



Kartlegging blant eksternevurderere av nasjonale prøver

**Rapport fra MMI
v/Håkon Kavli og Wenche Berntsen
August 2005**

Forord

På oppdrag fra Utdanningsdirektoratet har MMI Univero gjennomført denne undersøkelsen blant eksternvurdererne av de nasjonale prøvene i 2005.

Eksternvurdererne har bedømt prøver i som er gjennomført på 4., 7., 10. og 11. trinn. De har vurdert oppgaver innen lesing, skriving, engelsk skriving og matematikk.

Undersøkelsen er et ledd i Utdanningsdirektoratets oppsummering av erfaringene med de nasjonale prøvene. På sikt skal erfaringene danne grunnlag for å forbedre prosedyrene for gjennomføring og vurdering av de nasjonale prøvene.

Undersøkelsen er en oppfølging av et prosjekt der MMI har kartlagt hvordan rektorer, lærere og elever vurderer gjennomføringen av de nasjonale prøvene i 2005.

I MMI har Wenche Berntsen og Håkon Kavli vært ansvarlig for gjennomføring, analyse og rapportering. Kontaktperson ved Utdanningsdirektoratet har vært Annette Qvam.

Oslo, 31. august 2005

Metode og gjennomføring

Det ble foretatt en lignende kartlegging blant eksternvurderere også i 2004. Den gang var det et annet institutt som sto for gjennomføringen, og det ble brukt *postalt* spørreskjema. Problemstillingene fra 2004 er langt på vei de samme i denne undersøkelsen, men i år er ordlyden og svaralternativene i de fleste spørsmålene vesentlig justert i henhold til skjemaet som ble brukt mot lærerne i årets evalueringsundersøkelse. Det er også tatt inn nye spørsmål siden sist. Ulik metodebruk samt justeringer i spørreskjema gjør at sammenligning med tidligere resultater blir usikre. Det er fordi vi ikke kan vite om eventuelle observerte forskjeller er substansielle eller må tilskrives metodiske forhold. Årets utfall vil likevel bli sammenlignet med tall fra 2004 der vi anser det forsvarlig, men vi oppfordrer til forsiktig tolkning av endringene fra i fjor.

I år er dataene samlet inn elektronisk via web. MMI har mottatt lister over e-postadresser til alle eksternvurderere som har vurdert prøver. Samtlige fikk tilsendt en e-post med link til det elektroniske spørreskjemaet. Totalt sendte vi ut 411 henvendelser. Vi har mottatt 267 svar, og det gir en svarprosent på 65%. Dette er noe lavere enn fjorårets respons på 76%. Forskjellene i svarprosent kan tilskrives forskjellige datainnsamlingsmetoder, og ikke minst at feltperioden i år falt sammen med fellesferien. Mange av respondentene fikk tilsendt elektronisk skjema på sin jobbmail-adresse. Det er grunn til å tro at jobbrelatert e-post sjekkes med lavere frekvens i sommerferien. Vi anser likevel responsraten som tilfredsstillende.

Feltarbeidet foregikk i perioden 16.06.05 – 22.08.05. Det ble sendt ut to påminnelser i løpet av perioden.

Metode og gjennomføring

I tabellverket til denne rapporten finnes svarfordelingen på samtlige spørsmål i den rekkefølge spørsmålene er stilt. Hver tabell viser svarfordelingen totalt og etter om man er eksternvurderer 1 eller 2, kjønn, landsdel, årstrinn og type prøve som er vurdert. På neste side finnes en utvalgsbeskrivelse med hensyn på disse bakgrunnsvariablene. (Utvalgsbeskrivelse fremgår også av tabellene 2-7 i tabellverket.)

I og med at dette er en totalundersøkelse der vi har henvendt oss til alle eksternvurdere som har vurdert prøver, er resultatene i liten grad beheftet med usikkerhet utover den som knytter seg til det faktum at ikke alle har svart. Det sentrale spørsmålet vedrørende bortfallet er i hvilken grad det forekommer systematiske forskjeller mellom de som har svart og de som ikke har svart. Vi har kontrollert for utvalgsskjevheter, og finner ingen systematiske avvik mellom populasjonen og de som har svart. Dette fremgår av tabellen på neste side.

Kontroll for utvalgsskjevheter: Eksternvurderernes fordeling på sentrale bakgrunnsvariabler i populasjonen og i nettoutvalget

	Populasjon	Utvalg
Landsdel		
Oslo	5,6%	6%
Østlandet ellers	41%	41%
Vestlandet	26,8%	25%
Møre og Romsdal/ Trøndelag	15,8%	16%
Nord-Norge	10,9%	11%
Årstrinn		
4. Trinn	7,5%	10%
7. Trinn	33%	36%
10. Trinn	32%	30%
11. Trinn	28%	23%
Type prøve		
Lesing	19%	21%
Skriving	33%	32%
Engelsk skriving	33%	29%
Matematikk	15%	17%
Type eksternvurderer *		
Eksternvurderer 1	67%	68%
Eksternvurderer 2	33%	32%

Fordelingen i utvalget avviker såpass lite fra fordelingen i populasjonen at frafallet ser ut til å være tilfeldig. Statistisk vekting av resultatene er derfor ikke nødvendig, og ville ikke hatt noen innvirkning på svarfordelingen.

Alle resultatene som det vises til i rapporten er derfor uveide tall. Fordi datagrunnlaget ikke er beheftet med vesentlige utvalgsskjevheter, kan vi feste lit til at resultatene er representative for hele populasjonen av eksternvurderere.

* Ang. type eksternvurderer: Prøvene i skriving og engelsk skriving er åpne, og vurderes av to eksterne i tillegg til lærer. Begge eksterne vurderer prøven uavhengig av hverandre og sender resultatene til fagmiljøene. Skolen får bare tilbakemelding fra eksternvurderer 1.

Metode og gjennomføring

Generelt er det slik at de som svarer gjerne har større interesse for undersøkelsens tema enn de som unnlater å svare. Vi må m.a.o. forvente at de mest engasjerte eksternvurdererne har hatt en noe større tilbøyelighet til å svare enn de som ikke er opptatt av å fremme sine synspunkter. Vi vil likevel understreke at eventuelle forskjeller mellom de to gruppene må være svært store for at resultatene skulle ha blitt vesentlig annerledes enn de foreliggende.

Noen av spørsmålene i denne undersøkelsen er også brukt i kartleggingen blant rektorer, lærere og elever. Det er selvsagt eksternvurderernes synspunkter som er i fokus i denne rapporten, men vi tillater oss å gjøre noen sammenligner med lærernes svar der vi finner det relevant.

Før vi presenterer resultatene skal vi se nærmere på hvordan eksternvurdererne fordeler seg på de ulike kombinasjoner av årstrinn og type prøve.

	Antall	Prosent
LESING TOTALT	55	21%
Lesing 4. trinn	8	3%
Lesing 7. trinn	16	6%
Lesing 10. trinn	16	6%
Lesing 11. trinn	15	6%
SKRIVING TOTALT	85	32%
Skriving 4. trinn	11	4%
Skriving 7. trinn	31	12%
Skriving 10. trinn	30	11%
Skriving 11. trinn	13	5%
ENGELSK TOTALT	78	29%
Engelsk skriving 7. trinn	32	12%
Engelsk skriving 10. trinn	27	10%
Engelsk skriving 11. trinn	19	7%
MATEMATIKK TOTALT	45	18%
Matematikk 4. trinn	7	3%
Matematikk 7. trinn	16	6%
Matematikk 10. trinn	7	3%
Matematikk 11. trinn	15	6%
TOTAL	263*	100%

Tabellen viser at antallet observasjoner blir svært lite når vi bryter dem ned på kombinasjoner av trinn og prøvetype.

I rapporten konsentrerer vi oss primært om resultater brutt ned på et mindre detaljert nivå, slik at vi studerer undergruppene på henholdsvis årstrinn og prøvetype hver for seg. Likevel vil vi i flere tilfeller kommentere resultater for undergrupper som representerer kombinasjoner av trinn og prøvetype. Husk imidlertid at det i mange av disse tilfellene er et begrenset antall svar som ligger bak resultatet.

***) Det er 4 personer som ikke har latt seg plassere i tabellen pga manglende opplysninger**

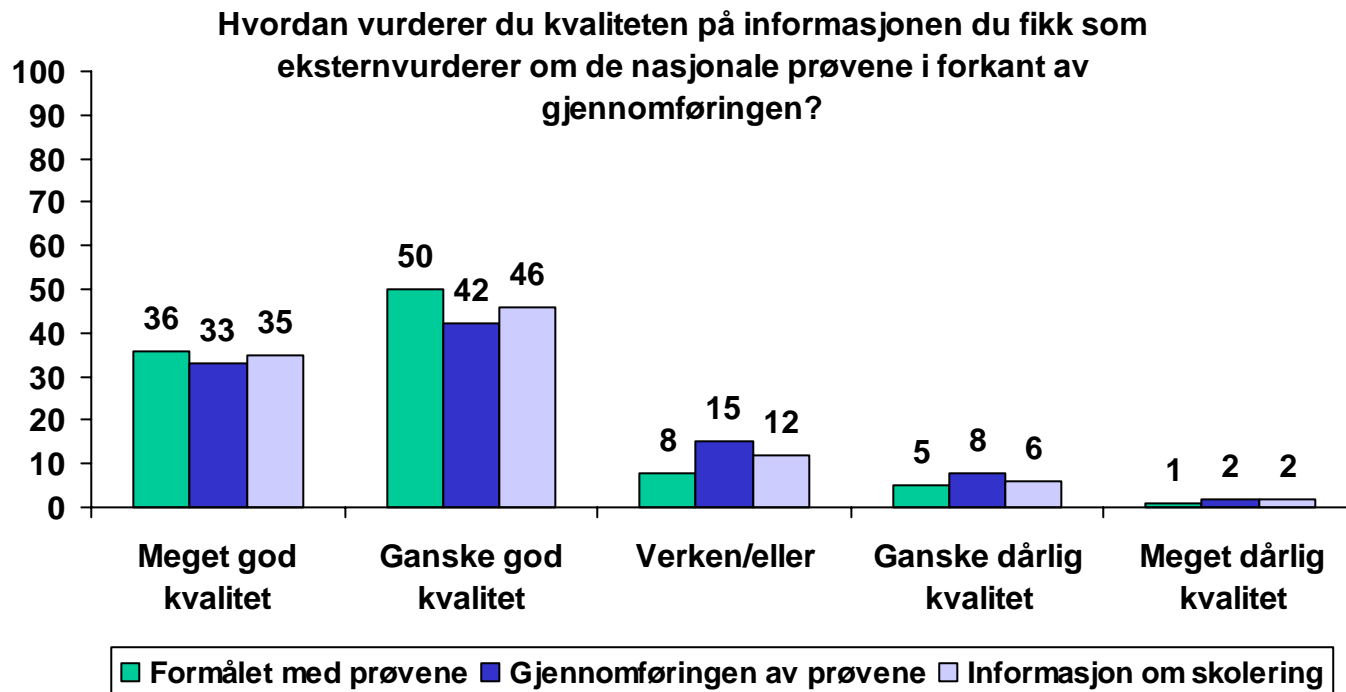
Informasjon om gjennomføring av prøvene

Spørsmål 8

Et klart flertall mener det var god kvalitet på den informasjonen de fikk i forkant av gjennomføringen av de nasjonale prøvene. Også i fjor var det utbredt tilfredshet med denne informasjonen.

Eksternvurdererne er mest fornøyd med informasjonen om *formålet* med prøven (86%), men de har også et godt inntrykk av informasjonen om *gjennomføringen* av prøvene og *skoleringen*. Det er verdt å merke seg at rundt 1 av 3 vurderer kvaliteten på alle tre informasjonsområder som ”meget god”.

Eksternvurdererne er klart mer fornøyd med informasjonen enn lærerne generelt.



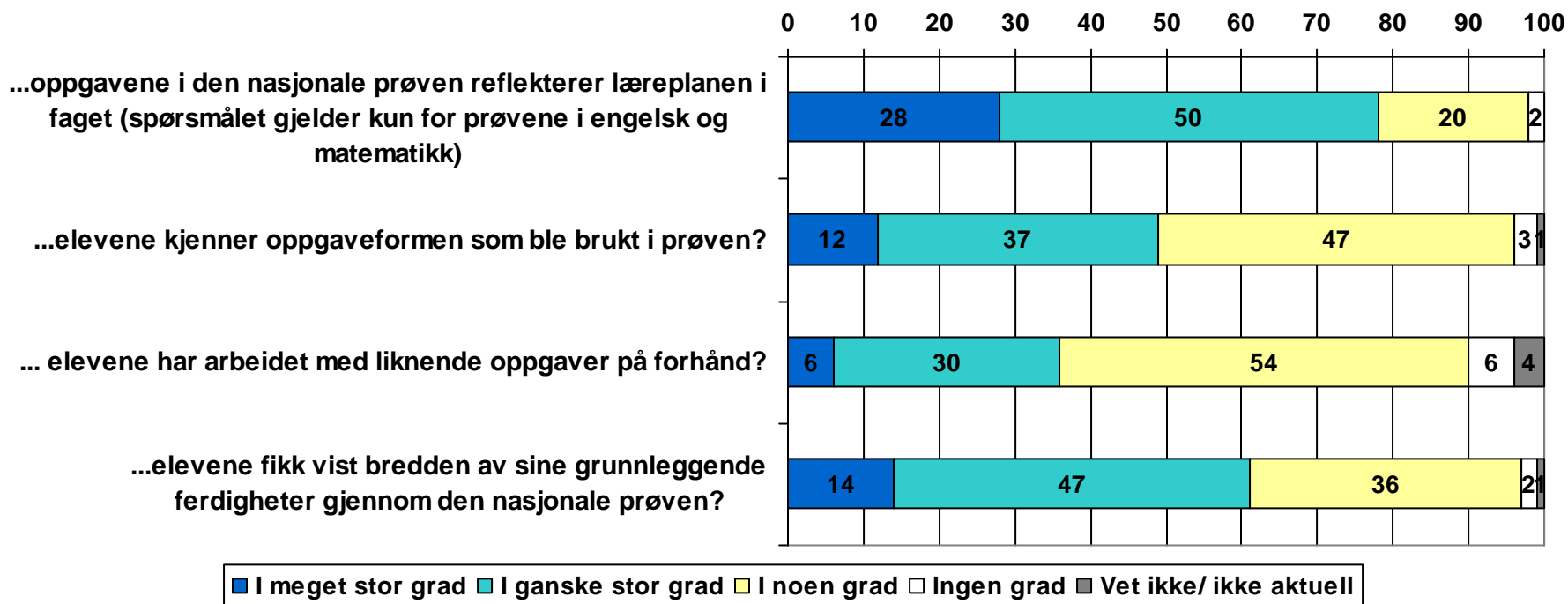
Innhold og utforming av prøvene

Spørsmål 9-12



Denne delen av undersøkelsen starter med spørsmål om i hvilken grad oppgavene reflekterer læreplanen, og i hvilken grad elevene kjenner oppgaveformen, har arbeidet med liknende oppgaver fra før, og om de fikk vist sine grunnleggende ferdigheter. Vi skal først se på hovedresultatene før vi studerer det enkelte spørsmål i detalj.

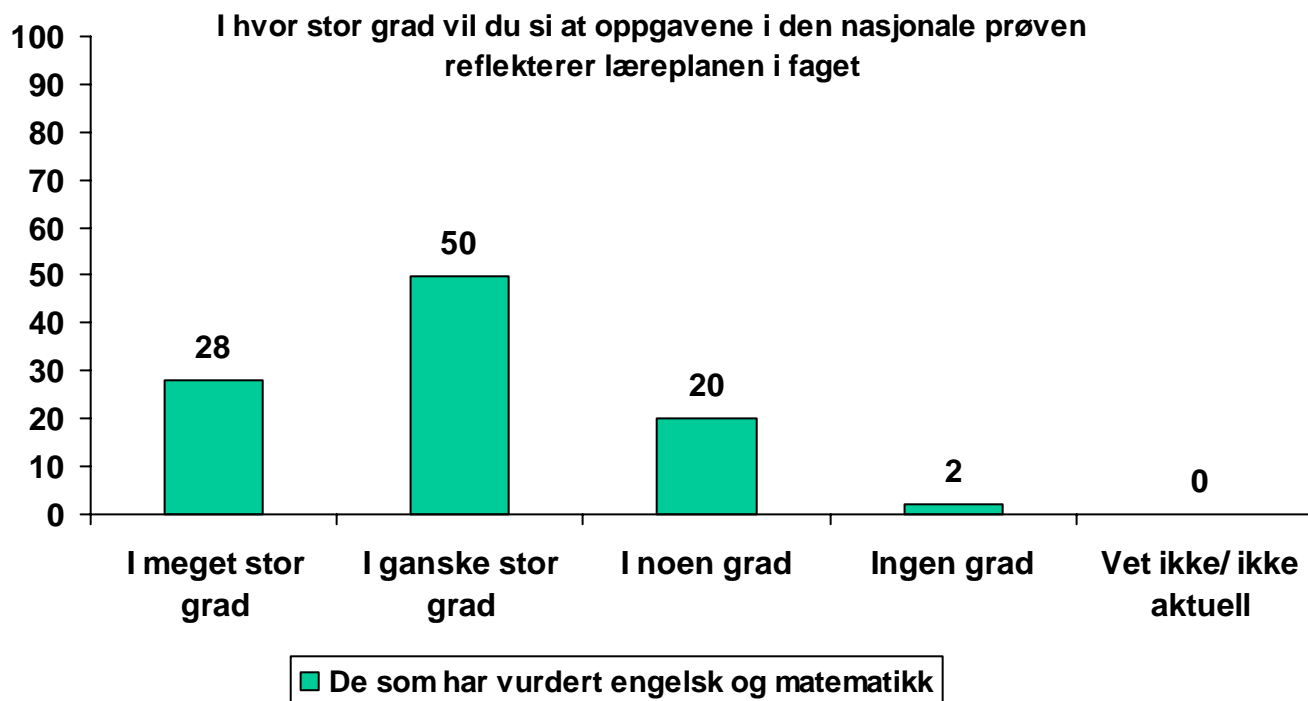
I hvor stor grad vil du si at...?



Det kan nevnes at eksternvurdererne gjennomgående har en mer positiv oppfatning enn lærerne på disse spørsmålene.

Et klart flertall (78%) av de som har vurdert oppgaver i engelsk og matematikk, opplever at prøven er i tråd med læreplanen.

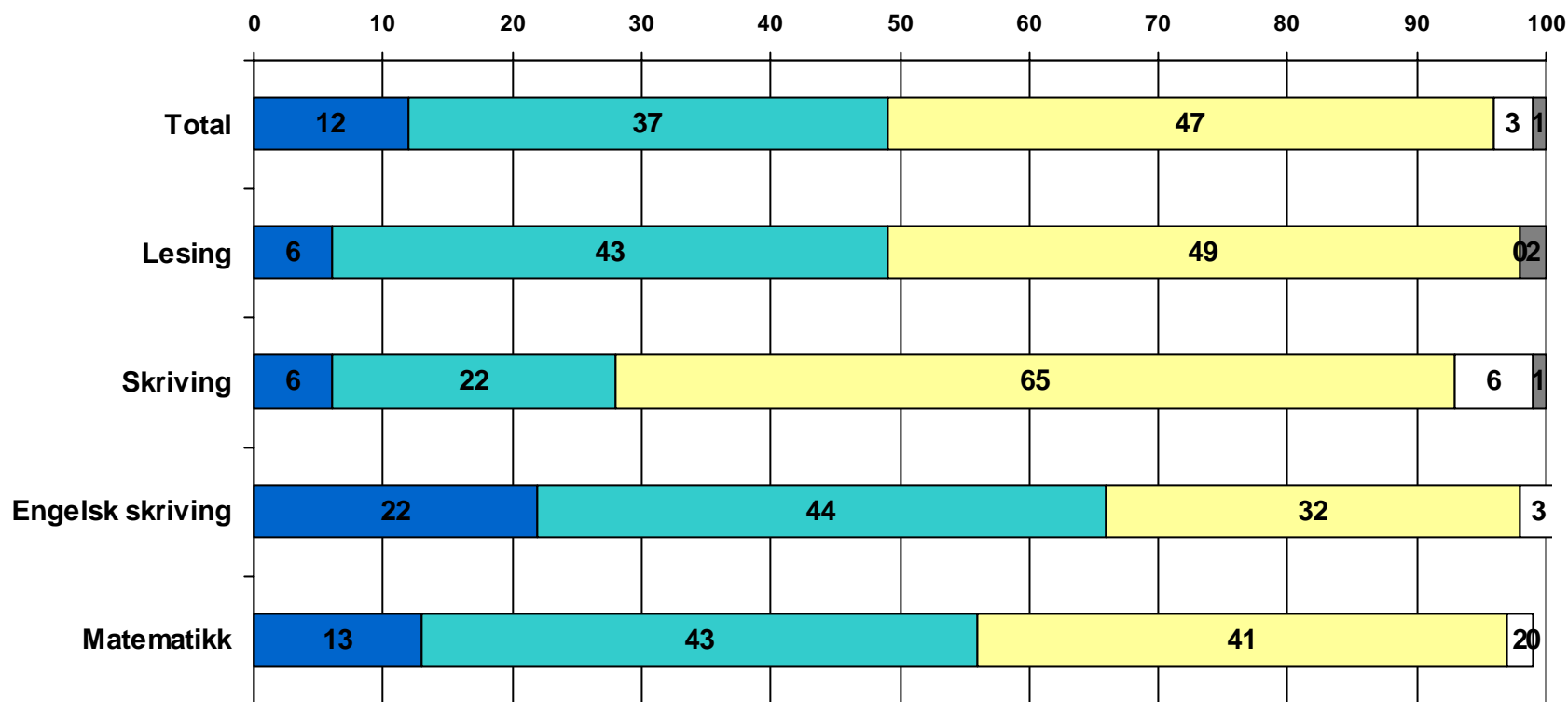
Eksternvurdererne synes å mene at prøvene i matematikk (unntatt matematikkprøvene på 11. trinn) reflekterer læreplanen bedre enn prøvene i engelsk.



Til sammen 49% mener at elevene *i stor grad* kjenner oppgaveformen som ble brukt i prøven. Omtrent like mange (47%) mener at elevene *i noen grad* kjenner oppgaveformen.

Eksternvurdererne opplever at elevene har klart bedre kjennskap til oppgavetyperne i engelsk (særlig på 10. og 11. trinn) og matematikk (særlig 4. trinn) enn til de øvrige oppgavetyperne (særlig skrivning).

I hvor stor grad vil du si at elevene kjenner oppgaveformen som ble brukt i prøven?

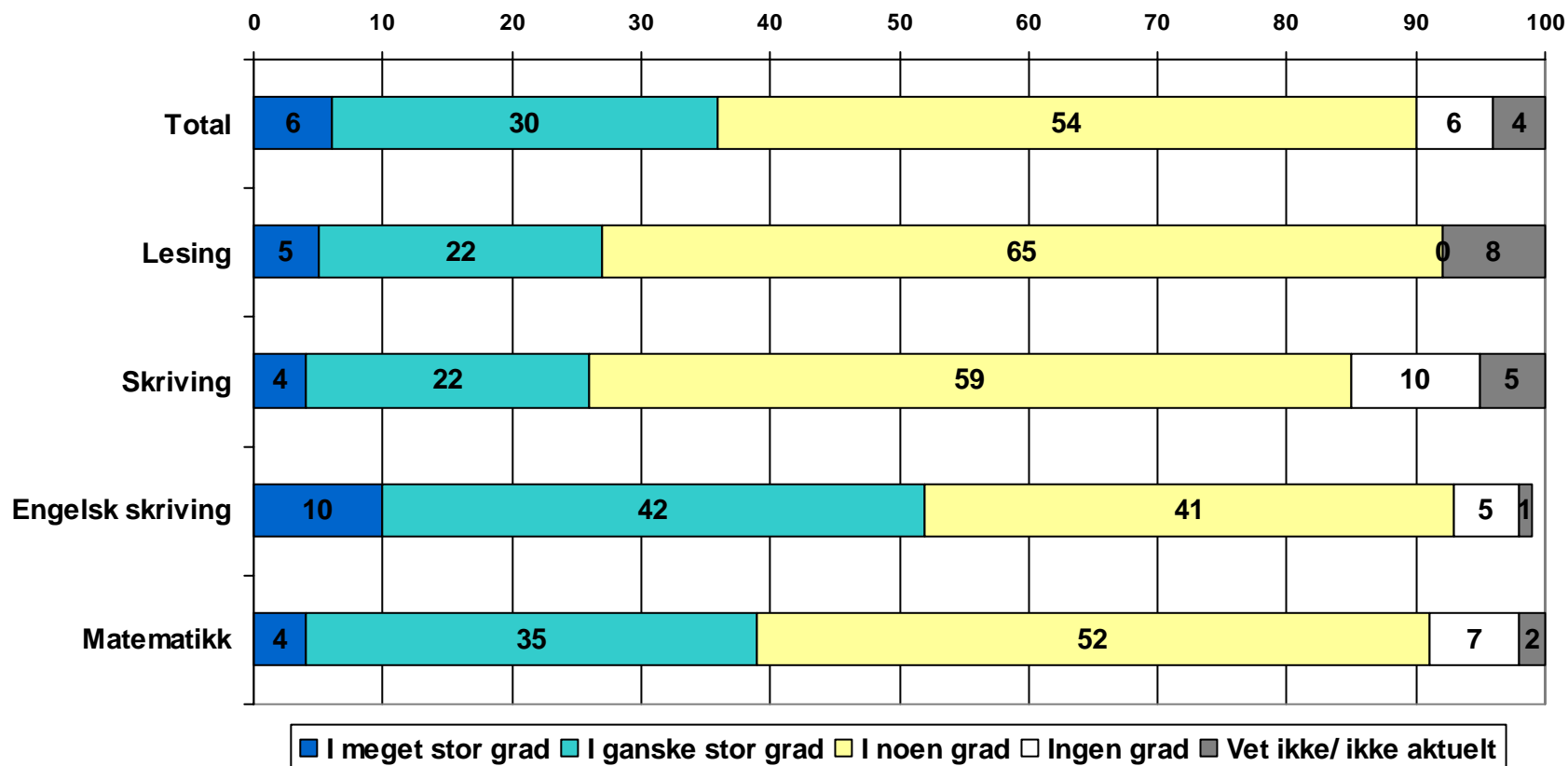


■ I meget stor grad ■ I ganske stor grad ■ I noen grad □ Ingen grad ■ Vet ikke/ ikke aktuelt

Det vanligste inntrykket er at elevene i noen grad har erfaring med liknende oppgaver fra før.

De som har vurdert engelskprøver skiller seg signifikant fra gjennomsnittet med en større andel som mener elevene i stor grad har forhåndserfaringer (særlig engelsk skrijving på 10. trinn). Denne holdningen er også utbredt blant de som har vurdert lesing på 4. trinn og matematikk på 4. og 7. trinn. Holdningen er minst utbredt blant de som har vurdert lesing (særlig 10. og 11. trinn) og skrijving (særlig 4. og 10. trinn).

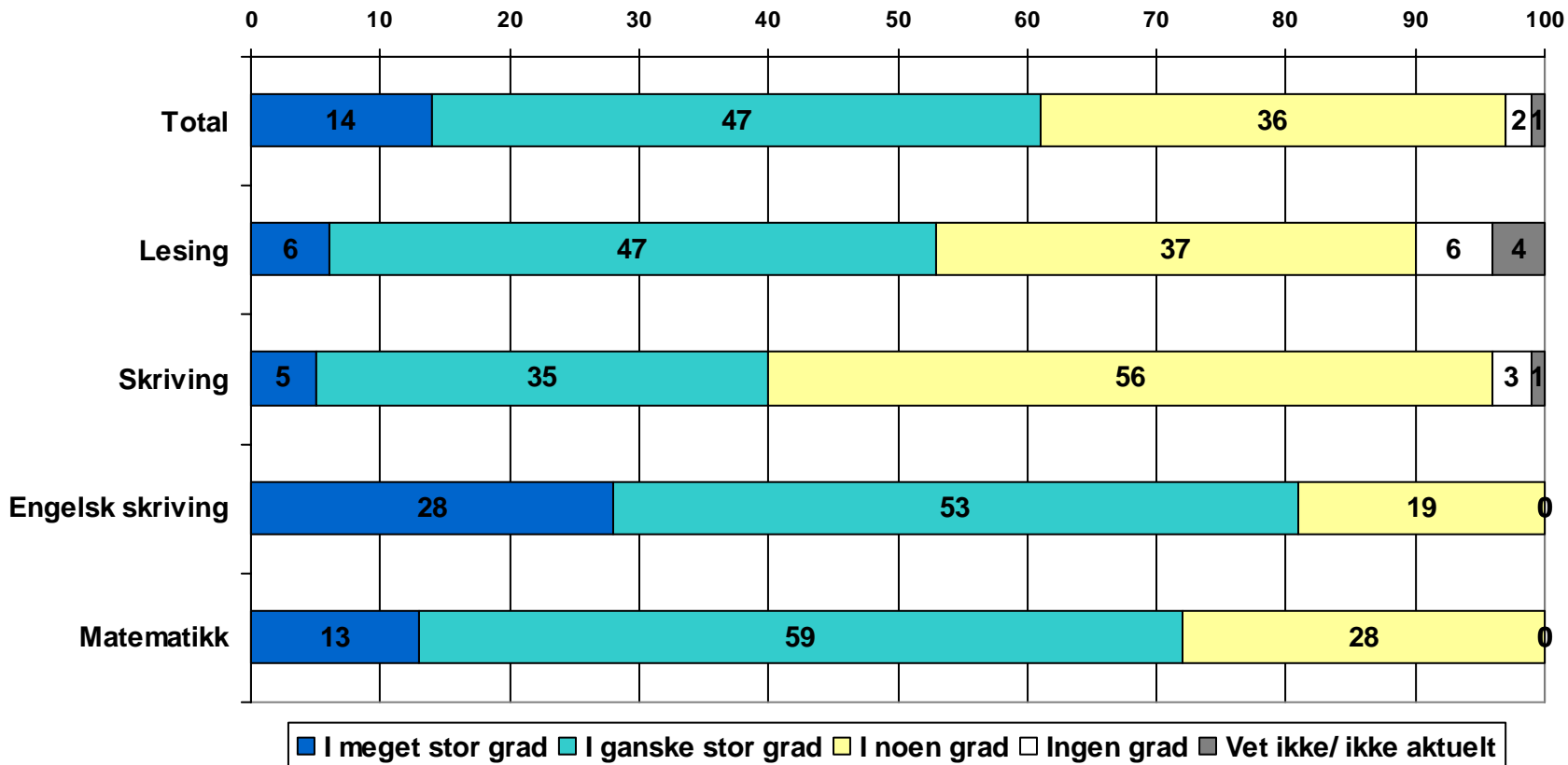
I hvor stor grad vil du si at elevene har arbeidet med liknende oppgaver på forhånd?



61% mener at elevene i stor grad fikk vist bredden av sine grunnleggende ferdigheter.

Eksternvurderer 1 (65%) er mer tilbøyelig enn eksternvurderer 2 (52%) til å mene at elevene fikk vist sine ferdigheter. De som har vurdert prøver på 7. trinn (74%), samt de som har vurdert engelsk (81%) og matematikk (72%) har også en klarere oppfatning enn øvrige spurte av at elevene i stor grad fikk vist sine ferdigheter. De som har vurdert skriving (4., 10. og 11. trinn) og de som har vurdert lesing på 4. trinn er noe mer forbeholdne enn gjennomsnittet.

I hvor stor grad vil du si at elevene fikk vist bredden av sine grunnleggende ferdigheter gjennom den nasjonale prøven?



Årets tall versus 2004-tall:

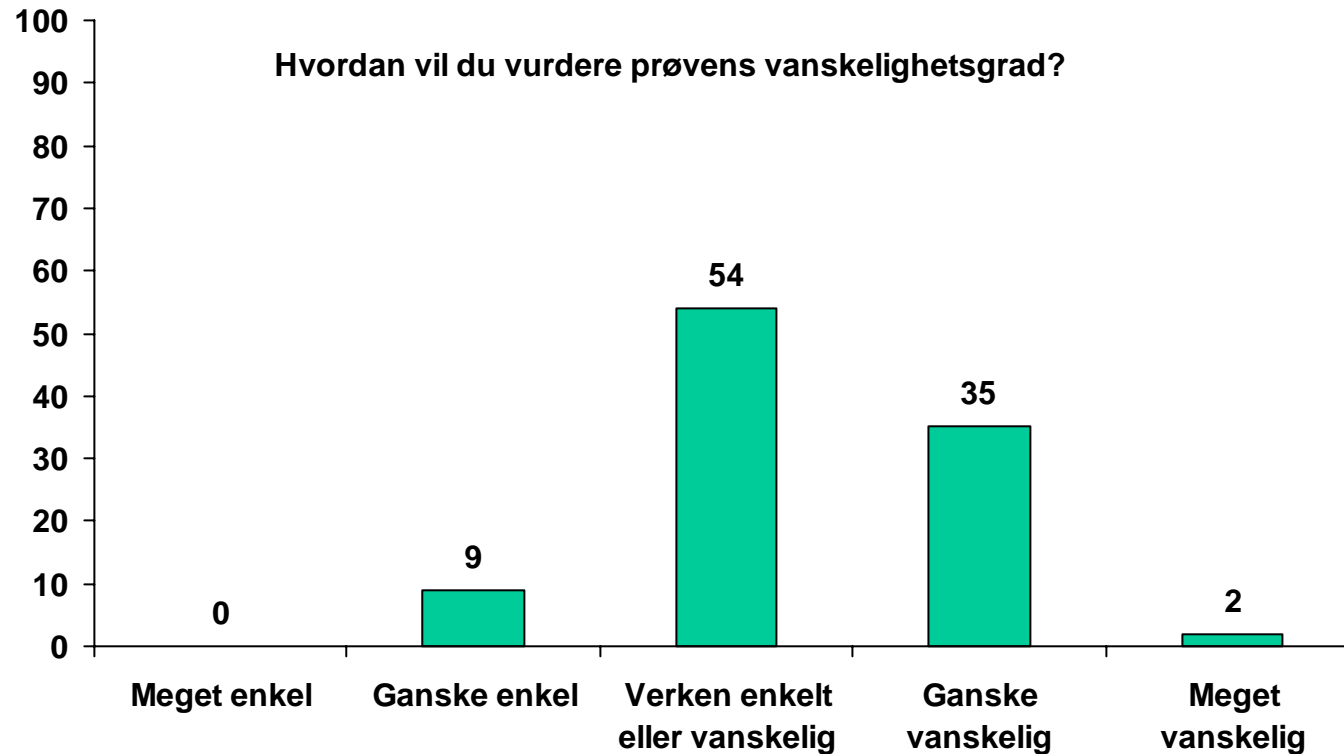
I 2004 ble det brukt en annen svarskala. Da mente hhv. 24% og 66% at elevene i svært stor eller noen grad fikk vist sine ferdigheter.

Også i fjor så man at de på 4. trinn var mer skeptiske. Man så også flere kritiske blant de som hadde vurdert engelsk på 10. trinn. Dette ser vi ikke i år.

Drøyt halvparten (54%) mener at prøvenes vanskelighetsgrad ligger ”midt på treet”. Forøvrig er det flere som mener at prøvene var vanskelige (37%) enn at de var enkle (9%).

Den største andelen som opplever at prøvene var vanskelig finner vi blant de som har vurdert skrijving på 10. og 4. trinn, samt lesing på 4., 7. og 11. trinn. De som har vurdert engelskprøver svarer i større grad enn andre at prøvene verken var enkle eller vanskelige.

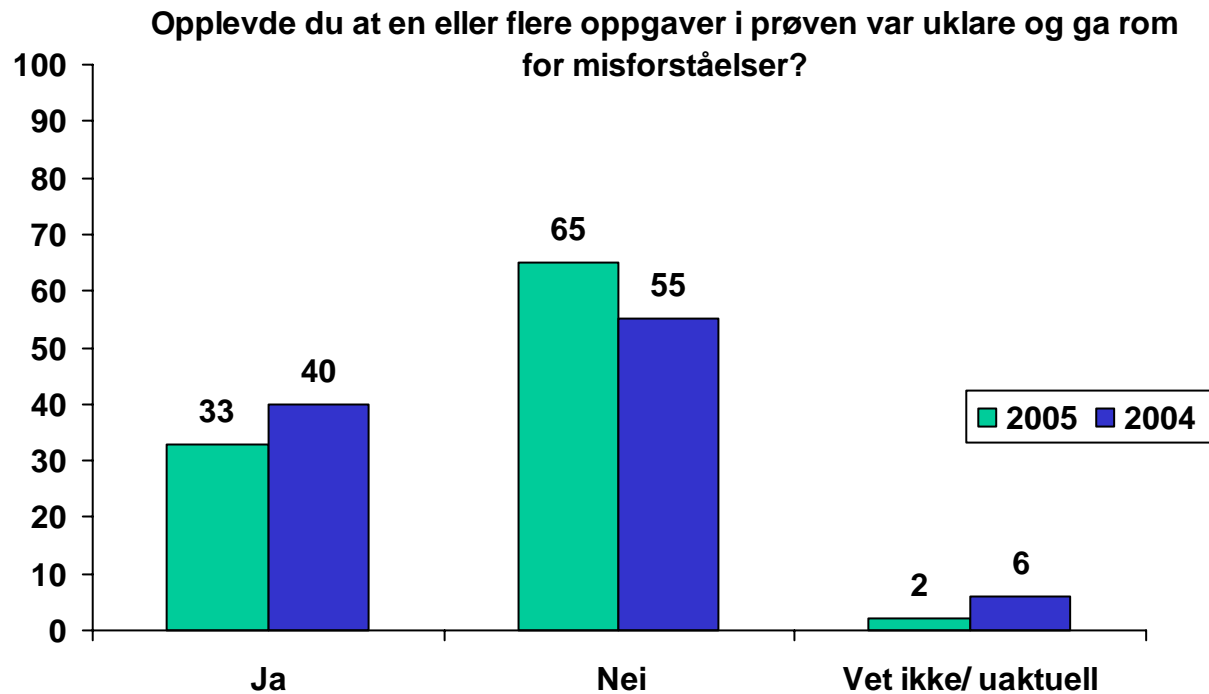
Lærerne synes i noe større grad enn eksternevurdererne at prøvene var vanskelige



Andelen som mener at det var uklarheter og rom for misforståelser i oppgavene er redusert fra 40% i fjor til 33% i år. Nedgangen er ikke signifikant, men resultatet peker i positiv retning.

Det oppleves å være mer uklarheter i oppgaver på 4. trinn enn på øvrige trinn, og mer i oppgaver innen lesing (særlig 11. trinn) og matematikk (særlig 4. og 7. trinn) enn i oppgaver innen skriving og engelsk. De som har vurdert 10. trinn rapporterer om mindre uklarheter enn gjennomsnittet.

Det ser ut til at eksternvurdererne har opplevd noe mindre uklarheter enn lærerne.



Skolering

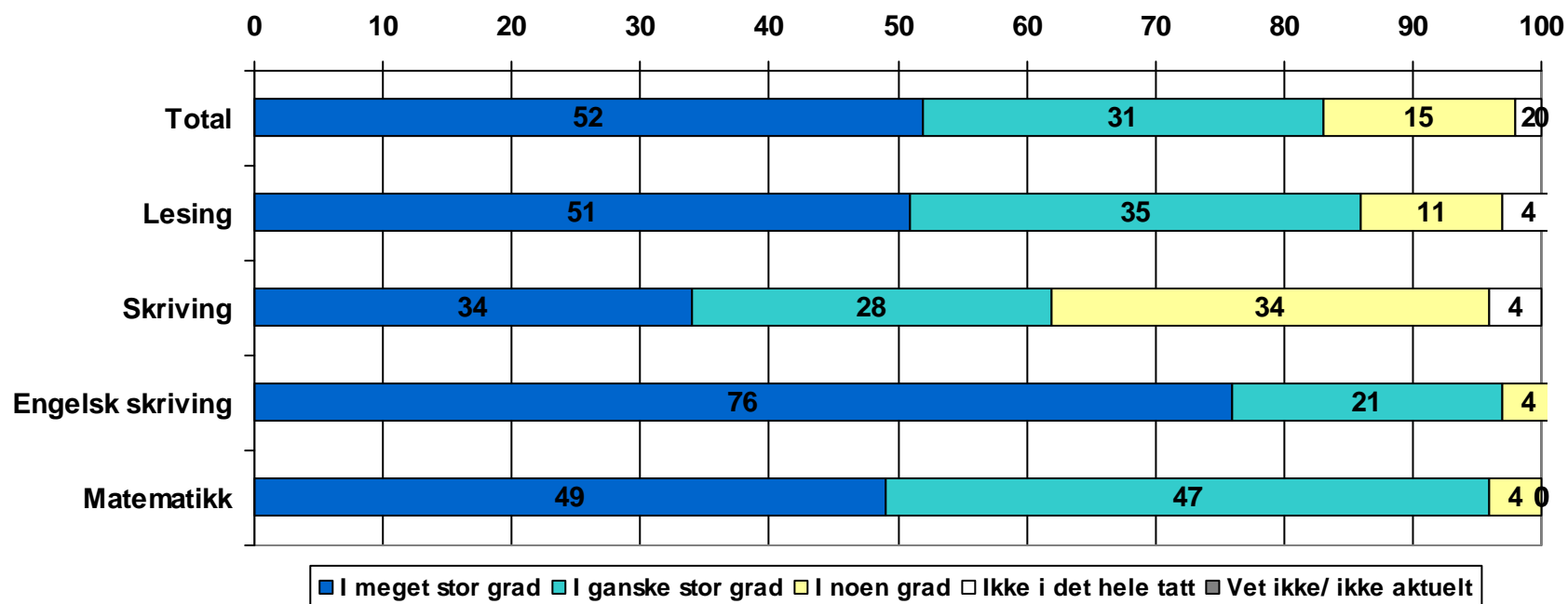
Spørsmål 13-20



De aller fleste har hatt nytte av skoleringen fra fagmiljøene. Det harmonerer med funnet fra i fjor.

Innen samtlige undergrupper av spurte er det et klart flertall som sier at de i stor grad har hatt nytte av skoleringen. Det kan nevnes at de som har vurdert prøver på 7. trinn synes å ha hatt størst utbytte, mens de som har vurdert prøver på 4. trinn er noe mer reserverte. De som har vurdert prøver i skrivning er også noe mer reserverte enn de som har vurdert andre prøver. Derimot rapporteres det om svært stor nytte blant de som har vurdert engelskprøver.

I hvor stor grad har du hatt nytte av skoleringen fra fagmiljøene?



Nytteverdien av skoleringen er også målt ved hjelp av noen utsagn som eksternvurdererne skulle si seg enig eller uenig i. Nedenfor kommenterer vi hovedresultatet, og på de neste sidene vises resultatene grafisk.

For samtlige påstander er det klart flere enige enn uenige. De fire første påstandene ble også testet i fjor, og det var en overvekt av enige også den gang.

Med andre ord mener majoriteten at skoleringen har gitt dem:

- grunnlag for å gi opplæring til lærerne (92% enige)
- en mal for vurdering av den aktuelle prøven (89% enige)
- kompetanse i vurderingsarbeid generelt (81% enige)
- innsikt i kompetanseprofilene og bruken av dem (74% enige)
- grunnlag for veiledning av lærerne i pedagogisk bruk av prøveresultatene (58% enige)

Vi finner mest uenighet (30%) for påstanden som er ny av året: ”Skoleringen ga grunnlag for veiledning av lærerne i pedagogisk bruk av prøveresultatene”.

- Det er særlig de som har vurdert oppgaver i skrivning som trekker resultatet i negativ retning.

Vi finner forskjeller etter årstrinn:

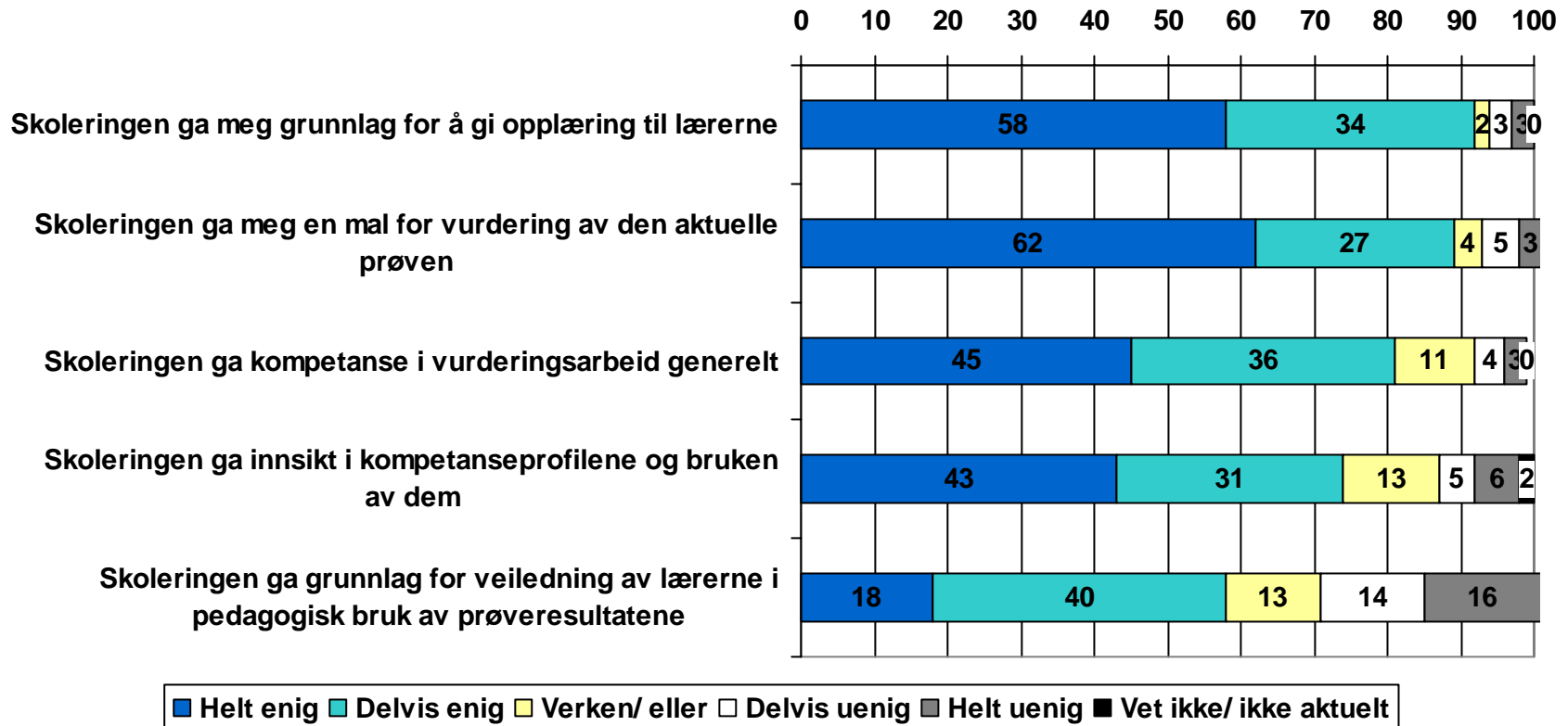
- det er generelt noe mer skepsis på 4. trinn, særlig når det gjelder opplæringsgrunnlag og kompetanse i vurderingsarbeid.
- 7. trinn utmerker seg ved at man fant særlig nytte med tanke på veiledningsgrunnlag for lærerne i pedagogisk bruk av prøveresultatene.

Vi finner forskjeller etter type prøve:

- De som har vurdert oppgaver i engelsk og matematikk, er gjennomgående mer positive til nytten av skoleringen.
- Motsatt ser vi at de som har vurdert prøver i skrivning er gjennomgående mer kritiske til skoleringen (kanskje særlig skrivning på 4. trinn).

Hovedresultater fra påstandssekvensen

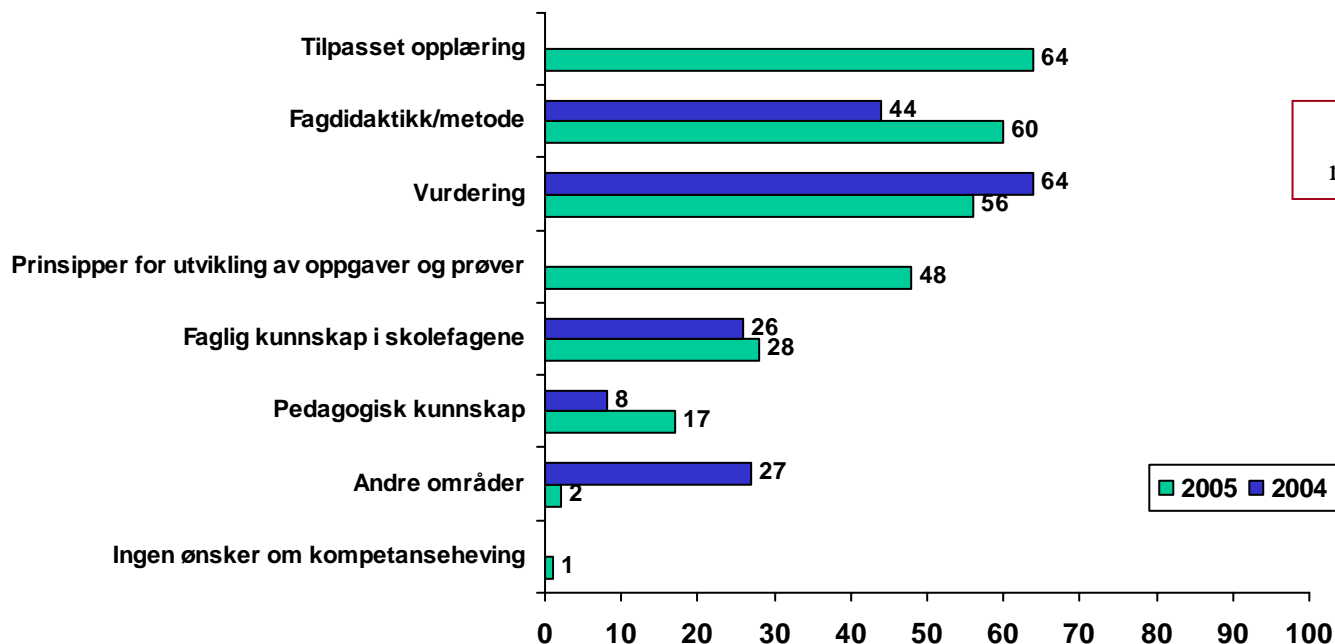
Hvor enig eller uenig er du i følgende utsagn?



Så godt som samtlige ønsker kompetanseheving. Det gjelder særlig tilpasset opplæring, fagdidaktikk/metode, vurdering og prinsipper for utvikling av oppgaver og prøver. I år er det noen flere enn i fjor som ønsker kompetanseheving innen fagdidaktikk/metode.

Behovet for kompetanseheving innen tilpasset opplæring og pedagogisk kunnskap er størst blant de som har vurdert lesing. Ønsket om mer kompetanse innen vurdering er størst blant de som har vurdert skrivning, mens behovet for fagdidaktikk/metode er størst blant de som har vurdert matematikkprøver.

Områder der eksternvurdererne ønsker kompetanseheving



Vi fant tilsvarende resultater blant lærerne.

Oversikt over hvor mange elevbesvarelser eksternevurdererne har vurdert for ulike fag og årstrinn.

	Gjennomsnitt antall prøver
LESING TOTALT	43
Lesing 4. trinn	45
Lesing 7. trinn	43
Lesing 10. trinn	44
Lesing 11. trinn	40
SKRIVING TOTALT	44
Skrijving 4. trinn	47
Skrijving 7. trinn	44
Skrijving 10. trinn	49
Skrijving 11. trinn	31
ENGELSK TOTALT	46
Engelsk skrijving 7. trinn	47
Engelsk skrijving 10. trinn	46
Engelsk skrijving 11. trinn	43
MATEMATIKK TOTALT	44
Matematikk 4. trinn	54
Matematikk 7. trinn	46
Matematikk 10. trinn	46
Matematikk 11. trinn	39
ALLE	43 (32 blant lærere)

•Det er liten variasjon i antall prøver mellom fagene.

•Eksternevurdererne har vurdert flere besvarelser enn lærerne generelt.

Oversikt over hvor lang tid eksternevurdererne mener de i gjennomsnitt brukte på vurderingen av den enkelte elevbesvarelse

	Gjennomsnitt minutter
LESING TOTALT	20
Lesing 4. trinn	15
Lesing 7. trinn	19
Lesing 10. trinn	20
Lesing 11. trinn	23
SKRIVING TOTALT	32
Skrijving 4. trinn	26
Skrijving 7. trinn	40
Skrijving 10. trinn	29
Skrijving 11. trinn	21
ENGELSK TOTALT	24
Engelsk skrijving 7. trinn	25
Engelsk skrijving 10. trinn	24
Engelsk skrijving 11. trinn	21
MATEMATIKK TOTALT	24
Matematikk 4. trinn	23
Matematikk 7. trinn	21
Matematikk 10. trinn	27
Matematikk 11. trinn	26
ALLE	24 (27 blant lærere)

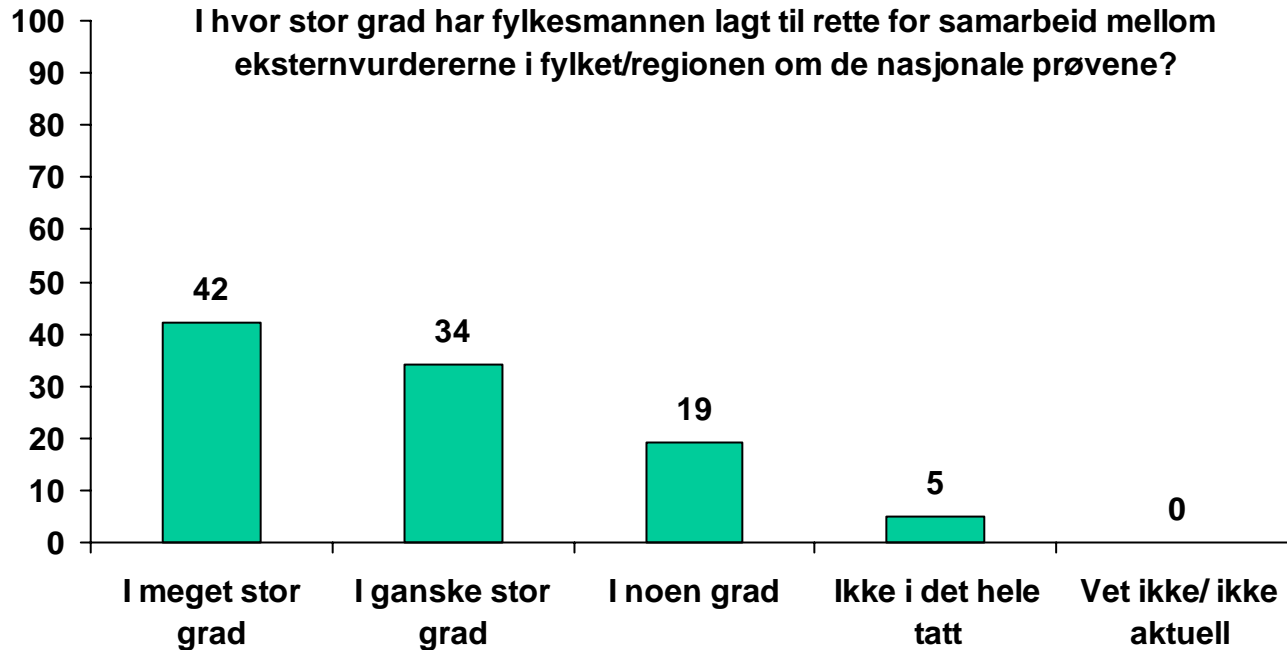
•Skrijving skiller seg ut med lengst tidsbruk.

•Lesing skiller seg ut med kortest tid.

Vi minner igjen om at det er få observasjoner i de undergrupper av spurte som tabellene ovenfor viser til.

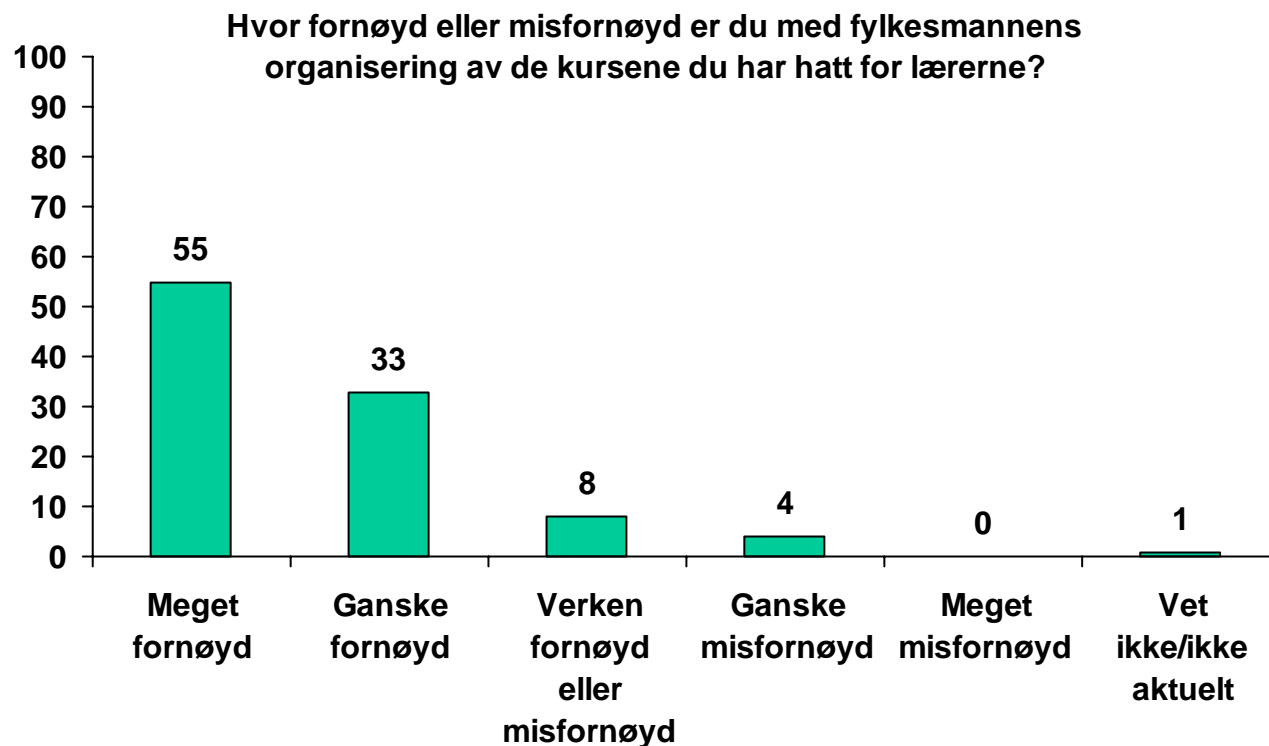
Det er en utbredt oppfatning av at fylkesmannen i stor grad har lagt til rette for samarbeid mellom eksternvurdererne i fylket/regionen. Det samme ble konstatert i fjorårets undersøkelse.

Eksternvurderere i Nord-Norge skiller seg fra gjennomsnittet med en signifikant større andel som mener at fylkesmannen ikke har lagt til rette for samarbeid (20%). Forøvrig er det ingen nevneverdige forskjeller mellom undergrupper av spurte.



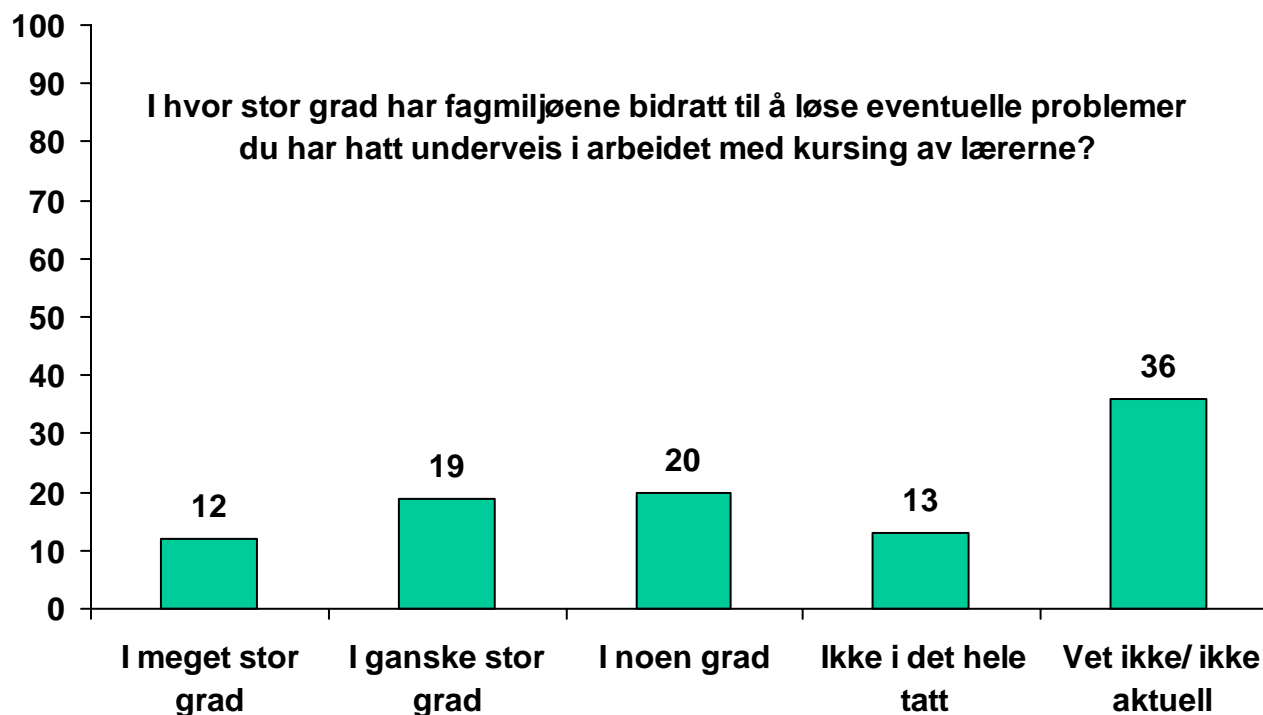
Eksternvurdererne er godt fornøyd med fylkesmannens organisering av kursene de har hatt for lærerne.

Igjen ser vi at eksternvurderere i Nord-Norge er noe mer kritiske. De skiller seg signifikant fra gjennomsnittet med større andeler som uttrykker misnøye (13%), eller forholder seg nøytrale (20% som svarer verken eller). Likevel er flertallet fornøyd også i denne undergruppen.



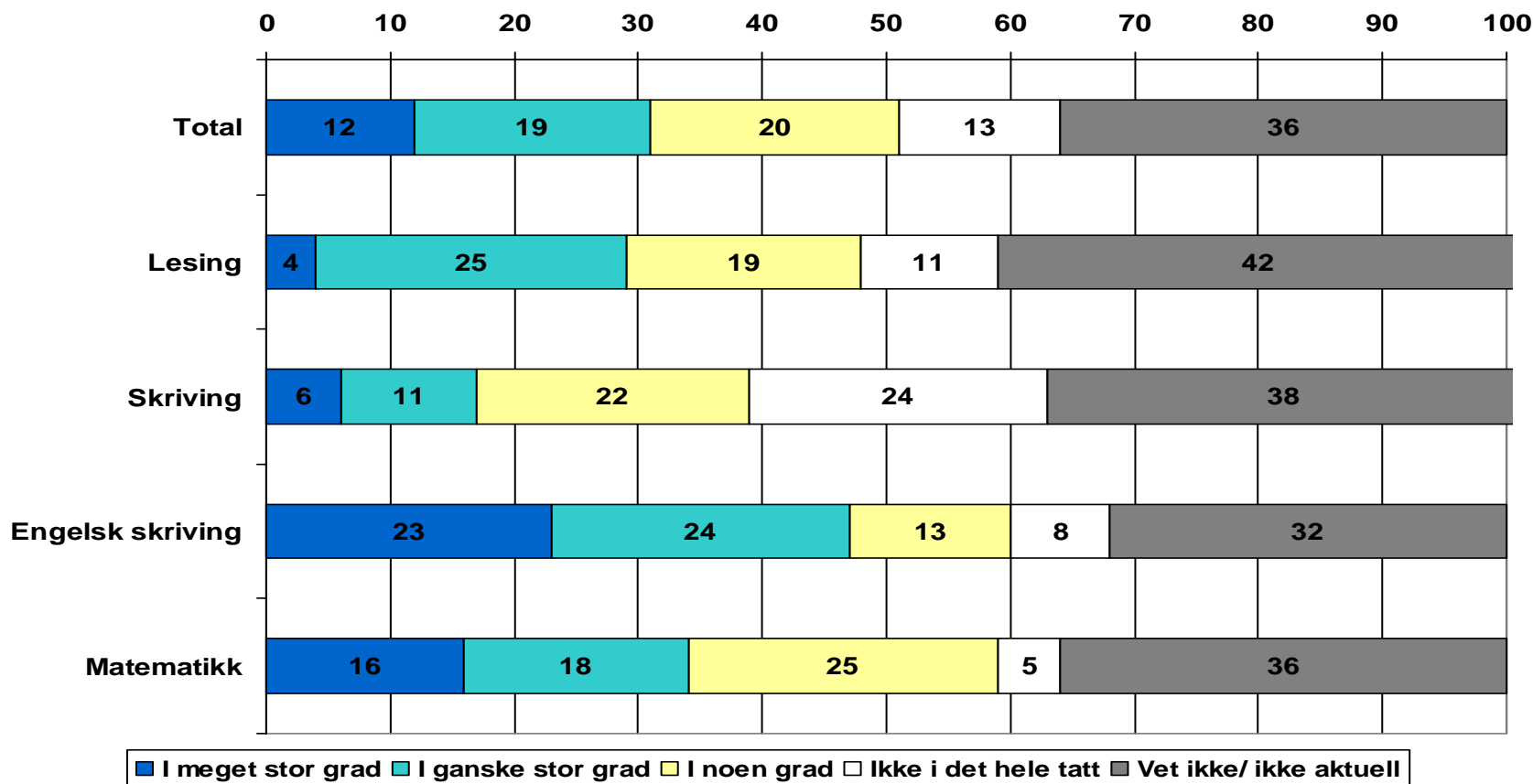
Det er delte meninger om i hvilken grad fagmiljøene har bidratt til å løse problemer i forbindelse med kursing av lærerne. 31% mener fagmiljøene har bidratt i stor grad, mens 20% sier i noen grad. 13% mener de ikke har bidratt i det hele tatt. Nokså mange (36%) synes ikke å ha brukt fagmiljøene ettersom de svarer ”vet ikke/ ikke aktuelt”.

Eksternvurderer 2 rapporterer i større grad (19%) enn eksternvurderer 1 (10%) om at fagmiljøene ikke har bidratt. Den største andelen negative (24%) ser vi blant de som har vurdert prøver i skrijving. Den største andelen som mener at fagmiljøene har bidratt i stor grad, finner vi blant dem som har vurdert prøver i engelsk skrijving, særlig 10. og 11. trinn (se grafisk oversikt på neste side). Det er også mange som opplever at fagmiljøene har bidratt blant de som har vurdert matematikkprøver på 7. trinn.



Det er særlig de som har vurdert engelsk skriving som opplever at fagmiljøene i stor grad har vært behjelpelige i arbeidet med kursing av lærerne

I hvor stor grad har fagmiljøene bidratt til å løse eventuelle problemer du har hatt underveis i arbeidet med kursing av lærerne?



Bruk av resultater i pedagogisk arbeid

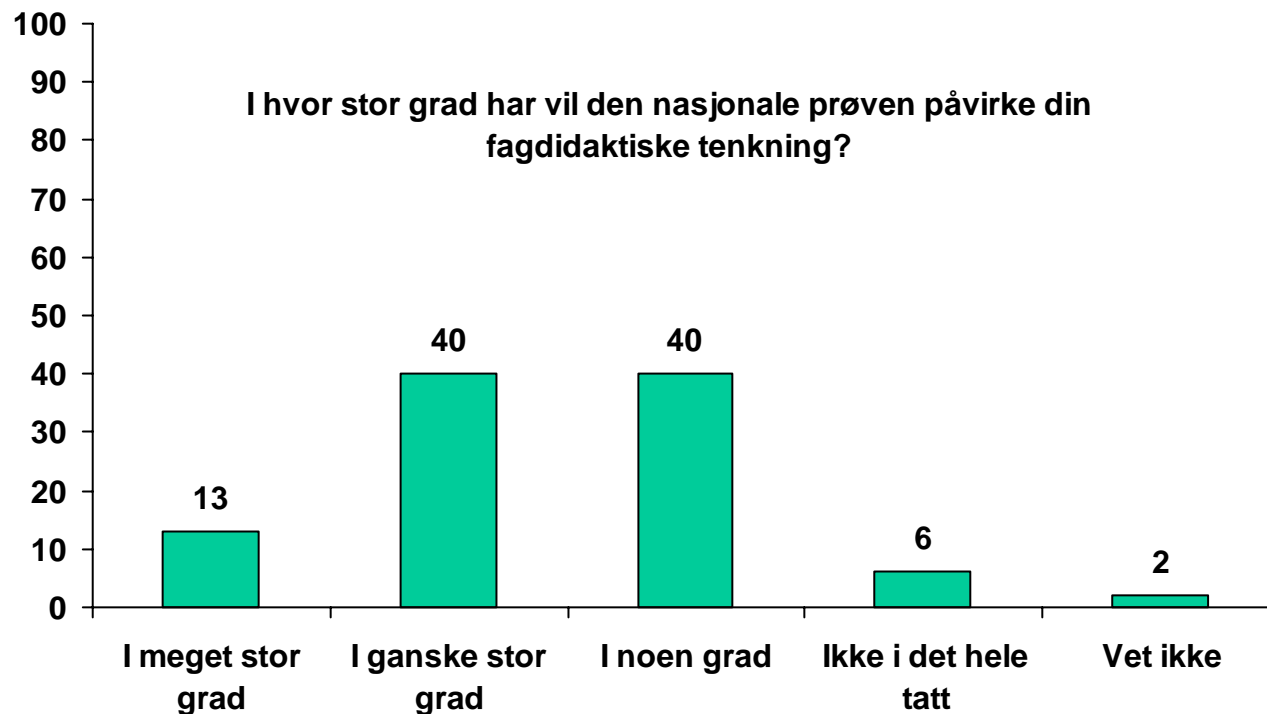
Spørsmål 21



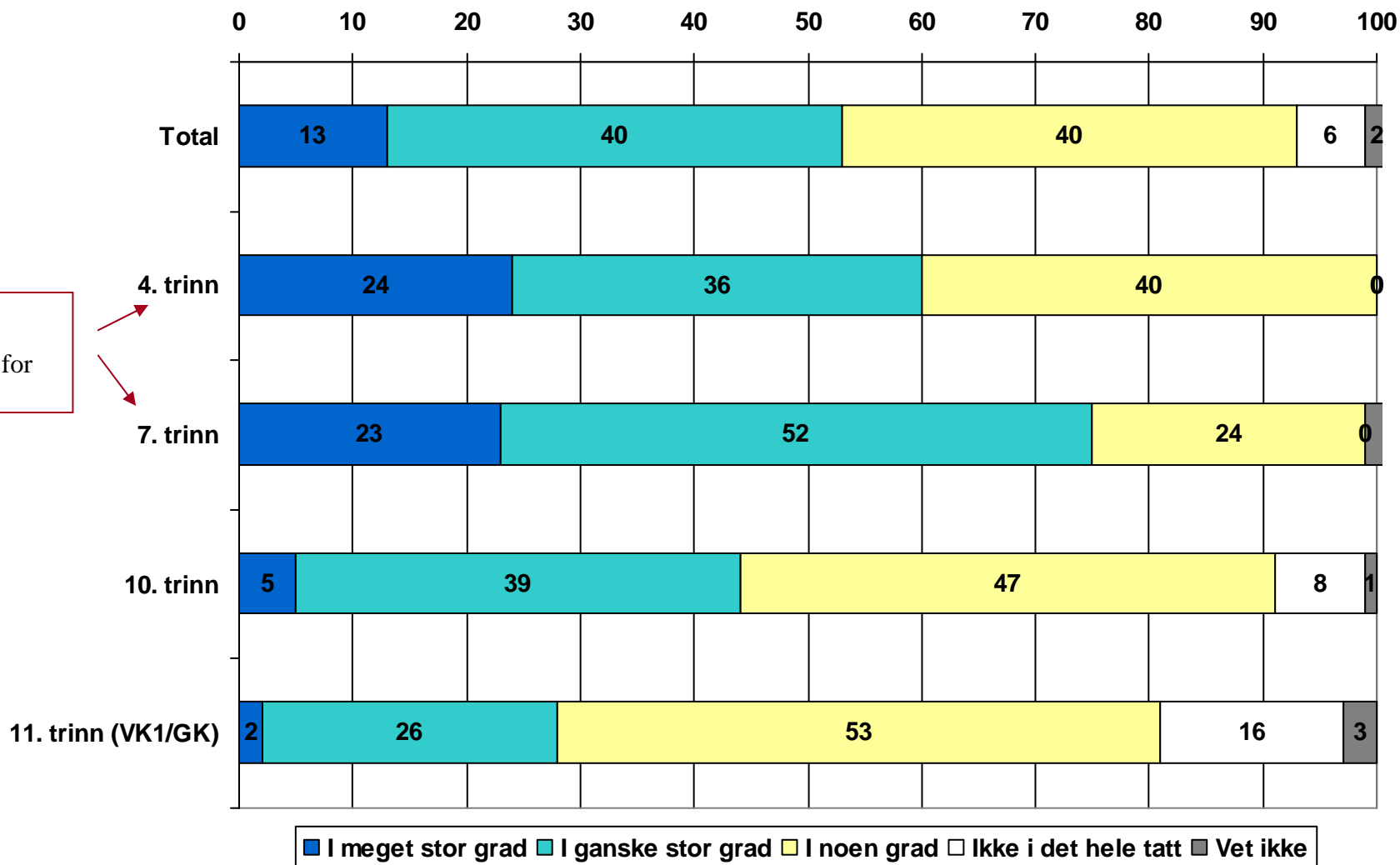
De fleste mener at den nasjonale prøven vil påvirke deres fagdidaktiske tenkning. Drøyt halvparten (53%) mener det vil påvirke tenkningen i stor grad.

Den største andelen som mener deres tenkning vil påvirkes i stor grad, finner vi blant de som har vurdert prøver på 7. trinn. Den laveste andelen (28%) finner vi på 11. trinn (primært skriving). Her finner vi også den største andelen som mener at de ikke vil bli påvirket i det hele tatt (16%). Dette er illustrert grafisk på neste side.

Det er klart flere eksternevurderere enn lærere som mener at deres fagdidaktiske tenkning vil påvirkes.



I hvor stor grad vil den nasjonale prøven påvirke din fagdidaktiske tenkning?



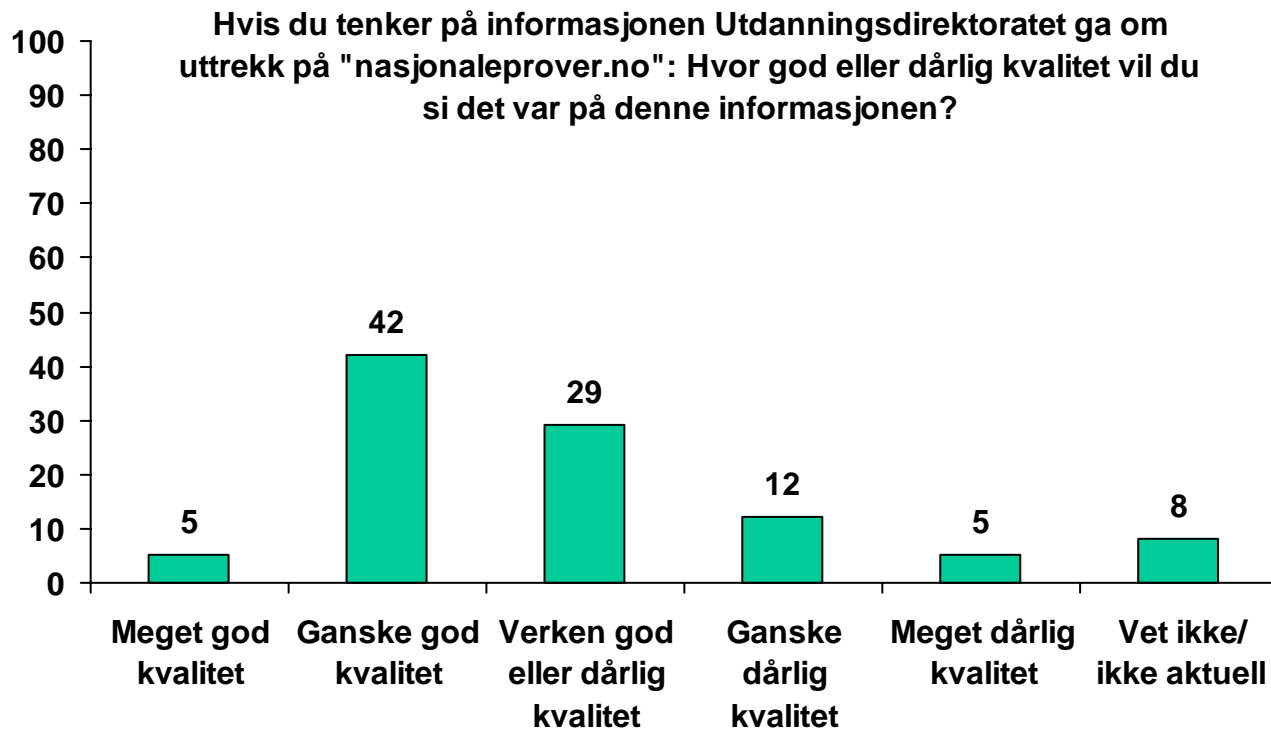
Gjennomføring av ekstern vurdering

Spørsmål 22-26

Denne delen av undersøkelsen er kun besvart av eksternvurderer 1

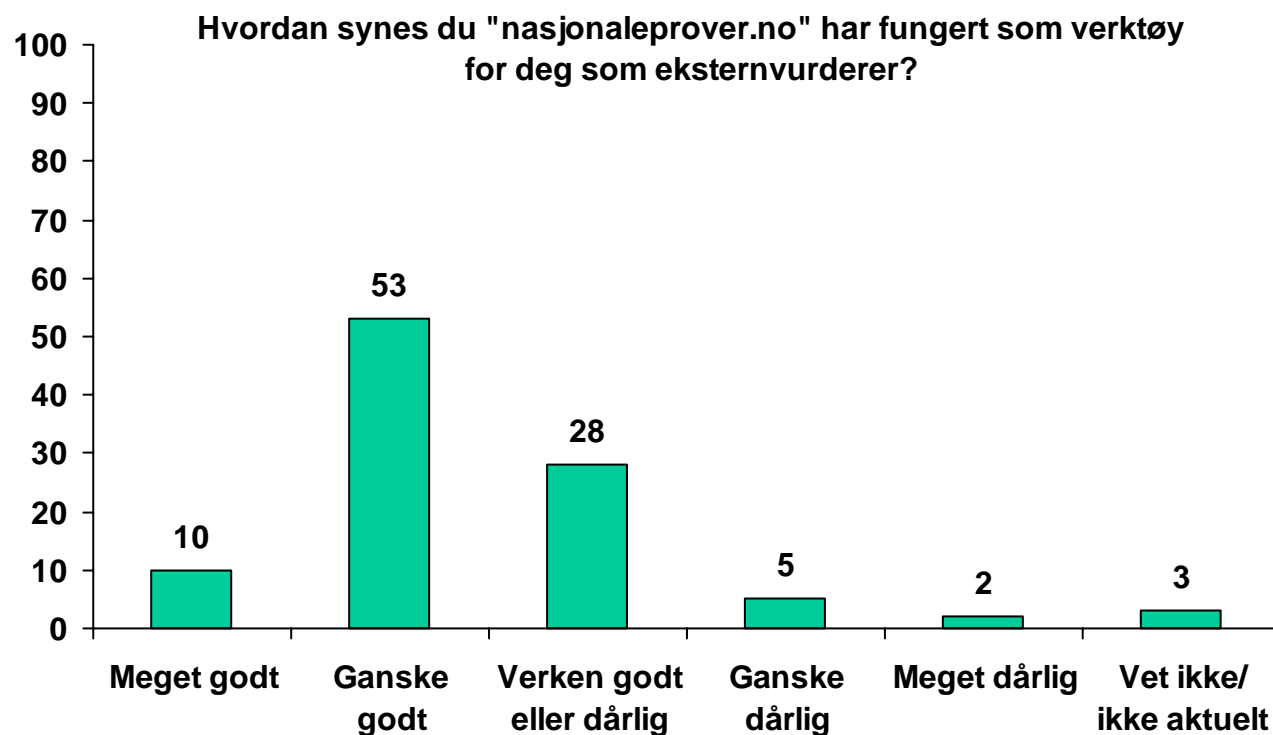
47% mener det var god kvalitet på informasjonen om uttrekk på nettstedet "nasjonaleprover.no", 29% vurderer informasjonen nøytralt, mens 17% mener kvaliteten var dårlig.

De som har vurdert prøver i lesing synes å være mer tilfreds enn de som har vurdert øvrige prøver.



3 av 5 synes "nasjonaleprover.no" har fungert som et godt verktøy.

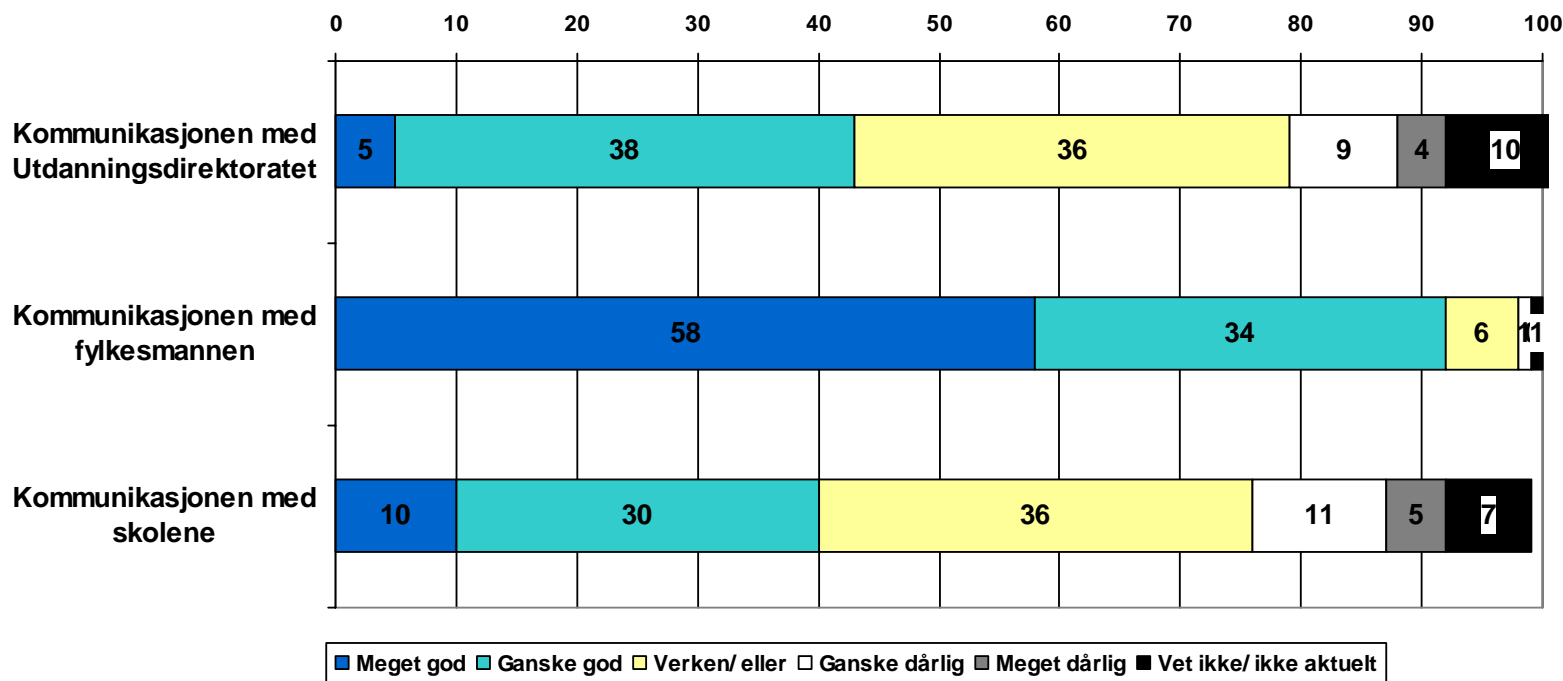
De som synes det har fungert best er de som har vurdert prøver på 7. trinn (lesing). Det er ingen undergrupper som skiller seg fra gjennomsnittet med spesielt store andeler som mener det har fungert dårlig.



Kommunikasjonen med fylkesmannen oppleves som klart bedre enn kommunikasjonen med skolene og Utdanningsdirektoratet. Det er mest misnøye med kommunikasjonen med skolene (16%)

Vi ser en tendens til at de som har vurdert lesing er mer tilfreds med kommunikasjonen med Utdanningsdirektoratet enn øvrige eksternvurderere. Når det gjelder kommunikasjonen med skolene, ser vi en tendens til størst tilfredshet blant de som har vurdert prøver i engelsk skrivning.

Hvor god eller dårlig vil du si at kommunikasjonen mellom de ulike instansene og deg som eksternvurderer har vært i forbindelse med de nasjonale prøvene?



De aller fleste fikk prøvene i posten som avtalt. 2 av 5 måtte følge opp skolen direkte for å få tilsendt prøvene.

De som vurderte prøver i lesing måtte i større grad enn andre følge opp skolene direkte (59%).

Hvordan foregikk forsendelsen fra skolene?

