

SAMMANFATTNING

Högskoleprovet har nu i 25 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 2002 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.

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 har antalet provdeltagare årligen varit 130 000–145 000 (fortfarande två femtedelar på hösten och tre femtedelar på våren). Till och med hösten år 2002 har sammanlagt 53 reguljära provtillfällen anordnats och totalt 1 599 342 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.

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

SYFTE

Syftet med denna redovisning är att beskriva provdeltagargruppen våren och hösten 2002 med avseende på sammansättning och resultat samt att jämföra med motsvarande resultat på proven år 2001. 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 poäng till fördel för män. Efter förändringen våren 1996 har skillnaderna minskat till i genomsnitt 5-6 poäng (se t.ex. Stage & Ögren, 2001).

RESULTAT

Provdeltagargruppens sammansättning

Antalet provdeltagare i det högskoleprov som genomfördes 6 april 2002 var 40 548, av vilka 18 288 var män och 22 260 var kvinnor. Sammanlagt 27 240 var yngre än 25 år och av dessa var 16 194 personer 20 år eller yngre.

Antalet provdeltagare i det högskoleprov som genomfördes 26 oktober 2002 var 28 589, av vilka 12 222 var män och 16 367 var kvinnor. Sammanlagt 19 373 var yngre än 25 år och av dessa var 12 010 personer 20 år eller yngre.

För att möjliggöra en jämförelse mellan de två senaste proven och föregående års två 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	00A	00B	01A	01B	02A	02B
Kön						
Män	46	44	46	44	45	43
Kvinnor	54	56	54	56	55	57
Ålder						
= 20 år	43	51	43	46	40	42
21–24 år	25	20	25	23	27	26
25–29 år	14	12	13	13	13	13
30–39 år	14	13	14	14	15	15
= 40 år	4	4	5	4	5	4
Utbildning						
Grund/real-skola	1	2	2	2	2	2
Folkhögskola	1	1	1	1	1	1
Gymnasium = 2 år	12	11	12	11	11	10
Gymnasium = 3 år	70	73	71	71	71	70
Högre utb. = 80 p	10	8	9	9	10	11
Högre utb. = 80 p	5	4	4	5	4	5
Ej angivet	1	1	1	1	1	1
Totalt antal	62 717	38 530	46 602	27 854	40 548	28 589

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.

Liksom åldersstrukturen har utbildningsbakgrund 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. Procentuell fördelning på högsta angivna utbildning för deltagarna i högskoleproven 01A–02B uppdelat på män och kvinnor.

Utbildning	01A	01B	02A	02B
	m/k	m/k	m/k	m/k
Grund/realskola	1/2	2/2	2/2	2/2
Folkhögskola	1/1	-/-	-/1	1/1
Gymnasium = 2 år	8/15	8/14	8/14	8/13
Gymnasium = 3 år	76/66	76/68	75/67	73/67
Högre utbildning = 80 p	9/10	9/10	10/11	10/11
Högre utbildning = 80 p	4/5	4/5	4/5	6/5
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. Procentuell fördelning på olika åldersgrupper för provdeltagarna i högskoleproven 01A–02B uppdelat på män och kvinnor.

Ålder	01A	01B	02A	02B
	M/k	m/k	m/k	m/k
= 20 år	47/40	49/43	43/38	44/41
21–24 år	27/23	25/21	30/25	28/24
25–29 år	13/13	13/13	13/13	14/13
30–39 år	10/18	10/18	11/18	11/17
= 40 år	3/6	2/5	3/6	3/5

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 prov-

tagargruppen 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 2002. 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 tre åren.

Tabell 4. Poängmedelvärde (M) och standardavvikelse (s) på högskoleproven 96A–02B totalt, för män respektive kvinnor samt skillnad mellan dem 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

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 provtillfällena. 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. Närmare beskrivning av normeringsförfarandet se sid 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 01A-02B för hela gruppen samt poängskillnad mellan män och kvinnor.

Delprov	01A			01B			02A			02B		
	M	s	m-k	M	s	m-k	M	s	m-k	M	s	m-k
ORD ¹	22.0	7.1	-.48	21.7	7.4	.19	21.8	7.2	-.59	21.6	7.5	.28
NOG ²	11.1	4.1	1.95	11.2	4.2	2.04	11.2	4.3	2.13	10.5	4.0	1.60
LÄS ³	12.9	3.6	.05	11.7	3.6	.16	12.0	3.9	.09	12.5	3.7	.28
DTK ⁴	11.3	3.8	1.98	11.0	3.7	1.67	11.4	3.6	2.07	10.4	3.5	1.67
ELF ⁵	11.3	4.4	1.29	11.5	4.2	1.30	11.5	4.4	1.39	11.5	4.3	1.18

Skillnaderna i resultat mellan män och kvinnor har alltid diskuterats mycket och utgångsläget har varit skillnaden i poängmedelvärde. 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 nästan 1.5 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 för hela gruppen och skillnaden mellan män och kvinnor.

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 2002 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 2001.

Utbildning	02A			02B			m-k	
	M	s	m-k	M	s	m-k	01A	01B
Grund/realskola	66.5	17.4	3.0	64.9	17.6	6.9	5.5	5.2
Folkhögskola	62.7	18.2	0.8	60.6	19.1	1.9	3.5	4.1
Gymn= 2 år	62.1	16.3	4.5	60.6	15.3	5.0	4.7	4.6
Gymn = 3 år	65.9	17.5	5.3	64.3	16.9	4.6	4.8	5.5
Högre utb =80p	79.7	15.8	6.1	77.4	16.0	6.0	5.5	5.8
Högre utb =80p	88.1	15.1	6.7	87.1	15.8	7.2	7.6	7.3
Totalt	67.7	18.2	5.1	66.5	17.9	5.0	4.8	5.4

Tabell 6 visar att sambandet mellan utbildning och resultat på högskoleprovet är tämligen starkt, d.v.s. 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 70 %). 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 2002 totalt och poängskillnaden mellan män och kvinnor med avslutad tre-årig gymnasieutbildning uppdelade efter gymnasieinriktning.

Våren 2002

Gren Kön	H		S		E		N		T	
	M	m-k	M	m-k	M	m-k	M	m-k	M	m-k
ORD	25.6	-1.95	22.5	.15	21.7	-1.40	24.1	-.38	23.5	-2.35
NOG	10.1	.76	10.7	1.25	11.0	1.16	14.0	1.39	13.9	.75
LÄS	13.3	-.62	12.6	.12	11.9	-.20	13.8	-.42	13.1	-.87
DTK	10.6	1.05	11.0	1.39	11.5	1.38	13.7	1.43	13.8	.75
ELF	12.8	.03	12.2	1.11	11.4	1.06	13.9	.65	13.2	.26
Totalt	72.5	-.72	69.1	4.02	67.6	2.17	79.2	2.55	77.5	-1.44

Hösten 2002

Gren Kön	H		S		E		N		T	
	M	m-k	M	m-k	M	m-k	M	m-k	M	m-k
ORD	24.4	-2.35	21.8	.93	21.0	-.38	23.4	1.01	23.7	-1.37
NOG	9.6	.01	9.9	.91	10.2	1.20	12.9	1.35	12.9	.83
LÄS	13.3	-1.04	12.9	.15	12.2	.16	14.1	.03	13.6	-.47
DTK	9.6	-.03	10.1	1.05	10.4	1.34	12.4	1.49	12.7	.95
ELF	12.8	-.30	12.1	.94	11.3	.76	13.6	.87	12.9	.26
Totalt	69.8	-3.70	66.8	3.96	65.1	3.07	76.5	4.74	75.8	.20

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.

Vad gäller skillnaderna mellan män och kvinnor kan konstateras att även inom gymnasiegren har kvinnorna genomgående något lägre resultat än männen. Kvinnor med naturvetenskaplig och teknisk gymnasieinriktning har dock klart högre resultat än män från övriga grenar.

Sammanfattningsvis kan man konstatera att resultatskillnaderna mellan gymnasieutbildningar är betydligt större än skillnaderna mellan kvinnor och män.

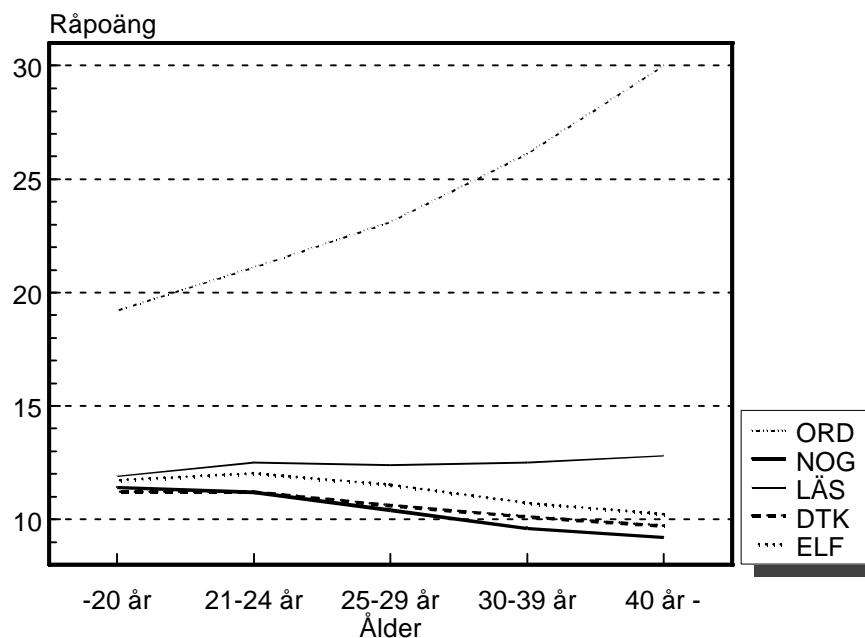
Provdeltagare som ännu ej avslutat sin gymnasieutbildning har genomgående lägre resultat än de som avslutat den.

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. Denna information har alltid insamlats i samband med provgenomförandet och har därmed varit lättillgänglig. 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 2002 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 2001.

Ålder	02A			02B			m-k (d)	
	M	S	m-k (d)	M	S	m-k (d)	01A	01B
= 20 år	65.9	17.6	5.8 (.33)	64.7	17.0	4.7 (.28)	4.9 (.28)	6.0 (.34)
21 – 24 år	68.4	18.1	4.1 (.23)	67.0	17.8	3.9 (.22)	3.7 (.21)	3.0 (.16)
25 – 29 år	68.6	19.1	5.1 (.27)	67.1	18.6	5.1 (.27)	5.3 (.29)	6.0 (.31)
30 – 39 år	69.1	18.7	6.5 (.35)	68.6	18.9	9.1 (.48)	6.9 (.39)	8.2 (.45)
= 40 år	71.5	18.8	10.1 (.54)	71.9	18.6	9.6 (.52)	9.4 (.52)	7.8 (.43)

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 ORD-provet utökades med 10 uppgifter och ORD-provet har visat sig ha ett positivt samband med ålder, dvs. ju högre ålder desto högre genomsnittligt resultat på ORD-provet. Man bör dock komma ihåg att detta är en mycket liten grupp provdeltagare, endast ungefär fyra procent av provdeltagargruppen. Ungefär 50 procent av provdeltagargruppen är 21 år eller yngre.



Figur 1. Genomsnittligt provresultat på de olika delproven för provdeltagare i olika åldersgrupper år 2002.

Av figuren framgår som redan nämnts att delprovet ORD har ett positivt samband med ålder. För två av delproven – NOG och DTK – har resultaten 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 uppföljningar kunnat genomföras på det sätt som har gjorts avseende kön, utbildning och ålder. I det av Riksbanken finansierade VALUTA-projektet⁶ har emellertid en databas sammanställts med utförlig information, där förutom testresultat även

⁶ VALUTA *Validering av den högre utbildningens antagningssystem* är ett samarbetsprojekt mellan Institutionen för pedagogik och didaktik vid Göteborgs universitet och Enheten för pedagogiska mätningar vid Umeå universitet. Projektet finansieras av Riksbankens jubileumsfond.

betyg och social bakgrund ingår. Detta möjliggör att jämförande studier kan genomföras även av resultat för olika sociala grupper.

NORMERING AV HÖGSKOLEPROVET

Alltsedan provet 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 grannliga uppgift, eftersom den normerade poängen ska vara jämförbar för tio provtillfällen.

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 möjligheten för ett visst normerat provresultat studeras även utfallet i två referensgrupper.

Referenspopulation I väljs ut 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 provtillfällen för denna referenspopulation kan relateras till provet. Denna referenspopulation bestod av 5 730 personer våren 2002 och 3 409 personer hösten 2002.

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, ändå är könsskillnaderna klart större i denna grupp. Denna referenspopulation bestod av 7 651 personer våren 2002 och 4 018 personer hösten 2002.

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 är på en skala från 0.0 till 2.0, där 2.0 är det högsta resultatet.

Utfall

Totalpopulationen

Medelvärdet vid vårens provtillfälle var $M_{tot} = 67.7$ för totalgruppen och standardavvikelsen var $s = 18.2$; medelvärdet för män var $M_M = 70.5$ och för kvinnor $M_{Kv} = 65.4$. Motsvarande värden våren 2001 var $M_{tot} = 68.7$ och $s = 17.6$. Medelvärdet för män var $M_M = 71.2$ och för kvinnor $M_{Kv} = 66.4$. Totalt var således medelvärdet en poäng lägre på vårens prov än på provet våren 2001 medan spridningen var 0.6 högre. En jämförelse med resultaten för män respektive kvinnor visar att detta prov var drygt en halv poäng svårare för män och exakt en poäng svårare för kvinnor, dvs. kvinnorna stod för en något större del av sänkningen. Medelvärdesskillnaden mellan män och kvinnor för totalgruppen provdeltagare var 5.1 poäng.

Medelvärdet vid höstens provtillfälle var $M_{tot} = 66.5$ för totalgruppen och standardavvikelsen var $s = 17.9$, medelvärdet för män var $M_M = 69.4$ och för kvinnor $M_{Kv} = 64.3$. Motsvarande värden hösten 2001 var $M_{tot} = 67.1$ och $s = 18.1$. Medelvärdet för män var $M_M = 70.1$ och för kvinnor $M_{Kv} = 64.8$. Totalt var således medelvärdet drygt en halv poäng lägre på detta prov än på provet hösten 2001. En jämförelse med resultaten för män respektive kvinnor visar att männen stod för en något större del av sänkningen. Medelvärdesskillnaden mellan män och kvinnor för totalgruppen provdeltagare var 5 poäng.

Referenspopulation I

Medelvärdet i referenspopulation I var vid vårens provtillfälle $M = 67.1$ och standardavvikelsen $s = 18.9$; motsvarande värden för våren 2001 var $M = 68.3$

och $s=17.9$. För samtliga delprov blev medelvärdet för män i denna grupp $M_M=69.5$ och för kvinnor $M_{KV}=65.1$, motsvarande värden våren 2001 var för män $M_M=70.4$ och för kvinnor $M_{KV}=66.3$. Totalt var medelvärdet på detta prov drygt en poäng lägre än på provet våren 2001 för referenspopulation I. En jämförelse med resultaten för män respektive kvinnor visar att detta prov var nästan en poäng svårare för män och drygt en poäng svårare för kvinnor i referenspopulation I. Spridningen på provet våren 2002 var högre än på provet våren 2001 även för referenspopulation I.

Medelvärdet i referenspopulation I var vid höstens provtillfälle $M=65.5$ och standardavvikelsen $s=18.8$; motsvarande värden hösten 2001 var $M=66.8$ och $s=18.6$. För samtliga delprov blev medelvärdet för män i denna grupp $M_M=67.9$ och för kvinnor $M_{KV}=63.4$, motsvarande värden hösten 2001 var för män $M_M=69.4$ och för kvinnor $M_{KV}=64.5$. I jämförelse med provet hösten 2001 var detta prov således en och en halv poäng svårare för män och lite drygt en poäng svårare för kvinnor i referenspopulation I. Totalt var medelvärdet på höstens prov drygt en poäng lägre medan spridningen var något högre än på provet hösten 2001 i referenspopulation I.

Tabell 9. Poängmedelvärde på de olika delproven för män respektive kvinnor och totalt för referenspopulation I för provtillfällena 96B–02B.

	ORD		NOG		LÄS		DTK		ELF		Tot	
	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k
96B	26.3	-0.3	12.3	1.9	12.9	0.3	11.1	2.2	12.2	1.5	74.8	5.6
97A	24.5	0.0	11.7	2.2	12.4	0.3	9.5	2.0	10.8	0.9	68.8	5.4
97B	25.0	-0.8	11.6	2.1	12.3	0.3	10.2	2.0	12.9	0.9	72.0	4.5
98A	25.7	-1.5	11.1	2.6	11.4	0.5	10.9	1.9	12.7	1.0	71.6	4.5
98B	25.1	-1.1	10.2	2.4	10.5	0.3	9.6	2.2	10.7	0.7	66.2	4.5
99A	25.1	-0.9	8.7	1.7	12.0	0.0	10.8	2.0	11.5	1.2	68.1	4.1
99B	24.4	-0.9	10.1	1.9	11.2	0.4	9.7	1.5	11.6	1.1	67.0	4.0
00A	24.3	-1.1	10.6	2.0	11.5	0.2	10.8	2.0	11.0	1.3	68.3	4.3
00B	24.9	-0.8	10.4	1.4	12.2	0.1	9.7	1.7	10.9	1.5	68.2	3.9
01A	24.6	-0.7	9.8	1.6	12.6	0.0	10.6	1.9	10.5	1.2	68.3	4.1
01B	24.4	-0.1	10.0	1.9	11.5	0.2	10.3	1.6	10.7	1.3	66.8	4.9
02A	24.2	-0.9	9.9	1.9	11.9	0.0	10.7	2.1	10.5	1.2	67.1	4.3
02B	24.0	0.0	9.4	1.3	12.0	0.3	9.4	1.7	10.7	1.2	65.5	4.5

Referenspopulation II

För referenspopulation II våren 2002 gäller att fördelningen av provdeltagare på olika gymnasiegrenar är ungefär densamma som våren 2001.

För hela referenspopulation II var medelvärdet $M=66.4$ och standardavvikelsen $s=17.0$; motsvarande värden våren 2001 var $M=67.7$ och $s=17.0$. För män var medelvärdet $M_M=69.6$ och för kvinnor $M_{Kv}=63.6$, motsvarande värden våren 2001 var för män $M_M=70.4$ och för kvinnor $M_{Kv}=65.1$. I jämförelse med provet våren 2001 var detta prov nästan en poäng svårare för män och en och en halv poäng svårare för kvinnor. Totalt var således medelvärdet på detta prov drygt en poäng lägre än provet våren 2001 för referenspopulation II medan spridningen var densamma som föregående vår.

För referenspopulation II hösten 2002 gäller att fördelningen av provdeltagare på olika gymnasiegrenar är ungefär densamma som hösten 2001.

För hela referenspopulation II var medelvärdet $M=64.7$ och standardavvikelsen $s=16.6$; motsvarande värden hösten 2001 var $M=65.8$ och $s=16.8$. För män var medelvärdet $M_M=67.9$ och för kvinnor $M_{Kv}=62.5$, motsvarande värden för denna referensgrupp hösten 2001 var för män $M_M=69.1$ och för kvinnor $M_{Kv}=63.1$. I jämförelse med provet hösten 2001 var provet hösten 2002 således drygt en poäng svårare för män och ungefär en halv poäng svårare för kvinnor. Totalt var således medelvärdet på detta prov drygt en poäng lägre än provet våren 2001 för referenspopulation II. Spridningen var i stort sett densamma som föregående vår.

Tabell 10. Poängmedelvärde på de olika delproven för män respektive kvinnor och totalt för referenspopulation II för provtillfällena 96B–02B.

	ORD		NOG		LÄS		DTK		ELF		Tot	
	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k	Tot	m-k
96B	22.3	0.2	14.1	2.4	13.2	0.1	12.4	2.1	12.9	1.3	74.9	6.1
97A	22.3	0.1	13.9	2.4	12.8	0.2	11.0	2.1	11.9	0.8	71.9	5.6
97B	20.0	0.0	13.8	2.0	12.4	0.1	11.8	1.8	13.3	0.8	71.3	4.7
98A	20.9	-0.6	13.2	2.5	11.4	0.1	12.3	1.8	13.9	0.9	71.7	4.7
98B	20.6	-0.1	12.5	2.5	10.7	0.5	10.9	2.0	11.7	0.7	66.5	5.6
99A	20.3	0.3	10.7	2.2	12.2	0.2	12.2	2.2	13.0	1.1	68.4	5.9
99B	20.0	0.3	11.9	2.0	10.9	0.1	10.9	1.6	12.6	1.3	66.4	5.3
00A	19.4	0.0	12.6	2.5	11.9	0.1	12.0	2.2	12.1	1.3	68.1	6.0
00B	19.2	0.5	12.8	1.6	12.3	0.0	11.2	1.6	11.4	1.5	66.9	5.2
01A	19.7	0.2	12.0	1.8	12.8	0.0	11.7	1.8	11.5	1.2	67.7	5.3
01B	19.1	1.1	12.1	1.9	11.7	0.2	11.5	1.5	11.4	1.4	65.8	6.0
02A	19.3	0.3	12.0	2.1	11.6	0.1	11.8	1.9	11.7	1.5	66.5	5.9
02B	19.0	0.7	11.0	1.6	12.3	0.3	10.8	1.5	11.5	1.3	64.7	5.4

Sammanfattningsvis kan sägas utifrån resultaten för både totalgruppen och de två referensgrupperna att provet våren 2002 tycks vara ungefär en poäng svårare än provet våren 2001. Även provet hösten 2002 tycks vara ungefär en poäng svårare än provet hösten 2001.

Normering baserad på IRT

Vid de senaste åtta provtillfällena 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 utprovningssmodellen.
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 BILOG.

4. Beroende på vilken skala som används för resultaten: IRT sann poäng ekvivalering eller observerad poäng ekvivalering.

Vid IRT-observerad-poäng-ekvivalering 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.

Provtillfället våren 2002 är det första där det finns uppgifter i samtliga delprov, som har ingått i ett och samma utprovningstillfälle, nämligen våren 2000. Det är därmed första gången ett användbart referenstest har funnits tillgängligt. Detta gör att utfallet av IRT-ekvivaleringen får större intresse än det tidigare har haft. De estimerade resultaten är att den population som genomförde provet våren 2000 hade medelvärdet 68.66 på detta test och skulle ha haft 67.16 på provet våren 2002. De estimerade medelvärdena för den population som genomförde provet våren 2002 blev 68.00 för provet våren 2000 och 67.10 på provet våren 2002. Enligt IRT är därmed provet våren 2002 ungefär en råpoäng svårare än provet våren 2000. Populationerna är däremot tämligen lika, möjligen är vårens population (2002) en liten aning svagare. Normeringsutfallet blev detsamma oberoende av vilken population som användes. Utfallet blev också, med undantag av ett fåtal poängintervall detsamma som med traditionell lika-percentil-ekvivalering.

Vid provtillfället hösten 2002 fanns uppgifter i samtliga delprov som, oförändrade, hade ingått i samma utprovningstillfälle, nämligen hösten 2000. Därmed kan IRT observerad-poäng-ekvivalering användas för att estimeras på detta sätt är, att den population som genomförde provet hösten 2000 hade medelvärdet 67.82 på det provet och, utifrån deras resultat på de gemensamma uppgifterna, skulle ha haft 65.60 på provet hösten 2002. Den population som gjorde föreliggande prov skulle, om de hade genomfört provet hösten 2000, ha haft 68.43 och estimerades på detta prov till 66.30. Enligt IRT skattas därmed detta prov som drygt två råpoäng svårare än provet hösten 2000. Eftersom den observerade skillnaden i medelvärden på dessa två prov var 2.55, skulle därmed två poäng av denna skillnad bero på skillnader i provens svårighetsgrad och en halv poäng skulle bero på skillnader mellan grupperna.

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. 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 2002 blev att den normerade poängen 1.0 erhålls i intervallet 68-71 råpoäng, vilket är en råpoäng lägre än såväl våren 2000 som våren 2001. För den normerade poängen 2.0 krävs 109 råpoäng eller mer, vilket är detsamma som våren 2001 men en poäng lägre än våren 2000. Den fullständiga normeringstabellen återfinns i bilaga 3a.

Resultatet av den genomförda normeringen av högskoleprovet hösten 2002 blev att den normerade poängen 1.0 erhålls i intervallet 67-70 råpoäng, vilket är detsamma som hösten 2001 men 2 poäng lägre än hösten 2000. För den normerade poängen 2.0 krävs 108 råpoäng eller mer, vilket är en råpoäng lägre än både hösten 2001 och hösten 2000. 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 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 en

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.

Skillnaden i medelvärde mellan dem som gjorde provet för första gången och dem som gjorde det för fjärde gången var 14 råpoäng vid provtillfället våren 2002 och 12 råpoäng vid provtillfället hösten 2002. I denna skillnad ingår både de faktiska skillnader som föreligger mellan de två grupperna som består av olika individer, de eventuella positiva effekterna av att upprepa provdeltagandet och den utveckling som individerna kan ha genomgått under tidsperioden. För de två procent av provdeltagarna som gjorde provet för fjärde gången var skillnaden i medelvärde till deras första provtillfälle cirka sju råpoäng och den genomsnittliga skillnaden i medelvärde mot deras tidigare bästa resultat var minus en tiondels råpoäng 02A och minus två och en halv råpoäng 02B.

Tabell 11. Andel provdeltagare som vid provtillfällena våren och hösten 2000, 2001 och 2002 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

En avvikelse från tidigare provtillfällen är att andelen provdeltagare som genomförde provet för första gången var något större än tidigare. Det tycks finnas en tendens till att intresset för att upprepa provdeltagandet kan ha minskat.

Männen är mer benägna att upprepa sitt provdeltagande. Den största andelen av dem som upprepar sitt provdeltagande återfinns i den yngsta åldersgruppen (20 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 medelvärdesresultat 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 6 april år 2002 var 40 548, av vilka 18 288 var män och 22 260 var kvinnor och den 26 oktober 2002 genomfördes provet av 28 589 personer, av vilka 12 222 var män och 16 367 var kvinnor. Sammanlagt under året var det drygt 69 000 som genomförde provet. Det är det lägsta antalet provdeltagare sedan de nya antagningsreglerna trädde i kraft 1991, vilka innebar att urvalet till studieplatser för samtliga sökande till högre utbildning sker antingen utifrån betyg eller resultat på högskoleprovet. De främsta anledningarna till denna minskning är förmodligen mindre årskullar och bättre arbetsmarknad. Från toppåret 1992 med cirka 145 000 provdeltagare har det skett en minskning med drygt 50 procent.

Sedan i början på 1990-talet har majoriteten av provdeltagarna varit ganska unga. Ungefär 70 procent av provdeltagarna år 2002 är 24 år eller yngre. Större andel män återfinns i de lägsta åldersgrupperna medan större andel kvinnor återfinns i de högsta åldersgrupperna.

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

De resultat som presenterats för högskoleprovet visar att provet våren 2002 var knappt en råpoäng lägre än det prov som genomfördes våren innan. Även resultatet på provet hösten 2002 var något lägre än provet hösten 2001. Medelvärde på provet våren 2002 var 67.72 råpoäng och på hösten 2002 var medelvärdet 66.49 råpoäng.

Vad gäller resultaten på 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 (cirka 5 råpoäng). 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 ungefär 1.5 poäng medan för delprovet LÄS är könsskillnaden liten.

För provet som helhet och för den totala gruppen provdeltagare är skillnaden mellan män och kvinnor 0.28 standardavvikelseenheter för såväl våren 2002 som hösten 2002. Resultaten vad avser könsskillnader i resultat avviker 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 skillna-

derna uppstår⁷. 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 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. Referensgrupp I (se sid 14) utgörs av en tämligen heterogen grupp provdeltagare där avsevärda skillnader föreligger mellan män och kvinnor avseende ålder och utbildning. Referensgrupp II är däremot en homogen grupp där åldern är densamma för män och kvinnor och där även utbildningens längd är densamma och det enda som skiljer är inriktning inom gymnasieskolan. Det är därför anmärkningsvärt att resultat-skillnaderna mellan män och kvinnor är något större i referensgrupp II. Det tycks alltså vara så att utbildningens innehåll har nog så stor betydelse som utbildningens längd.

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.

⁷ 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)

REFERENSER

- Andersson, K. (Red.). (1999). *Högskoleprovet. Konstruktion, resultat och erfarenheter* (Pm nr 153). 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. (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 universitet, Enheten för Pedagogiska mätningar.
- 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.