

Nasjonale prøvar i engelsk 2011 – rapport og analyse

v/Eli Moe, UiB, februar 2012

I rapporten gir vi opplysningar om dei nasjonale prøvane i engelsk 2011, både for prøvane som eit heile, men også for einskilditem. Hausten 2011 vart det kun gjennomført prøvar på 8. trinn.

Analysane er basert på data frå alle elevane som tok prøvane på åttande trinn hausten 2011.

Oppgåveformat

Alle item i engelskprøvene kan skårast automatisk. Nokre av itemformata kan likevel tolkast som "opne" i og med at svaralternativ ikkje er oppgitt. Åpne format er:

- *Click & drag* (eleven plasserer ein ting i eit bilde)
- *Who could say* (eleven klikkar på eit namn i teksten)
- *Click on a word* (eleven klikkar på eit ord i teksten)
- *Grammar* (eleven fullfører ein setning ved å skrive inn eitt eller fleire ord)

Lukka format er item der eleven vel mellom ulike svaralternativ. Lukka format er:

- *Click picture* (eleven klikkar på eitt av fleire bilde)
- *Click text* (eleven klikkar på ein av fleire tekstar)
- *Gap filling* (eleven klikkar på eitt av fire svaralternativ i løpande tekst - lukeprøve)
- *MC – multiple choice* - fleirval (eleven klikkar på eitt av fleire oppgitte svaralternativ)
- *Move paragraph* (eleven svarar ved å flytte avsnitta i ein tekst slik at rekkefølga blir korrekt)

Tabell 1 viser kor mange item av kvart format som finst i prøvane på 8. trinn.

Tabell 1: 8. trinn - fordeling av oppgåvetypar/itemformat i dei ulike prøveversjonane

Itemformat	Prøve 1	Prøve 2	Prøve 3
Click picture	3	3	3
Click text	3	3	3
Click& drag	4	4	4
Gap filling	5	5	5
Multiple choice	10	10	10
Who could say	7	7	7
Click word	5	5	5
Grammar	5	5	5
Move paragraph	4	4	4

Oppsummerande statistikk

Det blir utvikla 3 parallelle nasjonale prøvar i engelsk. Prøvene kan samanliknast med omsyn til innhald (t.d. tal på item, tal på item per format, tema osv), vanskegrad og diskriminering.

Tabell 2: Oppsummerande statistikk Nasjonale prøvar i engelsk 2010

Prøve	Ant. item	Andel fleirval	Reliabilitet	Gj.sn p-verdi	Gj.sn poeng	Std-avvik	Diskriminering	Tal på elevar
Eng 8-1	46	25/46	0.923	0.61	27.86	9.93	0.44	19380
Eng 8-2	46	25/46	0.927	0.61	28.14	10.09	0.45	19113
Eng 8-3	46	25/46	0.928	0.60	27.38	10.01	0.45	19014

Alle prøvar har tilfredstillande og tilnærma like reliabilitetsmål, frå 0,923 til 0,938.

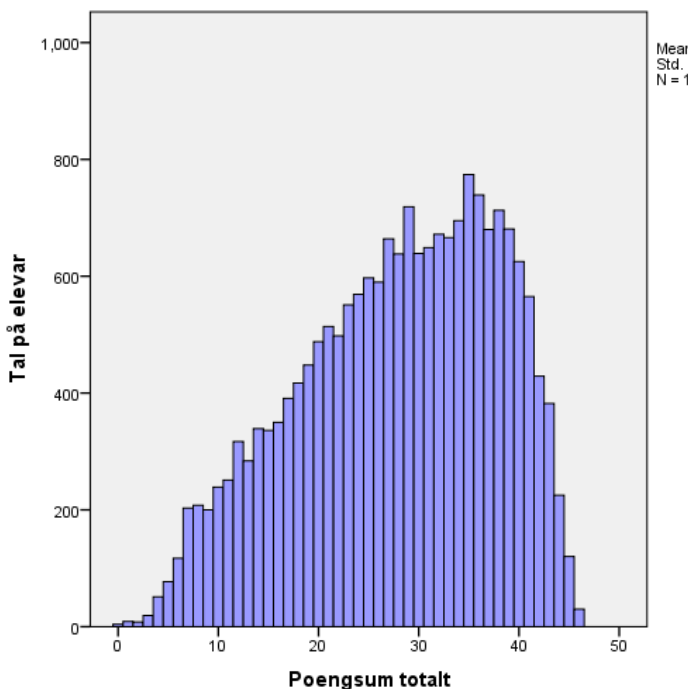
Prøvene sett under eitt, og dei aller fleste einskilditem, diskriminerer mellom sterke og svake elevar på ein fornuftig måte. Det er nokre få item som ikkje fungerer som forventa, dvs. dei fungerer annleis enn då dei vart pilotert.

Engelskprøvene i 2011 har gjennomsnittlege p-verdiar mellom 0.60 og 0.61. Det betyr at prøvene er litt lettare enn det vi siktar mot. Ideelt sett bør desse prøvene ha ein vanskegrad mellom 0.5 og 0.6.

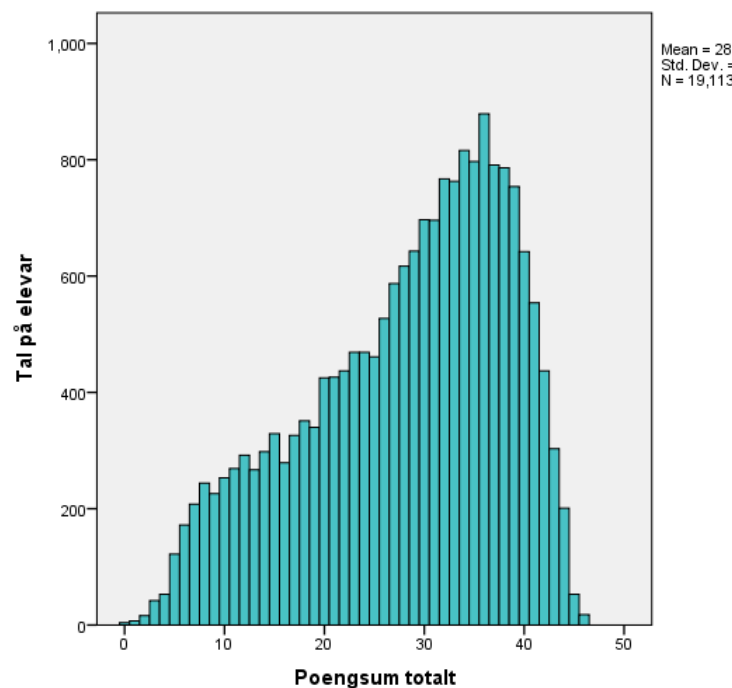
Fordeling av resultat på poengskala

Figur 1 – 3 viser korleis elevane i dei tre 8. trinnss prøvane fordeler seg på poengskalaen.

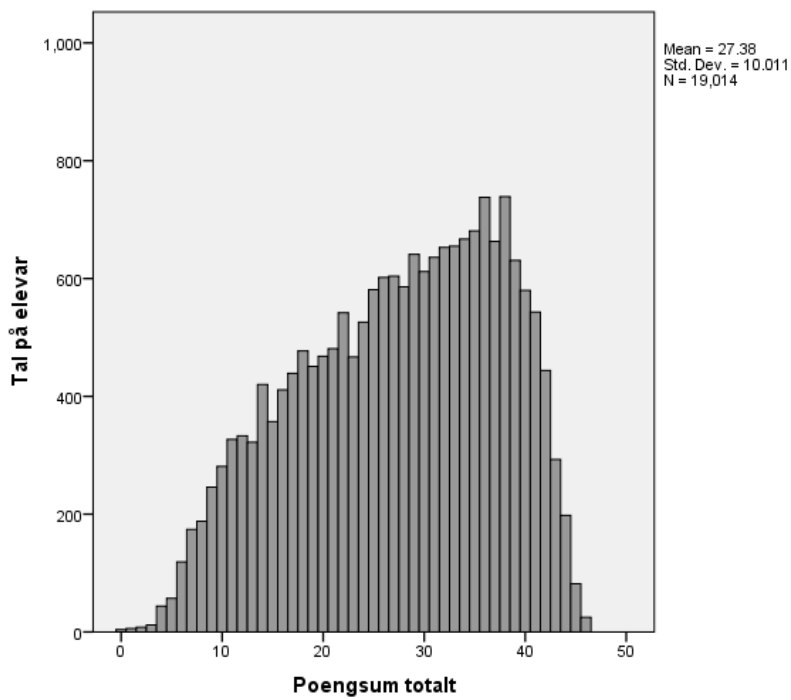
Figur 1: Totalpoengsum og elevar – prøve 8-1



Figur 2: Totalpoengsum og elevar – prøve 8-2



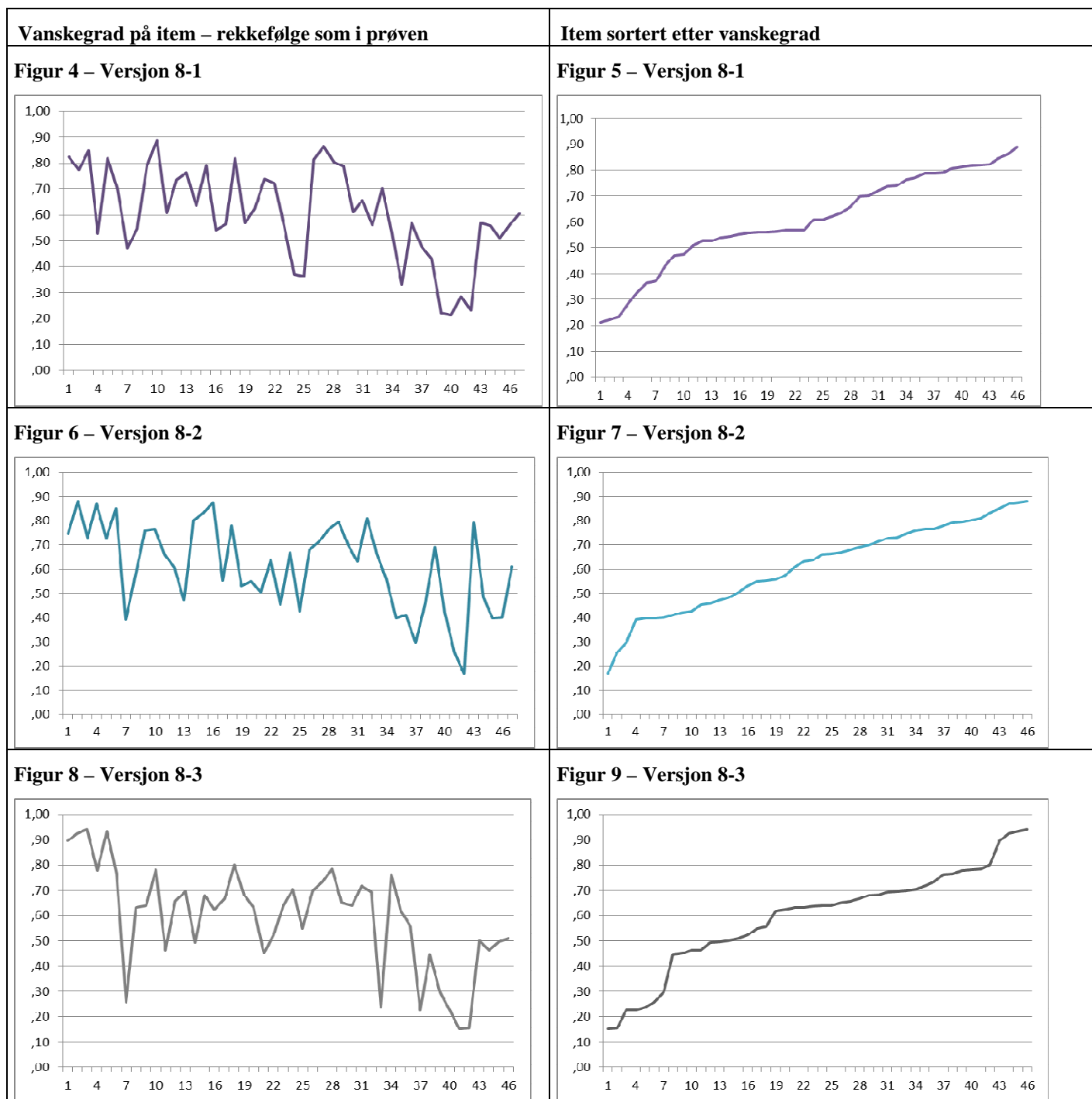
Figur 3: Totalpoengsum og elevar – prøve 8-3



Ingen av prøvane er heilt normalfordelte. Figur 1 – 3 viser at fordelinga av resultat heller litt mot høgre, det vil seie at litt fleire enn halvparten av elevane skårar i den høgare delen av skalaen.

Vanskegrad på item

Figurane 4 til 9 viser vanskegraden på itema i prøvane på 8. trinn, først fordelt etter rekkefølga itema opptrer i prøven (figur 4, 6, 8) og deretter sortert frå det vanskelegaste til det lettaste (figur 5, 7, 9).



Figur 4, 6 og 8 viser at i prøven er det ei veksling mellom lette og vanskelege oppgåver. Det er ein tendens til at dei fleste vanskelege oppgåvene kjem mot slutten av prøven. I figur 5, 7 og 9 er itema sortert etter vanskegrad. Vi ser at prøvane oppfører seg nokolunde likt. Dei vanskelegaste oppgåvene har ein p-verdi rundt 0.2 eller under, medan dei lettaste oppgåvene har ein p-verdi rundt 0.9 .

Tabell 3 summerer opp minimums-, maksimums- og gjennomsnittlige p-verdiar for alle tre versjonane.

Tabell 3: Minimum-, maksimum- og gjennomsnittlige p-verdiar 5. trinn

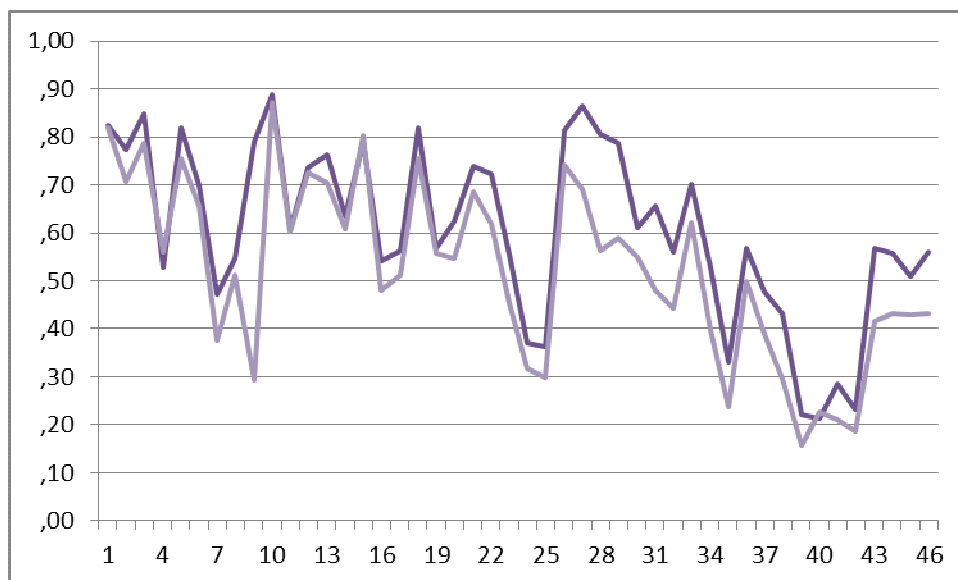
	Versjon 8-1	Versjon 8-2	Versjon 8-3
Minimum p-verdi	0.21	0.17	0.15
Maksimum p-verdi	0.89	0.88	0.94
Gjennomsnittleg p-verdi	0.61	0.61	0.60

Meir detaljerte opplysningar om format, kjønnskilnader og vanskegrad på itemnivå finst i tabell 11 - 13 til slutt i rapporten.

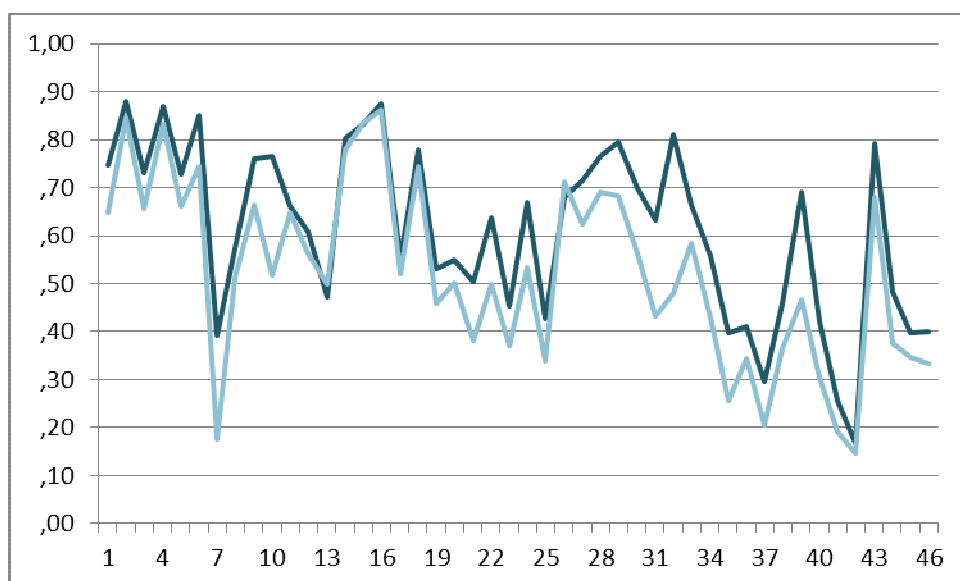
Vanskegrad - reelle prøvar vs pilotering

Figurane 10 til 12 viser vanskegraden på itema i piloteringa hausten 2010 og i dei reelle prøvane i 2011. X-aksen (den horisontale aksen) viser til itema i prøvane i den rekkefølga dei vart presenterte for elevane. Y-aksen (den vertikale aksen) viser til p-verdien for kvart item. Den mørkaste lina i dei tre figurane viser til vanskegrad (p-verdi) på dei nasjonale prøvane, medan den lysaste lina viser til vanskegrad ved pilotering. Jo nærare 0 p-verdien er, jo vanskelegare er itemet, og jo nærare 1 p-verdien er, jo lettare er itemet. Vi ser at itema i den reelle prøver er «lettare» enn i piloteringa.

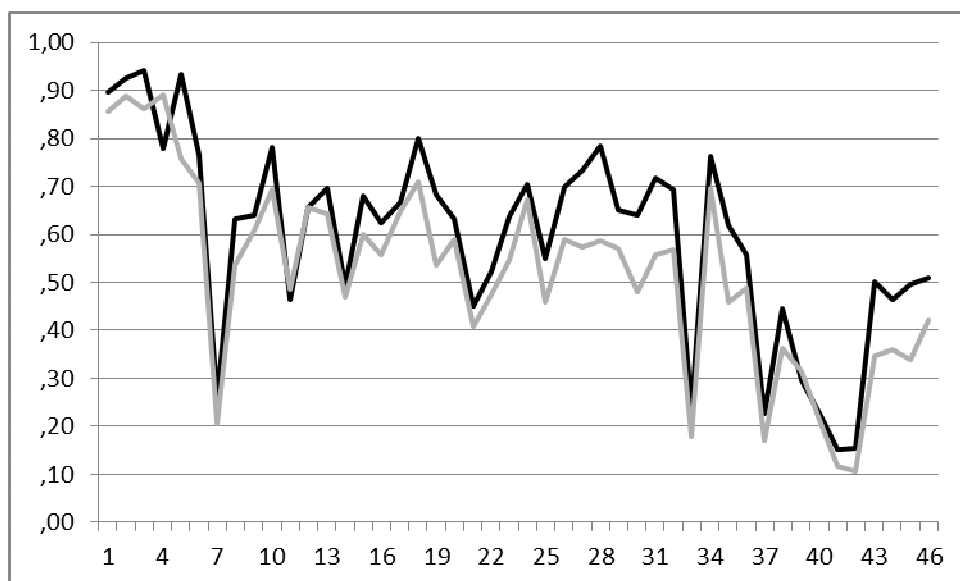
Figur 10: Vanskegrad nasjonale prøvar 2011 og pilotering 2010, versjon 8.1



Figur 11: Vanskegrad nasjonale prøvar 2011 og pilotering 2010, versjon 8.2



Figur 12: Vanskegrad nasjonale prøvar 2011 og pilotering 2010, versjon 8.3



Vi ser at fleire elevar svarar rett på itema i den reelle prøven enn i piloteringa. Dei same itema er altså lettare på dei reelle prøvane enn dei var eit år tidlegare i piloteringa. Dette er ein tendens vi ser kvart år. Elevane ser på dei reelle prøvane som meir viktige enn piloteringsprøvene. For item 26 til og med 34 (i alle tre versjonane) er skilnaden spesielt stor. Dette er item av typen *Who could say*. Elevane blir i dette formatet presentert for ein litt lengre tekst. For å svare rett på spørsmåla, må dei lese teksten nøye. Figurane viser ein relativt stor skilnad mellom reelle prøvar og pilotering. Elevane brukar sannsynlegvis meir tid og les teksten meir nøye på dei reelle prøvane, og dermed svarar ein større del av elevane rett på desse spørsmåla. Den same tendensen ser vi kvart år.

Kjønnskilnader

Ordet «signifikant» har to ulike tydingar. I vanleg språkbruk kan ordet bytast ut med adjektiva «stor», «betydeleg», «markert» osv. I statistisk samanheng betyr derimot ordet signifikant «ikkje tilfeldig». Ein t-test vil avsløre om skilnadene mellom gutar og jenter er signifikante. At ein skilnad er signifikant, vil seie at det er liten sjanse for at resultatet er tilfeldig. Dette blir uttrykt som ein p-verdi som kan tolkast som kor sannsynleg det er at resultatet kan oppstå ved eit reint tilfelle. Ein liten verdi er difor eit uttrykk for at skilnadene er store (t.d >0.01), dette kan også tolkast som ein prosent der ein multipliserer p-verdien med 100. Tradisjonelt brukar ein grenser på 5% (0.05) og 1% (0.01) som viktige markeringar.

Når vi har analysert data frå nasjonale prøvar i engelsk tidlegare år, har det i det store og heile ikkje vore særlege skilnader mellom gutane og jentene sine resultat. Jamnt over har jentene gjort det litt betre enn gutane utan at desse skilnadene har vore signifikante. Dersom ein baserer analysane sine på store utval (til dømes fleire tusen elevar), er det lett at trivielle skilnader blir signifikante. Desse analysane baserer seg på data frå omlag 19000 elevar for kvar av dei tre prøvane. Her vil vi meir eller mindre forvente at trivielle skilnader blir signifikante om vi samanliknar korleis gutar og jenter presterer.

8. trinn

Den gjennomsnittlege poengsummen viser at det er små skilnader mellom gutane og jentene. Jentene skårar gjennomsnittleg litt høgre enn gutane. Dette gjeld alle tre prøveversjonar. T-testar viser at kjønsskilnadene er signifikante.

Den gjennomsnittlege poengsummen og standardavviket viser at det er relativt små skilnader mellom gutane og jentene på 8. trinn. Dette gjeld alle prøveversjonar. Sjølv om skilnadene er små, viser altså t-testar at dei er signifikante for to prøvane, prøve 8-1 og 8-2. Resultatet for prøve 8-3 viser at skilnaden mellom gutar og jenter ikkje er signifikant. Sjølv om jentene i gjennomsnitt skårar litt høgre enn gutane, kan dette vere tilfeldig. Vi ser også at standardavviket i gutegruppene er litt større enn i jentegruppene. Det betyr at variasjonen i resultat er litt større blant gutane enn blant jentene.

Tabell 4: skilnader mellom jenter og gutar på 8. trinn – t-test

Prøve	Kjønn	Gjennomsnittleg poengsum	n	Standardavvik	t	Sig
8-1	Gutar	27.31	9820	10.21	7.839	0.000**
	Jenter	28.42	9560	9.60		
8-2	Gutar	27.74	9633	10.45	5.517	0.000**
	Jenter	28.54	9480	9.71		
8-3	Gutar	27.30	9618	10.28	1.142	0.254*
	Jenter	27.46	9396	9.73		

** skilnaden er signifikant på 0.01-nivå

* skilnaden er ikkje signifikant

Dersom vi deler jentene og gutane inn i ti grupper (skåringsgrupper) på grunnlag av oppnådd poengsum, kan vi samanlikne om jente- og gutegruppene oppfører seg likt. Skåringsgruppe 1 viser til gjennomsnittleg poengsum for dei 10% svakaste i gruppa, medan skåringsgruppe 10

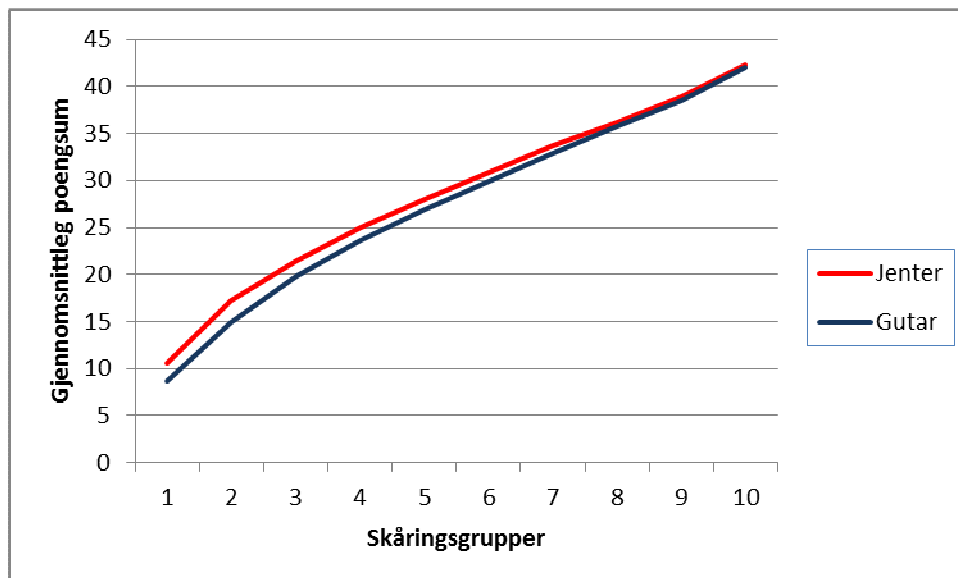
viser til resultatet for dei 10% av elevane med høgast totalpoengsum. Tabell 5 gir ei oversikt over kjønn og gjennomsnittleg poengsum for skåringsgruppene.

Tabell 5: 8. trinnsprøvar – gjennomsnittleg poengsum i skåringsgrupper

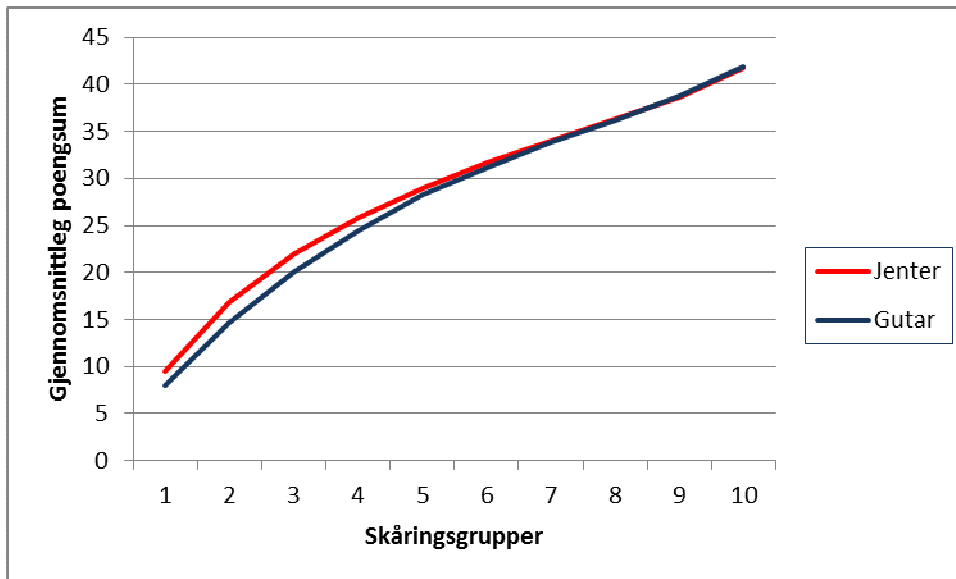
Skårings- grupper	Prøve 8-1		Prøve 8-2		Prøve 8-3	
	Jenter (n=9560)	Gutar (n=9820)	Jenter (n=9480)	Gutar (n=9633)	Jenter (n=9396)	Gutar (n=9618)
1	10.53	8.60	9.49	7.94	10.14	8.82
2	17.23	14.99	16.87	14.71	15.89	16.68
3	21.50	19.79	22.03	20.10	19.92	19.35
4	24.96	23.60	25.85	24.49	23.66	23.34
5	28.04	26.89	28.97	28.28	26.83	26.86
6	30.84	29.92	31.63	31.18	29.76	30.21
7	33.68	32.94	34.00	33.86	32.72	33.11
8	36.20	35.74	36.24	36.18	35.54	35.96
9	38.94	38.54	38.63	38.71	38.30	38.63
10	42.33	42.06	41.73	41.95	41.86	41.99

Opplysningane som kjem fram i tabell 5, er vist grafisk i figur 13 – 15. Desse figurane viser gjennomsnittleg poengsum for kvar gruppe på dei tre prøvane på 8. trinn. Den raude lina representerer jentegruppene og den blå gutegruppene.

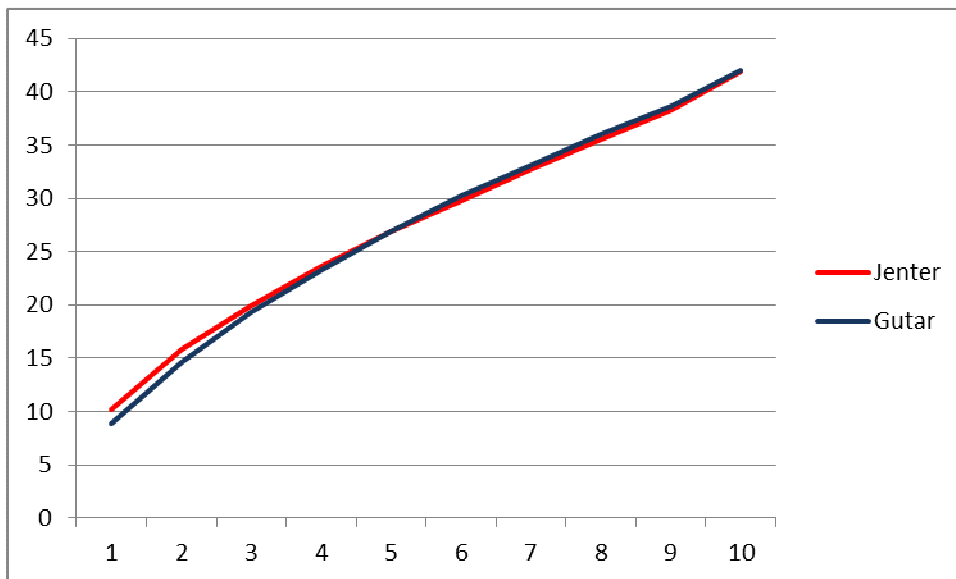
Figur 13: Prøve 8-1 – Gjennomsnittleg poengsum for ti skåringsgrupper – jenter og gutar



Figur 14: Prøve 8-2 – Gjennomsnittleg poengsum for ti skåringsgrupper – jenter og gutar



Figur 15: Prøve 8-3 – Gjennomsnittleg poengsum for ti skåringsgrupper – jenter og gutar



Figurane viser at skilnadene mellom gutar og jenter er små på 8. trinn. Tabell 5 viser det same. Dei svakaste jentene skårar høgre enn dei svakaste gutane. Samtidig så skårar dei flinkaste jentene og gutane tilnærma likt.

Tabell 6 - 8 viser korleis gutar og jenter fordeler seg på poengskalaen i dei tre prøvane. Desse tabellane viser om lag det same som figurane 13 – 15: Dei svakaste gutane er svakare enn dei svakaste jentene, medan dei sterkaste gutane og jentene skårar tilnærma likt.

Tabell 6: Fordeling av jenter og gutar i poenggrupper Prøve 8-1

Poeng	Jenter (n=9560)		Gutar (n=9820)	
	Antall	Prosent	Antall	Prosent
0 - 9	323	3.38	573	5.84
10 - 19	1572	16.44	1800	18.33
20 - 29	2926	30.61	2902	29.55
30 - 39	3516	36.78	3392	34.54
40 - 46	1223	12.79	1153	11.74

Tabell 7: Fordeling av jenter og gutar i poenggrupper Prøve 8-2

Poeng	Jenter (n=9480)		Gutar (n=9633)	
	Antall	Prosent	Antall	Prosent
0 - 9	435	4.59	659	6.84
10 - 19	1396	14.73	1608	16.69
20 - 29	2601	27.44	2460	25.54
30 - 39	3980	41.98	3766	39.09
40 - 46	1068	11.27	1140	11.83

Tabell 8: Fordeling av jenter og gutar i poenggrupper Prøve 8-3

Poeng	Jenter (n=9396)		Gutar (n=9618)	
	Antall	Prosent	Antall	Prosent
0 - 9	329	3.50	529	5.50
10 - 19	1908	20.31	1910	19.86
20 - 29	2854	30.37	2644	27.49
30 - 39	3299	35.11	3376	35.10
40 - 46	1006	10.71	1159	12.05

Av dei elevane som har fått 46 poeng (av 46 mogelege) er kjønnsfordelinga slik:

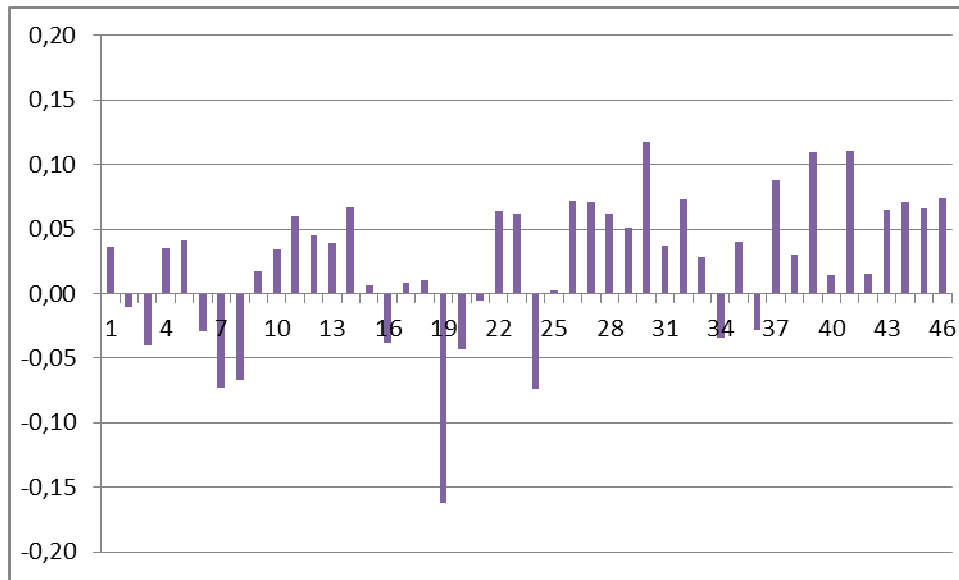
Tabell 9: Jenter og gutar med maksimal poengsum

Prøve	Jenter	Gutar	Elevar med 46 poeng totalt
8-1	15	15	30
8-2	10	8	18
8-3	13	12	25

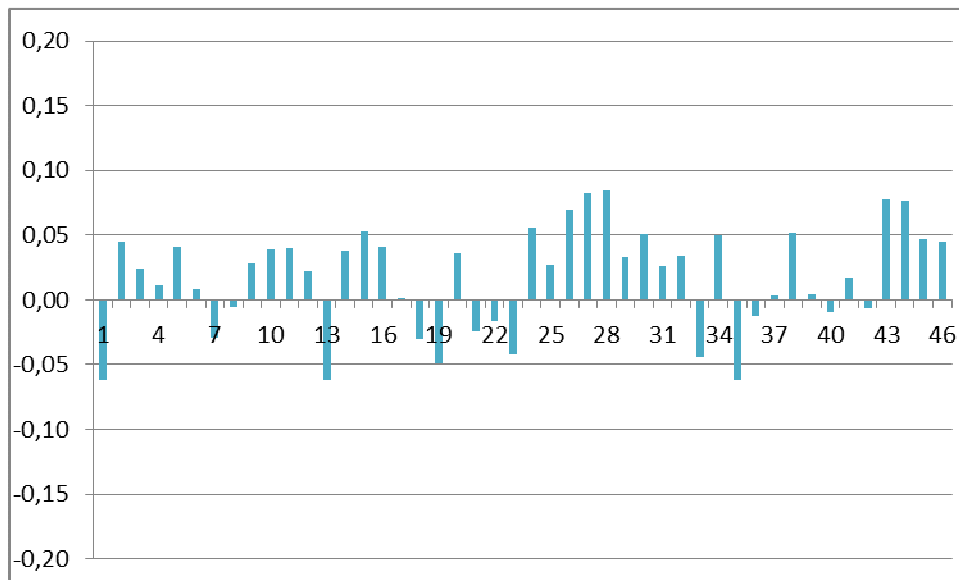
Tabell 9 viser at det er relativt få elevar som oppnår maksimal poengsum. Av dei 57507 som tok ein av dei nasjonale prøvane i engelsk for 8. trinn, har kun 73 personar skåra maksimalt, ca 0.13 prosent. Det kan også sjå ut som kjønnsfordelinga når det gjeld maksimal skåre, er tilnærma lik. Om lag like mange jenter og gutar skårar maksimalt.

Figur 16 – 18 viser differansen i p-verdi mellom gutar og jenter på itemnivå. Kolonnene som peikar oppover med positive verdiar viser til item der jentene gjer det best. Kolonnene som peikar nedover, med negative verdiar, viser til item der gutane gjer det best. Di høgre kolonnene er, di større er skilnaden mellom jenter og gutar. Vi ser at det er fleire «jentekolonner» enn «gutekolonner».

Figur 16 : Skilnad i p-verdi mellom gutar og jenter på itemnivå – Prøve 8-1

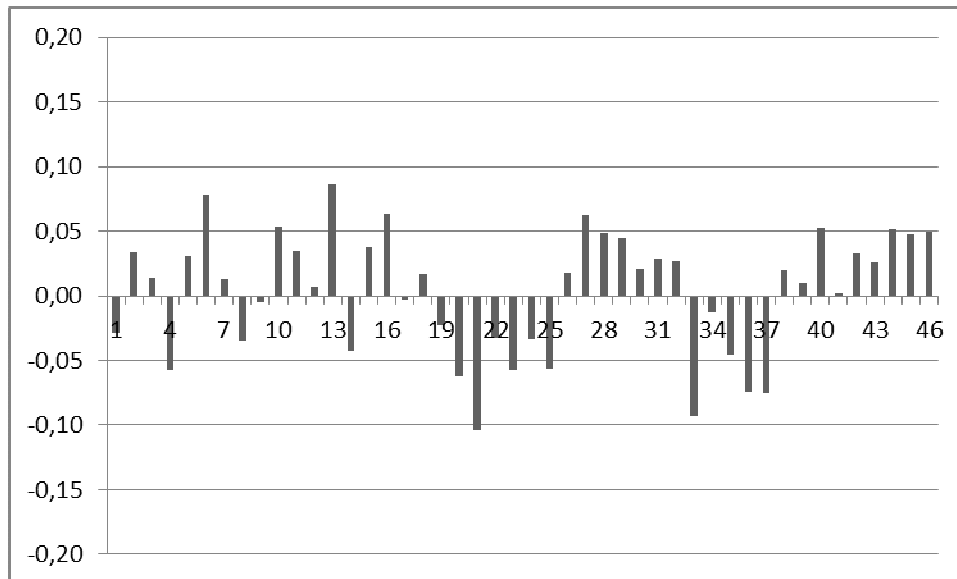


Figur 17: Skilnad i p-verdi mellom gutar og jenter på itemnivå – Prøve 8-2



Figur 18 viser til prøve 8-3. I denne prøven er det fleire item der gutane gjer det best enn i dei andre to prøvane.

Figur 18: Skilnad i p-verdi mellom gutar og jenter på itemnivå – Prøve 8-3



Vi ser at jentene i gjennomsnitt svarar rett på fleire item enn gutane (kolonnene med positiv verdi som peikar oppover). Dette gjeld i dei fleste itemformat. I fire av formatata gjer jentene så å seie reint bord: *Who could say*, *Gap*, *Grammar* og *Move paragraph*. I formatata *Who could say* og *Move paragraph* blir elevane presenterte for lengre tekstar, og dei treng å lese tekstane nøye for å kunne svare rett. Formatata *Gap* og *Grammar* krev kunnskap i grammatikk og ei kjensle for engelsk språkbruk.

I *Who could say*-item blir elevane presentert for ein litt lengre tekst, som beskriv ein del menneske. I tillegg får dei eit spørsmål av typen: "Who could say "I like to read books"?" For å svare rett på desse itema må elevane bruke tid og nærlese teksten ganske nøye. Det er til saman 21 *Who could say*-item i dei tre prøvane. Jentene gjer det best på alle desse itema.

Gap-itema presenterer elevane for ein tekst med fleire «luker». Kvar luke har ein nedtrekksmeny, og eleven skal velje eitt av fire alternative svar. Dette formatet testar vokabular og grammatikk. Av dei 15 itema av dette formatet som finst i prøvane, gjer jentene det best på 13 av dei.

I formatet *Grammar* må elevane følgje instruksjonar ved å skrive inn ord og fullføre ein setning. Til dømes:

Complete the second sentence so that the meaning is similar to that of the first sentence. Use 2-4 words in each gap.

- My mother would not let me go to the party in my shorts.
- My mother stopped _____ to the party in my shorts.

(I nokre tilfelle er fleire korrekte svar lagt inn i fasiten. Elevane blir ikkje straffa for feil i teiknsetjing.) For å svare rett på spørsmåla i dette formatet må elevane ha kunnskapar i /god kjensle for engelsk språkbruk. Jentene gjer det best på 13 av 15 grammatikkoppgåver.

Elevane må både nærlese og ha oversikt over heile teksten for å svare korrekt på itema i formatet *Move paragraph*. Oppgåva er å plassere avsnitta i ein tekst i korrekt rekkefølge ved å klikke på piltastar. Jentene taklar alle itema i dette formatet betre enn gutane, 12 item til saman i dei tre prøvane.

Det er ein tendens til at gutane svarar rett på fleire item enn jentene i formatet *Click word*. Av 15 item i dei tre prøvane, gjer gutane det best på 10 av desse. Jentene taklar dei andre 5 best.

I dei andre formata vekslar jentene og gutane på å gjera det best på einskilditema.

I formata *Who could say* og *Move paragraph* må elevane forhalde seg til litt lengre tekstar. Her gjer altså jentene det best. I formatet *Click word* derimot er utgangspunktet ein kort tekst. For å svare rett, må eleven forstå heile teksten og tydinga av ord i ein samanheng.

Vi gjentar til slutt at skilnadene mellom resultatet gutar og jenter oppnår på dei nasjonale prøvane i engelsk er små, trass i at skilnadene på to av prøvane er signifikante.

Sjå elles tabell 11 – 13 (heilt til slutt) for detaljerte skilnader på itemnivå.

Opplysningar på itemnivå

I tabell 11 – 13 presenterer vi opplysningar om einskilditem i dei tre prøvane for 8. trinn. For å vise korleis tabell 11 – 13 skal lesast, har vi tatt med tabell 10, som inneheld to item, som eit eksempel. Item 6 og 7 i denne tabellen er henta frå prøve 8-2 (tabell 12).

Tabell 10: Eksempel på resultattabell

Felt		1					2					3	4				Flagg/ Komm
Fleirval		A	B	C	D	Blank		A	B	C	D	Blank	D	p-verdi			
Rett/galt		0	1					0	1					Kjønsskilnader			
		Tal på elevar					Gjennomsnittleg skåre i svargrupper						alle	jenter	gutar	diff	
6	Click text	869	16252	1360	596	36		14.57	30.07	18.64	17.71	13.89	.428	.85	.85	.85	0.01
7	Click&drag	11449	7470			194		25.06	32.95			24.22	.340	.39	.38	.41	-0.03

Felt 1

Eit item er enten eit fleirvalsitem der eleven svarar ved å velja mellom ulike svaralternativ (A,B C osv.) eller eit item som blir skåra rett eller galt (t.d. *click & drag*, *colour*, *click word* osv.). Felt 1 viser korleis elevane i utvalet har svart på itemet. Vi ser at item 7 er eit rett/galt item, og at 11449 elevar i utvalet har avgitt galt svar, medan 7470 har svart rett. 194 elevar har ikkje svart på itemet. Item 6 er eit fleirvalsitem, og tabellen viser kor mange elevar som har valt dei ulike svaralternativa. 869 elevar har valt alternativ A, 16252 har valt alternativ B osv. 36 elevar har ikkje svart på item 6.

Felt 2

Dette feltet viser gjennomsnittspoengsum for ulike elevgrupper. Vi vil forvente at gjennomsnittspoengsummen er høgare for elevgruppa som har svart rett på eit item, enn gruppa som har svart feil på same item. Om vi ser på item 7, så ser vi at elevane som har svart feil på dette itemet, har ein gjennomsnittspoengsum på 25.06 poeng, mens elevane som svarar rett på itemet har ein gjennomsnittspoengsum på 32.95. Det er 194 elevar som ikkje har svart på dette itemet, og gjennomsnittleg poengsum i denne gruppa er 24.22. Når det gjeld item 6, så er gjennomsnittspoengsummen 14.57 for dei elevane som har svart alternativ A, for gruppa som har svart alternativ B er gjennomsnittet 30.07 (rett svar) osv. Gjennomsnittspoengsum for dei som har svart feil på dette itemet er 14.89.

Felt 3

D står for diskriminering, og dette er eit mål på kor godt itema diskriminerer mellom elevane. Vi forventar at fleire sterke enn svake elevar svarar rett på eit item. Diskrimineringsindeksen varierer mellom 0 og 1. Di nærare 1 verdien er, di betre diskriminerer itemet mellom svake og sterke elevar. Normalt vil ein seie at eit item bør ha ein diskrimineringsindeks over 0,25 (maxverdi =1).

Felt 4

Felt 4 presenterer vanskegrad i form av p-verdi for 1) alle elevane som har tatt prøven, 2) jentene og 3) gutane. I tillegg er det ei kolonne som viser skilnaden i p-verdi mellom jenter og gutar. P er eit mål på vanskegrad, og p varierer mellom 0 og 1. Di nærare 0 p-verdien er, di vanskelegare er itemet. Lette item har verdiar som nærmar seg 1. Item 6 har ein p-verdi på 0,85. Det betyr at 85% av elevane har svart rett på dette itemet. Og vi ser at dette itemet har vore like vanskeleg/lett for jentene og gutane. Vi ser at item 7 er vanskelegare. 39% av elevane har svart rett på itemet, 38% av jentene og 41% av gutane. Dette betyr at det er relativt noko fleire gutar enn jenter som har svart rett. Den siste kolonna som viser til skilnader mellom gutar og jenter, viser skilnad i vanskegrad mellom jentegruppa og gutegruppa. Positive verdiar viser til item der jentene gjer det best, medan negative verdiar viser til item der gutane gjer det best.

Det er vanskeleg å angi ideelle p-verdiar fordi prøvar har ulike formål. Dersom målet er å skilja mellom "dei fleste" og dei svakaste elevane, vil ein velje item med relativt høg p-verdi, det vil seia item dei fleste svarar rett på og dei svakaste ikkje klarar. Dersom målet er å skilja mellom personar i heile populasjonen (slik som dei nasjonale prøvane), både mellom dei flinkaste og dei svakaste, vil ein velje både lette og vanskelege item. Mange vil då ha som mål at gjennomsnitts p-verdi for heile prøven er rundt 0,5.

Vi reknar med at denne forklaringa gjer det lettare å lese dei tre neste tabellane.

Prøve 8-1

n = 19380 – gutar = 9820, jenter = 9560

Tabell 11: Informasjon om einskilditem – Prøve 8-1

Felt		1					2					3	4				Flagg/ Komm
		A	B	C	D	Blank	A	B	C	D	Blank	D	p-verdi				
Rett/galt		0	1				0	1					Kjønnskilnader				
		Tal på elevar					Gjennomsnittleg skåre i svargrupper						alle	jenter	gutur	diff	
1	Click pict.	272	15948	2923	205	32	20.38	30.15	16.82	18.14	21.97	.467	.82	.84	.81	0.04	
2	Click pict.	453	14979	1903	2010	35	19.80	30.28	20.05	19.29	14.11	.415	.77	.77	.78	-0.01	
3	Click pict.	1639	16422	239	1048	32	17.52	29.74	15.70	17.64	14.69	.418	.85	.83	.87	-0.04	
4	Click text	5127	2295	1704	10206	48	26.13	22.44	19.20	31.44	17.00	.336	.53	.54	.51	0.04	
5	Click text	1355	937	15881	1136	71	19.96	16.07	30.03	17.55	12.92	.435	.82	.84	.80	0.04	
6	Click text	13548	2078	2401	1223	130	31.60	18.74	20.20	18.35	15.03	.541	.70	.68	.71	-0.03	
7	Click&drag	10058	9130			192	24.10	32.07			24.16	.357	.47	.43	.51	-0.07	
8	Click&drag	8610	10576			194	23.74	31.29			23.47	.335	.55	.51	.58	-0.07	
9	Click&drag	3869	15324			187	21.19	29.59			23.86	.302	.79	.80	.78	0.02	
10	Click&drag	1953	17236			191	16.90	29.15			23.70	.340	.89	.91	.87	0.03	
11a	Gap	3445	2843	1236	11790	66	23.31	24.72	20.30	30.82	12.33	.329	.61	.64	.58	0.06	
11b	Gap	3952	696	14287	373	72	20.62	15.67	30.93	12.61	11.99	.485	.74	.76	.71	0.05	
11c	Gap	1482	14785	2786	198	129	16.51	29.82	25.04	14.72	14.19	.316	.76	.78	.74	0.04	
11d	Gap	2256	2079	2648	12311	86	19.51	19.84	25.03	31.46	12.03	.440	.64	.67	.60	0.07	
11e	Gap	15286	2430	1424	151	89	29.72	20.37	22.72	17.75	11.53	.326	.79	.79	.79	0.01	
12	Mc	5776	1791	1271	10480	62	24.35	18.31	19.21	32.55	14.03	.475	.54	.52	.56	-0.04	
13	Mc	5329	10918	628	2450	55	22.02	32.12	16.07	24.97	11.95	.448	.56	.57	.56	0.01	
14	Mc	1710	1128	15847	644	51	16.79	15.24	30.51	15.34	13.16	.538	.82	.82	.81	0.01	
15	Mc	11005	2835	1518	3956	66	31.67	22.02	21.41	24.16	13.50	.398	.57	.49	.65	-0.16	
16	Mc	2815	12076	3099	1258	132	22.83	31.21	23.53	19.01	14.27	.393	.62	.60	.64	-0.04	
17	Mc	2235	1495	14339	1241	70	18.19	18.73	31.06	20.18	12.49	.511	.74	.74	.74	-0.01	
18	Mc	1167	2198	1885	13991	139	19.62	21.34	19.51	30.84	13.32	.448	.72	.75	.69	0.06	
19	Mc	10729	2620	2459	3387	185	32.19	23.82	22.92	21.57	14.51	.447	.55	.58	.52	0.06	
20	Mc	4664	3870	3395	7204	247	29.63	23.72	23.75	31.25	16.67	.217	.37	.33	.41	-0.07	Diskr. dårleg
21	Mc	6082	4077	7031	1860	330	29.04	23.89	31.32	21.30	18.32	.217	.36	.36	.36	0.00	Diskr. dårleg

22	Who	3231	15756			393		16.53	30.56			12.75	.539	.81	.85	.78	0.07	
23	Who	2250	16736			394		14.31	30.05			12.03	.531	.86	.90	.83	0.07	
24	Who	3378	15610			392		16.53	30.71			11.78	.558	.81	.84	.78	0.06	
25	Who	3672	15266			442		16.90	30.94			12.33	.571	.79	.81	.76	0.05	
26	Who	7122	11820			438		22.78	31.51			11.93	.420	.61	.67	.55	0.12	
27	Who	6064	12723			593		20.76	31.86			14.60	.523	.66	.68	.64	0.04	
28	Who	7977	10868			535		22.85	32.24			13.61	.459	.56	.60	.53	0.07	
29	Click word	5371	13604			405		19.63	31.58			12.06	.543	.70	.72	.69	0.03	
30	Click word	8362	10218			800		21.40	34.01			16.80	.624	.53	.51	.54	-0.03	
31	Click word	11948	6397			1035		24,65	35,25			19.21	.487	.33	.35	.31	0.04	
32	Click word	7754	11017			609		21.79	32.86			14.69	.543	.57	.55	.58	-0.03	
33	Click word	8860	9231			1289		23.38	33.35			19.31	.490	.48	.52	.43	0.09	
34	Grammar	10716	8354			310		23,45	34.00			14.74	.502	.43	.45	.42	0.03	
35	Grammar	14554	4290			536		26.06	35.34			16.94	.366	.22	.28	.17	0.11	
36	Grammar	14614	4121			644		25.83	37.04			15.19	.448	.21	.22	.21	0.01	
37	Grammar	13255	5531			594		25.44	35.09			14.44	.423	.29	.34	.23	0.11	
38	Grammar	13377	4486			1517		26.18	35.86			18.99	.407	.23	.24	.22	0.02	
39a	Move par.	7379	11030			971		22.62	32.20			18.36	.464	.57	.60	.54	0.07	
39b	Move par.	7610	10799			971		22.14	32.74			18.36	.515	.56	.59	.52	0.07	
39c	Move par.	8564	9845			971		22.83	33.17			18.36	.507	.51	.54	.48	0.07	
39d	Move par.	7540	10869			971		22.56	32.38			18.36	.477	.56	.60	.52	0.07	

Prøve 8-2

n = 19113 – jenter = 9480, gutar = 9630

Tabell 12: Informasjon om einskilditem – Prøve 8-2

Felt		1					2					3	4				Flagg/ Komm
Fleirval		A	B	C	D	Blank	A	B	C	D	Blank	D	p-verdi				
Rett/galt		0	1				0	1					Kjønnskilnader				
		Tal på elevar					Gjennomsnittleg skåre i svargrupper						alle	jenter	gutur	diff	
1	Click pict.	14267	2357	1398	1060	31	30.90	19.92	20.32	19.80	20.55	.435	.75	.72	.78	-0.06	
2	Click pict.	187	1461	16821	616	28	14.53	14.68	29.96	14.65	20.96	.465	.88	.90	.86	0.04	
3	Click pict.	2606	1089	13962	1406	50	22.34	20.00	30.60	20.98	19.68	.365	.73	.74	.72	0.02	
4	Click text	456	725	1298	16605	29	12.73	13.44	19.46	29.89	19.79	.421	.87	.87	.86	0.01	
5	Click text	467	1824	2868	13905	49	13.57	19.45	20.77	31.32	18.94	.481	.73	.75	.71	0.04	
6	Click text	869	16252	1360	596	36	14.57	30.07	18.64	17.71	13.89	.428	.85	.85	.85	0.01	
7	Click&drag	11449	7470			194	25.06	32.95			24.22	.340	.39	.38	.41	-0.03	
8	Click&drag	7916	10978			219	23.28	31.74			23.16	.372	.57	.57	.58	-0.01	
9	Click&drag	4404	14512			197	18.82	31.02			23.69	.476	.76	.77	.75	0.03	
10	Click&drag	4266	14618			229	17.53	31.35			22.03	.541	.76	.78	.75	0.04	
11a	Gap	2627	12620	2204	1594	68	17.16	32.61	19.39	22.56	13.46	.587	.66	.68	.64	0.04	
11b	Gap	11632	2883	887	3629	82	32.15	23.75	16.74	21.86	14.43	.458	.61	.62	.60	0.02	
11c	Gap	2136	4546	9017	3319	95	19.78	26.15	33.01	23.40	14.07	.416	.47	.44	.50	-0.06	
11d	Gap	1319	915	1394	15369	116	15.80	21.27	17.61	30.64	17.18	.473	.80	.82	.79	0.04	
11e	Gap	1545	15885	684	890	109	15.72	30.50	14.74	19.42	14.93	.492	.83	.86	.81	0.05	
12	Mc	671	16701	840	848	53	12.58	30.12	12.66	17.69	12.30	.493	.87	.89	.85	0.04	
13	Mc	10571	3675	2180	2631	56	31.88	26.58	17.31	24.58	11.89	.371	.55	.55	.55	0.00	
14	Mc	2002	1469	14909	679	54	17.37	18.56	31.19	14.25	12.56	.540	.78	.76	.80	-0.03	
15	Mc	4871	1293	10166	2722	61	25.48	15.30	32.90	21.56	13.33	.464	.53	.51	.56	-0.05	
16	Mc	796	4661	10509	3081	66	14.69	23.20	32.54	24.35	14.38	.443	.55	.57	.53	0.04	
17	Mc	1911	4195	3302	9654	51	25.62	23.73	21.46	32.92	11.78	.439	.51	.49	.52	-0.02	
18	Mc	3754	12181	2337	732	109	23.50	31.09	22.44	23.01	14.16	.347	.64	.63	.65	-0.02	
19	Mc	1864	2940	5509	8862	138	24.77	19.60	28.30	31.86	15.30	.291	.45	.43	.47	-0.04	
20	Mc	12782	1877	2066	2248	140	32.40	20.43	21.09	17.63	14.99	.569	.67	.70	.64	0.06	
21	Mc	4262	4439	1995	8173	244	25.25	27.48	17.64	32.85	18.45	.361	.43	.44	.41	0.03	

22	Who	5740	12980			393		22.13	31.26			12.62	.413	.68	.71	.64	0.07
23	Who	5021	13683			409		18.78	32.05			12.20	.586	.72	.76	.68	0.08
24	Who	4061	14641			411		18.24	31.34			12.00	.544	.77	.81	.72	0.08
25	Who	3466	15212			435		16.18	31.32			12.00	.598	.80	.81	.78	0.03
26	Who	5302	13360			451		19.22	32.21			12.37	.585	.70	.72	.67	0.05
27	Who	6541	12077			495		21.01	32.62			13.04	.548	.63	.65	.62	0.03
28	Who	3160	15474			479		15.44	31.23			12.05	.607	.81	.83	.79	0.03
29	Click word	6031	12698			384		19.26	32.85			11.63	.630	.66	.64	.69	-0.04
30	Click word	7579	10671			863		20.81	34.23			17.25	.650	.56	.58	.53	0.05
31	Click word	10739	7595			779		23.26	36.21			16.65	.620	.40	.37	.43	-0.06
32	Click word	10327	7848			938		25.16	33.34			17.43	.389	.41	.40	.42	-0.01
33	Click word	12266	5668			1179		25.81	35.00			19.36	.404	.30	.30	.29	0.00
34	Grammar	9770	8747			596		23.74	33.95			14.79	.492	.46	.48	.43	0.05
35	Grammar	5450	13177			486		21.23	31.53			13.50	.466	.69	.69	.69	0.00
36	Grammar	10294	8046			773		25.37	32.94			15.03	.364	.42	.42	.43	-0.01
37	Grammar	13582	4902			629		26.08	35.64			14.03	.401	.26	.26	.25	0.02
38	Grammar	15111	3210			792		27.40	34.92			14.78	.268	.17	.16	.17	-0.01
39a	Move par.	3092	15152			869		20.71	30.25			17.72	.375	.79	.83	.75	0.08
39b	Move par.	9029	9215			869		24.98	32.21			17.72	.346	.48	.52	.44	0.08
39c	Move par.	10660	7584			869		26.42	31.75			17.72	.245	.40	.42	.37	0.05
39d	Move par.	10629	7624			869		26.27	31.93			17.72	.261	.40	.42	.38	0.04

Prøve 8-3

n = 19014 – jenter = 9396, gutar = 9618

Tabell 13: Informasjon om einskilditem – Prøve 8-3

Felt		1					2					3	4				Flagg/ Komm
Fleirval		A	B	C	D	Blank	A	B	C	D	Blank	D	p-verdi				
Rett/galt		0	1				0	1					Kjønnskilnader				
		Tal på elevar					Gjennomsnittleg skåre i svargrupper						alle	jenter	gutur	diff.	
1	Click pict.	801	17048	409	711	45	15.12	28.76	15.61	15.47	17.40	.380	.90	.88	.91	-0.03	
2	Click pict.	68	17617	519	784	26	11.35	28.35	16.39	14.54	14.27	.323	.93	.94	.91	0.03	
3	Click pict.	457	238	375	17909	35	14.90	12.18	13.36	28.22	15.09	.316	.94	.95	.94	0.01	
4	Click text	558	2581	14820	1007	48	17.42	18.26	29.89	19.76	16.90	.439	.78	.75	.81	-0.06	
5	Click text	437	397	17761	380	39	13.12	12.36	28.32	17.03	13.56	.331	.93	.95	.92	0.03	
6	Click text	1285	14558	1717	13.64	90	18.11	29.95	21.18	17.26	15.96	.430	.77	.81	.73	0.08	
7	Click&drag	13824	4912			176	25.32	33.37			22.75	.314	.26	.27	.25	0.01	
8	Click&drag	6783	12004			227	20.52	31.39			20.48	.488	.63	.61	.65	-0.04	
9	Click&drag	6606	12194			214	29.30	31.34			20.44	.493	.64	.64	.64	-0.01	
10	Click&drag	3938	14871			205	18.30	29.87			21.11	.438	.78	.81	.76	0.05	
11a	Gap	6640	8845	1599	1865	65	27.09	29.35	22.46	23.80	11.95	.135	.47	.48	.45	0.03	Diskr. dårleg
11b	Gap	2474	3270	12453	729	88	17.70	21.54	31.18	16.54	12.28	.487	.65	.66	.65	0.01	
11c	Gap	2321	2269	1959	13237	128	19.72	18.53	22.71	30.75	13.45	.474	.70	.74	.65	0.09	
11d	Gap	5721	2602	1174	9362	155	23.30	24.52	20.29	31.72	17.37	.385	.49	.47	.51	-0.04	
11e	Gap	3152	12929	614	2215	104	21.58	30.16	15.28	23.38	13.56	.365	.68	.70	.66	0.04	
12	Mc	6039	11866	897	164	48	20.98	31.70	16.70	13.23	11.96	.521	.62	.66	.59	0.06	
13	Mc	12658	3144	1423	1733	56	31.35	17.40	16.68	25.89	8.68	.526	.67	.66	.67	0.00	
14	Mc	2009	15206	803	841	65	15.10	30.28	15.60	18.20	9.98	.551	.80	.81	.79	0.02	
15	Mc	1649	3243	1083	12972	67	13.98	22.29	21.02	30.97	11.24	.490	.68	.67	.69	-0.02	
16	Mc	3439	12028	1634	1856	57	22.87	31.07	17.18	21.33	9.89	.445	.63	.60	.66	-0.06	
17	Mc	2987	8567	4181	3198	81	21.33	32.63	26.87	20.04	11.07	.435	.45	.40	.50	-0.10	
18	Mc	2832	2151	9924	3979	128	21.75	22.21	32.16	22.70	13.26	.460	.52	.51	.54	-0.03	
19	Mc	1231	5252	12127	318	86	18.29	22.80	30.74	13.78	12.49	.406	.64	.61	.67	-0.06	
20	Mc	2800	13366	1864	862	122	22.93	30.58	17.65	15.13	13.49	.457	.70	.69	.72	-0.03	
21	Mc	2982	10464	5223	254	91	20.10	31.94	23.25	14.86	12.77	.466	.55	.52	.58	-0.06	

22	Who	5260	13295			459		19.67	30.92			13.16	.506	.70	.71	.69	0.02	
23	Who	4598	13961			455		18.99	30.63			12.35	.508	.73	.77	.70	0.06	
24	Who	3624	14934			456		16.81	30.40			12.27	.550	.79	.81	.76	0.05	
25	Who	6123	12382			509		20.43	31.40			13.00	.515	.65	.67	.63	0.04	
26	Who	6262	12147			605		19.85	31.93			13.96	.573	.64	.65	.63	0.02	
27	Who	4825	13655			534		18.06	31.25			12.58	.588	.72	.73	.70	0.03	
28	Who	5257	13188			569		18.99	31.35			12.78	.566	.69	.71	.68	0.03	
29	Click word	13833	4510			671		25.13	35.93			16.29	.442	.24	.19	.28	-0.09	
30	Click word	4048	14491			475		17.20	30.71			12.32	.568	.76	.76	.77	-0.01	
31	Click word	6603	11751			660		18.91	32.88			14.12	.673	.62	.59	.64	-0.05	
32	Click word	7558	10600			856		20.77	33.01			15.95	.601	.56	.52	.59	-0.07	
33	Click word	13051	4316			1647		25.15	36.34			21.55	.452	.23	.19	.26	-0.07	
34	Grammar	10335	8442			237		24.46	31.37			12.23	.312	.44	.45	.43	0.02	
35	Grammar	13047	5622			345		24.67	34.47			14.19	.422	.30	.30	.29	0.01	
36	Grammar	13712	5622			237		25.80	35.04			16.02	.378	.23	.25	.20	0.05	
37	Grammar	15092	2873			1049		26.24	35.88			20.42	.326	.15	.15	.15	0.00	
38	Grammar	15243	2933			838		26.16	36.93			16.05	.377	.15	.17	.14	0.03	
39a	Move par.	8672	9514			828		22.41	32.75			17.72	.500	.50	.51	.49	0.03	
39b	Move par.	9348	8838			828		23.36	32.53			17.72	.440	.46	.49	.44	0.05	
39c	Move par.	8773	9413			828		22.48	32.80			17.72	.499	.50	.52	.47	0.05	
39d	Move par.	8481	9705			828		23.05	31.98			17.72	.429	.51	.54	.49	0.05	