

Högskoleprovet våren och hösten 2009

Provdeltagargruppens sammansättning och
resultat

Christina Stage
Gunilla Ögren



BVM Nr 40, 2009
ISSN 1652-7313

ABSTRACT

Since spring 1977 the Swedish Scholastic Assessment Test (SweSAT) has been used as an instrument for selection to higher education. The test is administered twice a year, once in spring and once in autumn, and it is allowed to take the test several times. The test scores are valid for five years, and for test-takers with more than one test result the best valid result is used for admittance. After each test administration the raw score is transformed to a standardized score on a scale from 0.0 to 2.0, where 2.0 is the best result. The aim of the standardization is to equate the scores between different test occasions, i.e. it should be just as easy or difficult to get a certain standardized score on one test occasion as on another. The aim of this report is to describe the groups of test-takers in spring and in autumn 2009 regarding composition and results. Results are presented for test-takers of different gender, age and background education. Further a description is given of the equating procedure and the outcome of the standardization. Finally some results are presented of the effects of repeated test-taking. The results presented in the report are based on data processing performed by Mats Hamrén, head of the data management of the SweSAT.

SAMMANFATTNING

Högskoleprovet har nu i drygt 30 år fungerat som urvalsinstrument för antagning till universitets- och högskolestudier. Provet genomförs en gång på hösten och en gång på våren och det är tillåtet att göra provet hur många gånger man önskar. Vid mer än ett provresultat är det alltid det bästa resultatet som gäller vid urval till en studieplats. Provresultaten är giltiga i fem år och efter varje provtillfälle överförs råpoängen (antal rätt besvarade uppgifter) till en normerad poäng, på en skala från 0.0 till 2.0, där 2.0 är det högsta resultatet. Denna normering görs för att säkerställa att det är lika lätt eller svårt att erhålla en viss normerad poäng oavsett prov och provgrupp. Avsikten med föreliggande rapport är att beskriva provdeltagargrupperna våren och hösten 2009 med avseende på sammansättning och resultat. Resultaten presenteras för provdeltagare med olika kön, ålder och utbildning. Vidare beskrivs hur normeringen av provresultaten genomförs och utfallet av normeringen. Slutligen presenteras några resultat avseende effekterna av att genomföra provet flera gånger. Resultaten i rapporten grundas på databearbetningar gjorda av Mats Hamrén, högskoleprovets systemchef.

Innehållsförteckning

BAKGRUND	1
SYFTE	2
RESULTAT	2
<i>Provdeltagargruppernas sammansättning</i>	<i>2</i>
<i>Resultat för olika grupper av provdeltagare</i>	<i>5</i>
NORMERING AV HÖGSKOLEPROVET	12
<i>Utfall</i>	<i>13</i>
Totalpopulationen.....	13
Referenspopulation I.....	14
Referenspopulation II.....	14
Normering baserad på IRT.....	16
<i>Normeringsgränser</i>	<i>17</i>
UPPREPAT PROVDELTAGANDE	18
AVSLUTNING	21
REFERENSER	25

BAKGRUND

Högskoleprovet har sedan våren 1977 fungerat som urvalsinstrument för antagning till universitets- och högskoleutbildning. Fram till våren 1991 kunde provet användas enbart av en begränsad grupp sökande nämligen de som tillhörde kvotgrupp fyra, d.v.s. var behöriga för högskolestudier genom att de var minst 25 år och hade minst fyra års arbetslivserfarenhet. År 1991 trädde nya regler för antagning till universitets- och högskolestudier i kraft. Dessa nya regler innebär att sedan höstterminen 1991 sker urvalet till studieplatser för samtliga sökande antingen utifrån betyg eller utifrån resultat på högskoleprovet. Högskoleprovet fick därmed en mycket större betydelse i antagningssammanhang än vad det tidigare haft. Denna ökade betydelse återspeglas även i antalet provdeltagare. Under hela 80-talet genomfördes högskoleprovet av cirka 10 000 personer årligen (två femtedelar på hösten och tre femtedelar på våren). Under 90-talet var antalet provdeltagare årligen 130 000–145 000 (fortfarande två femtedelar på hösten och tre femtedelar på våren). Till och med hösten år 2009 har sammanlagt 67 reguljära provtillfällen anordnats och totalt 2 117 546 prov administrerats. Åren 1992 och 1997 var det störst antal provdeltagare, 145 577 respektive 144 061 personer, därefter har antalet minskat till drygt 69 000 år 2002. Därefter ökade antalet provdeltagare så att år 2005 genomförde ungefär 79 500 personer högskoleprovet. Från en förväntad uppgång år 2006 skedde återigen en minskning av antalet provdeltagare, som även fortsatte 2007. År 2007 genomfördes provet av ungefär 65 000 personer, vilket innebär att antalet provdeltagare aldrig har varit så lågt sedan 1991. En marginell ökning av antalet provdeltagare skedde 2008 medan en rejäl ökning har skett under 2009. Sett över hela året har antalet ökat med ungefär en tredje del. Den största ökningen har dock skett mellan provtillfället hösten 2008 och provtillfället hösten 2009.

Högskoleprovet består sedan våren 1996 av 122 uppgifter fördelade på fem delprov: ORD, som mäter ordkunskap med 40 uppgifter, NOG som mäter logiskt, kvantitativt tänkande med 22 uppgifter, LÄS som mäter svensk läsförståelse med fem texter och 20 uppgifter, DTK som mäter förmågan att tolka diagram, tabeller och kartor med 20 uppgifter, samt ELF som mäter engelsk läsförståelse med 20 uppgifter.

Våren 1996 genomgick högskoleprovet tämligen stora förändringar vilka både gäller provets sammansättning och utprovning av nya provuppgifter. För utförligare beskrivning av provets tidigare sammansättning och förändringar av utprovningsrutinerna se t.ex. Stage & Ögren, 2001; Stage, 1995; Ögren, m.fl., 1995.

SYFTE

Syftet med denna redovisning är att beskriva provdeltagargrupperna våren och hösten 2009 med avseende på sammansättning och resultat samt att jämföra med motsvarande resultat på proven år 2008. Motsvarande beskrivningar av provdeltagarna och deras resultat har gjorts regelbundet även för tidigare prov och efter provet hösten år 2000 gjordes en sammanfattande beskrivning för åren 1977–2000. Denna resultatsammanställning omfattar förutom verksamhetsåret 2009, vissa jämförelser med föregående år och i viss mån även från och med året 2003.

Ett ytterligare syfte med denna redovisning är att följa upp hur könsskillnaderna i resultat har påverkats av de förändringar som har gjorts av högskoleprovet. Resultatskillnaderna mellan kvinnor och män har följts noggrant alltsedan högskoleprovet började användas (se t.ex. Stage, 1985, 1988, 1991, 1996). Den skillnad som har erhållits var fram till provtillfället våren 1996 i genomsnitt 8 råpoäng till fördel för män. Efter förändringen våren 1996 har skillnaderna minskat till i genomsnitt 5-6 råpoäng (se t.ex. Stage & Ögren, 2001).

Slutligen redovisas, för de provdeltagare som genomfört flera prov, resultaten på deras första, andra, tredje och fjärde provtillfälle.

RESULTAT

Provdeltagargruppernas sammansättning

Den 28 mars 2009 genomfördes högskoleprovet av 45 737 personer och den 24 oktober 2009 av 41 422 personer.

För att möjliggöra en jämförelse mellan provdeltagargrupperna vid de två senaste proven och vid föregående års genomförda prov återges i tabell 1 provdeltagarnas procentuella fördelning på kön, olika åldersgrupper och olika utbildningar. I följande redovisning anges provtillfälle på våren som A och på hösten som B.

Tabell 1. Provdeltagarnas procentuella fördelning på kön, ålder och utbildning.

År	03A+B	04A+B	05A+B	06 A+B	07A+B	08 A+B	09A	09B
Kön								
Män	45	44	46	46	46	48	49	48
Kvinnor	55	56	54	54	54	52	51	52
Ålder								
≤ 20 år	41	41	43	46	47	53	51	55
21–24 år	28	30	29	28	28	27	30	27
25–29 år	13	13	13	13	13	11	10	10
30–39 år	14	12	11	10	9	7	6	6
≥ 40 år	4	4	4	3	3	2	3	2
Utbildning								
Grund/real- skola	2	2	2	1	1	1	1	1
Folkhög- skola	1	1	1	1	1	1	1	1
Gymnasium ≤ 2 år	9	5	3	3	2	1	1	1
Gymnasium ≥ 3 år	71	74	76	78	79	81	81	81
Högre utb. ≤ 80 p	11	12	12	11	11	10	9	9
Högre utb. ≥ 80 p	5	5	5	5	5	5	5	6
Ej angivet	1	1	1	1	1	1	2	1
Totalt antal	72 341	77 521	79 498	71 317	65 079	65 288	45 737	41 422

På samma sätt som under 80- och 90-talet är antalet provdeltagare större på våren än på hösten. Andelen kvinnor som genomför provet är större än andelen män och har alltid varit så. Skillnaden har dock minskat de senaste åren.

Under hela 80-talet var andelen yngre än 25 år mindre än 10 procent. Därefter har åldersfördelningen förändrats så att den största andelen provdeltagare återfinns i den yngsta åldersgruppen. Andelen provdeltagare som är 20 år eller yngre är med något undantag större på hösten än på våren. Denna andel har aldrig varit så stor som vid de två senaste höstprovstillfällena, 56 respektive 55 %. Den kan jämföras med början på 2000-talet när motsvarande siffror var ungefär 40 %. Trenden är att provdeltagargruppen blir allt yngre.

Liksom åldersstrukturen har utbildningsbakgrunden förändrats sedan 80-talet. Majoriteten av provdeltagarna har tre- eller fyraårig gymnasieutbild-

ning medan dessförinnan en övervägande del hade lägre utbildning. I tabell 2 redovisas hur män respektive kvinnor är procentuellt fördelade på utbildningsbakgrund (högsta angivna utbildning) vid de fyra senaste provtillfällena.

Tabell 2. Män respektive kvinnor som deltog i högskoleproven 08A-09B, procentuellt fördelade efter högsta angivna utbildning.

Utbildning	08A	08B	09A	09B
	m/k	m/k	m/k	m/k
Grund/realskola	1/1	1/1	1/1	1/1
Folkhögskola	1/1	1/1	1/1	1/1
Gymnasium ≤ 2 år	1/2	1/1	1/1	1/1
Gymnasium ≥ 3 år	81/80	81/81	82/81	81/80
Högre utbildning ≤ 80 p	10/9	10/9	9/9	9/10
Högre utbildning ≥ 80 p	5/5	5/6	5/6	5/6
Ej angivet	1/2	1/1	1/1	1/1

Män har tidigare i större utsträckning än kvinnor haft tre- eller fyraårig gymnasieutbildning medan kvinnor i något större utsträckning har haft tvåårig gymnasieutbildning, något som utjämnats under senare år.

Även när det gäller åldersfördelningen mellan män och kvinnor har en utjämning och till och med en viss förändring skett. Tidigare var mönstret sådant att männen återfanns till störst andel i de lägsta åldersgrupperna och kvinnorna i de högsta åldersgrupperna. Fortfarande återfinns en något större andel kvinnor än män i den näst äldsta åldersgruppen men sedan 2008 återfinns kvinnorna också till störst andel i den yngsta åldersgruppen. Männen däremot återfinns till störst andel bland 21- till 29-åringarna. I tabell 3 framgår hur män respektive kvinnor är procentuellt fördelade på ålder.

Tabell 3. Män respektive kvinnor som deltog i högskoleproven 08A-09B, procentuellt fördelade på olika åldersgrupper.

Ålder	08A m/k	08B m/k	09A m/k	09B m/k
≤ 20 år	50/51	55/56	49/53	53/56
21–24 år	30/27	28/25	32/27	29/26
25–29 år	12/11	10/10	11/10	11/9
30–39 år	6/8	5/6	6/7	5/7
≥ 40 år	2/3	2/3	2/3	2/2

Resultat för olika grupper av provdeltagare

Könsskillnaderna i resultat på högskoleprovet har följts noggrant alltsedan provet började användas och en kort sammanfattning av hur dessa skillnader har sett ut återfinns t.ex. i Stage (1991):

På högskoleprovet totalt har skillnaderna mellan män och kvinnor varierat mellan 3 och 11 poäng (av 144 möjliga), men samtidigt har vissa skillnader förekommit avseende såväl andelen av den totala provtagargruppen som ålder och utbildning. Kvinnorna har vanligtvis utgjort en något större andel av provtagarna än männen; de har i genomsnitt varit något äldre och har haft något lägre bakgrundsutbildning.

Den beskrivning av provdeltagarna som gjordes 1991 avseende kvinnors och mäns ålder och utbildning är till viss del fortfarande giltig för de provdeltagare som genomförde högskoleprovet 2009. Kvinnorna förekommer i något större andel än män i de högsta åldersgrupperna (se tabell 3). Antalet kvinnor som genomfört provet har alltsedan provets begynnelse varit större än antalet män. Däremot har det skett en utjämning vad avser kvinnors och mäns utbildning (se tabell 2).

I tabell 4 visas resultatbilden för provet som helhet och för de olika delproven med avseende på medelvärdesresultaten för totalgrupperna provdeltagare och för män respektive kvinnor som har deltagit i högskoleprovet under de senaste fem åren. För åren 1996-1999 och 2000-2004 anges ett genomsnittligt värde för de fyra respektive fem åren.

Tabell 4. Poängmedelvärde (M) och standardavvikelse (s) på högskoleproven 96A–09B totalt, för män respektive kvinnor samt skillnaden mellan män och kvinnor, i råpoäng (m-k) och i standardavvikelseenheter (d).

Provtillfälle	Antal uppgifter	Totalt		Män	Kvinnor	m-k	d
		M	s	M	M		
96A-99B	122	71.65	17.74	74.93	68.93	6.00	.28
00A-04B	122	67.74	18.20	70.66	65.38	5.28	.29
05A	122	66.28	18.92	69.32	63.68	5.64	.30
05B	122	67.07	18.79	70.05	64.68	5.37	.28
06A	122	68.29	19.46	72.07	65.15	6.92	.36
06B	122	65.98	18.51	68.88	63.52	5.36	.29
07A	122	65.81	18.83	68.76	63.22	5.54	.29
07B	122	64.96	18.66	67.49	62.85	4.64	.25
08A	122	67.78	18.04	70.69	65.20	5.49	.30
08B	122	66.76	18.87	69.67	64.09	5.58	.30
09A	122	65.62	18.22	68.64	62.78	5.86	.32
09B	122	66.27	17.53	69.00	63.72	5.28	.30

Det förekommer vissa variationer i poängmedelvärdet. Det kan bero på att någon/några uppgifter inte fungerat på förväntat sätt eller på provdeltagargruppens sammansättning, men även att det var svårt att hitta rätt nivå när utvärderingsrutinerna förändrades våren 1996.

Eftersom provresultaten är giltiga i fem år är det nödvändigt att proven är jämförbara mellan provtillfällena. Därför måste proven vara parallella med avseende på både innehåll och svårighetsgrad. De ska vara lika svåra och de ska rangordna provdeltagarna på samma sätt. För att säkerställa kravet på parallellitet genomförs en normering av den reguljära provpoängen (råpoäng), som syftar till att jämställa resultaten från olika provtillfällen. Det ska vara lika lätt eller svårt att erhålla en viss normerad poäng oavsett sammansättningen på provdeltagargruppen även om provets svårighetsgrad skulle fluktuera. I samband med normeringen överförs provpoängen till en skala från 0.0 till 2.0 (s.k. normerad poäng). För närmare beskrivning av normeringsförfarandet se sidan 12.

I tabell 5 återges resultaten, dvs. poängmedelvärde och standardavvikelse, på varje delprov och på provet som helhet för hela gruppen samt skillnaden mellan män och kvinnor.

Tabell 5. Poängmedelvärde (M) och standardavvikelse (s) på de olika delproven 08A-09B för hela gruppen samt poängskillnad mellan män och kvinnor.

Delprov	08A			08B			09A			09B		
	M	s	m-k	M	s	m-k	M	s	m-k	M	s	m-k
ORD ¹	21.0	6.9	0.4	21.0	7.4	1.1	19.9	7.4	0.7	21.0	6.5	0.6
NOG ²	11.3	4.5	1.7	11.5	4.4	1.4	11.3	4.6	1.8	11.1	4.3	1.7
LÄS ³	12.8	3.8	0.3	12.8	3.9	0.2	12.2	3.5	0.2	12.4	3.8	0.1
DTK ⁴	11.0	3.6	1.5	9.8	3.6	1.6	11.1	3.6	2.1	10.0	3.7	1.8
ELF ⁵	11.6	4.0	1.6	11.6	4.0	1.3	11.2	4.0	1.1	11.7	4.0	1.0

Skillnaderna i resultat mellan män och kvinnor har alltid diskuterats mycket. Av tabell 5 framgår att det är de två delproven som har mest kvantitativa inslag, DTK och NOG, som svarar för den största könsskillnaden, ungefär 1.5-2 poäng vardera. Därefter kommer delprovet ELF som vanligtvis svarar för drygt 1 poäng av könsskillnaden, med undantag för våren 2008 då skillnaden var lika stor som för de kvantitativa delproven. För delproven ORD och LÄS brukar könsskillnaden vara liten. För de redovisade provtillfällena var könsskillnaden också liten för delprovet LÄS medan den har ökat de senaste åren på delprov ORD.

Det finns ett tämligen starkt samband mellan resultat på högskoleprovet och utbildningsbakgrund. I tabell 6 redovisas poängmedelvärde för provdeltagare med olika bakgrundsutbildning.

¹ ORD 40 uppgifter

² NOG 22 uppgifter

³ LÄS 20 uppgifter

⁴ DTK 20 uppgifter

⁵ ELF 20 uppgifter

Tabell 6. Poängmedelvärde (M) och standardavvikelse (s) på högskoleprovet våren och hösten 2009 för hela gruppen samt poängskillnad mellan män och kvinnor med olika bakgrundsutbildning (högsta angivna utbildning). Som jämförelse anges skillnaden mellan män och kvinnor våren och hösten 2008.

Utbildning	09A			09B			m-k	
	M	s	m-k	M	s	m-k	08A	08B
Grund/realskola	66.4	19.2	6.5	66.7	19.1	-1.0	4.2	8.5
Folkhögskola	62.1	18.9	2.4	62.5	18.1	-1.6	6.4	1.1
Gymn ≤ 2 år	60.2	18.4	4.8	59.6	18.0	5.0	2.7	8.8
Gymn ≥ 3 år	63.3	17.3	6.3	64.1	16.6	5.8	5.6	5.6
Högre utb ≤80p	77.1	15.7	4.5	76.1	15.3	3.8	5.2	5.0
Högre utb >80p	85.7	16.2	6.1	85.3	15.8	5.7	6.2	6.0
Totalt	65.6	18.2	5.9	66.3	17.5	5.3	5.5	5.6

Tabell 6 visar att sambandet mellan utbildning och resultat på högskoleprovet är tämligen starkt. Generellt gäller att ju längre bakgrundsutbildning desto högre provresultat. Det är inte förvånande med tanke på att provet ska mäta allmän studiefärdighet. De som har studerat längre tid bör rimligen ha såväl bättre studievana och studieteknik som större kunskaper än de som studerat kortare tid och därför bör de också ha högre resultat.

Det är inte bara utbildningens längd utan även utbildningens innehåll som har klart samband med provresultatet. Av tabell 1 framgick att den utan jämförelse största gruppen bland provdeltagarna är den med tre- eller fyraårig gymnasieutbildning (cirka 80 %). Eftersom tre- eller fyraårig gymnasieutbildning kan ha ganska olika innehåll finns det anledning att göra ytterligare uppdelningar av denna grupp.

Tabell 7. Poängmedelvärde (M) på högskoleprovet våren respektive hösten 2009 för män respektive kvinnor med avslutad treårig gymnasieutbildning uppdelade efter gymnasieinriktning⁶.

Våren 2009

	H		S		E		N		T	
	M	Kv	M	Kv	M	Kv	M	Kv	M	Kv
ORD	20.4	22.2	22.2	21.0	20.1	19.6	22.3	22.2	21.0	22.3
NOG	10.1	9.7	11.0	10.2	11.3	9.9	14.9	13.5	12.5	12.3
LÄS	11.7	12.5	12.7	12.6	12.1	11.9	13.6	13.7	12.2	12.7
DTK	10.8	9.7	11.8	10.1	12.1	10.1	13.6	12.1	12.5	11.2
ELF	11.5	11.7	12.5	11.5	11.4	10.4	13.0	12.3	11.7	11.1
Totalt	64.5	65.8	70.2	65.4	67.0	61.9	77.4	73.8	69.9	69.6

Hösten 2009

	H		S		E		N		T	
	M	Kv	M	Kv	M	Kv	M	Kv	M	Kv
ORD	21.1	22.8	22.3	21.6	20.9	20.4	23.0	22.7	21.5	21.0
NOG	10.3	9.6	10.8	9.9	11.0	9.4	14.4	13.0	12.1	11.5
LÄS	12.1	13.0	13.1	13.0	12.2	11.8	13.7	13.9	12.3	12.4
DTK	9.7	8.9	10.3	8.9	10.5	8.9	12.5	11.1	11.3	9.9
ELF	11.8	12.1	12.6	11.8	11.7	10.6	13.5	12.8	12.0	11.0
Totalt	64.9	66.5	69.1	65.2	66.3	61.0	77.2	73.5	69.2	65.9

På totalprovet har de provdeltagare som genomgått naturvetenskaplig utbildning högst resultat, och de provdeltagare som har genomgått ekonomisk utbildning har lägst resultat. Denna ordning mellan gymnasieutbildningarna har gällt alltsedan provet började användas 1977 och den tycks varken ha påverkats av förändringar av provet eller förändringar av gymnasieskolan.

En förändring har emellertid skett i skillnaderna mellan kvinnor och män för provdeltagare som har genomgått treårigt gymnasium. Under 80- och 90-talet hade kvinnor genomgående något lägre resultat än män från samma gymnasieinriktning, även om kvinnor från naturvetenskaplig utbildning hade högre resultat än män från övriga gymnasieinriktningar. Det senare gäller fortfarande, dvs. att kvinnor med naturvetenskaplig gymnasieutbildning har klart högre resultat än män från övriga gymnasieinriktningar, men lägre än

⁶ H=humanistisk, S=samhällsvetenskaplig, E=ekonomisk, N=naturvetenskaplig och T=teknisk inriktning

männen inom samma gymnasieutbildning. För provdeltagare från övriga gymnasieutbildningar har dock den förändringen skett, sedan början av 2000-talet, att kvinnor från humanistisk och teknisk utbildning tenderar att ha högre genomsnittligt resultat än männen från samma utbildningar. För teknisk utbildning gäller det dock inte för detta år. Sammanfattningsvis kan man säga att resultatskillnaderna mellan gymnasieutbildningar är betydligt större än skillnaderna mellan kvinnor och män.

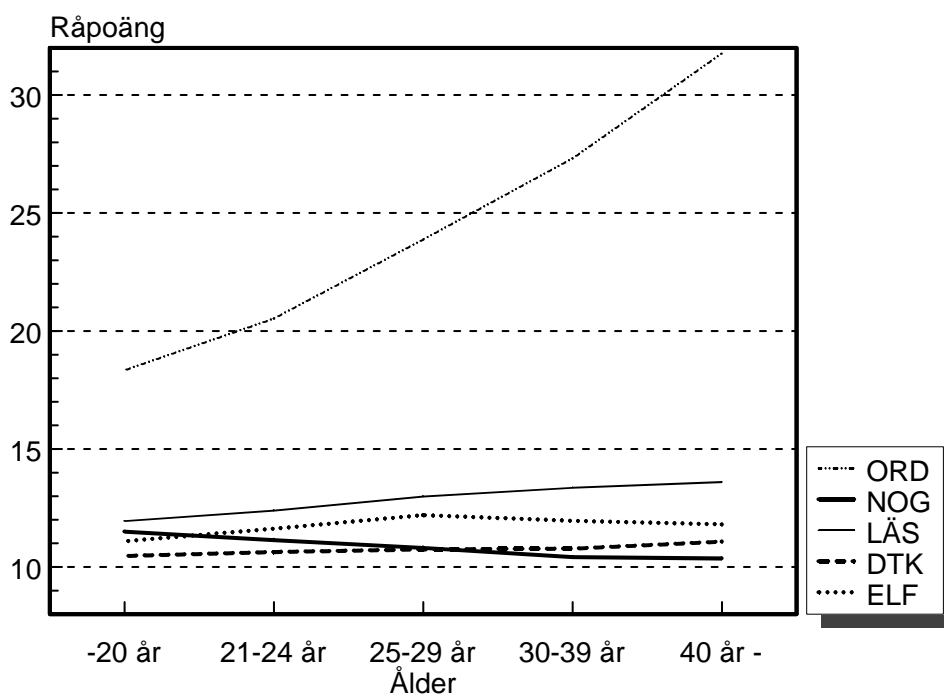
Provdeltagare som ännu ej avslutat sin gymnasieutbildning har genomgående lägre resultat än de som avslutat den.

Som tidigare nämnts har resultaten för män och kvinnor följts upp och analyserats alltsedan provet började användas år 1977. Resultat för provdeltagare med olika utbildning har också regelbundet jämförts och detsamma gäller resultat för provdeltagare i olika åldrar. I tabell 8 redovisas olika åldersgruppers resultat på hela provet och de olika delproven.

Tabell 8. Poängmedelvärde (M) och standardavvikelse (s) på högskoleprovet våren och hösten 2009 för hela gruppen samt poängskillnad i råpoäng (m-k) och i standardavvikelseenheter (d) mellan män och kvinnor i olika åldersgrupper. Som jämförelse anges motsvarande skillnader för våren och hösten 2008.

Ålder	09A			09B			m-k (d)	
	M	s	m-k (d)	M	s	m-k (d)	08A	08B
≤ 20 år	62.6	17.2	6.0 (.35)	64.0	16.7	6.2 (.37)	5.6 (.32)	5.7 (.32)
21 – 24 år	66.2	17.7	4.7 (.27)	66.5	17.0	3.0 (.18)	3.9 (.22)	4.9 (.26)
25 – 29 år	70.7	18.6	6.0 (.32)	70.5	17.8	4.2 (.23)	6.4 (.34)	4.7 (.23)
30 – 39 år	73.5	20.5	8.6 (.42)	74.1	19.6	7.0 (.36)	8.8 (.44)	8.7 (.40)
≥ 40 år	78.7	19.5	10.1 (.52)	78.9	20.0	11.4 (.57)	11.5 (.58)	10.4 (.51)

För totalgruppen provdeltagare har det generellt gällt att de yngsta och de äldsta provdeltagarna har haft de lägsta provresultaten. Efter 1996 när provets sammansättning genomgick relativt stora förändringar har de äldsta haft det högsta genomsnittliga provresultatet. Att provdeltagare som är 40 år eller äldre har det högsta provresultatet efter 1996 beror troligen på att delprovet ORD utökades med 10 uppgifter eftersom delprovet ORD har visat sig ha ett positivt samband med ålder, dvs. ju högre ålder desto högre genomsnittligt resultat på delprovet ORD. Man bör dock komma ihåg att detta är en mycket liten grupp provdeltagare, endast två-tre procent av provdeltagargruppen. Drygt 50 procent av provdeltagargruppen är 20 år eller yngre och drygt 80 procent är under 25 år.



Figur 1. Genomsnittligt provresultat på de olika delproven för provdeltagare i olika åldersgrupper.

Av figuren framgår som redan nämnts att delprovet ORD har ett starkt positivt samband med ålder. För två av delproven – NOG och ELF – har resultaten ett svagt negativt samband med ålder, dvs. ju högre ålder desto lägre genomsnittligt resultat och detta gäller för både män och kvinnor. Resultaten på delproven DTK och LÄS uppvisar inga klara tendenser vad gäller samband mellan resultat och ålder.

I samband med studier som har visat att den sociala snedrekryteringen till högre utbildning inte har förändrats har även en diskussion förts om betydelsen av social bakgrund för resultat på högskoleprovet. Provdeltagarnas sociala bakgrund ingår inte i den bakgrundsinformation som insamlas i samband med provgenomförandet och därför har inga regelbundna uppföljningar kunnat genomföras på det sätt som har gjorts avseende kön, utbildning

och ålder. I det av Riksbankens jubileumsfond finansierade VALUTA-projektet⁷ har emellertid en databas sammanställts med utförlig information, där förutom testresultat även betyg och social bakgrund ingår.

Studier har visat att provdeltagare från högre socialgrupp i genomsnitt har högre resultat på provet än provdeltagare från lägre socialgrupper⁸. Detta gäller för såväl totalprovet som samtliga delprov och varje enskild uppgift. För testet som helhet gäller att skillnaden mellan socialgrupp I och III är av samma storleksordning som skillnaden mellan kvinnor och män. Som nämndes på sidan 7 uppstår skillnaden mellan kvinnor och män i huvudsak på delproven NOG, DTK och ELF medan den är närmast obefintlig på delprovet LÅS. Från att skillnaden nästan varit obefintlig även på delprovet ORD har den de senaste åren ökat något. Vad gäller socialgrupper är skillnaderna av samma storleksordning på samtliga delprov. En viktig skillnad mellan socialgrupper finns också i andelen som genomför provet. Högre socialgrupper är klart överrepresenterade medan lägre socialgrupper är underrepresenterade. Den största skillnaden mellan socialgrupper är dock avseende uppreparandet. De provdeltagare som utnyttjar möjligheten att göra provet vid flera tillfällen kommer huvudsakligen från högre socialgrupper.

NORMERING AV HÖGSKOLEPROVET

Alltsedan högskoleprovet togs i bruk 1977 har provet rättats så att varje rätt svar har gett en råpoäng, råpoängen för samtliga delprov har summerats och summan av råpoängen har överförs till en normerad poäng.

Överföringen av råpoäng till normerad poäng har som främsta syfte att poängen skall vara jämförbar från ett provtillfälle till ett annat, dvs. det ska vara precis lika lätt eller svårt att erhålla en viss normerad poäng vid ett provtillfälle som ett annat. Eftersom högskoleprovets giltighetstid har förlängts från två till fem år skall denna likhet nu gälla för ännu längre tid än tidigare. Detta betyder att normeringen av högskoleprovet är en mycket grannlaga uppgift, eftersom den normerade poängen ska vara jämförbar för tio provtillfällen.

⁷ VALUTA *Validering av den högre utbildningens antagningssystem* var ett samarbetsprojekt mellan Institutionen för pedagogik och didaktik vid Göteborgs universitet och Institutionen för beteendevetenskapliga mätningar (tidigare Enheten för pedagogiska mätningar) vid Umeå universitet. Projektet finansierades av Riksbankens jubileumsfond och avslutades 2005.

⁸ För definition av socialgrupp se Svensson, A. (2002). Indelningen i socialgrupper. PM från Valuta-projektet. Göteborg: Göteborgs universitet, Institutionen för pedagogik och didaktik.

Den strategi som tillämpas för att fastställa poänggränser för den normerade poängen utgörs av en sammanvägning av olika jämförelser. Först studeras utfallet för samtliga provdeltagare, men eftersom sammansättningen av provdeltagargruppen inte får påverka det normerade provresultatet studeras även utfallet i två referensgrupper.

Referenspopulation I väljs ut bland provdeltagarna via proportionellt stratifierat urval för att ge en fördelning av kön, ålder och utbildning som exakt överensstämmer mellan provtillfällena. Detta innebär att eventuella skillnader i provutfall mellan olika provtillfällena för denna referenspopulation kan relateras till provet. Denna referenspopulation bestod av 3 901 personer våren 2009 och 2 513 personer hösten 2009.

Referenspopulation II utgörs av de provdeltagare som under året fyller 19 år och som vid provtillfället fortfarande går i årskurs tre i gymnasieskolan. I referenspopulation II har provdeltagarna således samma ålder och lika lång utbildning; den enda skillnaden mellan män och kvinnor i denna grupp är att de är olika representerade på olika linjer inom den teoretiska gymnasieskolan. Denna referenspopulation bestod av 9 306 personer våren 2009 och 6 166 personer hösten 2009.

Provkonstruktörerna har som målsättning att sammanställa parallella versioner av respektive delprov. Begreppet parallellitet vad avser delproven kan sägas ha två aspekter: dels skall varje delprov anpassas till en modell för hur provet skall se ut med avseende på, kognitiv nivå, innehållskategorier, uppgiftsformat, informationsmängd ämnesområden, informationsvärde, figurtyp och lösningsprocedur. Dels skall enskilda uppgifter och därmed hela delprovet ha kända definierade svårighetsnivåer, och varje uppgift skall diskriminera på ett tillfredsställande sätt och ha balans vad avser felaktiga svarsförslag.

Den normerade poängen ligger på en skala från 0.0 till 2.0, där 2.0 är det högsta resultatet.

Utfall

Totalpopulationen

Medelvärde vid vårens provtillfälle $M_{tot}=65.6$ för totalgruppen och standardavvikelsen $s=18.2$; medelvärde för män $M_M=68.6$ och för kvinnor $M_{Kv}=62.8$. En jämförelse med provet våren 2008 visar att medelvärde på provet våren 2009 var 2.2 poäng lägre och spridningen var 0.2 högre än på provet våren 2008 (se tabell 4). En jämförelse av resultaten för män respektive

tive kvinnor visar att skillnaden var ungefär densamma på detta prov som på provet våren 2008.

Medelvärden vid höstens provtillfälle var $M_{tot}=66.3$ för totalgruppen och standardavvikelse $s=17.5$. Medelvärdet för män var $M_M=69.0$ och för kvinnor $M_{KV}=63.7$. En jämförelse med provet hösten 2008 visar att medelvärdet på provet hösten 2009 var 0.5 poäng lägre och spridningen var 1.4 lägre än på provet hösten 2008 (se tabell 4). En jämförelse av resultaten för män respektive kvinnor visar att skillnaden till fördel för männen var något lägre på detta prov än på provet hösten 2008.

Referenspopulation I

Medelvärdet i referenspopulation I vid vårens provtillfälle var $M=68.5$ och standardavvikelsen $s=20.2$; motsvarande värden för våren 2008 var $M=68.4$ och $s=19.6$. För samtliga delprov blev medelvärdet för män i denna grupp $M_M=71.0$ och för kvinnor $M_{KV}=66.0$, motsvarande värden våren 2008 var för män $M_M=71.0$ och för kvinnor $M_{KV}=65.8$. I jämförelse med provet våren 2008 hade detta prov således ungefär samma svårighetsgrad för såväl män och kvinnor som totalgruppen.

Medelvärdet i referenspopulation I vid höstens provtillfälle var $M=67.7$ och standardavvikelsen $s=19.9$; motsvarande värden hösten 2008 var $M=67.9$ och $s=20.7$. För samtliga delprov blev medelvärdet för män i denna grupp $M_M=69.3$ och för kvinnor $M_{KV}=66.1$, motsvarande värden hösten 2008 var för män $M_M=70.7$ och för kvinnor $M_{KV}=65.1$. I jämförelse med provet hösten 2008 blev detta prov drygt en poäng svårare för män medan det blev en poäng lättare för kvinnor. För hela referenspopulationen blev provet ungefär lika svårt som provet föregående höst.

Referenspopulation II

För referenspopulation II våren 2009 gäller att fördelningen av provdeltagare på olika gymnasieinriktningar har förändrats såtillvida att gymnasister med teknisk inriktning har ökat med cirka två procentenheter jämfört med våren 2008, medan gymnasister med naturvetenskaplig inriktning har minskat med två procentenheter.

För referenspopulation II var medelvärdet på vårens prov $M=63.3$ och standardavvikelsen $s=16.6$; motsvarande värden våren 2008 var $M=66.9$ och $s=17.2$. För män var medelvärdet $M_M=66.4$ och för kvinnor $M_{KV}=60.3$, motsvarande värden våren 2008 var för män $M_M=69.5$ och för kvinnor $M_{KV}=64.4$. I jämförelse med provet våren 2008 var detta prov drygt tre råpoäng svårare för män och drygt fyra råpoäng svårare för kvinnor i denna

referenspopulation. För hela referenspopulation II blev detta prov drygt tre och halv råpoäng svårare än provet våren 2008.

För referenspopulation II hösten 2009 gäller att fördelningen av provdeltagare på olika gymnasieinriktningar är ungefär densamma som hösten 2008.

För hela referenspopulation II hösten 2009 var medelvärdet $M=64.8$ och standardavvikelsen $s=16.2$; motsvarande värden hösten 2008 var $M=64.5$ och $s=17.1$. För män var medelvärdet $M_M=68.5$ och för kvinnor $M_K=61.6$, motsvarande värden för denna referensgrupp hösten 2008 var för män $M_M=68.3$ och för kvinnor $M_K=61.4$. Detta prov hade således ungefär samma svårighetsnivå som provet föregående höst för såväl totalpopulationen som för män respektive kvinnor.

Sammanfattningsvis kan sägas att för totalgruppen och för referenspopulation II blev provet drygt två råpoäng respektive drygt tre och en halv råpoäng svårare våren 2009 än våren 2008. För referenspopulation I blev däremot provet ungefär lika svårt. Medelvärdet på provet hösten 2009 för referenspopulation I och II ligger 0.2 respektive 0.3 poäng högre än på provet hösten 2008 och för totalpopulationen ligger medelvärdet 0.5 poäng lägre. Utförligare beskrivning av resultaten se Stage, Hamrén & Jonsson, 2009 a och b.

Normering baserad på IRT

Sedan 1999 har, i studiesyfte, även normering med Item Response Theory (IRT) genomförts. Ekvivalering baserad på IRT har vissa fördelar i jämförelse med klassiska ekvivaleringsmetoder (se Cook & Eignor, 1989, 1991). IRT ekvivalering kan enligt Emons, (1998) ses som en process i fyra steg:

1. Val av en lämplig design. I praktiken bestäms designen ofta av praktiska begränsningar. I detta fall utprovningsmodellen.
2. Bestämning av lämplig IRT-modell. Den modell som i tidigare studier befunnits mest lämplig är tre-parameter-modellen (Stage, 1996, 1997a, 1997b, 1997c, 1997d).
3. Bestämning av en gemensam skala, dvs. alla parametrar måste överföras till samma skala. Detta görs med dataprogrammet.
4. Beroende på vilken skala som används för resultaten: IRT sann poäng ekvivalering eller observerad poäng ekvivalering.

Vid IRT-ekvivalering av observerad poäng används IRT-modellen för erhålla estimat av frekvensfördelningen för observerad poäng hos en population på både det gamla och det nya provet. Poängen kan ekvivaleras betingat på referenspopulationen eller den nya populationen.

För provtillfället våren 2009 fanns det inget prov, som var lämpligt som referensprov, eftersom uppgifterna hade hämtats från ett flertal olika utprovningversioner. Det prov som valdes blev provet våren 2006, vid vilket sex uppgifter i detta prov hade utprovats. Dessutom inkluderades denna gång nio uppgifter från samma utprovningssomgång trots att dessa uppgifter hade genomgått vissa förändringar. De estimerade resultaten blev att den population som genomförde provet våren 2006 hade medelvärdet 65.7 på det provet och skulle ha haft 62.0 på provet våren 2009. De estimerade medelvärdena för den population som genomförde provet våren 2006 var 68.5 på provet våren 2006 och 64.8 på provet våren 2009. Enligt IRT är därmed provet våren 2009 nästan fyra råpoäng svårare än provet våren 2006, och provdeltagarna vid detta provtillfälle nästan tre råpoäng bättre. De observerade medelvärdena var 68.3 våren 2006 och 65.6 våren 2009.

För provtillfället hösten 2009 bedömdes provet våren 2007 vara lämpligast som referensprov eftersom det var det provtillfälle där flest antal uppgifter (17) till provet hösten 2009 hade prövats ut. De estimerade resultaten var att den

population som genomförde provet våren 2007 hade medelvärdet 65.0 på det provet och skulle ha haft 65.2 på provet hösten 2009. De estimerade medelvärdena för den population som genomförde provet hösten 2009 var 65.0 på provet våren 2007 och 65.2 på provet hösten 2009. Enligt utfallet av IRT var därmed provet hösten 2009 ungefär lika lätt eller svårt som provet våren 2007 och provdeltagarna hösten 2009 var ungefär lika duktiga som provdeltagarna våren 2007. De observerade medelvärdena var 65.8 våren 2007 och 66.3 hösten 2009.

Normeringsgränser

Sedan våren 1996, när provet genomgick tämligen stora förändringar, används inte längre en färdig normeringstabell. Såväl tre- som fyrapoängsintervaller används numera mot tidigare enbart fyrapoängsintervaller (se Stage & Hamrén, 1996). Fördelen med en variabel normeringstabell är att en bättre anpassning kan göras till eventuella variationer i spridningen hos provdeltagargruppen. Resultaten analyseras såväl för totalpopulationen som för de två referensgrupperna. Den strategi som används för normering är att finna de normerade värden som bäst fördelar provdeltagarna (i samtliga tre grupper) på olika normerade poäng så att överensstämmelsen med tidigare fördelningar blir så god som möjligt, härvid används lika-percentil-metoden (se t.ex. Holland & Rubin, 1982).

De resultat som kan betraktas som viktigast är de normerade poängerna mellan 1.0 och 2.0, eftersom det är provdeltagare med dessa resultat som konkurrerar om de mest attraktiva studieplatserna. Detta är anledningen till att anpassningen av fördelningarna påbörjas uppifrån, dvs. från den normerade poängen 2.0. Därefter bestäms intervallen så att varje normerad poäng så väl som möjligt motsvarar de fördelningar för de två referenspopulationerna och totalpopulationen som har gällt för tidigare prov. Den procentuella fördelningen av provdeltagarna på olika normerade poäng för såväl totalgruppen som de båda referensgrupperna återfinns i bilaga 1a, b och c. De kumulativa frekvensfördelningarna illustreras dessutom i figurerna i bilaga 2a, b och c.

Resultatet av den genomförda normeringen av högskoleprovet såväl våren 2009 som hösten 2009 blev att den normerade poängen 1.0 erhålls i intervallet 68-71 råpoäng, vilket är detsamma som i proven våren och hösten 2008. För den normerade poängen 2.0 krävs 109 råpoäng eller mer, vilket också är detsamma som våren och hösten 2008. Den fullständiga normeringstabellen för provet våren respektive hösten 2009 återfinns i bilaga 3a och 3b.

UPPREPAT PROVDELTAGANDE

Som tidigare nämnts är det tillåtet att göra provet hur många gånger man önskar och när mer än ett provresultat finns är det alltid det bästa resultatet som räknas. Möjligheten att genomföra provet flera gånger utnyttjas tämligen flitigt och vid varje provtillfälle brukar cirka 30 procent (något högre på våren och något lägre på hösten) av provdeltagarna ha genomfört minst ett tidigare prov.

Ett flertal studier har genomförts av effekterna av upprepat provdeltagande (Henriksson, 1991, 1993, 1995; Henriksson & Bränberg, 1992, 1994; Törnqvist & Henriksson, 2004, 2006). Törnqvist och Henriksson (2004) jämförde resultaten avseende upprepat provdeltagande för proven hösten 1986, hösten 1991, hösten 1993, hösten 1997 och hösten 2002. De skillnader i normerad poäng som rapporterades mellan det första och andra provtillfället var 0.055, 0.083, 0.111, 0.093 respektive 0.083. Effekten av upprepat provdeltagande hade således ökat från hösten 1986 till hösten 1993 och denna ökning förklarades med att provdeltagargruppen hade förändrats under den studerade tidsperioden från att huvudsakligen ha bestått av personer över 24 år till huvudsakligen bestå av personer under 24 år. Yngre personer som fortfarande utbildar sig bör rimligen förbättra sin studieförmåga mer än äldre. Dessutom har förändringar i provets sammansättning skett så att 1992 utgick ett prov som mätte studiefärdighet och ersattes av ett engelskt prov som mäter mer skolkunskaper.

Gustafsson och Westerlund (1994) genomförde en studie av provet våren 1991 där resultaten för de provdeltagare som var födda 1972 och som även genomfört provet hösten 1990 var en ökning med i medeltal 3.4 råpoäng från första till andra provtillfället. Enligt normeringen av de två proven 90B och 91A var provet 91A en råpoäng lättare, vilket innebär att ökningen motsvarar 2.4 råpoäng.

Det föreligger avsevärda metodproblem när det gäller att bestämma effekterna av upprepat provdeltagande. Som visades i en studie av proven 98A och 98B (Stage & Ögren, 2001) blir resultaten helt olika om man jämför samma individer eller olika individer och om man enbart studerar poängökningen eller om man jämför senaste resultat med bästa tidigare resultat. Skillnaden i råpoäng mellan dem som gjorde provet för första och fjärde gången var ungefär 19 råpoäng, medan ökningen från första till fjärde provtillfället för samma individer var i genomsnitt ungefär nio råpoäng och skillnaden mellan resultatet vid fjärde provtillfället och bästa tidigare resultat var marginell. Det visade sig även att de initiala skillnaderna var avsevärda mellan de provdeltagare som under en tvåårsperiod hade genomfört ett, två, tre respektive fyra prov; skillnaden mellan dem som bara hade gjort ett prov och dem som hade gjort fyra prov uppgick till ungefär tio råpoäng.

I tabell 9 visas resultat för de provdeltagare våren 2009 som genomförde provet för första, andra, tredje respektive fjärde gången och i tabell 10 visas detsamma för provdeltagarna hösten 2009.

Tabell 9. Råpoäng och (normerad poäng) för de provdeltagare våren 2009, som genomfört provet en, två, tre eller fyra gånger sedan hösten 2007.

Antal prov	N	Provtillfälle			
		09A	08B	08A	07B
1	31 385	63.0 (0.8)			
2	6 346	68.7 (1.0)	66.1 (0.9)		
	2 613	68.9 (1.0)		66.5 (0.9)	
	1 043	70.7 (1.0)			63.4 (0.8)
3	1 952	74.4 (1.1)	72.7 (1.1)	70.2 (1.0)	
	559	75.5 (1.2)	73.5 (1.1)		66.2 (0.9)
	727	74.7 (1.1)		73.0 (1.1)	66.8 (0.9)
4	1 000	80.6 (1.3)	79.0 (1.2)	77.6 (1.2)	70.3 (1.0)

Tabell 10. Råpoäng (och normerad poäng) för de provdeltagare hösten 2009, som genomfört provet en, två, tre eller fyra gånger sedan våren 2008.

Antal prov	N	Provtillfälle			
		09B	09A	08B	08A
1	28 371	63.7 (0.9)			
2	6 039	68.8 (1.0)	64.9 (0.9)		
	1 575	68.8 (1.0)		63.9 (0.9)	
	1 363	70.0 (1.0)			65.8 (0.9)
3	1 823	74.4 (1.1)	71.1 (1.0)	68.6 (1.0)	
	712	74.7 (1.1)	71.4 (1.0)		68.5 (1.0)
	485	74.5 (1.1)		71.1 (1.0)	70.5 (1.0)
4	1 057	80.1 (1.3)	77.2 (1.2)	75.4 (1.1)	73.2 (1.1)

Av tabellerna 9 och 10 framgår att de provdeltagare som under den aktuella tvåårsperioden genomfört samtliga fyra prov i genomsnitt hade 1.0-1.1 i normerad poäng vid det första provtillfället. De som genomfört tre prov hade vid det första provtillfället, med två undantag, den normerade poängen 1.0. De som genomfört ett eller två prov hade vid det första provtillfället, med två undantag, den normerade poängen 0.9. Den genomsnittliga ökningen av resultatet från det första till det andra provtillfället är generellt 0.1 normerad poäng. Viktigt att poängtera är även att de värden som redovisats såväl här som i tidigare studier är medelvärden för tämligen stora grupper av individer. På individnivå kan effekterna av upprepat provdeltagande vara betydligt större i både positiv och negativ riktning än de medelvärdessresultat som redovisats här. En provdeltagare kan dock aldrig förlora i resultat genom att upprepa provdeltagandet eftersom det alltid är det bästa resultatet som gäller.

I tabell 11 visas andelen provdeltagare vid varje provtillfälle sedan våren 2001 som genomfört ett, två, tre respektive fyra prov.

Tabell 11. Andel provdeltagare vid provtillfällena 01A-09B, som genomförde provet för första, andra, tredje respektive fjärde gången.

	Antal genomförda prov			
	1	2	3	4
01A	60	27	10	3
01B	68	22	8	2
02A	67	24	7	2
02B	70	21	7	2
03A	69	23	6	2
03B	70	21	7	2
04A	68	23	7	2
04B	68	22	8	2
05A	66	23	7	2
05B	67	22	8	2
06A	66	23	8	3
06B	67	22	8	3
07A	67	23	7	3
07B	68	21	8	3
08A	69	22	7	2
08B	67	22	8	3
09A	69	22	7	2
09B	68	22	7	3

Det tycks finnas en tendens till att intresset för att upprepa provdeltagandet kan ha minskat. Fram till och med provet våren 2001 hade ungefär 40 procent av totalgruppen genomfört minst ett tidigare prov, därefter har andelen minskat till ungefär 30 procent.

AVSLUTNING

Antalet deltagare i det högskoleprov som genomfördes den 28 mars år 2009 var 45 737, av vilka 49 procent var män och 51 procent var kvinnor och den 24 oktober 2009 genomfördes provet av 41 422 personer, av vilka 48 procent var män och 52 procent var kvinnor. Sammanlagt under året var det därmed 87 159 personer som genomförde provet. En ökning av antalet provdeltagare med cirka 33 procent sedan föregående år, som var det år med näst lägsta antalet provdeltagare sedan de nya antagningsreglerna trädde i kraft 1991.

Sedan i början på 1990-talet har majoriteten av provdeltagarna varit ganska unga. Drygt 80 procent av provdeltagarna år 2009 är 24 år eller yngre. En något större andel män återfinns i de lägsta åldersgrupperna medan en något större andel kvinnor återfinns i de högsta åldersgrupperna.

Majoriteten av provdeltagarna har tre- eller fyraårig gymnasieutbildning. Män har vid tidigare provtillfällen alltid haft något högre utbildning än kvinnor något som börjat utjämnas under senare år.

De resultat som presenterats för högskoleprovet visar att medelvärdet på provet våren 2009 var drygt två råpoäng lägre än på det prov som genomfördes våren 2008 och medelvärdet på provet hösten 2009 var en halv råpoäng lägre än på provet hösten 2008. Medelvärdet på provet våren 2009 var 65.6 råpoäng och på hösten 2009 var medelvärdet 66.3 råpoäng.

Vad gäller resultaten för undergrupper av provdeltagare kan konstateras att sambandet mellan provresultat och utbildning är oförändrat; fortfarande gäller att ju högre utbildning desto bättre resultat. För provdeltagare med gymnasial utbildning gäller också som tidigare att de gymnasister som har naturvetenskaplig inriktning i genomsnitt klarar provet bäst och de med ekonomisk inriktning har de lägsta genomsnittliga resultaten.

Vad gäller resultaten för olika åldersgrupper kan konstateras att förhållandet mellan åldersgrupperna har förändrats något efter den förändring av provet som genomfördes våren 1996. Som framgår av tabell 8 hade den äldsta åldersgruppen det högsta genomsnittliga resultatet och det gäller för både män och kvinnor. Detta beror troligen på det ökade antalet uppgifter i delprovet ORD. Som också framgår i figur 1 är delprovet ORD det delprov som har det starkaste sambandet med ålder. Den äldsta gruppen är dock liten, endast två-tre procent av hela provdeltagargruppen.

Resultatskillnaden mellan totalgrupperna män och kvinnor var ungefär som den brukar, dvs. delproven NOG och DTK svarar vardera för ungefär 2 råpoäng, delprovet ELF för ungefär en råpoäng. För delproven ORD och LÄS brukar könsskillnaden vara liten men de senaste åren har denna skillnad ökat något på delprovet ORD.

För provet som helhet och för den totala gruppen provdeltagare är skillnaden mellan män och kvinnor 0.32 standardavvikelseenheter till männens fördel för våren 2009 och 0.30 för hösten 2009. När det gäller könsskillnader i resultat avviker de inte från de resultat som brukar erhållas internationellt vid användning av likartade test. Det mest använda urvalstestet, det amerikanska

Scholastic Assessment Test (SAT), som använts sedan 1926, ger skilda poäng för en verbal och en matematisk del. På den matematiska delen har den genomsnittliga poängen alltid legat högre för män än för kvinnor, skillnaden är ungefär en halv standardavvikelseenhet. På den verbala delen låg den genomsnittliga poängen något högre för kvinnor fram till år 1972, men därefter har män haft en genomsnittligt högre poäng även på denna del; skillnaden har varit lite drygt en tiondels standardavvikelseenhet (Wilder & Powell, 1989).

Forskare och pedagoger har länge varit medvetna om existensen av skillnaderna mellan män och kvinnor i resultat på objektiva test och det har förekommit intensiv forskning med syfte att förklara och förstå dessa skillnader. Några entydiga svar finns inte men att försöka bortförklara skillnaderna genom att säga att de orsakas av proven kommenterar Cleary (1991) på följande sätt:

”An obvious question to be asked..... is whether the observed differences are the result of defects in the tests. I believe that the pervasiveness of the pattern of gender differences on tests from so many different publishers makes this explanation unacceptable.” (p 74)

Alltsedan de första försöksversionerna inför ett blivande högskoleprov började prövas ut har resultatskillnaderna mellan män och kvinnor observerats och analyserats. Varje uppgift utsätts för en minutiös kontroll och en mängd speciella studier har genomförts för att söka klarlägga var och varför skillnaderna uppstår⁹. Betydligt mer forskning krävs dock på detta område innan alla frågor kan besvaras.

Det som gäller för deltagarna i högskoleprovet är att de utgör en utvald grupp, dvs. det är inte fråga om varken ett slumpmässigt eller representativt urval. Det är provdeltagarna själva som bestämmer att de ska genomföra provet. Ett flertal studier har visat att skillnader i resultat mellan män och kvinnor genomgående ökar till fördel för män när selektiviteten i grupperna är hög (se t.ex. Wilder & Powell, 1988, Cleary, 1991, Willingham & Cole, 1977). Deltagarna i högskoleprovet är inte bara totalt sett en utvald grupp, det tycks även vara så att det är hårdare selektion bland männen än bland kvinnorna (se Stage, 1992a).

⁹ Se t.ex. Wedman, I & Stage, C. (1983), Stage, C & Wedman, I. (1984), Stage, C. (1985), (1991), (1992), (1996), Bränberg, K., Henriksson, W., Nyquist, H. & Wedman, I. (1990), Wester-Wedman, A. (1992)

Resultatjämförelser mellan grupper är alltid svåra att göra på ett rättvisande sätt, eftersom grupperna kan vara olika i så många andra avseenden än just den man för tillfället är intresserad av. Viktigt att komma ihåg när det gäller gruppskillnader är också att skillnaden inom grupper oftast är större än skillnaden mellan grupper, dvs. i detta fall att skillnaden mellan de män som erhåller höga respektive låga poäng är mycket större än den genomsnittliga skillnaden mellan kvinnor och män.

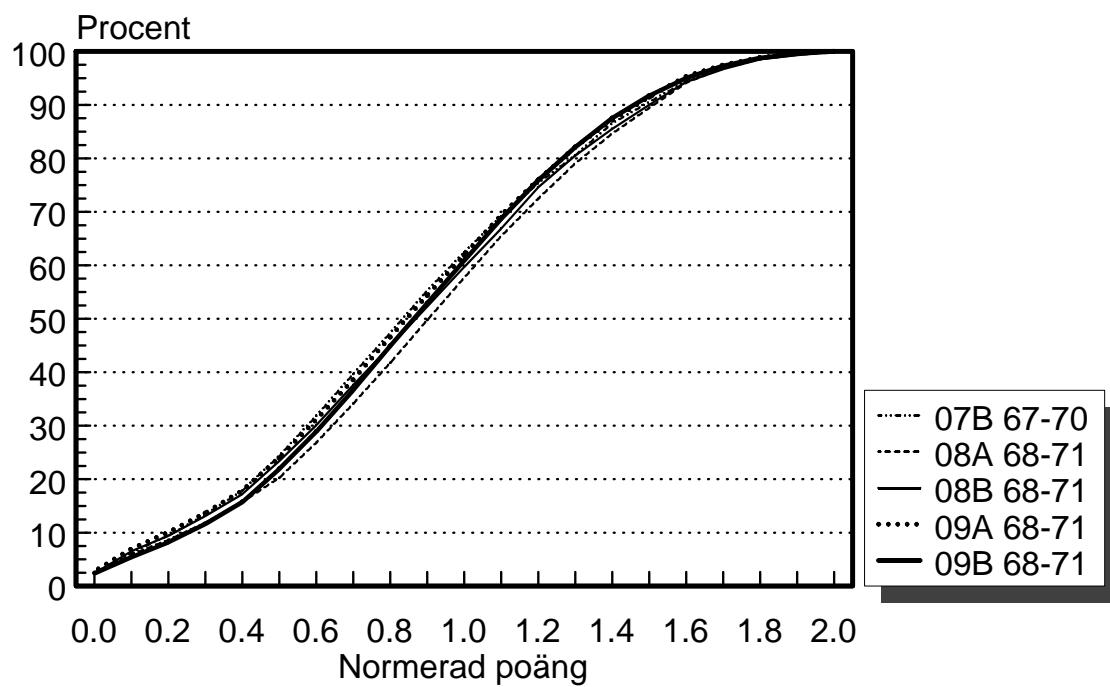
REFERENSER

- Bränberg, K., Henriksson, W., Nyquist, H. & Wedman, I. (1990). The Influence of Sex, Education and Age on Test Scores on the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 34, 189-203.
- Cleary, A. (1991). Gender Differences in Aptitude and Achievement Test Scores. I *Invitational Conference Proceedings, ETS*. p 51-90.
- Cook, L.L. & Eignor, D.R. (1989). Using item response theory in test score equating. *International Journal of Educational Research*, 13, 161-173.
- Cook, L.L. & Eignor, D.R. (1991). IRT Equating Methods. *Educational Measurement: Issues and Practices*, 10, 37-45.
- Emons, W.H.M. (1998). *Nonequivalent Groups IRT Observed Score Equating*. Its applicability and Appropriateness for the Swedish Scholastic Aptitude Test (Em No. 32). Umeå: Umeå University, Department of Educational Measurement.
- Gustafsson, J.-E. & Westerlund, A. (1994). Socialgruppskillnader i prestationer på högskoleprovet, I Eriksson, R. & J.O. Jonsson (Red.), *Sorteringen i skolan. Studier av snedrekrytering och utbildningens konsekvenser*. Stockholm: Carlsson Bokförlag.
- Henriksson, W. (1991). *Effekter av upprepat provdeltagande* (Pm nr 40). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Henriksson, W. (1993). *The problem of repeated test taking and the SweSAT* (Em No. 5). Umeå: University of Umeå, Department of Education, Division of Educational Measurement.
- Henriksson, W., & Bränberg, K. (1992). *Effekten av upprepat provtagande. En studie av poängförändringar från första till andra provgenomförandet* (Pm nr 70). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Henriksson, W., & Bränberg, K. (1994). The effects of practice on the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 38(2), 129-148.
- Henriksson, W. (1995). *Repeated test taking and the SweSAT* (Em No. 13). Umeå: Umeå University, Department of Educational Measurement.
- Holland, P.W., & Rubin, D. (1982). *Test Equating*. New York: Academic Press, Inc.

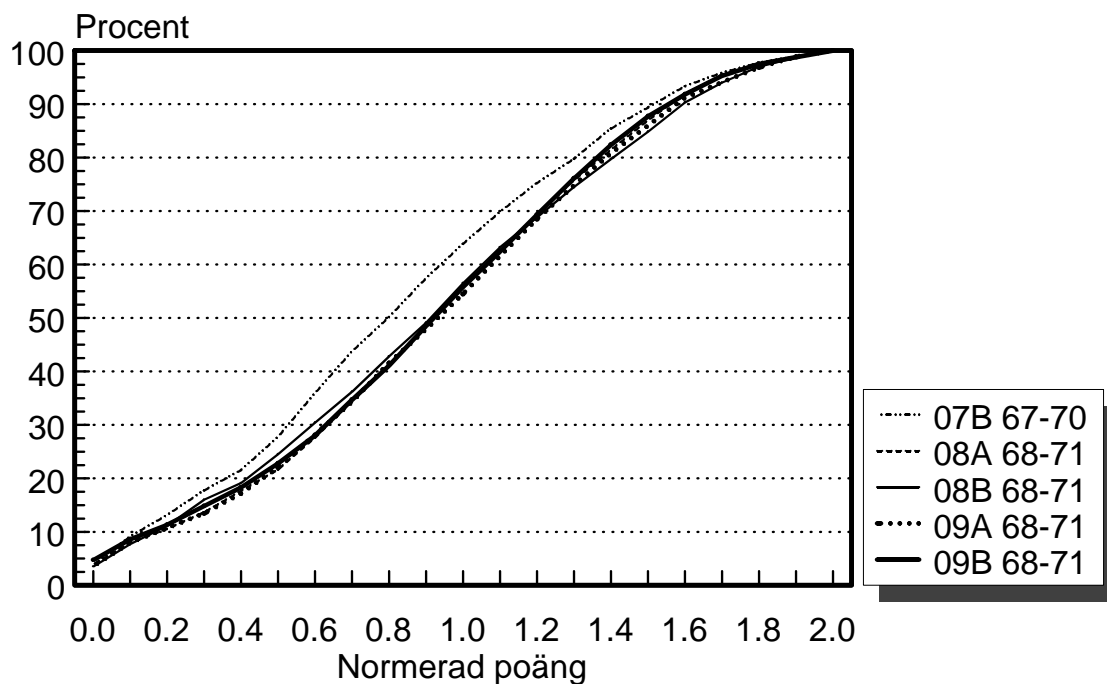
- Stage, C., & Wedman, I. (1984). Lika möjligheter till utbildning. I *Utbildningsstatistisk årsbok 1983/84* (s. 33-46). Stockholm: SCB.
- Stage, C. (1985). *Gruppskillnader i provresultat. Uppgiftsinnehållets betydelse för resultatskillnader mellan män och kvinnor på prov i ordkunskap och allmänorientering* (akademisk avhandling för doktorsexamen). Umeå: Umeå universitet, Pedagogiska institutionen.
- Stage, C. (1988). Gender Differences in Test Results. *Scandinavian Journal of Educational Research*, 36, 101-11.
- Stage, C. (1991). *Könsskillnader i resultat på högskoleprovet våren 1999* (Pm nr 42). Umeå: Umeå universitet, Pedagogiska institutionen.
- Stage, C. (1992a). *Högskoleprovet hösten 1991. Provdeltagargruppens sammansättning och resultat* (Pm nr 59). Umeå: Umeå universitet, Pedagogiska institutionen.
- Stage, C. (1995). *Utprovning av provuppgifter* (Pm nr 106). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Stage, C. (1996). *Samma poäng – samma prestation? En jämförelse på delprovs- och uppgiftsnivå mellan män och kvinnor med samma resultat på högskoleprovet* (Pm nr 113). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Stage, C. (1996). *An Attempt to fit IRT models to the DS subtest in the SweSAT* (Em No. 19). Umeå: University of Umeå, Department of Educational Measurement.
- Stage, C. & Hamrén, M. (1996). *Normering av högskoleprovet våren 1996* (Stencil). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Stage, C. (1997a). *The applicability of item response models to the SweSAT: A study of the DTM subtest* (Em No. 21). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1997b). *The applicability of item response models to the SweSAT: A study of the ERC subtest* (Em No. 24). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1997c). *The applicability of item response models to the SweSAT: A study of the READ subtest* (Em No. 25). Umeå: University of Umeå, Department of Educational Measurement.
- Stage, C. (1997d). *The applicability of item response models to the SweSAT: A study of the WORD subtest* (Em No. 26). Umeå: Umeå University, Department of Educational Measurement.

- Stage, C. & Ögren, G. (2001). *Högskoleprovets utveckling under åren 1977–2000. Provets sammansättning och provdeltagargruppens sammansättning och resultat* (Pm nr 169). Umeå: Umeå universitet, Enheten för Pedagogiska mätningar.
- Stage, C., Hamrén, M. & Jonsson, C. (2009a). *Normering av högskoleprovet våren 2009* (Arbetsrapport nr 33). Umeå: Umeå universitet, Institutionen för beteendevetenskapliga mätningar.
- Stage, C., Hamrén, M. & Jonsson, C. (2009b). *Normering av högskoleprovet hösten 2009* (Arbetsrapport nr 34). Umeå: Umeå universitet, Institutionen för beteendevetenskapliga mätningar.
- Svensson, A. (2002). *Indelningen i socialgrupper. PM från Valuta-projektet*. Göteborg: Göteborgs universitet, Institutionen för pedagogik och didaktik.
- Törnkvist, B. & Henriksson, W. (2004). *SweSAT repeat* (Em No. 24). Umeå: Umeå University, Department of Educational Measurement.
- Törnkvist, B. & Henriksson, W. (2006). *Validity issues concerning repeated test taking of the SweSAT* (Em No. 56). Umeå: Umeå University, Department of Educational Measurement.
- Wedman, I. & Stage, C. (1983). The Significance of Contents for Sex Differences in Test Results. *Scandinavian Journal of Educational Research*, 27, 49-71.
- Wester-Wedman, A. (1992). *Item Bias with Respect to Gender Interpreted in the Light of Problem-Solving Strategies*. Paper presented at IAEA's 18th Annual Conference at St Patrick's College, Dublin, Ireland.
- Wilder, G., & Powell, K. (1989). *Sex Differences in Test Performance: A Survey of the Literature* (College Board Report, No 89-3). Educational Testing Service: Princeton
- Willingham, W. W., & Cole, N.S. (1997). *Gender and Fair Assessment*. LEA. Mahwah, New Jersey.
- Ögren, G., Stage, C., Åström, K., & Lexelius, A. (1995). *Försök inför förändring av högskoleprovets utprövningsrutiner* (Pm nr 107). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.

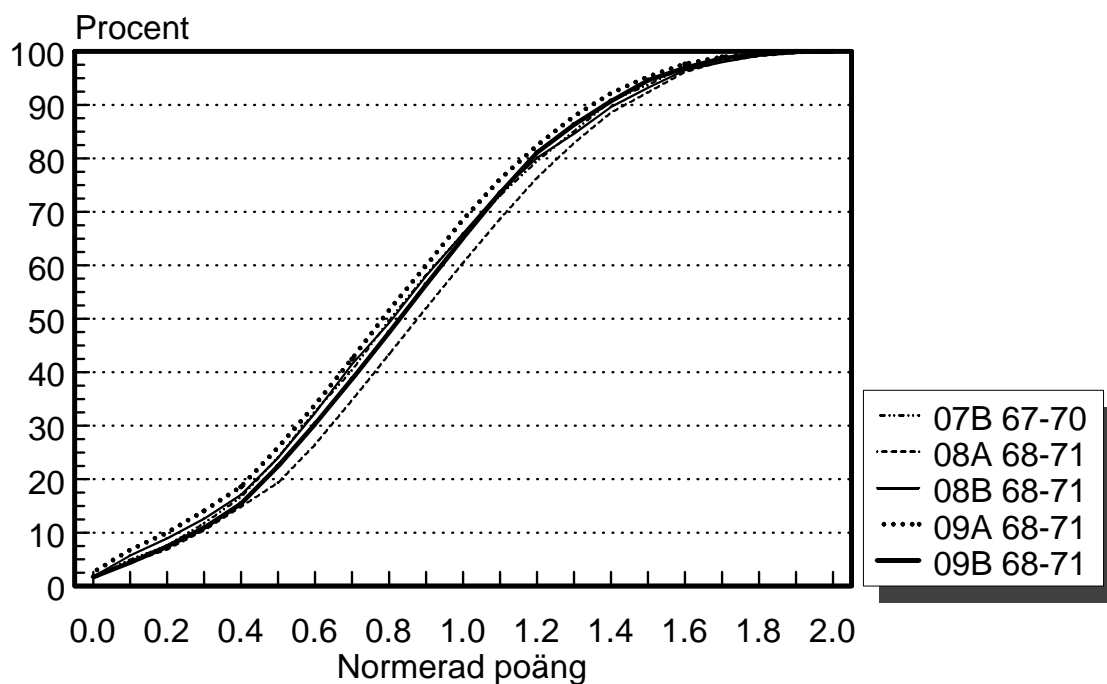
NORMERINGSUTFALL TOTALT



NORMERINGSUTFALL REF.POP. 1



NORMERINGSUTFALL REF.POP. 2



09/04/23

NORMERINGSTABELL GÄLLANDE HÖGSKOLEPROVET 2009-03-28 (09:A)

TABELL AVSEENDE ÖVERFÖRING AV ANTAL RÄTTA SVAR TILL NORMERAD
 POÄNG MED ANGIVANDE AV ANTAL, ANDEL OCH KUMULATIV ANDEL PROV-
 DELTAGARE FÖR RESPEKTIVE NORMERAD POÄNG.

ANTAL RÄTTA SVAR	NORMERAD POÄNG	ANTAL PROV- DELTAGARE	ANDEL PROV- DELTAGARE (%)	KUMULATIV ANDEL PROV- DELTAGARE (%)
0- 32	0.0	1248	2.7	2.7
33- 38	0.1	1902	4.2	6.9
39- 41	0.2	1411	3.1	10.0
42- 44	0.3	1659	3.6	13.6
45- 47	0.4	1902	4.2	17.8
48- 51	0.5	2830	6.2	23.9
52- 55	0.6	3205	7.0	31.0
56- 59	0.7	3429	7.5	38.5
60- 63	0.8	3704	8.1	46.5
64- 67	0.9	3523	7.7	54.3
68- 71	1.0	3471	7.6	61.8
72- 75	1.1	3396	7.4	69.3
76- 79	1.2	3026	6.6	75.9
80- 83	1.3	2776	6.1	82.0
84- 87	1.4	2352	5.1	87.1
88- 91	1.5	1990	4.4	91.4
92- 96	1.6	1768	3.9	95.3
97-100	1.7	953	2.1	97.4
101-104	1.8	638	1.4	98.8
105-108	1.9	354	.8	99.6
109-122	2.0	200	.4	100.0
		----- 45737	----- 100.0	

MEDELVÄRDE = 0.90

STANDARDAVVIKELSE = .46

09/11/11

NORMERINGSTABELL GÄLLANDE HÖGSKOLEPROVET 2009-10-24 (09:B)

TABELL AVSEENDE ÖVERFÖRING AV ANTAL RÄTTA SVAR TILL NORMERAD
 POÄNG MED ANGIVANDE AV ANTAL, ANDEL OCH KUMULATIV ANDEL PROV-
 DELTAGARE FÖR RESPEKTIVE NORMERAD POÄNG.

ANTAL RÄTTA SVAR	NORMERAD POÄNG	ANTAL PROV- DELTAGARE	ANDEL PROV- DELTAGARE (%)	KUMULATIV ANDEL PROV- DELTAGARE (%)
0- 33	0.0	985	2.4	2.4
34- 38	0.1	1267	3.1	5.4
39- 41	0.2	1144	2.8	8.2
42- 44	0.3	1423	3.4	11.6
45- 47	0.4	1668	4.0	15.7
48- 51	0.5	2593	6.3	21.9
52- 55	0.6	2906	7.0	28.9
56- 59	0.7	3207	7.7	36.7
60- 63	0.8	3403	8.2	44.9
64- 67	0.9	3324	8.0	52.9
68- 71	1.0	3311	8.0	60.9
72- 75	1.1	3219	7.8	68.7
76- 79	1.2	2969	7.2	75.9
80- 83	1.3	2575	6.2	82.1
84- 87	1.4	2247	5.4	87.5
88- 91	1.5	1762	4.3	91.7
92- 95	1.6	1360	3.3	95.0
96- 99	1.7	955	2.3	97.3
100-103	1.8	616	1.5	98.8
104-108	1.9	340	.8	99.6
109-122	2.0	148	.4	100.0
		----- 41422	----- 100.0	

MEDELVÄRDE = 0.91

STANDARDAVVIKELSE = .45

RAPPORTER FRÅN INSTITUTIONEN FÖR BETEENDEVETENSKAPLIGA MÄTNINGAR

Utgivna rapporter i föregående serie:

<http://www.umu.se/edmeas/publikationer/index.html>

2004

- BVM nr 1. LÄRARENKÄT OM DE NATIONELLA PROVEN I MATEMATIK. Kurs B, C och D hösten 2003. Maria Ericsson, Björn Sigurdsson
- BVM nr 2. ORDFÖRSTÅELSE. En litteraturstudie med anknytning till högskoleprovets ORD-prov. Sandra Scott
- BVM nr 3. SJÄLVVÄRDERING SOM METOD FÖR ATT MÄTA MÅLUPPFYLLELSE VIA PROV. Anna Sundström
- BVM nr 4. ATT MÄTA SKRIVFÖRMÅGA. En forskningspresentation om provformat, reliabilitet, validitet samt sociala aspekter. Marit Sigurdson
- BVM nr 5. KLASSISK OCH MODERN TESTTEORI. Analys av det teoretiska och det praktiska körkortsprovet. Marie Wiberg
- BVM nr 6. UTBYTESKOMPLETTERINGAR BLAND DEM SOM AVSLUTADE GYMNASIET 1997–2001. Differenser mellan avgångsbetyg från gymnasiet och betyg som har kompletterats efter den ordinarie gymnasieskolan. Kent Löfgren
- BVM nr 7. HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2004. Provdeltagargruppens sammansättning och resultat. Christina Stage, Gunilla Ögren
- BVM nr 8. HÖGSKOLEPROVET OCH DE MÅLRELATERADE BETYGEN. En studie av de första eleverna med de nya gymnasiebetygen. Anders Lexelius

2005

- BVM nr 9. LÄRARES SKATTNINGAR AV SINA ELEVERS PROVRESULTAT. Gunilla Näsström
- BVM nr 10. LÄRARENKÄT OM DE NATIONELLA PROVEN I MATEMATIK. Kurs B, C och D våren 2004. Maria Ericsson, Björn Sigurdsson
- BVM nr 11. SOCIALGRUPPSSKILLNADER I RESULTAT PÅ HÖGSKOLEPROVET. Christina Stage
- BVM nr 12. ÄR DET SVÅRARE ATT DELA MED FYRA ÄN MED TVÅ NÄR MAN LÄSER MATTE C? En jämförelse av svårighetsgrad mellan olika versioner av matematikuppgifter i Nationella kursprov. Ewa Bergqvist, Anna Lind

- BVM nr 13. DEN SVENSKA FÖRARPRÖVNINGENS RESULTAT. Sambandet mellan kunskapsprovet och körprovet för underkända och godkända provtagare. Anna Sundström, Marie Wiberg
- BVM nr 14. DATORBASERADE PROV – egenskaper, möjligheter och begränsningar. Christina Wikström
- BVM nr 15. PRESTATIONSSKILLNADER MELLAN FLICKOR OCH POJKAR I NO. En studie av uppgiftsformatets betydelse i TIMSS 2003. Niklas Eriksson
- BVM nr 16. FLICKOR, POJKAR, FYSIK OCH MATEMATIK. Skillnader i inställning mellan hög- och lågpresterande i TIMSS 1995. Lena Adolffsson
- BVM nr 17. HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2005. Provdeltagargruppens sammansättning och resultat. Christina Stage, Gunilla Ögren

2006

- BVM nr 18. MODELL FÖR BESKRIVNING AV KURSPLANEN FÖR DEN SVENSKA FÖRARUTBILDNINGEN: EN LITTERATURSTUDIE. Tova Stenlund
- BVM nr 19. VIKTNING AV DELPROVEN I HÖGSKOLEPROVET. Christina Stage, Mats Hamrén, Christina Jonsson
- BVM nr 20. TIMSS FIXPUNKTER. En analys av vad elever med olika resultat i TIMSS 2003 vet och kan göra. Peter Nyström
- BVM nr 21. MODELLPRÖVNING. Empirisk prövning av teoretiska modeller för beskrivning av kursplan för förarutbildning. Tova Stenlund
- BVM nr 22. UTVECKLING AV INSTRUMENT FÖR ATT MÄTA BEGREPPET UPPELVD FÖRARKOMPETENS. Anna Sundström
- BVM nr 23. HUR HÖGSKOLEINSTITUTIONER OCH ENHETER KAN GRANSKA SINA KURSER UR JÄMSTÄLLDHETS- OCH GENUSPERSPEKTIV. Nuläget och framåtblick med praktiska förslag. Kent Löfgren
- BVM nr 24. JÄMFÖRELSE AV INNEHÅLL I DEN GAMLA OCH DEN NYA KURSPLANEN FÖR DEN SVENSKA FÖRARUTBILDNINGEN. Tova Stenlund, Widar Henriksson, Anna Sundström
- BVM nr 25. HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2006. Provdeltagargruppens sammansättning och resultat. Christina Stage, Gunilla Ögren

2007

- BVM nr 26. UTVÄRDERING AV INSTRUMENT FÖR UPPLEVD TEORETISK OCH PRAKTISK KOMPETENS. Ett försök med en ny förarprovsmodell. Anna Sundström
- BVM nr 27. BEGREPPET KOMPLETTERING I ETT EXAMINATIONSPERSPEKTIV. Ett försök med det praktiska körkortsprovet. Tova Stenlund, Widar Henriksson, Charlotte Wahl, Bengt Holmberg
- BVM nr 28. NY KURSPLAN OCH NYA FORMER FÖR EXAMINATION. Jämförelse av förarprovet för körkort B före och efter förändringen. Marie Wiberg
- BVM nr 29. SEKVENTIELLA EFFEKTER I HÖGSKOLEPROVET AVSEENDE DELPROVET NOG. Anders Lexelius
- BVM nr 30. SAMSTÄMMIGHET MELLAN MÅL OCH PROV. En studie av det svenska förarutbildningssystemet. Tova Stenlund, Widar Henriksson, Anna Sundström
- BVM nr 31. HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2007. Provdeltagargruppens sammansättning och resultat. Christina Stage, Gunilla Ögren
- BVM nr 32. SVENSKA SKOLANS LÄRARE OCH UNDERVISNING I MATEMATIK OCH NO I ETT INTERNATIONELLT PERSPEKTIV. Några resultat från TIMSS 2003. Susanne Alger

2008

- BVM nr 33. UTVECKLING OCH UTVÄRDERING AV INSTRUMENT INFÖR EN STUDIE AV LIKVÄRDIGHETEN I KÖRPROVET. Susanne Alger, Widar Henriksson, Anna Sundström
- BVM nr 34. HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2008. Provdeltagargruppens sammansättning och resultat. Christina Stage, Gunilla Ögren
- BVM nr 35. URVALSPROV UR ETT SVENSKT OCH INTERNATIONELLT PERSPEKTIV. Christina Wikström

2009

- BVM nr 36. RAMVERK FÖR NATIONELLA ÄMNESPROV I BIOLOGI, FYSIK OCH KEMI I GRUNDSKOLANS ÅRSKURS 9. Peter Nyström, Maria Åström
- BVM nr 37. KUNSKAP, INTRESSE OCH STUDIEFRAMGÅNG. En litteraturgranskning. Ingela Eriksson, Widar Henriksson

- BVM nr 38. ELEVERS SYN PÅ NATIONELLA PROV. Resultat från elevenkät vid genomförandet av nationella ämnesprov i biologi, fysik och kemi i grundskolans årskurs 9 våren 2009. Maria Åström, Hanna Eklöf
- BVM nr 39. LIKVÄRDIGHETEN I KÖRPROVET. En studie av samstämmighet i bedömningen. Susanne Alger, Widar Henriksson, Anna Sundström