekker seg langt for å oppfylle disse rutinene, og vi som forskere spør om dette kan rokke ved grunnlaget for at elevene skal få en rettferdig standpunktkarakter. Uansett ser vi at elever som har svake forutsetninger i et fag eller som ikke møter frem til ordinære vurderingssituasjoner, blir gitt nye sjanser som andre elever med bedre faglige forutsetninger eller sterkere disiplin i prinsippet ikke får. De skolene som har innført slike muligheter også for elever med sterkere prestasjoner, skiller seg i tilfelle fra skoler som ikke har innført slike muligheter. Dette peker mot en differensiert behandling av elever, enten innfor den enkelte skole eller mellom skoler. Vi kommer nærmere tilbake til dette senere i drøftingen i dette kapittelet.

Det går klart fram av intervjuene at læreres erfaringer med det å være sensor kontra det å være faglærer gir to typer av uttalelser om de velkjente forskjellene mellom eksamenskarakter og standpunktkarakter. Lærerne som har sensorerfaring, forteller om det de kaller for «sensorbriller», og som her betegner vurdering av en anonym elevbesvarelse hvor det kun er vurderingskriteriene som kommer til anvendelse. Som faglærere med ansvar for å sette standpunktkarakter, sier imidlertid disse lærerne (i likhet med de øvrige lærerne vi har intervjuet), at kjennskap til eleven og dennes faglige utvikling vil påvirke læreren i vurderingssituasjonen på en overveiende positiv måte. Som faglærer vil læreren derfor anstrenge seg for å se elevens arbeid i et fordelaktig lys, gjenkjenne elevens måte å uttrykke seg på og tolke dette til beste for eleven. Norsklæreren kan lete etter «den røde tråden» i elevens besvarelse, mens matematikklæreren kan gjøre en helhetlig vurdering av elevens forståelse for matematikk, selv om ingen av disse elevene har utpregede evner til å uttrykke seg skriftlig. Dette er forhold som ifølge lærerne vil ha langt mindre innvirkning på en eksamenssensor, og som kan bidra til å forklare hvorfor det oppstår avvik mellom eksamens- og standpunktkarakterer. Det forklarer likevel ikke hvorfor *forskjeller i slike avvik* fremkommer mellom private og offentlige skoler i det statistiske materialet i vår analyse.

Det har vært en bemerkelsesverdig enighet blant elevene om at vurderingskriteriene i matematikk er klare, forståelige og forutsigbare. Vurderingskriteriene i norskfaget er imidlertid diffuse, sier mange av elevene. Vi finner en betydelig forskjell mellom to skoler med hensyn til elevenes formening om hvordan deres prestasjoner i norsk bedømmes. Ved den ene skolen utrykte flere av elevene en overbevisning om at kriteriene var personavhengige, de var umulige å forstå samt at anstrengelser for forbedring ikke ga noen uttelling. Ved den andre skolen mente elevene å ha god oversikt over vurderingskriteriene. Begge disse skolene er offentlige skoler. En viktig forskjell mellom dem er at ved den skolen hvor elevene mente de fikk god og nyttig veiledning i norskfaget, er en skole som har drevet Vurdering for læring i flere år. Her var også lærerne mest eksplisitte på hva de leter etter når de vurderer besvarelser i norsk. Dette betyr at norskfaget ikke nødvendigvis har mystiske eller uforståelige vurderingskriterier, men at det avgjørende spørsmålet er om lærere har trening i å gjøre kriteriene eksplisitte og forståelige for elevene.

6.2 Hvorfor er det avvik mellom eksamen og standpunktkarakterer mellom offentlige og private skoler?

Så langt har rapporten vist at de private allmenne skolene har et større avvik mellom standpunkt- og eksamenskarakter enn private kristne, private spesialiserte og offentlige skoler har. Vi har utførlig beskrevet det vell av rutiner og praksiser som lærerne bruker i vurdering og karaktersetting, og vi finner ikke her noen forskjeller mellom offentlig og private skoler i måten lærerne arbeider på eller hvordan de tillemper retningslinjer eller forholder seg til rutiner og praksiser. Dermed ser det ikke ut til at dette kan forklare hvorfor det er større avvik mellom standpunkt- og eksamenskarakter ved noen av de private skolene. Oppfølgingsspørsmålet blir da om dette kan forklares ved hvordan skolene arbeider for å forbedre vurderingspraksisen.

Vi har tidligere i denne rapporten vist to tabeller over de åtte skolene vi har besøkt. I tabell 2.1 viste vi skolenes avvik mellom standpunktkarakter og eksamenskarakter, mens vi i tabell 4.1 viste hvilke skoler som har egne satsninger eller prosjekter for vurdering. Hvis vi kombinerer opplysningene i disse to tabellene, får vi følgende resultat:

Tabell 6.1 Oversikt over skolenes eierform, deres avvik og utvikling i forholdet mellom standpunkt- og eksamenskarakterer og deres arbeid med vurdering

Skole ID	Eierform	Avvik og utvikling	Satsninger eller prosjekt for å forbedre vurderingspraksis
Skole 1	Privat	Stabilt høyt avvik	Skoleinternt prosjekt for vurdering
Skole 8	Privat	Endring fra høyt til gjennomsnittlig avvik	Ikke eget prosjekt, men uttalt oppmerksomhet rettet mot eventuelle avvik mellom eksamens- og standpunktkarakter.
Skole 2	Privat	Relativt stabilt lavt avvik	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering
Skole 5	Privat	Relativt stabilt lavt avvik	Skoleinternt prosjekt for vurdering, også samarbeid med UH-miljø.
Skole 6	Offentlig	Relativt høyt stabilt avvik	Satsning på vurdering i samarbeid med UH-miljø
Skole 4	Offentlig	Endring fra høyt til lavt avvik	Workshops knyttet til satsningen Vurdering for læring.
Skole 3	Offentlig	Stabilt gjennomsnittlig avvik	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering
Skole 7	Offentlig	Stabilt lavt avvik	Ikke eget prosjekt eller satsning, løpende arbeid med vurdering

Av tabell 6.1 går det frem at fire av de åtte skolene er involvert i satsninger eller prosjekter for vurdering. Av disse skolene har to av dem stabilt høyt eller relativt stort stabilt avvik mellom eksamens- og standpunktkarakter, mens en skole har gått fra høyt til lavt avvik og den siste skolen har hatt et relativt stabilt lavt avvik mellom disse to karaktertypene. På den andre siden kan vi se at skolene som ikke oppgir å ha prosjekter eller satsninger for å forbedre vurderingspraksis, kan vise til stabilt lavt eller stabilt gjennomsnittlig avvik mellom disse to karaktertypene. Hvis vi legger til grunn at disse skolene mener de har kontroll på vurderingsarbeidet slik at dette inngår i den løpende faglige diskusjonen i kollegiet, kan vi si at bevissthet om betydning av vurderingsarbeid synes gjennomgående høy i utvalget av skoler, og at skoler med høyt avvik har satt inn tiltak. Vi kan også slå fast at det ikke fremtrer noen tydelige forskjeller mellom offentlige og private skoler i dette.

I vårt utvalg av åtte skoler oppsummert i tabell 6.1 kan vi derfor hevde at ulik interesse, oppmerksomhet og bevissthet om vurderingsarbeid ikke synes å være årsaken til at det inntreffer forskjeller mellom offentlige og private skoler hva angår avvik mellom eksamen og standpunktkarakterer. I denne rapporten har vi også vist at vurderingsrutinene i seg selv og hvordan de blir praktisert på disse åtte skolene, heller ikke bidrar til å forklare slike forskjeller. Her tar vi forbehold for praktiseringen av ekstraprøver eller mindre formelle vurderingssituasjoner som kan skape forhandlingssituasjoner for elever som vil forbedre sin karakter, og som vi skal komme nærmere tilbake til i det følgende.

I siste del av dette kapittelet skal vi nemlig drøfte noen særlige problemstillinger som kommer til syne i rapporten, og vi søker å belyse og klargjøre disse problemstillingene ved å reise noen nye spørsmål som kan egne seg til videre forskning.

6.2.1 Betydningen av tolkningsfellesskap

Da vi under intervjuene spurte lærerne hvordan de ville definere eller beskrive hva som er en «rettferdig karakter», fikk vi som regel til svar at det er en karakter som viser elevens kompetanse i henhold til kompetansemålene. Svaret var påfallende ensartet på tvers av skoler, og er en gjenspeiling av vurderingsforskriften implementert som en designet rutine. Lærernes svar på dette spørsmålet er selvfølgelig gyldig, men ikke i seg selv analytisk interessant idet det ikke sier noe om hvordan man som lærer kan arbeide for å avspeile elevens kompetansenivå i karaktersetningen. Den ene læreren som gikk videre inn i dette resonnementet, utdypet ved å svare at en rettferdig karakter er et resultat av at lærerne bruker sitt faglige skjønn til å vurdere elevens kompetanse til et riktig nivå. Dette synes vi som forskere er et godt og analytisk interessant svar, også fordi det oppsummerer et av de mest sentrale fellestrekkene som vi kan finne i alle lærerintervjuene. Alle intervjuene viser nemlig at lærerne bruker sitt faglige skjønn til å sette det de mener er riktig karakter, og at dette faglige skjønnets utvikles gjennom et løpende og tidvis innfløkt samspill mellom designede og daglige rutiner og i samhandling i fagfellesskapet.

Gjennom intervjuene har lærerne på alle skoler, offentlige så vel som private, gitt utfyllende, brede og innholdsrike beskrivelser av hva slags faglig samarbeid som finner sted mellom lærerne på skolen. Noen steder har skoleledelsen satt inn konkrete tiltak for å utvikle og forme fagsamarbeidet mellom lærerne, andre steder er dette overlatt til seksjonene, og noen steder har vi hatt inntrykk av at fagsamarbeid «alltid» har eksistert. Dette kan følge av at enkeltlærere har overført sin tolkning og vurderingspraksis til yngre eller nyansatte kolleger, slik at den nyansatte har blitt sosialisert inn i et vurderingsfellesskap med bestemte normer og standarder som er i bruk. Etter et år eller så kan den nyansatte så ha følt seg «utlært», og fortsatt rette- og vurderingspraksis mer eller mindre på egenhånd – med innslag av kollegasamarbeid hvis det skulle oppstå tvil om enkeltbesvarelser. Andre skoler eller endog andre fagseksjoner på samme skole kan ha andre normer for fagsamarbeid om vurdering.

På flere skoler snakket lærerne om en «skolekultur» for retting, som de var mer eller mindre fornøyde med. Med skolekultur mente lærerne at det på skolen hersket normer for hvordan vurdering skal foregå, og at disse normene i noen grad kan skille seg fra praksis på andre skoler. Normene har blitt overført fra eldre til yngre eller i det minste nyansatte lærere. Noen av lærerne var bekymret for om måten de vurderer sine elever på, gir elevene det riktige grunnlaget for å konkurrere med elever fra andre skoler i søkning til høyere utdanning. Slike bekymringer dreide seg i mindre grad om hvorvidt skolen som «våre» lærere jobber på, vurderer elevene rett, men om lærere på andre skoler gjør det på samme måte med sine elever, slik at konkurransen blir rettferdig. Lærerne som vi fikk snakke med i intervjuene, var nok blant dem som mente at de var bærere av en sunn vurderingskultur, men på flere skoler var det lærere som i intervju antydte at «ikke alle» på samme skole fulgte de samme rutinene eller rettet og vurderte elevbesvarelser helt på samme måte som dem. Disse lærerne fremhevet betydningen av kommunikasjon og lærersamarbeid nettopp for å ha et tolkningsfellesskap.

I kapittel 4 og i tabell 6.1 går det fram at mange skoler har egne prosjekter eller deltar i satsninger for å utvikle og forbedre vurderingspraksisen. Gjennom intervjuer med lærere på private skoler har vi i noen grad fått inntrykk av at skoleledelsen der har satt inn tiltak for å endre vurderingspraksis etter at oppslag i riksmidia har satt søkelys på vurderingspraksis og avvik mellom eksamenskarakter og standpunkt karakter nettopp på private skoler. Med bakgrunn i e-postintervjuene med rektorer på de samme skolene kan vi ikke slå fast om det er riktig at arbeid for å forme en vurderingskultur på skolen har kommet i stand på grunn av medieoppslag, eller er ledd i en mer langvarig innsats. Likevel vil vi holde muligheten åpen for at designede rutiner fra skoleledelsen for å forme en skolekultur for vurdering på noen private skoler kan ha kommet til nokså nylig, og at lærerne på disse skolene i noe større grad har vært vant til å måtte lene seg på dagligrutiner for vurdering som har oppstått som følge

av opplevde behov og grad av velvilje for samarbeid i et kollegium. Kvalitetene på miljøet og skolekulturen som omgir disse dagligrutinene, får derfor stor betydning.

På denne bakgrunnen spør vi om lærere på private skoler har tilgang til tolkningsfellesskap som er av tilstrekkelig størrelse og som er sammensatt med den kompetanse og erfaring som kan tjene utvikling og anvendelse av vurderingskriterier på måter som er godt forankret i regelverk og som er i samsvar med en gjengs oppfatning om slike kriterier som hersker på andre skoler? I denne rapporten har vi ikke spesielt undersøkt kvaliteten på disse tolkningsfellesskapene, og lærerne som hører til de private skolene i materialet vårt, viser variasjon i alder så vel som ansiennitet. Med andre ord er det mye erfaring i lærerstaben i private skoler, hvor mange av lærere også kan fortelle om erfaring fra tidligere ansettelsesforhold i offentlig skole. Likevel spør vi om noen av disse skolene kanskje kan ha lite formaliserte tolkningsfellesskap, eventuelt begrenset til samarbeid mellom enkeltkolleger. Da kan lærere på disse skolene bli «øyer» av isolert vurderingspraksis i et «hav» av offentlige skoler underlagt et skoleleiernivå med samordnings- og koordineringsansvar, som også kan gi lærere muligheter for å møtes og å diskutere på tvers av skoletilhørighet.

Vårt spørsmål om hvorvidt tolkningsfellesskapene på de private skolene er store nok og har nødvendig kvalitet, beror i noen grad på intervjudata som viser at tre av de fire privatskolene i vårt materiale har innført rutiner for resultatmåling der lærerne blir målt på avvik mellom eksamens- og standpunkt karakterer. En del av lærerne vi har snakket med, opplever dette som problematisk. En hypotese kan derfor være at nokså nylig innførte designede rutiner for endret vurderingspraksis på disse skolene, kan rette seg mer mot *resultatet forstått som avvikene i vurderingspraksisen* fremfor det å forme og utvikle lærernes *tolkningsfellesskap som premiss for å skape slike resultater*. Det kan ellers være verd å merke seg at ingen av lærerne på skolene vi har besøkt, mener at de har fått noen egentlig opplæring i vurdering verken som del av utdannelsen eller når de ble ansatt som lærere. Nettopp derfor blir betydningen av kvaliteten på de eksisterende tolkningsfellesskapene, som nye lærere skal sosialiseres inn i, etter vår mening så viktig.

6.2.2 *Betydningen av sensorerfaring og andre muligheter for kalibrering av vurderingsrutiner*

I forlengelsen av lærernes generelle vektlegging av tolkningsfellesskapets betydning for vurderingspraksis, er lærerne med sensorerfaring og deres tilnærmet unisone uttrykk for glede over og tillit til det rette- og tolkningsfellesskap som oppstår i arbeidet med fellessensur. I kapittel 4 i denne rapporten har vi sett at lærere som har slik erfaring, mener sensorarbeidet ikke bare styrker dem i deres egen vurderingspraksis, men at de også styrker vurderingsfellesskapet på skolen sin ved å trekke erfaringene fra sensormøter inn i samarbeidet med kolleger. Lærere i materialet vårt som har erfaring som sensor, er kanskje mindre bekymret for at det utvikler seg en egen og avvikende skolekultur for vurdering enn lærere som ikke har slik erfaring, fordi de kjenner seg trygge på at både de selv og lærere fra andre skoler vil legge vekt på de samme tingene ved avsluttende eksamen. Dermed blir en del av jobben til disse lærerne å ta med seg de samme standardene inn i sin egen undervisning, noe de anser som en helt overkommelig og rimelig oppgave.

I denne rapporten har vi vist til at eksamen og vurderingen av eksamens karakterer fremstår som en designet rutine i nesten ekstrem grad, på den måten at sensor er avskåret fra nesten alle dagligpraksiser som preger en lærer i et kollegafellesskap når han eller hun skal sette standpunkt karakter. Dette vil i første rekke bety at sensor ikke legger kjennskap til eleven til grunn for vurderingen, noe som innebærer at elever ikke vil behandles differensielt ved å «belønnes» med en karakter fordi de har vist slik god faglig utvikling over tid eller står helt på kanten av stryk. I kapittel 4 har vi sett hvordan slike argumenter kan aktualiseres av faglærer i vurdering til standpunkt. Selvfølgelig må sensor vurdere besvarelser analytisk – alltid i norsk, og ofte også i matematikk – hvor de designede rutinene kan gli over i kvalitative og noe mindre formaliserte måter å arbeide på. Denne vurderingen skjer likevel på bakgrunn av sensurmøter med kolleger fra andre skoler, som ikke har de personlige bindinger og sympatier overfor de andre deltakerne som ellers kan være virksomme innenfor en hvilken som helst arbeidsplass. En lærer vi intervjuet, og som vi gjenga i kapittel 4,

understreket hvordan man som sensor og i sensurmøtene ble oppfordret til å legge fra seg sine «hangups» og få brynet seg mot kolleger som kanskje har et annet blikk på vektning av vurderingskriteriene i for eksempel norskfaget. Sensurmøtene blir derfor en slags nøytral grunn for disse lærerne hvor selve faget står i sentrum, og hvor hver besvarelse er anonym. Den rendyrkede formen som de designede rutinene for vurdering får ved disse sensurmøtene, styrker åpenbart lærernes opplevelse av trygghet og etterprøvbarhet i rettellesarbeidet både ved eksamenssensur og i det løpende arbeidet for vurdering som de kan justere på bakgrunn av disse erfaringene.

Uten å ha undersøkt fordelingen av lærere ansatt på private og offentlige skoler som deltar i fellessensur, spør vi derfor om sensorinstituttet generelt kan gjøres mer tilgjengelig for lærere i videregående skoler, og om lærere ved private skoler i særlig grad bør inkluderes i dette arbeidet? Lærere i materialet vårt som har deltatt på sensorskolering, anbefaler på generelt grunnlag at sensorskolering blir obligatorisk, fordi de mener at utbyttet for egen praksis er så høyt. På den annen side er det også lærere i materialet vårt som understreker at de ikke er interessert i sensurarbeid fordi arbeidsmengden er stor, og den lønnsmessige kompensasjonen liten. Enkelte lærere i materialet vårt som har sensurerfaring, peker på at det ofte er eldre og erfarne lærere som tar på seg slikt arbeid, og at selv om dette kan tjene de tolknings- og vurderingsfellesskapene som oppstår, kan yngre lærere bli utelukket fra slike muligheter både fordi de ofte har veldig mye å gjøre, og fordi eldre kolleger ikke villig gir fra seg plass i slike fellesskap. Følgelig spør vi om det er mulig å gjøre strukturelle endringer som kan åpne disse fellesskapene slik at flere lærere kan ha nytte av sensorerfaring og sensorskolering.

Arbeidsmengde og økonomisk kompensasjon har vært nevnt som hindringer, men på den annen side er det mange svært entusiastiske lærere i materialet vårt som åpenbart vil delta i fellessensur til tross for slike ulemper. Likevel skal vi ikke si at dette er forhold som ikke betyr noe, og kanskje særlig med tanke på unge og nyansatte lærere som har svært mange rutiner å sette seg inn i og liten erfaring å flyte på i det daglige arbeidet. Kanskje er det mulig å se for seg at innholdet i sensorskolering og arbeidsformen i sensurmøtene kan alminneliggjøres i større grad, og gjøres tilgjengelig for lærere som kurs- eller opplæringstilbud uten at de samtidig må ta for seg 200 eksamensbesvarelser. Dette er det vanskelig å svare på med bakgrunn i vårt materiale, som ikke har undersøkt sensorinstituttet i seg selv. Likevel er det et solid funn i vårt materiale at utbyttet av sensorskolering og sensurmøter er stort for lærerne som deltar, og at dette gir dem en særlig trygghet i egen vurderingspraksis. Vi ser mulige fordeler ved at disse erfaringsfellesskapene kan gjøres tilgjengelig for flere. Her nevner vi at en rektor på en privat skole sier i e-postintervju at de har spurt om å la lærere få delta i den offentlige skoleeierens arrangementer rundt eksamensforberedende prøver, men fått avslag. Om ikke offentlige skoleeiere skal pålegges ansvar for å la private skoler få delta i slike tilbud, kan kanskje nasjonale myndigheter i større grad legge til rette for at også private skoler får tilgang til slike større fagfellesskap som kan virke bekreftende eller justerende på egen vurderingspraksis.

6.2.3 *Betydningen av forhandlingsrom i karaktersetting*

Når vi i denne rapporten har sett på beskrivelser av designede rutiner og daglige praksiser i læreres vurderingsarbeid, er det et felt som har skilt seg ut i materialet med å være mangfoldig inntil det uoversiktlige i overgangen mellom formelle og ikke-formelle handlingsrepertoar for lærerne. Dette feltet omfatter de ekstraprøver og vurderingssituasjoner som skal hjelpe lærerne (og elevene deres) til en riktig standpunktvurdering, hvis eleven enten ikke møter frem til de ordinære prøvene og vurderingssituasjonene, eller resultatene fra disse vurderingssituasjonene er uklare. I intervjuene har vi mange beskrivelser fra lærerne om vanskeligheter som oppstår i disse gråsonene, hvor noe av ansvaret legges på et regelverk som krever at eleven får mange sjanser til sluttvurdering, og noe av ansvaret legges på eleven selv for å møte opp og være skikkelig forberedt. Mange lærere vi har snakket med, problematiserer de byråkratiske og juridiske sidene ved å gi en elev «ikke vurdert» istedenfor en karakter, og mener at regelverk og muligheter for klage pålegger lærerne å skape disse ekstra vurderingssituasjonene for elevene. Andre viser til et opplevd press for å få flest mulig elever gjennom videregående skole.

Samtidig har vi snakket med flere lærere som synes villige til å strekke seg langt for å tilrettelegge for elevene sine slik at de kan få vist hva de kan. Dette vil disse lærerne gjøre av omsorg og omtanke for elevene sine, ikke minst elever som står i fare for å stryke. De erfarer at disse elevene kan være veldig sårbare og vil ikke legge sten til byrden for dem når de skal videre ut i voksenlivet. Derfor har vi uttalelser fra lærere i rapportens kapittel 4 som viser at man vil gi elever stadig nye sjanser til å prestere på et minimum for å oppnå karakteren 2, eventuelt ved å be elevene forberede seg på bakgrunn av en liste som viser vurderingskriteriene man må oppfylle. Videre vil lærere be kolleger om å lete gjennom besvarelser og tidligere vurderinger som kan forsvare en ståkarakter. På den andre siden er det lærere i materialet vårt som mener at en slik praksis står på grensen til å være forkastelig, fordi dersom eleven kommer opp til eksamen i det aktuelle faget og stryker, vil han likevel ikke få bestått på vitnemålet. Disse lærerne problematiserer også det å «skyve problemet foran seg», slik at svake elever sendes videre i utdanningssystemet med åpenbart mangelfullt kompetansegrunnlag.

Omtalen elever ved en av de private allmenne skolene gjorde av hvordan det går an å be om vippeprøve og at noen lærere kan imøtekomme et slikt ønske ut fra at de vil eleven oppriktig vel, er gjengitt i kapittel 5. Strategisk fravær, som også ble tematisert ved en av de allmenne private skolene, peker også i retning av at det kan oppstå rom for forhandlinger og en ekstra vurderingssituasjon. For at vi skulle kunne konkludere tydeligere om sammenfall mellom disse fenomenene og eierforhold, måtte vi ha undersøkt dette systematisk ved samtlige skoler, og dette har ikke vært mulig fordi temaene kom opp i forløpet av intervjuene og etter at flere av skolebesøkene var gjennomført.

Ikke desto mindre viser materialet vårt at slik tilrettelegging skjer, og at dette i noen grad ser ut til å være opp til den enkelte lærer. Selv om materialet vårt ikke gir noe solid grunnlag for å si at slike ekstra vurderingssituasjoner er mer utbredte og omfattende på private skoler enn på offentlige skoler, vil vi likevel stille spørsmål om dette kanskje kan være tilfelle på noen private skoler.

6.2.4 *Betydningen av formelle tilpasninger til designede rutiner*

De fleste av lærerne vi har snakket med, mener at standpunkt-karakter og eksamens-karakter er to forskjellige ting, og at man ikke kan forvente nøyaktig sammenfall mellom de to karakterene. Dette er et gjennomgående funn, og gjelder følgelig for lærere på offentlige så vel som på private videregående skoler i vårt materiale. Likevel har vi oppfattet det som om problematiseringen av dette forholdet har vært mer fremtredende blant lærerne på private skoler sammenlignet med de offentlige skolene om vi har fått besøke. På de private skolene kan flere av lærerne fortelle at de blir målt på avvik mellom eksamens-karakter og standpunkt-karakter. Dette er det bare én offentlig videregående skole i materialet vårt som kan berette om, og vårt inntrykk var at på flere av de offentlige videregående skolene ble ikke slikt avvik tillagt stor vekt i den løpende kontakten mellom ledelse og kollegium.

Tidligere i dette kapittelet har vi nevnt de prosjekter og satsninger for bedre vurderingspraksis som mange av de åtte skolene i utvalget vårt har satt i verk, og vi har spurt om noen av disse prosjektene kan være mer innrettet mot å få ned selve avviket mellom de to karaktertypene enn å gripe inn i formingen og utviklingen av tolkningsfellesskap som skal lede til ønsket vurderingspraksis. Igjen understreker vi at vi ikke har særlig mye data om disse prosjektene som skolene deltar i, og at vi derfor trekker slike slutninger på et rent hypotetisk grunnlag.

Hypotesen styrkes likevel noe av det vi erfarte på en av de private skolene, hvor to lærere reflekterte over egen vurderingspraksis etter at det over tid hadde vært svært høy oppmerksomhet i lærerkollegiet om avvik mellom eksamen- og standpunkt-karakter. Disse lærerne sa at de satte standpunkt-karakter med tanke på hva den aktuelle eleven hadde fått på eksamen hvis eksamen var i dag, og så i dette altså bort fra alle de dagligrutinene og akkumulerte erfaringene som de som faglærere hadde om eleven. Vi kan si at de gjorde standpunkt-vurdering med «sensorbriller», og la en tilnærmet rendyrket designet vurderingsrutine til grunn for karakteren. Vi nevner dette her fordi de samme lærerne mente at denne måten å vurdere standpunkt-karakter ikke var i tråd med deres faglærerideal, men at fremgangsmåten bidrar til å oppfylle ledelsens mål om minst mulig avvik mellom

eksamen og standpunktkarakter. En parallell kan trekkes til lærere på en annen privat skole, som i intervju sa at dersom det er meningen at standpunktkarakter og eksamenskarakter skal være sammenfallende, kan standpunktkarakteren i realiteten sløyfes.

Disse to eksemplene understreker at lærerne basert på sitt profesjonelle skjønn mener at de to karakterene måler ulike ting, og at det å sette «standpunktkarakter med sensorbriller» blir en tilpasning til regelverk og designede rutiner fremfor å gjøre en reel inngripen i selve vurderingspraksisen. Som forskere ser vi at føringer for samsvar mellom eksamens- og standpunktkarakter kan skape tilpasninger i lærernes praksis, og vi har også beskrevet hvordan enkelte lærere tar med seg erfaringer fra blant annet sensorskolering inn i egen undervisning og i fagfellesskapet, og hvordan lærere utvikler rutiner for samarbeid. Dette vurderer vi som positive og ønskede tilpasninger. Samtidig ser vi at slike sammenligninger kan føre til andre og mer administrative vridninger i hvordan lærerne foretar vurderinger, som i dette eksempelet blir omtalt som å gå på akkord med profesjonelle prinsipper for hvordan lærerne mener at de skal jobbe.

Derfor spør vi om opplevde føringer om at eksamenskarakter og standpunktkarakter skal være like, kan resultere i at lærerne endrer sin vurderingspraksis på en rent administrativ måte uten at det blir gjort et reelt arbeid for å utvikle og forbedre de rutiner som ligger til grunn for vurderingsarbeidet? Hensikten med å ha to karakterer må nødvendigvis være at de skal gjenspeile to aspekter ved elevens læring og kompetanseoppnåelse, og at de to karakterene og vurderingsformene skal virke utfyllende og komplettere hverandre som mål på slik sluttkompetanse. Hvis man legger ensidig vekt på eksamen som «fasit», faller noe av hensikten med faglærers standpunktvurdering bort. Her kan det virke for oss som forskere at oppmerksomhet og måling av avvik har skapt noen grad av forvirring og frustrasjon på de enkelte lærerværelser, og at det å bruke slike avviksmål som designede rutiner ikke er godt nok forklart eller kommunisert helt ned på den enkelte skole. Dette kan også bero på tolkninger som er gjort av den enkelte skoleledelse, og av hensyn til for eksempel skolens omdømme eller konkurransesituasjon mer enn lærernes arbeidssituasjon og identitet knyttet til hevd og utvikling av profesjonelt skjønn.

Vi kan oppsummere dette med å si at selv om man oppnår lavere avvik mellom eksamens- og standpunktkarakter, trenger dette etter vår forståelse ikke å være ensbetydende med at man har fått høyere kvalitet i vurderingsrutinene. Man kan, forenklet sagt, også ha skapt bedre tilpasning i lærernes arbeid for å sette standpunktkarakter der malen blir eksamenskarakteren. Alt arbeidet vi har hørt om på skolene som tar sikte på å utvikle og vedlikeholde gode rutiner i vurderingsarbeidet, lar oss forstå at denne oppgaven blir tatt på alvor av lærere og av skoleledelse. Likevel ser vi at designede rutiner må utfylles av læreres tolkning, praksis og samarbeid for å bli forstått, levendegjort og anvendt i det daglige arbeidet slik at de blir aktive redskaper og ikke bare formelle pålegg. Ensidig vurderingsarbeid etter utelukkende designede rutiner uten å ta i bruk vellet av praksiser og dagligrutiner, synes like feil som å bare la dagligrutinene styre. De designede rutinene må utformes, formuleres og anvendes av skoleledelse og kollegafellesskap på skoler på måter som er til støtte for lærernes dagligrutiner, ikke til hinder eller frustrasjon. Likeens må arbeidet med å forme lærernes dagligrutiner ikke miste de designede rutinene av syne.

6.2.5 Oppsummering: tema for videre forskning

I denne siste delen av rapporten har vi vist til mulige nye forskningsspørsmål som kan åpne seg på grunnlag av analysen. Denne analysen slo fast at det på grunnlag av det innsamlede datamaterialet ikke er mulig å fastslå hvorvidt det er forskjeller i anvendelse av rutiner og praksiser for vurdering som kan forklare de forskjellene i avvik mellom standpunktkarakter og eksamenskarakter som vi finner mellom private og offentlige videregående skoler. Vi har kort vist til at interessen og oppmerksomheten om vurderingspraksis synes høy på de åtte skolene vi har besøkt, og har derfor lett etter andre mulige forhold som kan ha innvirkning på vurderingspraksis, og spurt om disse kan variere mellom skolene. Følgelig har vi i dette kapittelet spurt om tolkningsfellesskapene for lærere på private skoler er store nok og har de nødvendige kvalitetene. I forlengelse av dette har vi sett på læreres erfaringer med å

delta i sensoropplæring og fellessensur, og vi har spurt om sensorinstituttet kan gjøres tilgjengelig for flere lærere med særlig vekt på lærere ved private skoler.

Det eneste funnet i vår egen rapport som får oss til å stille spørsmål om det finnes systematiske forskjeller mellom rutiner og praksiser på offentlige og private skoler, gjelder ekstraprøver eller situasjoner for tilleggsvurderinger frem mot standpunktkarakter. Det er et mangfold av slike vurderingssituasjoner på offentlige så vel som på private skoler, men med bakgrunn i data fra elevintervjuene spør vi om forhandlingssituasjonen på de private skolene kan være mer åpen og inneholde flere ekstrasjanser på de private skolene. Dette er kun en hypotese, og må undersøkes nærmere før man kan trekke konklusjoner.

Til sist har vi i dette kapittelet omtalt formelle tilpasninger til designede rutiner som vi har fått beskrevet i enkelte av lærerintervjuene, forstått som at eksamen blir retningsgivende for det en faglærer vil gi eleven i standpunktkarakter. Dette har vi fulgt opp med å vise til at designede rutiner og ikke-formelle praksiser eksisterer på alle skoler, og de bindes sammen i et nett som delvis kan beskrives som en «skolekultur». Det er denne skolekulturen som må være gjenstand for utvikling av vurderingsrutinene, ikke bare administrative tilpasninger for at avviket mellom eksamens- og standpunktkarakter skal være så lavt som mulig. Dersom avvik mellom disse to karakterene skulle vise seg å gå betraktelig ned i fremtiden, er det derfor interessant med forskning som kan bidra til å avdekke hvorvidt lavere avvik kan sies å bero på slike formelle tilpasninger, eller om utviklingen skyldes reelle skoleutviklingsprosesser.

Referanser

- Angelo, T. A. (1998): *Classroom assessment and research: and update on uses, approaches, and research findings*. San Francisco: Jossey-Bass.
- Barnes, S. (1985): A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education* 36(4), 46-49
- Berge, Ø. & C. Hyggen (2011): *Privatskoler i Norden. Omfang, utvikling og den politiske debatten*. Fafo-notat 2011/01. Oslo: FAFO
- Bernstein, B. (1990). *Class, Code and Control: Volume 4 - The Structuring Of Pedagogic Discourse*. London: Routledge.
- Black, P., C. Harrison, C. Lee, B. Marshall & D. Wiliam (2003): *Assessment for Learning: Putting It into Practice*. Buckingham, U.K.: Open University Press.
- Black, P. (2003): Testing Times: Role of Assessment for Learning. *Curriculum Briefing* 2(1), 3-8
- Brookhart, S. M. (1991): Letter: Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35-36
- Brookhart, S. M. (1993): Teachers` Grading Practices: Meanings and Values. *Journal of Educational Measurement*, 30(2), 123-142
- Brookhart, S. M. (1999): *The art and science of classroom assessment: the missing part of pedagogy*. Washington, D. C.: George Washington University.
- Brookhart, S. M. (2013): Grading. I J. H. McMillan (red.), *SAGE Handbook of research on Classroom Assessment* (pp.257-272). USA: SAGE
- Burns, E. (2010): Developing Email Interview Practices in Qualitative Research. *Sociological Research Online*, 15(4), art. 8
- Eggen, A. E. (2004): *Alfa and Omega in Student Assessment; Exploring Identities of Secondary School Science Teachers*. PhD-thesis. Department of Teacher Education and School Research, University of Oslo. Oslo: University of Oslo.
- Galloway, T.A., L.J. Kirkebøen & M. Rønning (2011): *Karakterpraksis i grunnskoler. Sammenheng mellom standpunkt – og eksamenskarakter*. SSB rapport 4/2011. Oslo: Statistisk sentralbyrå.
- Galton, F. (1886): Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246–263.
- Grøgaard, J. B. (2012): *Hva kjennetegner barneskoler som oppnår høy skår på nasjonale prøver? Delrapport 5 fra prosjektet 'Ressurser og resultater i grunnopplæringen'*. NIFU rapport 38/2012. Oslo: NIFU
- Helland, H. & J. Lauglo (2007): Sosial segregering i private grunnskoler? I SSB, *Utdanning 2007. Muligheter, mål og mestring*, Statistiske analyser 90, Oslo/Kongsvinger: SSB
- Hodgson, J., W. Rønning & P. Tomlinson (2012): *Sammenhengen mellom undervisning og læring. En studie av læreres praksis og deres tenkning under Kunnskapsløftet. Sluttrapport*. NF-rapport nr. 4/2012. Bodø: Nordlandsforskning
- Hægeland T., L. J. Kirkebøen, O. Raaum & K.G. Salvanes (2005): *Skolebidragsindikatorer*. SSB rapport 33/2005. Oslo: Statistisk sentralbyrå.
- Kårstein, A. & T. C. Carlsten (2013): *Evaluerer av Internship-ordningen ved Universitetet i Tromsø*. Oslo: NIFU-rapport 26/2013. Oslo: NIFU

- Manke, M. P. & B. H. Loyd (1990): *An investigation of non-achievement related factors influencing teachers' grading practices*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Boston
- Meho, L. I. (2006): E-mail Interviewing in Qualitative Research: A Methodological Discussion. *Journal of the American Society for Information Science and Technology*, 57(10), 1284-1295
- Melograno, V. J. (2007): Grading and Report Cards for Standards-Based Physical Education. *Journal of physical education, recreation and dance*. 78(6), 45-53
- Morgan, D. L. (1996): Focus groups. *Annual Review of Sociology*, 22, 129-152.
- Opdenakker, R. (2006) Advantages and Disadvantages of Four Interview Techniques in Qualitative Research. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 7(4), art. 11
- Opheim, V., J.B. Grøgaard & T. Næss (2010). *De gamle er eldst? Betydning av skoleressurser, undervisningsformer og læringsmiljø for elevenes prestasjoner på 5., 8. og 10. trinn i grunnopplæringen*. NIFU STEP-rapport 34/2010. Oslo: NIFU STEP.
- Prøitz, T.S. & J.S. Borgen (2010): *Rettfærdig standpunktbedømming. Det (u)muliges kunst?* NIFU STEP rapport 16/2010. Oslo: NIFU STEP
- Prøitz, T. S. (2013) Variation and validity in final grading – subjects matter. *Education Inquiry*, 4(3), <http://dx.doi.org/10.3402/edui.v4i3.22629>
- Resh, N. (2009): Justice in grades allocation: teachers' perspective. *Social Psychology of Education*. 12(3), 315-325
- Riksaasen, R. & B. Vigeland (1994): *Basil Bernsteins kodeteori og nyere empiri*. Trondheim Tapir.
- Seland, I., N. Vibe & E. Hovdhaugen (2013): *Evaluerer av nasjonale prøver som system*. NIFU-rapport 4/2013. Oslo: NIFU
- Skolverket (2005): *Nationella prov i gymnasieskolan – ett stöd för likvärdig betygsättning?* Tilgjengelig på www.skolverket.se, nedlastet 6. juni 2013.
- Spillane, J. (2012): Data in Practice: Conceptualizing the Data Based Decision-Making Phenomena. *American Journal of Education*. 118(2), 113-141
- Stiggins, R. J., D. A. Frisbie & P. A. Griswold (1989): Inside High School Grading Practices: Building a Research Agenda. *Educational Measurement: Issues and Practice*, 8(2), 5-14.
- Stiggins, R. J. & N. F. Conclin (1992): *In teachers' hands: Investigating the practices of classroom assessment*. Albany: SUNY Press
- Thronsdén, I., T. Hopfenbeck, S. Lie & E. L. Dale (2009): *Bedre vurdering for læring – Rapport fra prosjektet Evaluerer av modeller for kjennetegn på måloppnåelse i fag*. Oslo: Universitetet i Oslo
- Vibe, N. & M. Evensen (2009): *Spørsmål til Skole-Norge høsten 2009*. NIFU STEP-rapport 45/2009. Oslo: NIFU STEP
- Vibe (2012): *Spørsmål til Skole-Norge Våren 2012*. NIFU-rapport 21/2012. Oslo: NIFU
- Vibe, N., P.O. Aamodt & T.C. Carlsten (2009): *Å være ungdomsskolelærer i Norge*. NIFU STEP-rapport 23/2009. Oslo: NIFU STEP
- Vibe, N. & J. S. Borgen (2007): *Å lese og lære det; ja. Men å skrive det, nei*. NIFU STEP-rapport 33/2007. Oslo: NIFU STEP
- Vibe, N, M.W. Frøseth, E. Hovdhaugen & E. Markussen (2012): *Strukturer og konjunkturer. Evaluerer av Kunnskapsløftet. Sluttrapport fra prosjektet «Tilbudsstruktur, gjennomføring og kompetanseoppnåelse i videregående opplæring*. NIFU-rapport 26/2012. Oslo: NIFU
- Vibe, N. & B. Lødning (2014) *Spørsmål til Skole-Norge våren 2014*. NIFU-rapport 26/2014. Oslo: NIFU

- Wiborg, Ø., C.Å. Arnesen, J.B. Grøgaard, L.A. Støren & V. Opheim (2011): *Elevers prestasjonsutvikling – hvor mye betyr skolen og familien?* Andre delrapport fra prosjektet «Ressurser og resultater». NIFU-rapport 35/2011. Oslo: NIFU.
- Wiliam, D. (1996): National Curriculum Assessments and Programmes of Study: validity and impact. *British Educational Research Journal*, 22(1), 129-141.
- Vlachos, J. (2010): *Betygets värde. En analys av hur konkurrens påverkar betygssättningen vid svenska skolor*, Rapport 2010:6, Stockholm: Konkurrensverket.
- Wyatt-Smith, C. & V. Klenowski (2012): Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assessment in Education: Principles, Policy & Practice*, 20(1), 35-52.
- Aasen, P., J. Møller, E. Rye, E. Ottesen, T. S. Prøitz & F. Hertzberg (2012): *Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen*. NIFU-rapport 20/2012. Oslo: NIFU

Nettbaserte referanser

Offentlige dokumenter/nettsider:

Forskrift til opplæringsloven 2013, <http://lovdata.no/dokument/SF/forskrift/2006-06-23-724>

KD (2014): *Endringer i privatskolelova*. Pressemelding 13.01.2014
<http://www.regjeringen.no/nn/dep/kd/Pressesenter/pressemeldingar/2014/endringer-i-privatskolelova.html?id=749098>

Utdanningsdirektoratet (2013a): *Endringer fra standpunkt til eksamen*. Statistikknotat 06, 2013.
http://www.udir.no/PageFiles/79637/Statistikknotat_13_6.pdf?epslanguage=no

Utdanningsdirektoratet (2013b): *Utdanningsspeilet 2013*.
<http://www.udir.no/Tilstand/Utdanningsspeilet/Utdanningsspeilet/Utdanningsspeilet-2013/1-Fakta-om-grunnskole-og-videregaende-opplaring/110-Fortsatt-nedgang-i-antall-videregaende-skoler-i-Norge/>

Medieoppslag:

Aftenposten 05.02.2013: *Svenske privatskoler setter bedre karakterer enn offentlige skoler*
<http://www.aftenposten.no/nyheter/iriks/Svenske-privatskoler-setter-bedre-karakterer-enn-offentlige-skoler-7111072.html>

Aftenposten 15.02.2013: *Se om skolen din er streng eller snill med karakterene*
<http://www.aftenposten.no/nyheter/iriks/Se-om-skolen-din-er-streng-eller-snill-med-karakterene-7119112.html>

Vedlegg 1 Guide fokusgruppe lærere

Intervjuguide fokusgruppeintervjuer med lærere

Prosjektets problemstillinger/formål:

Generere kunnskap om ulike **praksiser** og **rutiner** som ligger til grunn for standpunktkaraktersetting ved ulike videregående skoler

For å belyse disse problemstillingene vil følgende delproblemstillinger ligge til grunn for intervjuene:

Hvilke rutiner og retningslinjer ligger til grunn for vurdering ved de ulike skolene?

Er det forskjeller i rutiner og retningslinjer ved de ulike skolene?

Lærerinformantene skal ha blitt bedt om å ta med seg evt. hjelpemidler, verktøy de bruker i sitt arbeid med standpunktkaraktersetting.

Bakgrunnsinformasjon: informanter (oppvarming)

For å få et bilde av hvem dere er og hvilken skolefaglig bakgrunn dere har vil jeg gjerne starte med en kort runde rundt bordet om deres skolefaglige bakgrunn, hvilke elevgrupper dere har ansvar for nå. Kan dere si noen ord om:

Hvor lenge har dere jobbet på denne skolen?

Har dere jobbet ved andre skoler/tidligere yrkesbakgrunn?

Hvilken faglig bakgrunn har dere?

Hvilket fag og hvilke elevgrupper underviser dere i dag?

Vi vil også gjerne vite litt om hvilke erfaringer dere har med karaktersetting

Hvor mye erfaring har dere med standpunktkaraktersetting i videregående opplæring?

I hvilke fag?

Har dere vært ekstern sensor ved eksamen?

Skriftlig eksamen? Muntlig eksamen?

Har dere deltatt på sensor opplæring?

Hvordan var det?

Hvis ikke hvorfor?

Har dere deltatt i annen kursing i vurdering/karaktersetting?

Hvem har arrangert dette?

Så vil vi gjerne vite litt mer om hva dere konkret gjør når dere setter standpunktkarakter.

Kan dere beskrive hvordan dere arbeider med å sette standpunktkarakter i Norsk og Matematikk her på skolen?

Hva slags informasjon bruker dere når dere setter standpunktkarakter?

Hvor får dere den informasjonen fra?

Bruker dere noen hjelpemidler for å støtte karaktersettingen av elevene - for eksempel kartleggingsprøver, tester, egne utformede prøver, hjemmeoppgaver, fremføringer, gruppeoppgaver, Udirs eksamensoppgaver.

Hvilke andre hjelpemidler bruker dere?

Hvor får dere de fra, har dere utviklet dem selv her på skolen?

(I hvilken grad bruker du nasjonal læreplan/lokal læreplan som målestokk/forankringspunkt når du setter standpunktkarakter?)

Hvordan går dere frem for å bestemme hvor en elevs arbeid ligger på karakterskalaen?

Hva legger dere vekt på når dere skal bestemme en karakter?

Er det noe dere vektlegger mer/mindre for eksempel om en elev står i vippeposisjon mellom to karakterer? (kompetanse, prestasjon, oppmøte, innsats, progresjon)

Eksempler?

Bruker dere hele karakterskalaen?

Hva gjør dere når en elev må strykes?

Hva gjør dere når det ikke er grunnlag nok til å sette karakter?

Mener dere at det er mulig å få rettferdig vurdering av elevene?

Hva er eventuelt vanskelig?

Hvordan kan dette forklares for elever, foresatte, publikum utenfor skolen?

Hva gjør dere her for å sikre en mest mulig rettferdig vurdering av elevene?

Samarbeider dere om karaktersettingen?

Er det viktig å samarbeide om karaktersetting?

Eksempler?

Hva kan hindre rettferdig vurdering av elevene, mener du?

Eksempler?

Standpunkt karakterer og eksamen

Ved flere anledninger har forholdet mellom standpunktkarakterer og eksamenskarakterer blitt diskutert i media (Aftenposten spesielt – ta med utskrift). Myndighetene har sagt lite om dette.

Hvordan mener dere at forholdet mellom standpunktkarakterer og eksamenskarakterer skal være?

Vi dere si at eksamen er en test på hvordan dere har satt standpunktkarakterer?

Er dette en tematikk dere har diskutert her på skolen?

Hva har dere diskutert?

Eksempler?

Har det ført til endringer i hvordan dere sette standpunktkarakter her hos dere?

(Bruker dere hjelpemidler, eksempeloppgaver o.l. som Utdanningsdirektoratet legger ut på sine nettsider? Hvorfor?)

Eksempler?

Er dere kjent med Utdanningsdirektoratets eksamensveiledning til elever og lærere?

Bruker dere dette dokumentet aktivt selv i undervisningsplanleggingen evt som informasjon til elevene? (ha med og vise)

Eksempler?

Hva slags hjelp eller støtte får dere i arbeidet med karaktersetting?

Mener dere at dere har god nok støtte i dette arbeidet?

Er det noe viktig vi ikke har spurt om som dere mener vi bør ha med oss inn i undersøkelsen?

Tusen takk for at dere har stilt opp til intervju.

Vedlegg 2 Guide fokusgruppe elever

Intervjuguide fokusgruppeintervjuer med elever

Bakgrunn (oppvarming)

Vi vil gjerne vite litt om dere og starte med noen spørsmål om hvilket trinn dere går, hvilke fag dere tar og hvor gamle dere er.

Hvorfor går dere på denne skolen?

Har noen av dere byttet skole eller studieprogram underveis?

Trives dere her?

Er det noe som er spesielt bra med denne skolen?

Er det noe som ikke er så bra?

Om standpunktkarakterer – elevenes synspunkter

Vi vil gjerne vite hva dere tenker om det å få standpunktkarakterer. Vi er spesielt opptatt av fagene norsk og matematikk men vi kan også snakke om andre fag.

Hva synes dere om det å få standpunktkarakterer i Norsk og Matematikk?

Er dere opptatt av karakterer?

Snakker dere mye om karakterer? Hva snakker dere om da?

Dere får standpunktkarakterer og eksamenskarakterer. Hadde det vært nok med eksamenskarakterer?

Hvorfor/hvorfor ikke?

Mener dere at det er nødvendig?

Kunne vi hatt en skole uten standpunktkarakterer?

Er det noe som er bra med det å få standpunktkarakterer?

Er det noe som ikke er så bra?

I hvilke fag er det viktigst å få gode karakterer?

Hvorfor disse fagene?

Hva med andre fag?

Er det viktigere å få gode karakterer i noen fag enn andre?

Hvorfor?

Om standpunkt og strukturer (formelle og uformelle krav, eksamen, oppfatninger om rettferdighet)

Hva må til for å få gode karakterer i norsk og matematikk?

OBS: Sørge for at elevene holder fokus på standpunkt.

Eksempler?

Hvordan vet dere hva som må til?

Får dere informasjon fra læreren om hva som må til?

Vet dere hva som teller til standpunktkaraktersetting, får dere beskjed om det?

Hva er det som teller?

Sammenlikner dere resultater på prøver, besvarelser?

Kan man jukse?

Hvordan?

Vet dere om noen som har jukset?

Når det nærmer seg jul/sommer vet dere da hva slags karakter dere får i faget?

Har dere noen gang blitt overrasket, fått en helt annen karakter enn det du trodde?

Eksempler?

Skjønte du da hva det var som gjorde at du fikk en slik karakter?

Noen mener at standpunktkarakterer skal være så lik eksamenskarakterer som mulig - hva mener dere?

Vet dere hvordan en eksamensoppgave i Norsk eller matematikk ser ut?

Har dere tatt en eksamensoppgave i Norsk og matematikk?

Hva har dere gjort da?

Hvorfor gjør dere det mener dere?

Utdanningsdirektoratet har laget en eksamensveiledning til elever og lærere om eksamen i Norsk og matematikk som ligger på deres nettsider – vet dere om det? (ha med denne og vise)

Mener dere at lærere setter rettferdige standpunktkarakterer?

Hva er en rettferdig karakter?

Har noen av dere klaget på en karakter dere har fått?

Eksempler?

Hvem klaget du til da?

Hva skjedde?

Hva mener dere at lærere/skolen gjør for å sette så rettferdige standpunktkarakterer i Norsk og Matematikk som mulig?

Mitt siste spørsmål til dere er om det er noe jeg burde ha spurt dere om som jeg ikke har gjort?

Tusen takk for at vi fikk komme og intervju dere, det har vært til stor hjelp.

Vedlegg 3 Karakterforskjeller – delrapport levert til Udir i desember 2013

Karakterforskjeller mellom private og offentlige skoler

RESULTATER AV ANALYSER AV KARAKTERDATA
NILS VIBE OG ELISABETH HOVDHAUGEN, NIFU

DELRAPPORT LEVERT TIL UDIR 2. DESEMBER 2013,
REVIDERT UTGAVE JULI 2014

Innhold

1	Innledning.....	3
2	Beskrivelse av datamaterialet	7
2.1	Samlet oversikt over karakterdata.....	7
2.2	Vårt utvalg av fag.....	7
2.3	Kjennetegn ved offentlige og private skoler.....	9
3	Karakterforskjeller i norskfaget	11
3.1	Norsk skriftlig hovedmål.....	11
3.1.1	Hovedresultatene	11
3.1.2	Karakterdifferanse kontrollert for standpunktkarakter	12
3.1.3	Hovedforskjell mellom offentlige og private skoler.....	15
3.1.4	Ulike elevgruppers resultater	16
3.1.5	Regresjonsanalyse: Hva har effekt på differansen mellom karakterene?	17
3.1.6	Foreløpig konklusjon.....	20
3.2	Norsk skriftlig sidemål.....	22
3.3	Norsk muntlig.....	24
3.3.1	Samvariasjon mellom norskkarakterene	24
3.3.2	Karakter i norsk muntlig	24
3.4	Oppsummering	27
4	Karakterforskjeller i andre fag med skriftlig eksamen	29
4.1	Rettslære.....	29
4.2	Realfagsmatematikk	31
4.3	Matematikk for påbyggingselever.....	34
4.4	Oppsummering	35
5	Karakterforskjeller i andre fag med muntlig eksamen	37
5.1	Religion og etikk for elever ved studiespesialisering	37
5.2	Historie for elever ved studiespesialisering	38
5.3	Historie for påbyggingselever.....	41
5.4	Oppsummering	42
6	Konklusjoner.....	43
	Referanser	45
	Vedleggstabeller til kapittel 3	46

1 Innledning

Dette vedlegget er en revider utgave av delrapporten Karakterforskjeller mellom private og offentlige skoler som ble levert til Utdanningsdirektoratet i desember 2013. Delrapporten var ment til internt bruk i Utdanningsdirektoratet frem til prosjektet ble ferdigstilt sommeren 2014. Analysene som er presentert i dette vedlegget er oppsummert i kapittel 3 i denne sluttrapporten. Delrapporten fungerer dermed som et grunnlag for de data og konklusjoner som er presentert i kapittel 3.

I dette vedlegget fokuserer vi på prosjektets første problemstilling: Om det er større avvik mellom standpunktkarakter og eksamenskarakter når private og offentlige skoler sammenlignes. For å gjøre dette sammenligner vi gjennomsnittsavstanden mellom standpunktkarakter og eksamenskarakter for de elever som har vurdering fra både standpunkt og eksamen i samme fag.

Det er viktig å understreke at det å sammenlikne standpunktkarakterer og eksamenskarakterer ikke er uproblematisk. Avvik mellom karakterer gitt til standpunkt og eksamen ikke en ny problemstilling i norsk grunnopplæring (Prøitz & Borgen 2010, Galloway et al 2011). Problemstillingen handler dels om ulike praksiser i standpunktkaraktersetting, eksamensutforming og sensorordninger, dels handler det også om i hvilken grad det er hensiktsmessig å sammenlikne karakterene ettersom de er basert på ulikt vurderingsgrunnlag og per definisjon skal måle ulike sider ved elevens oppnådde kompetanse (Forskrift til opplæringsloven 2013). Samtidig er det et faktum at både nasjonale og regionale skolemyndigheter, skoleledere, journalister og forskere sammenlikner standpunkt og eksamen for ulike formål. Eksamen blir brukt som rettesnor og har på denne måten en kalibrerende funksjon i forhold til standpunktkaraktersettingen (Prøitz & Borgen 2010, Prøitz 2013). Praksis er imidlertid ulik og oppfatningene om hvorvidt karakterene for standpunkt og eksamen ideelt sett skal være like eller hvor store avvik som er akseptable varierer.

Hovedhensikten med delrapporten basert på kvantitative data er å fokusere på forskjeller og likheter mellom private og offentlige skoler i karaktersetting. Analyser av materialet vi baserer studien på viser at i alt 8,3 prosent av elevene går på en privat skole og denne andelen har vært konstant over tid for de fire årene vi studerer. Vi skal dermed sammenligne en relativt liten gruppe av elever, i alt ca 3 300 hvert år, med det store flertallet som går i offentlige skoler. For å kunne gjøre dette er det dermed avgjørende at analysene baseres på tilstrekkelig store utvalg, og vi vil derfor konsentrere oss om noen få fag der vi har karakteropplysninger om tilstrekkelig mange elever ved både offentlige og private skoler. Hvilke fag som har mange nok elever ved både offentlige og private skoler er redegjort for i kapittel 2.

Selv om oppdraget går ut på å se på forskjeller mellom private og offentlige skoler kan likevel begrepet «privatskole» være for snevert i en del sammenhenger ettersom det innenfor denne betegnelsen finnes svært ulike skoler ut fra formål og idegrunnlag. Imidlertid er det ikke alle typer privatskoler som er egnet i en slik analyse: flere av de videregående skolene med alternativ pedagogikk har ikke eksamen eller karakterer, og er uaktuelle i denne sammenhengen. De internasjonale skolene som følger en annen læreplan i tillegg til den norske har eksamen og karakter i andre fag, og vil derfor heller ikke inngå i utvalget av skoler. Til slutt er det mest aktuelt å se på skoler og fag som stort sett gir studiekompetanse fordi det er her vi finner store nok elevgrupper til å kunne gjøre sammenligninger innenfor enkeltfag. Vi har i datamaterialet identifisert i alt 59 privatskoler som lar seg dele inn i følgende grupper etter en grovmasket (horisontal) og en mer finmasket (vertikal) inndeling:

Tabell 1.1 Ulike typer private skoler. Antall skoler (gjennomsnittlig antall elever pr år i parentes).

	Kristne skoler	Allmenne skoler	Spesialiserte skoler
Kristne skoler	19 (1456)		
Idrettsgymnas			11 (668)
Mediegymnas			4 (81)
Naturbruk			5 (150)
Musikk			2 (39)
Allmenne, Norge		14 (1169)	
Allmenn, utland		4 (49)	
Sum	19 (1456)	18 (1219)	22 (939)

De tre største typene private skoler er 19 kristne skoler med i gjennomsnitt 1456 avgangselever pr år i vårt materiale, 11 idrettsgymnas med 668 avgangselever og en gruppe videregående skoler vi har valgt å kalle allmenne private skoler, 14 i Norge med 1169 avgangselever og fire små skoler i utlandet som til sammen har i gjennomsnitt 49 avgangselever i året. Den sistnevnte gruppen er skoler som stort sett tilbyr eksamensrettet privatist opplæring innenfor ulike former for studiespesialiserende programmer, ofte i en kombinasjon med et tilbud med rene videregående skoler med spesielt fokus på for eksempel E-læring. Flere av disse skolene ble godkjent på grunnlag av den såkalte Friskoleloven før omlegging til ny privatskolelov i 2005, og vi vil derfor omtale dem som «allmenne private skoler» i rapporten. Det er også i denne gruppen vi gjenfinner skoler som har markedsført seg med «karaktergaranti».

I tillegg finner vi fem skoler som er spesialisert i naturbruk med 150 avgangselever, fire mediegymnas med 81 avgangselever og to musikkgympnas med 39 avgangselever. I våre analyser kommer vi til enten å bruke hovedskillet mellom private og offentlige skoler, eller benytte den grove inndelingen i offentlige skoler og tre typer private.

I analysene vil vi også trekke inn andre forklaringsfaktorer, som kjønn, alder og programområde elevene går på, da tidligere studier har vist at det er forskjeller i både standpunkt- og eksamenskarakterer mellom gutter og jenter, mellom elever i ulike aldre og etter hvilket studieprogram eleven følger. Vi ser også på forskjeller mellom år, da vi har tilgang på data for skoleårene 2009/10, 2010/11, 2011/12 og 2012/13. Imidlertid er fire år kort tid i forhold til å si noe om utvikling over tid, vi skal heller se på ev. endringer mellom år, og hvilken retning disse endringene har. Vi kontrollerer også for størrelsen på skolene og om elevene går på en skole med høy inntakskvalitet. Skoler med høy inntakskvalitet er definert som den sjettedelen av skolene som har høyest gjennomsnittlig inntakskarakterer fra grunnskolen, det vil si at gjennomsnittskarakterene til elevene ligger godt over landsgjennomsnittet. Argumentet for disse to variablene er at vi fra tidligere studier vet at små skoler kan ha mindre erfaring med karaktersetting og forberedelse til eksamen, hvilket kan påvirke karaktersettingen. Videre har tidligere forskning også vist at det er gunstig for elever å gå på skoler med mange faglige dyktige elever (Grøgaard 2002, 2012, Opheim et al. 2010, Wiborg et al. 2011). Alle forklaringsvariabler vil bli nærmere introdusert når de blir brukt i regresjonsanalysene.

Metodene som brukes er å sammenligne forskjell mellom standpunktkarakter og eksamenskarakter. Dette gjøres dels gjennom å sammenligne de to karakterene direkte, men fremfor alt gjennom å se på gjennomsnittlig forskjell. Den gjennomsnittlige forskjellen mellom de to karakterene kommer til uttrykk som et tall som kan være både negativt og positivt, og kan i prinsippet variere mellom -5 og +5. Men i praksis vil forskjellen i fagene med skriftlig eksamen være negativ for flertallet av elevene fordi de har fått en eksamenskarakter som er lavere enn standpunktkarakteren. Tallet for gjennomsnittlig forskjell brukes også som avhengig variabel i regresjonsanalysene, der vi måler effekten av ulike forhold samtidig som vi kontrollerer for andre forhold. Hoveddelen av figurene viser bivariate sammenhenger. I de tilfeller der figurene er basert på regresjonsanalyse er dette angitt i figurteksten.

I figurene som sammenligner standpunktkarakter og eksamenskarakter presenteres gjennomsnittet og konfidensintervallet for gjennomsnittet, som viser den statistiske usikkerheten for gjennomsnittet. Gjennomsnittet er marker med et punkt, mens konfidensintervallet vises med en strek som går ut fra punktet. Jo mindre antallet er som ligger til grunn for gjennomsnittet, desto større blir konfidensintervallet. Derfor har gjennomgående de private skolene større konfidensintervall enn de offentlige skolene, fordi det er færre elever på de private skolene.

2 Beskrivelse av datamaterialet

I dette kapitlet beskriver vi datamaterialet som er benyttet for analysene av elevenes karakterer og forskjellen mellom eksamenskarakter og standpunktkarakter.

2.1 Samlet oversikt over karakterdata

Datamaterialet består av komplette karakterdata for i alt fire skoleår fra 2009 – 2010 til 2012 – 2013 som Utdanningsdirektoratet har stilt til disposisjon. I utgangspunktet består materialet av tre datalinjer for hvert enkelt fag som er relevant for den enkelte elev. Datamaterialet for karakterer våren 2013 består av 3 814 709 linjer som dekker 1 271 569 fag for 184 993 elever, i gjennomsnitt 6,87 fag pr elev. Det er gitt i alt 1 422 039 karakterer, eller 7,69 pr elev. Når antallet karakterer er høyere enn antallet fag, skyldes dette at det vil være både standpunktkarakterer og eksamenskarakterer i enkelte fag. Tabellen viser omfanget for et enkelt skoleår.

Tabell 2.1 Karakterdata for skoleåret 2012 – 2013. Avsluttende vurdering.

	Antall	Andel med karakter	Gj. snitt pr elev	Modus = Vanligste antall	Median = Midtpunkt
Antall datalinjer	3 814 707		20,62	30	24
Antall fagopplysninger	1 271 569		6,87	10	8
Antall elever	184 993				
Antall karakterer i alt	1 422 038	98,8 %	7,69	6	6
Antall standpunkt	1 178 830	96,9 %	6,37	5	6
Antall skriftlig eksamen	190 438	59,5 %	1,03	0	1
Antall muntlig eksamen	52 770	27,0 %	0,27	0	0
Antall ulike registrerte fag	979		6,67	5	6

Det samlede materialet inneholder elever på alle nivåer i videregående opplæring. Dette er forklaringen på at det er relativt få eksamenskarakterer pr elev. Sluttvurderingen gis oftest i form av en standpunktkarakter for fag på de lavere nivåene. Hoveddelen av elevene har vurdering i minst et fag (98,8 %), det kun er en liten gruppe elever (1,2 %) som ikke har vurdering i noe fag.

2.2 Vårt utvalg av fag

Vi bruker bare en liten del av det samlede datamaterialet i våre analyser, og det er flere grunner til dette. For det første er det bare relevant å gjøre analyser der vi har både standpunktkarakter og eksamenskarakter for de samme elevene, siden hensikten med prosjektet er å måle forskjell mellom standpunktkarakter og eksamenskarakter. For det andre må vi avgrense materialet slik at vi får store nok elevgrupper til å kunne sammenligne elever i private og offentlige skoler. For det tredje har vi valgt å begrense analysene til avsluttende eksamener på Vg3. Ut fra disse kriteriene har vi valgt ut i alt ni fag, fem av disse har skriftlig eksamen mens fire har muntlig. I praksis vil de kriteriene vi har benyttet begrense utvalget til å gjelde elever som tar sikte på studiekompetanse.

Tabell 2.2 Karakterdata for ni utvalgte fag i fire skoleår. Avsluttende vurdering. Antall elever

	Vår 2010	Vår 2011	Vår 2012	Vår 2013
NOR1211 + 1231 Norsk hovedmål	38126	39816	39239	40444
NOR1212 + 1232 Norsk sidemål	18270	18147	16919	18132
REA3024 Realfagsmatematikk	3579	3353	3149	3758
MAT1005 Matematikk påbygg	4087	4791	4711	4520
SAM3023 Rettslære	2433	2369	2415	2196
NOR1213 + 1233 Norsk muntlig	6287	7948	7345	8041
REL1001 Religion og etikk	3535	4158	3915	4310
HIS1002 Historie studiespesialisering.	4385	4562	4586	5048
HIS1003 Historie påbygg	1906	2292	2148	2336

Tabell 2.2 viser at antallet elever som har vært oppe til eksamen hvert år varierer fra 1900 i historie for påbyggselever i 2010 til 40.444 for norsk skriftlig hovedmål i 2013. Det er kun små variasjoner i antallet elever fra år til år innenfor hvert enkelt fag.

Tabell 2.3 gjengir noen nøkkeltall for det materialet vi har brukt i analysene når vi har gjort avgrensingen til ni fag der det foreligger både standpunktkarakter og eksamenskarakter i minst ett fag for hver enkelt elev.

Tabell 2.3 Karakterdata for ni utvalgte fag i fire skoleår. Avsluttende vurdering.

	Antall	Andel med karakter	Gj. snitt pr elev	Modus = Vanligste antall	Median = Midtpunkt
Antall elever	172 852				
Antall karakterer i alt	1 037 945	100,0 %	6,00	5	6
Antall standpunkt	687 518	96,9 %	3,98	3	4
Antall skriftlig eksamen	277 082	96,0 %	1,60	2	2
Antall muntlig eksamen	73 345	42,0 %	0,42	0	0

Datamaterialet omfatter i alt 172.852 elever¹, eller i gjennomsnitt 43 000 for hvert år. Disse har til sammen over en million karakterer fra det året vi ser på, eller i gjennomsnitt seks karakterer for hver elev. Av disse seks er i gjennomsnitt fire standpunktkarakterer og to eksamenskarakterer. Det er fire ganger så mange skriftlige eksamenskarakterer som muntlige. Når vi sammenligner tabell 2.2 og 2.3, kan vi regne oss fram til at det foreligger karakterdata i norsk skriftlig hovedmål for omtrent 93 prosent av elevene, mens bare 5 prosent har karakterdata for historie påbygg.

I tabell 2.4 summerer vi opp noen nøkkeltall for hvert av de ni fagene vi ser på. Det er 57 prosent jenter i materialet og et klart jenteflertall i alle fag unntatt realfagsmatematikk der det er 61 prosent gutter. De aller fleste fyller 19 år det året vi har sett på, men påbyggsfagene har litt høyere gjennomsnittsalder. Dette er helt i tråd med det vi kan forvente siden det blant påbyggselevne vil være noen som først har oppnådd yrkeskompetanse.

¹ 5149 av disse opptre i to ulike år, mens 47 er registrert tre ganger. Antallet individer er dermed 167.539.

Tabell 2.4 Karakterdata for ni utvalgte fag i fire skoleår. Avsluttende vurdering. Antall elever

	Jenter	Alder	Privat	Standpunkt	Eksamen	Antall
NOR1211 + 1231 Norsk hovedmål	57 %	19,36	9 %	3,68	3,15	157 625
NOR1212 + 1232 Norsk sidemål	59 %	19,27	8 %	3,42	3,03	71 468
REA3024 Realfagsmatematikk	39 %	19,09	9 %	3,94	3,58	13 839
MAT1005 Matematikk påbygg	59 %	19,96	10 %	3,31	2,62	18 109
SAM3023 Rettslære	63 %	19,08	14 %	4,01	3,51	9 413
NOR1213 + 1233 Norsk muntlig	57 %	19,41	9 %	4,08	4,11	29 621
REL1001 Religion og etikk	56 %	19,12	10 %	4,13	4,34	15 918
HIS1002 Historie studiespesialisering	56 %	19,13	8 %	4,12	4,24	18 581
HIS1003 Historie påbygg	60 %	20,06	8 %	3,61	3,69	8 682
Alle ni fag	57 %	19,48	8 %	3,76	3,26	172 852

Det er i faget rettslære vi finner høyest andel ved private skoler, 14 prosent, mens andelen ligger mellom 8 og 10 prosent for de andre fagene. De høyeste karakterene finner vi for religion og etikk, mens de laveste er i påbyggsmatematikk. Dette gjelder både standpunktkarakterene og eksamenskarakterene.

2.3 Kjennetegn ved offentlige og private skoler

Vårt oppdrag er å studere forskjeller i karakterer ved private og offentlige videregående skoler, og som nevnt har vi valgt å operasjonalisere dette på to forskjellige måter. I våre analyser kommer vi til enten å bruke hovedskillet mellom private og offentlige skoler, eller benytte den grove inndelingen i offentlige skoler og tre typer private.

Tabell 2.5 Kjennetegn ved avgangselever ved offentlige og private skoler. Gjennomsnittstall for fire år (2010 – 2013)

	Offentlige	Private kristne	Private allmenne	Private spesialiserte	Alle
Alder	19,5	19,9	19,3	19,4	19,5
Andel jenter	57 %	60 %	59 %	46 %	57 %
Andel i studiespesialisering	53 %	65 %	59 %	68 %	54 %
Andel i påbygg for stud.komp.	25 %	25 %	27 %	11 %	24 %
Gjennomsnittskarakter	3,55	3,86	3,62	3,62	3,56
Gj.snittskarakter standpunkt	3,75	4,07	3,85	3,79	3,76
Gj.snittskarakter eksamen	3,26	3,45	3,20	3,27	3,26

Tabell 2.5 viser forskjeller mellom offentlige skoler og de tre typene private skoler. Det er kun små forskjeller i gjennomsnittsalder, og ved de offentlig, private kristne og private allmenne skoler er jenter i flertall. De private spesialiserte skolene derimot har et flertall av gutter blant elevene. Ved alle skoletypene er det over halvparten av elevene som er registrert på studiespesialisering og denne andelen er høyest på de private spesialiserte skolene, noe som skyldes at elever på idrettsgymnasene som regel er registrert på studiespesialisering. Det er også forskjell mellom skoler i forhold til hvor stor andel av elevene som er registrert på påbygg for studiekompetanse. Ved offentlige, private kristne og private allmenne skoler er denne andelen omtrent en av fire, mens det bare er 11 prosent av elevene på de private spesialiserte skolene som er registrert på påbygg til studiekompetanse.

Det er små variasjoner i gjennomsnittskarakterer mellom de ulike skoleslagene, men de private kristne skolene har høyest gjennomsnitt generelt og også høyest gjennomsnittskarakter for standpunkt og eksamen.

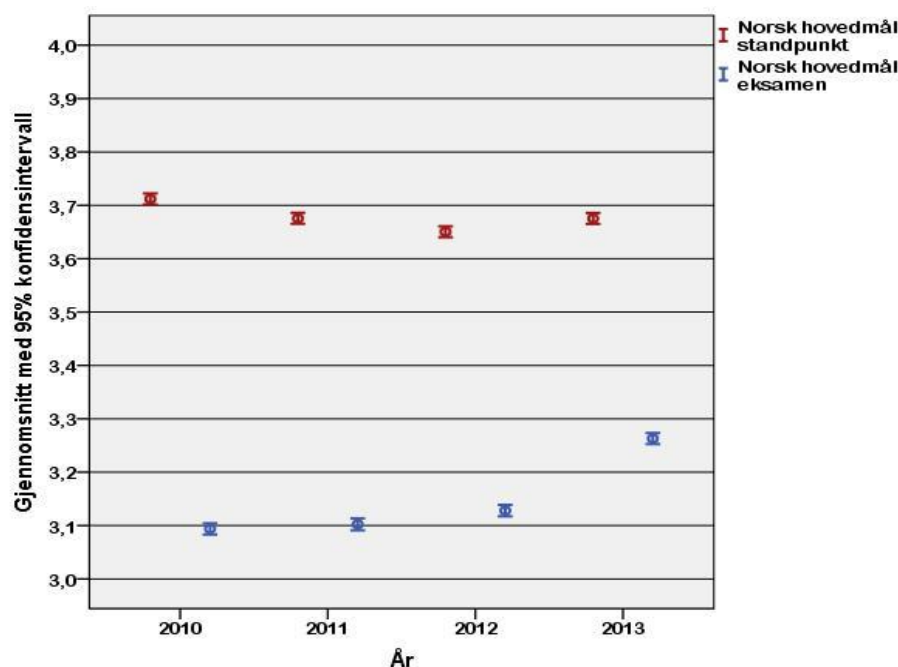
3 Karakterforskjeller i norskfaget

Dette kapittelet tar for seg karakterene i norsk skriftlig hovedmål, skriftlig sidemål og norsk muntlig. Disse er de tre eksamensfagene der vi har størst sammenligningsgrunnlag, siden alle kommer opp i eksamen i norsk hovedmål, og det i tillegg er mange som blir trukket til eksamen i sidemål eller muntlig.

3.1 Norsk skriftlig hovedmål

3.1.1 Hovedresultatene

Det var resultatene i norsk skriftlig hovedmål som fikk fokus i media vinteren 2013 og som er bakgrunnen for at dette prosjektet ble igangsatt (se for eksempel Aftenposten 05.02.2013, 07.02.2013, 15.02.2013, Stavanger Aftenblad 19.02.2013). Vi skal derfor først presentere hovedtallene for dette faget i de fire skoleårene. Figur 3.1 viser gjennomsnittlige standpunkts- og eksamenskarakterer for de fire skoleårene med statistisk feilmargin.



Figur 3.1 Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk skriftlig hovedmål i fire skoleår.

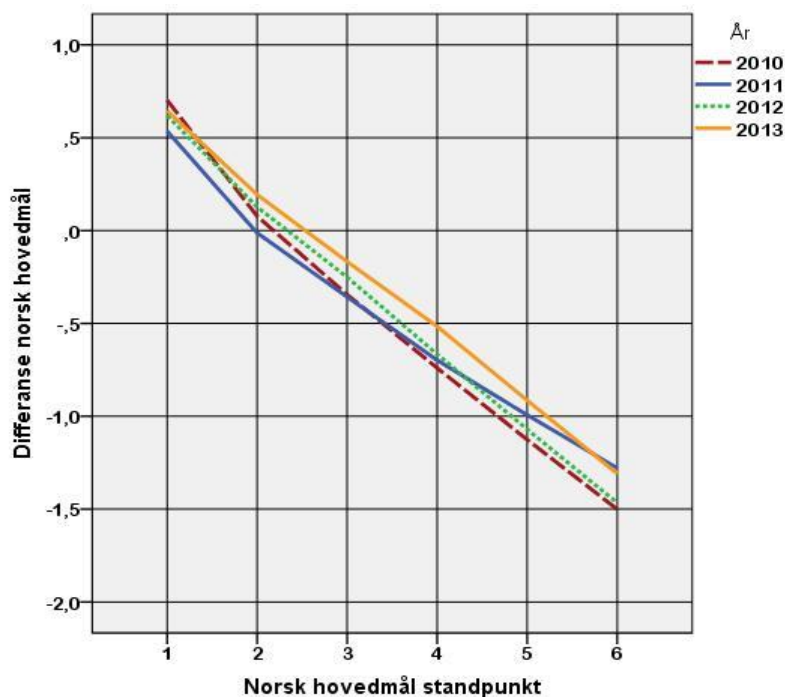
Standpunktkarakterene ligger betydelig over eksamenskarakterene, men forskjellen har minsket over tid, fra 0,62 karakterpoeng i 2010 til 0,42 i 2013. Endringen skyldes først og fremst en betydelig økning i eksamenskarakteren fra 2012 til 2013, med 0,13 karakterpoeng, men også at standpunktkarakteren sank med 0,06 karakterpoeng fra 2010 til 2012. Endringen fra 2010 til 2013 er imidlertid bare -0,03 karakterpoeng for standpunktkarakteren, mens den er på 0,17 karakterpoeng for eksamenskarakteren. Konklusjonen er dermed at det har vært en betydelig større stabilitet i standpunktkarakteren enn i eksamenskarakteren, med 2013 som et særlig avvikende år. Tallene bak gjennomsnittstallene i figuren viser at andelen med eksamenskarakterene 1 og 2 sank med 5,5 prosentpoeng fra 2010 til 2013, mens andelen med karakterene 4, 5 og 6 økte med 6,4 prosentpoeng (ikke vist i egen tabell).

Vi har ingen holdepunkter for å angi årsaker til utviklingen i eksamenskarakteren over tid, men vil vise til at eksamenskarakterer ikke er uttrykk for objektive standarder og at det alltid vil være noe tilfeldig variasjon knyttet til slik vurdering. To særlig viktige faktorer spiller inn, nemlig utformingen av og

endringer i eksamensoppgavene og det skjønne sensorene utøver, og kanskje ikke minst en kombinasjon av disse to faktorene. Det normale er uansett at gjennomsnittlig eksamenskarakter i norsk skriftlig hovedmål er svakere enn gjennomsnittlig standpunktkarakter. For de fire årene sett under ett er denne forskjellen i gjennomsnitt -0,53 karakterpoeng. Sett fra elevenes side betyr dette at halvparten av dem har en eksamenskarakter som er dårligere enn standpunktkarakteren. Denne andelen har sunket fra 54,7 prosent i 2010 til 46,1 prosent i 2013. I gjennomsnitt har 13 prosent en eksamenskarakter som er bedre enn standpunktkarakteren og denne andelen har økt fra 10,6 til 14,9 prosent over tid. Dette funnet vil bli videre fulgt opp i den kvalitative delen av prosjektet.

3.1.2 Karakterdifferanse kontrollert for standpunktkarakter

Dette generelle mønsteret med svakere eksamenskarakter enn standpunktkarakter gjelder imidlertid ikke når vi ser på forskjellen mellom de to karakterene avhengig av hvilken standpunktkarakter elevene har i utgangspunktet, som vises i figur 3.2.



Figur 3.2 Gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål i fire skoleår etter standpunktkarakter.

Når det generelle mønsteret er en negativ differanse, vil sannsynligheten for dette være større for de elevene som har de beste karakterene, mens den er minst for de med de dårligste. De som har standpunktkarakteren 1 kan ikke få en dårligere karakter enn dette til eksamen, mens de som har 6 ikke kan få en bedre karakter. Figur 3.2 viser denne sammenhengen og det skilles her mellom de fire årene vi studerer.

Vi gjenkjenner mønsteret fra figur 3.1, men her fortelles historien på en annen måte. Den oransje heltrukne linja i diagrammet ligger over de øvrige tre linjene, med unntak for standpunktkarakterene 1 og 6. Denne linja gjelder 2013 og viser at de som hadde standpunktkarakter mellom 2 og 5 dette året fikk en eksamenskarakter som avvek mindre fra standpunktkarakteren enn det som gjelder for de tre øvrige årene.

Den eneste gruppa som har klart bedre karakter til eksamen enn standpunktkarakteren er de med karakteren 1 i standpunkt. Forskjellen er i gjennomsnitt litt mer enn et halvt karakterpoeng og det betyr at omtrent halvparten av dem står til eksamen, mens den andre halvparten stryker. Tre fjerdedeler av de som har 1 og som står får karakteren 2. I motsatt ende av karakterskalaen ser vi at de med

standpunktkarakteren 6 har en negativ differanse på mellom 1,25 og 1,50 karakterpoeng. Det betyr at bare 15 prosent av dem får karakteren 6 til eksamen, mens resten går tilbake. 88 prosent av de som ikke beholder karakteren 6, får 5 eller 4, litt flere femmere enn firere. Mens de som har 1 og de som har 6 kan oppleve endring i bare én retning, kan de øvrige elevene få både bedre og dårligere karakterer til eksamen sammenlignet med standpunktkarakteren. Sannsynligheten for å oppleve framgang eller tilbakegang er imidlertid høyst forskjellig. Tabell 3.1 viser dette.

Tabell 3.1 Forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål. Gjennomsnittstall for fire år (2010 – 2013).

		Norsk hovedmål standpunkt						
		1	2	3	4	5	6	Alle
Endring norsk	Negativ		22,5%	40,3%	57,3%	71,7%	85,3%	51,0%
skriftlig	Uendret	50,5%	49,6%	43,0%	33,7%	24,8%	14,7%	36,3%
hovedmål	Positiv	49,5%	27,9%	16,8%	9,0%	3,6%		12,7%
Sum		100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

For dem med standpunktkarakteren 2 er det mest sannsynlig å beholde denne karakteren og halvparten gjør det. For de øvrige toerne er sannsynligheten for en bedre karakter 24 prosent høyere enn en dårligere karakter, som i praksis betyr at de stryker. 43 prosent av treerne beholder denne karakteren, mens sannsynligheten for en dårligere karakter enn 3 er 2,5 ganger så høy som sannsynligheten for å få en bedre karakter. En av tre med karakteren 4 beholder denne, mens sannsynligheten for tilbakegang er 6,4 ganger så høy som sannsynligheten for framgang. En av fire med karakteren 5 beholder denne. Her er sannsynligheten for en dårligere karakter enn 5 tjue ganger så høy som sannsynligheten for å få en sekser.

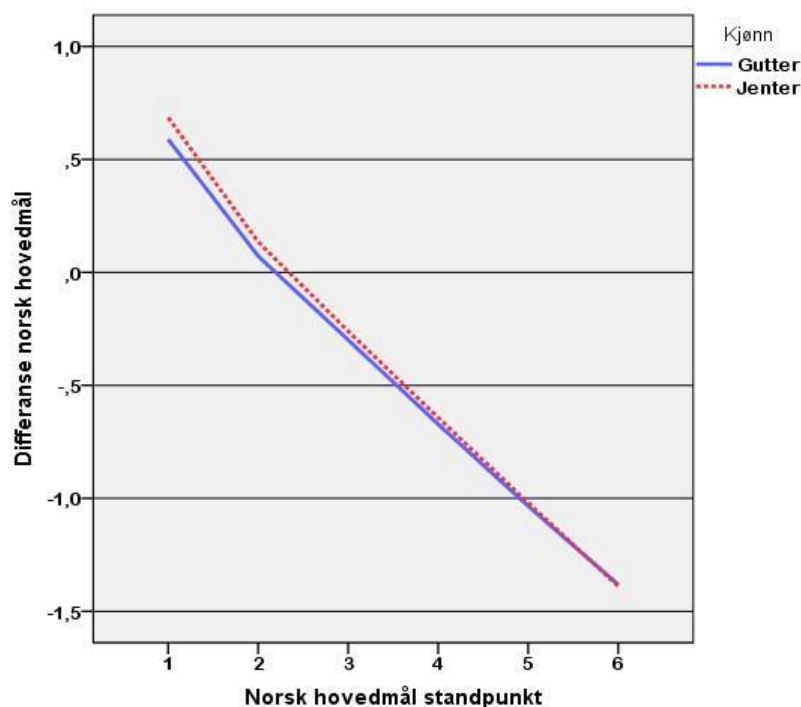
Tabell 3.1 og figur 3.2 uttrykker det som kalles regresjon mot gjennomsnittet (se for eksempel Galton 1886, Grøgaard 2012). Når det er en statistisk samvariasjon mellom to forhold, slik det er mellom standpunktkarakter og eksamenskarakterer i norsk skriftlig hovedmål (Pearsons $r = 0,591$), og en av disse karakterene avviker fra gjennomsnittet, så vil det være størst sannsynlighet for at den andre karakteren ligger nærmere gjennomsnittet enn den første. Når gjennomsnittlig standpunktkarakter er 3,7 og gjennomsnittlig eksamenskarakter er 3,2, vil en elev med standpunktkarakteren 5 ha større sannsynlighet for å få en firer enn en femmer til eksamen.

Det er viktig å ta høyde for dette fenomenet i sammenligning av gjennomsnittsverdier, i vårt tilfelle spesielt når vi sammenligner grupper av elever med ulikt karaktergjennomsnitt, for eksempel jenter og gutter eller skoler med høye inntakskrav og skoler som også tar inn elever med svakere resultater. Tabell 3.2 viser resultatene for jenter og gutter. Mens differansen for jenter er 0,58 karakterpoeng er den 0,47 for guttene. Dermed utlignes 30 prosent av den forskjellen vi ser mellom gutter og jenter fra standpunkt til eksamen. Er dette fordi guttene gjør det bedre enn jentene til eksamen? Dette kan vi undersøke dersom vi kontrollerer for hvilken standpunktkarakter de har. Det viser vi i figur 3.3.

Tabell 3.2 Eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål for gutter og jenter. Gjennomsnittstall for fire år (2010 – 2013).

Kjønn		Norsk hovedmål	Norsk hovedmål	Differanse norsk
		standpunkt	eksamen	hovedmål
Gutter	Gjennomsnitt	3,47	3,00	-0,47
	Antall (N=)	67948	67948	67948
	Standardavvik	1,041	1,094	,962
Jenter	Gjennomsnitt	3,83	3,26	-0,58
	Antall (N=)	89676	89676	89676
	Standardavvik	1,001	1,072	,958
Alle	Gjennomsnitt	3,68	3,15	-0,53
	Antall (N=)	157624	157624	157624
	Standardavvik	1,034	1,089	,961

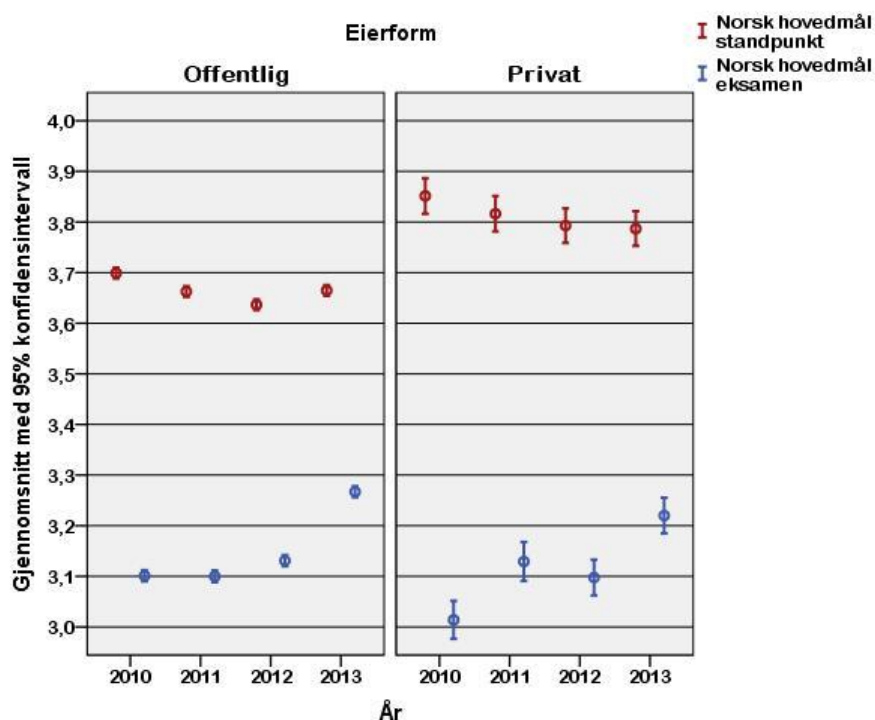
Figur 3.3 viser at gutter og jenter gjør det omtrent like godt til eksamen når vi kontrollerer for standpunktkarakteren de har. I den grad det er noen forskjell, så er det jentene som gjør det relativt sett best, spesielt når utgangspunktet er som dårligst. Jenter med 1 i standpunkt går fram med 0,69 karakterpoeng mot guttenes 0,59. Sagt på en annen måte: 55 prosent av disse jentene får bestått til eksamen mot 47 prosent av guttene. Så avtar forskjellen mellom gutter og jenter gradvis inntil den forsvinner helt for dem med standpunktkarakteren 5 og 6.



Figur 3.3 Gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål for gutter og jenter etter standpunktkarakter.

3.1.3 Hovedforskjell mellom offentlige og private skoler

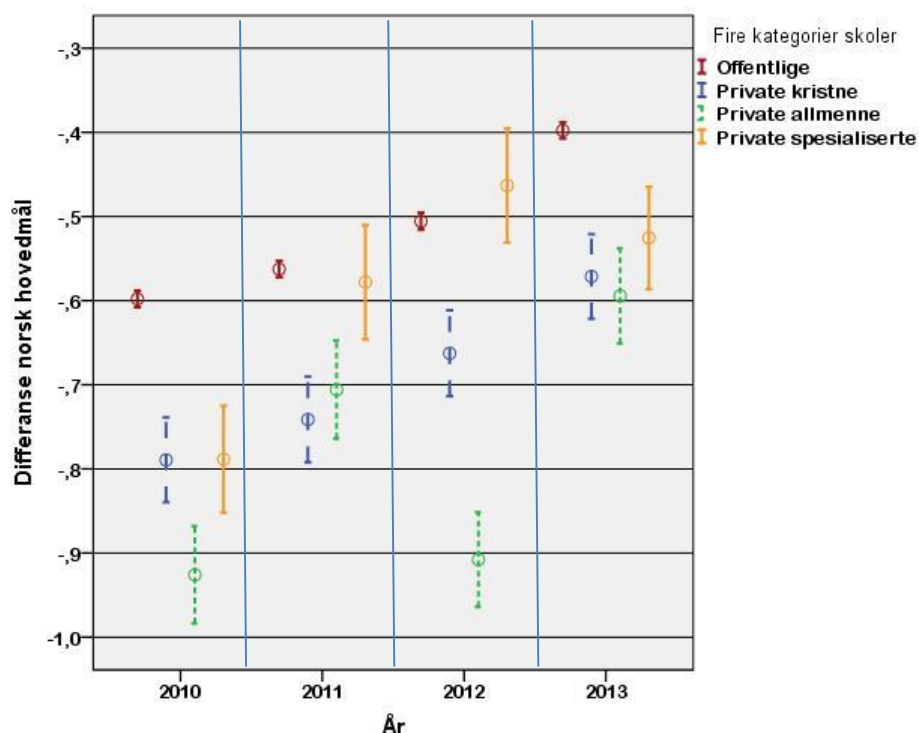
Vi skal nå se på hvilke forskjeller det er i standpunktkarakterer og eksamenskarakterer mellom offentlige og private skoler. Figur 3.4 viser at det er tydelige forskjeller mellom skoleslagene, samtidig som vi også kan få øye på noen endringstrekk over tid. Standpunktkarakterene er noe høyere ved de private skolene enn ved de offentlige, og forskjellen er i gjennomsnitt 0,14 karakterpoeng, varierende mellom 0,16 og 0,13 karakterpoeng. Forskjellene er mye mindre for eksamenskarakterene, samtidig som bildet framstår som noe mer uklart. I tre av fire år er forskjellen i avvik størst hos de offentlige skolene og mellom 0,03 og 0,09 karakterpoeng. Unntaket gjelder 2011 da de private skolene oppnådde et gjennomsnitt som var 0,03 karakterpoeng bedre enn de offentlige.



Figur 3.4 Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk skriftlig hovedmål i fire skoleår for offentlige og private skoler.

Hovedinntrykket er at forskjellen mellom eksamenskarakterer og standpunktkarakterer er større ved private skoler enn ved offentlige. Dette utdypes videre i figur 3.5 der vi dessuten skiller mellom tre ulike former for private skoler. De røde, korte stolpene viser utviklingen for de offentlige skolene, som stort sett tilsvarer bildet for alle skoler, ettersom 92 prosent av elevene befinner seg her. Forskjellen mellom eksamenskarakter og standpunktkarakter er redusert fra ca. 0,6 til 0,4 prosentpoeng over tid. De private kristne skolene (blå brutte streker) og de private spesialiserte (oransje) følger stort sett samme utvikling, mens de private allmenne skolene (stiplet grønn) viser et mer uklart mønster. Her ble forskjellen kraftig redusert fra 2010 til 2011, for så å øke nesten like mye fra 2011 til 2012. Fra 2012 til 2013 foregår det så igjen en kraftig reduksjon i forskjellen mellom standpunktkarakter og eksamenskarakter ved de private allmenne skolene. Hovedbildet er likevel at det fortsatt er en større forskjell mellom de to karakterene ved private enn ved offentlige skoler, men at forskjellen er redusert over tid.

Siden det var 2012-tallene som fikk offentlig oppmerksomhet og utløste dette prosjektet, kan det være grunn til å knytte en ekstra kommentar til disse. Vi ser av figur 3.5 at det dette året ikke var noen forskjell av betydning mellom offentlige skoler og de private spesialiserte, mens det var noe større forskjeller mellom offentlige skoler og de to øvrige private skoleslagene. Særlig gjaldt dette de allmenne private skolene. Det er også for disse skolene vi finner en kraftig endring fra 2012 til 2013.



Figur 3.5 Gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål i fire skoleår for offentlige og tre typer private skoler.

3.1.4 Ulike elevgruppers resultater

Vi har allerede vist at det er forskjeller mellom elever ved offentlige og private skoler når det gjelder karakterer i norsk skriftlig hovedmål og at det er en endring over tid i nivået for eksamenskarakterene. Dessuten har vi vist til forskjellen mellom jenter og gutter. Vi vil også se på om det er andre bakgrunnsegenskaper ved elevene som samvarierer med karakterene før vi gjør en multivariat analyse av hva det er som har effekt på forskjellen mellom standpunktkarakter og eksamenskarakter. I egne vedleggstabeller viser vi eksamenskarakterer og standpunktkarakterer etter kjønn, alder, hvilket utdanningsprogram eller kompetansetype eleven tar sikte på, skoletype, antall avgangselever ved skolen og år. Vi skiller også mellom det vi kaller skoler med høy inntakskvalitet der elevenes gjennomsnittskarakterer fra grunnskolen er særlig høye, og andre skoler.

Ved siden av de tre forholdene vi allerede har beskrevet, ser vi tydelige forskjeller mellom elever fra de ulike utdanningsprogrammene. Realfagselevne ligger et stykke foran de øvrige både når det gjelder standpunkt- og eksamenskarakter, mens differansen mellom standpunkt og eksamen er 0,54 og marginalt forskjellig fra gjennomsnittet på 0,53. Det er elevene på Påbygging for studiekompetanse (gruppen som går over til studiekompetanse etter to år i yrkesfag) som gjennomsnittlig oppnår de svakeste resultatene. Størst gjennomsnittsforskjell mellom standpunkt og eksamenskarakter finner vi blant elevene på Naturbruk. I tillegg er det store forskjeller i strykprosent, i gjennomsnitt strøk 14 respektive 13 prosent av elevene på Naturbruk og Påbygging til studiekompetanse til norsk skriftlig hovedmål-eksamen alle de fire årene, mens dette bare gjaldt to prosent av realfagselevne (ikke vist i tabell).

Elever som er 20 år og eldre oppnår svakere resultater enn 19-åringene. Dette er ikke overraskende ettersom vi her vil finne en del som har tatt et ekstra år for å forbedre resultatene. Vi har dessverre ikke hatt ressurser til å gå nærmere inn i denne problematikken i dette prosjektet. Det er særlig 20 og 21-åringene som oppnår svakere resultater enn 19-åringene, mens dette i noe mindre grad gjelder de som er 22 år og eldre. De eldste har på sin side den største differansen mellom standpunktkarakter og eksamenskarakter.

Det er en sammenheng mellom skolestørrelse og resultater. Eksamenskarakteren er svakere for elever ved skoler med færre enn 50 avgangselever enn ved større skoler og også differansen mellom de to karakterene er størst her. Mens differansen mellom karakteren er 0,68 karakterpoeng ved de minste skolene, er den 0,44 ved de største.

Den siste variabelen vi skal ta med i analysene er også en kontekstvariabel. Vi skiller her mellom skoler som har et karaktergjennomsnitt ved inntak fra grunnskolen som ligger godt over gjennomsnittet og resten av skolene. Den sjettedelen av skolene som har høyest inntakskarakterer fra grunnskolen defineres som skoler med høy inntakskvalitet, og sammenlignes med resten som har normal inntakskvalitet. Drøyt 20 prosent av elevene går ved skoler med høy inntakskvalitet.² Siden skoler med høy inntakskvalitet har et høyere karaktergjennomsnitt, skulle vi forvente at sannsynligheten for en negativ differanse mellom eksamenskarakter og standpunktkarakter ville være større enn på andre skoler. Tabell 3.1 viser jo at jo bedre standpunkt karakter eleven har desto større er sjansen for at eksamenskarakteren blir lavere enn standpunktkarakter. Men dette er ikke tilfelle for skoler med høy inntakskvalitet. Dette funnet vil vi prøve å undersøke videre i den kvalitative delen av studien.

3.1.5 Regresjonsanalyse: Hva har effekt på differansen mellom karakterene?

Vi skal nå gjennomføre en regresjonsanalyse der forskjellen mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål er avhengig variabel og der de sju forholdene vi til nå har trukket inn i analysene er uavhengige variabler. Dette gjelder kjønn, alder, programområde, skoletype, skolestørrelse, skole med høy inntakskvalitet og år. I tillegg tar vi med en sjuende variabel, nemlig standpunktkarakteren. Vi så allerede i figurene 3.2 og 3.3 at det er en sterk lineær sammenheng mellom hvilken standpunktkarakter elevene har og hvor mye eksamenskarakteren skiller seg fra standpunktkarakteren. Tabell 3.3 viser at standpunktkarakteren har en koeffisient på $-0,437$. Dette betyr at for hvert karakterpoeng standpunktkarakteren øker, vil eksamenskarakteren reduseres med i gjennomsnitt 0,437 karakterpoeng sammenlignet med standpunktkarakteren. Dette stemmer godt med bildet de to figurene (3.2 og 3.3) viste, også etter at vi har kontrollert for effekten av en rekke andre forhold. Ingen andre forhold har en tilnærmet like sterk effekt som standpunktkarakteren. Modellen forklarer riktignok bare drøyt 20 prosent av variasjonen i forskjellen mellom de to karakterene, men 16,5 prosent av dette er det standpunktkarakteren som står for.

² Vi har basert oss på opplysninger om standpunktkarakterer fra grunnskolen for avgangskullene 2004 til 2007. Data er hentet fra evalueringen av Kunnskapsløftet.

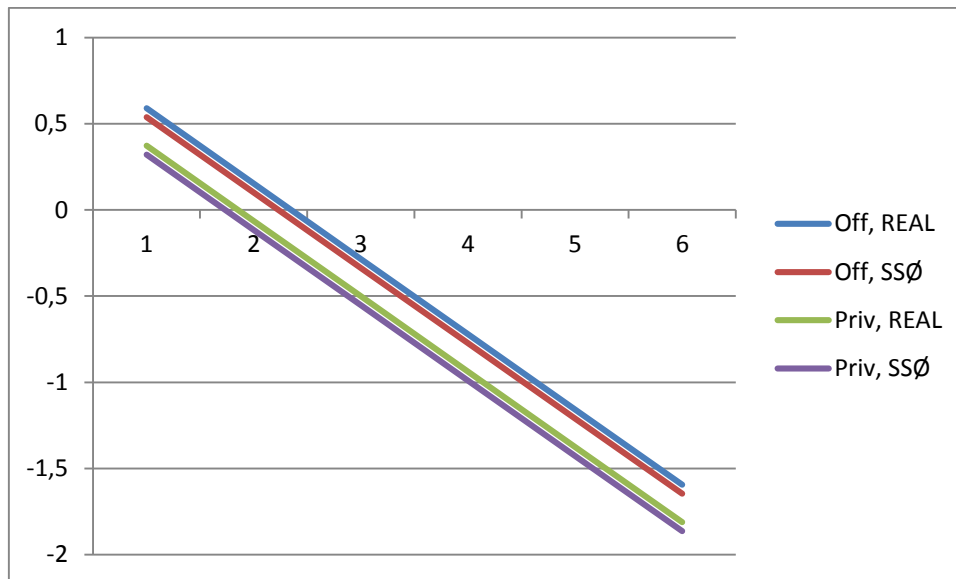
Tabell 3.3 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål som avhengig variabel. N=157 624.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign. nivå.
	B	Std. feil	Beta	t-verdi	
Konstantledd	0,899	0,033		27,036	0,000
<u>Norsk hovedmål standpunkt</u>	-0,437	0,002	-0,470	-194,896	0,000*
<u>Jente</u>	0,068	0,005	0,035	15,036	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,050	0,009	-0,013	-5,606	0,000*
Alder 21 år	-0,081	0,014	-0,014	-5,761	0,000*
Alder 22 år pluss	-0,064	0,013	-0,013	-4,849	0,000*
<u>Programområde: Ref = Annet</u>					
Realfag	0,128	0,031	0,055	4,120	0,000*
Språk, samf.fag, økonomi	0,076	0,031	0,037	2,450	0,014*
Formgivingsfag	-0,078	0,034	-0,012	-2,293	0,022*
Musikk	0,101	0,033	0,017	3,013	0,003*
Dans	-0,072	0,040	-0,006	-1,789	0,074*
Drama	0,059	0,037	0,006	1,606	0,108
Idrett	-0,130	0,032	-0,038	-4,107	0,000*
Media	-0,021	0,032	-0,005	-0,656	0,512
Naturbruk	-0,342	0,038	-0,035	-8,990	0,000*
Påbygg uten yrkeskompetanse	-0,295	0,031	-0,123	-9,556	0,000*
Påbygg etter yrkeskompetanse	-0,204	0,033	-0,037	-6,181	0,000*
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,083	0,012	-0,016	-6,718	0,000*
Privat Allmenn	-0,217	0,013	-0,038	-16,802	0,000*
Privat Spesialisert	-0,043	0,015	-0,007	-2,799	0,005*
<u>Skole med høy inntakskvalitet</u>	0,153	0,006	0,065	27,387	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,062	0,011	0,023	5,910	0,000*
Elevtall 100 til 199	0,082	0,010	0,043	8,385	0,000*
Elevtall 200 til 299	0,066	0,010	0,029	6,303	0,000*
Elevtall 300 og mer	0,118	0,012	0,033	9,644	0,000*
<u>År: Referansekategori = 2010.</u>					
År 2011	0,036	0,006	0,016	5,868	0,000*
År 2012	0,078	0,006	0,035	12,641	0,000*
År 2013	0,193	0,006	0,088	31,442	0,000*

Andel forklart varians: 21 %. * : Statistisk signifikant på 5 % nivå eller lavere.

Det neste viktigste forholdet målt i B-verdi er å være elev på Naturbruk, med -0,342. Også de to påbyggvariantene har relativt store negative koeffisienter. Når vi sammenligner de ulike programområdene, kan vi regne oss fram til at elever på Realfag har en B-verdi som er 0,47 større enn elever på Naturbruk. Det betyr at forskjellen mellom de to programområde tilsvarer 0,47 karakterpoeng når alt ellers er likt.

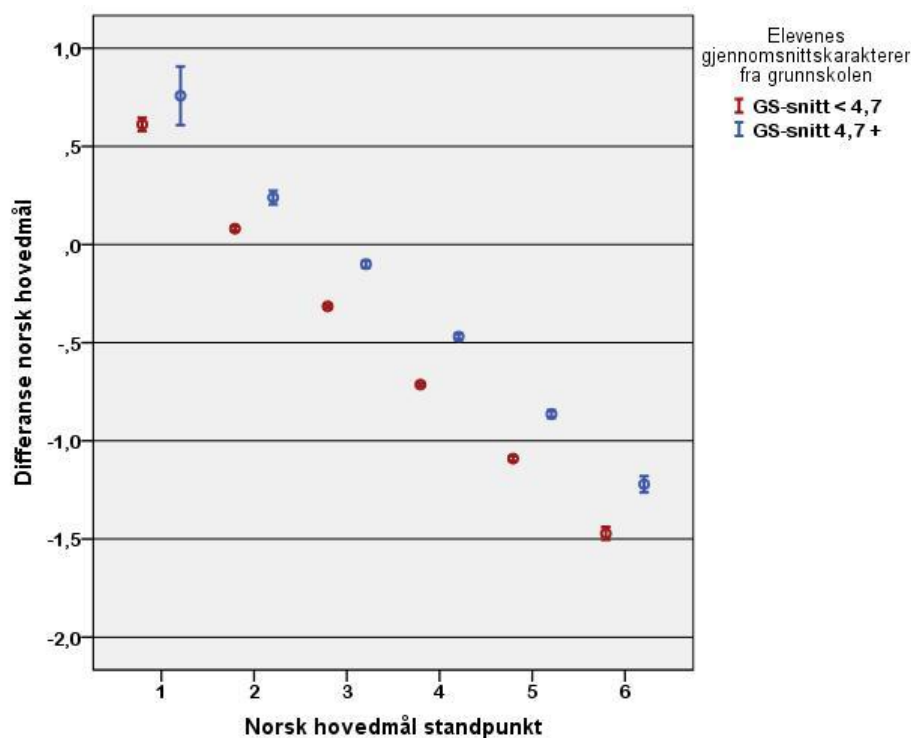
Viktigst etter standpunktkarakter og programområde er skoletype, og her er det de allmenne private skolene som skiller seg ut, selv om de to andre private skoletypene også avviker i forhold til referansekategori, som er offentlige skoler. Alle tre private skoletyper har negative koeffisienter, som betyr at forskjellen mellom eksamenskarakter og standpunktkarakter i gjennomsnitt er større på alle typer private skoler sammenlignet med offentlige, men den er størst på allmenne private skoler. Dersom vi sammenligner to av de største programområdene: språk, samfunnsfag og økonomi med realfag på offentlige skoler og allmenne private skoler ser vi at det er mye større effekt av hvilken type skole man går på, enn hvilket program man går på (figur 3.6).



Figur 3.6 Beregnet gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål, etter type studieprogram og om eleven går på en offentlig eller en allmenn privat skole. Basert på regresjonsanalysen i tabell 3.3.

Alder, skolestørrelse og årstall har også betydning, men utslagene er relativt beskjedne. Når de eldre elevene avviker fra 19-åringene, kan dette henge sammen med at de kan ha tatt et ekstra år eller tatt opp igjen enkeltfag, men vi har ikke hatt anledning til å undersøke dette nærmere. Det ser ellers ut til å være en lineær sammenheng mellom elevtall og differanse mellom standpunktkarakter og eksamenskarakter. 2013 skiller seg klart fra de andre årene, selv om også 2011 og 2012 avviker fra referanseåret 2010.

Den siste faktoren av særlig betydning er det vi kaller skole med høy inntakskvalitet, eller nivået på grunnskolekarakterene elevene hadde med seg til den videregående skolen, her målt som et gjennomsnitt for hele elevgruppa. Elever på slike skoler har en mindre differanse mellom eksamenskarakter og standpunktkarakter, til tross for at de i gjennomsnitt har bedre karakterer og dermed skulle forventes å ha en større differanse. Gjelder dette løftet, som det å være elev ved en skole med høy inntakskvalitet gir, for både svake og sterke elever?



Figur 3.7 Gjennomsnittlig forskjell mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål ved skoler med høye gjennomsnittskarakterer fra grunnskolen og andre skoler.

Figur 3.7 viser at det synes å være en fordel å være elev ved en skole med særlig mange faglig dyktige elever sammenlignet med andre skoler, uansett hvor gode karakterer den enkelte elev selv har. På skoler med høy inntakskvalitet er det av naturlige grunner bare en håndfull elever som har 1 som standpunktkarakter, så denne verdien kan vi se bort fra i sammenligningen. For alle andre standpunktkarakterer ser vi et tydelig mønster ved at forskjellen mellom eksamenskarakter og standpunktkarakter er betydelig mindre ved skoler med høy inntakskvalitet enn ved andre skoler. Hvordan forklarer vi dette? Setter lærerne ved disse skolene standpunktkarakterer som gir et riktigere bilde av elevenes kunnskapsnivå eller setter de karakterer som egentlig er for lave for å nærme seg en normalfordeling? Forberedes disse elevene bedre på eksamen? Eller har sensorene spesielle forventninger til elever fra disse skolene og lar dette påvirke eksamenskarakteren? Siden sensorene vet hvilke skoler elevene kommer fra, kan vi ikke utelukke den siste forklaringen.

Et annet viktig spørsmål er om dette gjelder alle skoler med et høyt nivå for grunnskolekarakterene. Når vi går nærmere inn i dette, finner vi at det er bare 13 av de 67 skolene vi har valgt å kalle skoler med høy inntakskvalitet som har en særlig liten forskjell mellom eksamenskarakter og standpunktkarakter. Disse skolene forklarer dermed det meste av forskjellen vi finner mellom skole med høy inntakskvalitet og skoler med normal inntakskvalitet. Ingen av disse skolene er private. I den andre enden av fordelingen finner vi fire skoler med høy inntakskvalitet som avviker i den andre retningen og har en særlig stor forskjell mellom de to karakterene og av disse er tre private skoler.

3.1.6 Foreløpig konklusjon

Vårt hovedanliggende har vært å dokumentere mulige forskjeller mellom private og offentlige skoler når det gjelder skolens gjennomsnittlige standpunkt- og eksamenskarakter for ulike grupper av elever. Forskjellene som ble påvist i bivariate analyser består når vi bruker regresjonsanalyse, selv om de er noe svekket. Tabell 3.4 viser de opprinnelige forskjellene mellom offentlige og tre typer private skoler og forskjellene vi finner når vi har kontrollert for standpunktkarakter, kjønn, alder, programområde, skolestørrelse, effekten av å gå på en skole med høy inntakskvalitet og år.

Tabell 3.4 Forskjell i differanse mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål når tre typer private skoler sammenlignes med offentlige skoler. N=157 623.

	Forskjell bivariat analyse	Kontrollert for sju forhold	Rest forskjell
Privat Kristen	-0,176	-0,083	47 %
Privat Allmenn	-0,268	-0,217	81 %
Privat Spesialisert	-0,067	-0,043	64 %

Tabell 3.4 viser at halvparten av forskjellen mellom kristne skoler og offentlige skoler forsvinner når vi kontrollerer for andre forhold. En viktig forklaring på dette er at de kristne skolene har en høyere standpunktkaraktersnitt enn de offentlige, en annen at mange av elevene på de kristne skolene går på små skoler. Det er dermed det høye karaktersnittet og skolestørrelsen som står for mye av forklaringen på det relativt store avviket mellom standpunktkarakterer og eksamenskarakterer her og ikke bare i seg selv at de er kristne skoler. For de spesialiserte private skolene opprettholdes det meste av den lille forskjellen sammenlignet med de offentlige skolene.

For de allmenne private skolene opprettholdes det meste av forskjellen i forhold til de offentlige skolene når vi kontrollerer for andre forhold. Imidlertid er det slik at det er noen få skoler med særlig stor gjennomsnittsforskjell som forklarer mye av dette. Forskjellen for alle private allmenne skoler er i gjennomsnitt på -0,217 karakterpoeng, sammenlignet med offentlige skoler. Vi kan illustrere dette på følgende måte. Dersom vi har en offentlig skole og en allmenn privat med samme kjønns- og aldersfordeling, samme standpunktkarakterer og elevtall og like mange elever på hvert programområde, og sammenligner dem de samme årene, så vil i gjennomsnitt 52 % av elevene på den offentlige skolen gå tilbake en karakter til eksamen, sammenlignet med standpunkt, mens det vil gjelde 73 % av elevene på den allmenne private skolen. Tabell 3.5 viser hvordan fordelingene i praksis kunne se ut på skoler med 100 elever når vi tar utgangspunkt i de dokumenterte effektene i tabell 3.4.

Tabell 3.5 Differanse mellom eksamenskarakter og standpunktkarakter i norsk skriftlig hovedmål når private skoler sammenlignes med offentlige skoler. Tenkt eksempel med identiske skoler der det kontrolleres for effekten av sju andre forhold.

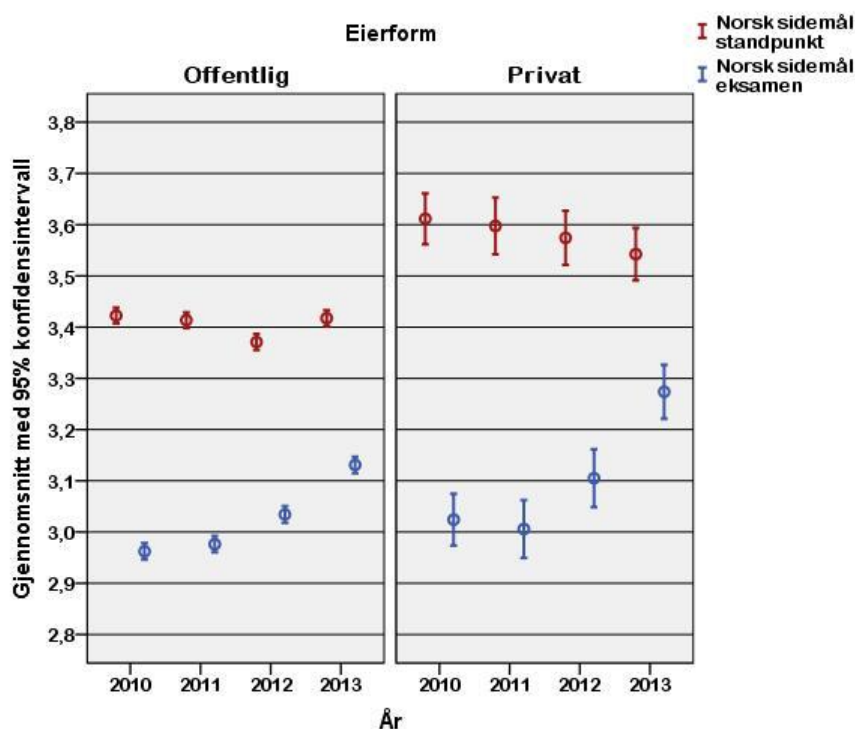
Differanse standpunkt- eksamen	Offentlig	Privat kristen	Privat spesia- lisert	Privat allmenn
-3	2	2	2	4
-2	12	13	13	16
-1	36	39	37	40
0	37	36	36	30
1	12	9	11	9
2	1	1	1	1
Sum	100	100	100	100
Snitt forskjell	-0,52	-0,60	-0,56	-0,73

Av de 100 elevene ved den offentlige skolen ville halvparten få en eksamenskarakter som er dårligere enn standpunktkarakteren, mens 13 ville få en karakter som er bedre. Ved den kristne skolen ville 54 elever få en dårligere karakter, mens 11 ville få en bedre karakter. Ved den allmenne, private skolen ville 60 av 100 få en dårligere karakter, mens 10 ville få en bedre karakter. Fordelingen ved den spesialiserte private skolen avviker kun marginalt fra det vi finner ved den offentlige skolen.

3.2 Norsk skriftlig sidemål

Vi skal nå se på resultatene i norsk skriftlig sidemål, men dette vil ikke bli gjort like inngående som for hovedmålet. Grunnen til dette er at de resultatene vi får peker helt i samme retning. En forklaring på sammenfallende karakterer i hovedmål og sidemål kan være at det i en del tilfeller er samme lærere og sensorer i begge fag³. Elevene er også de samme, men de er betydelig færre. Vi har eksamens- og standpunktkarakterer i norsk skriftlig sidemål bare for 45 prosent av elevene vi har hovedmålsopplysninger om.

Når vi sammenligner figur 3.8 med tilsvarende figur for hovedmålet (figur 3.4), finner vi flest likheter, men også noen forskjeller. Likhetene er for det første at nivået på standpunktkarakteren er temmelig likt fra år til år, mens eksamenskarakteren viser en stigende tendens. Dermed ser vi også her at standpunktkarakteren er mer stabil enn eksamenskarakteren over tid. For det andre er forskjellen mellom standpunktkarakteren og eksamenskarakteren større på private skoler enn på offentlige. For det tredje har de private skolene noe høyere standpunktkarakter enn de offentlige. Forskjellene er at nivået på standpunktkarakteren er lavere i sidemålet enn i hovedmålet. Dessuten er tendensen til økende nivå over tid for eksamenskarakteren tydeligere for sidemålet enn for hovedmålet.



Figur 3.8 Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk skriftlig sidemål i fire skoleår for offentlige og private skoler.

Vi har gjort en regresjonsanalyse også for sidemålet og resultatene av denne gjengis i tabell 3.6. Resultatene er til forveksling lik de vi finner for hovedmålet. Dette er et uttrykk for at vi her står overfor to vurderinger innenfor samme fag som er foretatt av de samme personene. Effekten av standpunktkarakteren er fortsatt den klart viktigste og praktisk talt den samme som for hovedmålet. Realfag avviker i positiv retning, mens Naturbruk og de to variantene av påbygg til studiekompetanse har de sterkeste negative koeffisientene av programområdene. Effekten av eksamensår er også tydelig med positive tall for 2012 og 2013. Også effekten av å gå på en skole med elever med høyt karaktersnitt fra grunnskolen er omtrent den samme.

³ I norskfaget generelt har man langt tradisjoner for utvikling av eksamensoppgaver, sensorveiledninger og organisert sensoroppfølging. Dette kan bidra til at man utvikler et sterkt tolkingsfelleskap.

Noen avvik fra mønsteret for hovedmålet er det likevel. Effekten av å gå på en allmenn privatskole er sterkere, mens vi ikke finner noen effekt av å gå på en kristen skole. Endelig er det ingen effekt av skolestørrelse. Det siste kan henge sammen med en faktor vi ikke har med i modellen, nemlig målform. Hvis det er slik at elever med nynorsk som målform i stor grad er å finne på små skoler, og nynorskelever gjør det bedre til sidemåls eksamen enn bokmåls elever, kan dette ha nøytralisert effekten av skolestørrelse. De kristne skolene er ofte små, og flere av dem har nynorsk som målform, så da kan dette forklare at effekten av å gå på en kristen skole forsvinner.

Tabell 3.6 Resultat av multipl lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i norsk skriftlig sidemål som avhengig variabel. N=71.468.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	1,005	0,049		20,663	0,000
<u>Norsk sidemål standpunkt</u>	-0,426	0,003	-0,471	-130,964	0,000*
<u>Jente</u>	0,094	0,007	0,050	14,460	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,040	0,013	-0,011	-3,072	0,002*
Alder 21 år	-0,043	0,021	-0,007	-2,002	0,045*
Alder 22 år pluss	-0,038	0,021	-0,007	-1,844	0,065
<u>Programområde: Ref = Annet</u>					
Realfag	0,117	0,046	0,052	2,550	0,011*
Språk, samf.fag, økonomi	-0,003	0,046	-0,002	-0,064	0,949
Formgivingsfag	-0,147	0,049	-0,026	-2,995	0,003*
Musikk	0,080	0,049	0,015	1,639	0,101
Dans	-0,038	0,056	-0,004	-0,677	0,498
Drama	0,056	0,057	0,005	0,988	0,323
Idrett	-0,142	0,046	-0,052	-3,071	0,002*
Media	-0,075	0,048	-0,015	-1,544	0,123
Naturbruk	-0,476	0,063	-0,037	-7,586	0,000*
Påbygg uten yrkeskompetanse	-0,330	0,046	-0,140	-7,210	0,000*
Påbygg etter yrkeskompetanse	-0,160	0,049	-0,028	-3,251	0,001*
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	0,027	0,017	0,006	1,583	0,113
Privat Allmenn	-0,240	0,021	-0,038	-11,219	0,000*
Privat Spesialisert	-0,043	0,021	-0,007	-2,041	0,041*
<u>Skole med høy inntakskvalitet</u>	0,126	0,008	0,056	15,923	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	-0,005	0,015	-0,002	-0,333	0,739
Elevtall 100 til 199	-0,013	0,014	-0,007	-0,950	0,342
Elevtall 200 til 299	-0,036	0,015	-0,017	-2,428	0,015*
Elevtall 300 og mer	-0,035	0,018	-0,010	-2,007	0,045*
<u>År: Referansekategori = 2010.</u>					
År 2011	0,017	0,009	0,008	1,990	0,047*
År 2012	0,111	0,009	0,051	12,619	0,000*
År 2013	0,181	0,009	0,086	20,986	0,000*

Andel forklart varians: 21 %. * : Statistisk signifikant på 5 % nivå eller lavere.

3.3 Norsk muntlig

3.3.1 Samvariasjon mellom norskkarakterene

Før vi skal se på muntligkarakterene i norsk, skal vi ta et nærmere blikk på norskkarakterene under ett. Tabell 3.7 viser samvariasjonen mellom de seks karakterene for elever der vi har opplysninger om alle seks. Norsk er ett fag og det er samme lærer som setter standpunktkarakteren. Derfor er det ikke overraskende at vi finner de sterkeste korrelasjonene mellom nettopp disse karakterene. Aller sterkest er den mellom standpunktkarakterene i hovedmålet og sidemålet. 54 prosent av elevene får samme standpunktkarakter i sidemål og hovedmål og siden sidemålskarakteren i gjennomsnitt ligger noe under hovedmålskarakteren, vil det være vanligere med en dårligere karakter i sidemål enn i hovedmål enn det motsatte (ikke vist i egen tabell).

Tabell 3.7. Korrelasjon mellom seks karakterer i norskfaget. Pearsons R. N = 7441.

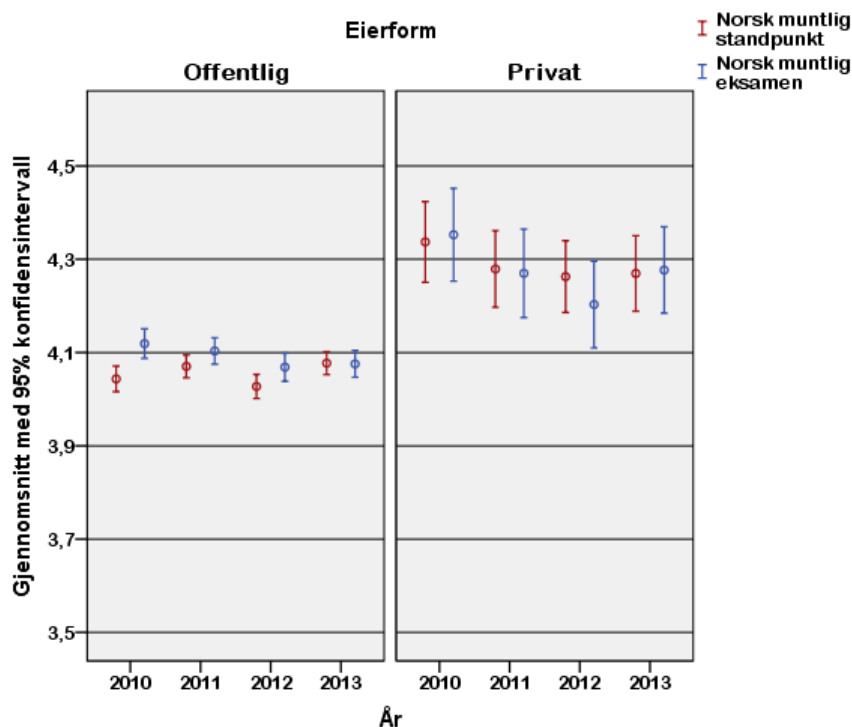
	Norsk hovedmål standpunkt	Norsk sidemål standpunkt	Norsk muntlig standpunkt	Norsk hovedmål eksamen	Norsk sidemål eksamen
Norsk sidemål standpunkt	,776**				
Norsk muntlig standpunkt	,729**	,700**			
Norsk hovedmål eksamen	,566**	,558**	,502**		
Norsk sidemål eksamen	,577**	,612**	,510**	,543**	
Norsk muntlig eksamen	,597**	,566**	,624**	,489**	,478**

** . Korrelasjonen er signifikant på 1 prosentnivå.

De nest sterkeste sammenhengene finner vi mellom standpunkt og eksamenskarakter i muntlig og i sidemål, mens korrelasjonen mellom standpunkt og eksamen er litt svakere for hovedmål. De svakeste sammenhengene finner vi mellom muntlig eksamen på den ene siden og eksamen i hovedmål og sidemål på den andre. Dette forklares med et høyere gjennomsnitt og mindre spredning i muntlig enn i skriftlig eksamen. Bare 20,5 prosent får samme karakter i muntlig eksamen og sidemåls eksamen, og bare 7,5 prosent får en dårligere karakter til muntlig eksamen. Alt i alt viser analysen at det til tross for ulikt gjennomsnittsnivå på de seks karakterene er en svært nær statistisk sammenheng mellom dem, spesielt for de tre standpunktkarakterene, som er satt av samme lærer.

3.3.2 Karakter i norsk muntlig

Når vi skiller mellom offentlige og private skoler, ser vi et helt annet mønster for standpunktkarakterer og eksamenskarakterer enn vi fant for de to skriftligkarakterene. Figur 3.9 viser dette.



Figur 3.9 Gjennomsnittlig standpunktkarakter og eksamenskarakter i norsk muntlig i fire skoleår for offentlige og private skoler.

Gjennomsnittskarakterene ved de private skolene ligger noe over de offentlige, både for standpunkt og eksamen og for alle fire år, og forskjellene er statistisk signifikante. Karakterene ved de offentlige skolene varierer mellom 4,0 og 4,1 som gjennomsnitt, mens de varierer mellom 4,2 og 4,35 ved de private. Dessuten er det, med unntak av tallene for de offentlige skolene i 2010, ingen signifikant forskjell mellom standpunktkarakterene og eksamenskarakterene. Det betyr slett ikke at de fleste får samme karakter til eksamen som de fikk som standpunktkarakter, dette gjelder bare 41 prosent av elevene. Det er omtrent like vanlig å gå fram som tilbake, men sannsynligheten for å gå fram er sterkest for de med standpunktkarakteren 2 eller 3, mens de som har 5 har større sannsynlighet for å gå tilbake. Er standpunktkarakteren 4, så er sannsynligheten for å gå fram like stor som å gå tilbake (ikke vist i egen tabell).

Tabell 3.8 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i norsk muntlig som avhengig variabel. N=29.620.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	1,099	0,071		15,527	0,000
<u>Norsk muntlig standpunkt</u>	-0,292	0,005	-0,314	-53,156	0,000*
<u>Jente</u>	0,033	0,012	0,016	2,838	0,005*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,027	0,021	-0,008	-1,301	0,193
Alder 21 år	-0,055	0,032	-0,010	-1,723	0,085
Alder 22 år pluss	-0,050	0,031	-0,011	-1,623	0,105
<u>Programområde: Ref = Annet</u>					
Realfag	0,175	0,062	0,072	2,796	0,005*
Språk, samf.fag, økonomi	0,082	0,062	0,038	1,316	0,188
Formgivingsfag	-0,023	0,072	-0,003	-0,320	0,749
Musikk	0,138	0,070	0,022	1,957	0,050*
Dans	0,092	0,091	0,008	1,010	0,313
Drama	0,197	0,080	0,021	2,445	0,014*
Idrett	-0,014	0,065	-0,004	-0,224	0,823
Media	0,125	0,068	0,022	1,827	0,068
Naturbruk	-0,155	0,100	-0,011	-1,555	0,120
Påbygg uten yrkeskompetanse	-0,123	0,062	-0,053	-1,995	0,046*
Påbygg etter yrkeskompetanse	0,080	0,067	0,016	1,202	0,229
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	0,026	0,033	0,005	0,798	0,425
Privat Allmenn	-0,058	0,030	-0,011	-1,910	0,056
Privat Spesialisert	0,181	0,041	0,025	4,383	0,000*
<u>Skole med høy inntakskvalitet</u>	0,126	0,015	0,050	8,498	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,057	0,028	0,021	2,024	0,043*
Elevtall 100 til 199	0,087	0,026	0,044	3,314	0,001*
Elevtall 200 til 299	0,057	0,028	0,024	2,043	0,041*
Elevtall 300 og mer	0,138	0,033	0,036	4,238	0,000*
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,034	0,016	-0,015	-2,110	0,035*
År 2012	-0,043	0,016	-0,019	-2,611	0,009*
År 2013	-0,063	0,016	-0,028	-3,951	0,000*

Andel forklart varians: 9 %.* : Statistisk signifikant på 5 % nivå eller lavere.

Regresjonsanalysen kan bare forklare 9 prosent av variasjonen i forskjellen mellom standpunktkarakteren og eksamenskarakteren. Vi ser at bare halvparten av faktorene vi har tatt med har signifikante effekter. Standpunktkarakteren er også her den viktigste faktoren, men koeffisienten er redusert med omtrent 30 prosent i forhold til det vi så for de skriftlige karakterene. Kjønn har liten betydning og alder ingen betydning. Bare tre av programområdene har signifikante verdier, nemlig Realfag og Drama, med positive verdier og Påbygg uten yrkeskompetanse, som er negativ.

Når det gjelder skoletype, finner vi et helt annet mønster enn det som gjelder for karakterene i skriftlig hovedmål og sidemål ved at de spesialiserte privatskolene har en positiv verdi, mens de to andre private skoletypene ikke avviker fra de offentlige skolene. Når vi undersøker den bivariate sammenhengen nærmere, finner vi at elevene ved de spesialiserte skolene har lavere gjennomsnittlige standpunktkarakterer enn elevene ved de to andre typene privatskoler, mens det ikke er noen forskjell i eksamensresultatet (ikke vist i egen tabell).

Skolestørrelse har omtrent samme betydning som vi så for norsk skriftlig hovedmål og vi finner også her en positiv effekt av å gå på en skole med et høyt karaktersnitt fra grunnskolen blant elevene. Endelig finner vi en effekt av årskull, men ulik den som gjaldt for karakterer i skriftlig. Her er det 2010 som skiller seg ut, ved at det da var flere som fikk en bedre karakter enn standpunktkarakteren til eksamen.

3.4 Oppsummering

Dette kapitlet har vist at standpunktkarakterene i skriftlig norsk hovedmål og sidemål ligger betydelig over eksamenskarakterene, men forskjellen har minsket over tid. Vi har også vist at det er stabilitet over tid i standpunktkarakterene mens eksamenskarakterene endrer seg. Forskjellen mellom standpunkt og eksamen er mindre i 2013 enn den var i 2010. I muntlig norsk derimot er det liten gjennomsnittlig forskjell mellom standpunkt og eksamen.

Bildet av at gjennomsnittsforskjellen mellom standpunkt og eksamen minsker over tid er tydelig for offentlige skoler, private kristne og spesialiserte skoler, mens mønsteret er mer uklart for de private allmenne skolene. Generelt ser vi at det er noen enkeltskoler som utmerker seg med særlig stor forskjell mellom standpunkt og eksamenskarakter.

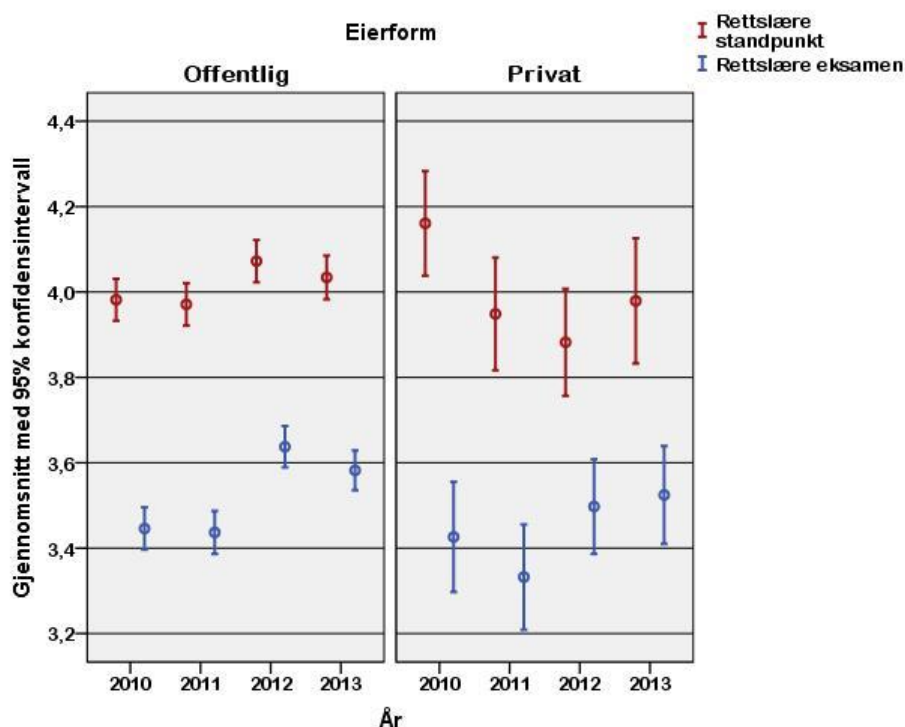
Standpunktkarakter er den variabelen som forklarer mest av gjennomsnittsforskjellen, og dette henger sammen med regresjon mot gjennomsnittet, det vil si at elever som har høy standpunktkarakter har høyere sannsynlighet for å gå ned ved eksamen enn elever med svak standpunktkarakter. Denne variabelen forklarer mer enn hvilken type skole eleven har gått på. Analysene så langt tyder på at sammenhengene mellom gjennomsnittlig standpunktkarakterer og eksamenskarakterer er langt mer komplekse og sammensatte enn at dette bare handler om forskjeller mellom offentlige og private skoler.

4 Karakterforskjeller i andre fag med skriftlig eksamen

I dette kapittelet sammenlignes karakterene til eksamen og standpunkt i de andre fagene med skriftlig eksamen, det vi si rettslære og matematikk. I disse fagene er det årlig mellom 2100 og 4700 elever som blir trukket til eksamen.

4.1 Rettslære

Rettslære (SAM3023) er et av de største avgangsfagene på Studiespesialisering, når vi tar utgangspunkt i hvor mange av Vg3-elevene som er oppe i dette til skriftlig eksamen. I vårt materiale finner vi mellom 2200 og 2400 elever som trekkes til eksamen hvert år og mellom 300 og 400 av disse går på private skoler.



Figur 4.1 Gjennomsnittlig standpunktkarakter og eksamenskarakter i rettslære (SAM3023) i fire skoleår for offentlige og private skoler.

De aller fleste av disse elevene går på programområdet Språk, samfunnsfag og økonomi, men det er også mulig for realfagselever å ta faget. Figur 4.1 viser at forskjellene mellom offentlige og private skoler synes å være marginale, og fordi elevtallet er lavt, særlig på de private skolene, vil den statistiske usikkerheten rundt gjennomsnittet bli tilsvarende stor. Det kan ellers synes som å være en heving i karakternivået på de offentlige skolene, spesielt til eksamen, når vi sammenligner de to første årene med de to siste.

Tabell 4.1 Resultat av multipl lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i rettslære som avhengig variabel. N=9.413.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	0,711	0,065		10,882	0,000
<u>Rettslære standpunkt</u>	-0,448	0,008	-0,500	-54,127	0,000*
<u>Jente</u>	0,098	0,020	0,046	4,977	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,142	0,048	-0,027	-2,973	0,003*
Alder 21 år	-0,210	0,095	-0,020	-2,214	0,027*
Alder 22 år pluss	-0,102	0,112	-0,008	-0,913	0,361
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,101	0,058	-0,017	-1,752	0,080
Privat Allmenn	-0,346	0,053	-0,058	-6,476	0,000*
Privat Spesialisert	0,249	0,039	0,064	6,326	0,000*
<u>Skole med høy inntakskvalitet</u>	0,134	0,022	0,057	6,029	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,348	0,058	0,110	6,016	0,000*
Elevtall 100 til 199	0,424	0,055	0,203	7,680	0,000*
Elevtall 200 til 299	0,538	0,058	0,213	9,278	0,000*
Elevtall 300 og mer	0,493	0,065	0,122	7,624	0,000*
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,006	0,026	-0,002	-0,218	0,828
År 2012	0,157	0,026	0,066	6,098	0,000*
År 2013	0,109	0,026	0,045	4,119	0,000*

Andel forklart varians: 25 %. * : Statistisk signifikant på 5 % nivå eller lavere.

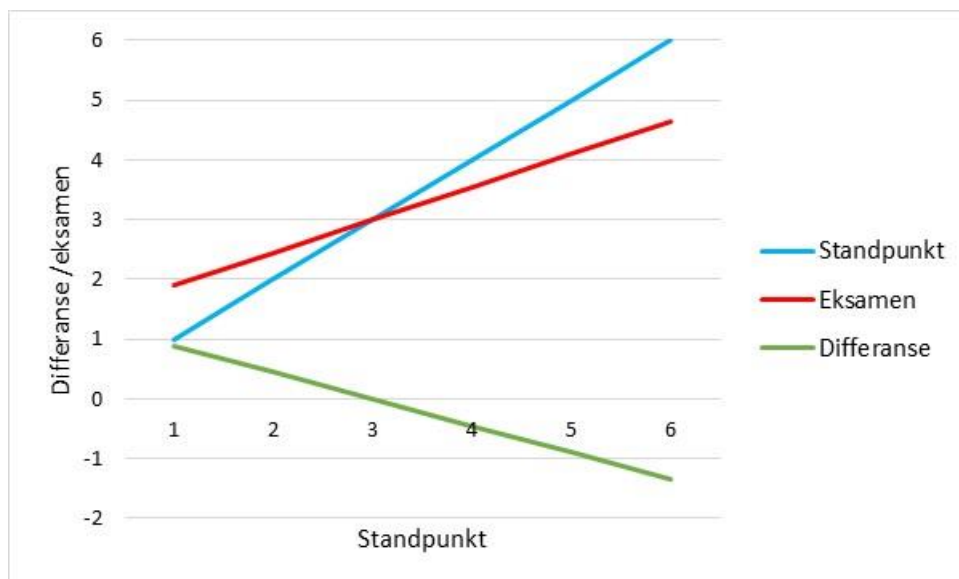
For rettslære kan vi forklare litt mer av variasjonen i differansen mellom standpunktkarakter og eksamenskarakter enn for norskfaget, 25 prosent. Viktigst som forklaringsfaktor er standpunktkarakteren, og effekten er enda litt sterkere enn for norsk skriftlig hovedmål og sidemål. For hver hele karakter standpunktkarakteren øker, øker den gjennomsnittlige forskjellen mellom eksamenskarakteren og standpunktkarakteren med -0,448 karakterpoeng. Vi ser ellers at det er både en kjønnseffekt og en alderseffekt slik at jenter har mindre forskjell mellom eksamenskarakteren og standpunktkarakteren enn gutter, mens 20- og 21-åringer har en større forskjell enn 19-åringer.

Vi finner dessuten tydelige effekter av skoleslag. Elever på allmenne private skoler har større forskjell mellom eksamenskarakteren og standpunktkarakteren enn elever på offentlige skoler, mens det motsatte er tilfelle for elever på spesialiserte skoler. Når vi går inn og ser på de bivariate sammenhengene, finner vi at elevene på begge de to typene privatskoler har en eksamenskarakter som ligger så vidt under nivået på de offentlige skolene. Ser vi på standpunktkarakteren, så er den lavere på de spesialiserte private skolene enn på de offentlige, mens den er høyere på de allmenne private skolene (ikke vist i tabell). Dette bildet med større forskjeller mellom karakterene på de allmenne private skolene og mindre på de spesialiserte skolene opprettholdes også når vi kontrollerer for andre forhold i regresjonsanalysen.

Også om eleven går på en skole med høy inntakskvalitet og skolestørrelse spiller inn. For skolestørrelsen er det viktigste at skolen har minst 50 elever, mens det har mindre betydning hvor mange den da har. Endelig finner vi tydelige effekter av kull ved at de to siste årene skiller seg fra de to første ved en mindre forskjell mellom eksamenskarakter og standpunktkarakter.

I figur 4.2 illustrerer vi effekten av standpunktkarakteren mens vi holder de andre forholdene som har signifikante effekter konstant. Vi varierer standpunktkarakteren for 19 år gamle jenter på offentlige skoler med mellom 100 og 199 avgangselever i 2013. Skolene er ikke skoler med høy inntakskvalitet.

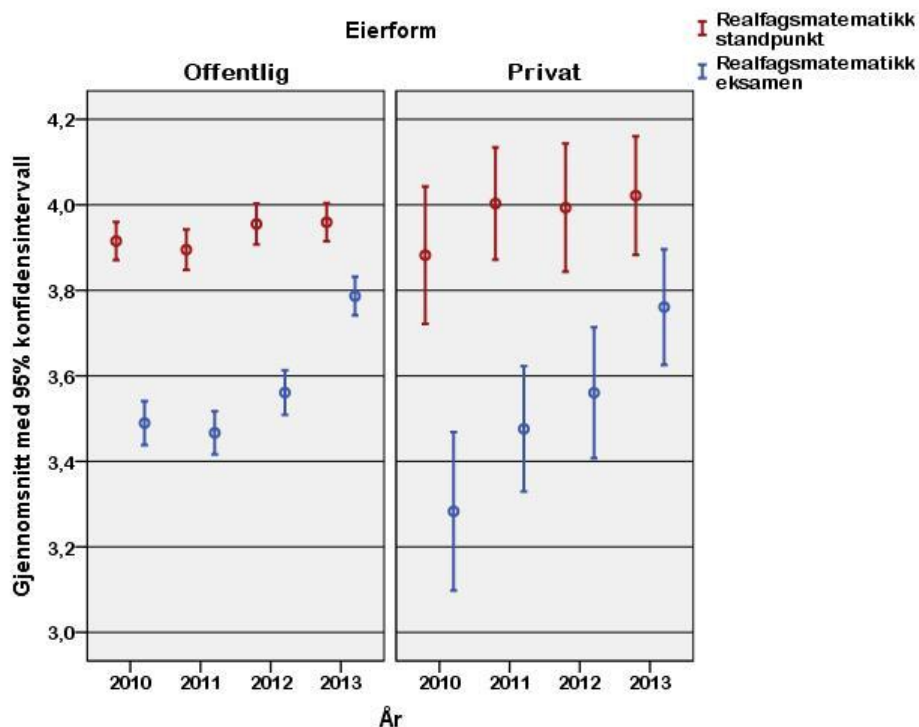
Vi ser at de elevene som har en standpunktkarakter på 1 i gjennomsnitt får eksamenskarakteren 1,89, altså en tydelig framgang, mens elever med standpunktkarakteren 6 går i gjennomsnitt tilbake med 1,35 karakterpoeng til 4,65. De som har standpunktkarakteren 3 får i gjennomsnitt ingen endring.



Figur 4.2 Estimert effekt av standpunktkarakteren på eksamenskarakteren i rettslære for ei jente på 19 år som går i Vg3 på en offentlig skole uten spesielt høyt karakternivå og med mellom 100 og 199 avgangselever i 2013. Lineær regresjon, estimer basert på tabell 4.1.

4.2 Realfagsmatematikk

Skriftlig eksamen i realfagsmatematikk (REA3024) berører omtrent 3.500 avgangselever på Studiespesialisering hvert år og 10 prosent av disse går på private skoler. Figur 4.3 viser at det gjennomsnittlige nivået for standpunktkarakterene har ligget svært stabilt i de fire årene vi har tatt for oss, og det er heller ingen statistisk signifikant forskjell mellom private og offentlige skoler. Når det gjelder eksamenskarakterene, er situasjonen en helt annen. Generelt har det vært en heving av gjennomsnittsnivået og spesielt for de private skolene. Her har gjennomsnittskarakteren steget fra ca.3,3 til 3,75. Det betyr samtidig at forskjellen mellom private og offentlige skoler, som vi ser for 2010, er helt forsvunnet i 2013. Vi så samme tendens for norsk hovedmål og sidemål, men utviklingen for matematikk synes å være enda tydeligere. Dette funnet vil bli fulgt opp nærmere i den kvalitative delen av prosjektet, våren 2014.



Figur 4.3 Gjennomsnittlig standpunktkarakter og eksamenskarakter i realfagsmatematikk (REA3024) i fire skoleår for offentlige og private skoler.

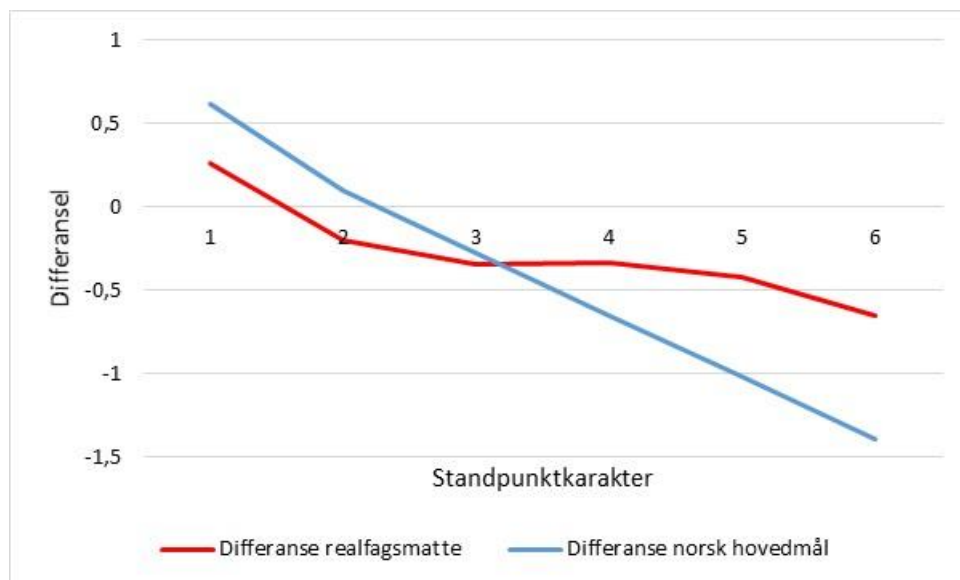
Det er gjort en regresjonsanalyse også for realfagsmatematikken og resultatene gjengis i tabell 4.2. Det mest påfallende resultatet i denne analysen er at standpunktkarakteren spiller en langt mindre rolle for karakterutviklingen enn det vi har sett for norsk og rettslære. Riktignok finner vi også her en negativ effekt, men den er bare en fjerdedel av det vi har sett for de andre fagene med skriftlig eksamen. Det er også her en regresjon mot gjennomsnittet, slik at elever med de beste standpunktkarakterene har større sannsynlighet for en dårligere karakter til eksamen enn de med dårlige karakterer, men mønsteret er et helt annet enn det vi så for norskfaget og rettslære.

I figur 4.4 sammenligner vi de observerte forskjellene mellom eksamenskarakter og standpunktkarakter i realfagsmatematikk og norsk skriftlig hovedmål. I norskfaget er det en nesten perfekt lineær sammenheng med en stigningskoeffisient på $-0,40$. For realfagsmatematikken finner vi et helt annet mønster. Her er det bare ved ytterpunktene av karakterskalaen at vi finner verdier som avviker. De som har standpunktkarakteren karakteren 1 kan ikke få dårligere karakter og får en gjennomsnittlig eksamenskarakter som er 0,26 karakterpoeng høyere. De med en sekser kan ikke gå fram og får en eksamenskarakter som er 0,65 karakterpoeng lavere. For de med standpunktkarakter mellom 2 og 5, er eksamenskarakteren mellom 0,20 og 0,42 karakterpoeng lavere. Den gjennomsnittlige stigningskoeffisienten blir $-0,18$, altså under halvparten av det vi fant for norsk hovedmål. Når tabell 4.2 viser at koeffisienten er $-0,114$, betyr det at enda noe av effekten av standpunktkarakteren forsvinner når vi kontrollerer for andre variabler. Vi ser for eksempel at skolestørrelse har betydning, spesielt når det er minst 100 avgangselever. Året 2013 skiller seg dessuten ut fra de andre tre årene. Det betyr også noe å gå på en skole med mange elever med gode grunnskolekarakterer. 21-åringene kommer noe dårligere ut enn andre og det er en liten negativ, men signifikant negativ effekt for jentene.

Tabell 4.2 Resultat av multipl lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i realfagsmatematikk som avhengig variabel. N=13.839.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	-0,353	0,039		-9,160	0,000
<u>Realfagsmatematikk standpunkt</u>	-0,114	0,005	-0,184	-22,263	0,000*
<u>Jente</u>	-0,042	0,014	-0,025	-3,020	0,003*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,044	0,036	-0,010	-1,208	0,227
Alder 21 år	-0,194	0,077	-0,021	-2,528	0,011*
Alder 22 år pluss	-0,088	0,065	-0,011	-1,356	0,175
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,028	0,033	-0,007	-0,831	0,406
Privat Allmenn	-0,063	0,045	-0,012	-1,420	0,156
Privat Spesialisert	0,103	0,047	0,018	2,176	0,030*
<u>Skole med høy inntakskvalitet</u>	0,145	0,015	0,084	9,784	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,168	0,035	0,074	4,751	0,000*
Elevtall 100 til 199	0,373	0,033	0,229	11,434	0,000*
Elevtall 200 til 299	0,402	0,034	0,212	11,759	0,000*
Elevtall 300 og mer	0,382	0,039	0,128	9,803	0,000*
<u>År: Referansekategori = 2010.</u>					
År 2011	0,002	0,019	0,001	0,121	0,904
År 2012	0,043	0,019	0,022	2,264	0,024*
År 2013	0,260	0,018	0,142	14,059	0,000*

Andel forklart varians: 8 %.* : Statistisk signifikant på 5 % nivå eller lavere.



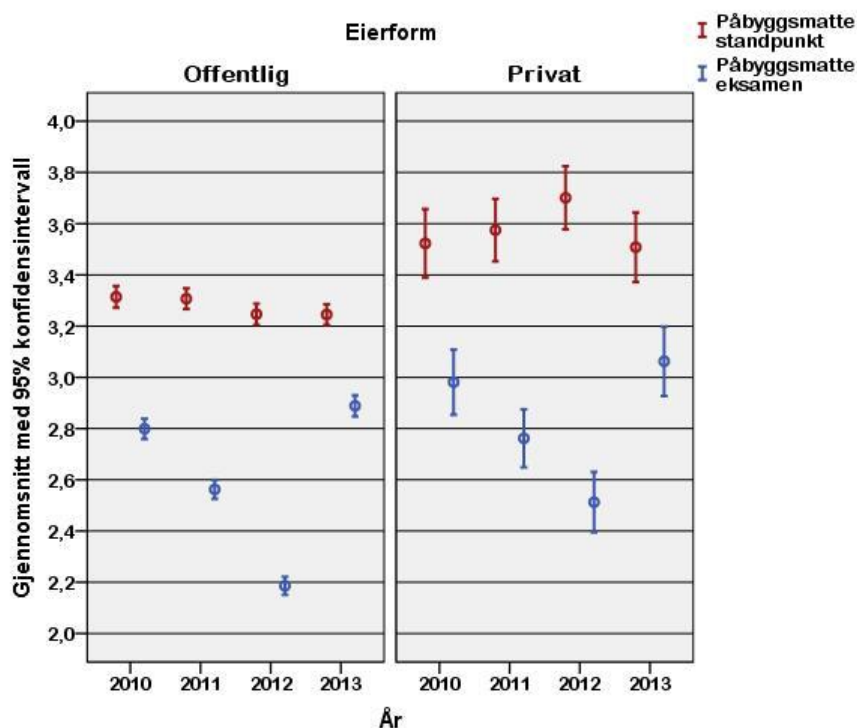
Figur 4.4 Gjennomsnittlig differanse mellom eksamenskarakter og standpunktkarakter i fire skoleår for realfagsmatte og norsk skriftlig hovedmål. Etter standpunktkarakter.

Av skoletypene er det bare de private spesialiserte skolene som skiller seg signifikant fra de offentlige skolene, men det er bare snakk om en ganske svak effekt på 0,103. Som vi har sett for norsk muntlig og rettslære, så ligger forklaringen på dette i at standpunktkarakterene er lavere ved de spesialiserte

private skolene enn ved de øvrige skoleslagene, mens det ikke er noen forskjell av betydning mellom skoleslagene for eksamenskarakterenes del (ikke vist i egen tabell).

4.3 Matematikk for påbyggings elever

Matematikk for påbyggings elever (MAT1005) omfatter en stor elevgruppe og ca. 4500 er oppe til skriftlig eksamen hvert år, hvorav omtrent 10 prosent går på private skoler. 90 prosent kommer fra påbygg for studiekompetanse og de aller fleste av disse tar påbygg etter to år på yrkesfag, mens 10 prosent tar påbygg etter å ha oppnådd yrkeskompetanse. De gjenværende 10 prosentene fordeler seg med 8 prosent fra Medier og kommunikasjon og to prosent fra Naturbruk.



Figur 4.5 Gjennomsnittlig standpunktkarakter og eksamenskarakter i påbyggsmatematikk (MAT1005) i fire skoleår for offentlige og private skoler.

Utviklingen i påbyggsmatematikk viser helt særegne trekk i forhold til de andre fagene med skriftlig eksamen. Mens nivået for standpunktkarakterene sank markert fra 2010 til 2012, økte det kraftig fra 2012 til 2013 (figur 4.5). Mønsteret er det samme i offentlige og private skoler med 2012 som et bunnår og 2010 og 2013 på omtrent samme nivå. Karakternivået er hvert år litt høyere i de private skolene enn i de offentlige og dette gjelder begge karakterer. Eksamen i 2012 framstår som svært spesiell ved at hele 35 prosent strøk dette året, mot «bare» 16 og 17 prosent i 2010 og 2013 (ikke vist i egen tabell). Det er flere som stryker i offentlige enn i private skoler. Ut fra det vi her ser er det lite som tyder på at vi skal finne noen negativ effekt på karakterutviklingen fra standpunkt til eksamen for de private skolene.

Regresjonsanalysen (tabell 4.3) kan forklare 27 prosent av variasjonen i forskjellen mellom eksamenskarakteren og standpunktkarakteren i påbyggsmatematikk. Tabell 4.3 viser at det også for påbyggsmatematikken er standpunktkarakteren som har størst betydning for karakterutviklingen. Koeffisienten på -0,275 er høyere enn for realfagsmatematikken, men lavere enn for norsk og rettslære. Vi ser ellers at private spesialiserte skoler kommer ut med en betydelig positiv effekt. En nærmere analyse viser at elevene fra disse skolene har en bedre karakterutvikling (mindre negativ) fra standpunkt til eksamen enn elever fra andre typer private skoler og fra offentlige skoler når standpunktkarakteren er mellom 1 og 5 (ikke vist i egen tabell).

Tabell 4.3 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i påbyggsmatematikk som avhengig variabel. N=18.109.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	0,276	0,027		10,252	0,000
<u>Påbyggsmatematikk standpunkt</u>	-0,275	0,004	-0,413	-63,868	0,000*
<u>Jente</u>	-0,127	0,012	-0,069	-10,823	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,052	0,017	-0,020	-3,077	0,002*
Alder 21 år	-0,073	0,021	-0,024	-3,407	0,001*
Alder 22 år pluss	-0,097	0,020	-0,036	-4,841	0,000*
<u>Programområde: Ref = Påbygg 3</u>					
Medier og kommunikasjon	0,143	0,021	0,044	6,702	0,000*
Naturbruk	-0,083	0,048	-0,011	-1,708	0,088
Påbygg etter studiekompetanse	0,188	0,022	0,066	8,748	0,000*
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,016	0,032	-0,003	-0,489	0,625
Privat Allmenn	-0,071	0,031	-0,015	-2,307	0,021*
Privat Spesialisert	0,349	0,043	0,056	8,087	0,000*
<u>Skole med høy inntakskvalitet</u>	0,192	0,021	0,060	9,114	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,205	0,021	0,098	9,657	0,000*
Elevtall 100 til 199	0,190	0,020	0,104	9,429	0,000*
Elevtall 200 til 299	0,250	0,023	0,108	11,003	0,000*
Elevtall 300 og mer	0,200	0,034	0,044	5,912	0,000*
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,242	0,016	-0,119	-14,778	0,000*
År 2012	-0,580	0,016	-0,283	-35,167	0,000*
År 2013	0,117	0,017	0,057	7,043	0,000*

Andel forklart varians: 27 %.* : Statistisk signifikant på 5 % nivå eller lavere.

Vi finner ellers en klart negativ effekt av å være jente og svakere negative effekter av å være over 19 år. Elever fra Medier og kommunikasjon og elever som tar påbygg etter yrkeskompetanse har en mer positiv karakterutvikling enn elever som tar påbygg etter to år i yrkesfag. Det er en svak negativ effekt for elevene ved de allmenne private skolene, mens det er en klar fordel å gå på en skole med høy inntakskvalitet. Videre finner vi positive effekter av skolestørrelse, men det spiller ingen rolle hvor stor skolen er bare den har mer enn 50 elever. Endelig finner vi effekter av skoleår som samsvarer med det som er vist i figur 4.5, ved at 2011 og 2012 kommer ut med negative effekter sammenlignet med 2010, mens det er en svakere positiv effekt for 2013.

4.4 Oppsummering

Dette kapitlet har tatt for seg forskjellen mellom standpunktkarakter og eksamenskarakter i tre ulike fag: rettslære, realfagsmatematikk og matematikk på påbygging til studiekompetanse. Alle tre fagene viser samme mønster som norsk hovedmål og sidemål, at avstanden mellom standpunkt og eksamen er mindre i 2013 enn de tidligere årene, og at dette skyldes at gjennomsnittlig eksamenskarakter er bedre i 2013 mens standpunktkarakterene er relativt stabile over tid.

Gjennomsnittsforskjellene i rettslære er i stor grad den samme som for norsk hovedmål og sidemål, det vil si at det er relativt stor avstand mellom standpunktkarakter og eksamen og at flertallet av elevene har en eksamenskarakter som er lavere enn standpunkt.

I realfagsmatematikk derimot er avstanden mellom standpunkt og eksamen mindre, og dersom vi ser den gjennomsnittlige differansen i forhold til karakter ser vi at mønsteret indikerer ingen sammenheng: elever med standpunktkarakter mellom 2 og 5 har alle mellom 0,2 og 0,4 karakterpoeng lavere eksamenskarakter. Det er kun de med karakterene 1 og 6 som avviker fra mønsteret, men dette kommer av at de ikke kan gå ned respektive opp.

Mens utviklingen over tid for de andre fagene er positiv øker avstanden mellom standpunkt og eksamenskarakter i matematikk for påbyggingselever fra 2010 til 2012, for siden å øke kraftig i 2013. Dette forklares i stor grad av at det var 35 prosent som strøk det året, mens vanlig strykandel i faget er 16-17 prosent. Det er mulig å anta at den kraftige nedgangen i eksamenskarakterer i 2012 kommer av endringer i eksamensutformingen eller sensorordningen det året..

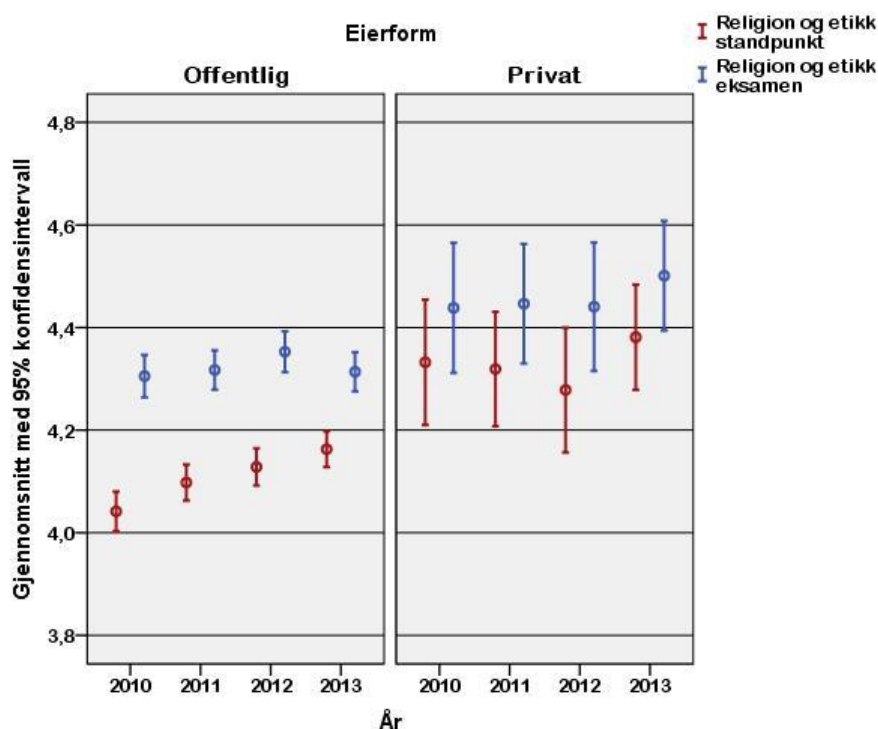
Det er generelt stor effekt av skolestørrelse på gjennomsnittsforskjellen mellom standpunkt og eksamen i disse fagene, ved at det er en entydig ulempe i form av større avvik å gå på en liten skole, som har mindre enn 50 avgangselever det aktuelle året.

5 Karakterforskjeller i andre fag med muntlig eksamen

Dette kapitlet tar før seg muntlig eksamen i religion og etikk for studiespesialisering samt historie for studiespesialisering og historie for elever på påbygging til studiekompetanse.

5.1 Religion og etikk for elever ved studiespesialisering

Religion og etikk (REL1001) er et av fellesfagene i studieforberedende utdanningsprogram. I vårt materiale er mellom 3500 og 4300 elever trukket ut til eksamen hvert år i dette faget, ca. 11 prosent fra private skoler. 12 prosent er fra Idrettsfag, 6,5 prosent fra Musikk, dans og drama og resten fra Studiespesialisering.



Figur 5.1 Gjennomsnittlig standpunktkarakter og eksamenskarakter i religion og etikk (REL1001) i fire skoleår for offentlige og private skoler.

For religion og etikk finner vi en gjennomsnittlig *positiv* forskjell mellom eksamenskarakteren og standpunktkarakteren (figur 5.1). Mens 41 prosent beholder standpunktkarakteren til eksamen, får 37 prosent en bedre karakter og bare 22 prosent en dårligere karakter (ikke vist i egen tabell). Karakternivået ligger generelt litt høyere ved private skoler enn ved offentlige, mens vi ser en utvikling over tid i retning av bedre standpunktkarakterer ved de offentlige skolene.

Regresjonsanalysen (tabell 5.1) kan bare forklare 13 prosent av variasjonen i forskjellen mellom de to karakterene og viser at standpunktkarakteren har en negativ effekt på karakterforskjellen; jo bedre standpunktkarakter, desto større sannsynlighet for en dårligere eksamenskarakter. Vi ser videre at det ikke er noen kjønnseffekt, samtidig som tabellen viser at det ikke er noen fordel å være over 19 år.

Tabell 5.1 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i påbyggsmatematikk som avhengig variabel. N=15.918.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign. nivå.
	B	Std. feil	Beta	t-verdi	
Konstantledd	1,655	0,047		35,496	0,000
<u>Religion og etikk standpunkt</u>	-0,330	0,007	-0,372	-47,229	0,000*
<u>Jente</u>	0,006	0,016	0,003	0,417	0,677
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,157	0,032	-0,036	-4,907	0,000*
Alder 21 år	-0,185	0,070	-0,020	-2,656	0,008*
Alder 22 år pluss	-0,203	0,069	-0,022	-2,940	0,003*
<u>Programområde: Ref = Realfag</u>					
Språk, samf.fag, økonomi	-0,093	0,018	-0,047	-5,250	0,000*
Formgivingsfag	-0,194	0,038	-0,041	-5,159	0,000*
Musikk	-0,002	0,041	0,000	-0,044	0,965
Dans	-0,005	0,072	-0,001	-0,074	0,941
Drama	0,002	0,057	0,000	0,034	0,973
Idrett	-0,198	0,026	-0,065	-7,685	0,000*
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,017	0,042	-0,003	-0,393	0,694
Privat Allmenn	-0,091	0,040	-0,017	-2,253	0,024*
Privat Spesialisert	0,107	0,046	0,018	2,345	0,019*
<u>Skole med høy inntakskvalitet</u>	0,051	0,018	0,022	2,840	0,005*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	-0,026	0,036	-0,009	-0,723	0,470
Elevtall 100 til 199	0,017	0,033	0,009	0,533	0,594
Elevtall 200 til 299	0,027	0,035	0,012	0,771	0,441
Elevtall 300 og mer	0,029	0,041	0,008	0,712	0,476
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,023	0,021	-0,010	-1,069	0,285
År 2012	-0,008	0,021	-0,004	-0,380	0,704
År 2013	-0,064	0,021	-0,029	-3,023	0,003*

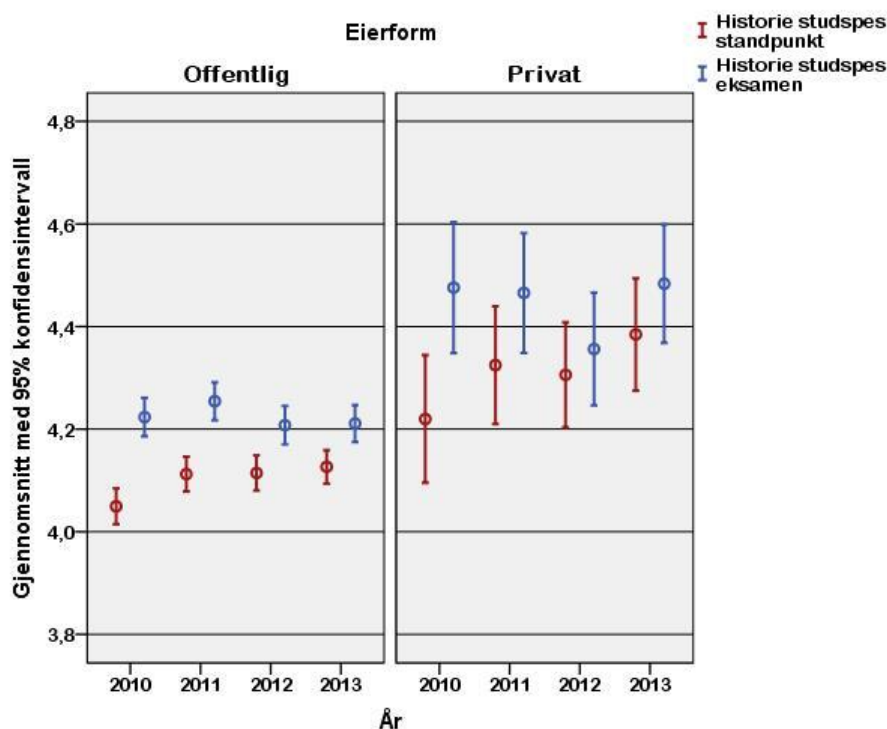
Andel forklart varians: 13 %. * : Statistisk signifikant på 5 % nivå eller lavere.

Realfagselevne er referansekategori for programområde, og vi ser at elevene på språk, samfunnsfag og økonomi, formgivingsfag og idrettsfag har en negativ karakterutvikling i forhold til realfagselevne. Videre ser vi en negativ effekt for elever på allmenne private skoler og en positiv effekt for de spesialiserte private skolene. Vi finner også en svak positiv effekt av å gå på en skole med høy inntakskvalitet, mens elevtallet ikke har noen betydning. Endelig skiller 2013 seg fra 2010 med en svakere karakterutvikling, noe som henger sammen med en økning i nivået på standpunktkarakteren og en utjevning av forskjellen over tid.

Som en kuriositet kan til slutt nevnes at det er elevene ved de kristne private skolene som har best resultater i religion og etikk med i gjennomsnitt 4,7 både som standpunktkarakter og til eksamen.

5.2 Historie for elever ved studiespesialisering

Historie (HIS1002) er også et av fellesfagene i studieforberedende utdanningsprogram. I vårt materiale er mellom 4400 og 5000 elever trukket ut til eksamen hvert år i dette faget, ca. 8 prosent fra private skoler. 12 prosent er fra Idrettsfag, 7 prosent fra Musikk, dans og drama og resten fra Studiespesialisering.



Figur 5.2 Gjennomsnittlig standpunktkarakter og eksamenskarakter i historie (HIS1002) i fire skoleår for offentlige og private skoler.

Også for historie finner vi en gjennomsnittlig positiv forskjell mellom eksamenskarakteren og standpunktkarakteren (figur 5.2). Mens 42 prosent beholder standpunktkarakteren til eksamen, får 34 prosent en bedre karakter og 24 prosent en dårligere karakter (ikke vist i egen tabell). Karakternivået ligger generelt litt høyere ved private skoler enn ved offentlige, mens vi ser en generell utvikling over tid i retning av bedre standpunktkarakterer.

Regresjonsanalysen (tabell 5.2) kan bare forklare 12 prosent av variasjonen i forskjellen mellom de to karakterene og viser at standpunktkarakteren har en negativ effekt på karakterforskjellen; jo bedre standpunktkarakter, desto større sannsynlighet for en dårligere eksamenskarakter. Vi ser videre at det er en negativ effekt av å være jente, samtidig som det ikke er noen fordel å være over 20 år.

Realfagselevne er referansekategori for programområdene, og vi ser at elevene på språk, samfunnsfag og økonomi, formgivingsfag og idrettsfag har en negativ karakterutvikling i forhold til realfagselevne, akkurat slik det var for religion og etikk. Koeffisienten for Dans er bare på grensen til å være statistisk signifikant, noe som skyldes at det her gjelder en svært liten gruppe elever. Videre ser vi positive effekt for elever både ved private kristne og spesialiserte skoler. Vi finner også en svak positiv effekt av å gå på en skole med høy inntakskvalitet, mens elevtallet ikke har noen betydning. Endelig skiller 2012 og 2013 seg fra 2010 og 2011 med en svakere karakterutvikling, noe som henger sammen med en økning i nivået på standpunktkarakteren og en utjevning av forskjellen over tid.

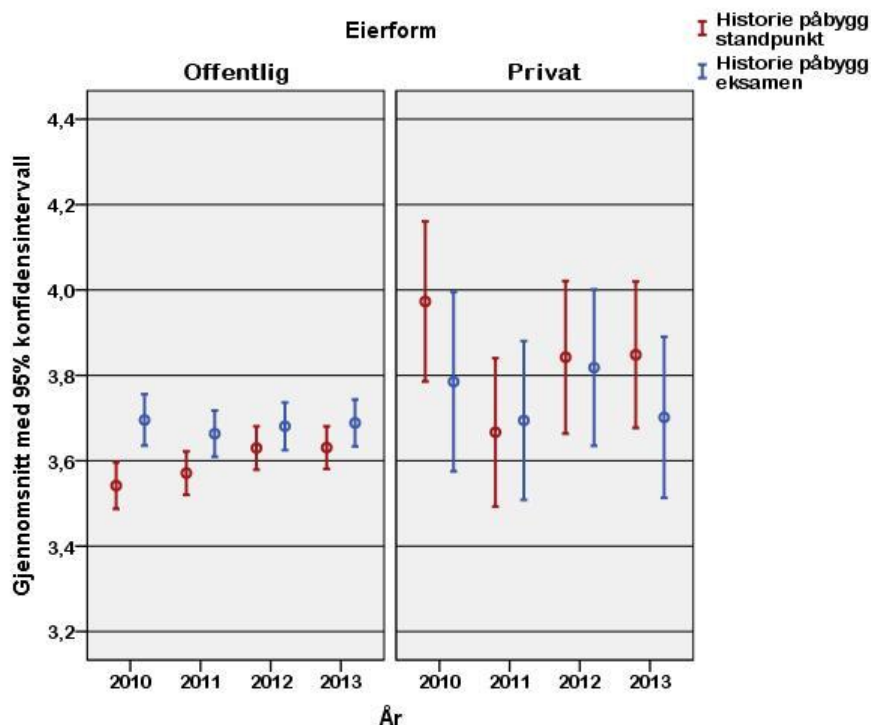
Tabell 5.2 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i påbyggsmatematikk som avhengig variabel. N=18.580.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign. nivå.
	B	Std. feil	Beta	t-verdi	
Konstantledd	1,464	0,045		32,476	0,000
<u>Religion og etikk standpunkt</u>	-0,293	0,006	-0,336	-46,843	0,000*
<u>Jente</u>	-0,138	0,014	-0,070	-9,803	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,059	0,031	-0,013	-1,910	0,056
Alder 21 år	-0,164	0,062	-0,018	-2,637	0,008*
Alder 22 år pluss	-0,145	0,064	-0,016	-2,285	0,022*
<u>Programområde: Ref = Realfag</u>					
Språk, samf.fag, økonomi	-0,046	0,016	-0,023	-2,811	0,005*
Formgivingsfag	-0,196	0,037	-0,039	-5,257	0,000*
Musikk	-0,050	0,037	-0,010	-1,366	0,172
Dans	-0,122	0,064	-0,013	-1,888	0,059
Drama	-0,008	0,053	-0,001	-0,145	0,885
Idrett	-0,087	0,024	-0,029	-3,676	0,000*
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	0,098	0,043	0,016	2,262	0,024*
Privat Allmenn	-0,009	0,040	-0,002	-0,237	0,813
Privat Spesialisert	0,134	0,049	0,020	2,768	0,006*
<u>Skole med høy inntakskvalitet</u>	0,059	0,016	0,026	3,611	0,000*
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,004	0,036	0,001	0,100	0,921
Elevtall 100 til 199	0,024	0,033	0,012	0,716	0,474
Elevtall 200 til 299	-0,004	0,035	-0,002	-0,124	0,901
Elevtall 300 og mer	0,080	0,039	0,023	2,039	0,041
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,023	0,020	-0,010	-1,179	0,239
År 2012	-0,076	0,020	-0,033	-3,856	0,000*
År 2013	-0,071	0,019	-0,032	-3,693	0,000*

Andel forklart varians: 12 %. * : Statistisk signifikant på 5 % nivå eller lavere.

5.3 Historie for påbyggingsselever

Historie for påbyggingsselever (HIS1002) omfatter mellom 1900 og 2300 elever som er oppe til muntlig eksamen hvert år, hvorav omtrent 8 prosent går på private skoler. 87 prosent kommer fra påbygg for studiekompetanse og de aller fleste av disse tar påbygg etter to år på yrkesfag, mens 12 prosent tar påbygg etter å ha oppnådd yrkeskompetanse. De gjenværende 13 prosentene fordeler seg med 10 prosent fra Medier og kommunikasjon og tre prosent fra Naturbruk.



Figur 5.3 Gjennomsnittlig standpunktkarakter og eksamenskarakter i påbyggshistorie (HIS1003) i fire skoleår for offentlige og private skoler.

Bildet for påbyggshistorie minner mye om det vi fant for historie for studieforberedende. Nivået for standpunktkarakterene har økt marginalt over tid (figur 5,3) og det er en tendens til høyere karakternivå ved private skoler. Tallene for de private skolene er imidlertid så små at usikkerheten blir stor. Den gjennomsnittlige forskjellen mellom standpunktkarakter og eksamenskarakter er liten. 38 prosent av elevene får samme karakter til eksamen som standpunktkarakteren, 33 prosent får en bedre karakter, mens 29 prosent får en karakter som er dårligere.

Regresjonsanalysen (tabell 4.3) kan bare forklare 14 prosent av variasjonen i forskjellen mellom de to karakterene og viser at det også for påbyggshistorie er standpunktkarakteren som har størst betydning for karakterutviklingen. Utover dette er det bare to forhold som viser signifikante effekter. Vi finner en relativt sterk negativ effekt for jentene og en positiv effekt for Medier og kommunikasjon. Når vi ser på forskjellen mellom gutter og jenter i en bivariat analyse, finner vi at guttene har en tydelig framgang fra standpunkt til eksamen, mens jentene står på stedet hvil (ikke vist i egen tabell). Med unntak for de elevene som standpunktkarakteren 1, har guttene en bedre karakterutvikling enn jentene, uansett standpunktkarakter (ikke vist i egen tabell).

Tabell 5.3 Resultat av multippel lineær regresjonsanalyse med differansen mellom eksamenskarakter og standpunktkarakter i påbyggshistorie som avhengig variabel. N=8682.

	Ikke-standardiserte koeffisienter		Standardiserte koeffisienter		Sign.
	B	Std. feil	Beta	t-verdi	B
Konstantledd	1,417	0,053		26,682	0,000
<u>Påbyggshistorie standpunkt</u>	-0,326	0,009	-0,355	-35,074	0,000*
<u>Jente</u>	-0,241	0,022	-0,110	-10,955	0,000*
<u>Alder: Ref.kat. = 19 år / yngre</u>					
Alder 20 år	-0,061	0,032	-0,020	-1,930	0,054
Alder 21 år	-0,075	0,041	-0,020	-1,840	0,066
Alder 22 år pluss	-0,052	0,037	-0,016	-1,411	0,158
<u>Programområde: Ref = Påbygg 3</u>					
Medier og kommunikasjon	0,103	0,037	0,029	2,829	0,005*
Naturbruk	-0,019	0,072	-0,003	-0,268	0,789
Påbygg etter studiekompetanse	-0,016	0,039	-0,005	-0,409	0,682
<u>Skoletype: Ref. = Offentlig</u>					
Privat Kristen	-0,082	0,063	-0,013	-1,289	0,198
Privat Allmenn	-0,035	0,063	-0,006	-0,548	0,584
Privat Spesialisert	-0,165	0,090	-0,020	-1,837	0,066
<u>Skole med høy inntakskvalitet</u>	0,020	0,039	0,005	0,511	0,609
<u>Avgangselever: Ref. = < 50.</u>					
Elevtall 50 til 99	0,022	0,040	0,008	0,534	0,593
Elevtall 100 til 199	0,046	0,038	0,021	1,207	0,227
Elevtall 200 til 299	0,045	0,042	0,017	1,071	0,284
Elevtall 300 og mer	0,036	0,061	0,007	0,593	0,553
<u>År: Referansekategori = 2010.</u>					
År 2011	-0,030	0,031	-0,012	-0,964	0,335
År 2012	-0,050	0,032	-0,020	-1,589	0,112
År 2013	-0,055	0,031	-0,022	-1,748	0,080

Andel forklart varians: 14 %.* : Statistisk signifikant på 5 % nivå eller lavere.

5.4 Oppsummering

I dette kapittelet har vi tatt for oss forskjeller mellom standpunktkarakter og muntlig eksamen i religion og etikk for studiespesialisering samt historie for studiespesialisering og historie for elever på påbygging til studiekompetanse. Alle analysene viser at det generelt er liten forskjell mellom standpunkt og eksamenskarakter i fag med muntlig eksamen, og mønsteret er dermed det samme som for norsk muntlig.

Det er generelt små forskjeller mellom offentlige og private skoler i forhold til muntlig eksamen, også her viser alle regresjonsanalysene at det som forklarer mest av differansen er hvilken standpunktkarakter eleven hadde, selv om også denne forklarer relativt mindre enn variabelen gjorde i analysene av skriftlige eksamener.

6 Konklusjoner

Vi har her gått igjennom og undersøkt gjennomsnittlig forskjell mellom standpunkt og eksamenskarakter i fem fag med skriftlig eksamen og fire fag med muntlig eksamen i videregående skole. Gjennomgående er det slik at avstanden mellom standpunkt og eksamenskarakter er større i fag med skriftlig eksamen enn i fag med muntlig eksamen.

Et viktig funn i undersøkelsen er at vi ser at standpunktkarakterene i fagene med skriftlig eksamen er forholdsvis konstant over tid, mens eksamenskarakterene varierer mer. Dermed er eksamenskarakterer den minst stabile av de to. Dette åpner for en rekke spørsmål knyttet til eksamen som målepunkt. Det er tidligere vist til at det er en forventning i samfunnet til at eksamen skal fungere som rettesnor og har en kalibrerende funksjon overfor læreres standpunktkaraktersetting (Prøitz & Borgen 2010, Prøitz 2013). Vi ser imidlertid her at begge former for karakterer varierer, men altså mest på eksamen. Det blir således nødvendig å stille spørsmål ved eksamens funksjon som objektivt målepunkt. Disse foreløpige resultatene indikerer også at sammenlikning av standpunkt- og eksamenskarakterer med forventninger eller et ideal om sammenfall ikke er uproblematisk.

Problemstillingen for dette prosjektet er om det er forskjeller mellom offentlige og private skoler i karaktersetting, men regresjonsanalyser av hvilke forhold som påvirker gjennomsnittsforskjellen mellom standpunkt og eksamenskarakter viser at den viktigste forklaringsfaktoren for forskjellen er nivået på standpunktkarakteren. Selv om effekten varierer noe fra fag til fag, finner vi hele tiden en tydelig tendens til regresjon mot gjennomsnittet, ved at de med best karakter går mest tilbake, mens de med dårligst karakter går fram. Dette funnet er noe det alltid må tas hensyn til når grupper skal sammenlignes.

I analysene finner vi imidlertid noen forskjeller mellom offentlige og private skoler, ved at gjennomsnittsforskjellen mellom standpunkt og eksamen er noe større for private skoler enn for offentlige, og at dette særlig gjelder de private allmenne skolene. I fem av ni fag finner vi større gjennomsnittsforskjell ved allmenne private skoler, mens vi i seks av ni fag ved de spesialiserte skolene finner mindre gjennomsnittsforskjell enn ved de offentlige skolene. Ved den siste kategorien skoler ser forklaringen ut til å være at det her settes relativt lave standpunktkarakterer i forhold til hva som oppnås ved eksamen, mens det motsatte synes å være tilfelle ved de allmenne private skolene.

Men det er også andre faktorer som spiller rolle i forhold til gjennomsnittsforskjell mellom standpunkt og eksamenskarakter, som gjelder uavhengig av om skolen er privat eller offentlig. Skolestørrelse ser ut til å ha en negativ effekt på karakterforskjellen mellom standpunkt og eksamen, særlig i fag med skriftlig eksamen. Det ser ut til å være en entydig ulempe å gå på svært små skoler, ved at gjennomsnittsforskjellen mellom standpunkt og eksamenskarakter er større der.

Tilsvarende finner vi at det i åtte av ni undersøkte fag er mindre gjennomsnittsforskjell mellom standpunkt og eksamen ved skoler som har høy inntakskvalitet, sammenlignet med skoler med normal inntakskvalitet, og dette gjelder også når det kontrolleres for hvilken standpunktkarakter eleven har.

Funnene i den kvantitative delen av prosjektet gir viktige innspill til det videre arbeidet i den kvalitative delen av prosjektet, i tråd med studiedesignet. Opplegget for den kvalitative delen av studien er å gjennomføre intervjuer ved fire offentlige og fire private skoler. I tillegg til å se etter forskjeller i praksis og rutiner i standpunktkaraktersetting mellom offentlige og private skoler vil vi også se etter forhold som kan bidra til å forklare eller tolke mange av de andre funnene vi har kommet frem til i den kvantitative delen av undersøkelsen. Vi må derfor også se etter forklaringer som kan kaste lys over hvorfor det er forskjeller mellom små skoler og større skoler og hvorfor det er forskjeller mellom skoler som har høy inntakskvalitet og skoler med normal inntakskvalitet.

Referanser

- Galloway, T.A., Kirkebøen, L.J. og Rønning, M (2011): *Karakterpraksis i grunnskoler. Sammenheng mellom standpunkt – og eksamenskarakter*. SSB rapport 4/2011. Oslo: Statistisk sentralbyrå.
- Galton, Francis (1886): Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246–263.
- Grøgaard, Jens B. (2002): Integrerte eller segregerte undervisningsopplegg i videregående opplæring: hvilke gir best resultater? *Tidsskrift for ungdomsforskning*, 2002, 2 (2):83–108
- Grøgaard, Jens B. (2012): *Hva kjennetegner barneskoler som oppnår høy skår på nasjonale prøver? Delrapport 5 fra prosjektet 'Ressurser og resultater i grunnopplæringen'*. NIFU rapport 38/2012. Oslo: NIFU
- Opheim, Vibeke, Jens B. Grøgaard & Terje Næss (2010): *De gamle er eldst? Betydning av skoleressurser, undervisningsformer og læringsmiljø for elevenes prestasjoner på 5., 8. og 10. trinn i grunnopplæringen*. NIFU rapport 34/2010. Oslo: NIFU
- Prøitz, Tine Sophie & Borgen, Jorunn Spord (2010): *Rettferdig standpunktbedømming. Det (u)muliges kunst?* NIFU STEP rapport 16/2010. Oslo: NIFU STEP
- Prøitz, Tine Sophie (2013): Variations in grading practice – subjects matter. *Education Inquiry*, 4(3): 555-575
- Wiborg, Øyvind, Clara Åse Arnesen, Jens B. Grøgaard, Liv Anne Støren, Vibeke Opheim (2011): *Elevers prestasjonsutvikling – hvor mye betyr skolen og familien? Andre delrapport fra prosjektet «Ressurser og resultater»*. NIFU rapport 35/2011. Oslo: NIFU

Offentlige dokumenter:

Forskrift til opplæringsloven 2013, <http://lovdata.no/dokument/SF/forskrift/2006-06-23-724>

Medieoppslag:

- Aftenposten 05.02.2013: *Svenske privatskoler setter bedre karakterer enn offentlige skoler*
<http://www.aftenposten.no/nyheter/iriks/Svenske-privatskoler-setter-bedre-karakterer-enn-offentlige-skoler-7111072.html>
- Aftenposten 07.02.2013: *Kristin Halvorsen vil undersøke karakterforskjeller*
<http://www.aftenposten.no/nyheter/iriks/Kristin-Halvorsen-vil-undersoke-karakterforskjeller-7113343.html>
- Aftenposten 15.02.2013: *Se om skolen din er streng eller snill med karakterene*
<http://www.aftenposten.no/nyheter/iriks/Se-om-skolen-din-er-streng-eller-snill-med-karakterene-7119112.html>
- Stavanger Aftenblad 19.02.2013: *Størst karakterforskjell hos de private skolene*
<http://www.aftenbladet.no/nyheter/okonomi/jobb/Storst-karakterforskjell-hos-de-private-skolene-3125366.html>

Vedleggstabeller til kapittel 3

		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)	
Kjønn	Gutter	Gjennomsnitt	3,47	3,00	-0,47	67948
		Standardavvik	1,041	1,094	0,962	
Jenter	Gjennomsnitt	3,83	3,26	-0,58	89676	
		Standardavvik	1,001	1,072	0,958	
		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)	
Alder	19 og yngre	Gjennomsnitt	3,72	3,21	-0,52	135961
		Standardavvik	1,024	1,080	0,956	
20	Gjennomsnitt	3,37	2,81	-0,55	11056	
		Standardavvik	1,073	1,094	0,986	
21	Gjennomsnitt	3,29	2,66	-0,63	4489	
		Standardavvik	1,010	1,031	0,975	
22+	Gjennomsnitt	3,56	2,81	-0,75	6118	
		Standardavvik	1,042	1,066	0,994	
		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)	
Kompetansetype	Samfunnsfag, økonomi, språk	Gjennomsnitt	3,72	3,29	-0,43	51479
		Standardavvik	0,987	1,058	0,956	
Realfag	Gjennomsnitt	4,09	3,56	-0,54	34816	
		Standardavvik	0,975	1,034	0,946	
Formgivingsfag	Gjennomsnitt	3,62	3,07	-0,55	3322	
		Standardavvik	1,016	1,040	0,948	
Musikk	Gjennomsnitt	3,97	3,48	-0,49	4289	
		Standardavvik	0,999	1,070	1,001	
Dans	Gjennomsnitt	3,85	3,29	-0,56	1096	
		Standardavvik	0,951	1,064	1,012	
Drama	Gjennomsnitt	3,96	3,45	-0,51	1712	
		Standardavvik	0,992	1,077	1,041	
Idrettsfag	Gjennomsnitt	3,53	2,95	-0,57	13841	
		Standardavvik	0,986	0,992	0,946	
Media og kommunikasjon	Gjennomsnitt	3,72	3,15	-0,57	8214	
		Standardavvik	0,992	1,045	1,005	
Naturbruk	Gjennomsnitt	3,41	2,60	-0,81	1508	
		Standardavvik	1,068	1,034	0,959	
Påbygg uten studiekompetanse	Gjennomsnitt	3,18	2,57	-0,62	31448	
		Standardavvik	0,993	0,984	0,951	
Påbygg etter studiekompetanse	Gjennomsnitt	3,53	2,83	-0,70	5099	
		Standardavvik	0,983	1,006	0,963	
Annet	Gjennomsnitt	3,93	3,23	-0,70	800	
		Standardavvik	1,023	1,113	1,034	

		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)
Skoleslag					
Offentlige	Gjennomsnitt	3,67	3,15	-0,51	143991
	Standardavvik	1,034	1,088	0,957	
Private kristne	Gjennomsnitt	3,97	3,28	-0,69	5413
	Standardavvik	0,974	1,094	0,954	
Private allmenne	Gjennomsnitt	3,75	2,97	-0,78	4690
	Standardavvik	1,078	1,097	1,009	
Private spesialiserte	Gjennomsnitt	3,65	3,07	-0,58	3530
	Standardavvik	1,021	1,063	0,994	

		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)
Antall avgangselever					
Færre enn 50	Gjennomsnitt	3,58	2,90	-0,68	9820
	Standardavvik	1,057	1,084	0,981	
50 - 99	Gjennomsnitt	3,56	2,98	-0,58	24379
	Standardavvik	1,067	1,082	0,970	
100 - 199	Gjennomsnitt	3,69	3,18	-0,52	73654
	Standardavvik	1,020	1,087	0,957	
200 - 299	Gjennomsnitt	3,75	3,24	-0,52	37653
	Standardavvik	1,026	1,085	0,954	
300 og flere	Gjennomsnitt	3,68	3,24	-0,44	12118
	Standardavvik	1,032	1,065	0,963	

		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)
År					
2010	Gjennomsnitt	3,71	3,09	-0,62	38126
	Standardavvik	1,026	1,055	0,953	
2011	Gjennomsnitt	3,68	3,10	-0,57	39815
	Standardavvik	1,033	1,141	0,970	
2012	Gjennomsnitt	3,65	3,13	-0,52	39239
	Standardavvik	1,036	1,073	0,974	
2013	Gjennomsnitt	3,68	3,26	-0,41	40444
	Standardavvik	1,041	1,074	0,936	

		Norsk hovedmål standpunkt	Norsk hovedmål eksamen	Differanse norsk hovedmål	Antall (N=)
År					
Ordinær skole	Gjennomsnitt	3,60	3,05	-0,55	124029
Snitt GS < 4,58	Standardavvik	1,036	1,074	0,964	
Skole med høy inntakskvalitet	Gjennomsnitt	3,98	3,51	-0,47	33595
Snitt GS 4,58 +	Standardavvik	0,969	1,064	0,948	

Nordisk institutt for studier av
innovasjon, forskning og utdanning

Nordic Institute for Studies in
Innovation, Research and Education

www.nifu.no