

# HÖGSKOLEPROVET VÅREN OCH HÖSTEN 2004

Provdeltagargruppens sammansättning  
och resultat

Christina Stage  
Gunilla Ögren

BVM nr 7, 2004



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 group of test-takers in spring and in autumn 2004 regarding composition and results. Results are presented for test-takers of different gender, ages 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 28 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 er-hålla en viss normerad poäng oavsett prov och provgrupp. Avsikten med föreliggande rapport är att beskriva provdeltagargruppen våren och hösten 2004 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

<b>BAKGRUND</b> .....	<b>1</b>
<b>SYFTE</b> .....	<b>1</b>
<b>RESULTAT</b> .....	<b>2</b>
<i>Provdeltagargruppens sammansättning</i> .....	<b>2</b>
<i>Resultat för olika grupper av provdeltagare</i> .....	<b>5</b>
<b>NORMERING AV HÖGSKOLEPROVET</b> .....	<b>12</b>
<i>Utfall</i> .....	<b>14</b>
Totalpopulationen.....	14
Referenspopulation I .....	14
Referenspopulation II .....	15
Normering baserad på IRT .....	16
<i>Normeringsgränser</i> .....	<b>17</b>
<b>UPPREPAT PROVDELTAGANDE</b> .....	<b>18</b>
<b>AVSLUTNING</b> .....	<b>20</b>
<b>REFERENSER</b> .....	<b>24</b>

## **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 2004 har sammanlagt 57 reguljära provtillfällen anordnats och totalt 1 749 204 prov administrerats. Åren 1992 och 1997 var det störst antal provdeltagare, 145 577 respektive 144 061 personer, där efter har antalet minskat till drygt 69 000 år 2002. År 2004 genomförde ungefär 77 500 personer högskoleprovet, vilket innebär att antalet provdeltagare har ökat med drygt 12 procent jämfört med år 2002 då antalet var som lägst.

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, LÅS 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 utprövningsrutinerna se t.ex. Stage, 1995; Ögren, m.fl., 1995 eller Andersson, 1999.

## **SYFTE**

Syftet med denna redovisning är att beskriva provdeltagargruppen våren och hösten 2004 med avseende på sammansättning och resultat samt att jämföra

med motsvarande resultat på proven år 2003. 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. Fortsättningsvis kommer resultatsammanställningarna att omfatta hela verksamhetsåret, vissa jämförelser med föregående år och i viss mån även från och med året 2000.

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).

## **RESULTAT**

### ***Provdeltagargruppens sammansättning***

Den 3 april 2004 genomfördes högskoleprovet av 42 189 personer och den 23 oktober 2004 av 35 332 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	00 A+B	01 A+B	02A	02B	03A	03B	04A	04B
<b>Kön</b>								
Män	45	45	45	43	45	44	44	45
Kvinnor	55	55	55	57	55	56	56	55
<b>Ålder</b>								
≤ 20 år	47	45	40	42	41	41	39	43
21–24 år	23	24	27	26	28	28	30	29
25–29 år	13	13	13	13	13	14	14	13
30–39 år	13	14	15	15	14	13	13	12
≥ 40 år	4	4	5	4	4	4	4	3
<b>Utbildning</b>								
Grund/real-skola	2	2	2	2	2	2	2	1
Folkhögskola	1	1	1	1	1	-	1	1
Gymnasium ≤ 2 år	11	11	11	10	9	9	5	4
Gymnasium ≥ 3 år	72	71	71	70	71	70	74	75
Högre utb. ≤ 80 p	9	9	10	11	11	12	12	12
Högre utb. ≥ 80 p	4	5	4	5	5	6	5	6
Ej angivet	1	1	1	1	1	1	1	1
<b>Totalt antal</b>	<b>101 247</b>	<b>74 456</b>	<b>40 548</b>	<b>28 589</b>	<b>40 614</b>	<b>31 727</b>	<b>42 189</b>	<b>35 332</b>

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.

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 större på hösten än på våren, med undantag för provet 2003 då andelen 20-åringar eller yngre var densamma vid både vårens och höstens provtillfälle.

Liksom åldersstrukturen har utbildningsbakgrunden förändrats sedan 80-talet. Majoriteten av provdeltagarna har tre- eller fyraårig gymnasieutbildning 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 03A-04B, procentuellt fördelade efter högsta angivna utbildning.

<b>Utbildning</b>	<b>03A</b>	<b>03B</b>	<b>04A</b>	<b>04B</b>
	<b>m/k</b>	<b>m/k</b>	<b>m/k</b>	<b>m/k</b>
Grund/realskola	2/2	2/2	1/2	1/2
Folkhögskola	-/1	1/1	1/1	-/1
Gymnasium $\leq$ 2 år	7/11	6/10	4/6	3/4
Gymnasium $\geq$ 3 år	75/69	73/69	77/73	77/74
Högre utbildning $\leq$ 80 p	10/11	11/12	11/12	12/12
Högre utbildning $\geq$ 80 p	5/5	6/5	5/5	6/6
Ej angivet	1/1	1/1	1/1	1/1

Män har i större utsträckning än kvinnor tre- eller fyraårig gymnasieutbildning medan kvinnor i större utsträckning har tvåårig gymnasieutbildning.

Även avseende ålder föreligger en viss skillnad mellan män och kvinnor eftersom kvinnorna återfinns i större andel i de högsta åldersgrupperna medan männen överväger i de lägsta åldersgrupperna. I tabell 3 framgår hur män respektive kvinnor är procentuellt fördelade på ålder.

**Tabell 3.** Män respektive kvinnor som deltog i högskoleproven 03A-04B, procentuellt fördelade på olika åldersgrupper.

<b>Ålder</b>	<b>03A</b>	<b>03B</b>	<b>04A</b>	<b>04B</b>
	<b>m/k</b>	<b>m/k</b>	<b>m/k</b>	<b>m/k</b>
$\leq$ 20 år	42/39	43/40	41/38	46/41
21–24 år	30/26	29/26	32/28	29/28
25–29 år	14/13	14/14	14/13	13/13
30–39 år	11/16	11/16	10/15	9/14
$\geq$ 40 år	3/6	3/4	3/6	3/4



### **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.*

Som framgår ovan är den beskrivning av provdeltagarna som gjordes 1991 avseende kvinnors och mäns ålder och utbildning giltig även för de provdeltagare som genomförde högskoleprovet år 2004. Kvinnorna har som tidigare något lägre utbildning (se tabell 2), och de förekommer i 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.

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 åtta åren. För åren 1996-1999 anges ett genomsnittligt värde för de fyra åren.

**Tabell 4.** Poängmedelvärde (M) och standardavvikelse (s) på högskoleproven 96A–04B 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	122	68.97	18.37	71.99	66.39	5.60	.30
00B	122	68.32	17.24	70.89	66.28	4.61	.27
01A	122	68.66	17.64	71.23	66.45	4.78	.27
01B	122	67.14	18.07	70.13	64.79	5.36	.30
02A	122	67.72	18.24	70.52	65.42	5.10	.28
02B	122	66.49	17.88	69.36	64.34	5.02	.28
03A	122	66.84	19.14	69.91	64.30	5.61	.29
03B	122	67.97	17.95	71.23	65.43	5.80	.32
04A	122	67.99	18.15	70.38	66.09	4.29	.24
04B	122	68.00	17.87	71.09	65.48	5.61	.31

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 utprovningrutinerna 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. Normeringen innebär att provpoängen överförs till en skala från 0.0 till 2.0 (s.k. normerad poäng), vilket innebär att det är lika lätt eller svårt att erhålla en viss normerad poäng oavsett sammansättningen på provdeltagargruppen även om provets svårighetsgrad skulle fluktuera. 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 03A-04B för hela gruppen samt poängskillnad mellan män och kvinnor.

Delprov	03A			03B			04A			04B		
	M	s	m-k	M	s	m-k	M	s	m-k	M	s	m-k
ORD <sup>1</sup>	22.2	7.4	-.62	21.4	6.6	.12	21.8	7.5	-.12	21.5	7.2	.30
NOG <sup>2</sup>	11.1	4.3	1.86	11.4	4.3	2.11	11.4	4.3	1.65	10.9	4.1	1.68
LÄS <sup>3</sup>	12.0	4.0	.05	11.9	3.9	.45	13.2	3.6	.29	12.5	3.6	.35
DTK <sup>4</sup>	10.1	4.0	1.93	11.2	3.7	2.07	10.4	3.5	1.59	11.1	3.9	2.03
ELF <sup>5</sup>	11.4	4.4	1.15	12.1	4.2	1.05	11.3	4.2	.88	11.9	4.1	1.25

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 2 poäng vardera. Delprovet ELF svarar för drygt 1 poäng av könsskillnaden. För delproven ORD och LÄS är könsskillnaden liten.

En förklaring till könsskillnaden kan vara att män i större utsträckning än kvinnor har tre- eller fyraårig gymnasieutbildning (tabell 2). Det finns ett tämligen starkt samband mellan resultat på högskoleprovet och utbildningsbakgrund. I tabell 6 redovisas poängmedelvärde för provdeltagare med olika bakgrundsutbildning.

<sup>1</sup> ORD 40 uppgifter

<sup>2</sup> NOG 22 uppgifter

<sup>3</sup> LÄS 20 uppgifter

<sup>4</sup> DTK 20 uppgifter

<sup>5</sup> ELF 20 uppgifter

**Tabell 6.** Poängmedelvärde (M) och standardavvikelse (s) på högskoleprovet våren och hösten 2004 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 2003.

Utbildning	04A			04B			m-k	
	M	s	m-k	M	s	m-k	03A	03B
Grund/realskola	67.5	18.6	7.5	67.5	17.9	6.3	7.0	8.6
Folkhögskola	62.5	18.0	3.4	62.8	18.1	6.0	2.9	1.7
Gymn ≤ 2 år	62.3	15.9	5.5	61.8	15.6	5.9	4.8	6.9
Gymn ≥ 3 år	65.3	17.0	4.0	65.1	16.7	5.6	5.6	5.1
Högre utb ≤80p	79.5	16.1	5.2	79.2	15.5	5.9	5.5	5.6
Högre utb ≥80p	88.9	15.6	6.7	87.7	15.2	6.9	7.7	7.4
Totalt	68.0	18.2	4.3	68.0	17.9	5.6	5.6	5.8

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 (75 %). Eftersom tre- eller fyraårig gymnasieutbildning kan ha ganska många olika innehåll finns det anledning att göra ytterligare uppdelningar av denna grupp.

**Tabell 7.** Poängmedelvärde (M) på högskoleprovet våren respektive hösten 2004 för män respektive kvinnor med avslutad treårig gymnasieutbildning uppdelade efter gymnasieinriktning<sup>6</sup>.

**Våren 2004**

	H		S		E		N		T	
	Män	Kv	Män	Kv	M	Kv	M	Kv	M	kv
ORD	23.1	25.5	23.5	22.7	20.7	21.4	23.9	24.3	22.8	25.1
NOG	10.4	10.3	11.3	10.4	11.4	10.3	14.9	13.9	13.7	13.7
LÄS	13.0	14.2	14.0	13.7	13.0	12.8	14.7	14.9	13.9	14.6
DTK	10.1	9.7	10.8	9.5	11.1	9.8	12.9	11.7	12.3	11.9
ELF	11.9	12.5	12.5	11.8	11.3	10.6	13.3	13.0	12.2	12.4
Totalt	68.6	72.2	72.1	68.3	67.5	64.9	79.6	77.8	74.9	77.7

**Hösten 2004**

	H		S		E		N		T	
	M	Kv	M	Kv	M	Kv	M	Kv	M	kv
ORD	22.3	23.9	22.5	21.6	20.2	20.2	23.1	23.2	22.5	24.4
NOG	10.2	9.7	10.7	9.8	10.8	9.7	14.1	12.9	12.8	12.5
LÄS	12.4	13.2	13.2	12.8	12.1	12.0	13.8	13.9	12.8	13.7
DTK	10.7	9.8	11.5	10.0	11.9	10.5	14.1	12.7	13.1	12.5
ELF	12.8	12.8	13.3	12.1	12.0	11.0	14.0	13.3	12.7	12.7
Totalt	68.5	69.3	71.2	66.4	67.0	62.9	79.2	76.0	73.8	75.9

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 de 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

<sup>6</sup> 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, under de första åren av 2000-talet, att kvinnor från humanistisk och teknisk utbildning tenderar att ha högre genomsnittligt resultat än männen från samma utbildningar. Sammanfattningsvis kan man säga att resultatskillnaderna mellan gymnasieutbildningar är betydligt större än skillnaderna mellan kvinnor och män.

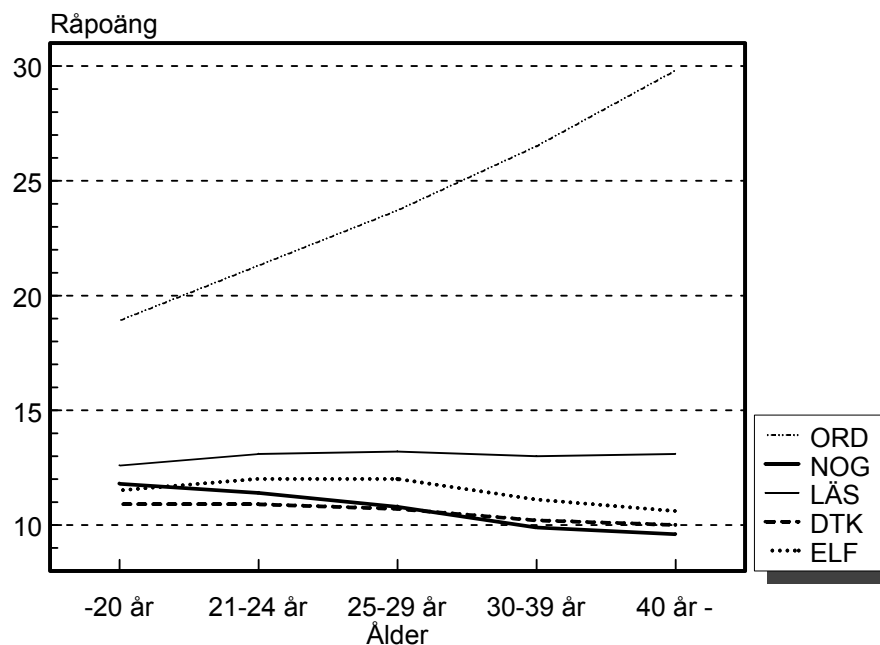
Provdeltagare som ännu ej avslutat sin gymnasieutbildning har genomgående lägre resultat än de som avslutat den.

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 2004 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 2003.

Ålder	04A			04B			m-k (d)	
	M	s	m-k (d)	M	s	m-k (d)	03A	03B
≤ 20 år	65.3	17.1	3.7 (.22)	65.9	16.9	5.6 (.33)	5.7 (.31)	5.3 (.31)
21 – 24 år	68.5	17.9	2.7 (.15)	68.4	17.8	4.4 (.25)	4.3 (.23)	4.3 (.24)
25 – 29 år	70.5	19.1	5.5 (.32)	70.0	18.9	6.1 (.32)	5.9 (.29)	6.6 (.35)
30 – 39 år	70.5	19.1	9.2 (.48)	70.9	19.1	9.9 (.52)	8.8 (.44)	9.9 (.52)
≥ 40 år	72.8	19.0	11.0 (.58)	73.1	18.9	11.0 (.58)	10.4 (.53)	12.3 (.64)

För totalgruppen provdeltagare har det generellt gällt att de yngsta och de äldsta provdeltagarna 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 ungefär fyra procent av provdeltagargruppen. Ungefär 40 procent av provdeltagargruppen är 20 år eller yngre och ungefär 70 procent är under 25 år.



**Figur 1.** Genomsnittligt provresultat på de olika delproven för provdeltagare i olika åldersgrupper år 2004.

Av figuren framgår som redan nämnts att delprovet ORD har ett positivt samband med ålder. För två av delproven – NOG och DTK – har resultaten ett svagt negativt samband med ålder, dvs. ju högre ålder desto lägre genomsnittligt resultat och detta gäller för både män och kvinnor. Även delprovet ELF uppvisar ett svagt negativt samband med ålder. Resultaten på delprovet 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 Riksbanken finansierade VALUTA-projektet<sup>7</sup> har emellertid en databas sammanställts med utförlig information, där förutom testresultat även betyg och social bakgrund ingår.

Preliminära resultat har visat att provdeltagare från högre socialgrupp i genomsnitt har högre resultat på provet än provdeltagare från lägre socialgrupper<sup>8</sup>. 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 de är närmast obefintliga på delproven ORD och LÅS. 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 upprepan- det. 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.

---

<sup>7</sup> VALUTA *Validering av den högre utbildningens antagningssystem* är 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 finansieras av Riksbankens jubileumsfond.

<sup>8</sup> 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 5 123 personer våren 2004 och 3 099 personer hösten 2004.

Referenspopulation I motsvarar också den tidigare 25:4-gruppen, vilket innebär att det är stor spridning inom denna provdeltagargrupp beträffande ålder och utbildning. Ungefär hälften av provdeltagarna i denna grupp är äldre än 30 år och mer än hälften har mindre än tre års gymnasiestudier. I denna grupp är det också avsevärda skillnader mellan män och kvinnor både vad gäller ålder och utbildning; kvinnorna är i genomsnitt något äldre och har något lägre utbildning än männen.

Referenspopulation II utgörs av de provdeltagare som är 18 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 6 814 våren 2004 och 4 211 personer hösten 2004.

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 ämnesområden, kognitiv nivå, innehållskategorier, uppgiftsformat, informationsmängd, 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  $M_{\text{tot}} = 68.0$  för totalgruppen och standardavvikelsen  $s = 18.2$ ; medelvärdet för män var  $M_M = 70.4$  och för kvinnor  $M_{K_V} = 66.1$ . En jämförelse med provet våren 2003 visar att medelvärdet på vårens prov var drygt en poäng högre än på provet våren 2003 medan spridningen var 0.9 lägre (se tabell 4). Jämförelsen visar också att detta prov var en halv poäng lättare för män och nästan två poäng lättare för kvinnor, dvs. kvinnorna stod för en större del av ökningen. Medelvärdesskillnaden mellan män och kvinnor för totalgruppen provdeltagare var 4.3 poäng vid detta prov, den minsta skillnaden någonsin.

Medelvärdet var vid höstens prov provtillfälle  $M_{\text{tot}} = 68.0$  för totalgruppen och standardavvikelse  $s = 17.9$ , medelvärdet för män var  $M_M = 71.1$  och för kvinnor  $M_{K_V} = 65.5$ . En jämförelse med medelvärdet och spridningen på detta prov visar exakt samma som på provet hösten 2003 (se tabell 4). Jämförelsen visar att även resultaten för män och kvinnor är nästan exakt desamma denna höst som föregående höst.

### **Referenspopulation I**

Medelvärdet i referenspopulation I vid vårens provtillfälle var  $M = 67.8$  och standardavvikelsen  $s = 19.1$ ; motsvarande värden för våren 2003 var  $M = 65.4$  och  $s = 19.6$ . För samtliga delprov blev medelvärdet för män i denna grupp  $M_M = 71.0$  och för kvinnor  $M_{K_V} = 65.0$ , motsvarande värden våren 2003 var för män  $M_M = 68.2$  och för kvinnor  $M_{K_V} = 63.0$ . Totalt var medelvärdet på detta prov 2.4 råpoäng högre än på provet våren 2003 för referenspopulation I. En jämförelse med resultaten för män respektive kvinnor visar att detta prov var 2.8 råpoäng lättare för män och två poäng lättare för kvinnor i referenspopulation I. Spridningen på provet våren 2004 var lägre än på provet våren 2003 även för referenspopulation I.

Medelvärdet i referenspopulation I vid höstens provtillfälle var  $M = 67.1$  och standardavvikelsen  $s = 18.9$ ; motsvarande värden hösten 2003 var  $M = 66.5$  och  $s = 18.7$ . För samtliga delprov blev medelvärdet för män i denna grupp  $M_M = 70.3$  och för kvinnor  $M_{K_V} = 64.4$ , motsvarande värden hösten 2003 var för män  $M_M = 70.3$  och för kvinnor  $M_{K_V} = 63.3$ . I jämförelse med provet hösten 2003 var detta prov lika svårt för män, medan det var 1.1 råpoäng lättare för kvinnor i referenspopulation I. Totalt var medelvärdet på höstens prov ungefär en halv poäng högre medan spridningen var ungefär densamma som på provet hösten 2003 i referenspopulation I.

## Referenspopulation II

För referenspopulation II våren 2004 gäller att fördelningen av provdeltagare på olika gymnasiegrenar har ändrats något i jämförelse med våren 2003. Andelen provdeltagare från det naturvetenskapliga programmets naturvetenskapliga inriktning har minskat med 1.5 procentenheter, medan teknisk inriktning har ökat med 1 procentenhet.

För hela referenspopulation II våren 2004 var medelvärdet  $M=66.0$  och standardavvikelsen  $s=16.7$ ; motsvarande värden våren 2003 var  $M=65.5$  och  $s=17.8$ . För män var medelvärdet  $M_M=67.8$  och för kvinnor  $M_{Kv}=64.6$ , motsvarande värden våren 2003 var för män  $M_M=68.6$  och för kvinnor  $M_{Kv}=62.9$ . I jämförelse med provet våren 2003 var detta prov 0.8 råpoäng svårare för män och 1.7 råpoäng lättare för kvinnor. Totalt var således medelvärdet på detta prov en halv poäng högre än provet våren 2003 för referenspopulation II medan spridningen var 1.1 lägre än föregående vår.

För referenspopulation II hösten 2004 gäller att fördelningen av provdeltagare på olika gymnasiegrenar har ändrats något i jämförelse med hösten 2003. Andelen provdeltagare från det naturvetenskapliga programmets har minskat med drygt 1 procentenhet, medan det samhällsvetenskapliga programmet har ökat med drygt 1 procentenhet.

För hela referenspopulation II hösten 2004 var medelvärdet  $M=65.8$  och standardavvikelsen  $s=16.3$ ; motsvarande värden hösten 2003 var  $M=65.5$  och  $s=16.3$ . För män var medelvärdet  $M_M=68.9$  och för kvinnor  $M_{Kv}=63.4$ , motsvarande värden för denna referensgrupp hösten 2003 var för män  $M_M=69.1$  och för kvinnor  $M_{Kv}=63.0$ . I jämförelse med provet hösten 2003 var provet hösten 2004 något svårare för män medan det var något lättare för kvinnor. Totalt var således medelvärdet på detta prov 0.3 råpoäng högre än provet hösten 2003 för referenspopulation II. Spridningen var densamma som föregående höst.

Sammanfattningsvis kan sägas utifrån resultaten för både totalgruppen och de två referensgrupperna att medelvärdet på provet våren 2004 ligger högre än medelvärdet på provet våren 2003. Provet hösten 2004 ligger däremot ungefär på samma nivå som provet hösten 2003. För utförligare beskrivning av resultaten se Stage, Hamrén & Jonsson, 2004 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 utprövningsmodellen.
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 2004 bedömdes provet hösten 2001 vara lämpligast som referensprov eftersom det var det provtillfälle där flest antal uppgifter till provet våren 2004 hade prövats ut. De estimerade resultaten var att den population som genomförde provet hösten 2001 hade medelvärdet 67.5 på det provet och skulle ha haft 68.1 på provet våren 2004. De estimerade medelvärdena för den population som genomförde provet våren 2004 var 66.6 på provet hösten 2001 och 66.9 på provet våren 2004. Enligt IRT är därmed provet våren 2004 ungefär en halv råpoäng lättare än provet hösten 2001. Normeringsutfallet blev detsamma oberoende av vilken population som användes.

För provtillfället hösten 2004 bedömdes provet hösten 2002 vara lämpligast som referensprov. De estimerade resultaten var att den population som genomförde provet hösten 2002 hade medelvärdet 65.2 på det provet och skulle ha haft 65.5 på provet hösten 2004. De estimerade medelvärdena för den population som genomförde provet hösten 2004 var 67.0 på provet hösten 2002 och 67.3 på provet hösten 2004. Enligt IRT är därmed provet hösten 2004 mindre än en halv råpoäng lättare än provet hösten 2002. Populationen provdeltagare

hösten 2004 estimeras samtidigt vara något (drygt en råpoäng) bättre än provdeltagarpopulationen hösten 2002. Normeringsutfallet blev detsamma oberoende av vilken population som användes.

### **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 våren 2004 blev att den normerade poängen 1.0 erhålls i intervallet 68-71 råpoäng, vilket är en råpoäng högre än våren 2003, men densamma som hösten 2003. För den normerade poängen 2.0 krävs 110 råpoäng eller mer, vilket också är en råpoäng högre än våren 2003. Den fullständiga normeringstabellen återfinns i bilaga 3a.

Resultatet av den genomförda normeringen av högskoleprovet hösten 2004 blev att den normerade poängen 1.0 erhålls i intervallet 68-71 råpoäng, vilket är detsamma som hösten 2003 och våren 2004. För den normerade poängen 2.0 krävs 109 råpoäng eller mer, vilket är detsamma som hösten 2003, men en råpoäng lägre än våren 2004. Den fullständiga normeringstabellen återfinns i bilaga 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 40 procent (något högre på våren och något lägre på hösten) av provdeltagarna ha genomfört minst ett tidigare prov. Vid de senaste sju provtillfällena har dock denna andel minskat till cirka 30 procent.

Ett flertal studier har genomförts av effekterna av upprepat provdeltagande (Henriksson, 1991, 1993, 1995; Henriksson & Bränberg, 1992, 1994). I sin senaste studie jämförde Henriksson (1995) resultaten avseende upprepat provdeltagande för proven hösten 1986, hösten 1991 och hösten 1993. De skillnader i normerad poäng som rapporterades mellan det första och andra provtillfället var 0.055, 0.083 respektive 0.111. 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.

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 13-14 råpoäng, medan ökningen från första till fjärde provtillfället för samma individer var i genomsnitt drygt sju råpoäng och skillnaden mellan resultatet vid fjärde provtillfället och bästa tidigare resultat var obefintlig. 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 sju råpoäng.

I tabell 9 visas resultat för de provdeltagare hösten 2004 som genomförde provet för första, andra, tredje respektive fjärde gången.

**Tabell 9.** Normerad poäng för provdeltagare hösten 2004 som genomfört provet vid ett till fyra tillfällena under åren 2003 och 2004.

Antal prov	N	Provtillfälle			
		04B	04A	03B	03A
1	23 976	0.9			
2	4 724	1.0	0.9		
	1 617	1.0		0.9	
	1 324	1.0			0.9
3	1 582	1.1	1.1	1.0	
	619	1.1	1.1		1.0
	531	1.1		1.1	1.0
4	862	1.2	1.2	1.2	1.1

Av tabell 9 framgår att de provdeltagare som under den aktuella tvåårsperioden genomfört samtliga fyra prov i genomsnitt hade 1.1 i normerad poäng vid det första provtillfället. De som genomfört tre prov hade vid det första provtillfället den normerade poängen 1.0 och de som genomfört ett eller två prov hade vid det första provtillfället den normerade poängen 0.9. För samtliga grupper har en genomsnittlig ökning av resultatet med 0.1 normerade poäng skett från det första till det andra provtillfället. I tabell 10 visas andelen provdeltagare vid varje provtillfälle under 2000-talet som genomfört ett, två, tre respektive fyra prov.

**Tabell 10.** Andel provdeltagare som vid provtillfällena 00A-04B genomförde provet för första, andra, tredje respektive fjärde gången.

	Antal genomförda prov			
	1	2	3	4
00A	59	28	10	3
00B	63	25	9	3
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

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.

Männen är något mer benägna att upprepa sitt provdeltagande, med undantag för denna vår då skillnaden mellan andelen kvinnor och män som upprepat sitt provdeltagande var mycket liten. Den största andelen av dem som upprepar sitt provdeltagande återfinns i de två yngsta åldersgrupperna (24 år och yngre).

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 medvärdesresultat som redovisats här. En provdeltagare kan dock aldrig förlora i resultat genom att upprepa provdeltagandet eftersom det alltid är det bästa resultatet som gäller.

## AVSLUTNING

Antalet deltagare i det högskoleprov som genomfördes den 3 april år 2004 var 42 189, av vilka 44 procent var män och 56 procent var kvinnor och den 23 oktober 2004 genomfördes provet av 35 332 personer, av vilka 45 procent var män och 55 procent var kvinnor. Sammanlagt under året var det 77 521 som genomförde provet. År 2002 var det år då det var det lägsta antalet



provdeltagare sedan de nya antagningsreglerna trädde i kraft 1991. Jämfört med åren 1992 och 1997 som hade flest antal provdeltagare hade det skett en minskning med drygt 50 procent. Antalet provdeltagare har dock ökat de två senaste åren med totalt 12 procent jämfört med år 2002.

Sedan i början på 1990-talet har majoriteten av provdeltagarna varit ganska unga. Ungefär 70 procent av provdeltagarna år 2004 är 24 år eller yngre. En större andel män återfinns i de lägsta åldersgrupperna medan en större andel kvinnor återfinns i de högsta åldersgrupperna, framför allt i åldersgruppen 30-39 år.

Majoriteten av provdeltagarna har tre- eller fyraårig gymnasieutbildning och män har något högre utbildning än kvinnor.

De resultat som presenterats för högskoleprovet visar att medelvärdet på provet våren 2004 var drygt en råpoäng högre än på det prov som genomfördes våren 2003, medan medelvärdet på provet hösten 2004 var detsamma som på provet hösten 2003. Medelvärdet på provet våren 2004 var 67.99 råpoäng och på hösten 2004 var medelvärdet 68.00 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 fyra procent av hela provdeltagargruppen.

Resultatskillnaden mellan totalgrupperna män och kvinnor var ungefär som vid de senaste provtillfällena. Förklaringen till den minskade könsskillnaden i resultat, jämfört med före våren 1996, är troligen förändringen av sammansättningen av provet; utökningen av delprovet ORD med 10 uppgifter och borttagandet av delprovet AO (allmänorientering). Delprovet ORD har alltid uppvisat små könsskillnader medan delprovet AO uppvisade skillnader upp till 2 råpoäng. Könsskillnaden för de övriga delproven är som den varit vid

tidigare tillfällen. Delproven DTK och NOG svarar för vardera ungefär 2 råpoäng, delprovet ELF för drygt 1 råpoäng till männens fördel medan för delprovet LÄS är könsskillnaden liten.

För provet som helhet och för den totala gruppen provdeltagare är skillnaden mellan män och kvinnor 0.24 standardavvikelseenheter till männens fördel för våren 2004 och 0.31 för hösten 2004. 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 inom utbildningsområdet och det har förekommit intensiv forskning med syfte att förklara och förstå dessa skillnader. Några entydiga svar finns ej 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 uppstår<sup>9</sup>. 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

---

<sup>9</sup> 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)

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 och 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).

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

- Andersson, K. (Red.). (1999). *Högskoleprovet. Konstruktion, resultat och erfarenheter* (Pm nr 153). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- 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, p. 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, p. 161-173.
- Cook, L.L. & Eignor, D.R. (1991). IRT Equating Methods. *Educational Measurement: Issues and Practices*, p. 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: University of Umeå, 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. p 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: University of Umeå, 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: University of Umeå, 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: University of Umeå, 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. (2004a). *Normering av högskoleprovet våren 2004* (Arbetsrapport nr 5). Umeå: Umeå universitet, Institutionen för beteendevetenskapliga mätningar.
- Stage, C., Hamrén, M. & Jonsson, C. (2004b). *Normering av högskoleprovet hösten 2004* (Arbetsrapport nr 6). 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.
- Wedman, I. & Stage, C. (1983). The Significance of Contents for Sex Differences in Test Results. *Scandinavian Journal of Educational Research*, 27, p. 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 Patricks College, Dublin, Ireland.
- Wilder, G., & Powell, K. (1989). Sex Differences in Test Performance: A Survey of the Litterature. *College Board Report*, No 89-3.
- 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.

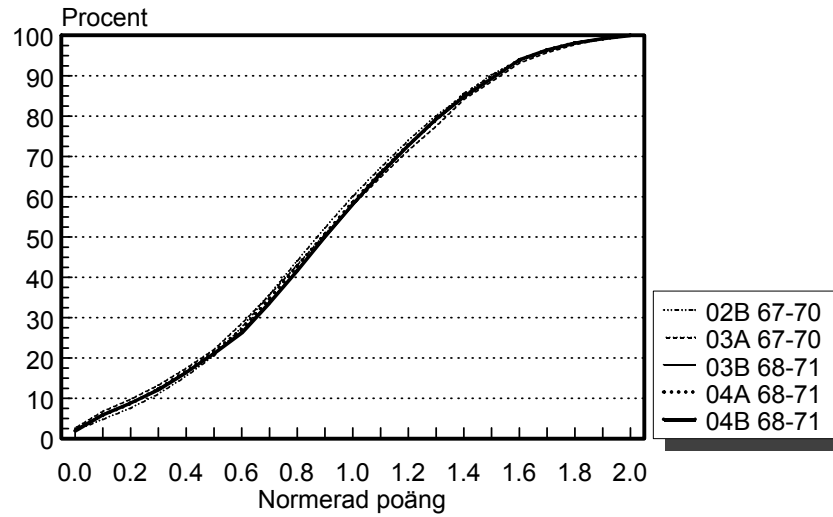






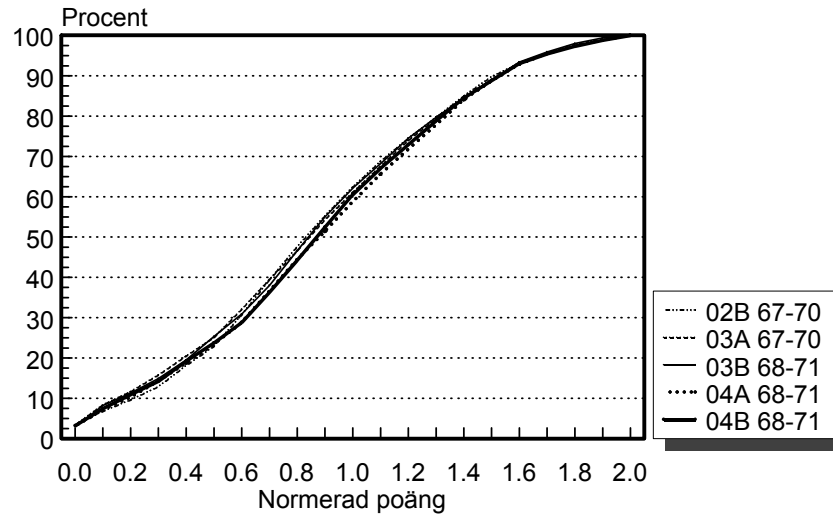


## NORMERINGSUTFALL TOTALT



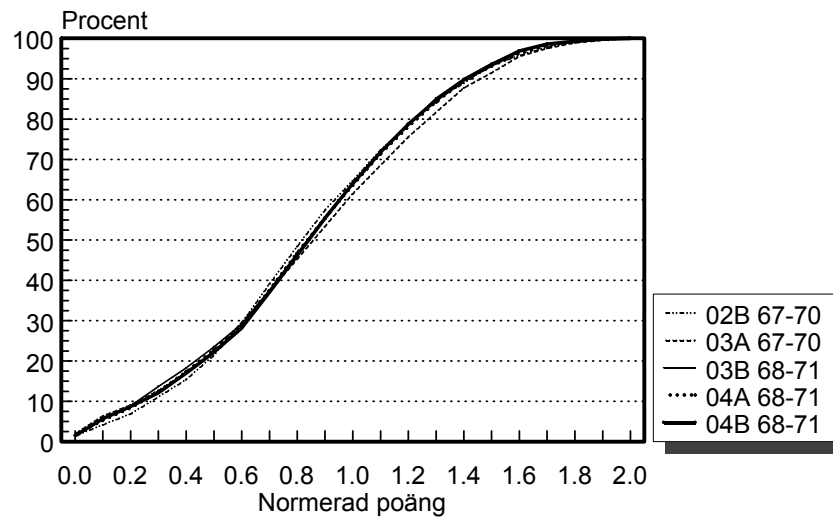
**Figur.** Den kumulativa fördelningen av normerad poäng för totalpopulationen, proven 02B–04B.

## NORMERINGSUTFALL REF.POP. 1



**Figur.** Den kumulativa fördelningen av normerad poäng för referenspopulation I, proven 02B–04B.

## NORMERINGSUTFALL REF.POP. 2



**Figur.** Den kumulativa fördelningen av normerad poäng för referenspopulation II, proven 02B–04B.

NORMERINGSTABELL GÄLLANDE HÖGSKOLEPROVET 2004-04-03 (04: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- 34	0.0	877	2.1	2.1
35- 40	0.1	1539	3.6	5.7
41- 43	0.2	1224	2.9	8.6
44- 46	0.3	1527	3.6	12.2
47- 49	0.4	1843	4.4	16.6
50- 52	0.5	2028	4.8	21.4
53- 55	0.6	2343	5.6	27.0
56- 59	0.7	3229	7.7	34.6
60- 63	0.8	3424	8.1	42.7
64- 67	0.9	3294	7.8	50.6
68- 71	1.0	3346	7.9	58.5
72- 75	1.1	3131	7.4	65.9
76- 79	1.2	2849	6.8	72.7
80- 83	1.3	2627	6.2	78.9
84- 88	1.4	2713	6.4	85.3
89- 92	1.5	1860	4.4	89.7
93- 97	1.6	1719	4.1	93.8
98-101	1.7	1072	2.5	96.3
102-105	1.8	733	1.7	98.1
106-109	1.9	448	1.1	99.1
110-122	2.0	363	.9	100.0
		----- 42189	----- 100.0	

MEDELVÄRDE = 0.9400  
 STANDARDAVVIKELSE = .4652

NORMERINGSTABELL GÄLLANDE HÖGSKOLEPROVET 2004-10-23 (04: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	665	1.9	1.9
34- 40	0.1	1405	4.0	5.9
41- 43	0.2	1040	2.9	8.8
44- 46	0.3	1163	3.3	12.1
47- 49	0.4	1504	4.3	16.4
50- 52	0.5	1683	4.8	21.1
53- 55	0.6	1811	5.1	26.2
56- 59	0.7	2596	7.3	33.6
60- 63	0.8	2829	8.0	41.6
64- 67	0.9	2993	8.5	50.1
68- 71	1.0	2856	8.1	58.1
72- 75	1.1	2689	7.6	65.8
76- 79	1.2	2463	7.0	72.7
80- 83	1.3	2307	6.5	79.3
84- 87	1.4	1946	5.5	84.8
88- 91	1.5	1607	4.5	89.3
92- 96	1.6	1606	4.5	93.9
97-100	1.7	890	2.5	96.4
101-104	1.8	607	1.7	98.1
105-108	1.9	390	1.1	99.2
109-122	2.0	282	.8	100.0
		----- 35332	----- 100.0	

MEDELVÄRDE = 0.9449  
 STANDARDAVVIKELSE = .4640