



Matematikksenteret

Nasjonalt senter for matematikk i opplæringen

RAPPORT

NASJONAL PRØVE I REGNING 8. og 9. TRINN 2012

**Grethe Ravlo
Ole Harald Johansen**

NSMO/NTNU mars 2013

Nasjonal prøve i regning 8. og 9. trinn 2012

Rapporten er basert på resultatene fra versjon 1 av prøven

Grethe Ravlo og Ole Harald Johansen
Nasjonalt senter for matematikk i opplæringen (NSMO)
NTNU mars 2013

Innhold

Innledning.....	3
Validitet.....	3
Resultater innen områdene tall, måling og statistikk	7
Resultater på enkeltoppgaver	7
Oppgaver med størst forskjell i p-verdi for 8. og 9. trinn	8
Oppgaver med lav p-verdi	10
Oppgaver med høy p-verdi.....	11
Mestringsnivå 1 og p-verdi.....	12
Oppgaveformat og p-verdi	13
Kjønnsforskjeller	15
Kjønnsforskjeller innen områdene tall, måling og statistikk.....	18
Kjønnsforskjeller i p-verdier på enkeltoppgaver.....	18
Kjønnsforskjeller og oppgaveformat.....	20
Ubesvarte oppgaver.....	22
Teknisk rapport.....	25
Item-analyse av alle oppgavene	25
<i>Kilder:</i>	28

Innledning

Denne rapporten er basert på resultatene fra alle elevene som gjennomførte versjon 1 av nasjonal prøve i regning for 8. og 9. trinn høsten 2012. Dette gjelder 19 609 elever på 8. trinn og 19 370 elever på 9. trinn. Rapporten handler i utgangspunktet om resultatene for 8. trinn, men i tillegg sammenlignes resultatene for 8. og 9. trinn.

Den nasjonale prøven for 2012 foreligger i tre versjoner, versjon 1, 2 og 3. Oppgavene er de samme, men rekkefølgen til de ti første og de ti siste oppgavene er forskjellig i de tre versjonene. Resultatene viser at de tre versjonene av prøvene fungerte tilnærmet likt. Dette gjelder både for 8. trinn og 9. trinn. Veiledningsmateriellet som ligger på www.udir.no, og analyse materialet som ligger i PAS, baserer seg på rekkefølgen oppgavene har i versjon 1 av prøven. Det er årsaken til at denne rapporten bare inneholder resultatene for versjon 1.

Det var til sammen 59 327 elever på 8. trinn og 58 056 elever på 9. trinn som gjennomførte nasjonal prøve i regning for 2012. Prøven tester målene etter 7. trinn i Kunnskapsløftet (LK06) for grunnleggende ferdighet i å kunne regne i alle fag.

Prøven for 2012 er en elektronisk prøve med 58 oppgaver og prøvetid 90 minutter. Elevene hadde ark til å klaffe på, men svarte elektronisk. Riktig svar ble honorert med 1 poeng per oppgave, og det ble ikke gitt delpoeng. Resultatene i denne rapporten blir presentert i form av p-verdier som tilsvarer den prosentandelen elever som fikk poeng på hver oppgave. Rapporten inneholder gjennomsnittlige p-verdier for alle elevene og for jenter og gutter, både samlet for hele prøven og for hver enkelt oppgave.

Oppgavene er kategorisert etter spesielle kriterier. Vi har sammenlignet løsningsprosenter (p-verdier) i flervalgsoppgaver og åpne oppgaver, løsningsprosenter innenfor områdene tall, måling og statistikk, og undersøkt hvor stor andel av elevene som ikke har svart på oppgavene. Videre har vi sett på hvilke oppgaver som er riktig besvart av mange elever (høy p-verdi), hvilke oppgaver som elevene i stor grad har hatt problemer med å løse riktig (lav p-verdi), og hvilke oppgaver elever med lav gjennomsnittlig poengsum har løst. Oppgavene innen disse kategoriene er analysert for alle elevene og for jentene og guttene hver for seg for begge klassetrinn.

Validitetⁱ

I Kunnskapsløftet (LK06) presiseres det at regning er en grunnleggende ferdighet i alle fag. Dette skal i så stor grad som mulig, gjenspeiles i den nasjonale prøven i regning. Prøven skal kartlegge i hvilken grad elevenes regneferdigheter er i samsvar med kompetansemål der regneferdigheter er integrert (Jfr. Rammeverk for nasjonale prøver, Utdanningsdirektoratet 2010). Siden prøven tester målene for de ulike fagene etter 7. trinn i LK06, blir den gjennomført tidlig på 8. trinn. Gjennomføringen på 9. trinn er for å se effekten av et års ekstra skolegang.

Innholdsmessig er den nasjonale prøven i regning knyttet til områdene tall, måling og statistikk. Rammeverket definerer hva som inngår i de ulike områdene (Utdanningsdirektoratet 2010). Det å kunne bruke regning i ulike sammenhenger skal vektlegges, og prøven skal inneholde både åpne oppgaver og flervalgsoppgaver. Noen av oppgavene er interaktive. I åpne oppgaver skal elevene skrive et svar. I flervalgsoppgaver skal elevene enten velge et av flere svaralternativer, sortere tall til riktig rekkefølge eller koble ulike representasjoner av samme tallverdi. I interaktive oppgaver skal elevene flytte på ett eller flere objekter.

ⁱ Validitet betyr at prøven måler det den er ment å skulle måle, -grunnleggende ferdigheter i regning innenfor områdene tall, måling og statistikk i alle fag, relatert til LK06.

Prøven for 8. og 9. trinn 2012 består av 34 flervalgsoppgaver og 24 åpne oppgaver.

Mange av oppgavene i prøven er relevante for mer enn ett fag. Oppgavene i prøven for 2012 knyttes til følgende fag:

1. Oppgaver innenfor området tall: Norsk, matematikk, naturfag, samfunnsfag, religion/livssyn/etikk, mat & helse, musikk og kroppsøving
2. Oppgaver innenfor området måling: Norsk, matematikk, naturfag, samfunnsfag, religion/livssyn/etikk, kunst & håndverk, mat & helse, kroppsøving og engelsk
3. Oppgaver innenfor området statistikk: Norsk, matematikk, naturfag, samfunnsfag, religion/livssyn/etikk, mat & helse, musikk, kroppsøving og engelsk

Prøven for 8. og 9. trinn består av 23 oppgaver i området tall, 21 innenfor måling og 14 statistikkoppgaver. For at alle elever skal få vist sin kompetanse, uavhengig av faglig nivå, inneholder prøven oppgaver av ulik vanskelighetsgrad. Det er ferdigoppstilte regneoppgaver, oppgaver som tester evne til tolkning og til å løse problemer i kontekst, og evne til resonnering, analyse og vurdering.

Tabell 1 Oppgaverekkefølge i versjon 1 av nasjonal prøve i regning 2012. Tallverdiene er resultater for 8. trinn, $N(8)=19\ 609$ og 9. trinn, $N(9)=19\ 370$

Oppgaver	Innhold	Område	Format	Relevans til fag	P-verdi 8.trinn %	Diff. j-g 8. trinn prosent-poeng	P-verdi 9. trinn %	Diff. j-g 9. trinn prosent-poeng	Diff. (9.-8.) prosent-poeng	Forventet vanskegrad
1	Regneoperasjon med hele tall	Tall	Åpen	Mat, Saf, M&H	67,0	5,5	70,2	4,2	3,3	1
2	Tabell/diagram	Statistikk	Åpen	Mat, No, Saf, Eng, Nat, Rle	92,4	2,5	94,5	2,2	2,0	1
3	Måleenheter	Måling	Flervalg	Mat, No, M&H	43,4	2,7	49,0	-0,3	5,6	4
4	Regning m/desimaltall	Tall	Flervalg	Mat	48,6	-0,3	57,4	-4,0	8,9	3
5	Brøk/prosent	Tall	Flervalg	Mat, Saf	71,2	-8,9	76,7	-8,4	5,5	3
6	Tabell/diagram	Statistikk	Åpen	Mat, No, Saf, Eng, Nat, Rle	86,4	1,4	89,7	1,9	3,3	2
7	Klokke/tid	Måling	Flervalg	Mat, Krø, Nat, Saf	54,8	-19,0	63,6	-17,8	8,8	2
8	Regneoperasjon med prosent	Tall	Flervalg	Mat, Saf, Nat	57,5	-4,7	66,3	-6,7	8,8	3
9	Tabell/diagram	Statistikk	Flervalg	Mat, No, Eng, Nat, Saf, Rle	56,9	-3,9	62,9	-5,7	6,0	3
10	Sammensatt problem - gjennomsnitt og tabell/diagram	Statistikk	Åpen	Mat, No, Eng, Nat, Saf, Rle	19,9	1,3	26,7	2,3	6,8	5
11	Tabell/diagram	Statistikk	Åpen	Mat, No, Eng, Nat, Saf, Rle	56,7	7,1	66,7	7,1	10,0	3
12	Posisjonssystemet	Tall	Flervalg	Mat	72,0	-10,3	77,5	-10,1	5,5	2
13	Brøk/prosent	Tall	Flervalg	Mat, Mu, M&H	24,3	-4,0	30,7	-4,2	6,5	2

Oppgaver	Innhold	Område	Format	Relevans til fag	P-verdi 8. trinn %	Diff. j-g 8. trinn prosentpoeng	P-verdi 9. trinn %	Diff. j-g 9. trinn prosentpoeng	Diff. (9.-8.) prosentpoeng	Forventet vanskegrad
14	Areal	Måling	Åpen	Mat, K&H	39,5	2,8	49,7	3,8	10,2	2
15	Brøk/prosent	Tall	Flervalg	Mat, Nat, Rle, M&H, Saf	50,8	-9,7	62,2	-8,8	11,4	3
16	Regneoperasjon med hele tall	Tall	Flervalg	Mat	74,8	4,4	77,1	2,3	2,3	1
17	Enkelt sammensatt problem - brøk og tabell/diagram	Statistikk	Åpen	Mat, No, Nat, Rle, Saf, Mu, Eng	62,6	-1,1	69,2	-0,7	6,7	3
18	Enkelt sammensatt problem - hele tall	Tall	Flervalg	Mat, M&H	58,1	-8,8	68,6	-9,3	10,5	2
19	Enkelt sammensatt problem - kjøp og salg	Statistikk	Flervalg	Mat, No, Nat, Rle, Saf, Mu, Eng, M&H	63,9	2,8	71,2	2,0	7,3	3
20	Brøk	Tall	Flervalg	Mat, Mu	90,6	-1,5	93,1	-0,6	2,5	1
21	Tabell/diagram	Statistikk	Flervalg	M&H, Mat, No, Saf, Nat	76,6	-0,4	80,1	-2,8	3,5	3
22	Enkelt sammensatt problem - klokke/tid	Måling	Åpen	Mat, No, Saf, Nat	23,0	0,9	30,5	0,3	7,5	3
23	Tabell/diagram	Statistikk	Flervalg	No, Mat, Saf, Nat, Rle, Eng	75,9	-10,5	81,8	-8,2	6,0	3
24	Regneoperasjon med hele tall	Tall	Åpen	Mat, M&H	75,7	8,9	79,6	5,3	3,9	1
25	Tabell/diagram	Statistikk	Flervalg	No, Mat, Saf, Nat, Rle, Eng	76,4	-0,3	82,7	-1,1	6,3	3
26	Måleenheter	Måling	Åpen	Mat, Nat, K&H	24,1	-4,1	31,0	-7,1	6,9	4
27	Brøk/prosent	Tall	Flervalg	Mat, M&H, Mu	43,0	-3,8	50,9	-5,0	7,9	4
28	Sammensatt problem - forhold	Måling	Åpen	No, Mat, Saf, K&H	10,9	-4,8	15,8	-7,3	4,9	5
29	Sammensatt problem - sammenligne data	Tall	Flervalg	Mat, No, Saf	26,3	-10,9	33,8	-14,1	7,6	5
30	Temperatur	Måling	Flervalg	Mat, Nat, Saf	47,0	-0,7	56,7	-1,5	9,7	3
31	Klokke/tid	Måling	Åpen	Mat, Nat, Saf, M&H, Krø	48,9	-4,4	56,8	-3,7	8,0	2
32	Sammensatt problem - måleenheter og problemløsning	Måling	Åpen	Mat, K&H	28,0	-4,5	38,9	-5,1	11,0	4
33	Enkelt sammensatt problem - hele tall	Tall	Flervalg	Mat, Saf, M&H	66,4	5,7	72,9	4,2	6,5	2
34	Sammensatt problem - hele tall og gjennomsnitt	Tall	Åpen	Mat, Krø, M&H	17,6	1,0	26,5	2,1	8,9	4
35	Tabell/diagram	Statistikk	Flervalg	Mat, Nat, Saf, Rle, Eng	72,7	0,8	79,7	0,1	7,0	3
36	Sammensatt problem - målestokk	Måling	Åpen	Mat, Saf, Krø, K&H	8,2	-5,3	14,2	-7,8	6,0	5
37	Brøk/prosent	Tall	Flervalg	Mat, Saf, Nat	38,9	0,3	43,8	-1,4	4,9	4
38	Tabell/diagram	Statistikk	Åpen	Mat, nat, Saf, Eng, Rle	24,3	3,9	33,3	2,3	9,0	4
39	Posisjonssystemet	Tall	Flervalg	Mat	70,4	-0,1	72,8	0,1	2,3	2
40	Klokke/tid	Måling	Åpen	Mat, Nat, Saf, Krø	28,4	2,9	39,5	3,6	11,1	4
41	Enkelt sammensatt problem - vei/fart/tid	Måling	Flervalg	Mat, Krø, Nat	56,9	-3,8	64,2	-4,8	7,3	3
42	Enkelt sammensatt problem - forhold	Måling	Flervalg	Mat, Nat, M&H	50,1	-3,2	58,4	-2,2	8,3	3
43	Gjennomsnitt	Statistikk	Åpen	Mat, No, Krø	11,3	-6,2	16,0	-8,9	4,8	5
44	Regneoperasjon med desimaltall - valuta	Måling	Åpen	Mat, Saf, Eng, M&H	39,0	2,4	46,3	-0,9	7,3	3
45	Brøk/prosent	Tall	Åpen	Mat	60,8	-7,7	68,2	-5,6	7,3	3
46	Måleenheter	Måling	Flervalg	Mat, Nat, M&H	70,5	-7,9	77,1	-7,5	6,6	2
47	Måleenheter	Måling	Flervalg	Mat, Nat, M&H	64,8	-4,2	71,4	-5,6	6,6	4

Oppgaver	Innhold	Område	Format	Relevans til fag	P-verdi 8. trinn %	Diff. j-g 8. trinn prosentpoeng	P-verdi 9. trinn %	Diff. j-g 9. trinn prosentpoeng	Diff. (9.-8.) prosentpoeng	Forventet vanskegrad
48	Måleenheter	Måling	Flervalg	Mat, Nat	35,7	-7,7	43,9	-9,5	8,2	4
49	Regneoperasjon med hele tall	Tall	Flervalg	Mat	41,3	0,4	47,1	-1,8	5,8	3
50	Måleenheter	Måling	Flervalg	Mat, M&H, Nat	35,9	-15,8	41,6	-17,1	5,7	4
51	Areal	Måling	Åpen	Mat, K&H	51,9	3,6	56,4	1,5	4,5	1
52	Enkelt sammensatt problem, med hele tall - kjøp og salg	Tall	Flervalg	Mat, Saf, M&H	53,5	7,2	61,0	6,3	7,5	2
53	Klokke/tid	Måling	Flervalg	Mat, Saf, Eng, Nat	26,7	-5,3	35,7	-7,3	8,9	4
54	Regneoperasjon med desimaltall	Tall	Flervalg	Mat, Saf, M&H	37,4	-7,5	46,0	-9,0	8,6	2
55	Gjennomsnitt	Statistikk	Flervalg	Mat, Saf, Nat	63,4	-2,0	71,1	-1,9	7,7	3
56	Klokke/tid	Måling	Åpen	Mat, Saf	61,6	0,7	68,0	2,7	6,4	1
57	Regneoperasjon med desimaltall	Tall	Åpen	Mat	12,9	-2,8	20,3	-6,1	7,5	2
58	Sammensatt problem	Tall	Åpen	Nat, Krø	2,4	-1,1	5,1	-2,6	2,7	5
Gjennomsnitt					49,1	-2,2	55,9	-3,1	6,8	

Tabell 1 viser rekkefølgen til oppgavene i versjon 1 av nasjonal prøve i regning for 8. og 9. trinn 2012. Tabellen viser hva oppgavene handler om, hvilket område hver oppgave hører inn under, om oppgaven er åpen eller om det er en flervalgsoppgave, og hvilke fag hver oppgave har relevans til. Siden nasjonal prøve for 8. og 9. trinn tester målene etter 7. trinn for grunnleggende ferdigheter i å kunne regne, må valideringen hovedsakelig skje i forhold til hva som kan forventes av elever som nettopp har begynt på 8. trinn. Verdiene for 9. trinn er tatt med til sammenligning.

P-verdiene angir hvor mange prosent av elevene som fikk rett svar på oppgavene. Det er p-verdier for alle de 19 609 elevene på 8. trinn og 19 370 elevene på 9. trinn som gjennomførte versjon 1 av prøven. To kolonner i tabellen viser forskjellen i prosentpoeng mellom jenter og gutter for hvert av trinnene (Diff j-g). Positiv verdi betyr at jentene gjorde det bedre enn guttene på en oppgave, og negativ verdi at guttene gjorde det bedre enn jentene. En kolonne viser forskjellen i prosentpoeng mellom 8. og 9. trinn. Positiv verdi betyr at det er flere elever på 9. trinn enn på 8. trinn som har løst oppgaven.

Oppgavene er plassert på ulike vanskelighetsgrader, relatert til kompetansemålene i LK06. Forventet vanskelighetsgrad 1 betyr at oppgaven bør kunne besvares av de fleste elevene, mens oppgavene på vanskegrad 5 krever normalt både analyse og evne til vurdering på et høyt kognitivt nivå for alderen, og oppgaven forventes løst bare av elever med høy måloppnåelse. Tabell 1 viser at det ikke alltid er samsvar mellom det man kan forvente ut fra LK06 og hva elevene faktisk mestrer (p-verdier).

Poengsummen som en elev oppnår på prøven, relateres til et mestringsnivå. På 8. og 9. trinn er det fem mestringsnivåer, nivå 1 til nivå 5, hvor 5 er det høyeste nivået.

En lærergruppe på fire personer har foretatt ekstern validering av prøvesettet. I tillegg fikk vi tilbakemelding i form av logg fra lærere da vi piloterte oppgavene. Hver oppgave er pilotert på ca. 1 000 elever.

Resultater innen områdene tall, måling og statistikk

Den gjennomsnittlige p-verdien for hele prøven er 49,1 prosent på 8. trinn. Det betyr at elevene i gjennomsnitt løste 49,1 prosent av alle oppgavene riktig. Tilsvarende for 9. trinn er 55,9 prosent (se tabell 1 og 2).

Tabell 2 Gjennomsnittlige p-verdier, forskjeller i prosentpoeng og effektstørrelser for tall, måling og statistikk for 8. og 9. trinn. $N(8) = 19\ 609$ og $N(9) = 19\ 370$

Område	Gjennomsnittlig p-verdi		Diff 9. – 8. trinn	Effektstørrelse ⁱⁱ 8. / 9. trinn
	8. trinn	9. trinn		
Tall (n=23)	50,5	56,9	6,4	0,30
Måling (n=21)	40,4	48,0	7,6	0,32
Statistikk (n=14)	59,9	66,1	6,2	0,31
Hele prøven	49,1	55,9	6,8	0,33

Det er området statistikk som har høyest gjennomsnittlig løsningsprosent på denne prøven, og måling som har lavest. Vi ser at tendensen er den samme for 9. som for 8. trinn, men at elevene på 9. trinn i gjennomsnitt har løst flere oppgaver riktig enn elevene på 8. trinn. Det er også minst forskjell i løsningsprosent innen statistikk når vi sammenligner trinnene, ca. seks prosentpoeng, mens forskjellen for tall og måling er 6,4 og 7,6 prosentpoeng, alle i favør 9. trinn. Betydningen av å være elev på 8. eller 9. trinn i denne prøven ($e = 0,33$) viser middels effekt siden forskjellen i gjennomsnittspoengsum utgjør litt mer enn 30 prosent av standardavviketⁱⁱⁱ (se tabell 2). Av enkeltområdene er det måling som viser størst forskjell i resultatene for trinnene i favør 9. trinn.

En enkeltoppgave kan inneholde elementer fra flere områder, og det kan derfor noen ganger være vanskelig å plassere oppgaver på et bestemt område. I alle oppgavene behandler vi tall i en eller annen sammenheng. Det typiske for måling er imidlertid behandling av enheter for vekt, lengde og volum, beregning av areal og regning med temperatur og tid. Statistikk handler om tabeller og grafer og tolkninger og beregninger i slike sammenhenger. Alle områdene har oppgaver med ulik vanskelighetsgrad.

Det blir ikke riktig å sammenligne resultatene for prøven i 2012 med resultatene fra tidligere prøver. Oppgavene er ikke like, og det er heller ikke like mange oppgaver innenfor hvert av områdene fra år til år. De nasjonale prøvene er ikke laget ut fra de forutsetningene at de skal kunne sammenligne resultater over tid. Det vi kan si noe om, er resultatene på hver enkelt oppgave, og antatte misoppfatninger som resulterer i feil svar på oppgaver. I tillegg har det siden 2010 vært interessant å sammenligne resultatene for 8. og 9. trinn for å se i hvilken grad et års ekstra skolegang har ført til bedre prestasjoner.

Resultater på enkeltoppgaver

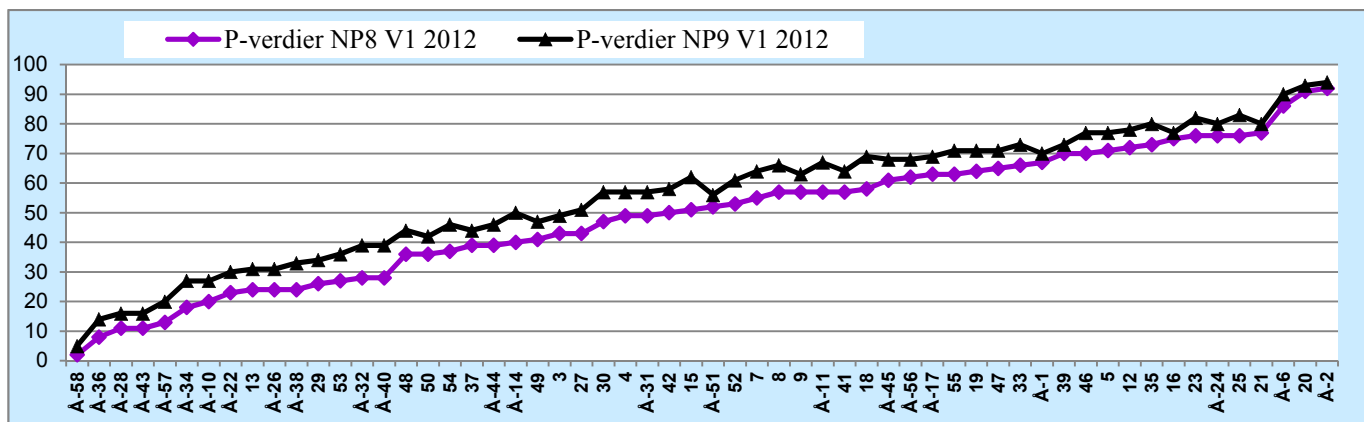
Oppgavenes p-verdier varierer fra 2,4 til 92,4 prosent på 8. trinn og fra 5,1 til 94,5 prosent på 9. trinn. For alle oppgavene gjelder at de elevene som har løst en oppgave riktig, har høyere gjennomsnittlig poengsum på hele prøven enn de som ikke har greid å løse oppgaven (se teknisk rapport tabell 17).

Resultatene for 8. og 9. trinn samsvarer når det gjelder hvilke oppgaver som har falt vanskelig for elevene, og hvilke oppgaver som har høy p-verdi (se figur 1).

ⁱⁱ Effektstørrelse (e) har benevnelse standardavvik, og sier noe om forskjeller i gjennomsnittsverdier til to grupper. For eksempel i hvilken grad antall år i skolen har betydning for et resultat. For $e \leq 0,3$ er effekten liten. Hvis $0,3 < e < 0,8$ er effekten middels, og hvis $e \geq 0,8$ er effekten av hvilken gruppe man tilhører stor.

ⁱⁱⁱ Standardavviket (sd) forteller om spredningen i et datamateriale. Jo større standardavvik, jo mer avviker enkeltobservasjoner i et datamateriale fra gjennomsnittsverdien. 68 % av observasjonene finnes innenfor $\pm 1sd$.

Figur 1 Oppgavene i nasjonal prøve i regning for 8. og 9. trinn 2012, sortert fra laveste til høyeste p-verdi ut fra resultatene til elevene på 8. trinn. Fiolett kurve viser p-verdier for 8. trinn og svart kurve for 9. trinn. Å betyr åpne oppgaver. De andre er flervalgsoppgaver. $N(8) = 19\ 609$ og $N(9) = 19\ 370$. 58 oppgaver

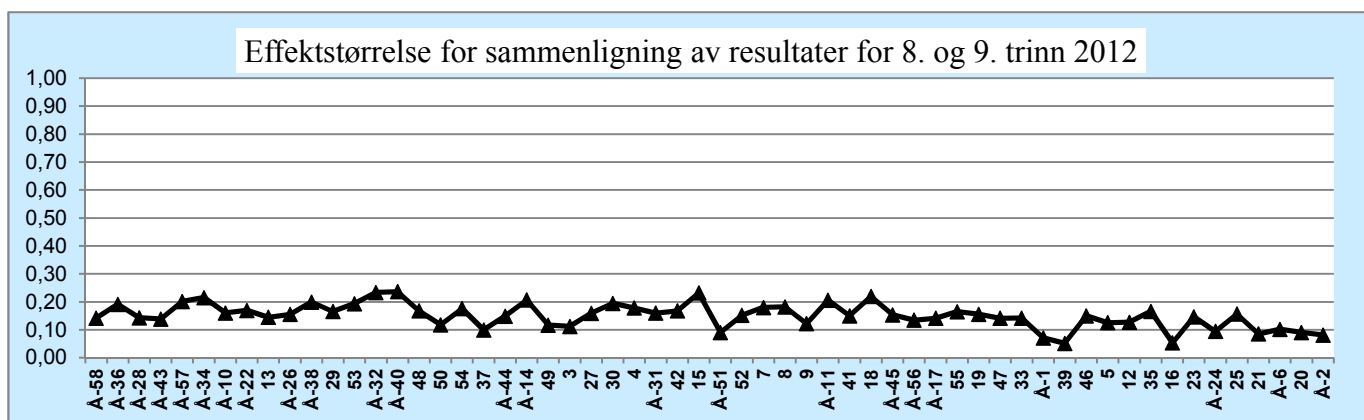


Det er små variasjoner når det gjelder plassering etter p-verdi for resultatene på 8. og 9. trinn. For alle oppgavene gjelder at løsningsprosenten på 9. trinn var like høy eller høyere enn løsningsprosenten på 8. trinn (se figur 1).

Oppgaver med størst forskjell i p-verdi for 8. og 9. trinn

Effektstørrelsen regnes i prosent av standardavviket. Størst effektstørrelse er for oppgave 40 hvor verdien er 24 prosent av standardavviket ($e = 0,24$). Oppgave 40 er en sammensatt oppgave som handler om planlegging av tid. Det var 28 prosent av elevene på 8. trinn og 40 prosent på 9. trinn som løste oppgaven (se figur 1 og 2).

Figur 2 Effektstørrelsen for hver oppgavene når vi sammenligner resultatene for 8. og 9. trinn. Oppgavene er sortert fra laveste til høyeste p-verdi ut fra resultatene for 8. trinn og har samme rekkefølge som i figur 1. Å betyr åpne oppgaver. De andre er flervalgsoppgaver. $N(8) = 19\ 609$ og $N(9) = 19\ 370$. 58 oppgaver



I sammensatte oppgaver må elevene forholde seg til flere opplysninger og velge riktig regneart og metode for å finne løsningen (se figur 3 og 4).

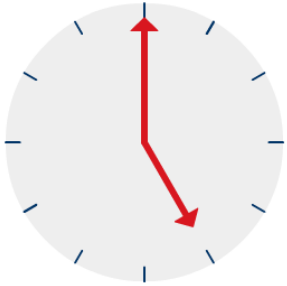
Figur 3 Eksempel på sammensatt oppgave som handler om planlegging av tid. Oppgave 4 i eksempeloppgaver 2, nasjonal prøve i regning for 8. trinn, www.udir.no

Torill skal reise med båt fra Oslo til København. Båten går fra Oslo kl. 17.00, og reisetiden er 16 t og 30 min.

Torill skal vekkes 1 t og 15 min før båten kommer til København.

Når skal Torill vekkes?

Still viserne slik at de viser riktig klokkeslett.



Alle oppgavene i prøven for 8. og 9. trinn 2012 har altså effektstørrelse lavere enn 0,30 og viser at det har liten betydning for resultatet om man er elev på 8. trinn eller 9. trinn. Oppgavene nærmest oppgave 40 i effektstørrelse er 32, 15 og 18 (se figur 2). Oppgave 32 er en sammensatt oppgave med tema måling ($e = 0,23$), mens oppgave 15 handler om ulike representasjoner av tall ($e = 0,23$). Oppgave 18 handler om kjøp og salg ($e = 0,22$).

Det er en tendens til at elever på 9. trinn behersker sammenhengen mellom brøk og prosent, og regning med desimaltall, prosent og tid, bedre enn elevene på 8. trinn. I tillegg løser elevene på 9. trinn sammensatte oppgaver i større grad enn 8. trinns elevene (se tabell 3).

Figur 4 Eksempel på en sammensatt oppgave fra prøven i 2011. Oppgaven stod sist i oppgavesettet og har høy vanskelighetsgrad. P-verdi 10 for 8. trinn og p-verdi 14 for 9. trinn i 2011

En bonde skal sprøyte en åker som er 200 m bred og 400 m lang. Bonden skal bruke 4 liter sprøytemiddel per 1000 m² med åker.

Hvor mange liter sprøytemiddel trenger bonden for å sprøyte åkeren?

Svar: liter



Oppgaver om målestokk faller ofte vanskelig for elevene. Årets oppgave (36) inneholder både regning med målestokk og omgjøring av enheter. Det var nesten dobbelt så mange elever på 9. som på 8. trinn som løste oppgaven, men p-verdiene er svært lave, bare 8 prosent på 8. trinn og 14 prosent på 9. trinn. Effektstørrelsen er 19 prosent av standardavviket.

Tabell 3 viser oppgaver med størst økning i p-verdi og størst effektstørrelse fra 8. til 9. trinn. De fleste av disse oppgavene krever regning i flere trinn og er sammensatte oppgaver. Hvis vi sammenligner forventet vanskelighetsgrad med p-verdien til oppgaven, ser vi at det ikke alltid er samsvar mellom det vi kan forvente ut fra læreplanen i fag, og det elevene greier å løse. Vi ser også at den største forskjellen mellom trinnene ikke er å finne blant de vanskeligste oppgavene.

Tabell 3 Oppgaver med mer enn ni prosentpoeng økning i p-verdi fra 8. til 9. trinn.
 $N(8) = 19\ 609$ og $N(9) = 19\ 370$

Oppgave	Innhold	Område	Format	P-verdi 8.trinn	P-verdi 9.trinn	Diff. p-verdi 9. – 8. trinn	Effektstørrelse	Forventet vanskelighetsgrad
15	Brøk, desimaltall og prosent	Tall	Flervalg	50,8	62,2	11,4	0,23	3
40	Klokke / tid	Måling	Åpen	28,4	39,5	11,1	0,24	4
32	Sammensatt problem, måleenheter	Måling	Åpen	28,0	38,9	11,0	0,23	4
18	Enkelt sammensatt problem, hele tall	Tall	Flervalg	58,1	68,6	10,5	0,22	2
14	Areal	Måling	Åpen	39,5	49,7	10,2	0,21	2
11	Tabell / diagram	Statistikk	Åpen	56,7	66,7	10,0	0,21	3
30	Temperatur	Måling	Flervalg	47,0	56,7	9,7	0,19	3
38	Tabell / diagram	Statistikk	Åpen	24,3	33,3	9,0	0,20	4

Oppgaver med lav p-verdi

De ti oppgavene som elevene på 8. trinn skåret dårligst på, har p-verdi fra 2,4 til 24,3. Alle disse oppgavene finner vi også blant de ti med lavest skår på 9. trinn med p-verdier fra 5,1 til 30,7 (se tabell 4). I sju av disse ti oppgavene gjør guttene på 8. og 9. trinn det bedre enn jentene.

Tabell 4 Prøvens ti oppgaver med lavest p-verdi på 8. trinn, og resultatet for de samme oppgavene på 9. trinn. $N(8) = 19\ 609$ og $N(9) = 19\ 370$

Oppgave	Innhold	Område	Format	8. trinn		9. trinn		Forventet vanskelighetsgrad
				P-verdi	Diff j-g	P-verdi	Diff j-g	
58	Sammensatt problem	Tall	Åpen	2,4	-1,1	5,1	-2,6	5
36	Sammensatt problem, målestokk	Måling	Åpen	8,2	-5,3	14,2	-7,8	5
28	Sammensatt problem, forholdstall	Måling	Åpen	10,9	-4,8	15,8	-7,3	5
43	Gjennomsnitt	Statistikk	Åpen	11,3	-6,2	16,0	-8,9	5
57	Regneoperasjon med desimaltall	Tall	Åpen	12,9	-2,8	20,3	-6,1	2
34	Sammensatt problem med hele tall og gjennomsnitt	Tall	Åpen	17,6	1,0	26,5	2,1	4
10	Sammensatt problem med gjennomsnitt og tabell	Statistikk	Åpen	19,9	1,3	26,7	2,3	5
22	Enkelt sammensatt problem med klokke / tid	Måling	Åpen	23,0	0,9	30,5	0,3	3
26	Måleenheter	Måling	Åpen	24,1	-4,1	31,0	-7,1	4
13	Brøk / prosent	Tall	Flervalg	24,3	-4,0	30,7	-4,2	2

Oppgave 58 har lavest p-verdi (se tabell 4 og figur 1). Oppgaven står sist i settet og har 15,5 prosent ubesvart for 8. trinn og 16,3 prosent ubesvart for 9. trinn. Den er en sammensatt oppgave som krever høy grad av analyse, vurdering og resonnement i tillegg til ferdigheter i å regne med tid. De andre oppgavene handler om å sortere brøker, regne med negative desimaltall, regne ut volum, omgjøring av enheter, beregne klokkeslett og å regne med forholdstall i ulike sammenhenger. Åtte av de ti oppgavene med lavest p-verdi på begge trinn, er innenfor områdene tall og måling. Dette stemmer med tendensen fra tidligere års prøver. Det er én flervalgsoppgave og ni åpne oppgaver, og tre av oppgavene er interaktive.

Når vi ser på de ti oppgavene som færrest elever løser riktig, er det i stor grad samsvar med hva vi forventet (se tabell 4). Sju av de ti oppgavene har forventet vanskelighetsgrad 4 eller 5. I oppgave 22 hadde vi ventet høyere p-verdi. Oppgaven er interaktiv og består i å stille viserne på ei klokke etter planlegging og resonnement sett i sammenheng med tid (se figur 3

som eksempel). Hvert år ser vi at elever har problemer med å regne ut tidsintervall mellom to klokkeslett, eller med å finne nytt klokkeslett noen timer og minutter fram eller tilbake i tid. Sammenhengen mellom enheter og å vurdere størrelser med ulike enheter, er også oppgaver elever har stor utfordring med å løse (se figur 5).

Figur 5 Eksempel på målingsoppgave som tester enheter for volum. Oppgave 44 i prøven for 2011. P-verdi 37 for 8. trinn og 43 for 9. trinn i 2011

Nadia skal reise med fly. Hun får ikke lov til å ta med seg flasker som rommer mer enn 100 ml.

Hvilken av flaskene kan hun ta med på flyet?

en flaske som rommer 0,25 liter

en flaske som rommer 25 cl

en flaske som rommer 15 dl

en flaske som rommer 0,75 dl



For å løse oppgaven i figur 5, kreves formelle kunnskaper i matematikk. Her kan ikke eleven resonnerer seg fram til riktig svar bare ved å bruke sunn fornuft. Hvis man imidlertid har forstått strukturen i systemet med milli, centi, desi, osv. vil det kunne ha stor overføringsverdi også til andre systemer av enheter.

Oppgaver med høy p-verdi

De ti oppgavene med høyest løsningsprosent har p-verdier fra 72,0 til 92,4 på 8. trinn og 77,5 til 94,5 på 9. trinn. Her er det ingen målingsoppgaver, men seks statistikk- og fire tallopgaver. I tillegg er sju av oppgavene flervalgsoppgaver. Alle disse oppgavene ligger i første halvdel av prøvesettet, med unntak av oppgave 35. De ti letteste er de samme oppgavene på 8.- og 9.trinn. Det er godt samsvar mellom hva vi har forventet og elevenes prestasjoner på disse ti oppgavene, men liten forskjell i p-verdi på resultatene for 8. og 9. trinn (se tabell 5).

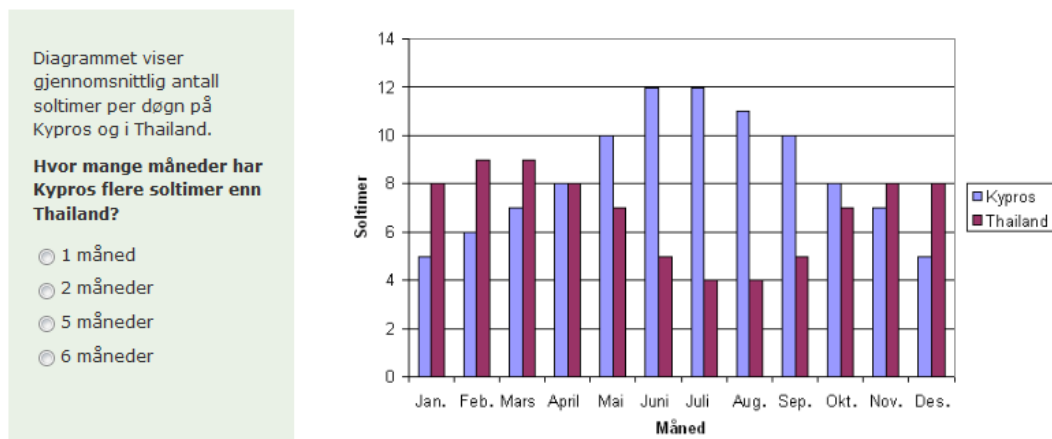
Tabell 5

Prøvens ti oppgaver med høyest p-verdi på 8. trinn, og resultater for de samme oppgavene på 9. trinn. $N(8) = 19\ 609$ og $N(9) = 19\ 370$

Oppgave	Innhold	Område	Format	8. trinn		9. trinn		Forventet vanskelighetsgrad
				P-verdi	Diff j-g	P-verdi	Diff j-g	
2	Tabell/diagram	Statistikk	Åpen	92,4	2,5	94,5	2,2	1
20	Brøk	Tall	Flervalg	90,6	-1,5	93,1	-0,6	1
6	Tabell / diagram	Statistikk	Åpen	86,4	1,4	89,7	1,9	2
21	Tabell/diagram	Statistikk	Flervalg	76,6	-0,4	80,1	-2,8	3
25	Tabell / diagram	Statistikk	Flervalg	76,4	-0,3	82,7	-1,1	3
23	Tabell / diagram	Statistikk	Flervalg	75,9	-10,5	81,8	-8,2	3
24	Regneoperasjon med hele tall	Tall	Åpen	75,7	8,9	79,6	5,3	1
16	Regneoperasjon med hele tall	Tall	Flervalg	74,8	4,4	77,1	2,3	1
35	Tabell / diagram	Statistikk	Flervalg	72,7	0,8	79,7	0,1	3
12	Posisjonssystemet	Tall	Flervalg	72,0	-10,3	77,5	-10,1	2

Oppgave 6 er den eneste interaktive oppgaven blant de ti letteste oppgavene. Elevene skal lese av en tabell og dra punkter til riktig verdi. Ca. 86 prosent (8. trinn) og 90 prosent (9. trinn) av elevene løste oppgaven. Alle de ti oppgavene med høyest p-verdi er oppgaver som løses i ett trinn. Statistikkoppgavene handler om å lese av en tabell eller en graf, eller å tegne et diagram (se figur 6). Talloppgavene handler om representasjoner av tall eller enkle regneoperasjoner.

Figur 6 Eksempel på statistikkoppgave som mange elever løser. Oppgave 22 i prøven for 2011, med p-verdi 69 på 8. trinn og 73 på 9. trinn



Blant de ti oppgavene med høyest løsningsprosent er det like mange oppgaver i jente- som i guttefavør. De største forskjellene i enkeltoppgaver er imidlertid i guttefavør, og i de samme oppgavene for 8. og 9. trinn. Størst forskjell er det i oppgave 23 som handler om avlesning av en litt komplisert tabell med klokkeslett, og oppgave 12 som handler om posisjonssystemet og desimaltall. Jentene gjør det bedre enn guttene i ferdig oppstilte oppgaver, i å lese av gjenkjennelige grafer og tabeller og i enkle praktiske oppgaver hvor de kan bruke kjente algoritmer. Dette samsvarer med resultater fra tidligere års nasjonale prøver i regning.

Mestringsnivå 1 og p-verdi

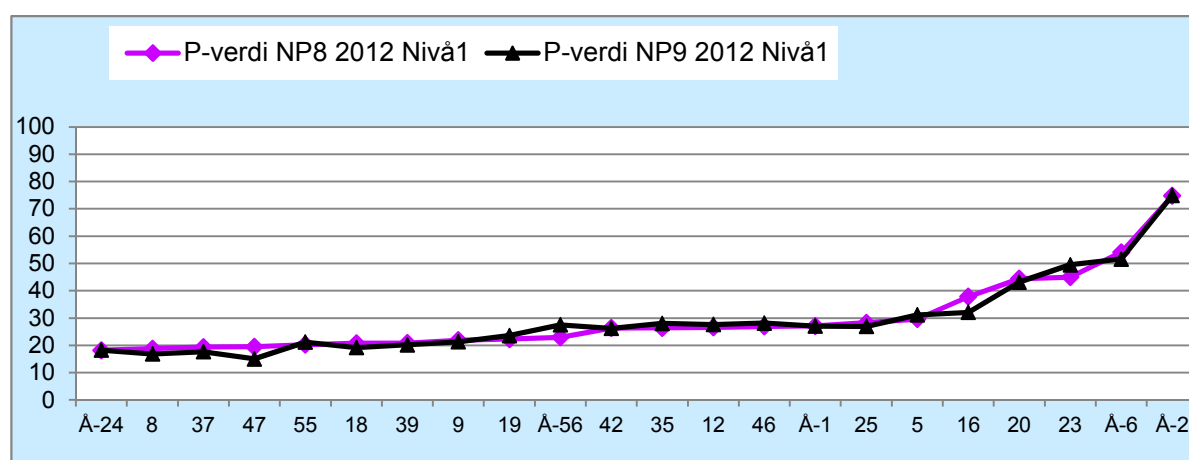
Elevene på mestringsnivå 1 på 8. trinn (1 189 av 19 609 elever) og på 9. trinn (713 av 19 370 elever) har fra 0 til 11 poeng. Analyser viser at disse elevene fikk de fleste av poengene sine på seks av de ti oppgavene med høyest p-verdi for alle elevene (se tabell 5 og 6).

Tabell 6 De ti oppgavene som har høyest p-verdi for elever på **mestringsnivå 1** på 8. trinn, og resultatet for de samme oppgavene på 9. trinn. P-verdi 74,8 betyr at 74,8 prosent av elevene på nivå 1 har løst oppgaven riktig. Forskjellen, j-g, gjelder elever på nivå 1. $N(8) = 1\,189$ og $N(9) = 713$

Oppgave	Innhold	Område	Format	8. trinn		9. trinn		Forventet vanskelighetsgrad
				P-verdi	Diff j-g	P-verdi	Diff j-g	
2	Tabell / diagram	Statistikk	Åpen	74,8	12,9	74,9	10,4	1
6	Tabell / diagram	Statistikk	Åpen	54,1	-3,8	51,6	-2,3	2
23	Tabell / diagram	Statistikk	Flervalg	45,0	-1,2	49,5	5,7	3
20	Brøk	Tall	Flervalg	44,4	-10,2	43,2	-1,0	1
16	Regneoperasjon med hele tall	Tall	Flervalg	37,8	-2,3	32,1	4,4	1
5	Brøk / prosent	Tall	Flervalg	29,6	-5,8	31,1	-8,8	3
25	Tabell / diagram	Statistikk	Flervalg	28,3	-0,5	26,9	-0,6	3
1	Regneoperasjon med hele tall	Tall	Åpen	27,1	4,1	27,1	8,7	1
46	Måleenheter	Måling	Flervalg	26,9	-4,3	28,2	-0,2	2
12	Posisjonssystemet	Tall	Flervalg	26,6	-6,1	27,6	-9,2	2

Av de ti oppgavene som elevene på nivå 1 fikk flest poeng på, er det fire oppgaver uten kontekst; 12, 16, 20 og 46. Det har vist seg tidligere at hvis elevene kjenner seg igjen i en situasjon, er ikke tekstens lengde den største hindringen. Dette kan vi se av resultatene på oppgavene 1, 2, 5, 6 og 23 som er oppgaver i kontekst. Sju av oppgavene er flervalgsoppgaver. I fire av oppgavene er p-verdien for 8. trinn nivå 1 høyere enn for 9. trinn nivå 1. Et eksempel er oppgave 16 der forskjellen mellom p-verdiene for 8. og 9. trinn er 5,7 prosentpoeng i favør 8. trinn. En oppgave i guttefavør på 8. trinn, men jentefavør på 9. trinn (se tabell 6 og figur 7). Oppgaven er ferdig oppstilt og handler om subtraksjon av to hele tall. Dette løser vanligvis jentene bedre enn guttene, men ikke denne gang på 8. trinn nivå 1.

Figur 7 De 22 oppgavene som elevene på **mestringsnivå 1** på 8. trinn har fått flest poeng på, sortert fra laveste til høyeste p-verdi og i sammenligning med resultatene for elever på 9. trinn på samme nivå. Fiolett kurve viser p-verdier for 8. trinn og svart kurve for 9. trinn. $N(8) = 1\,189$ og $N(9) = 713$



De oppgavene elevene som skårer dårligst har fått flest poeng på, handler om å lese av tabell, tolke og sammenligne data og gjøre enkle beregninger med hele tall i praktiske sammenhenger. Jentene gjør det bedre enn guttene i noen oppgaver av regneteknisk art, og når det gjelder å lese av enkle tabeller og grafer. Guttene gjør det imidlertid bedre enn jentene i åtte av ti oppgaver på 8. trinn og i seks av ti på 9. trinn. Guttene gjør det spesielt bedre enn jentene i oppgaver hvor det er omgjøring av enheter, plassere desimaltall på tallinja og forstå prosent og brøk, oppgave 5, 12, 20 og 46 (se tabell 6 og figur 7).

I oppgavene 8, 16, 37 og 47 gjør elevene på 9. trinn nivå 1 det litt dårligere enn elevene på samme nivå på 8. trinn. Oppgave 8 og 37 handler om prosent, 16 er en ferdig oppstilt oppgave med subtraksjon av hele tall og oppgave 47 innebærer både omgjøring fra g til kg og beregninger i den sammenheng.

For de ti oppgavene som har høyest p-verdi på nivå 1, er andel ubesvare av guttene på 8. trinn i gjennomsnitt ca. 5 prosent og for jentene ca. 4 prosent. Tilsvarende andel er ca. 8 prosent for guttene og ca. 5 prosent for jentene på 9. trinn. Det betyr at jentene har svart i noe større grad enn guttene på 9. trinn, mens det er en ubetydelig forskjell på 8. trinn. Oppgavene 25 (tolke diagram) og 46 (sammenligne enheter) har størst andel ubesvarte på nivå 1 både på 8. trinn (8 og 14 prosent) og 9. trinn (9 og 16 prosent). Det er større andel ubesvarte oppgaver på nivå 1 for 9. trinn enn for 8. trinn.

Oppgaveformat og p-verdi

Gjennomsnittlige p-verdier for åpne oppgaver og flervalgsoppgaver er vist i tabell 7. Det er 34 flervalgsoppgaver og 24 åpne oppgaver i prøven for 2012. På 8. trinn har flervalgsoppgavene

(p-verdi 55,8) falt lettere ut sammenlignet med gjennomsnittet for hele prøven (p-verdi 49,1), mens de åpne oppgavene har falt klart vanskeligere ut (p-verdi 39,7).

På 9. trinn er tendensen den samme. Flervalgsoppgavene har p-verdi 62,6 (p-verdi 55,9 for hele prøven) og de åpne oppgavene har p-verdi 46,4. Blant de ti oppgavene med lavest p-verdi er det kun én flervalgsoppgave, og blant de ti med høyest p-verdi er det syv flervalgsoppgaver (se tabell 4 og 5). At flervalgsoppgavene har høyere gjennomsnittlig p-verdi enn de åpne oppgavene, kan ha en naturlig årsak. Prøven skal ha overvekt av flervalgsoppgaver, men de skal ikke ha lavere p-verdi enn 25 fra piloteringen. Samtidig skal prøven ikke ha «takeffekt». Det betyr at det skal være med oppgaver med lavere p-verdi enn 25, og da må alle disse være åpne oppgaver (se figur 1).

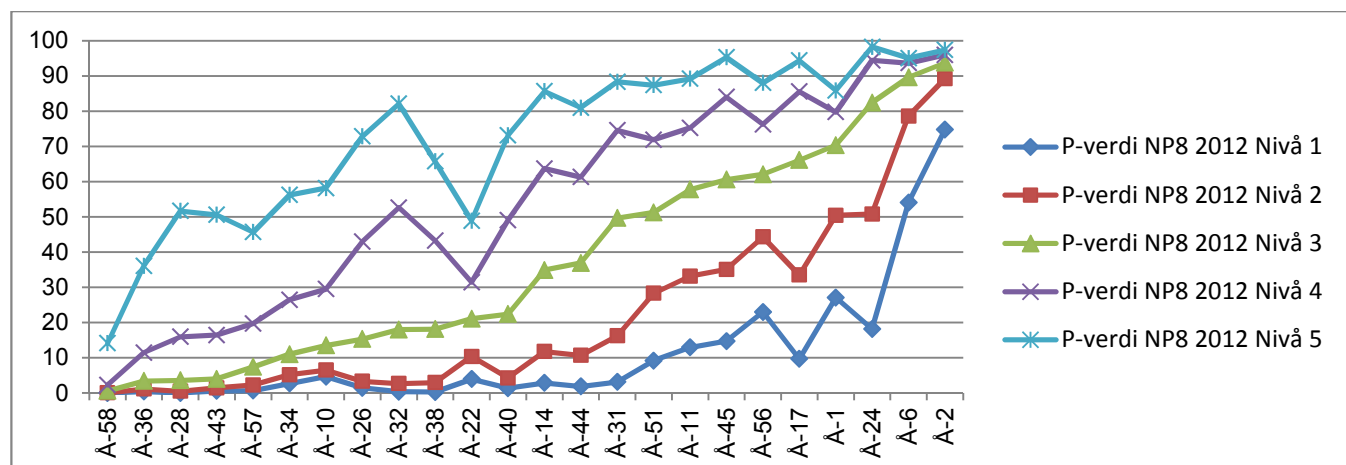
P-verdien for flervalgsoppgavene er høyere enn for de åpne oppgavene på alle fem mestringsnivåer (se tabell 7). Forskjellen i p-verdi er størst på nivå 4 i favør av flervalgsoppgavene. På 8. trinn er forskjellen 21,9 prosentpoeng, og på 9. trinn er forskjellen 20,9 prosentpoeng. Samtidig som p-verdien er høyest for flervalgsoppgavene, finner vi også færrest andel ubesvarte innenfor dette oppgaveformatet. Dette gjelder begge trinn (se tabell 14).

Tabell 7 Løsningsprosent på åpne oppgaver og flervalgsoppgaver for 8. og 9. trinn, både på mestringsnivå og for alle elever. $N(8) = 19\ 609$ og $N(9) = 19\ 370$

Mestringsnivå	Gjennomsnittlig p-verdi			
	8. trinn		9. trinn	
	Åpen	Flervalg	Åpen	Flervalg
1 (0 - 11) p	11,2	18,5	11,3	18,5
2 (12 - 20) p	21,8	32,3	22,4	32,5
3 (21 - 34) p	37,2	54,4	38,1	54,9
4 (35 - 43) p	54,0	75,9	54,9	75,8
5 (44 - 58) p	72,5	90,1	74,5	90,4
Alle elever	39,7	55,8	46,4	62,6

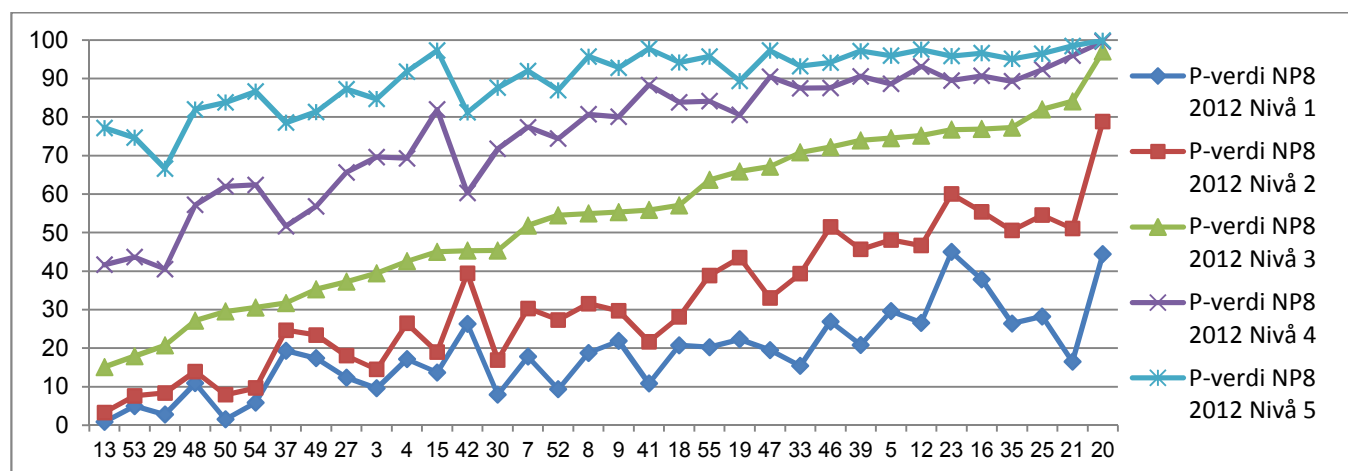
Analysen viser at elevene på mestringsnivå 1 på 8. trinn i gjennomsnitt gjør det like godt som elevene på 9. trinn innenfor begge oppgaveformatene. På de andre nivåene er det i de åpne oppgavene elevene på 9. trinn gjør det bedre enn elevene på 8. trinn.

Figur 8 P-verdier på mestringsnivå for de åpne oppgavene, 8. trinn 2012. Sortert fra laveste til høyeste p-verdi ut fra resultatet til elevene på mestringsnivå 3



Figur 8 viser i hvor stor grad elevene på de ulike mestringsnivåene har løst de åpne oppgavene, og figur 9 viser tilsvarende for flervalgsoppgavene. Resultatet for 9. trinn er omtrent identisk og er derfor ikke tatt med her. Det er godt samsvar mellom mestringsnivået elevene er plassert på og mestringsnivåets p-verdi for de enkelte oppgavene. Elevene på mestringsnivå 5 får færrest poeng på de åpne oppgavene som hadde lavest p-verdi på hele prøven både for elevene på 8. og 9. trinn (se figur 8 og 1). Oppgavene er omtalt i forbindelse med tabell 4. Det er interessant å følge oppgave 22 som handler om å beregne nytt klokkeslett ut fra et resonnement rundt tid. Det var 49 prosent av elevene på 8. trinn nivå 5 som løste oppgaven og på 9. trinn nivå 5 var p-verdien 55. Målestokkoppgaven (oppgave 36) har p-verdi 36 på 8. trinn nivå 5 og p-verdi 44 tilsvarende på 9. trinn.

Figur 9 P-verdier på mestringsnivå for flervalgsoppgavene, 8. trinn 2012. Sortert fra laveste til høyeste p-verdi ut fra resultatet til elevene på mestringsnivå 3



Den vanskeligste flervalgsoppgaven for elevene på mestringsnivå 5 (oppgave 29), er en interaktiv oppgave med gjennomsnitt som elever med evne til analyse, resonnement og refleksjon, kan løse på en effektiv måte. Det var 66 prosent av elevene på 8. trinn nivå 5 som greide denne oppgaven og 67 prosent av elevene på 9. trinn nivå 5. En av de vanskeligste flervalgsoppgavene for elevene på nivå 4 (oppgave 13) handler om å sortere fire brøker (to ekte og to uekte) etter størrelse. I denne oppgaven gjorde elevene på 8. trinn nivå 4 (p-verdi 42) det bedre enn elevene på 9. trinn nivå 4 (p-verdi 38).

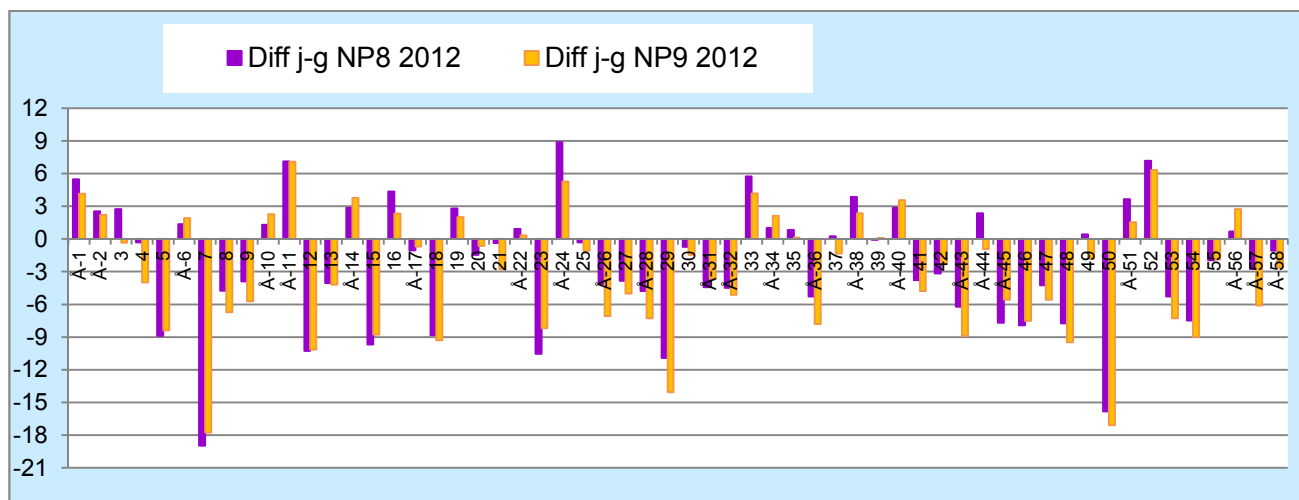
Kjønnforskjeller

Guttene gjorde det signifikant^{iv} bedre enn jentene på prøven som helhet. Forskjellen er 2,2 prosentpoeng på 8. trinn og 3,1 prosentpoeng på 9. trinn i guttenes favør.

På 8. trinn er guttene signifikant bedre enn jentene i 30 oppgaver og i 33 oppgaver på 9. trinn. I mer enn dobbelt så mange flervalgsoppgaver som åpne oppgaver gjør guttene det signifikant bedre enn jentene. Jentene gjør det signifikant bedre enn guttene i 16 oppgaver på 8. trinn og i 17 oppgaver på 9. trinn, og her er tendensen motsatt. Jentene gjør det signifikant bedre enn guttene i mer enn dobbelt så mange åpne oppgaver som flervalgsoppgaver (se figur 10).

^{iv} Signifikant – forskjellen er så stor at den ikke kan skyldes tilfeldigheter. 5 % signifikansnivå.

Figur 10 Forskjell i p-verdi for jenter og gutter for hver oppgave. Positive søyler for oppgaver hvor jentene gjør det bedre enn guttene, og negative søyler hvor guttene gjør det bedre enn jentene. Fiolette søyler for resultater på 8. trinn og gule for 9. trinn. Å betyr åpne oppgaver. Oppgavenummer på horisontal akse og prosentpoeng på vertikal akse. $N(8)j = 9\ 548$, $N(8)g = 10\ 061$, $N(9)j = 9\ 482$ og $N(9)g = 9\ 888$



Største forskjell i jentefavør er 9 prosentpoeng på 8. trinn og 7 prosentpoeng på 9. trinn. ”Gutteoppgavene” har opp til 19 prosentpoeng forskjell i guttefavør på 8. trinn, mens den største forskjellen på 9. trinn er ca. 18 prosentpoeng. Den gjennomsnittlige kjønnsforskjellen på prøven for 2012 er litt mindre enn i prøven for 2011, men stabil i guttefavør som den har vært alle år siden 2007, og på begge trinn siden 2010. Små kjønnsforskjeller kan vi også lese av resultatet for de norske elevene i matematikk i TIMSS 2007 (Grønmo mfl. 2008), mens rapporten etter gjennomføring av TIMSS i 2011, viser ingen signifikant forskjell mellom kjønn, verken i matematikk eller naturfag (Grønmo mfl. 2012).

En nasjonal prøve i regning er ikke en prøve i matematikk, men har fokus på anvendelse av grunnleggende ferdigheter. Derfor er det kanskje mer naturlig å sammenligne med resultatene fra PISA, som ikke er en test i oppnådde kompetansemål i fag. I PISA 2006 (Kjærnsli mfl. 2007) gjør guttene i Norge det litt bedre enn jentene, og forskjellen er statistisk signifikant. Det har vist seg at guttene ofte er flinkere enn jentene til å anvende kunnskap, mens jentene er flinkere rent regneteknisk (Kjærnsli mfl. 2007). De fleste oppgavene i den nasjonale prøven i regning tester anvendelse av kunnskap. Dette kan kanskje være en medvirkende årsak til at guttene gjør det litt bedre enn jentene på prøven.

Også resultatene etter PISA 2009 viser at guttene skårer bedre enn jentene i matematikk. Dette gjelder så å si alle OECD-landene (Kjærnsli mfl. 2010). Gjennomgående er kjønnsforskjellene i de nordiske landene små og ikke-signifikante, men unntaket er Danmark hvor guttene skårer 16 poeng høyere enn jentene. Kjærnsli mfl. 2010, sier at dette er et trekk som har vist seg i alle PISA-undersøkelsene. Det er ingenting som tyder på at kjønnsforskjeller i matematikk innenfor OECD-området totalt sett har minket i de årene PISA har vært gjennomført. I de aller fleste land har kjønnsforskjellene ikke endret seg vesentlig over tid.

Tabell 8 viser at den gjennomsnittlige kjønnsforskjellen i guttenes favør på 8. trinn er 1,3 poeng (2,2 prosentpoeng) og på 9. trinn 1,8 poeng (3,1 prosentpoeng). Dette utgjør 11 prosent av standardavviket på 8. trinn og 15 prosent av standardavviket på 9. trinn, og viser at betydningen av kjønn er lav på begge trinnene. Forskjellen er allikevel signifikant. Det er større spredning i resultatene for guttene enn jentene både på 8. og 9. trinn. På 8. trinn er standardavviket til guttene 8 prosent større enn jentenes, og på 9. trinn 7 prosent større.

Tabell 8 Gjennomsnittspoeng, p-verdier og forskjeller i skåre for jenter og gutter på 8. og 9. trinn. $N(8)j = 9\,548$, $N(8)g = 10\,061$, $N(9)j = 9\,482$ og $N(9)g = 9\,888$

	Gjennomsnittsverdier				Effektstørrelse 8. trinn i forhold til 9. trinn
	Poeng 8. trinn	P-verdi 8. trinn	Poeng 9. trinn	P-verdi 9. trinn	
Alle	28,5 p	49,1	32,4 p	55,9	0,33
Jenter	27,8 p	48,0	31,5 p	54,3	0,32
Gutter	29,1 p	50,2	33,3 p	57,4	0,34
*Diff. j-g	-1,3 p	-2,2	-1,8 p	-3,1	

*Signifikant forskjell mht. kjønn

Effektstørrelsene viser at det på denne prøven har middels betydning om du er elev på 8. eller 9. trinn, jente på 8. eller 9. trinn eller gutt på 8. eller 9. trinn. Effektstørrelsene viser imidlertid at betydningen av et års ekstra skolegang er litt større for guttene enn for jentene (se tabell 8).

I versjon 1 av prøven deltok 9 548 jenter og 10 061 gutter på 8. trinn og 9 482 jenter og 9 888 gutter på 9. trinn. På høyeste nivå, mestringsnivå 5, finner vi 9,8 prosent av jentene på 8. trinn og 14,2 prosent av guttene. På 9. trinn finner vi 17,8 prosent av jentene og 24,0 prosent av guttene på det høyeste nivået (se tabell 9). Det er 5 gutter og 1 jente på 8. trinn og 10 gutter og 3 jenter på 9. trinn som har alt rett på versjon 1 av prøven.

Mestringsnivå 1 består av 5,9 prosent av jentene og 6,2 prosent av guttene på 8. trinn og 3,7 prosent av både jentene og guttene på 9. trinn. Tabell 9 viser at det er små kjønnsforskjeller på alle nivåene. Guttene på nivå 5 gjør det litt bedre enn jentene. På alle nivå unntatt nivå 1, presterer i gjennomsnitt elevene på 9. trinn litt bedre enn elevene på 8. trinn.

Tabell 9 Gjennomsnittspoengsum og prosentvis fordeling av jenter og gutter på mestringsnivå for 8. og 9. trinn. $N(8)j = 9\,548$, $N(8)g = 10\,061$, $N(9)j = 9\,482$ og $N(9)g = 9\,888$

Mestringsnivå	Poeng i gjennomsnitt 8. trinn		Prosent av jentene 8. trinn	Prosent av guttene 8. trinn	Poeng i gjennomsnitt 9. trinn		Prosent av jentene 9. trinn	Prosent av guttene 9. trinn
	Jenter	Gutter			Jenter	Gutter		
1 (0 - 11) p	9,0	9,0	5,9	6,2	9,1	8,9	3,7	3,7
2 (12 - 20) p	16,3	16,2	23,1	21,8	16,5	16,3	16,1	14,4
3 (21 - 34) p	27,4	27,5	42,1	37,6	27,7	27,9	38,7	33,6
4 (35 - 43) p	38,6	38,9	19,0	20,2	38,9	39,0	23,8	24,3
5 (44 - 58) p	47,7	48,3	9,8	14,2	48,2	48,9	17,8	24,0

Den største forskjellen i gjennomsnittlig poengsum for jenter og gutter finner vi i gruppen med mestringsnivå 5. Det er vanskelig å si noe om årsaken til dette, men tendensen til at kjønnsforskjellene er tydeligst i gutters favør på det øverste nivået, ser man også i undersøkelser fra USA (Carr mfl. 2007). I undersøkelsen fra 2007 understreker forskerne at dette er bekymringsfullt i forhold til rekruttering av jenter til høyere utdanning i matematikk og realfag generelt. Flere faktorer antas å ha betydning for en gryende kjønnsforskjell allerede blant elever i barneskolen. Det kan for eksempel være hvilke strategier man velger når man skal løse oppgaver, for eksempel hukommelsesstrategier eller enkle oppskrifter (telle på fingrene), det kan handle om tempo, "mental rotasjon" av informasjon, elevenes faglige selvtillit og hvilke forventninger elevene selv og omverdenen har til den enkelte elev.

En del forskning viser resultater som tyder på at gutter oppfatter seg selv som bedre i matematikk enn jenter (Kimball 1994, Li 1999), og annen forskning viser til resultater om at

gutter har mer positiv holdning enn jenter til matematikk (Hannula mfl. 2007). Lovisa Sumpter (2009) har forsket på lærernes oppfatning av hvordan jenter og gutter arbeider med matematikkfaget, og hun påpeker ulikheter i bruk av strategier.

Både Carr (2007) og Grønmo (2005) nevner i sine rapporter at automatisering av grunnleggende ferdigheter bør vektlegges. Da kan mental kapasitet frigis og tankekapasiteten brukes på mer kognitivt krevende oppgaver.

I artikkelen *Kjønnsforskjeller i motivasjon, læringsstrategibruk og selvregulering i naturfag* (Elstad & Turmo 2007), peker Elstad og Turmo på at gutter har en tendens til å ha større fokus på såkalt "prestasjonsmotivasjon" enn jenter, det vil si at de er mer motiverte for å yte sitt beste i det som kan oppfattes som en konkurranse situasjon. Dette støttes av Grønmo (2008) som nevner at gutter både kan være mer konkurranseorienterte enn jenter, og mer opptatt av å bruke kvantitative begreper. Gjennom daglige aktiviteter dannes mentale representasjoner som igjen overføres til matematisk læring (Pitta-Pantazi mfl. 2004). Derfor kan viktige begreper innenfor den tidlige matematikkopplæringen, slike som større, mindre og lengst, i større grad bli innlært hos gutter enn hos jenter. Dette kan gi gutter et forsprang når det gjelder regning.

Kjønnsforskjeller innen områdene tall, måling og statistikk

Siden en oppgave kan inneholde elementer fra flere områder, er det noen ganger vanskelig å plassere en oppgave på ett bestemt område. Ut i fra de plasseringene vi har gjort, gjør guttene det bedre enn jentene innen alle de tre områdene. Kjønnsforskjellene er signifikante bortsett fra på området statistikk 8. trinn. Forskjellen er størst innenfor måling på begge trinn. Effektanalysen viser at det har liten betydning for resultatet innen alle områdene på begge trinn, om du er en gutt eller jente ($e < 0,3$).

Når vi sammenligner resultatene for 8. og 9. trinn, er betydningen av hvilket trinn du er på, middels for alle tre områdene, $0,3 < e < 0,8$ (se tabell 10).

Tabell 10 Gjennomsnittlige p-verdier, effektstørrelser og forskjeller i prosentpoeng for alle og for jenter og gutter innenfor tall, måling og statistikk, både på 8. og 9. trinn.
 $N(8)j = 9\ 548$, $N(8)g = 10\ 061$, $N(9)j = 9\ 482$ og $N(9)g = 9\ 888$

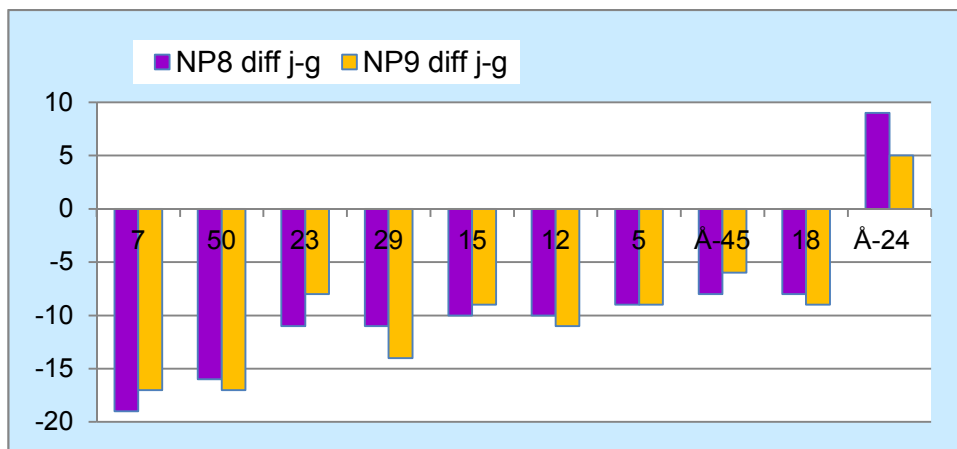
Område	8. trinn				9. trinn				8. trinn	9. trinn	Effektstørrelse 8. / 9. trinn alle elever
	P-verdi alle	P-verdi jenter	P-verdi gutter	Diff j-g	P-verdi alle	P-verdi jenter	P-verdi gutter	Diff j-g	Effektstørrelse j/g	Effektstørrelse j/g	
Tall* (n=23)	50,5	49,4	51,5	-2,1	56,9	55,3	58,4	-3,1	0,10	0,15	0,30
Måling* (n=21)	40,4	38,6	42,2	-3,6	48,0	45,8	50,2	-4,4	0,16	0,18	0,32
Statistikk* (n=14)	59,9	59,8	60,1	-0,3	66,1	65,7	66,5	-0,8	0,02	0,04	0,31

*Signifikant forskjell mht. kjønn for alle områder bortsett fra statistikk 8. trinn

Kjønnsforskjeller i p-verdier på enkeltoppgaver

Den største forskjellen i guttefavør sortert etter resultatene på 8. trinn, er i to målingsoppgaver (oppgave 7 og 50), en statistikkoppgave (oppgave 23) og sju tallopgaver (se figur 11). For 8. trinn er 19 prosentpoeng største forskjell i p-verdi mellom jenter og gutter i en oppgave, og for 9. trinn er største forskjell 17 prosentpoeng. Oppgavene har p-verdier fra 26 til 76. Det er altså ikke blant de aller vanskeligste oppgavene vi har de største forskjellene mellom kjønnene.

Figur 11 De ti oppgavene med størst forskjell i p-verdi for jenter og gutter sortert etter resultatene på 8. trinn. Positive søyler for oppgaver hvor jentene gjør det bedre enn guttene, og negative søyler hvor guttene gjør det bedre enn jentene. Fiolette søyler for resultater på 8. trinn og gule for 9. trinn. Å betyr åpne oppgaver. Oppgavenummer på horisontal akse og prosentpoeng på vertikal akse



I oppgave 7 (p-verdi 55 på 8. trinn og 64 på 9. trinn) skal man regne med minutter og sekunder og i oppgave 50 (p-verdi 36 på 8. trinn og 42 på 9. trinn) er det ren omgjøring av måleenheter. Omgjøring av måleenheter og beregninger i slike sammenhenger (se figur 12), å regne med prosent og desimaltall, sammenligne brøker og foreta beregninger ved hjelp av målestokk, er områder der guttene har skåret bedre enn jentene. Det er en tendens til at forskjellen er størst i problemløsningsoppgaver hvor regneteknikk også kan erstattes av resonnement og strategitenkning, men det er vanskelig å si noe generelt om dette.

Figur 12 Eksempel på en oppgave med stor forskjell i p-verdi i guttefavør både på 8. og 9. trinn. Oppgave 53 fra 2011. P-verdi 57 for gutter og 41 for jenter på 8. trinn, og 64 for gutter og 47 for jenter på 9. trinn

En trailer er lastet med 15 like tunge kasser. Lasset veier til sammen 3 tonn.

Hvor mye veier én kasse?

2 kg
 5 kg
 200 kg
 500 kg



Blant de ti oppgavene med størst forskjell i p-verdi mellom kjønnene, er det en oppgave i jentefavør (oppgave 24, p-verdier 76 og 80 på 8. og 9. trinn). Oppgaven er en tradisjonell «kjøp og salg» oppgave. Det samme kan sies om oppgavene 1, 33 og 52 hvor jentene gjør det bedre enn guttene på begge trinn. Å lese, tolke og gjøre beregninger ut fra tabeller og grafer, løse ferdig oppstilte oppgaver og utføre beregninger i praktiske sammenhenger hvor kjente algoritmer kan benyttes, synes å være typisk for oppgaver hvor jentene gjør det bedre enn guttene på denne prøven. Er det derimot måleenheter som skal bearbeides i slike oppgaver, blir det raskt en oppgave i guttefavør. Dette stemmer med tidligere års resultater. Se figur 6

som eksempel på oppgave i jentefavør, - forskjellen i jentefavør var 7 prosentpoeng både på 8. og 9. trinn i denne oppgaven i 2011.

Kjønnsforskjeller og oppgaveformat

Analysen av resultatene viser at guttene på både 8. og 9. trinn i gjennomsnitt løser flere flervalgsoppgaver enn jentene, og at guttene på 9. trinn også løser flere åpne oppgaver enn jentene. Effekttørrelsene for åpne oppgaver er 1 og 3 prosent av standardavviket på henholdsvis 8. og 9. trinn, og forteller at kjønn har ubetydelig betydning når det gjelder å løse åpne oppgaver. For flervalgsoppgavene er forskjellen 18 og 22 prosent av standardavviket. Forskjellen er større enn i de åpne oppgavene, men innenfor rammen av lav betydning. Selv om effekttørrelsene viser lav betydning, er imidlertid forskjellen mellom kjønnene signifikant for flervalgsoppgavene og også for åpne oppgaver på 9. trinn. Vi kan allikevel slå fast at oppgaveformatet bidrar i svært liten grad til å skape kjønnsforskjellene (se tabell 11).

Tabell 11 P-verdier, forskjeller i prosentpoeng og effekttørrelser for jenter og gutter for åpne oppgaver og flervalgsoppgaver, både for 8. og 9. trinn. De to siste kolonnene sammenligner trinnene. $N(8)j = 9\ 548$, $N(8)g = 10\ 061$, $N(9)j = 9\ 482$, $N(9)g = 9\ 888$

Område	8.trinn				9.trinn				8.trinn	9.trinn	Effektstørrelse 8./9. jenter	Effektstørrelse 8./9. gutter
	P-verdi alle	P-verdi jenter	P-verdi gutter	Diff. j-g	P-verdi alle	P-verdi jenter	P-verdi gutter	Diff. j-g	Effektstørrelse j/g	Effektstørrelse j/g		
Åpne* (24)	39,7	39,8	39,7	0,1	46,4	46,0	46,7	-0,7	0,01	0,03	0,33	0,34
Flervalg* (34)	55,8	53,8	57,6	-3,8	62,6	60,2	64,9	-4,7	0,18	0,22	0,30	0,32

*Signifikant forskjell mht. kjønn for alle områder bortsett fra åpne oppgaver 8. trinn

Når vi sammenligner resultatene til jentene på 8.trinn og 9. trinn, er effekttørrelsen for de åpne oppgavene 33 prosent av standardavviket. For guttene er tilsvarende resultat 34 prosent av standardavviket. Det har derfor middels betydning hvilket trinn elevene er på når det gjelder å løse åpne oppgaver. For flervalgsoppgavene er effekttørrelsene litt lavere, men fremdeles innenfor området middels betydning. Størst forskjell når det gjelder å løse flervalgsoppgaver er det mellom gutter på 8. og 9. trinn (se tabell 11).

Det kan det være interessant å se på forskjeller i p-verdier for de oppgavene hvor forskjellen er størst (se tabell 12).

Tabell 12 Åpne oppgaver hvor forskjellen i p-verdier for jenter og gutter er størst, sortert etter 8. trinn. $N(8) = 19\ 609$, $N(9) = 19\ 370$

Oppgave	Innhold	Område	Format	8. trinn		9. trinn		Forventet vanskelighetsgrad
				P-verdi	Diff j-g	P-verdi	Diff j-g	
24*	Regneoperasjoner med hele tall	Tall	Åpne	75,7	8,9	79,6	5,3	1
45*	Brøk/prosent	Tall	Åpne	60,8	-7,7	68,2	-5,6	3
11*	Tabell/diagram	Statistikk	Åpne	56,7	7,1	66,7	7,1	3
43*	Gjennomsnitt	Statistikk	Åpne	11,3	-6,2	16,0	-8,9	5
1*	Regneoperasjoner med hele tall	Tall	Åpne	67,0	5,5	70,2	4,2	1
36*	Sammensatt problem med målestokk	Måling	Åpne	8,2	-5,3	14,2	-7,8	5

*Signifikant forskjell mht. kjønn på både 8. og 9.trinn

Resultatene viser at jentene gjør det bedre enn guttene i oppgaver med regneoperasjoner med hele tall, samt én tabell/diagramoppgave. Dette gjelder begge trinn. To av disse oppgavene er, ut fra kompetansemålene i Kunnskapsløftet (LK06), vurdert som lett og en som middels vanskelig. Guttene på begge trinn gjør det bedre enn jentene i en oppgave av middels vanskelighetsgrad som handler om sammenhengen brøk/prosent, og i to oppgaver av høy vanskelighetsgrad med begrepet gjennomsnitt og temaet målestokk.

Tabell 13 viser flervalgsoppgaver med størst forskjell i løsningsprosent for jenter og gutter, sortert etter 8. trinn. Guttene gjør det bedre enn jentene i alle de seks oppgavene. Det er to oppgaver med lav vanskelighetsgrad, to oppgaver med middels og to med høy vanskelighetsgrad. Oppgavene er omtalt i forbindelse med figur 11.

Tabell 13 Flervalgsoppgaver hvor forskjellen i p-verdier mellom jenter og gutter er størst, sortert etter 8. trinn. $N(8)j = 9\,548$, $N(8)g = 10\,061$, $N(9)j = 9\,482$ og $N(9)g = 9\,888$

Oppgave	Innhold	Område	Format	8. trinn		9. trinn		Forventet vanskelighetsgrad
				P-verdi	Diff j-g	P-verdi	Diff j-g	
7*	Klokke/tid	Måling	Flervalg	54,7	-19,0	63,6	-17,8	2
50*	Måleenheter	Måling	Flervalg	35,9	-15,8	41,6	-17,1	4
29*	Sammensatt problem med sammenligning av data	Tall	Flervalg	26,3	-10,9	33,8	-14,1	5
23*	Tabell/diagram	Statistikk	Flervalg	75,9	-10,5	81,8	-8,2	3
12*	Posisjonssystemet	Tall	Flervalg	72,0	-10,3	77,5	-10,1	2
15*	Brøk/prosent	Tall	Flervalg	50,8	-9,7	62,2	-8,8	3

*Signifikant forskjell mht. kjønn på både 8. og 9. trinn

Ubesvarte oppgaver

Tabell 14 viser prosent ubesvarte oppgaver i gjennomsnitt for alle elevene og for jenter og gutter hver for seg på begge trinnene. I gjennomsnitt er det 4,0 prosent ubesvarte oppgaver for 8. trinn og 3,9 prosent ubesvarte for 9. trinn. Største andel ubesvarte finner vi blant de åpne oppgavene. På 8. trinn har elevene i gjennomsnitt 4,9 prosentpoeng flere ubesvarte åpne enn flervalgsoppgaver. Tilsvarende forskjell på 9. trinn er 5,1 prosentpoeng. På 8. trinn har jentene i gjennomsnitt 1,6 prosentpoeng flere ubesvarte åpne oppgaver enn guttene, mens på 9. trinn har de i gjennomsnitt 1,8 prosentpoeng flere ubesvarte. Forskjellen mellom jenter og gutter når det gjelder ubesvarte flervalgsoppgaver, er marginal (se tabell 14). Totalt er forskjellen mellom jenter og gutter på ubesvarte oppgaver så liten, at det er vanskelig å bruke dette som en årsak til at guttene skårer bedre enn jentene på hele prøven. Vi legger imidlertid merke til at guttene på 9. trinn har færrest ubesvarte oppgaver innenfor alle områdene når vi sammenligner alle gruppene.

Tabell 14 Prosent ubesvarte oppgaver i gjennomsnitt på hele prøven og på åpne oppgaver, flervalgsoppgaver, tall-, måling- og statistikkoppgaver for 8. og 9. trinn.
 $N(8)j = 9\ 548$, $N(8)g = 10\ 061$, $N(9)j = 9\ 482$ og $N(9)g = 9\ 888$

Format og område	8. trinn			9. trinn		
	Prosent ubesvarte alle elever	Prosent ubesvarte jenter	Prosent ubesvarte gutter	Prosent ubesvarte alle elever	Prosent ubesvarte jenter	Prosent ubesvarte gutter
Åpne (24)**	6,9	7,7	6,1	6,9	7,8	6,0
Flervalg (34)*	2,0	2,1	1,9	1,8	1,9	1,7
Tall (23)*	2,7	2,9	2,5	2,6	2,8	2,4
Måling (21)*	5,9	6,6	5,2	5,8	6,6	5,0
Statistikk (14)**	3,4	3,7	3,1	3,2	3,6	2,9
Alle oppgaver (58)*	4,0	4,4	3,6	3,9	4,4	3,5

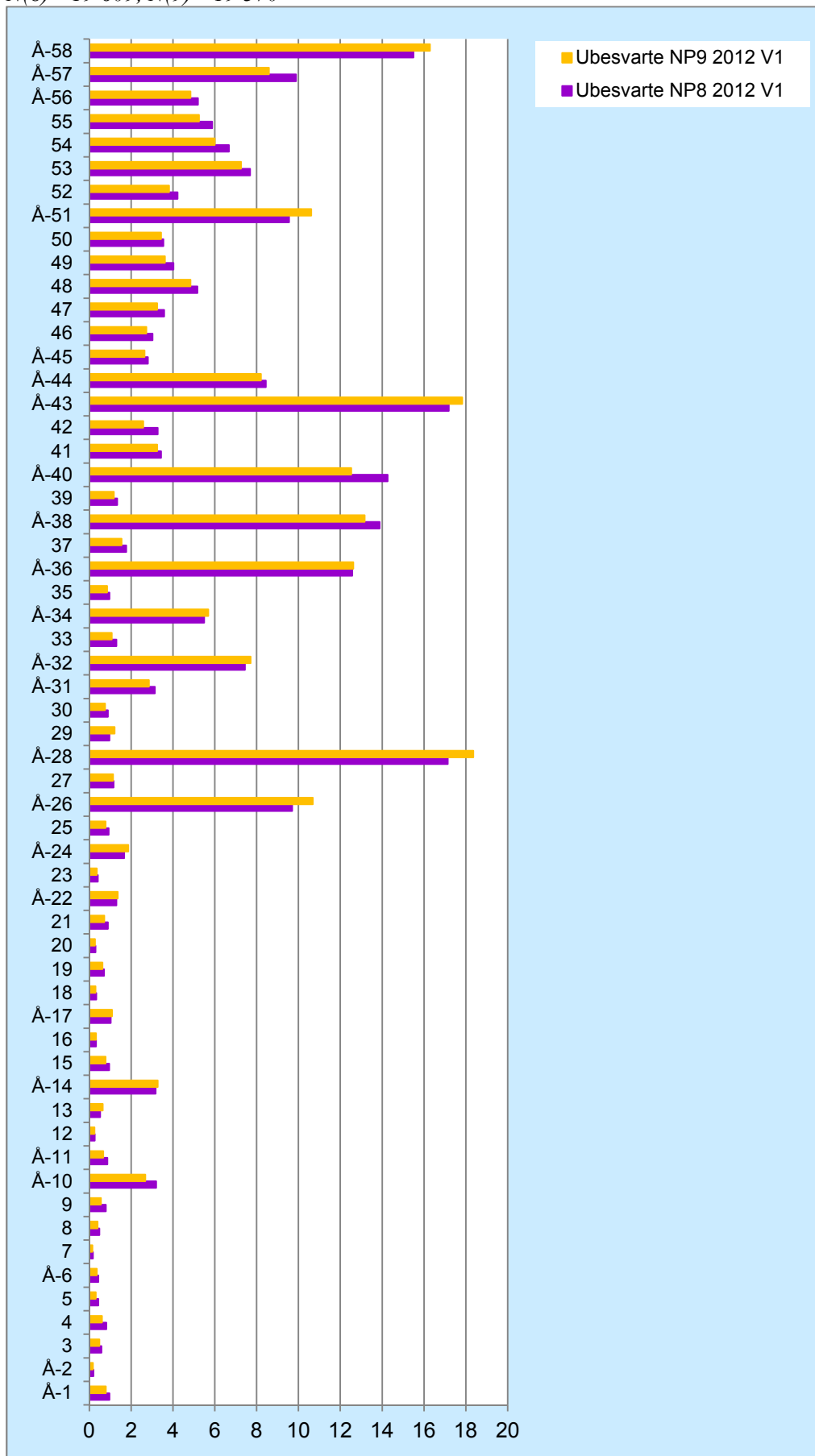
*Signifikant forskjell mht. kjønn på 8. og 9. trinn

**Signifikant forskjell mht. kjønn bare på 9. trinn

Det er en tendens til at andelen ubesvarte oppgaver øker mot slutten av prøven (se figur 13 og 14). Det er også en tendens til at andelen jenter som ikke svarer, øker mer i siste del av prøven enn andel gutter som ikke svarer. Analysen viser ni oppgaver hvor andelen ubesvarte er mer enn 10 prosent, og hvor jentene har til dels betydelig flere ubesvarte enn guttene. Alle unntatt oppgave 57 er åpne oppgaver i kontekst som krever regneoperasjoner i flere trinn for å komme frem til svaret, og alle unntatt oppgave 51 har p-verdi lavere enn 28. Det er fem talloppgaver og fire målingsoppgaver, og guttene gjør det bedre enn jentene i seks av oppgavene (se tabell 1 og 17).

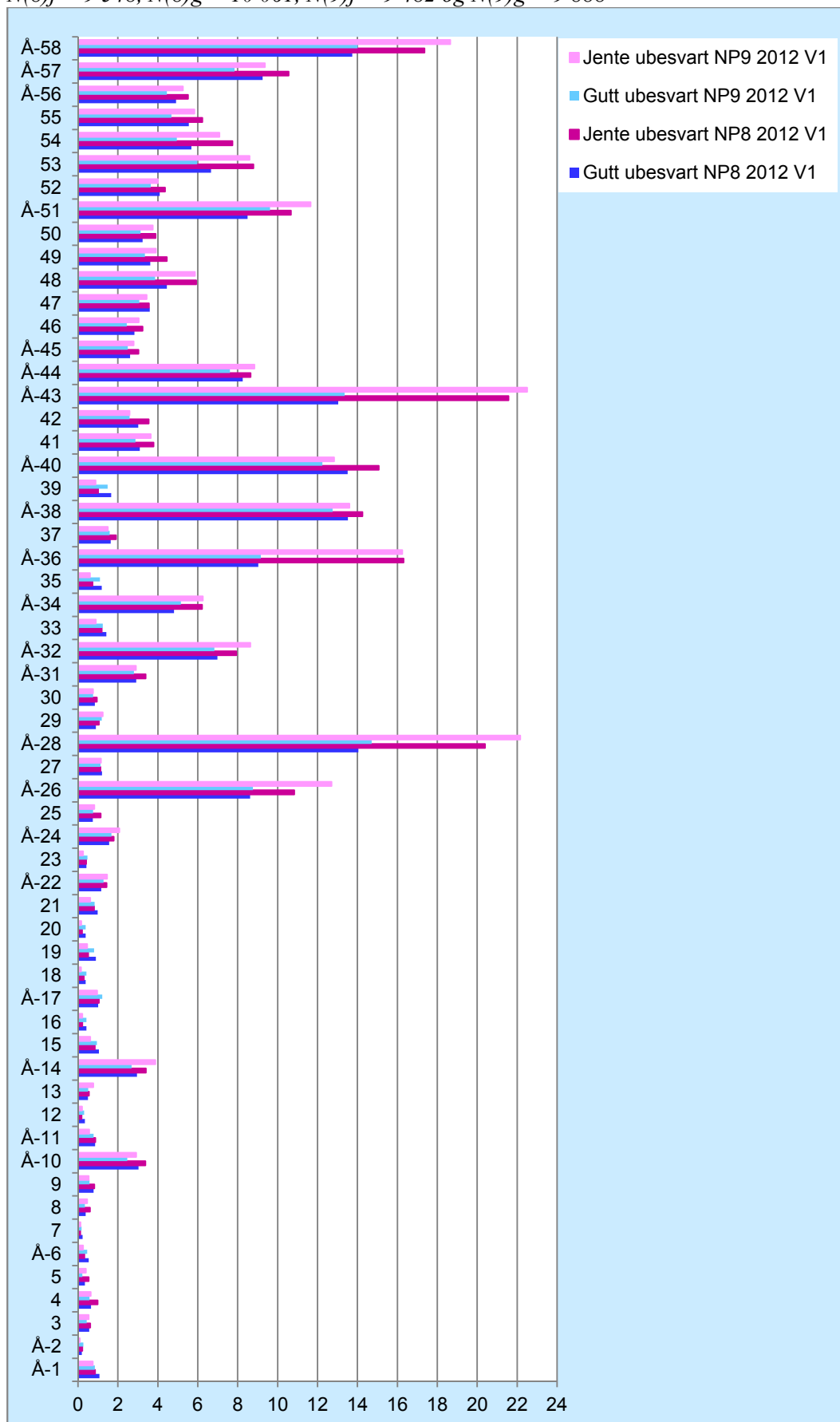
Figur 13

Prosent ubesvart for hver oppgave for elever på 8. og 9. trinn. Fiolet, prosent 8. trinn ubesvart og gult 9. trinn. Å betyr åpen oppgave. De andre er flervalgsoppgaver. Oppgavenummer langs vertikal akse, prosent langs horisontal akse. $N(8) = 19\ 609$, $N(9) = 19\ 370$



Figur 14

Prosent ubesvart for hver oppgave for jenter og gutter på 8. og 9. trinn. Rosa og rødt for jenter og blåfarger for gutter. Å betyr åpen oppgave. De andre er flervalgsoppgaver. Oppgavenummer langs vertikal akse, prosent langs horisontal akse. $N(8)j = 9\ 548$, $N(8)g = 10\ 061$, $N(9)j = 9\ 482$ og $N(9)g = 9\ 888$



Teknisk rapport

De tekniske rapportene for 8. og 9. trinn består av to tabeller med sammenfattende informasjon (tabell 15 og 16), og en detaljert item-analyse av alle oppgavene (tabell 17). Tabell 16 viser at gjennomsnittlig poengsum for guttene på prøven er høyere enn for jentene på både 8. og 9. trinn. Forskjellen er henholdsvis 2,2 prosentpoeng på 8. trinn og 3,1 prosentpoeng på 9. trinn.

Tabell 15 Tekniske data for prøven på 8. og 9. trinn

Trinn	Antall oppgaver	Flervalgsoppgaver	Interaktive oppgaver	Reliabilitet ^y	Gjennomsnittlig p-verdi	Gjennomsnittlig poengsum	Std.avvik	Std.feil til gj.sn.
8.trinn*	58	34	11	0,93	49,1	28,5	11,6	0,08
9.trinn*	58	34	11	0,94	55,9	32,4	12,0	0,09

*Signifikant forskjell mht. kjønn

Tabell 16 Antall elever i utvalget for 8. og 9. trinn. Gjennomsnittlige resultater som poeng og prosent

	8. trinn	9. trinn
Antall elever	19 609	19 370
Antall gutter	10 061	9 888
Antall jenter	9 548	9 482
Gjennomsnitt gutter, poeng og prosent	29,1 (50,2 %)	33,3 (57,4 %)
Gjennomsnitt jenter, poeng og prosent	27,8 (48,0 %)	31,5 (54,3%)
Gjennomsnitt alle, poeng og prosent	28,5 (49,1 %)	32,4 (55,9 %)

Item-analyse av alle oppgavene

Guttene gjør det signifikant bedre enn jentene på prøven, både på 8. og 9. trinn. Tallene i analysen viser at oppgavene har fungert meget bra.

- Av 58 oppgaver fungerer 54 meget bra ($D > 0,30$). De fire andre har diskriminering^{vi} fra 0,17 til 0,28. På 9. trinn er det bare to av de samme oppgavene som har $D < 0,30$, oppgave 2 og 6, med diskriminering henholdsvis 0,18 og 0,25. Begge disse oppgavene har svært høy p-verdi
- I teknisk rapport er p-verdiene skrevet som hele prosenttall. I kolonnen Diff. J-G er enheten prosentpoeng. Positive verdier betyr at jentene gjør det bedre enn guttene, og negative at guttene gjør det bedre enn jentene
- FV betyr flervalgsoppgave. A, B, C og D er svaralternativene
- Å betyr åpen oppgave. 0 betyr galt svar og 1 betyr rett svar

Tabell 17 Resultater for 8. og 9. trinn. Item-analyse for alle oppgavene. Svarfordeling angitt i prosent og elevenes dyktighet i poeng. Med dyktighet menes gjennomsnittlig poengsum for de elevene som har svart det bestemte alternativet i flervalgsoppgavene, og har fått rett eller galt i de åpne oppgavene. D-verdi betyr oppgavens diskriminering (korrelasjon med summen av poeng). Guttene har høyest løsningsprosent i de oppgavene hvor differansen er negativ.

^y Reliabilitet – pålitelighet - et mål for kvalitet – målt i form av Chronbachs alpha $> 0,80$

^{vi} Diskriminering: Diskr (D) betyr oppgavens diskriminering, dvs. hvordan oppgaven samsvarer med summen av alle oppgavene.

NP Regning 8.trinn 2012																	
Item	Svarfordeling i %						Dyktighet poeng					D-verdi	p-verdi	Jente	Gutt	Diff. J-G b)	Merk
	FV	A	B	C	D	Ubesvart	A	B	C	D	Ubesvart						
Nr.	Å	0	1				0	1									
1	Å	32	67			1	23	31			15	0,32	67	70	64	5,5	c
2	Å	8	92			0	22	29			18	0,17	92	94	91	2,5	ac
3	FV	16	22	19	43	1	20	23	26	35	18	0,52	43	45	42	2,7	c
4	FV	49	28	13	10	1	34	24	22	24	17	0,45	49	48	49	-0,3	
5	FV	7	16	71	5	0	19	21	32	24	18	0,42	71	67	76	-8,9	c
6	Å	14	86			0	21	30			13	0,26	86	87	86	1,4	ac
7	FV	55	14	23	8	0	33	19	25	22	12	0,46	55	45	64	-19,0	c
8	FV	19	57	11	13	1	25	33	22	17	19	0,48	57	55	60	-4,7	c
9	FV	8	18	57	17	1	22	21	33	24	15	0,47	57	55	59	-3,9	c
10	Å	77	20			3	26	38			19	0,39	20	21	19	1,3	c
11	Å	42	57			1	23	33			19	0,43	57	60	53	7,1	c
12	FV	28	72			0	20	32			13	0,47	72	67	77	-10,3	c
13	FV	76	24			1	25	40			21	0,54	24	22	26	-4,0	c
14	Å	57	40			3	24	36			20	0,53	40	41	38	2,8	c
15	FV	16	51	21	12	1	22	35	20	25	14	0,57	51	46	56	-9,7	c
16	FV	5	75	9	11	0	20	31	20	22	11	0,39	75	77	73	4,4	c
17	Å	36	63			1	21	33			13	0,51	63	62	63	-1,1	
18	FV	4	6	58	31	0	20	21	33	22	12	0,50	58	54	62	-8,8	c
19	FV	6	64	23	7	1	22	32	24	18	13	0,39	64	65	63	2,8	c
20	FV	91	5	1	3	0	30	14	14	16	10	0,40	91	90	91	-1,5	c
21	FV	77	11	9	3	1	32	17	18	20	13	0,51	77	76	77	-0,4	
22	Å	76	23			1	27	35			16	0,30	23	23	23	0,9	
23	FV	76	8	12	4	0	31	19	23	22	11	0,34	76	70	81	-10,5	c
24	Å	24	76			0	18	32			15	0,49	76	80	71	8,9	c
25	FV	8	76	6	8	1	22	31	18	20	13	0,43	76	76	77	-0,3	
26	Å	66	24			10	25	39			21	0,53	24	22	26	-4,1	c
27	FV	15	27	14	43	1	22	24	25	35	16	0,49	43	41	45	-3,8	c
28	Å	72	11			17	27	43			24	0,45	11	8	13	-4,8	c
29	FV	73	26			1	26	37			23	0,43	26	21	32	-10,9	c
30	FV	14	15	24	47	1	21	20	26	35	13	0,52	47	47	47	-0,7	
31	Å	48	49			3	22	35			16	0,54	49	47	51	-4,4	c
32	Å	65	28			7	24	39			20	0,59	28	26	30	-4,5	c
33	FV	6	15	66	11	1	17	20	32	25	15	0,47	66	69	64	5,7	c
34	Å	77	18			6	26	39			21	0,40	18	18	17	1,0	
35	FV	5	13	73	9	1	16	21	31	23	13	0,42	73	73	72	0,8	
36	Å	79	8			13	27	42			21	0,36	8	6	11	-5,3	c
37	FV	16	34	39	10	2	28	25	34	24	19	0,35	39	39	39	0,3	
38	Å	62	24			14	25	39			22	0,49	24	26	22	3,9	c
39	FV	16	70	8	5	1	22	32	20	18	13	0,47	70	70	70	-0,1	
40	Å	58	28			14	25	38			21	0,51	28	30	27	2,9	c
41	FV	6	11	23	57	3	21	20	21	34	18	0,59	57	55	59	-3,8	c
42	FV	30	50	13	4	3	26	32	25	29	19	0,28	50	48	52	-3,2	ac
43	Å	72	11			17	27	43			24	0,43	11	8	14	-6,2	c
44	Å	53	39			8	24	36			19	0,51	39	40	38	2,4	c
45	Å	36	61			3	22	33			18	0,48	61	57	65	-7,7	c
46	FV	15	8	70	4	3	22	21	31	24	17	0,40	70	66	74	-7,9	c
47	FV	65	9	8	14	4	33	21	19	21	17	0,53	65	63	67	-4,2	c
48	FV	8	37	14	36	5	22	27	22	36	20	0,48	36	32	39	-7,7	c
49	FV	17	41	23	15	4	28	34	25	22	19	0,39	41	42	41	0,4	
50	FV	61	36			4	24	37			18	0,56	36	28	44	-15,8	c
51	Å	38	52			10	23	34			20	0,45	52	54	50	3,6	c
52	FV	53	18	14	10	4	33	23	21	26	19	0,46	53	57	50	7,2	c
53	FV	25	23	18	27	8	27	25	24	38	22	0,48	27	24	29	-5,3	c
54	FV	19	24	13	37	7	27	23	21	37	20	0,55	37	34	41	-7,5	c
55	FV	6	63	12	12	6	22	33	23	20	21	0,46	63	62	64	-2,0	c
56	Å	33	62			5	23	32			21	0,35	62	62	61	0,7	
57	Å	77	13			10	27	40			22	0,38	13	11	14	-2,8	c
58	Å	83	2			16	28	47			24	0,24	2	2	3	-1,1	ac

Gj.snitt ubesvart: 4,0 %

Gj.snitt: 0,44 49,1 48,0 50,2 -2,2

a) Svak diskriminering D<0,30

b) Positive verdier er kjønnsforskjeller i jenters favør

c) Signifikante forskjeller

Negative verdier er kjønnsforskjeller i gutters favør

NP Regning 9.trinn 2012																	
Item	Svarfordeling i %					Ubesvart	Dyktighet poeng					D-verdi	p-verdi	Jente	Gutt	Diff. J-G b)	Merk
	FV	A	B	C	D		A	B	C	D	Ubesvart						
Nr.	Å	O	1			0	1			Ubesvart							
1	Å	29	70			1	27	35			17	0,30	70	72	68	4,2	c
2	Å	6	94			0	24	33			13	0,18	94	96	93	2,2	ac
3	FV	12	21	17	49	1	22	26	30	39	20	0,52	49	49	49	-0,3	
4	FV	57	22	10	10	1	37	26	24	28	20	0,47	57	55	59	-4,0	c
5	FV	6	14	77	4	0	20	24	35	28	22	0,41	77	72	81	-8,4	c
6	Å	10	90			0	24	33			15	0,25	90	91	89	1,9	ac
7	FV	64	12	19	6	0	37	22	27	24	12	0,48	64	55	72	-17,8	c
8	FV	16	66	10	8	0	27	37	24	19	22	0,49	66	63	70	-6,7	c
9	FV	6	14	63	16	1	25	24	37	26	15	0,49	63	60	66	-5,7	c
10	Å	70	27			3	29	41			21	0,44	27	28	26	2,3	c
11	Å	32	67			1	25	36			17	0,41	67	70	63	7,1	c
12	FV	22	78			0	22	35			13	0,46	78	72	83	-10,1	c
13	FV	68	31			1	28	42			21	0,55	31	29	33	-4,2	c
14	Å	47	50			3	26	39			22	0,53	50	52	48	3,8	c
15	FV	14	62	14	9	1	24	38	21	26	15	0,60	62	58	67	-8,8	c
16	FV	4	77	8	10	0	23	35	23	25	11	0,39	77	78	76	2,3	c
17	Å	30	69			1	23	36			15	0,51	69	69	70	-0,7	
18	FV	3	5	69	23	0	21	24	37	24	11	0,52	69	64	73	-9,3	c
19	FV	5	71	19	4	1	25	35	27	20	14	0,38	71	72	70	2,0	c
20	FV	93	3	1	3	0	34	15	16	18	10	0,38	93	93	93	-0,6	
21	FV	80	9	8	2	1	35	19	22	22	13	0,51	80	79	82	-2,8	c
22	Å	69	30			1	30	38			18	0,33	30	31	30	0,3	
23	FV	82	5	10	3	0	34	22	25	26	11	0,33	82	78	86	-8,2	c
24	Å	20	80			0	21	35			17	0,49	80	82	77	5,3	c
25	FV	6	83	5	5	1	26	35	19	20	14	0,41	83	82	83	-1,1	c
26	Å	58	31			11	28	42			24	0,56	31	27	34	-7,1	c
27	FV	12	23	13	51	1	25	27	27	38	18	0,51	51	48	53	-5,0	c
28	Å	66	16			18	30	46			27	0,48	16	12	19	-7,3	c
29	FV	65	34			1	29	40			25	0,45	34	27	41	-14,1	c
30	FV	8	11	23	57	1	23	23	28	38	14	0,52	57	56	57	-1,5	c
31	Å	40	57			3	25	38			17	0,52	57	55	59	-3,7	c
32	Å	53	39			8	27	42			22	0,61	39	36	41	-5,1	c
33	FV	4	13	73	9	1	18	23	36	28	16	0,45	73	75	71	4,2	c
34	Å	67	27			6	29	42			25	0,47	27	28	25	2,1	c
35	FV	3	10	80	7	1	17	23	35	27	15	0,40	80	80	80	0,1	
36	Å	73	14			13	30	45			23	0,43	14	10	18	-7,8	c
37	FV	15	33	44	8	2	31	28	38	27	22	0,40	44	43	45	-1,4	
38	Å	54	33			13	28	41			24	0,51	33	34	32	2,3	c
39	FV	13	73	8	5	1	25	36	23	21	15	0,49	73	73	73	0,1	
40	Å	49	39			13	27	40			23	0,54	39	41	38	3,6	c
41	FV	5	9	19	64	3	23	22	23	38	21	0,60	64	62	67	-4,8	c
42	FV	23	58	11	5	3	28	36	27	31	20	0,35	58	57	60	-2,2	c
43	Å	66	16			18	30	46			27	0,48	16	11	20	-8,9	c
44	Å	46	46			8	27	39			21	0,51	46	46	47	-0,9	
45	Å	29	68			3	24	36			20	0,48	68	65	71	-5,6	c
46	FV	11	6	77	3	3	24	23	35	26	20	0,39	77	73	81	-7,5	c
47	FV	71	8	7	11	3	37	23	20	23	20	0,56	71	69	74	-5,6	c
48	FV	6	33	12	44	5	24	29	24	39	22	0,52	44	39	49	-9,5	c
49	FV	16	47	21	12	4	31	38	29	24	22	0,42	47	46	48	-1,8	c
50	FV	55	42			3	27	41			22	0,57	42	33	50	-17,1	c
51	Å	33	56			11	26	37			23	0,45	56	57	56	1,5	c
52	FV	61	16	11	8	4	37	27	23	29	23	0,45	61	64	58	6,3	c
53	FV	21	20	16	36	7	30	28	27	41	26	0,52	36	32	39	-7,3	c
54	FV	18	20	11	46	6	30	25	23	40	23	0,57	46	41	50	-9,0	c
55	FV	5	71	10	8	5	25	36	25	21	24	0,46	71	70	72	-1,9	c
56	Å	27	68			5	26	35			24	0,35	68	69	67	2,7	c
57	Å	71	20			9	30	41			24	0,38	20	17	23	-6,1	c
58	Å	79	5			16	32	48			27	0,31	5	4	6	-2,0	c

Gj.snitt ubesvart: 3,9 %

Gj.snitt: 0,46 55,9 54,3 57,4 -3,1

a) Svak diskriminering D<0,30

b) Positive verdier er kjønnsforskjeller i jenters favør

c) Signifikante forskjeller

Negative verdier er kjønnsforskjeller i gutters favør

Kilder:

- Carr, M., Steiner, H. S., Kyser, B. & Biddlecomb, B. (2007). A comparison of predictors of early emerging gender differences in mathematics competency. *I Learning and Individual Differences*, 18 (2008), 61 – 75.
- Elstad, E. & Turmo, A. (2007). Kjønnforskjeller i motivasjon, læringsstrategibruk og selvregulering i naturfag. *I NorDiNa 1*.
- Grønmo, L. S. (2005). Ferdighetenes plass i matematikkundervisningen. *I Namnaren*, 4.
- Grønmo, L. S., Bergem, O. K., Nylehn, J. & Onstad, T (2008). *Fortsatt store utfordringer for norsk skole*. ILS, Universitet i Oslo.
- Grønmo, L. S., Onstad, T., Nilsen, T., Hole, A., Aslaksen, H. & Borge, I. C. (2012). *Framgang, men langt fram. Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2011*. Akademika forlag
- Hannula, M.S, Juuti, K. & Ahtee, M. (2007). Gender Issues in Finnish Mathematics and Physics Education. I E. Pehkonen, M. Ahtee & J. Lavonen (Eds.), *How finns learn mathematics and science* (pp. 85-96). Rotterdam: Sense
- Kimball, M. (1995). *Feminist Visions of Gender Similarities and Differences*. Harrington Park Press
- Kjærnsli, M., Lie, S., Olsen, R.V. & Roe, A. (2007). *Tid for tunge løft: norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006*. Oslo: Universitetsforlaget.
- Kjærnsli, M. & Roe, A. (2010). *På rett spor: norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009*. Oslo: Universitetsforlaget.
- Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: a review. *Educational Research*, 41(1):63–76.
- Pitta-Pantazi, D., Gray, EM. & Christou, C (2004). Elementary School Students' Mental Representation of Fractions. *I Mathematics Education*, 4, 41–48.
- Sumpter, L. (2009). *On Aspects of Mathematical Reasoning. Affect and Gender*, Doctoral Thesis No. 41. Department of Mathematics and Mathematical statistics, Umeå University, Sweden
- Utdanningsdirektoratet (2010). Rammeverk for Nasjonale prøver, s. 7 – 8. Lastet ned fra: http://www.udir.no/Upload/Nasjonale_prover/2010/5/Rammeverk_NP_22122010.pdf 19. februar 2012, kl. 10.10.