

Högskoleprovet

Genom elva forskares ögon

Högskoleverket 1996

Högskoleprovet – Genom elva forskares ögon

Producerad av Högskoleverket i augusti 1996

Högskoleverkets rapportserie 1996:22 R

ISSN 1400-948X

ISRN HSV-R--96/22--SE

Innehåll: Bedömnings- och utredningsavdelningen

Form: Informationsavdelningen

Tryck: AB PP Print, Stockholm, augusti 1996

Innehållsförteckning

Förord	5
1 Introduktion. <i>Eivor Johansson</i>	6
2 Kvinnor och män. <i>Christina Stage</i>	9
3 Är män bättre än kvinnor i engelsk läsförståelse? <i>Jan Hellekant</i>	18
4 Analys av engelska läsförståelseuppgifter. <i>Gudrun Erickson</i>	23
5 Forskning kring ordförrådet i det engelska läsförståelseprovet. <i>Sölve Ohlander</i>	27
6 Selektionsmekanismernas betydelse för gruppskillnader på Högskoleprovet. <i>Sven-Eric Reuterberg</i>	32
7 Högskoleprovet, social bakgrund och upprepad provtagning. <i>Jan-Eric Gustafsson & Maria Benjegård</i>	42
8 Den sociala bakgrundens betydelse för prestationer på Högskoleprovet. <i>Sven-Eric Reuterberg</i>	50
9 Kulturfria instrument – finns sådana hybrider? <i>Allan Svensson</i>	56
10 Intervju som urvalsinstrument. <i>Christina Stage</i>	66
11 Högskoleprovet som prognosinstrument. <i>Ewa Andersson & Simon Wolming</i>	70
12 Vad mäter betyg och högskoleprov. <i>Jan-Eric Gustafsson</i>	82
13 Testträning. <i>Gunilla Ögren</i>	89
Författarpresentation	93

Förord

Högskoleprovet har funnits sedan 1977. Från och med 1991 kan samtliga sökande till högskolan konkurrera om platserna med hjälp av resultat från högskoleprovet och 35-40 procent av platserna vid de längre programmen besätts på grundval av provresultat. Högskoleprovet kan således ha stor betydelse för den enskildes möjligheter att nå en attraktiv utbildningsplats. Årligen anmäler sig cirka 140 000 individer till provet, dvs. tre gånger så många som lämnar gymnasieskolans samhällsvetenskapliga och naturvetenskapliga program. Det är därför naturligt att stor uppmärksamhet riktas mot hur provet fungerar.

Alltsedan provets tillkomst har forskare granskat högskoleprovet och vetandet kring provet har successivt ökat. Denna kunskap har kunnat användas i samband med konstruktion och revidering av provet. Vid ett seminarium på Högskoleverket i slutet av januari 1996 redovisades forskningsresultat och förslag till ytterligare forskning över en rad aktuella aspekter på provet. I denna publikation presenteras anförandena vid detta seminarium av forskarna själva med en inledande kommentar över kunskapsläget av Eivor Johansson, ansvarig för högskoleprovet vid Högskoleverket.

Agnetta Bladh

I Introduktion

Eivor Johansson

Den största uppmärksamheten har under de senaste åren kommit att fästas vid det förhållandet att kvinnor i genomsnitt har något lägre provresultat än män. Samma mönster upprepas i andra länder med liknande prov. Det har bedömts som sannolikt att testformatet är en delförklaring. Flervalsfrågor har en tendens att gynna män medan egen produktion, såsom att formulera egna svar på frågor, gynnar kvinnor. Detta skulle också kunna förklara att kvinnor har bättre betyg än män då förmågan att formulera sig på egen hand har stor betydelse vid de flesta av skolans prov.

Erfarenheterna från högskoleprovet har visat att också de områden som texterna är hämtade från har betydelse för kvinnors respektive mäns prestationer. Redan 1985 kunde konstateras att män var bäst på frågor inom områdena fysik/teknik, sport/idrott, geografi, ekonomi/affärliv, kemi och politik. Kvinnorna hävdade sig bäst när det handlade om vård, hemkunskap, biologi, religion, litteratur och konst. Även när det handlar om att ange betydelsen av enskilda ord slår denna tendens igenom. Ord som *konsortium*, *inkognito* och *emittera* klarar fler män än kvinnor. Kvinnor har däremot lättare än män för ord som *trauma*, *plafond*, *preferens*, *ekivok* och *intonation*. Dessutom har det visat sig att uppgifter som innehåller index eller procent tenderar att bli svårare för kvinnor än för män. Ansträngningarna att utjämna skillnaderna i kvinnors och mäns resultat har till stor del byggt på denna kunskap. Det finns dock rimliga gränser för i vilken utsträckning man kan utnyttja s.k. kvinnliga miljöer för att kompensera de skillnader som testformatet eventuellt kan ge upphov till.

Pojkar lyckas bättre i det engelska läsförståelseprovet i de centrala proven i gymnasieskolan trots att flickor har bättre betyg i engelska. Även i högskoleprovets delprov engelsk läsförståelse lyckas männen som regel något bättre än kvinnor trots att man här gjort medvetna ansträngningar att kompensera för de skillnader som testformatet antas ge upphov till för att åstadkomma ett prov med lika resultat. Någon skillnad i läsvanorna bland pojkar och flickor som skulle kunna förklara skillnaden har inte kunnat upptäckas. Orsaken kan vara av helt annat slag nämligen att de duktigaste kvinnorna helt enkelt inte finns med bland dem som gör provet.

I forskningen kring provet tillämpas olika slag av analyser för att utröna orsakerna till könsskillnaderna i resultaten. Analyserna har visat att inte bara texternas innehåll påverkar resultaten, utan att även *frågornas* innehåll och karaktär har betydelse. De uppgifter där män lyckas bäst handlar ofta om att dra slutsatser om sakinnehåll snarare än om personer. Frågorna i provet är också av olika slag, dels faktabetonade, dels texttolkande. Det förefaller som om män och kvinnor lyckas olika väl på de olika frågetyperna.

Ytterligare kunskaper om provet kommer att erhållas genom lingvistiska studier. Ordförrådet i framför allt delprovet i engelsk läsförståelse kommer att kartläggas. Detta kan ge en uppfattning om hur avancerade texterna i provet är och om ordförrådet skiljer sig mellan lätta och svåra uppgifter. Möjligen kan man också få svar på frågan om det finns en manlig respektive kvinnlig vokalbulärprofil.

Även andra faktorer, som ligger utanför provets innehåll och konstruktion, kan påverka resultaten. Forskarna har visat att genom att göra högskoleprovet flera gånger förbättras provresultaten något. Detta beror dels på träningseffekter, dels på att chansen att få ett bra resultat ökar på grund av slumpfaktorer. Barn från tjänstemannahem gör provet fler gånger än andra och män gör provet fler gånger än kvinnor. Detta innebär att även om det åtminstone teoretiskt är möjligt att åstadkomma ett prov som i sig är neutralt i förhållande till social bakgrund och kön så kan skillnaderna i det upprepade provtagandet leda till skillnader i provresultat mellan individer från olika grupper.

När högskoleprovet infördes fanns förhoppningar om att provet skulle bidra till att minska den sociala snedrekryteringen till högskolan. Nu kan forskare konstatera att det finns betydande sociala skillnader i resultaten på provet. Man kan också slå fast att dessa skillnader är mer uttalade bland de kvinnliga provdeltagarna än bland de manliga. Att därmed hävda att provet missgynnar vissa sociala grupper är emellertid förhastat. Samma sociala skillnader ger också betyg och test av andra slag. Den centrala frågan är om provet trots allt minskar denna skillnad i samband med antagningen till högskolan.

Forskare ifrågasätter också om man skall förvänta sig samma genomsnittliga resultat för manliga och kvinnliga högskoleprovstagare. Provet är frivilligt och varje enskild individ avgör själv om han/hon skall genomgå det eller inte. Därmed utgör både de manliga och de kvinnliga provtagarna så kallade självselekterade grupper. Jämförelser som baseras på sådana grupper måste

alltid tolkas med stor försiktighet eftersom man normalt inte vet hur selektionsmekanismerna verkar i det enskilda fallet. Klart är att de bästa kvinnorna ofta väljer att inte göra provet. De män som gör provet är en mera positivt selekterad grupp och har bättre resultat än kvinnorna på ett tidigare genomgången test och har nästan lika bra betyg som kvinnorna. Fortsatt forskning får visa om det finns otillbörliga skillnader mellan olika gruppers resultat på provet.

Högskoleprovet används för att göra urval bland sökande till högre studier. Det är därför naturligt att ställa frågan: Vilken förmåga har provet att förutsäga framgång i studierna? En aktuell rapport visar att högskoleprovet och betyg har ungefär samma prediktiva värde mätt med avlagda poäng. Meriter för arbetslivserfarenhet ger en sämre prognos. Men kriterierna på studieframgång är inte självklara varken bland studenter eller bland lärare. I en kommande doktorsavhandling kommer begreppen att redas ut.

Den psykometriska forskningen har kunnat visa att högskoleprovet mäter olika förmågor: en generell analytisk förmåga, verbala kunskaper och färdigheter och delprovsspecifika förmågor. Den generellt analytiska förmågan mäts framför allt i högskoleprovets delprov NOG och DTK. Verbala kunskaper och färdigheter påverkar framför allt delproven ORD, LÄS och ELF i högskoleprovet. Resultaten ger anledning till reflektion över ytterligare användningsområden för provresultaten.

Om det nya betygssystemet på sikt inte kan användas kommer högskoleprovet att utgöra det huvudsakliga urvalsinstrumentet. Detta kan försämra kvinnornas möjlighet att konkurrera om de attraktiva platserna. Möjligen kan antagning i flera steg då vara en alternativ modell. Att använda mindre säkra test, och i synnerhet intervjuer, kan emellertid också leda till större problem än dem man vill komma till rätta med.

På grund av den stora betydelse ett provresultat kan ha för individen finns ett stort intresse för hur man kan förbereda sig för att uppnå så goda resultat som möjligt. Klart är att högre utbildning ger högre provresultat. Man kan också göra sig förtrogen med provets innehåll och uppläggning genom att studera gamla prov. Forskarna har också kunnat visa att genom att göra provet mer än en gång förbättrar man i allmänhet sina resultat.

I de följande utvecklas de aktuella studierna med reflexioner över högskoleprovet, dess innehåll, användning och konsekvenser.

2 Kvinnor och män

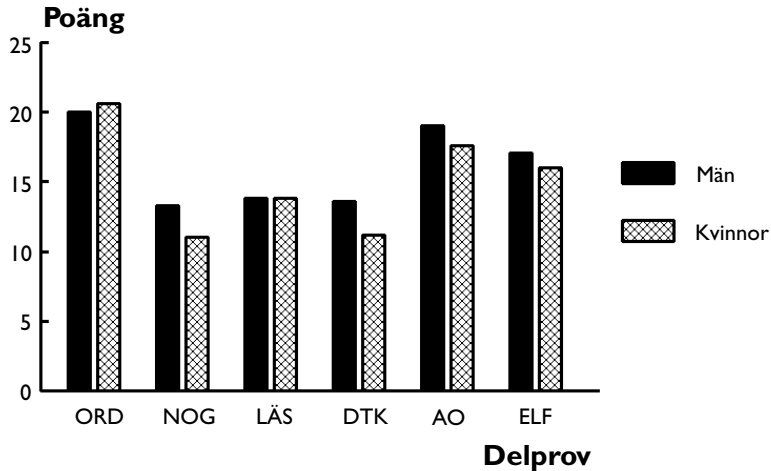
Christina Stage

Resultatskillnade mellan män och kvinnor på högskoleprovet har studerats intensivt. Redan 1985 kunde konstateras att män var bäst på frågor inom områdena fysik/teknik, sport/lidrott, geografi, ekonomilaffärliv, kemi och politik. Kvinnorna hävdade sig bäst när det handlade om vård, hemkunskap, biologi, religion, litteratur och konst. Även bland dem som har samma totalpoäng på provet finns skillnader i hur kvinnor och män klarat olika slag av uppgifter. I denna studie har en jämförelse gjorts, på delprovs- och uppgiftsnivå mellan män och kvinnor med samma totalresultat på provet. Även mellan dessa grupper av män och kvinnor finns tydliga skillnader i hur de klarat olika slag av uppgifter. Uppgifter som innehåller index eller procent tenderar att alltid bli svårare för kvinnor än för män och detsamma gäller ord som konsortium, inkognito och emittera. Kvinnor har däremot lättare än män för ord som trauma, plafond, preferens, ekivok, intonation.

Den första stora utprövningen av uppgifter och delprovstyper inför ett blivande högskoleprov genomfördes 1970 på 2 980 gymnasieelever. I samband med analyserna av resultaten av denna utprövning studerades även resultatskillnader mellan olika grupper. I detta skede av arbetet med (det blivande) högskoleprovet var skillnader mellan socialgrupper av större allmänt intresse än skillnader mellan könen. Resultaten från de 11 delprov som ingick i denna första utprövning analyserades dock även med avseende på könsskillnader i resultat. Dessa första analyser och resultaten av dem redovisas i ett psykologexamensarbete (Stage, 1975). Sedan dess och fram till 1993 publicerades c:a 41 arbeten om könsskillnader i resultat på högskoleprovet; dessa arbeten finns sammanfattade i en rapport, *Gender Differences on the SweSAT. A Review of Studies since 1975* (Stage, 1993). Även en avhandling, som publicerades under denna tid (Stage, 1985), handlade om könsskillnader i resultat på högskoleprovet. Efter 1993 har ytterligare ett tiotal rapporter skrivits om ämnet.

Vad är då kunskapsläget? En illustration kan vara en av de senare studierna på området: *Samma poäng - samma prestation?* (Stage, 1996). Den övergripande frågeställningen för denna studie var om den totala provpoängen, som generellt kan tolkas som ett mått på förväntad studieframgång, innebär detsamma för olika grupper av provdeltagare, i detta fall män och kvinnor.

I det högskoleprov som genomfördes hösten 1994 deltog 55 869 personer, 26 137 män och 29 732 kvinnor. I genomsnitt hade männen 96.9 råpoäng vilket motsvarar en normerad poäng på 1.0, medan kvinnorna hade 90.3 råpoäng och en normerad poäng på 0.9. Resultaten för män och kvinnor på de sex delproven visas i figur 1.



Figur 1. Resultat för män och kvinnor på de sex delproven i högskoleprovet hösten 1994.

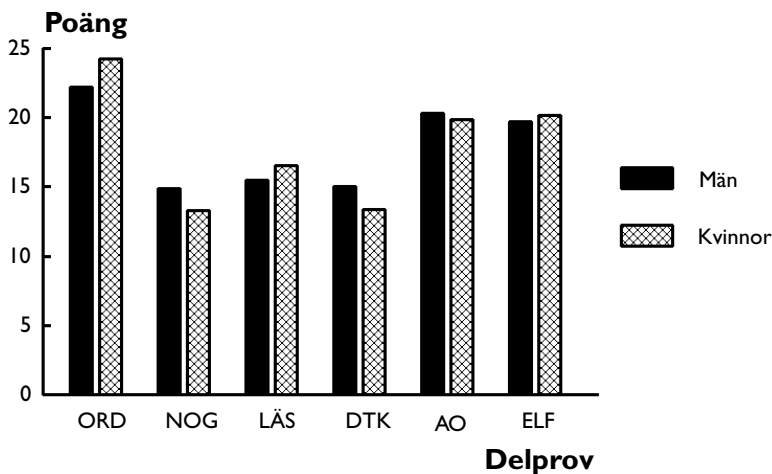
De skillnader som visas i figur 1 är de som alltid brukar återfinnas på högskoleprovet; män har högre genomsnittligt resultat på delproven NOG, DTK, AO och ELF medan kvinnor har något högre resultat på delprovet ORD och resultaten på delprovet LÄS är ungefär desamma för män och kvinnor. Dessa skillnader är i sig inte speciellt intressanta, eftersom grupperna män och kvinnor skiljer sig åt avseende såväl ålder som bakgrundsutbildning.

Syftet med studien var att jämföra resultaten på delprovs- och uppgiftsnivå mellan män och kvinnor som hade samma totalresultat på högskoleprovet. Därför utvaldes de provdeltagare som hade en normerad poäng = 1.3, vilket vid detta provtillfälle motsvarade en råpoäng mellan 106 och 109 poäng. Denna poäng hade erhållits av 3781 personer, 1985 män och 1796 kvinnor, vilket utgjorde 6.8 procent av samtliga provdeltagare hösten 1994. Anledningen till valet av just den normerade poängen 1.3 var att detta kan betraktas som ett klart konkurrensdugligt resultat, samtidigt som antalet provdeltagare är tämligen stort och fördelningen mellan män och kvinnor ganska jämn.

För grupperna män respektive kvinnor med den normerade poängen 1.3 undersöktes och jämfördes fördelning på ålder och bakgrundsutbildning, resultat på de sex delproven samt resultat på enskilda uppgifter.

Fördelningen av män och kvinnor på bakgrundsutbildning och ålder överensstämde tämligen väl med vad som gällde för totala gruppen provdeltagare

I figur 2 visas resultaten på de sex delproven i högskoleprovet hösten 1994 för de män och kvinnor som hade den normerade poängen 1.3.



Figur 2. Resultat på de sex delproven i högskoleprovet hösten 1994 för de män och kvinnor som hade den normerade poängen 1.3.

Av figur 2 framgår att även mellan män och kvinnor, som har samma totalresultat på provet föreligger skillnader på delprovsnivå. En jämförelse med figur 1 visar att skillnaderna snarast har ökat. Skillnaden till fördel för kvinnor har ökat på delproven ORD och LÄS, skillnaden till fördel för män kvarstår på delproven NOG och DTK, medan skillnaden på delprovet AO nästan försvunnit och skillnaden till fördel för män på delprovet ELF ändrats till en liten skillnad till fördel för kvinnor.

Vid analysen av resultatskillnaderna mellan män och kvinnor på enskilda uppgifter valdes den kategoriindelning av skillnaderna, som tillämpats i tidigare studier (se Stage, 1985, 1986, 1992). Uppgifter där skillnaden i

lösningensproportion mellan män och kvinnor varit +/- 0.02 eller mindre har betecknats som neutrala uppgifter eftersom detta motsvarar gränsen för en statistiskt signifikant skillnad när antalet individer är 250 och spridningen är maximal (dvs om $p = .5$). Övriga gränser har satts till .10 och .20 vilket har gett sju kategorier av uppgifter utifrån skillnaden mellan män och kvinnor i lösningensproportioner:

1. Extremt kvinnliga uppgifter: Uppgifter som mer än 20% fler kvinnor än män har besvarat korrekt
2. Klart kvinnliga uppgifter: Uppgifter som 11-20% fler kvinnor än män har besvarat korrekt
3. Kvinnliga uppgifter: Uppgifter som 3-10% fler kvinnor än män har besvarat korrekt
4. Neutrala uppgifter: Uppgifter där skillnaden är 2% eller mindre mellan andelarna män och kvinnor som svarat rätt
5. Manliga uppgifter: Uppgifter som 3-10% fler män än kvinnor har besvarat korrekt
6. Klart manliga uppgifter: Uppgifter som 11-20% fler män än kvinnor har besvarat korrekt
7. Extremt manliga uppgifter: Uppgifter som mer än 20% fler män än kvinnor har besvarat korrekt

Tabell 1 visar resultatet av denna kategoriindelning av uppgifterna för de provdeltagare som hade den normerade provpoängen 1.3

Tabell 1. Fördelning av provuppgifterna i högskoleprovet hösten 1994 utifrån könsskillnader i lösningensproportioner mellan män och kvinnor med den normerade poängen 1.3.

Kategori	Extremt						Extremt manliga	Σ
	kvinnliga	2	3	4	5	6		
ORD	-	9	14	4	2	1	-	30
NOG	-	-	-	2	15	3	-	20
LÄS	-	1	17	5	1	-	-	24
DTK	-	-	-	4	9	7	-	20
AO	1	3	4	6	12	4	-	30
ELF	-	-	11	11	1	1	-	24
Σ	1	13	46	32	30	16	0	148

Av samtliga 148 uppgifter var endast 32 neutrala, medan 46 var mer eller mindre manliga och 60 var mer eller mindre kvinnliga. Ingen av uppgifterna var extremt manlig medan en AO-uppgift var extremt kvinnlig. Denna AO-uppgift var på det hela taget svår; 38 procent av kvinnorna och 15 procent av männen i den här aktuella gruppen hade besvarat uppgiften korrekt. Uppgiften gällde *syftet med de ultraljudsundersökningar som görs i ett tidigt skede av graviditeten*. De tre "klart kvinnliga" uppgifterna behandlade respektive *Spanska ridskolan i Wien, en dikt av Stig Dagerman* och *växtfärgning med koschenill*. De fyra "klart manliga" uppgifterna handlade om *en gas som orsakar växthuseffekt, varor undantagna från moms, Alcazar-projektet och vad en pulsar är*.

Anmärkningsvärt är att ingen enda uppgift i varken delprovet NOG eller DTK blev kvinnlig trots att jämförelsen gjordes mellan män och kvinnor med samma totalresultat på provet. En av de neutrala uppgifterna i delprovet NOG löd

Olle köpte 1 kg äpplen, 1,5 kg potatis och 0,5 kg tomater samt ett salladshuvud. **Hur mycket kostade varorna tillsammans?**

- (1) Alla fyra varorna kostade lika mycket
- (2) Det genomsnittliga kilopriset för de fyra varorna var 14 kronor.

Och den av de "klart manliga" uppgifterna som gav upphov till den största skillnaden mellan män och kvinnor löd:

Index används för att visa förändring över tid. År 1991 fanns det i vårt land 42 erkända arbetslöshetskassor. **Vilket indextal motsvarar den utbetalda ersättningen från dessa kassor år 1991 om indextalet för basåret 1980 = 100?**

- (1) År 1984 (index = 346) var summan av de utbetalda ersättningarna hälften av vad den var år 1991.
- (2) År 1991 var de utbetalda ersättningarna 11 886 miljoner kronor.

Detta är ett exempel på att uppgifter som innehåller index eller procent tenderar att alltid bli svårare för kvinnor än för män.

Även i delprovet DTK kan två uppgifter på ett mycket bra sätt illustrera dessa skillnader som vanligtvis erhålls mellan män och kvinnor. Den senare var mycket manlig medan den förra var neutral.

Till en figursida, som innehöll diagram över jaktstatistik för räv och mård - ett diagram för Norrland och ett annat för Götaland och Svealand - ställdes två frågor varav en visade sig bli neutral och den andra ”klart manlig”; den neutrala uppgiften löd:

Ungefär hur många rävar respektive mårdar sköts i Sverige 1978?

- A 42 000 respektive 7 000
- B 64 000 respektive 7 000
- C 64 000 respektive 28 000
- D 68 000 respektive 7 000
- E 68 000 respektive 28 000

Och den andra löd:

Omkring år 1980 drabbades Götaland och Svealand av rävska. Ungefär hur mycket mindre var antalet skjutna rävar i området 1988 än 1981?

- A 50 procent
- B 60 procent
- C 70 procent
- D 80 procent
- E 90 procent

I delproven ORD, LÄS och ELF påverkas kvinnors och mäns resultat av det innehåll som orden respektive texterna handlar om .

De nio ”klart kvinnliga” ORD-uppgifterna hade huvudorden: *trauma, plafond, anbelanga, preferens, ekivok, intonation, kalligrafi, gracil och galej*. De tre ”manliga” ORD-uppgifterna hade huvudorden: *konsortium, inkognito och emittera*.

Den ”klart kvinnliga” LÄS-uppgiften hörde till en text med rubriken *Modern* och den ”manliga” uppgiften på detta prov hörde till en text med rubriken *Skola och samhälle*.

Den ”klart manliga” ELF-uppgiften bestod av en kort text med frågan: *What is the conclusion of the IMF (International Monetary Fund)?* Den uppgift som gav störst skillnad i lösningsproportion av de ”något kvinnliga” uppgifterna var till en lång text med rubriken *The Love Letters* och frågan löd: *How does the reviewer explain the fact that Einstein and Mileva Maric parted?*

Diskussion

Vid den analys av könsskillnader i resultat på 450 allmänorienteringsuppgifter som gjordes redan 1985 (Stage, 1985) konstaterades att de ämnesområden som var mest ”manliga” var fysik/teknik, sport/idrott, geografi, ekonomi/affärliv, kemi och politik. Ämnesområden som var kvinnliga var vård, hemkunskap, biologi, religion, litteratur och konst. De slutsatser som kan dras av den analys, som genomförts i denna studie, avviker inte på något sätt från tidigare studier.

Det som är anmärkningsvärt i resultaten av denna studie är att samma normerade poäng - i detta fall 1.3 - har så olika innebörd för män och kvinnor. Trots att män och kvinnor har samma totalresultat återkommer - på både delprovs- och uppgiftsnivå - de resultatskillnader som alltid brukar observeras på högskoleprovet.

Begreppet item-bias definieras så att en provuppgift är biased om provtagare som har samma prestationsnivå, men tillhör olika undergrupper, har olika sannolikhet att besvara uppgiften korrekt. Den för närvarande mest använda metoden för att undersöka förekomsten av item-bias är Mantel-Haenszel-metoden. Lite förenklat innebär denna metod att på varje totalpoängnivå jämförs grupper, utifrån totalpoängen, förväntade resultat med det faktiskt erhållna resultatet på varje enskild uppgift och om det föreligger skillnader mellan grupperna bedöms den aktuella uppgiften som biased.

Den analys som har gjorts här kan ses som en del av en M-H-analys. Skillnaden är att i denna studie har jämförelsen mellan grupper endast gjorts på en poängnivå, nämligen 1.3, medan vid en fullständig M-H-analys motsvarande jämförelse skulle ha gjorts på samtliga poängnivåer. Det kan dock diskuteras om det är av speciellt stort intresse om uppgifter skiljer mellan grupper på de lägsta poängnivåerna. I detta fall har jämförelsen enbart gjorts på en poängnivå som är av intresse i urvalssammanhang. Slutsatsen blir dock att av de studerade 148 uppgifterna är 16 biased till

fördel för män och 14 till fördel för kvinnor, enligt de kriterier som används av ETS¹.

Resultaten ger även anledning att fundera över vad testpoängen egentligen innebär. Resultatet på högskoleprovet används för att rangordna provdeltagarna och denna rangordning görs på grundval av den normerade poängen, som i sin tur bestäms av summan av poängen på de sex delproven.

Studieförmåga är ett komplext begrepp; den operationella definition som erhålls genom högskoleprovet är att denna förmåga består av ordförståelse, logisk, kvantitativ förmåga, svensk läsförståelse, förmåga att tolka diagram, tabeller och kartor, allmänorientering och engelsk läsförståelse. Det är fullt tänkbart och t. o. m troligt att dessa förmågor fungerar på ett kompensatoriskt sätt, så att t ex sämre logisk, kvantitativ förmåga skulle kunna kompenseras genom bättre läsförståelse och omvänt. Det sätt som högskoleprovet används förutsätter denna form av kompensation. Det skulle dock vara av betydelse att analysera begreppet studieförmåga mer ingående för att klarlägga dessa kompensatoriska egenskaper.

Det alternativa urvalsinstrumentet till högskolestudier, d. v. s. gymnasiebetygen, utgörs i ännu högre utsträckning av en sammanslagning av olika förmågor. Medelbetyg från gymnasieskolan beräknas på olika ämnen på olika gymnasielinjer, men räknas ändå som likvärdiga i konkurrens om studieplatser vid högskolan. Även här finns verkligen anledning att närmare fundera över delkomponenterna och deras kompensatoriska egenskaper.

Det som mäts med högskoleprovet, eller genom betygen, kan dock bara ses som nödvändiga men inte tillräckliga förutsättningar för studieframgång. Bland många andra självklara förutsättningar för studieframgång kan nämnas motivation, ambition, intresse och yttre omständigheter.

¹ Educational Testing Service är den organisation i USA, som ansvarar för Scholastic Assessment Test. Här genomförs rutinmässigt M-H-analyser på samtliga test och man klassificerar uppgifterna efter utfallet i: A- (försumbar DIF), B- (måttlig DIF) och C-uppgifter (svår DIF), som ungefär överensstämmer med A = 3 och 5, B = 2 och 6 samt C = 1 och 7 i denna studie.

Referenser

Berk, R.A. (Ed.) (1982). *Handbook of Methods for Detecting Test Bias*. Johns Hopkins University Press, Baltimore, Maryland.

Hambleton, R. & Rogers, J. (1989). Detecting Potentially biased test items: Comparison of IRT and Mantel-Haenszel methods. *Applied Measurement in Education*. 2 (4). p 313-334.

Holland, P.W. & Thayer, D.T. (1988). Differential item performance and the Mantel-Haenszel procedure. In Wainer, H. & Braun, H.I. (Eds.) *Test Validity*. Hillsdale NJ: Lawrence Erlbaum Associates p 129-145.

Stage, C. (1975). *Intellektuella könsdifferenser*. Psykologexamensarbete. Pedagogiska institutionen, Umeå universitet.

Stage, C. (1985). *Gruppskillnader i provresultat*. Uppgiftsinnehållets betydelse för resultatskillnader mellan män och kvinnor på prov i ordkunskap och allmänorientering. Akademiska avhandlingar nr 17. Pedagogiska institutionen, Umeå universitet.

Stage, C. (1986). Könsskillnader i resultat på sex högskoleprov. *Pm nr 9*, Pedagogiska institutionen, Umeå universitet.

Stage, C. (1988). Gender Differences in Test Results. *Scandinavian Journal of Educational Research* 32, p 101-111.

Stage, C. (1992) How Important are Age and Education for Gender Differences in Test Results? *Scandinavian Journal of Educational Research* 36, p 223-235.

Stage, C. (1993). Gender Differences on the SweSAT. A Review of Studies since 1975. *Em, No 7*. Division of Educational Measurement, Umeå university.

Stage, C. (1996). Samma poäng - samma prestation? *Pm nr 113*, Enheten för pedagogiska mätningar, Umeå universitet.

3 Är män bättre än kvinnor i engelsk läsförståelse?

Jan Hellekant

Flickor har bättre betyg i engelska, men pojkar har som regel haft bättre resultat på läsförståelseuppgifterna i de centrala proven i gymnasieskolan. Männen presterar också bättre på högskoleprovets delprov i engelsk läsförståelse, trots att provkonstruktörerna bemödar sig om att åstadkomma ett prov som män och kvinnor skall lyckas lika bra på. Det finns ingenting som tyder på att pojkars och flickors läsvanor skulle kunna förklara resultatskillnaden. En sannolik förklaring är att den manliga gruppen provdeltagare är mer kvalificerad på grund av att många duktiga kvinnor avstår från att göra högskoleprovet.

Ett engelskt läsförståelseprov (ELF) infördes i högskoleprovet våren 1992. Eftersom högskoleprovet redan innehöll ett läsförståelseprov i svenska - det viktigaste nationella språket - ansågs det relevantt att även i detta sammanhang testa engelskan - det för oss viktigaste internationella språket. Man hyste därjämte vissa förhoppningar att en engelsk läsförståelsedel skulle kunna bidra till att lösa problemet med de i varje prov återkommande skillnaderna i resultat mellan män och kvinnor till de senares nackdel. Dessa förhoppningar var grundade på erfarenheten att kvinnor är duktigare i språk och att flickor i skolan har bättre språkbetyg än pojkar.

Resultat på de centrala proven i gymnasieskolan

Vad dock inte alla kände till var att pojkarna presterade bättre på läsförståelsedelarna i gymnasieskolans centrala prov. Läsförståelseprovet i de centrala proven i engelska för åk 2:3 / 2:4 bestod av en längre text med c:a 10 frågor försedda med fyra svarsalternativ och c:a 10 texter, s k minitexter, var och med en fråga och fyra svarsalternativ. Pojkarna lyckades undantagslöst bättre på läsförståelseprovet än flickorna. Detta gällde i synnerhet den del som omfattade minitexterna. 1993 låg pojkarnas medelvärde 5 procentenheter över flickornas på läsförståelseprovets första del men hela 14 procentenheter över på minitexterna. 1994 var skillnaderna 4 resp 6 procentenheter till pojkarnas fördel och 1995 var motsvarande siffror 7 resp 11 procentenheter.

Det är intressant att jämföra pojkars och flickors resultat på en annan del av samma centralprov. Denna del är en längre text i vilken vissa ord strukits och där eleverna själva skall skriva in ett ord i varje lucka för att visa dels att de förstått textens innehåll, dels att de kan det utelämnade ordet på engelska. På denna del som alltså i hög grad testar läsförståelse var det *flickorna* som lyckades bättre än pojkarna.

En analys av utfallet på lucktexten och en annan del av centralprovet vilken består av korta lucktexter med fasta alternativ visar att i de prov som gavs åren 1986-1993 flickorna presterade bättre än pojkarna när de själva skulle skriva in det ord som saknades i luckorna, medan pojkarna lyckades bättre än flickorna när eleverna skulle välja mellan fasta alternativ. Medelvärden för pojkarna låg i genomsnitt 8 procentenheter över medelvärdet för flickorna. Slutsatsen var att det finns skäl att anta att skillnaden kan bero på testformatet och att flervalformatet förefaller missgynna flickor som grupp (Hellekant, 1994).

Det första ELF-provet

När arbetet med ELF-provet hunnit så långt att den första utprovningen av 108 items (uppgifter) fördelade på sex utprovningversioner avslutats våren 1991, visade det sig att det skulle bli svårt att infria de förhoppningar som avsåg skillnaderna i resultat mellan män och kvinnor. Männens resultat låg nämligen genomsnittligt drygt fyra procentenheter över kvinnornas. Erfarenheterna från centralproven upprepades alltså: på engelska läsförståelseuppgifter av flervalstyp lyckas män i regel bättre.

När ett ELF-prov komponeras, beaktas ett flertal principer. Provet skall ha ett p-värde på drygt .60 (den genomsnittlige provdeltagaren skall förväntas klara drygt 60 procent av uppgifterna), det skall ha hög reliabilitet (slumpen skall spela så liten roll som möjligt för resultatet) och det skall representera skilda ämnesområden. Till detta kom det dåvarande vetenskapliga rådets rekommendation att provresultat helst borde utfalla till den kvinnliga gruppens fördel. När det första ELF-provet (92A) konstruerades valdes uppgifterna så, att de nämnda kriterierna skulle uppfyllas så långt det var möjligt. Det gick dock inte med utgångspunkt från det material som då fanns utprovat att konstruera ett prov där prognosen för kvinnor var mer gynnsam än den för män. Faciliteten (svårighetsgraden) beräknades dock till .64 för både män och kvinnor. Utfallet i det verkliga provet .77 för män och

.69 för kvinnor blev en överraskning. Männens resultat låg 13 procentenheter över prognosen och kvinnornas 5 enheter. Skillnaden till männens fördel var alltså hela 8 procentenheter.

En fråga som ställdes, när resultatet blev känt, var om uppgifterna konstruerats av män. Svaret var att av konstruktörerna var två män och en kvinna. Den kvinnliga konstruktören är mycket kvalificerad. Hon är fil dr i engelsk litteratur och har lång erfarenhet som konstruktör av läsförståelseprov. Hon är dessutom uppmärksam på jämställdhetsfrågor och mycket medveten i sitt val av texter. Utfallet på de uppgifter som vår kvinnliga konstruktör tagit fram avvek dock inte från dem som de manliga konstruerat. Skillnaden i vid utprövningen var även här ca fyra procentenheter till männens fördel.

Selektionseffekter

De uppgifter som ELF-provet innehåller är valda så, att män och kvinnor skall få ungefär samma medelvärde. Det kan därför hävdas att vi i ELF-provet genom vårt val av uppgifter har gjort vad som går för att kompensera kvinnorna för den negativa effekt flervalssystemet kan tänkas ha på deras resultat. Varför ligger då männens resultat trots våra ansträngningar alltid 4-8 procentenheter högre än kvinnornas när provet väl har genomförts? En rimlig förklaring till detta är s k selektionseffekter, något som Sven-Eric Reuterberg utreder i ett annat kapitel av denna publikation. Det har visat sig att de män som söker sig till provlokalerna på det hela taget är mer kvalificerade än kvinnorna och därför sannolikt bättre i engelsk läsförståelse. Med andra ord: männens medelvärde inte bara *är* utan *skall också vara* högre än kvinnornas på det verkliga provet. Man kan förstås fråga sig om selektionseffekterna förklarar hela medelvärdeskillnaden eller spekulera i om "flervalseffekten" trots kompensatoriska åtgärder orsakar en del av skillnaden genom att den slår hårdare i avgörande provsituationer (*high-stake situations*).

Kvinnors och mäns läsvanor

En annan orsak till att männen presterar bättre än kvinnorna på ELF-provet skulle kunna vara att pojkar läser mer engelska utanför skolan än flickor och därmed blir bättre på att hämta information ur de engelska texter som de möter i ELF-provet.

För att få en uppfattning om vad pojkar och flickor läser på engelska utanför skolan lät vi ett antal gymnasieklasser besvara en enkät angående sina läsvanor (Hellekant 1995). I enkäten skulle eleverna ange hur ofta de läste olika typer av engelsk text utanför skolan. Enkäten gavs i åk 3 till drygt 800 elever som våren 1995 deltog i en utprovning av nyskrivna uppgifter till högskoleprovet. Eleverna var någorlunda jämnt fördelade över linjerna H, S, E, N och T. Genom att de deltog i en av våra utprovningar för högskoleprovet, vilka sker två gånger årligen och i vilka vissa uppgifter alltid är desamma, s k ankaruppgifter, var det möjligt att jämföra gruppens nivå med tidigare grupper. Av resultatet på ankaruppgifterna framgick att gruppen inte i någon högre grad skilde sig från de grupper som tidigare deltagit i utprovningar av engelska läsförståelseuppgifter för högskoleprovet. Medelvärdet för pojkarna var som vanligt något högre på dessa uppgifter än motsvarande medelvärde för flickorna.

Antagandet att pojkar i betydligt större utsträckning än flickor skulle läsa engelsk text utanför skolan får inte något stöd i denna undersökning. Det visar sig att flickorna oftare läser skönlitteratur och att pojkarna oftare läser texter om data, teknik och sport. I övrigt är inte skillnaderna särskilt stora.

Man måste dock fråga sig om de skillnader i läsvanor som enkätsvaren faktiskt avspeglar ändå är av den arten att de kan vara en bidragande orsak till resultat skillnaderna mellan pojkar och flickor i ELF-provet.

Detta förefaller föga troligt. Texterna i ELF-provet speglar olika ämnesområden, t ex kultur, samhälle, medicin, naturvetenskap men är trots detta föga specialiserade. Just sådana texter förekommer regelbundet i tidningar och tidskrifter. Det är därför intressant att konstatera att både pojkar och flickor på det hela taget är lika flitiga tidningsläsare och att de lika ofta även läser nyhetsmagasin av typ *Newsweek*. Det är svårt att tro att pojkarnas större intresse för data och teknik skulle kunna förklara en del av resultat skillnaderna.

Enkäten ger även besked om att dagens gymnasister på sin fritid läser engelsk text av olika karaktär utanför skolan och detta i en utsträckning som skulle förvånat forna tiders gymnasister. Engelskan är inte ett främmande språk på samma sätt som då utan börjar mer få karaktären av ett andraspråk. Då var det ett skolämne. Nu är det dessutom ett kommunikationsmedel utanför skolan.

Slutsats

I de centrala proven i engelska för åk 2:3 / 2:4 presterar pojkar bättre än flickor på flervalsuppgifter. Det motsatta förhållandet gäller produktiva uppgifter - även när endast ett ord skall produceras per uppgift. På de uppgifter som prövas ut till ELF-proven (alla av flervalstyp) har pojkarna också på det hela taget lyckats något bättre. En något kontroversiell hypotes som har ett visst internationellt stöd är att skillnaden främst beror på testformatet och att flervalsformatet tenderar att missgynna flickor som grupp (se t ex Bolger & Kellaghan 1990).

Till skillnad från hur man arbetar med centralproven väljs uppgifter som prövats ut i gymnasieskolan till ELF-provet på så sätt att flickornas medelvärde beräknas bli ungefär lika högt som pojkarnas. Genom valet av uppgifter försöker vi kompensera flickorna för den negativa effekt som flervalsformatet eventuellt har på deras resultat.

Att pojkar trots dessa kompensatoriska åtgärder lyckas bättre än flickor på ELF-provet torde främst kunna förklaras av selektionseffekter. Just den grupp män som *gör ELF-provet* är helt enkelt bättre i engelsk läsförståelse än kvinnogruppen. Ingenting tyder dock på att de pojkar i gymnasieskolan som *deltagit i utprovningarna* skulle vara bättre i engelsk läsförståelse än flickorna. I varje fall läser de inte oftare än flickorna engelskspråkig text av den art som förekommer i ELF-provet.

Referenser

Bolger, N. & Kellaghan, T.(1990). Method of Measurement and Gender Differences in Scholastic Achievement. *Journal of Educational Measurement*, Summer 1990, Vol. 27, No. 2, pp. 165-174.

Hellekant, J. (1994) Are multiple-choice tests unfair to girls? *System*, 22, 3.

Hellekant, J (1995) Engelsk fritidsläsning. *LMS-Lingua* 4/1995, bilaga *Språkpedagogik* 51.

4 Analys av engelska läsförståelseuppgifter

Gudrun Erickson

Olika typer av analyser av ELF-provet görs för att utröna orsakerna till könsskillnaderna i resultaten. Analyserna visar att texternas innehåll har stor betydelse för utfallet. Vidare undersöks effekten av olika slags rubriker, liksom av frågornas innehåll och karaktär. Frågorna i provet är dels faktabetonade, dels mera texttolkande. Av stort intresse är att studera i vad mån män och kvinnor lyckas olika väl på de olika frågetyperna och hur detta kan sättas i relation till texternas innehåll.

Analysen görs kontinuerligt av de uppgifter i högskoleprovet som testar engelsk läsförståelse (ELF). Så studeras till exempel innehåll, validitet, reliabilitet, svårighetsgrad och spridning. Särskild uppmärksamhet ägnas skillnader mellan kvinnliga och manliga testtagares resultat på ELF-provet som helhet och på enskilda uppgifter inom detta delprov.

Jan Hellekant har i det förra kapitlet fokuserat frågan om könsskillnader i ELF-provet och ställt frågan om man, på basis av dessa skillnader, kan dra slutsatsen att män är bättre än kvinnor i engelsk läsförståelse. På ett generellt plan är svaret på denna fråga med allra största sannolikhet 'nej'. Att män trots detta lyckas bättre på ELF-provet går inte att förklara på något enkelt sätt. Orsakerna är många och komplexa, och analyserna av resultatskillnaderna måste därför vara mångfasetterade och spänna över ett vitt fält.

Utöver de undersökningar som redan nämnts skall här några ytterligare analysmöjligheter kort beskrivas och kommenteras.

Analysansatser

Olika statistiska analysmetoder har använts för att fastställa vilka uppgifter och uppgiftstyper i ELF-provet som givit upphov till stora könsskillnader. Så har t ex medelvärdes- och spridningsskillnader mellan kvinnors och mäns svar på uppgifter av skilda slag studerats. Med hjälp av multivariata metoder har analyserna också kunnat fördjupas genom att olika dimensioner av begreppet engelsk läsförståelse satts i relation till utfallet för kvinnor och män.

Olika aspekter av läsförståelse

I Jan-Eric Gustafssons kapitel i denna skrift diskuteras olika aspekter av flerdimensionalitet i högskoleprovet. Det påpekas bl a att två gemensamma kunskapsdimensioner ingår i samtliga delar av provet, nämligen en analytisk och en generellt verbal faktor. Dessutom finns i varje del en mera specifik dimension.

Begreppet engelsk läsförståelse är följaktligen flerdimensionellt. Den analytiska komponenten visar sig här vara av underordnad betydelse, medan den generellt verbala förmågan är tydlig, liksom den mera specifikt engelska läsförståelsedimensionen. För att kunna läsa och förstå en engelsk text krävs alltså både en övergripande förmåga att läsa och förstå texter av olika slag och en mera engelskspecifik färdighet, som främst hänger samman med förståelse av ord och uttryck på det främmande språket.

Analyser har gjorts av utfallet på uppgifterna i samtliga ELF-prov i relation till den generella, verbala faktorn och den specifikt engelska dimensionen. Det visar sig att uppgifterna i olika hög grad hänger samman med respektive faktor. Flertalet uppgifter tycks dock ha en tydligare relation till den generellt verbala än till den specifika faktorn. Detta kan vidare jämföras med de könsskillnader som föreligger på de olika uppgifterna och uppgiftstyperna. Bland annat kan frågan ställas hur mycket av dessa skillnader som förklaras av den generellt verbala dimensionen och hur mycket som hänger samman med den mera specifikt engelska faktorn.

Arbetet med multivariata analyser av ELF-uppgifter har nyligen påbörjats, varför några definitiva slutsatser av vad som hittills framkommit inte kan dras. Sättet att studera flerdimensionalitet på item-nivå tycks dock vara mycket fruktbart. Fördjupade analyser av detta slag förefaller ytterst angelägna, inte minst i valideringen av proven. Det är därför glädjande att konstatera att det nya utprövningsförfarandet ger möjlighet till ett fortsatt utvecklingsarbete inom detta område.

Textinnehåll och rubriker

Med hjälp av olika analysmetoder har, som tidigare nämnts, vissa uppgifter i ELF-provet kunnat pekas ut som tydligt "manliga" respektive "kvinnliga". Uppgifter där fler män än kvinnor svarat korrekt dominerar. Innehållsliga analyser av de texter som ingår i uppgifterna ger inga entydiga svar på frågan

om vilka typer av innehåll som medför att män eller kvinnor lyckas bäst. Det kan dock noteras att texter som berör naturvetenskap, samhällsfrågor och ekonomi ofta tycks vara lättare för män, medan kvinnor inte sällan lyckas bättre på uppgifter som handlar om människor, mänskliga relationer och forskning kring dessa frågor. Några konkreta exempel på "manliga" uppgifter är texter som handlat om dinosaurier, mexikansk inrikespolitik och den internationella monetära fonden. Uppgifter som kvinnor lyckats lika bra eller bättre på har t ex berört Martin Luther Kings arbete, Albert Einsteins relationer till kvinnor och olika yttringar av modersinstinkten.

Det finns enstaka undantag från den ovan beskrivna stereotypen, men mönstret tycks trots allt tydligt. Det förefaller dock angeläget att även fortsättningsvis ingående analysera uppgifters innehåll och att i görligaste mån försöka undvika texter som resulterar i kraftiga könsskillnader. Detta får dock inte innebära att vissa ämnesområden tas bort ur provet.

Rubrikerna till de olika texterna i ELF-provet bestäms i de flesta fall av testkonstruktörerna, medan själva texterna är helt autentiska. En viktig faktor som kommer att analyseras ytterligare är effekten av olika rubriksättningar på könsskillnader i resultaten. Jämförelser av olika rubriker, "manliga", "neutrala" och "kvinnliga", på samma text kan t ex göras och utfallet studeras. Givetvis kommer detta emellertid enbart att göras i de fall de olika rubrikerna bedöms som relevanta för sina respektive texter.

Uppgifts- och frågetyper

ELF-provet innehåller tre olika uppgiftstyper: långa texter med fyra till sju frågor, korta texter med en eller två frågor och en sammanhängande text med ungefär sex luckor, där uppgiften är att komplettera texten med korrekta ord eller uttryck. I samtliga fall skall det riktiga svaret väljas bland fyra alternativ. Uppgifter baserade på längre texter utgör i regel ungefär halva provet.

Frågorna formuleras på olika sätt. I vissa fall efterfrågas mera faktabetonad information (ex: *Why was the 1794 law made?* eller *What did Ellen do after Dickens's death?*), medan frågorna i andra fall har en mera texttolkande eller inferensartad karaktär (ex: *What is the main message of this text?* eller *What is the general impression we get of the writer?*).

De hittills gjorda analyserna har visat att frågornas innehåll mer än deras form tycks ha betydelse för resultaten. Det kan t ex nämnas att den första

faktabetonade frågan ovan (1794) var lättare för män, medan den andra (*Dickens*) besvarades korrekt av något fler kvinnor. Samma förhållande gäller de inferensartade frågorna: den fråga som krävde att testtagaren skulle avgöra textens huvudbudskap var väsentligt enklare för män, medan den där inferensen gällde intrycket av en person besvarades korrekt av fler kvinnor.

Analysen av extremt "manliga" uppgifter visar att frågorna i dessa fall ofta har en inferensartad karaktär, och att testtagaren i regel skall dra slutsatser om sakinhåll snarare än om person.

De olika frågetyperna kräver olika lästekniker eller 'läsarter'. Faktabetonade frågor förutsätter detaljerad läsning, där ofta möjligheten att svara rätt handlar om förståelse av enstaka engelska ord eller fraser. Frågor av inferenskaraktär kräver däremot en mera övergripande läsning med syfte att förstå och dra slutsatser av ett sammanhang. Förståelse av enskilda ord spelar här inte samma avgörande roll som i de mera detaljerade frågorna.

Vissa inledande analyser har även gjorts av utfallet på de olika fråge- och uppgiftstyperna i relation till den generella verbala faktorn respektive den specifika engelska läsfaktorn. Det finns tecken på att den uppgiftstyp där texter skall kompletteras med ord mera kan relateras till den specifika faktorn än övriga uppgiftstyper, men det är för tidigt att uttala sig om detta mer än i ytterst tentativa termer. Vidare analyser får också visa på eventuella relationer mellan de olika faktorerna och könsskillnader på de olika uppgiftstyperna i provet.

Sammanfattning

Många olika typer av analyser görs och planeras för den engelska delen av högskoleprovet. Dessa analyser har olika fokus och kräver olika metoder. Det gemensamma syftet med samtliga är dock att skapa en djupare kunskap om begreppet engelsk läsförståelse och också om de faktorer som betingar olika individers resultat på provet. Förhoppningen är att denna kunskap skall bidra till den vidare utvecklingen av provet som helhet liksom av enskilda provtyper och uppgifter.

En väsentlig del av forsknings- och utvecklingsarbetet kring ELF-provet kommer att utgöras av lingvistiska analyser av de texter som ingår. Dessa analyser diskuteras av Sölve Ohlander i det följande kapitlet.

5 Forskning kring ordförrådet i det engelska läsförståelseprovet

Sölve Ohlander

För att kunna tillgodogöra sig innehållet i autentiska texter av den art som används i ELF-provet krävs ett relativt stort och varierat engelskt ordförråd. Genom att jämföra det ordförråd som används i ELF-provet med andra undersökta texter kan man få en bild av hur avancerade ELF-texterna är. Vidare kan sådana lingvistiska studier ge svar på om ordförrådet skiljer sig mellan lätta och svåra uppgifter och om det möjligen finns något som skulle kunna betecknas som en typiskt kvinnlig respektive manlig vokabulärprofil.

De autentiska texter som ELF-provet är baserat på är genomgående av sakprosekaraktär; rent skönlitterära texter förekommer således inte. De kan t.ex. bestå av artiklar hämtade ur engelskspråkiga tidningar och tidskrifter, riktade till en relativt allmän läsekrets snarare än till specialister inom olika fackområden. Såväl brittisk som amerikansk engelska förekommer. De individuella texternas längd kan variera från en dryg sida till några få rader. Som typiska exempel på sådana breda ämnesområden som behandlas i texterna kan nämnas naturvetenskap, samhälle, ekonomi, kultur och vård.

Förståelsen av de i provet ingående textavsnitten testas genom uppgifter i form av fyrvälsfrågor som på olika sätt berör sådant som kan utläsas ur texterna. Vid sidan av mer specifika frågor om innehållet i en text kan en uppgift också gälla den huvudsakliga ”poängen” eller avsikten med en text, eller vilken generell slutsats man som läsare kan dra ur den. Som en allmän princip gäller att uppgifterna skall vara formulerade på ett relativt enkelt språk; svårighetsgraden i en uppgift bör bestämmas av det aktuella textavsnittets karaktär snarare än av formuleringar i frågan och svarsalternativen.

Läsförståelse och ordförråd

Huruvida en skriven text uppfattas som lätt eller svår att förstå bestäms till betydande del av de ord som ingår, i förhållande till läsarens (passiva) ordförråd. Andra faktorer som spelar in är grammatisk komplexitet, semantiska och logiska relationer mellan olika delar av texten, dess uppbyggnad

m.m. Svårighetsgraden i en läsförståelseuppgift av det slag som ingår i ELF-provet bestäms dessutom av hur själva frågan är formulerad samt hur svarsalternativen ser ut, dvs. av samspelet mellan text och uppgift.

För att kunna tillgodogöra sig innehållet i autentiska texter av den art och svårighetsgrad som används i ELF-provet – och rätt lösa uppgifterna krävs ett relativt stort och varierat engelskt ordförråd inom centrala ämnesområden.

Det faller sig därför naturligt att lingvistisk forskning kring ELF-provet åtminstone i en inledande fas fokuseras på språkets vokabulär- eller lexikonsida.

Hur ser ordförrådet i hela ELF-provet ut?

Detta grundläggande undersökningsmoment syftar till en genomlysning, frekvensmässigt och semantiskt, av hela det ordförråd som kommit till användning i ELF-provet sedan starten 1992. Vilka individuella ord, liksom även typer av ord (semantiskt och grammatiskt), är vanligast förekommande respektive ovanligast i ELF-provet? Vilka frekvensmässiga skillnader finns vidare mellan själva textmaterialet och frågor/svarsalternativ? I en frekvensundersökning av detta slag fokuseras intresset huvudsakligen på de ”öppna”, s.k. lexikala ordklasserna, dvs. substantiv, verb, adjektiv och vissa adverb. Det är sådana ord som till största delen bestämmer en texts betydelseinnehåll, vad den ”handlar om”, till skillnad mot t.ex. pronomen och prepositioner. (Dock kan exempelvis frekvensuppgifter för ord som *he* och *she* ge intressanta upplysningar om texterna i ett kvinnligt/manligt perspektiv.)

Förutom en undersökning av de relativa frekvenserna hos orden i ELF-provet är det också av intresse att studera ELF-ordens allmänna frekvens i engelskan. Sådana frekvensuppgifter för engelskans ordförråd i stort finns att tillgå i de mycket omfattande, representativa, redan frekvensundersökta textsamlingar som idag finns tillgängliga. En studie av detta slag kan ge besked om den relativa svårighetsgraden i ELF-provens ordförråd. Även jämförelser med tidigare gjorda vokabulärundersökningar avseende läromedel i engelska för gymnasieskolan är här av intresse: hur avancerade är ELF-texterna i ett sådant perspektiv?

En kartläggning av ELF-ordens semantiska egenskaper är också en angelägen forskningsuppgift, med direkt anknytning till texternas innehåll. Detta förutsätter en semantisk klassificering av ordförrådet i provet. Bl.a. följande frågor kan ställas: Vilka betydelsefält finns representerade, med hur många och hur frekventa ord? Vilket samband finns mellan de grova innehållskategorier (naturvetenskap, samhälle, vård m.m.; jfr ovan) i vilka ELF-texterna i förväg klassificerats och resultatet av en mer noggrann semantisk analys av texternas ordförråd? Är det möjligt att klassificera ELF-texternas ordförråd som relativt könsneutralt, eller kan det vara berättigt att tala om en dominans för t.ex. ”manliga” ord i provet som helhet?

Sammantaget bör en frekvensundersökning av ovan skisserat slag samt en genomlysning av ordförrådets semantiska egenskaper kunna ge intressanta upplysningar om ELF-provets ”vokabulärprofil” i stort. I detta sammanhang är det också av intresse att undersöka hur stora variationerna i vokabulär är mellan olika provtillfällen. Hur konstant är ELF-provet i detta avseende?

Med den ovan skisserade undersökningen av hela ELF-provets ordförråd som grund blir det möjligt att gå vidare till mer specifika frågor kring förhållandet mellan provets vokabulärprofil och provtagarnas resultat. Några sådana frågor berörs nedan.

Hur ser ordförrådet ut i de ”svåraste” respektive ”lätteste” uppgifterna?

Det kan vara av intresse att undersöka eventuella samband mellan en uppgifts relativa svårighetsgrad och det aktuella ordförrådet i det tillhörande textavsnittet liksom i själva frågeapparaten. Finns det exempelvis några gemensamma drag beträffande frekvens och semantiska egenskaper i den vokabulär som finns i de svåraste respektive lättaste uppgifterna? Är det möjligt att t.ex. hävda att en lätt respektive svår uppgift har en vokabulärprofil av en viss karakteristisk typ, i själva textavsnittet och/eller i frågeapparaten? Eller är den relativa svårighetsgraden hos en uppgift mer betingad av andra faktorer, i texterna eller i frågorna och svarsalternativen? Den grundläggande frågan är med andra ord i vad mån det på basis av själva ordförrådet i en uppgift (textavsnitt, fråga/svarsalternativ) går att med någon säkerhet förutsäga uppgiftens relativa svårighetsgrad.

Hur ser ordförrådet ut i olika innehållskategorier, särskilt ”kvinnliga” respektive ”manliga”?

För varje ELF-prov görs, som en del i konstruktionsarbetet, en innehållslig grovklassificering av de texter som används i provet. Grundtanken är att inget särskilt ämnesområde vid varje provtillfälle skall vara markant över- eller underrepresenterat. Indelningen i ämnesområden sker på intuitiv, ”holistisk” grund. Här infinner sig den generella frågan i vad mån detta slags innehållsklassificering också motsvaras av en för varje ämnesområde typisk vokabulärprofil, semantiskt och frekvensmässigt. I detta sammanhang kan särskild uppmärksamhet riktas mot ämnesområden som med rätt eller orätt brukar uppfattas som ”typiskt kvinnliga” (t.ex. ”vård”) respektive ”typiskt manliga” (t.ex. ”naturvetenskap”). Uppvisar ordförrådet i dessa slags texter, även i jämförelse med texter som anses mer könsneutrala, specifika vokabulärprofiler? Om inte, vilka andra faktorer kan i så fall tänkas spela in, i texterna och/eller i frågeapparaten, som förklaring till att vissa texter/uppgifter intuitivt uppfattas på det ena eller andra sättet?

Hur ser ordförrådet ut i uppgifter med kvinnlig respektive manlig övervikt i resultatet?

Detta undersökningsmoment har vissa drag gemensamma med det nyss nämnda. Den huvudsakliga skillnaden består i att det här inte rör sig om i förväg klassificerade innehållskategorier, utan om sådana uppgifter – oavsett ämnesområde – där kvinnliga provtagare lyckats klart bättre än manliga, och vice versa.

Grundfrågan är även här vilken roll ordförrådet i de aktuella uppgifterna kan tänkas spela: finns det något generellt samband mellan olika egenskaper hos ordförrådet i en text och de olikheter i resultat som observerats mellan kvinnliga och manliga provtagare? Mer specifikt gäller det huruvida det finns något som skulle kunna betecknas som en ”typiskt kvinnlig” respektive ”typiskt manlig” vokabulärprofil. Om så är fallet, hur skiljer de sig åt beträffande ordfrekvenser och semantiska egenskaper, även i jämförelse med resultatmässigt mer könsneutrala uppgifter?

Några ytterligare undersökningsområden

Förutom de ovan skisserade ordförrädsstudierna kring ELF-provet ligger bl.a. följande undersökningsområden nära till hands, med fokus på såväl texter/frågeapparat som provtagare:

- skillnader i vokabulär mellan *olika text - och uppgiftstyper*: långa texter jämfört med korta texter; frågor med svarsalternativ kontra luckprov;
- ordförrådet i *rubrikerna* – ett slags innehållsdeklaration – till de texter som ligger till grund för olika delprov;
- provtagare med *invandrarbakgrund*: vilka uppgifter/textavsnitt är svårast respektive lättast – vilka vokabulärprofiler?; jämförelse med motsvarande resultat från ”svenska” provtagare.

Slutligen är det av givet intresse att undersöka i vad mån det finns ett samband mellan resultaten på ELF-provet och resultaten på proven i *svensk* ord- och läsförståelse. Det är rimligt att tro att läsförståelse till en betydande del bestäms av en allmän förmåga att tillgodogöra sig innehållet i en text, oavsett språk. För att en sådan förmåga skall kunna fungera i ett givet sammanhang, t.ex. i högskoleprovet, krävs bl.a. ett någorlunda adekvat ordförråd. Eftersom svenska och engelska är närbesläktade språk med ett delvis gemensamt eller snarlikt ordförråd – inte minst inom det mer ”avancerade” lager som är av latinskt/romanskt ursprung – så är det också rimligt att tro att ett relativt stort svenskt ordförråd kan ge en viss draghjälp även i ELF-provet. Omvänt skulle ett begränsat svenskt ordförråd generellt kunna vara ett handikapp även vid engelsk läsförståelse, exempelvis för vissa – men långt ifrån alla – provtagare med invandrarbakgrund. Sådana (eventuella) samband är självfallet värda att undersöka mer i detalj.

Ytterst handlar det om att försöka isolera det specifikt *engelska* i det engelska läsförståelseprovet. Att ordförrådet härvidlag spelar en stor roll är knappast någon vågad gissning. Den närmare karaktären hos denna roll är inte lika självklar.

6 Selektionsmekanismernas betydelse för gruppskillnader på Högskoleprovet

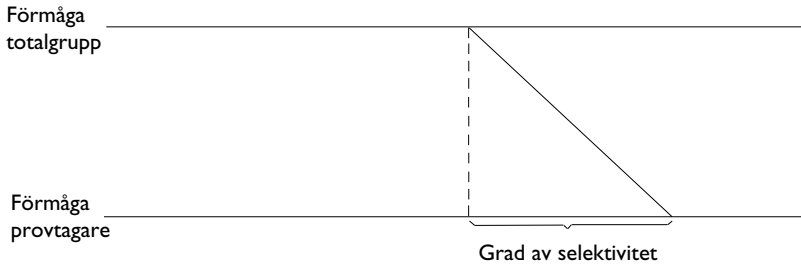
Sven-Eric Reuterberg

Vad är det egentligen som säger att man skall förvänta sig samma genomsnittliga resultat för manliga och kvinnliga högskoleprovsstagare? Provet är ju frivilligt och varje enskild individ avgör själv om han/hon skall genomgå det eller inte. Därmed utgör både de manliga och de kvinnliga provtagarna självselektade grupper. Jämförelser som baseras på sådana grupper måste alltid tolkas med stor försiktighet eftersom man normalt inte vet hur selektionsmekanismerna verkat i det enskilda fallet. Detta förhållande har knappast beaktats i den allmänna debatten. Syftet med denna artikel är att uppmärksamma den inverkan som selektionsmekanismerna kan ha för gruppskillnader i provresultat.

När man under senare år diskuterat gruppskillnader i resultat på Högskoleprovet har diskussionen främst gällt skillnader mellan manliga och kvinnliga provtagare. Sällan har de skillnader i resultat som föreligger mellan provtagare med olika social bakgrund tagits upp till allmän debatt trots att dessa skillnader också är ansenliga (Gustafsson & Westerlund, 1994).

Vad beträffar könsskillnaderna är det ett välkänt faktum att manliga provtagare får ett högre medelvärde på provet än kvinnliga. Enligt Stage och Jarl (1995) har t.ex. männens medelvärden legat mellan 6.6 och 8.2 poäng högre än kvinnornas på de prov som gavs under 1994 och våren 1995. Dessa skillnader till männens fördel har ofta tolkats så i den allmänna debatten att Högskoleprovet på ett orättfärdigt sätt gynnat de manliga provtagarna. Provet har ansetts vara "biased" till männens fördel.

Eftersom Högskoleprovet vänder sig enbart till dem som planerar en högskoleutbildning är det uppenbart att provtagarna utgör en positivt selektad grupp ur hela befolkningen när det gäller sådana egenskaper som mäts av provet. Graden av selektivitet kan då uttryckas som skillnaden i provresultat mellan provtagargruppen och ett representativt urval av alla i befolkningen.

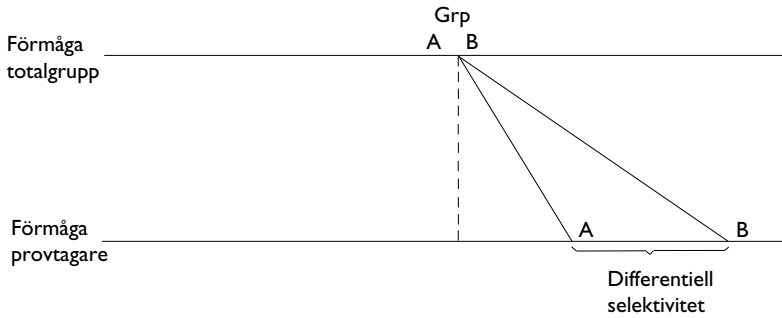


Figur 1. Principmodell för positiv selektion till Högskoleprovet.

Som nämndes tidigare är provet frivilligt och detta innebär också att de som planerar en högre utbildning kan skilja sig åt vad gäller den nytta de kan uppleva sig ha av att genomgå provet. Somliga som fått mycket höga betyg från gymnasieskolan kan känna sig säkra på att bli antagna. De har lite att vinna på att genomgå provet. Konkurrensen om utbildningsplatserna varierar mellan olika utbildningar. Detta kan medföra att två sökande med samma gymnasiebetyg, men som söker till olika utbildningar, kan uppleva olika stark press på sig att ta provet. Den som söker till en utbildning med hård konkurrens bör ha mera att vinna på ett provtagande än den som söker till en utbildning med svagare konkurrens.

När det gäller könsjämförelser visar Wernersson (1991, s. 45) att kvinnorna har ett högre medelbetyg än männen från strängt taget alla gymnasielinjer. När det gäller de 3- och 4-åriga linjerna, från vilka huvudparten av de högskolestuderande rekryteras, är skillnaderna i betygsmedeltal mellan 0.11 och 0.25 betygssteg till flickornas fördel. Enligt SCB (1989) är de sociala skillnaderna i betyg ännu större än könsskillnaderna. Samtidigt föreligger det också klara skillnader såväl mellan män och kvinnor som mellan socialgrupper vad gäller val av högskoleutbildning (Erikson & Jonsson, 1993).

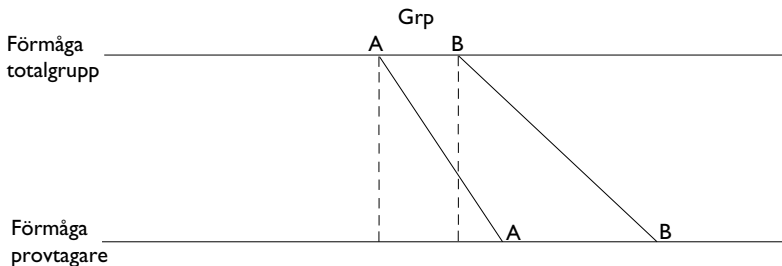
Mot denna bakgrund är det inte orimligt att förvänta sig att selektionsmekanismerna till högskoleprovet varierar mellan kön och mellan olika sociala grupper. Om så är fallet har vi inte enbart en generell positiv selektion till Högskoleprovet utan också en **differentiell selektion**, vilket innebär att en grupp kan vara starkare positivt selekterad än en annan grupp.



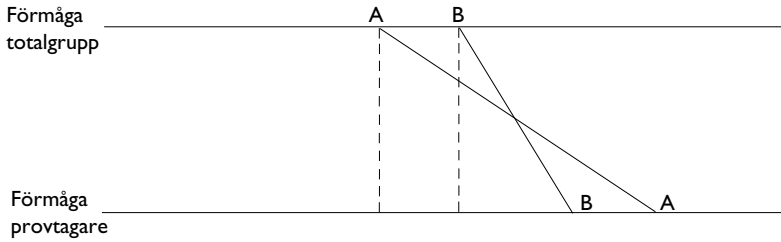
Figur 2. Exempel, där differentiell selektion till högskoleprovet leder till grupp-skillnader i provresultat trots att grupperna ej skiljer sig i totalbefolkningen.

I detta fall finns inga skillnader mellan grupperna i totalbefolkningen, men den differentiella selektionen gör att de provtagare som tillhör grupp B kan förväntas få ett högre resultat på provet än de som tillhör grupp A eftersom den förstnämnda gruppen är starkast positivt selekterad.

Bilden kan givetvis göras mer komplicerad än vad figur 2 visar såtillvida att det också kan finnas skillnader mellan grupperna i totalbefolkningen, vilka antingen förstärks av den differentiella selektionen så som i figur 3 eller blir omvända på grund av den differentiella selektionen (figur 4).



Figur 3. Exempel, där differentiell selektion förstärker den skillnad som finns mellan två grupper i totalbefolkningen.



Figur 4. Exempel, där differentiell selektion leder till att skillnaden mellan två grupper i totalbefolkningen blir den omvända bland provtagarna.

När det gäller selektionsmekanismerna till Högskoleprovet är kunskaperna ännu så länge begränsade, vilket beror på att dessa inte kan studeras utan tillgång till longitudinella data om såväl totalgrupp som provtagare. Dessa data måste dessutom vara relevanta och jämförbara mellan grupperna. Det sistnämnda är ofta ett stort problem eftersom utbildningsvalen sker successivt och de två grupperna går olika vägar genom utbildningssystemet.

Inom ramen för projektet Utvärdering genom Uppföljning (UGU) finns emellertid denna typ av longitudinella data tillgängliga för stora och riksrepresentativa stickprov (Härnqvist et al., 1994). För dem som är födda 1972 har UGU-data samkörts med högskoleprovsuppgifter för samtliga provtillfällen från och med våren 1990 till och med hösten 1992. De sistnämnda uppgifterna är hämtade från databasen Betyg, inskrivningsprov och högskoleprov (Binh-72) (Gustafsson & Westerlund, 1994). Bland dessa sammanslagna uppgifter ingår:

- Bakgrundsvariabler såsom kön och social bakgrund.
- Tidigare utbildning på grundskolans högstadium och gymnasieskolan
- Prestationsvariabler såsom betyg, begåvningsmått och resultat på ett antal standardiserade kunskapsprov i olika ämnen
- Högskoleprovresultat från de ovan nämnda provtillfällena

Selektionen bland kvinnor respektive män

På basis av dessa data har hittills två undersökningar presenterats rörande selektionsmekanismer till Högskoleprovet. I den ena undersökningen har Mäkitalo (1994) granskat den differentiella selektionen bland manliga och kvinnliga provtagare och som exempel på resultat redovisas i tabell 1 selektionseffekterna samt den differentiella selektionen mätt med medelbetyg från grundskolans årskurs 9 samt med det begåvningsstest som bjöds i anslutning till UGU:s datainsamling i årskurs 6. Begåvningsstestet omfattar tre deltest: språkligt, numeriskt-logiskt samt spatialt test.

Selektionseffekterna har här mätts genom att beräkna medelvärdeskillnaden mellan högskoleprovstagarna och övriga inom respektive kön. Differensen har sedan standardiserats genom att dividera skillnaden uttryckt i råpoäng med standardavvikelsen för respektive variabel. Den differentiella selektion utgörs av skillnaden mellan kvinnornas och männens selektionseffekter.

Tabell 1. Selektion och differentiell selektion till Högskoleprovet i relation till kön. (Standardiserade medelvärdesdifferenser mellan provtagare och övriga).

	Män	Kvinnor	Differentiell selektion
Medelbetyg åk 9	1.33	1.08	0.25
Testresultat	0.99	0.83	0.16

De positiva talen i tabellen ovan indikerar att såväl manliga som kvinnliga högskoleprovstagare utgör en positivt gallrad grupp ur sin respektive totalgrupp. Såväl vid analys av betygen som av resultatet av begåvningsstestet framträder således en selektionseffekt enligt den modell som visats i figur 1 ovan. Medelbetygen för de manliga provtagarna ligger ett och en tredjedels betygssteg över de män som inte gjort provet. Bland kvinnorna är skillnaden mellan de båda grupperna drygt ett steg. Även testresultaten visar en motsvarande selektionseffekt, om än något svagare.

Vi har också klara differentiella selektionseffekter, (jfr fig 2) vilka anger att de manliga provtagarna utgör en starkare positivt selekterad grupp bland samtliga män än vad de kvinnliga provtagarna är bland samtliga kvinnor. Mätt med medelbetyg utgör den differentiella selektionen 0.25 och enligt testresultaten uppgår den till 0.16.

Betyder då detta att de manliga provtagarna också har högre medelbetyg och högre resultat på testet än de kvinnliga provtagarna? Den frågan ger selektions-effekterna inget svar på. Som visats i figurerna 3 och 4 måste vi också ta hänsyn till de skillnader som finns inom totalgruppen. I tabell 2 redovisas de faktiska skillnaderna i dessa två variabler dels för provtagarna, dels för totalgruppen.

Tabell 2. Betyg och resultat på testet i relation till kön för provtagare respektive totalgrupp. Medelvärden.

	Provtagare		Totalgrupp	
	Män	Kvinnor	Män	Kvinnor
Medelbetyg	3.88	3.99	3.15	3.44
Testsumma	83.44	80.70	70.44	70.55

De manliga provtagarna har, som framgår av tabellen, inte bättre betyg än de kvinnliga men väl bättre resultat på begåvnings-testet. Den differentiella selektionen ur totalgrupperna män respektive kvinnor har medfört att kvinnornas stora försprång när det gäller betyg har reducerats betydligt, och när det gäller testresultat har de små könsskillnaderna i totalgruppen förvandlats till en skillnad till männens fördel i provtagargruppen.

Selektionen bland sociala grupper

Den andra undersökningen, baserad på samma datamaterial, har utförts av Reuterberg (1994) och i den studeras selektionseffekter bland provtagare från olika sociala grupper. Metodiken är den samma som i Mäkitalos studie.

Tabell 3. Selektion till Högskoleprovet i relation till social bakgrund. (Standardiserade medelvärdesdifferenser mellan provtagare och övriga).

	Socialgrupp		
	I	II	III
Medelbetyg åk 9	0.89	1.06	1.28
Testresultat	0.66	0.73	1.00

Även inom de olika sociala grupperna utgör provtagarna genomgående en klart positivt gallrad grupp och särskilt gäller detta om selektionseffekterna mäts med betygen. Det föreligger också en mycket klar trend såtillvida att

dessa effekter ökar successivt när vi går från socialgrupp I över till grupp III. Följaktligen kan vi också tala om mycket tydliga differentiella selektions-effekter. Ju lägre socialgrupp desto starkare positivt selekterade är prov-tagarna.

Eftersom såväl medelbetyg som testresultat uppvisar klara sociala skillnader i totalgruppen med högst medelvärde bland dem i grupp I och lägst bland dem i grupp III, betyder resultaten i tabell 3 att den differentiella selektionen gör gruppkillnaderna mindre bland provtagarna än de är i totalgruppen. Detta betyder emellertid inte att de sociala skillnaderna helt uttraderats bland dem som tar Högskoleprovet, vilket framgår av tabell 4. Bland dem som gör högskoleprovet har således fortfarande socialgrupp I högst betyg och högre medelvärde på begåvningsstestet, medan grupp III har de lägsta medel-värdena.

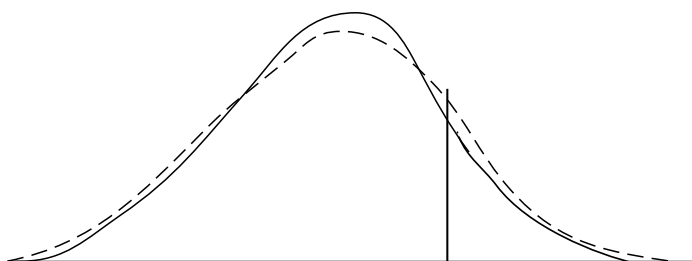
Tabell 4. Betyg och testresultat i relation till social bakgrund bland provtagarna. Medelvärden.

	Socialgrupp		
	I	II	III
Medelbetyg åk 9	4.02	3.91	3.84
Testresultat	83.75	81.39	80.37

Effekter av spridning inom grupperna

Så här långt har diskussionen förts på en beskrivande nivå när det gäller selektionsmekanismerna. Med tillgång till relevanta data kan såväl de generella som de differentiella selektionseffekterna mätas på det sätt som beskrivits. Om man emellertid vill komma ett steg vidare och förstå på vilket sätt selektionsmekanismerna verkar är medelvärdesjämförelser av det slag som beskrivits ovan inte tillräckliga. Då måste vi nämligen också beakta eventuella skillnader grupperna emellan i variabilitet.

Antag att vi har två totalgrupper med exakt samma medelvärde på den variabel, med vars hjälp vi mäter selektionseffekterna. Grupperna skiljer sig emellertid åt vad gäller variation i resultat på variabeln i fråga. Från vardera totalgruppen utväljs en positivt selekterad delgrupp som går till Högskoleprovet så som visas i figur 5. Avgränsningen sker här vid samma värde för båda grupperna.

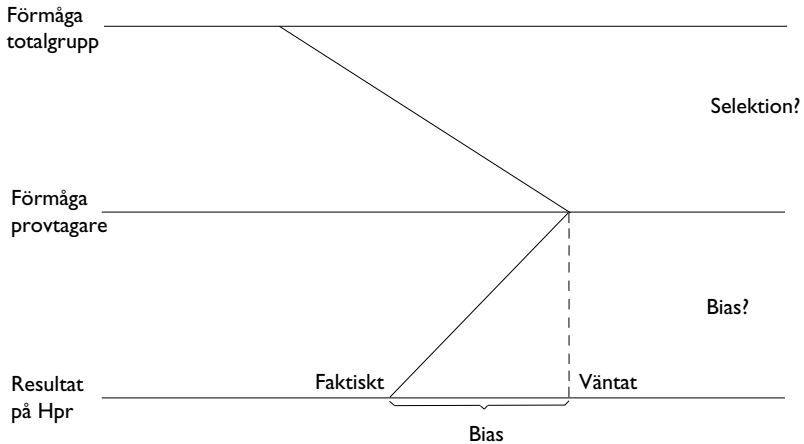


Figur 5. Variabilitetens betydelse för selektiva urval.

Eftersom nu den mest variabla gruppen har flera individer på de högsta nivåerna (och flera på de lägsta) kommer vi att få vissa skillnader mellan grupperna bland de positivt selekterade. För det första får vi med flera individer från den mest variabla gruppen i det positiva urvalet. Dessutom kommer vi att få en medelvärdeskillnad grupperna emellan bland de positivt gallrade såtillvida att de från den mest variabla gruppen får högst medelvärde. Vi får alltså en differentiell selektionseffekt enbart beroende på dessa gruppkillnader i variabilitet. Detta kan betraktas som en statistisk effekt och den bör givetvis separeras ut från de effekter, vilka härrör från faktorer verksamma i den situation, då individen fattar sitt beslut om ett eventuellt provtagande. En metod för att kontrollera denna statistiska effekt har utvecklats av Reuterberg och Mäkitalo (1996).

Rättvisa skillnader?

Studiet av den differentiella selektionen är emellertid enbart ett första steg i den analys som skall ge svar på frågan om gruppkillnader i högskoleprovresultat ligger på en rimlig nivå eller om det finns "bias" i provresultat, dvs om det finns otillbörliga skillnader. Som framgår av figur 6 innebär nästa steg att vi med ledning av provdeltagarnas faktiska förutsättningar skattar det medelvärde en viss grupp skulle uppnå under förutsättning av ett fullständigt rättvist (unbiased) prov. Slutligen jämförs gruppen faktiska medelvärde med detta förväntade värde. Om skillnaden då är positiv kan provet sägas gynna gruppen i fråga; är skillnaden däremot negativ missgynnas gruppen av provet. För ett unbiased prov skall det förväntade och faktiska resultatet sammanfalla.



Figur 6. Principmodell för mätning av "bias".

Avslutningsvis förtjänar att påpekas, att den differentiella selektionen till Högscoleprovet har betydande effekter och att dessa självfallet måste beaktas då man tolkar gruppskillnader i provresultat. På grund av männens starkare selektion till provet har de manliga provtagarna ett högre testresultat än de kvinnliga, men samtidigt kvarstår en skillnad till de kvinnliga provtagarnas fördel vad gäller medelbetyg från årskurs 9. Om man väljer testresultaten som en prognosvariabel för resultat på högscoleprovet skulle vi vänta oss en viss skillnad till männens fördel på provet. Väljer vi däremot betyg som prognosvariabel skulle vi förvänta oss att de kvinnliga provtagarna presterade högst resultat. Vi har följaktligen två frågor att besvara i detta sammanhang:

- Vilket mått är lämpligast att utgå ifrån då vi skall beräkna det förväntade provresultatet?
- Går gruppskillnader i detta mått i samma riktning som de i provresultat och är de i så fall så stora att de kan förklara skillnaderna i provresultat?
- När det gäller socialgruppskillnaderna visar ju såväl betyg som testresultat gruppskillnader av samma natur som de i högscoleprovsresultat. Frågan är i detta fall om skillnaderna i betyg respektive testresultat är så stora att de motiverar de sociala skillnaderna i högscoleprovsresultat.

Detta är frågor som kräver ytterligare analyser för att kunna besvaras.

Referenser:

Erikson, R., & Jonsson, J. O. (1993). *Ursprung och utbildning. Social snedrekrytering till högre studier*. SOU 1993:85. Stockholm: Utbildningsdepartementet.

Gustafsson, J-E., & Westerlund, A. (1994). Socialgruppskillnader i prestationer på Högskoleprovet. I R. Erikson & J. O. Jonsson (Red.), *Sorteringen i skolan* (s 264-284). Stockholm: Carlssons Bokförlag.

Härnqvist, K., Emanuelsson, I., Reuterberg, S-E., & Svensson, A. (1994). *Dokumentation av projektet "Utvärdering genom Uppföljning"*. (Rapport 1994:03). Göteborg: Göteborgs universitet, Institutionen för pedagogik.

Mäkitalo, Å. (1994). *Non-comparability of female and male admission test takers*. (Report 1994:06). Göteborg: Göteborgs universitet, Institutionen för pedagogik.

Reuterberg, S-E. (1994). *Selection of Swedish Scholastic Aptitude Test takers as a function of socioeconomic background and ability*. (Report 1994:05). Göteborg: Göteborgs universitet, Institutionen för pedagogik.

Reuterberg, S-E., & Mäkitalo, Å. (1996). *Who takes the SweSAT?* (Report 1996:03) Göteborg: Göteborgs universitet, Institutionen för pedagogik.

SCB (1989). *Social skiktning i utbildningsväsendet*. (Sveriges officiella statistik. Siffror om högskolan, 8). Stockholm: Statistiska centralbyrån.

Stage, C., & Jarl, C. (1995). *Högskoleprovet våren 1995. Provdeltagargruppens sammansättning och resultat*. (PM 102). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.

Wernersson, I. (1991). *Könsskillnader i gymnasieskolan. En kunskapsöversikt*. (F91:2 Vad säger forskningen?). Stockholm: Skolöverstyrelsen.

7 Högskoleprovet, social bakgrund och upprepad provtagning

Jan-Eric Gustafsson & Maria Benjégard

Den sociala bakgrunden påverkar resultaten på högskoleprovet i mindre omfattning än betyget. Men genom att genomgå högskoleprovet flera gånger förbättras provresultaten något på grund av träningseffekter. Detta utnyttjas av olika grupper i olika omfattning. Barn från tjänstemannahem t ex gör provet flera gånger än andra. Detta kan bidra till ökade skillnader i provresultat mellan studerande från olika social bakgrund.

Introduktion

Systemet för antagning av studerande har tillmätts en betydande roll i strävandena att minska den sociala snedrekryteringen till högre utbildning och det är uppenbart att förhoppningar även knutits till Högskoleprovets möjligheter i detta avseende. Införandet av Högskoleprovet var sålunda delvis motiverat av syftet att minska den sociala snedrekryteringen till högre utbildning. I sitt slutbetänkande (SOU 1970:21 och 55) betonade Kompetensutredningen att urvalssystemets struktur har betydelse för rekryteringen av studerande ur olika samhällskategorier samt att:

Urval enbart på grundval av betyg tenderar att gynna högre socialgrupper. Barn från högre socialgrupper har genomsnittligt högre betyg, även om intelligensen, mätt med intelligenstest, hålls konstant. (SOU 1970:21, sid 265)

År 1983 tillsattes Tillträdesutredningen och i utredningsdirektiven framhölls att enklare tillträdes- och urvalsregler bör eftersträvas och återigen betonades målet att tillträdesreglerna skall bidra till en jämnare social rekrytering till högskolan. Tillträdesutredningen föreslog (SOU 1985:87) att en viss andel (1/3-2/3) av platserna på allmänna utbildningslinjer tillsätts på grundval av medelbetyg från gymnasieskolan och att återstående platser tillsätts på grundval av poäng för Högskoleprov jämte arbetslivserfarenhet. Bland argumenten för denna kraftigt vidgade användning av Högskoleprovet framhölls att:

Möjligheten för alla att bli antagna på ett prov som man kan genomgå hur många gånger som helst, gör att man kan förbättra sin studieförmåga.

Provrvalet tjänar också på det sättet ett viktigt syfte genom att beroendet av betygen som det enda och en gång för alla givna måttet på studieförmåga försvinner (SOU 1985:57, s 20).

I stora drag kom de nya antagningsreglerna, vilka trädde i kraft från höstterminen 1991, att följa Tillträdesutredningens förslag.

Gustafsson och Westerlund (1994) undersökte om socialgruppskillnaderna är mindre i resultat på Högskoleprovet än i betyg. Studien visade att det inte finns några stora skillnader i sambandet mellan social bakgrund och betyg jämfört med sambandet med Högskoleprov, även om det fanns en liten skillnad i den förväntade riktningen. En mer detaljerad analys, där samtidig hänsyn togs till betyg, Högskoleprovsresultat, gymnasielinje och kön, visade emellertid att ungdomar med en eller två föräldrar ur gruppen högre tjänstemän m fl. hade ett betygsmedelvärde som var 0,14 enheter högre än förväntat. Denna skillnad är relativt stor, och kan jämföras med medelvärdeskillnaden i betyg till förmån för kvinnor, som i denna analys uppgick till 0,20. För övriga kategorier i den analyserade socialgruppsvariabeln förelåg inga, eller mycket små, skillnader i betyg då hänsyn togs till provresultat.

Denna analys visar sålunda att socialgruppskillnader är mindre för Högskoleprovet än för betygen. Jämfört med mer renodlade intelligenstest med en större andel icke-verbala uppgifter, som t ex de prov som ingår i Värnpliktsverkets Inskrivningsprov ger dock Högskoleprovet större socialgruppskillnader (Gustafsson & Westerlund, 1994). Det tycks sålunda som om de kunskaper och färdigheter som Högskoleprovet mäter påverkas av hemmiljön i nästan samma grad som de kunskaper och färdigheter som återspeglas i gymnasiebetygen. En förklaring till detta kan vara den rikliga förekomsten av uppgifter med verbalt och numeriskt innehåll i Högskoleprovet.

Gustafsson och Westerlund (1994) genomförde sina analyser på det högskoleprov som gavs våren 1991 med studerande födda 1972 som undersökningsgrupp. Vissa studerande hade tagit högskoleprovet hösten 1990 och en analys gjordes av effekterna av att tidigare ha genomgått Högskoleprovet hösten 90 på resultaten på Högskoleprovet våren 91. Denna analys visade på en ökning med i medeltal ca 3,4 råpoäng.

Upprepad provtagning och övning

Det faktum att det är möjligt förbättra resultaten på Högscoleprovet kan ses som något positivt och av citatet ovan från Tillträdesutredningens betänkande framgår att möjligheterna till upprepad provtagning och förbättring av resultat var ett av huvudargumenten för införande av Högscoleprovet. Om det emellertid finns ett samband mellan den sociala bakgrunden och omfattningen av det upprepade provtagandet kommer en social orättvisa att byggas in i urvalssystemet med successivt ökande skillnader till förmån för de grupper som i stor utsträckning utnyttjar möjligheterna till upprepat provtagande.

Det bör av uppenbara skäl inte vara möjligt att i hög grad påverka resultatet på ett urvalsprov genom en kort, mot själva provet, inriktad träning eller genom upprepad provtagning. Det finns ingen anledning att förvänta sig att Högscoleprovet och andra liknande instrument skall vara resistent mot övnings- och träningseffekter, men det finns inte heller någon grund för en stor optimism vad gäller möjligheterna att på kort tid åstadkomma dramatiska förbättringar. Anledningen till detta är att proven i stor utsträckning mäter kunskaper och färdigheter som tillägnats under lång tid, vilket gör att övning under kort tid knappast kan förväntas ha mer än marginell effekt.

Internationellt har en stor mängd studier genomförts av effekter av träning på urvalsprov (se Bond, 1989; Gustafsson & Undheim, 1996). Träning kan ske på många olika sätt, såsom upprepad provtagning, träning i allmän testförtroenhet, övning i att lösa ett visst slag av uppgifter, eller undervisning i ett innehåll som kan förekomma på provet. Studier av den amerikanska motsvarigheten till högscoleprovet (SAT) visar att övning har effekter, även om dessa är relativt blygsamma. Messick and Jungeblut (1981) demonstrerade att effekterna är icke-linjärt relaterade till övningens omfattning på så sätt att störst effekt nås initialt, och att tillskotten sedan succesivt avtar. Svenska studier av upprepad provtagning ger en likartad resultatbild. Henriksson et. al. (1985) redovisade en ökning av i medeltal ca 2,5 råpoäng för de provdeltagare som genomförde två prov under perioden våren 1983 - hösten 1984, men endast marginella ökningar som en följd av ytterligare upprepning av provet. De effekter som upprepad provtagning medför går givetvis att uppnå även på andra sätt, och det kan antas att ungefär samma inlärningsprinciper är gällande även när provtagning tränas i mer eller mindre organiserade former.

Social bakgrund och upprepad provtagning

Det faktum att effekterna av den upprepade provtagningen är relativt små och tycks vara begränsade till den första omtagningen av provet innebär att riskerna för att detta skulle utgöra en källa till sociala orättvisor inte framstår som överhängande. Det har dock förelegat betydande svårigheter att med precision studera effekterna av upprepad tagning av högskoleprovet eftersom den grupp som tar provet två eller flera gånger kan misstänkas skilja sig från andra provtagare i olika hänseenden. Det är sålunda rimligt anta att de som tar provet flera gånger har högre motivation för högre studier än de som endast tar provet en gång, men också att de som upprepar provet tenderar att ha lägre resultat vid det initiala provtillfället. Eftersom en grupp "normala" provtagare reduceras med de som blir antagna till högre studier finns även här en anledning till bristande jämförbarhet mellan de som tar provet många gånger och de som tar det få gånger.

Den år 1991 genomförda regeländring som gjorde det möjligt även för studerande på gymnasienivå att ta provet har emellertid som konsekvens både att problemet med upprepad provtagning får större betydelse och att möjligheterna att studera dess effekter ökar. Mängden data som kan nyttjas för att studera effekter av upprepad provtagning har sålunda ökat dramatiskt. Gymnasistgruppen är inte heller utsatt för den systematiska avtappning genom antagning till högre studier som gäller för äldre studerandegrupper.

I anslutning till forskning om högskoleprovet vid Göteborgs universitet har databaser byggts upp som omfattar samtliga individer födda 1972 och 1973. Databaserna är byggda bl a från register vid SCB och omfattar bland mycket annat uppgifter om resultat på högskoleproven givna från våren 1990 till hösten 1992. Vissa analyser har genomförts av dessa data i syfte att undersöka dels i vilken utsträckning det föreligger ett samband mellan social bakgrund och frekvens av provtagning, dels om upprepad provtagning leder till bättre provresultat. Resultaten från dessa analyser redovisas kortfattat nedan.

Elever födda 1972 lämnade normalt grundskolan våren 1988 och treåriga gymnasiestudier 1991. Under gymnasiets åk 2 hade sålunda elever i denna grupp möjlighet att ta provet våren 90, under åk 3 hösten 90 och våren 91, samt efter gymnasiet (utom för elever i fyraårig utbildning) hösten 91, våren 92 och hösten 92. Enligt samma resonemang kunde elever födda 1973 ta

provet våren 90 under åk 1, våren och hösten under åk 2 och 3, samt hösten 92 efter avslutade gymnasiestudier. Tabell 1 redovisar för båda årskullarna medelvärdet av antalet genomgångna prov för elever från olika social bakgrund. Elever med föräldrar som är tjänstemän på hög- eller mellannivå har ett betydligt högre medelvärde av genomgångna prov än övriga grupper.

Tabell 1. Medelvärde av antal genomgångna prov för respektive kohort och social klass

Social bakgrund	72 (max 6 prov)	73 (max 4 prov)	Antal
Tjänsteman (högre)	1,03	0,90	12 137
Tjänsteman (mellan)	0,79	0,68	30 368
Tjänsteman (lägre)	0,56	0,52	30 767
Egna företagare	0,37	0,36	10 901
Jordbrukare m.fl	0,39	0,36	4 911
Arbetare (kval)	0,31	0,32	29 509
Arbetare (okval)	0,23	0,23	79 105
Uppgift saknas	0,25	0,22	13 856
			211 554

För att studera effekterna av upprepat provtagande på provresultat har en serie regressionsanalyser genomförts, där i tur och ordning provresultat vår och höst 91 -92 utgjort beroendevariabler, med i stort samma teknik som användes av Gustafsson och Westerlund (1994). I Tabell 2 redovisas skattningar av effekten av olika antal tidigare genomgångna prov på prestationer (uttryckt i råpoäng), med alla andra variabler konstanthållna.

Tabell 2. Skattad effekt av tidigare provtagning på provresultat

Antal prov	Vt 91	Ht 91	Vt 92	Ht 92
1 tidigare prov	3,56	3,79	4,01	2,80
2 tidigare prov	4,69	6,24	6,37	5,49
3 tidigare prov		7,04	7,67	6,67
4 eller fler tidigare prov				6,97

Som framgår av tabellen är skattningarna av effekten av tidigare provtagning sådan att varje ytterligare genomgången prov ger en förbättring av provresultatet, även om den positiva effekten minskar. Både 1 och 2 tidigare prov ger emellertid ett betydande tillskott till resultatet. Dessa resultat avviker

sålunda från vad som framkommit i den tidigare svenska forskningen, enligt vilken endast den första omtagningen har någon positiv effekt.

Om det föreligger ett samband mellan social bakgrund och frekvens av provtagning torde detta leda till ökande prestationsskillnader mellan provtagare med olika social bakgrund. Denna effekt illustreras i Tabell 3.

Tabell 3. Medelvärden för provtagare med olika social bakgrund

Social bakgrund	Vt 91	Vt 92
Tjänsteman (högre)	10,19	12,00
Tjänsteman (mellan)	6,19	7,26
Tjänsteman (lägre)	3,20	4,24
Egna företagare	2,06	1,89
Jordbrukare m.fl	4,71	2,40
Arbetare (kval)	-0,81	0,63

De i tabellen presenterade värdena är beräknade som skillnader mellan mellan de angivna socialgrupperna och gruppen okvalificerade arbetare. Skillnaderna till förmån för framförallt tjänstemän på hög- och mellannivå ökar från våren 91 till våren 92. Det kan naturligtvis finnas flera förklaringar till detta resultat, men den högre frekvensen av provtagande i högre socialgrupper torde åtminstone vara en bidragande orsak.

Slutsatser

De analysresultat som redovisats här antyder att ett eller två tidigare genomgångna högskoleprov har betydande effekter på provresultat, och även ytterligare genomgångna prov ger en viss förbättring. Tidigare studier av effekter av upprepad provtagning har visat på en positiv effekt av ett tidigare prov, men inte på några effekter av ytterligare prov. En anledning till de kraftigt avvikande resultaten i denna studie kan vara det faktum att undersökningsgruppen här utgörs av gymnasiegrupper vilka inte påverkas av den avtappning av provtagare med höga resultat genom antagning till högre studier som gör sig gällande i andra grupper.

Ett samband har också, icke oväntat, påvisats mellan social bakgrund och frekvens av provtagning. Detta samband är ett uttryck för de skillnader i intresse för högre studier som föreligger mellan studerande från olika social

bakgrund, men är förmodligen också delvis grundat i det faktum att deltagande i högskoleprovet är förknippat med en relativt hög kostnad. Eftersom upprepat provtagande leder till bättre provresultat medför emellertid detta samband ökade prestationsskillnader mellan studerande med olika social bakgrund.

Vid sidan om den systematiska effekt av upprepat provtagande som påvisats här, finns även en ytterligare omständighet som bidrar till att upprepat provtagande kan medföra ökande skillnader mellan provtagare med olika social bakgrund. Den regel som säger att provtagarna har rätt att tillgodoräkna sig sitt bästa provresultat medför nämligen att upprepat provtagande medför ytterligare chanser att erhålla ett gott resultat genom de slumpmässiga mätfel (t ex gissning och uppgiftsurval) som i viss utsträckning vidlåder högskoleprovet. Sådana osystematiska effekter har ingen inverkan på de medelresultat som presenterats här, men på den individuella nivån kan effekten vara betydande. Den systematiska och osystematiska effekten av upprepat provtagande ger sålunda i kombination med sambandet mellan social bakgrund och upprepat provtagande anledning till oro att högskoleprovet snarare kommer att öka än att minska den sociala snedrekryteringen till högre utbildning.

Fortsättningsvis kommer studenterna att kunna tentera upp slutbetyget från gymnasieskolan även om de har godkänt. Mot bakgrund av de ovan redovisade resultaten kring högskoleprovet och olika gruppers intresse för upprepade provtagning reses även frågan om möjligheten att tentera upp betygen kan komma att få samma effekt.

Referenser

Bond, L. (1989). *The effects of special preparation on measures of scholastic ability*. In R. Linn (Ed.), *Educational Measurement* (3rd ed.), (pp. 429-444). New York: Macmillan.

Gustafsson, J.-E. & Undheim, J. O. (1996). *Individual differences in cognitive functions*. In D. C. Berliner & R. C. Calfee (Eds.) *The Handbook of Educational Psychology*, (pp.186-242). New York: Macmillan.

Gustafsson, J.-E., & Westerlund, A. (1994). *Socialgruppskillnader i prestationer på Högskoleprovet*. I R. Erikson & J. O. Jonsson (red.) *Sorteringen i skolan*. Stockholm: Carlssons.

Henriksson, W., Henrysson, S., Stage, C., & Wedman, I. (1985). *Prov för urval till högskolan*. Stockholm: Utbildningsdepartementet, SOU 1985:59.

Messick, S., & Jungeblut, A. (1982). *Time and method in coaching for the SAT*. *Psychological Bulletin*, 89, 191-216.

SOU 1970:21. *Vägar till högre utbildning*. Behörighet och urval. Betänkande av kompetensutredningen, VI:1

SOU 1970:55. *Vägar till högre utbildning*. Organisation och information. Betänkande av kompetensutredningen, VI:2

SOU 1974:71. *Om behörighet och antagning till högskolan*. Betänkande av kompetenskommittén.

SOU 1985:57. *Tillträde till Högskolan*. Betänkande av tillträdesutredningen.

8 Den sociala bakgrundens betydelse för prestationer på Högskoleprovet

Sven-Eric Reuterberg

När högskoleprovet infördes fanns förhoppningar om att provet skulle bidra till att minska den sociala snedrekrytering till högskolan. Nu kan konstateras att det finns betydande sociala skillnader i resultaten på högskoleprovet. Man kan också slå fast att dessa skillnader är mer uttalade bland de kvinnliga provdeltagarna än bland männen. Att därmed hävda att provet missgynnar vissa sociala grupper är emellertid förhastat. Samma sociala skillnader ger också betyg och test av andra slag. Den centrala frågan att besvara är om provet trots allt minskar denna skillnad i samband med antagningen till högskolan.

När Högskoleprovet introducerades för snart 20 år sedan knöts en del förhoppningar till provet om att det skulle vara mindre påverkat av social bakgrund än vad som var fallet med gymnasiebetygen (Wedman & Henrysson, 1992). Under senare år har emellertid de sociala skillnaderna ägnats liten uppmärksamhet. Den allmänna debatten har mera fokuserats på könsskillnaderna i provresultat. En anledning till detta är att man inte vetat lika mycket om de sociala skillnaderna som om könsskillnaderna i provresultat.

Den enda tidigare undersökning som mig veterligen studerat socialgruppskillnaderna är den, som genomfördes av Gustafsson och Westerlund (1994) på uppdrag av Utredningen om den sociala snedrekryteringen till högre studier. Enligt denna undersökning visade resultaten på Högskoleprovet ett något svagare samband med den sociala bakgrunden än vad som var fallet för gymnasiebetygen. Dock visade även provresultaten på klara skillnader mellan provtagare med olika social bakgrund.

I den ovan nämnda undersökningen studerades socialgruppskillnaderna med manliga och kvinnliga provtagare sammanslagna. Här kommer däremot skillnaderna att studeras separat för manliga och kvinnliga provtagare och ett syfte är att granska om dessa skillnader varierar med kön.

De bearbetningar som här skall presenteras är baserade på data hämtade från två olika databaser, nämligen UGU-projektet (Utvärdering genom uppföljning) samt Binh-basen (Betyg, inskrivningsprov och högskoleprov). Den grupp som här är aktuell utgörs av ett tioprocentigt urval av individer födda 1972. För denna grupp finns högskoleprovsresultat tillgängliga för de prov som gavs fr.o.m. våren 1990 t.o.m. hösten 1992. Detta innebär att varje individ har haft möjlighet att ta provet mer än en gång och åtskilliga har också gjort så. För de individer som tagit provet mer än en gång gäller här den högsta provpoäng som individen uppnått. Detta innebär att skillnaderna mellan olika grupper inkluderar också sådana effekter som kan tillskrivas övning genom upprepat provtagande.

Socialgruppsindelningen innefattar följande tre kategorier:

- I: Högre tjänstemän och storföretagare
- II: Övriga tjänstemän och småföretagare
- III: Arbetare

Skillnaderna mellan de sociala grupperna uttrycks i standardiserade provpoäng (z), vilket innebär att skillnaderna dividerats med standardavvikelsen för totalgruppen. Genomgående gäller att socialgrupp III utgör referensgrupp. De standardiserade differenserna anger följaktligen hur stor skillnaden är mellan grupp I och III respektive mellan grupp II och III.

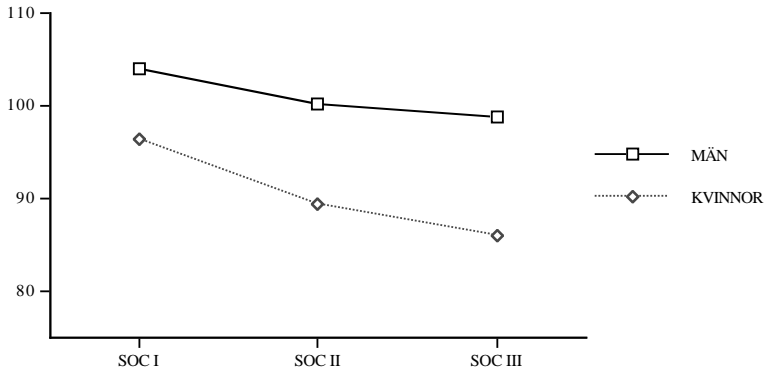
Tabell 1. Socialgruppskillnader i totalpoäng på Högskoleprovet uttryckt i standardiserade provpoäng (z).

Jämförelse	Män	Kvinnor	Totalt
I - III	0.28*	0.56*	0.50*
II - III	0.08	0.18*	0.16*

* signifikant skillnad på 5%-nivån

Om vi först ser på totalgruppen, uppgår skillnaden i provresultat mellan socialgrupp I och III till en halv spridningsenhet vilket motsvarar 8 à 9 provpoäng. Framför allt är det socialgrupp I som avviker från de övriga. Skillnaden mellan grupp II och III är nämligen klart mindre, 0.16 spridningsenheter.

Om man studerar socialgruppskillnaderna för män och kvinnor var för sig, finner man att de är dubbelt så stora bland kvinnorna jämfört med bland männen. Av figur 1 nedan kan vi utläsa att det framförallt är kvinnorna i grupp III som har presterat låga provresultat.



Figur 1. Sociala skillnader i total provpoäng bland manliga och kvinnliga provtagare.

Hur stabila är då socialgruppskillnaderna över olika provtillfällen? För att få en viss uppfattning om detta har socialgruppskillnaderna beräknats för vart och ett av provtillfällena fr.o.m. hösten 1990 t.o.m. hösten 1992. (Vid det första provtillfället våren 1990 togs provet av mycket få individer i vårt stickprov.)

Tabell 2. Socialgruppskillnader vid olika provtillfällen uttryckt i standardiserade provpoäng (z).

Jämförelse	Hösten 1990	Våren 1991	Hösten 1991	Våren 1992	Hösten 1992
I - III	0.40*	0.45*	0.47*	0.41*	0.63*
II - III	0.02	0.12	0.20	0.08	0.20
Antal	589	1171	547	902	409

Som tabellen visar, är mönstret det samma hela tiden även om vi har en viss variation i skillnadernas storlek. Skillnaderna mellan grupp I och III är genomgående signifikanta (markerade med *) vilket inte är fallet för dem mellan grupp II och III. Socialgruppskillnaderna i prestationer är således tämligen stabila över tiden.

Delproven

De resultat som hittills redovisats har gällt totalpoäng på Högskoleprovet. I tabell 3 granskas socialgruppskillnaderna för vart och ett av delproven. De största skillnaderna visar ELF-, NOG- och AO-proven och minst skillnader finner vi för studiefärdighetsprovet (STUF), som nu utgått ur Högskoleprovet. Av stort intresse är att konstatera att AO uppvisar förhållandevis stora socialgruppskillnader. Detta delprov kommer ju att utgå ur Högskoleprovet fr.o.m. våren 1996 och sett ur socialgruppsperspektiv är detta således knappast någon nackdel.

Tabell 3. Socialgruppskillnader på de olika delproven uttryckt i standardiserade provpoäng (z)

Jämförelse	ELF	NOG	AO	LÄS	ORD	DTK	STUF
I - III	0.50*	0.49*	0.46*	0.40*	0.35*	0.39*	0.32*
II - III	0.15	0.22*	0.17*	0.14*	0.06	0.18*	0.13

I figur 1 tidigare har kunnat konstaterats att männen i socialgrupp I har de högsta poängen och kvinnorna i socialgrupp III de lägsta. Gäller då detta för samtliga delprov eller är skillnaderna framför allt kopplade till vissa delprov.

Tabell 4. Socialgruppskillnader i relation till kön på ELF-, NOG- och AO-provet uttryckt i standardiserade provpoäng (z).

	ELF		NOG		AO	
	Män	Kvinnor	Män	Kvinnor	Män	Kvinnor
I - III	0.42*	0.52*	0.33*	0.50*	0.28*	0.50*
II - III	0.21	0.06	0.17	0.22*	0.05	0.23*

Tabell 5. Socialgruppskillnader i relation till kön på LÄS-, ORD-, DTK- och STUF-provet uttryckt i standardiserade provpoäng (z).

	LÄS		ORD		DTK		STUF	
	Män	Kvin.	Män	Kvin.	Män	Kvin.	Män	Kvin.
I - III	0.22*	0.50*	0.19	0.46*	0.16	0.42*	0.07	0.45*
II - III	0.01	0.21*	0.03	0.06	0.07	0.21*	-0.03	0.22*

Tabellerna 4 och 5 visar på ett mycket enhetligt mönster såtillvida att på varje delprov är socialgruppskillnaderna större bland kvinnliga provtagare än bland manliga. Även på delprovsnivå förefaller det vara kvinnorna i socialgrupp III som får låga resultat. Genomgående har dessa ett signifikant lägre medelvärde än kvinnor från grupp I och skillnaderna uppgår till mellan 0.4 och 0.5 standardavvikelseenheter. På fem av de sju delproven presterar kvinnor i grupp III också signifikant lägre resultat än kvinnor i grupp II.

Bland de manliga provtagarna går medelvärdesjämförelserna mellan grupperna I och III genomgående till grupp I-s fördel, men för tre av delproven i tabell 5 är denna skillnad inte statistiskt säkerställd. Däremot visar ingen av jämförelserna mellan grupperna II och III på signifikanta skillnader även om de flesta jämförelser går till grupp II-s fördel.

Sammanfattningsvis kan man således konstatera att det föreligger betydande sociala skillnader i resultat på Högskoleprovet och detta gäller såväl totalt som på delprovsnivå. Vi kan också slå fast att de sociala skillnaderna är mera uttalade bland kvinnliga provtagare än bland manliga.

Man kan då fråga sig om detta innebär att provet missgynnar de provtagare som kommer från socialgrupp III och då särskilt grupp III-s kvinnor. Som diskuterats i avsnittet "Selektionsmekanismernas betydelse för gruppskillnader på Högskoleprovet" kan en sådan slutsats inte dras på grundval av dessa medelvärdesjämförelser. Provtagare från socialgrupp III är visserligen mera positivt selekterade ur sin totalgrupp än vad som är fallet för provtagare från de båda övriga socialgrupperna, men eftersom det finns betydande sociala skillnader i såväl betyg som testresultat i totalgruppen kvarstår det även bland provtagarna en skillnad till grupp III-s nackdel (jfr. tabell 4 i ovan nämnda avsnitt). Frågan är om denna kvarstående skillnad bland provtagarna är så stor att den kan vara orsak till de sociala skillnaderna i resultat på Högskoleprovet, eller om provet är "biased" till de högre socialgruppernas fördel. Denna fråga finns all anledning att granska närmare i ett senare sammanhang.

Referenser:

Gustafsson, J-E., & Westerlund, A. (1994). Socialgruppskillnader i prestationer på Högskoleprovet. I R. Erikson & J. O. Jonsson (Red.), *Sorteringen i skolan* (s 264-284). Stockholm: Carlssons Bokförlag.

Härnqvist, K., Emanuelsson, I., Reuterberg, S-E., & Svensson, A. (1994). *Dokumentation av projektet "Utvärdering genom Uppföljning"*. (Rapport 1994:03). Göteborg: Göteborgs universitet, Institutionen för pedagogik.

Wedman, I., & Henrysson, S. (1992). Överväganden om könsmässig och social snedrekrytering vid Högskoleprovets tillkomst och utformning. *I Betyg och högskoleprov för män och kvinnor*. (UHÄ-rapport 1992:03). Stockholm: Universitets- och Högskoleämbetet.

9 Kulturfria instrument – finns sådana hybrider?

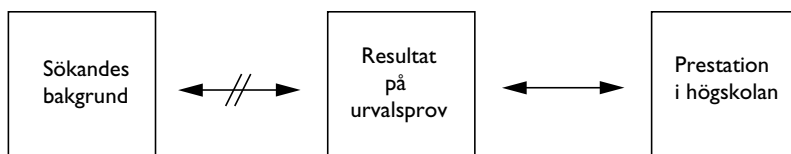
Allan Svensson

Med nuvarande urvalssystem antas till de längre tekniska och naturvetenskapliga utbildningarna en mindre andel ur socialgrupp III är vad som skulle kunna förväntas. Kvinnornas andel av de antagna till dessa utbildningar står i proportion till deras andel av N- och T-linjerna i gymnasieskolan. Om det nya betygssystemet på sikt inte kan användas utan högskoleprovet kommer att utgöra det huvudsakliga urvalsinstrumentet kan kvinnornas möjlighet att konkurrera om de attraktiva platserna försämrats än mer. I stället kan antagning i flera steg vara en alternativ modell.

Är det möjligt att åstadkomma socialt och könsmässigt neutrala urvalsinstrument? Vad skulle hända med den sociala rekryteringen och mäns respektive kvinnors möjligheter om specialinriktade test och intervjuer började användas i stor skala?

Frågan i huvudrubriken har två stora förtjänster. Formuleringen är klatschig (jag har inte hittat på den själv) och frågan är lätt att besvara - med ett tveklöst NEJ.

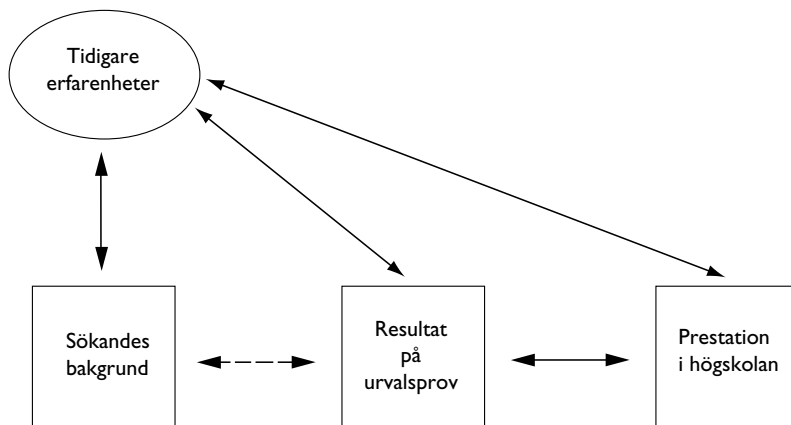
Jag skall dock kortfattat redogöra för, varför sådana hybrider ej existerar och sedan ägna mig åt en av underrubrikerna - den som handlar om specialinriktade test och intervjuer. Vi börjar med att betrakta figur 1.



Figur 1. Resultaten på högskoleprovet samvarierar med prestationerna i högskolan men ej med de sökandes bakgrund

Vad man skulle önska är ett urvalsinstrument, vars resultat har klara samband med prestationerna i högskolan. Däremot bör det inte finnas några samband mellan resultaten på urvalsprovet och de sökandes bakgrund i form av kön och socialgruppstillhörighet.

Svårigheterna uppstår genom att det finns starka samband mellan de sökandes tidigare erfarenheter - i form av tidigare skolprestationer och genomgången gymnasieutbildning - och deras bakgrund. Dessa tidigare erfarenheter är också mycket betydelsefulla för hur man lyckas i högskolan. Vill man ha ett urvalsinstrument som korrelerar med resultaten i högskolan, får man följaktligen också höga samband mellan urvalsinstrumentet och de sökandes bakgrund. Resonemanget åskådliggörs i figur 2.



Figur 2. Tidigare erfarenheter samvarierar såväl med de sökandes bakgrund, resultaten på urvalsinstrument som prestationerna i högskolan, vilket också ger upphov till samband mellan bakgrund och provresultat.

Låt oss då ställa frågan så här - finns det några mera påtagliga samband mellan kön respektive social bakgrund om man kontrollerar för tidigare erfarenheter i form av genomgången gymnasieutbildning. Jag väljer de elever som genomgått N- eller T-linjen. Varför just dessa? Jo, eleverna från dessa linjer erhåller de högsta resultaten på högskoleprovet (Stage, 1993). Härtill kommer att de har den bredaste kompetensen och kan direkt söka till de högskoleutbildningar, som i allmänhet tillhör de mest eftersökta.

I tabell 1 visas hur stor andel av dem som är födda 1972, som fullföljt dessa linjer. Materialet är hämtat från UGU-projektet, ett longitudinellt forskningsprojekt, där det finns utförlig information om stora och riksrepresentativa stickprov av elever (Härnqvist et al., 1994).

Tabell 1. Den procentuella andelen som genomgått N- eller T-linjen bland samtliga män och kvinnor från olika socialgrupper.

Kön	Socialgrupp			Samtliga
	1	2	3	
Män	45	24	9	22
Kvinnor	21	11	5	11
Män och kvinnor	33	18	7	17

Anm. I den socialgruppsindelning som används särhålls följande tre grupper:

1. Högre tjänstemän och storföretagare
2. Övriga tjänstemän och småföretagare
3. Arbetare

Som man kan se av tabellen finns det betydande skillnader såväl mellan könen som mellan socialgrupperna, vad gäller valet av gymnasieskolans N- eller T-linje. Sålunda är det *inom* varje socialgrupp ungefär dubbelt så stor andel bland männen som gjort detta. Skillnaderna *mellan* socialgrupperna är än större - bland eleverna från grupp 1 är det var tredje som skaffat sig den aktuella utbildningen jämfört med var femtonde från grupp 3, d.v.s. sannolikheten är fem gånger högre om man kommer från grupp 1. De adderade effekterna av kön och socialgrupp gör att skillnaderna mellan män från grupp 1 och kvinnor från grupp 3 blir mycket stora - nästan varannan av de förra jämfört med endast var tjugonde av de senare har genomgått naturvetenskaplig eller teknisk linje.

Finns det då några skillnader med avseende på högskolestudier mellan dem som fullföljt N- eller T-linjen? Granskar vi dem, som fram till 22 års ålder påbörjat någon form av NT-utbildning inom högskolan, har drygt hälften gjort detta. Här uppstår det dock *ingen ytterligare skärpning av den könsmissiga och sociala selektionen*. Inom alla undergrupper sker nämligen en halvering av respektive andelar. Jämför man t.ex. elever från grupp 1 och grupp 3 är det således 18 procent av de förra och 4 procent av de senare som går vidare till en NT-utbildning i högskolan - fortfarande är alltså oddsen ungefär fem mot ett (tabell 2).

Tabell 2. Den procentuella andelen som påbörjat någon typ av naturvetenskapliga eller tekniska utbildningar i högskolan bland samtliga män och kvinnor från olika socialgrupper.

Kön	Socialgrupp			Samtliga
	1	2	3	
Mä	22	13	5	12
Kvinnor	12	6	2	6
Män och kvinnor	18	9	4	9

De uppgifter som presenteras i tabell 2 tyder inte på att urvalet till högskolan missgynnade kvinnor eller elever från lägre socialgrupper - under förutsättning att de skaffat sig "rätt" gymnasieutbildning. Vad som krävs är alltså inte förändringar av högskoleprovet. Däremot behöver rekryteringen till NT-utbildningarna i gymnasieskolan förändras. Detta är dock en fråga som jag ej går närmare in på i detta sammanhang utan hänvisar i stället till tidigare gjord undersökning (Svensson, 1995a).

Att högskoleprovet inte skulle behöva förändras, för att säkra en rättvisare fördelning av platserna i högskolan, är dock en sanning med modifikationer. Hittills har vi granskat de NT-elever som har påbörjat *någon typ* av naturvetenskapliga eller tekniska utbildningar i högskolan. Vad händer om vi endast ser till dem som påbörjat *de längre och mer eftersökta utbildningarna*, t.ex. läkar- och civilingenjörsutbildningarna? Svaret finns i tabell 3.

Tabell 3. Den procentuella andelen som påbörjat längre naturvetenskapliga eller tekniska utbildningar i högskolan bland samtliga män och kvinnor från olika socialgrupper.

Kön	Socialgrupp			Samtliga
	1	2	3	
Män	13	5	2	5
Kvinnor	8	3	1	3
Män och kvinnor	11	4	1	4

Om vi endast betraktar de elever som satsat på en längre högskoleutbildning med NT-inriktning, sker totalt sett en ny halvering. Denna slår ganska jämnt vad gäller män och kvinnor - cirka varannan av vardera könet som valt en

NT-utbildning i högskolan har valt en längre sådan. Däremot kan man konstatera en ökad social snedrekrytering. Skillnaderna mellan grupp 1 och 3 växer och sannolikheten är nu tio gånger högre att man skall få en avancerad teknisk eller naturvetenskaplig högskoleutbildning, om man tillhör grupp 1.

Ytterligare en omständighet bör nämnas. De elever som ingår i denna undersökning kunde söka in på högskoleprov eller betyg från gymnasieskolan. De förra gynnar männen och de senare kvinnorna, såtillvida att männen har ett högre medeltal på högskoleprovet, medan kvinnorna uppvisar högre medelbetyg från gymnasieskolan (Reuterberg et al, 1992, s 22).

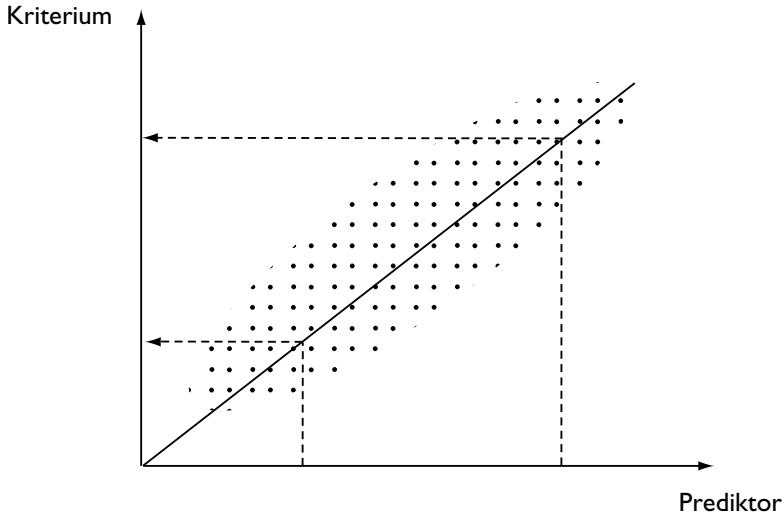
De betyg som undersökningsdeltagarna erhöll var s.k. relativa betyg. Dessa kommer inom kort att ersättas med målrelaterade betyg. Enligt ett förslag från Utbildningsdepartementet är det tänkt att även dessa skall fungera som urvalsinstrument. Detta framgår klart av följande uttalande:

”Enligt arbetsgruppens bedömning bör - utan att ogrundade jämförelser görs med tidigare system - det nya betygssystemet mycket väl kunna användas i högskoleurvalet. Detta bör även i fortsättningen vara den viktigaste urvalsgrunden.”
(Utbildningsdepartementet, 1995a, s 41).

Enligt min bedömning är det dock ytterst tveksamt om de målrelaterade betygen kan fungera i urvalssammanhang. Orsaken till denna tveksamhet är den bristande jämförbarheten mellan betyg satta av olika lärare samt risken för betygsinflation - bland annat kommer ju s.k. konkurrenskomplettering att tillåtas i det nya systemet. För en utförlig diskussion i denna fråga hänvisa till Wedman (1995) och Svensson (1995b).

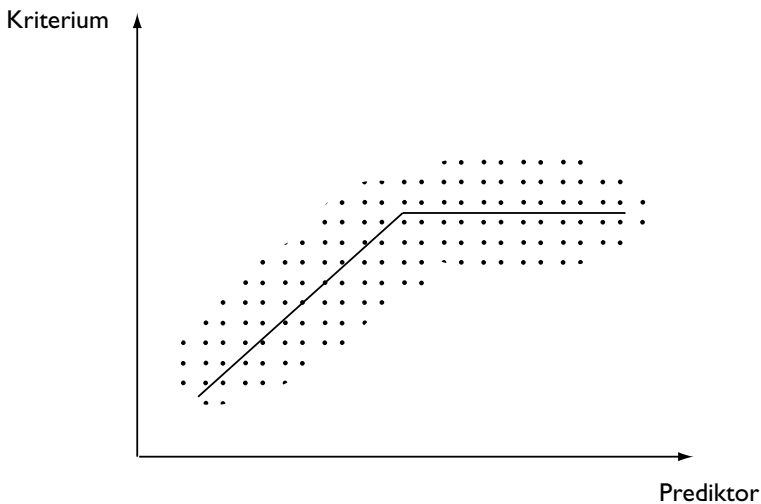
Om mina farhågor besannas skulle enbart högskoleprovet återstå som urvalsinstrument - i varje fall till de mer eftersökta högskoleutbildningarna. Detta skulle i sin tur innebära att kvinnornas möjligheter att få tillträde hit avsevärt minskas. Vid en sådan situation skulle jag vilja tala för användningen av specialinriktade test och intervjuer.

Innan jag närmare beskriver hur dessa skulle kunna genomföras, skall vi granska några figurer. I figur 3 visas grafiskt hur förhållande mellan en prediktor (poäng på högskoleprovet) och ett kriterium (studieframgång i högskolan) ser ut då korrelationskoefficienten är relativt hög - cirka 0.60 - och sambandet *linjärt*. Det senare innebär att ett tillskott i provpoäng i allmänhet åtföljs av en något bättre prestation i högskolan.



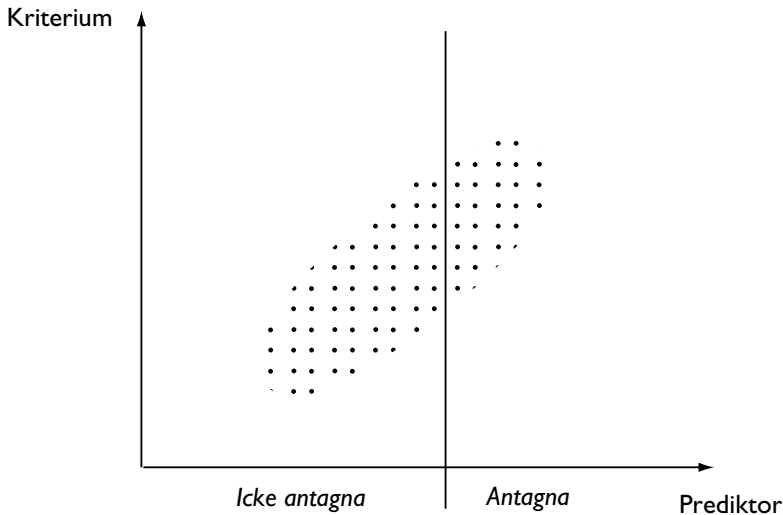
Figur 3. Ett linjärt samband mellan prediktor och kriterium.

Figur 4 åskådliggör ett *icke linjärt* samband, d.v.s. att studieprestationen i genomsnitt förbättras upp till en viss provpoäng, men att en ytterligare ökning av provpoängen inte ger något systematiskt tillskott i studieframgång. Om man kan konstatera ett icke linjärt samband, innebär det således att en extremt hög poäng på högskoleprovet ej har bättre prognosvärde än en något lägre.



Figur 4. Ett icke linjärt samband mellan prediktor och kriterium.

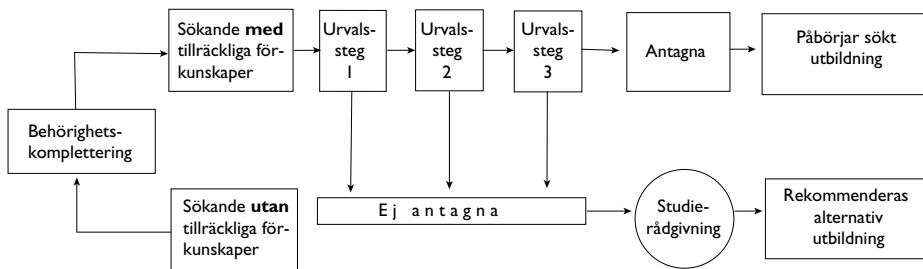
Figur 5 slutligen visar fenomenet *beskuren spridning*, d.v.s. att det råder ett relativt starkt samband mellan prediktor och kriterium, men på grund av att man endast antar dem som ligger över en viss nivå, kan sambandet mellan resultatet på provet och prestationerna i högskolan bli mycket lågt eller t.o.m. obefintligt. Problemet med beskuren spridning i prediktorn växer, ju mindre andel av de sökande som antas till en utbildning.



Figur 5. Sambandet mellan prediktor och kriterium reduceras genom att endast de över ett visst värde i prediktorn väljs ut.

Om sådana förhållanden som illustrerats i figur 4 eller 5 råder, skulle jag föreslå att man går över till ett stegvis antagningsförfarande, varvid högskoleprovet kompletteras med specialinriktade test och eventuellt med intervjuer.

I ett första steg används högskoleprovet som urvalsinstrument. Härfter låter man ett relativt stort antal sökande få gå vidare till ett andra steg, där särskilda konstruerade test sätts in. Så kunde man exempelvis vid urvalet till civilingenjörsutbildningarna välja prov av matematisk karaktär och vid urvalet till vissa humanistiska utbildningar s.k. produktiva prov (uppsatser). Till de mest eftersökta utbildningarna kan man dessutom tänka sig ett tredje steg, varvid noggrant strukturerade intervjuer får bli det avgörande urvalsmåttet. Det skisserade förslaget åskådliggörs i figur 6.



Figur 6. Förslag till nytt antagningsförfarande.

Vilka vinster skulle detta förfarande kunna ge?

1. I och med att högskoleprovet kompletteras med urvalsinstrument som mäter delvis andra förmågor än allmän studielämplighet - förmågor som är av speciell vikt för den sökta utbildningen bör den prognostiska validiteten förbättras. Detta borde i sin tur innebära att studietiderna förkortas och antalet studieavbrott minskar. För närvarande är situationen i dessa avseenden ingalunda tillfredsställande. Så t.ex. var det knappt 60 procent av dem som påbörjade en utbildning till civilingenjör i slutet av 80-talet som avlagt examen efter sju år (Statistiska centralbyrån, 1995).
2. Sökande med god studiebegåvning, men som av andra skäl är mindre lämpade för en viss utbildning och det yrke som utbildningen syftar till, kan stoppas innan de förspiller tid och resurser på en mindre lyckad karriär.
3. Enligt en nyligen gjord undersökning upplever många sökande det nuvarande systemet som opersonligt och byråkratiskt (Utbildningsdepartementet, 1995b). Genom att tillämpa det föreslagna systemet torde dessa negativa attityder elimineras.
4. Man erhåller en hel del information även om de sökande som ej antas. Denna borde kunna ligga till grund för någon form av studievägledning, där de icke antagna får råd och upplysningar om alternativa utbildningar.

Vilka konsekvenser skulle detta urvalsförfarande få för kvinnor och elever från lägre socialgrupper?

Här är jag inte helt säker - men resultaten kommer troligen att variera från en utbildning till en annan. Jag tror således inte, att det skulle missgynna dessa kategorier vid antagningen till en av de mest eftersökta utbildningarna, läkarutbildningen. Erfarenheter från en liknade antagningsprocedur vid Karolinska institutet i Stockholm visar nämligen att de grupper som hittills varit underrepresenterade snarast har gynnats av den nya typen av antagning (Åberg, 1995).

Vad gäller de högre tekniska utbildningarna är jag mer tveksam - vissa erfarenheter från den speciella provantagningen till arkitektutbildningen tyder på ökad snedrekrytering. Å andra sidan hörde jag nyligen rektorn för Chalmers tekniska högskola med kraft framhålla hur viktigt det är, att dagens civilingenjörer inte enbart intresserar sig för tekniska frågor utan även för samhällsvetenskap och humaniora - något som satt spår i utbildningen vid högskolan. (Sjöberg, 1996). Vidare påpekade han vikten av, att tekniker inte isolerar sig utan tar del av kultur- och samhällslivet, samt nödvändigheten av att de måste kunna samarbeta med människor med olika erfarenheter och från skilda miljöer. Möjligen kan denna programförklaring tyda på, att man vid de tekniska högskolorna nu blivit mer angelägna om att bredda sitt rekryteringsunderlag - något som den föreslagna urvalsmodellen möjliggör.

Avslutningsvis vill jag dock poängtera att den föreslagna urvalsmodellen, liksom urval till högskolan överhuvudtaget, måhända får en förhållandevis blygsam betydelse framöver - det stora problemet kan i stället bli att förmå tillräckligt många ungdomar att söka sig till högre utbildning.

Referenser:

- Härnqvist, K., Emanuelsson, I., Reuterberg, S-E. & Svensson, A. (1994). Dokumentation av projektet ”Utvärdering genom uppföljning”. *Rapporter från Institutionen för pedagogik, Göteborgs universitet*. 1994:03.
- Reuterberg, S-E., Westerlund, A. & Gustafsson, J-E. (1992). Beror gruppskillnaderna på tidigare selektion i utbildningssystemet? I *Betyg och högskoleprov för män och kvinnor*. UHÄ-rapport 1992:3.
- SCB (1995) Universitet och högskolor. Grundutbildning: Genomströmning. Resultat t.o.m. 1993/94. *Statistiska meddelanden* U20 SM 9502.
- Sjöberg, A. (1996). Tal vid Chalmers examensceremoni för civilingenjörer i Göteborgs Konserthus, 21 januari 1996.
- Stage, C. (1993). Gender differences on the SweSAT. A review of studies since 1975. *Educational Measurement 1993*:7. Department of Education. University of Umeå.
- Svensson, A. (1995a). Att välja eller välja bort naturvetenskap och teknik. Skolverket & VHS: *NOT-häfte nr 3* 1995.
- Svensson, A. (1995b). Kommentar till Ingemar Wedmans uppsats. I *RUT-93. Arbetsrapport nr 3*. Stockholm: Utbildningsdepartementet.
- Utbildningsdepartementet (1995a). *Gymnasieskolan - en bred grund för högskolan*. Ds 1995:56.
- Utbildningsdepartementet (1995b). Antagningen i praktiken. *RUT-93: Arbetsrapport nr 4*. Stockholm: Utbildningsdepartementet.
- Wedman, I. (1995). Studieframgång, prov och betyg. I *RUT-93. Arbetsrapport nr 3*. Stockholm: Utbildningsdepartementet.
- Åberg, H. (1995). Kommentar till Carl Holmbergs uppsats. I *RUT-93. Arbetsrapport nr 3*. Stockholm: Utbildningsdepartementet.

10 Intervju som urvalsinstrument

Christina Stage

Det är anmärkningsvärt att intervju som urvalsmetod har accepterats så lättvindigt och kritiklöst med tanke på alla diskussioner som har ägt rum om både betygen och högskoleprovet (och även andra prov) som urvalsinstrument. Det brukar framhållas som främsta fördel med intervjun att den kan fånga in motivation, ambition, intresse etc. Kan dessa svårsmätbara variabler verkligen mätas i en intervjusituation?

Den frihet som universitet och högskolor fick 1993 att själva bestämma urvalsmetod har lett till att intervjuer har ökat i användning (se t ex RUT-93). Detta gäller speciellt för de medicinska utbildningarna, där samtliga universitet antingen redan har startat eller planerar att starta med intervjuer som ett andra eller tredje steg i urvalsprocessen.

Nedan följer vissa överväganden som bör göras och frågor som bör besvaras innan man accepterar intervju som den bästa metoden i en urvalssituation:

Intervjun som urvalsinstrument innefattar:

- (Liksom alla urvalssituationer) att ett begränsat antal personer skall väljas från ett större antal kandidater till ett antal tillgängliga platser.
- (Liksom alla urvalssituationer) ett antagande att skillnader mellan individer är relaterade till deras möjlighet till framgång på de platser till vilka de söker.
- Ett samtal mellan en intervjuare och kandidaterna.
- På basis av detta samtal skall intervjuaren utvärdera varje kandidat och denna utvärdering skall leda till en prediktion (explicit eller implicit) beträffande den sannolika framgång denne kandidat kommer att ha på den sökta platsen i jämförelse med de övriga kandidaterna.

Vad krävs av en intervjuare?

Han/hon behöver ha grundlig kunskap om vad som krävs av de utvalda. Om inte en noggrann kriterieanalys redan har utförts måste han/hon göra en sådan.

- Han/hon måste veta vilka variabler som är relevanta: dvs vilka förmågor, karakteristika och egenskaper som är relaterade till en bra kriterieprestation, och således fungerar som prediktorvariabler.
- Han/hon måste veta hur han/hon i intervjun ska välja och få fram de beteenden som är valida indikatorer på de relevanta prediktorvariablerna.
- Han/hon måste veta hur han/hon ska utvärdera och väga varje respons (såväl verbala som andra) som uppvisats i intervjun för att erhålla den mest valida skattningen av de relevanta variablerna.
- En intervjuare bör mao klara av att genomföra: arbets/studieanalys, testkonstruktion, testadministration, resultatbedömning och statistisk analys.

Edwards, J. et al (1990) gjorde en genomgång av forskningsresultat angående intervju som urvalsmetod och konstaterade bland annat:

Bias can arise from a number of sources, including, but not limited to, rater tendencies, stereotyping, and interviewer background. The following findings are supported by research:

1. *Unfavorable information carries more weight than favorable information.*
2. *Rating errors can be corrected. Rating errors include the halo effect and distribution errors (leniency, severity, central tendency) ...*
3. *Racial bias against blacks has not been documented in research interviews. Lack of research evidence, however, does not necessarily mean that problems in accepting blacks and other minorities to medical school no longer exist, ...*
4. *Gender differences exist in interview ratings. Women are rated lower than men by both women and men interviewers.*
5. *Nonverbal as well as verbal interactions influence decisions.*
6. *Interviewers develop a stereotype of a good applicant and then try to match the candidates to the stereotype.*
7. *If several outstanding or marginal candidates have been interviewed in succession, an average candidate following these will appear better or worse in comparison. This is called the "contrast effect". (p 170)*

Enligt Trost (1992) bör för intervjun som för andra mätinstrument gälla att dess värde prövas mot dess:

- Objektivitet - vid utfrågning, registrering, utvärdering och tolkning
- Reliabilitet - stabilitet, inre konsistens
- Validitet - innehållsvaliditet, kriterievaliditet, begreppsvaliditet
- Ekonomi - intervjuer är dyrbara
- Fairness - behandlar/bedömer intervjuaren alla lika oberoende av kön, socialgrupp etnisk bakgrund osv.
- Etik - samma etiska krav måste ställas på en intervju, som på skriftliga test och frågeformulär avseende känsliga eller förtroliga frågor.
- Acceptans - upplever samtliga intervju-personer intervjun som en rimlig/rättvis urvalsmetod?

Förutom att det ovan beskrivna bör beaktas vid användning av intervju som urvalsinstrument, bör naturligtvis resultaten av ett intervjuförfarande utvärderas lika noggrant som andra urvalsmetoder.

Det som brukar framhållas som främsta fördel med intervju är möjligheten att få fram fakta som inte kan erhållas genom andra instrument. Som nämndes inledningsvis kan ett högt provresultat eller ett högt betyg beskrivas som nödvändigt, men ej tillräckligt för en bra studieprestation. Det som krävs därutöver är motivation, ambition, intresse osv. Om dessa svåråttbara variabler verkligen kan erhållas i en intervjusituation kan en intervju naturligtvis fungera som ett komplement till mer objektiva urvalsinstrument.

Referenser

Edwards, J., Johnson, E. & Molidor, J. (1990) The interview in the Admission Process. *Academic Medicin*, 65, p 167-177.

Hindbeck, H., Hagenfeldt, K., & Åberg, H. (1994) Lokal antagning till läkar-utbildning vid Karolinska institutet. Genomströmning och studieresultat för studerande antagna vårterminen 1992. *Rapport 59*. Karolinska, Institutionen för medicinsk informatik och pedagogik.

Kelly, E. L. (1953) An Evaluation of the Interview as a Selective Technique. In *Proceedings of the 1953 Invitational Conference on Testing Problems*. Princeton. N. J. Educational Testing Service. p 116-123.

Trost, G. (1992) Interview. In *Enzyklopädie der Psychologie. Grundlagen und Methoden der Differentiellen Psychologie*. Hogrefe. Verlag für Psychologie. Göttingen. p 463-505.

Weiss, M., Lotan, H. K. & Ben-Shakhar, G. (1988) Selecting candidates for medical school: an evaluation of a selection model based on cognitive and personality predictors. *Medical Education*, 22, p. 492-497.

II Högskoleprovet som prognosinstrument

Ewa Andersson & Simon Wolming

Högskoleprovet används för att göra urval bland sökande till högre studier. Det är därför naturligt att ställa frågan: Vilken förmåga har provet att förutsäga framgång i studierna? En aktuell studie visar att högskoleprovet och betyg har ungefär samma förmåga att förutsäga framgångsrika studier mätt med avlagda poäng. Meriter för arbetslivserfarenhet ger en sämre prognos.

Vad är då studieframgång? I dagsläget anses studieframgång vara avklarad examen på stipulerad utbildningstid. Men varken de faktorer som kan antas ha betydelse för studieframgång eller kriterier på studieframgång kan anses självklara. I en kommande doktorsavhandling baserad på bl.a. intervjuer med studenter och lärare problematiseras frågan om kriterier på studieframgång.

Innan vi går vidare med två artiklar, baserad på den forskning som utifrån olika perspektiv försöker besvara frågan om högskoleprovets prognosförmåga, görs en kort resumé av övrig prognosforskning vid Enheten för pedagogiska mätningar vid Umeå universitet. Avsnittet är således att betrakta som en introduktion till de båda artiklarna som baseras på data från det s.k. prognosprojektet. Mot bakgrund av att de båda artiklarna till viss del använder sig av gemensamma referenser anges dessa avslutningsvis i en gemensam referenslista.

Redan innan högskoleprovet började användas genomfördes en tämligen omfattande prognosstudie (Lexelius et al, 1979). Resultaten av den studien visade att betyg ger den bästa prognosen och att prov, jämförbara med högskoleprovet, är i det närmaste likvärdiga med betyg medan andra instrument genomgående har mycket måttliga samband med resultat i högre studier. Detta är även något som är i överensstämmelse med resultat från internationella studier.

Att genomföra en prognosstudie är dock behäftat med en rad problem, varav de främsta kanske är att definiera och mäta studieframgång (Henrysson 1984).

Hösten 1992 påbörjades en prognosstudie, i form av ett antal delstudier, som med skilda utgångspunkter och perspektiv tillsammans försöker besvara frågan om provets prognosförmåga.

Förutom de delstudier som presenteras i de båda artiklarna har en litteraturstudie, av den internationella prognosforskningen under perioden 1980-95, genomförts. Litteraturstudien fokuserade tre prediktorvariabler; betyg, urvalstest och antagningsintervjuer. Resultaten skiljer sig inte från tidigare studier, nämligen att betyg och urvalstest fungerade bäst som prediktorer av studieframgång och att antagningsintervjuer fungerade sämre, men det genomgående problemet att finna generella kriterier för studieframgång kvarstår (Wolming, 1995).

OLIKA URVALSGRUPPERS STUDIEPRESTATION

Simon Wolming

Det övergripande syftet med studien var att se hur studenter i olika urvalsgrupper hade presterat vid de aktuella utbildningarna samt om signifikanta skillnader mellan gruppernas prestationer kunde urskiljas. En annan aspekt som beaktades var gruppernas genomströmning, d v s i detta fall studieavbrott/uppehåll vid utbildningarna. Kunde några avvikelser mellan urvalsgrupperna uppvisas i detta avseende?

Studiens uppläggning

Studiens försökspersoner utgjordes av de 840 studenter som antogs hösten 1992 till ekonomlinjen, sociala linjen, läkarlinjen samt teknisk fysik vid universiteten i Lund och Umeå, samt till ekonomlinjen vid Högskolan i Karlstad.

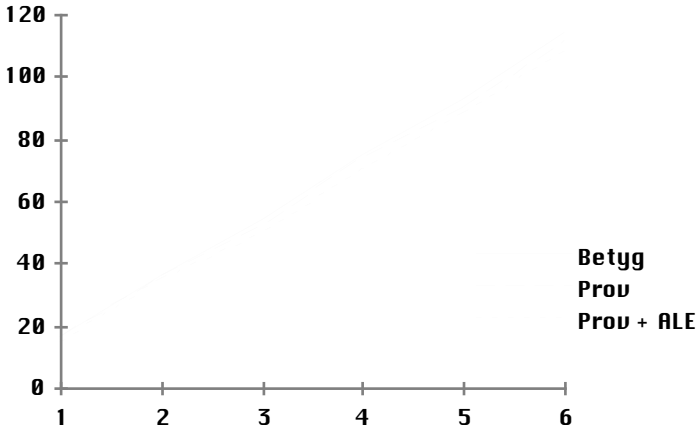
De utvalda studenternas resultat på högskoleprovet, betyg samt information om studenternas prestationer vid utbildningarna inhämtades från LADOK-registret vid respektive utbildningsort samt högskoleprovets eget register över provtagarna. LADOK-registret gav information om studieprestationer, betyg samt olika bakgrundsvariabler som kön och tidigare gymnasieutbildning och högskoleprovets register tillhandahöll information om studenternas resultat på högskoleprovet.

De antagna studenterna delades in i tre olika grupper, där indelningen berodde på hur de blivit antagna till utbildningarna. Alla som antagits via betyg från någon gymnasieutbildning eller motsvarande, sammanfördes till en betygsgrupp. Den andra gruppen utgjordes av de som antagits utifrån resultat på högskoleprovet och den tredje gruppen bestod av de studenter som tillgodoräknat sig 0,5 poäng för tidigare arbetslivserfarenhet (ALE) till sitt högskoleprovsresultat.

Studien följde dessa tre grupper under de sex första terminerna vid respektive utbildning. Efter varje termin samlades information in om studenternas prestationer i form av antal tagna högskolepoäng.

Hur klarar sig studenterna som antagits med högskoleprov jämfört med de som antagits med betyg?

Den centrala frågeställningen i studien var att se hur de studenter som antagits med högskoleprov klarade sig i jämförelse med de betygsantagna studenterna. I figur 1 visas de tre urvalsgruppernas prestationer i form av tagna högskolepoäng under de sex terminerna.



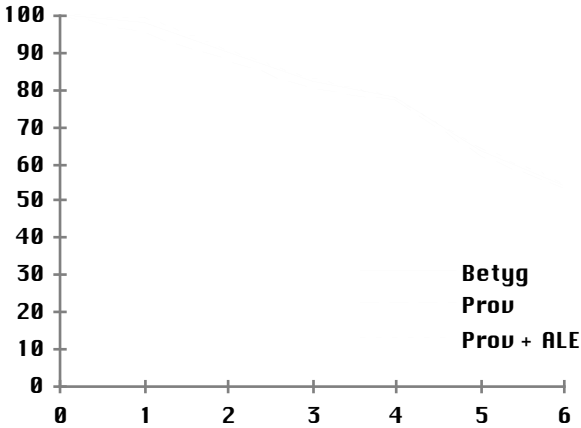
Figur 1. Genomsnittligt antal tagna högskolepoäng vid respektive termin för de tre urvalsgrupperna. Samtliga utbildningar.

Vid anblicken av figur 1 kan skillnaderna mellan urvalsgrupperna tyckas relativt små. De olika gruppernas prestationer avviker inte nämnvärt från varandra. Man ser dock att de betygsantagna studenterna under studietiden hade den bästa prestationen av de tre grupperna och att de ALE-antagna uppvisade en något sämre prestation än vad de övriga grupperna gjorde. Den fortsatta analysen visade att statistiskt signifikanta skillnader fanns mellan de betygsantagna studenterna och de studenter som antagits med ALE. Signifikanta skillnader uppvisades även vid vissa tillfällen mellan de två provgrupperna. De ALE-antagna studenterna uppvisade generellt en sämre prestation än vad de betygs- och högskoleprovsantagna studenterna gjorde. Inga signifikanta skillnader kunde urskiljas mellan de betygsantagna och de studenter som antagits med högskoleprovet utan ALE-tillägg.

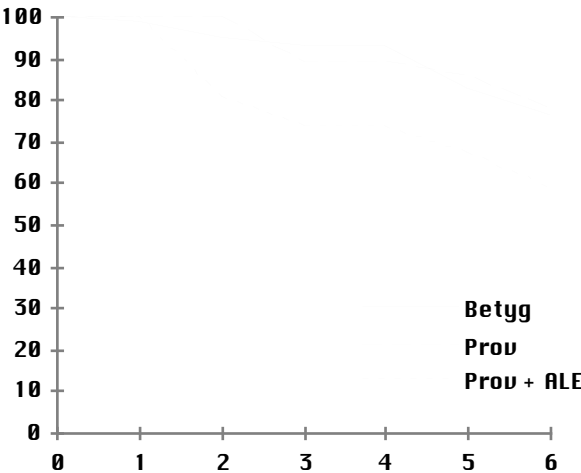
Studieavbrott och studieuppehåll för de tre urvalsgrupperna

Analysen av bortfallet från utbildningarna visade inte på några signifikanta skillnader mellan urvalsgrupperna. Undantaget var läkarlinjen där studieavbrott/uppehåll i högre grad uppvisades av de ALE-antagna än de övriga

urvalsgrupperna. Inga skillnader kunde identifieras mellan de betygsantagna och de som antagits utifrån det rena provresultatet. Intressant att notera är att de ALE-antagna vid läkarlinjen inte presterade sämre än vad de övriga urvalsgrupperna gjorde, utan avvek bara när det gällde bortfallet. I figur 2 och 3 visas bortfallet vid samtliga utbildningar och för läkarlinjen.



Figur 2. De tre urvalsgruppernas procentuella bortfall per termin. Samtliga utbildningar.



Figur 3. De tre urvalsgruppernas procentuella bortfall per termin. Läkarlinjen.

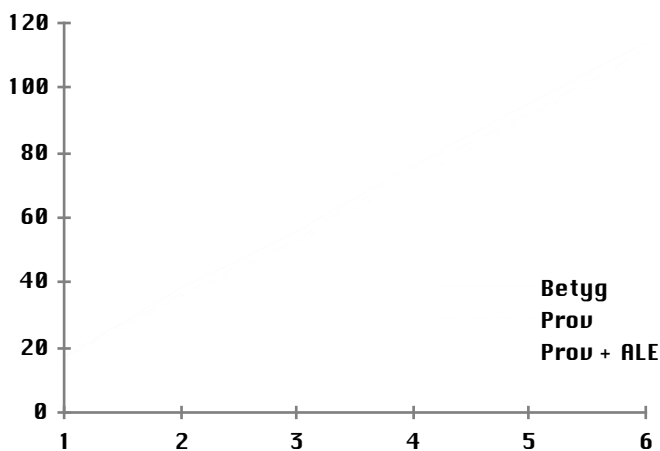
Utifrån det redovisade materialet är det omöjligt att identifiera vad orsakerna till studenternas bortfall egentligen är. Bortfallet behöver inte betyda att studenterna har misslyckats med sina studier utan det kan ha flera möjliga

orsaker. Hindbeck, Hagenfeldt & Åberg (1994), visade i en studie att de som antagits med ALE var äldre, bildat familj etc innan studierna påbörjats och att detta möjligtvis påverkade deras möjligheter till goda studieresultat på den krävande läkarutbildningen. Det är möjligt att dessa förhållanden kan vara en av förklaringarna till de ovan redovisade ALE-antagna studenternas bortfall.

Vad som däremot är av intresse är att bortfallet av studenter påverkar utbildningarnas ekonomiska tilldelning. Ju fler studenter som utexamineras, ju mer resurser erhåller institutionerna. Sett ur detta perspektiv är det av vikt för utbildningarna att rekrytera studenter som fullföljer den valda utbildningen med normal studietakt och med goda studieresultat.

Studieprestationen för kvarvarande studenterna

För att ge en "sann" bild av urvalsgruppernas prestationer gjordes en analys baserad endast på resultaten för de studenter som fanns kvar efter de sex terminerna. På detta vis har hänsyn tagits till det terminsvisa bortfall som uppvisats för urvalsgrupperna. I figur 4 visas de kvarvarande studenternas prestationer.



Figur 4. Studieprestation för kvarvarande studenter efter sex terminer. Samtliga utbildningar.

Som synes skiljer sig inte urvalsgruppernas prestationer nämnvärt åt när bortfallet tagits bort. Den statistiska analysen uppvisade i stort sett ett liknande resultat som när samtliga studenters resultat fokuserades. De

signifikanta skillnaderna återfanns mellan betygs- och ALE gruppens prestationer.

Sammanfattning

Utifrån de resultat som framkommit i studien kan man dra följande generella slutsatser. Inga direkta skillnader i studieprestation kan identifieras mellan de som antagits på betyg och resultatet på högskoleprovet. Dessa två urvalsmetoder väljer ut studenter som i de följande studierna presterar på ett ungefär likvärdigt sätt. När extrapoäng för ALE läggs till provresultatet avviker däremot resultatet för de ALE-antagna studenterna gentemot övriga grupper. Studenterna som antagits med ALE, uppvisade en genomgående sämre studieprestation än övriga grupper. Resultaten bekräftar de tidigare erfarenheterna av ALE i samband med urval, se t ex Henrysson, Kriström & Lexelius (1985). ALE sänker prognosvärdet för urvalsinstrumentet, beroende på att de som antas med ALE ofta har ett lägre medelbetyg och provresultat än vad övriga antagna har. Tillägget som ALE utgör till urvalsinstrumentet, i det här fallet till provresultatet, ger i samband med urvalet en skev bild av studentens förmåga. De studenter som tillgodoräknar sig ALE överskattas vid urvalet i jämförelse med övriga studenter som inte får något tillägg.

KRITERIER FÖR STUDIEFRAMGÅNG

Ewa Andersson

Traditionella prognosstudier är ofta inriktade mot att studera sambandet mellan s.k. prediktorer som betyg och/eller provresultat och s.k. kriterier för studieframgång som ofta likställs med avklarad examen, alternativt godkända kurser inom ramen för beräknad studietid. Prognosstudier förutsätter således en variation i såväl prediktorerna (betyg eller provresultat) som i kriterierna (avlagda examina eller poäng). För att ett urvalsinstrument ska fungera bra innebär det mer konkret att enskilda personer eller grupper av personer som har ett högt medelbetyg och/eller ett högt högskoleprovresultat även förväntas lyckas väl med sina universitetsstudier i den bemärkelsen att de i högre grad än andra studerande kommer att avlägga examen. Men studenterna har inte alltid för avsikt att avlägga examen. Detta kan vara en tänkbar förklaring till att genomströmningen inte stämmer överens med den tänkta. Det finns också exempel på utbildningar där så gott som samtliga studenter avlägger examen oberoende av prediktorvärde, vilket försvårar prognosstudier.

Denna studie tar därför sin utgångspunkt i att studieframgång inte enbart kan begränsas till avlagd examen inom de tidsramar som en utbildning omfattar. Studiens syfte är således att problematisera begreppet studieframgång. Inledningsvis redovisas huvudresultaten från en pilotstudie som genomfördes med i huvudsak studerande från linjen för personal- och arbetslivsfrågor (P-linjen) i ett förberedande syfte.

Studenterna om förutsättningar för framgång

Pilotstudien utgörs av tre delstudier baserad på intervju- och enkätdata. I den första studien som omfattar 30 studenter ombads de studerande i enkätform att ange de faktorer som de anser vara av betydelse för studieframgång respektive avsaknad av studieframgång. Sammanlagt erhöles 283 faktorer, varav 103 faktorer kan beskrivas i termer av *personliga egenskaper* medan 87 faktorer närmast kan beskrivas som *studietekniska färdigheter*, 51 faktorer kan beskrivas som *prestationsinriktade* egenskaper samt 20 egenskaper som kan sammanfattas i termer av att vara *socialt inriktad*.

De personliga egenskaperna dominerar även när 28 studerande i enkätform, (delstudie II) bland ett antal i förväg listade egenskaper, ombads att rangordna de fem viktigaste egenskaperna för studieframgång. *Motivation* nämns av samtliga studenter, därefter kommer *självdisciplin* som 23 av de 28 studenterna anser vara mest betydelsefullt för studieframgång. Även studieteknik har nämnts av ett flertal studenter. Intressant att notera är dock att endast 4 av de 28 studenterna har nämnt *skriftlig förmåga* respektive *tidigare utbildning* (förkunskaper) som betydelsefullt i sammanhanget. På motsvarande sätt anser endast 2 studenter att *samarbetsförmåga* är betydelsefullt för studieframgång.

Med hjälp av intervjuer gavs 20 av de studerande (delstudie III) tillfälle att utveckla sin syn på förutsättningen för studieframgång. Förutom motivation och självdisciplin, bedöms *att trivas med sig själv*, *självförtroende* och *någon form av intelligens* vara viktiga. Den framtida yrkesrollen bedöms dock kräva andra förutsättningar. Vilka dessa är varierar mellan studenterna. Den egenskap som främst nämns är *god samarbetsförmåga*, men även *lyhördhet* samt en *helhetssyn*. Andra egenskaper som anses betydelsefulla för den framtida yrkesfunktionen är t.ex. *självförtroende*, *självkänedom*, *analytisk förmåga* och *att kunna uttrycka sig i tal och skrift*. En egenskap som är gemensam med de faktorer som nämns för studietiden är *självdisciplin*.

Trots att de studerande huvudsakligen anger skilda egenskaper som förutsättning för framgång i studier respektive yrkesliv och att den verkliga utbildningen sker först i och med inträdet i yrkeslivet, ansåg samtliga (med undantag för en person) att de kommer att ha stor nytta av de kunskaper och färdigheter som utbildningen ger. Flertalet ansåg också att utbildningen utgör en bra grund för vidare utveckling och ett intressant arbete i framtiden.

Framgång i studier - inte bara att klara tentamina

I en fördjupad studie över kriterierna på studieframgång har intervju- och enkätdata inhämtats från sammanlagt 430 studenter antagna höstterminen 1992 respektive 1993 vid linjerna teknisk fysik, ekonomlinjen, läkarlinjen, sociala linjen och ämneslärarlinjen. Vidare har ett antal lärare vid de fem utbildningarna intervjuats om sin syn på kriterier för studieframgång.

En övergripande beskrivning av gruppens bakgrund och de studerandes syn på studierelaterade faktorer har erhållits med hjälp av ett frågeformulär, till

exempel de studerandes förberedelser inför högskolestudierna, deras syn på nuvarande utbildning, studiemiljö, examination och betygssystem samt framtidsplaner.

Enkätdata har legat till grund för planering och utformning av de intervjuer som sedan har gjorts med ett slumpmässigt urval studerande, antagna höstterminen 1992, vid de fem utbildningarna. De studerande har intervjuats vid två tillfällen varav det första intervjutillfället ägde rum efter ungefär ett år på utbildningen. Den andra intervjun genomfördes efter ungefär ytterligare ett år. De studerande har bl. a. intervjuats om motiv för val av utbildning, undervisnings- och examinationsformer, studievanor och studieprestation. De har också getts i uppgift att beskriva en framgångsrik student. Utan att föregripa den framtida resultatredovisningen kan redan nu sägas att de studerandes syn på en framgångsrik student inte enbart kan beskrivas som en student som klarar tentorna. Preliminära analyser av enkät- och intervjudata indikerar att flertalet studerande har en uppfattning av vad som krävs i förhållande till utbildningarnas explicita kriterier och att detta till viss del får konsekvenser för deras studier, t.ex. för omfattningen av den tid som ägnas åt självstudier.

En s.k. skattningsstudie har även genomförts bland de studerande antagna höstterminen 1993 på ekonomlinjen och teknisk fysiklinjen. De studerande har i denna studie ombetts att skatta den egna studieprestationen ett år senare.

Figuren nedan visar fördjupningsstudiens fokus, vilket schematiskt kan beskrivas med hjälp av en fyrfältstabell i termer av explicita och implicita kriterier för studieframgång. Med explicita kriterier avses således de kriterier som är att betrakta som traditionella i den meningen att de hänför sig till ett uttalat formellt betygssystem. Med implicita kriterier avses sådana kriterier som inte är direkt uttalade men som existerar likaväl.

	Traditionella kriterier	
	Ja	Nej
Explicita kriterier		
Implicita kriterier		

Figur 2. Schematisk modell av fördjupningsstudiens syfte.

De explicita kriterierna i de fem undersökta utbildningarna har beskrivits i en särskild rapport (Andersson & Grysell 1995). Utbildningarna skiljer sig från varandra i ett flertal avseenden. Vissa utbildningar har t. ex. en mer fmggraderad betygsskala än andra och utbildningarna har olika undervisnings- och examinationsformer. Beträffande examinationsformer kan exempelvis sägas att dessa kan variera från obligatorisk närvaro till individuell skriftlig tentamen. Andra exempel på examinationsformer är seminariedeltagande, skriftliga och/eller muntliga gruppredovisningar. Denna typ av skillnader mellan utbildningar kommer dock sällan till uttryck vid traditionella prognosstudier utan utbildningarna eller snarare de studerandes prestationer betraktas som jämförbara.

För närvarande pågår arbetet med att analysera de data som närmast kan beskrivas som den implicita delen av studieframgångsbegreppet och som baseras på de enkät- och intervjustudier som har genomförts tillsammans med Tomas Grysell under 1992-95. Resultaten kommer att presenteras i form av en doktorsavhandling under 1996/97.

Referenser

Andersson, E., Grysell, T. (1992) Studenters syn på studieframgång. En pilotstudie inför studier av studieframgångskriteriet inom ramen för prognosprojektet vid Högskoleprovet. *Pm nr 67*. Avdelningen för pedagogiska mätningar, Pedagogiska institutionen, Umeå universitet.

Andersson, E., Grysell, T. (1995) En beskrivning av fem utbildningar vid Umeå universitet. En genomgång av ekonomlinjen, sociala linjen, teknisk fysik, läkarlinjen samt ämneslärarlinjen med avseende på organisation och utformning. *Pm nr 104*. Avdelningen för pedagogiska mätningar, Pedagogiska institutionen, Umeå universitet.

Henrysson, S. (1984) Prognos av studieframgång i högre studier, metoder och problem. *Arbetsrapport nr 14*. Pedagogiska institutionen, Umeå universitet.

Henrysson, S., Kriström, M., & Lexelius, A. (1985) Meritvärdering och studie-prognos. Några undersökningar av antagningssystemets effekter. *Arbetsrapporter från pedagogiska institutionen, Umeå universitet nr 21*, Pedagogiska institutionen, Umeå universitet.

Lexelius, A., Mattsson, H., Nordlund, G. & Wedman, I. (1979) Prognosförmåga hos ett provbatteri för urval till högskoleutbildning. Uppföljningsförsök med en förlaga till Högskoleprovet. *Spånor från Spint, nr 14*, Pedagogiska institutionen, Umeå universitet.

RUT-93 (1995) Urvalsmetoder. *Arbetsrapport Nr 3*. Utbildningsdepartementet.

Wolming, S. (1995) Prediktion av studieramgång. Reflektioner utifrån en litteraturgranskning. *Pm nr 101*. Enheten för pedagogiska mätningar, Umeå universitet.

Wolming, S. (1996) Hur klarar sig studenter som antagits via högskoleprov? *Pm nr 111*. Enheten för pedagogiska mätningar, Umeå universitet.

12 Vad mäter betyg och högskoleprov

Jan-Eric Gustafsson

Den psykometriska forskningen har kunnat visa att högskoleprovet mäter olika förmågor: en generell analytisk förmåga, verbala kunskaper och färdigheter och delprovsspecifika förmågor. Den generella analytiska förmågan mäts framför allt i högskoleprovets delprov NOG och DTK. Den breda verbala kunskaps- och färdighetsförmågan påverkar framför allt delproven ORD, LÄS och ELF i högskoleprovet.

Introduktion

Forskning kring mätning av mänskliga förmågor och funktioner brukar benämnas "psykometrisk" forskning, och under denna rubrik faller studier av urvalsinstrument. Syftet med denna uppsats är att med utgångspunkt i aktuell psykometrisk forskning diskutera vad betyg och högskoleprovet mäter när de används för urval till högre utbildning.

En av de frågor som rönt störst intresse i den mer teoretiskt inriktade psykometriska forskningen gäller de kognitiva förmågornas struktur. Flera studier (Carroll, 1993; Gustafsson, 1988) antyder dock att föremågsdimensioner på tre nivåer av generalitet kan separeras. Den högsta nivån omfattar en generell intellektuell förmåga, och en hel del forskning visar att denna förmåga har sin tyngdpunkt i en analytisk, icke-verbal, resonemangs-förmåga, eller en förmåga att uppfatta och hantera komplexa relationer (Gustafsson & Undheim, 1996). Den näst högsta nivån omfattar ett drygt halvdussin dimensioner, där de följande utgör de viktigaste:

- *Verbal Förmåga och Kunskap*, som har sin tyngdpunkt på allmänbildning och skolprestationer, där läsförmåga också är central.
- *Visuell Förmåga*, som är en förmåga att uppfatta och hantera rumslig och bildlig information.
- *Minnesförmåga*, som avser tillägnande och reproduktion av material under kortare tidsperioder.
- *Auditiv Förmåga*, som avser förmåga att uppfatta och differentiera ljudbilder, inklusive musik.
- *Snabbhet*, som i första hand avser utförande av enkla uppgifter.

På den tredje och lägsta nivån återfinns en stor mängd specifika förmågor, vilka motsvarar de så kallade "primära mentala förmågor" vilka först identifierades av Thurstone (1938).

Förmågor som mäts av högskoleprovet

Mot bakgrund av den forskning som beskrivits ovan är det naturligt att ställa frågan vilken eller vilka förmågor Högskoleprovet mäter. Högskoleprovet består från och med våren 1996 av fem delprov:

- *ORD*: (30 uppgifter). Detta prov avser att mäta förståelse av olika ord och begrepp. Till ett givet ord finns fem svarsalternativ varav ett skall markeras som synonym till det givna.
- *NOG*: (20 uppgifter). Provet avser att mäta numerisk-logisk förmåga. Provtagaren skall bedöma om tillräcklig information föreligger för att ett givet problem skall kunna lösas. Svarsalternativen är fasta.
- *LÄS*: (24 uppgifter). Läsförståelseprovet avser att mäta läsförståelse i vid mening. Provet innehåller 6 texter med vardera 4 frågor.
- *DTK*: (20 uppgifter). Provet avser att mäta förmågan att tolka information i diagram, tabeller, kartor och andra grafiska framställningar. Till 10 uppsättningar information skall 20 frågor besvaras.
- *ELF*: (24 uppgifter). Provet mäter engelsk läsförståelse med texter av varierande längd.

Till och med 1995 ingick ytterligare ett delprov, *AO*, med 30 uppgifter. Det allmänorienterade provet avsåg att återspegla kunskap från många olika områden. Provet var bredare till sitt innehåll än konventionella skolkunskaper. Kunskaperna förväntades förvärvas under en längre tidsperiod.

Åtminstone 4 av delproven har en naturlig hemvist i den breda prestationsdimension som ovan kallats Verbal Förmåga och Kunskap (*ORD*, *LÄS*, *AO*, och *ELF*) och delvis gäller detta även de två återstående delproven (*NOG* och *DTK*). De två senare delproven ställer emellertid också mer direkta krav på problemlösning och hantering av komplex information, vilket antyder att de i större utsträckning återspeglar den generella analytiska förmågan.

En analys av 10 tidiga omgångar Högskoleprov med multivariata statistiska metoder (Gustafsson, Wedman & Westerlund, 1992) gav stöd för en uppdelning av delproven i två kategorier: *NOG* och *DTK* föll ut i en dimension som kan benämnas *A* (för "analytisk" förmåga), och *ORD*, *LÄS*,

AO och STUF (Studietekniska färdigheter; ELF-provets föregångare), föll ut i en annan dimension som kan benämnas *K* (för "verbala Kunskaper och färdigheter"). Det förelåg emellertid också ett betydande samband mellan de två dimensionerna vilket stödjer tanken att det också finns en generell förmåga bakom A och K. I anslutning till mycket annan forskning kan det antas att A-dimensionen i stort sammanfaller med den generella förmågan.

Tabell 1 visar andelen variation i de 6 delproven i Högskoleprovet våren 92 som bestäms av tre kategorier av förmågor: den generella *Analytiska förmågan*; *verbala Kunskaper och färdigheter*, och *delprovsspecifika förmågor* (de senare dock sammanblandade med rena mätfel).

Tabell 1. Procent variation i delproven som förklaras av olika dimensioner

Prov	Analytisk förmåga	Verbala kunskaper och färdigheter	Specifik förmåga
ORD	21	49	30
NOG	64		36
LÄS	25	29	46
DTK	58		42
AO	18	27	55
ELF	27	35	38

Som framgår av Tabell 1 har den generella *Analytiska förmågan* betydelse för prestationer på samtliga delprov men framförallt på NOG och DTK. *Verbala Kunskaper och färdigheter* har framförallt inflytande över ORD och ELF-proven. Det bör också understrykas att en betydande mängd varians är oförklarad av A och K. Den återstående variationen kan delas upp i en del som har med det specifika delprovet att göra, och en del som är att betrakta som slumpmässiga mätfel. De data vi arbetar med här tillåter inte uppdelning i dessa två komponenter, men för åtminstone några delprov är den delprovsspecifika förmågan av stor betydelse. För LÄS och AO utgör denna variationskälla sålunda ungefär hälften av den observerade variationen.

Det måste betonas att en analys av delprovets mättegenskaper inte är detsamma som en analys av Högskoleprovets mättegenskaper. Anledningen till detta är att summering av delprovresultaten till ett totalresultat tenderar att ge dominans åt de komponenter som återfinns i flera delprov, och reducera betydelsen av komponenter som är representerade i ett enda eller

få delprov. I det summerade Högskoleprovresultat svarar därför den generella A-dimensionen för ett bidrag som är ungefär dubbelt så stort som bidraget från K. Inflytandet från de delprovspecifika komponenterna och de rena mätfelen är också mindre än i något enskilt delprov.

Förmågor som mäts av betyg

En motsvarande analys skulle kunna göras av gymnasiebetygen, vilket dock ännu inte skett. I väntan på empiriska resultat kan det dock vara intressant att spekulera i vilka resultat som kan förväntas. En rimlig hypotes är att variationen i gymnasiebetyg också går att dela upp i A och K, men att den relativa betydelsen av K är större, liksom att det föreligger systematiska skillnader mellan olika gymnasielinjer. Det är också rimligt att förvänta sig ett inflytande från motivationella faktorer, liksom från intressfaktorer.

Westerlund (1990) studerade medelvärdena på A och K för provtagare med bakgrund i olika gymnasieutbildningar, och fann en variation i resultatprofilerna som var mycket större än variationen i totalresultat. Provtagare som genomgått humanistisk linje hade sålunda ett lågt medelvärde på A och ett högt medelvärde på K, medan provtagare från teknisk linje uppvisade motsatt mönster. Det måste observeras att dessa resultat inte kan generaliseras att gälla alla studerande på gymnasielinjerna eftersom gruppen av provtagare utgör ett självselekerat urval. Icke desto mindre demonstrerar dessa resultat att de två dimensionerna i Högskoleprovet förmår att differentiera mellan studerandegrupper med olika inriktning.

Sambanden mellan Högskoleprovets dimensioner och studieresultat kan också tänkas variera över utbildningar. I avsaknad av studieresultat från högre utbildning kan vi återigen utnyttja resultat från gymnasiet. I Tabell 2 presenteras sambanden mellan A och K (bestämde från Högskoleprovet våren 92), och betygsmedelvärde inom fem gymnasielinjer för elever födda 1973.

Tabell 2. Samband mellan A och K förmågorna mätt med Högskoleprovet och betygsmedelvärde

Prov	Analytisk förmåga	Verbala kunskaper och färdigheter	Anatal
Ekonomisk	0,47	0,21	3928
Humanistisk	0,44	0,38	970
Naturvetenskaplig	0,51	0,31	3993
Samhällsvetenskaplig	0,51	0,31	4106
Teknisk (åk 3)	0,39	0,34	3619

Det är tydligt att sambanden varierar över linjerna, och i synnerhet tycks K ha högre samband med studieresultat på humanistisk linje än på ekonomisk linje. Även för A finns skillnader, med lägst samband inom teknisk linje. Resultaten är inte helt lättolkade och kräver också en analys av det sätt på vilket olika ämnen influerar betygsmedelvärdet inom de olika gymnasielinjerna.

Det är också rimligt att anta att många utbildningar ställer krav på förmågor som inte mäts av Högskoleprovet, vilket kan vara en förklaring till det låga sambandet inom teknisk linje. Mer uppenbara exempel är här de konstnärliga utbildningarna, som också sedan länge anordnat speciella urvalsprov. Inom andra utbildningsområden pågår också försök med alternativa urvalsinstrument (t ex vid flera läkarutbildningar, vid arkitektutbildningarna, och vid journalistutbildningarna). De här nämnda utbildningarna antyder några förmågeområden som Högskoleprovet inte representerar: spatial förmåga, skriftlig uttrycksförmåga, och förmåga att relatera till andra människor. Det är naturligtvis lätt att finna exempel på andra utbildningar där dessa förmågor skulle vara betydelsefulla.

Könsskillnaderna i betyg är till kvinnornas fördel. Även här finns forskningsresultat som ger åtminstone vissa uppslag till förklaringar av orsakerna till dessa skillnader. Vad som framstår som den kanske viktigaste förklaringsgrunden är det faktum att kvinnor har en betydligt bättre förmåga att uttrycka sig i skriftlig form än män (se t ex Breland et al., 1994). Enkelt uttryckt kan man säga att kvinnor i betydligt större utsträckning än män skriver läsligt och begripligt, och eftersom resultat på skriftliga prov utgör den kanske viktigaste grunden för betygssättning i många ämnen har detta ett kraftigt genomslag på betygsmedelvärdet. Förmåga att uttrycka sig i skriftlig form är förvisso en mycket viktig förmåga i högskolestudier, och

detta är inte heller något som låter sig mätas med instrument av Högskoleprovstyp. Det kan sålunda hävdas att betygen i detta avseende mäter en relevant förmåga, och att det bland annat är detta som gör att betyg predicerar studieframgång något bättre än Högskoleprov. Det kan dock också hävdas att genomslaget av skrivförmåga är för stort i betygen, genom att det influerar de flesta ämnen.

Skillnader mellan betyg och högskoleprov

I kort sammanfattning kan följande hypoteser framföras om skillnader mellan vad som mäts av betyg och högskoleprov:

- Högskoleprovet mäter i större utsträckning den generella analytiska förmågan, medan betygen i större utsträckning mäter verbala kunskaper och färdigheter.
- Betygen påverkas i stor utsträckning av produktiva färdigheter (skrivning, muntlig produktion), vilket inte gäller för högskoleprovet.
- Betygen återspeglar resultat under en treårsperiod, medan resultat på högskoleprovet bestäms av prestationer under några timmar. Det är sålunda troligt att betygen i större utsträckning än högskoleprovet återspeglar skolmotivation och förmåga att arbeta långsiktigt. För goda resultat på högskoleprovet krävs i stället en förmåga till kraftsamling under kortare tid, liksom sannolikt även en förmåga att arbeta snabbt.

Eftersom olika utbildningar i olika hög grad tar i anspråk de olika förmågorna och färdigheterna är det rimligt att förvänta sig att betyg och högskoleprov i varierande grad förmår predicera studieresultat i olika utbildningar. Det framstår som angeläget att dessa sambandsmönster utsätts för systematiska studier nu när successivt allt större grupper som antagits på högskoleprovsresultat examineras från högskolan.

Referenser

Breland, H. M., Danos, D. O., Kahn, H. D., Kubota, M. Y., & Bonner, M. W. (1994). Performance versus objective testing and gender: An exploratory study of an Advanced Placement history examination. *Journal of Educational Measurement*, 31, 275-294.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Gustafsson, J.-E. (1988). *Hierarchical models of individual differences in cognitive abilities*. In R. J. Sternberg, *Advances in the psychology of human intelligence*, Vol. 4, (pp. 35-71). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Gustafsson, J.-E. & Undheim, J. O. (1996). *Individual differences in cognitive functions*. In D. C. Berliner & R. C. Calfee (Eds.) *The handbook of educational psychology*. New York: Macmillan.

Gustafsson, J.-E., Wedman, I., & Westerlund, A. (1992). The dimensionality of the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 36, 21-39.

Westerlund, A. (1990). *Högskoleprovet och utbildningsbakgrund*. 80-poängs-uppsats i pedagogik.

13 Testträning

Gunilla Ögren

Genom utbildning kan man påverka sitt resultat på högskoleprovet. Den som skall genomgå högskoleprovet bör också förbereda sig genom att träna på gamla prov. Därigenom blir man förtrogen med provets innehåll och uppläggning och upptäcker sina svagheter och kan förbättra sina kunskaper just på dessa områden. Man kan också göra provet mer än en gång för att få erfarenhet av testsituationen.

Föreliggande presentation är baserad på den forskning som gjorts av docent Widar Henriksson, Enheten för pedagogiska mätningar, Umeå universitet. (Henriksson var vid tiden för konferensen utomlands.) Han började redan under 70-talet att intressera sig för testträning vilket resulterade i en avhandling 1981 med titeln *Effekter av övning och instruktion på testprestation. Några empiriska studier och analyser avseende övningens och instruktionens betydelse för testprestation*. De empiriska studierna i avhandlingen utgörs av ett antal undersökningar om effekter av övning och instruktioner på kvantitativa numeriska prov.

Därefter har han intresserat sig för uppföljningar och studier av upprepat provdeltagande. Resultat från två senare studier finns redovisade i *Effekter av upprepat provtagande. En studie av poängförändring från första till andra provgenomförandet* (Henriksson & Bränberg, 1992) och *Repeated Test Taking and the Swe SAT* (Henriksson, 1995).

Under 80-talet och början av 90-talet hade ungefär 1/3 av provdeltagarna tagit provet vid tidigare provtillfälle och andelen var större vid höstprovet än vid vårprovet. Vid de senaste årens provtillfällen hade 40 till 45 procent tagit provet tidigare, även då var andelen större på hösten än på våren. Man kan också parentetiskt nämna att fram till 1990 hade en person upprepat sitt provdeltagande 18 gånger, tre personer hade genomfört provet 17 gånger och fem personer hade genomfört provet 16 gånger.

Henrikssons studier om upprepat provdeltagande har visat att den största förändringen i provpoäng sker mellan det första och andra provtillfället medan förändringen, jämfört med bästa tidigare provresultat, i princip var

marginell vid det tredje eller fjärde provtillfället. Studier visar att ungefär 60 procent av provdeltagarna förbättrar sitt resultat mellan första och andra provtillfället medan för resterande 40 procent resultatet är oförändrat eller försämrat. (Henriksson et.al.,1985; Henriksson, 1995).

Resultat från studier gjorda under 80- och början av 90-talet visar att det på gruppnivå sker en poängförändring på ungefär 2.5 råpoäng mellan första och andra provtillfället, vilket motsvarar 0.1 normerad poäng (Henriksson, 1990). De senaste studierna visar dock en poängförändring på drygt 3 råpoäng (Henriksson,1995).

Det finns två troliga förklaringar till denna förhöjda ökning av råpoängen. Den ena är att den utökning av antalet provdeltagare, som skedde i början på 90-talet, i och med de ändrade antagningsreglerna, huvudsakligen utgjordes av gymnasister. Gymnasieelever genomför ofta högskoleprovet flera gånger under sin gymnasietid. Eftersom poäng på högskoleprovet påverkas av ökad utbildning så bidrar också den utbildning som genomgått mellan två provtillfällen till att provpoängen ökar. Det bör uppmärksammas att denna poängökning inte är oegentlig eftersom den återspeglar en ökning av bakomliggande kunskaper och färdigheter. Den andra förklaringen är att ELF-provet infördes 1992. Detta delprov avser att mäta kunskaper och färdigheter i engelska och eftersom engelska är ett skolämne i gymnasieskolan påverkas poängen på detta delprov (och därmed också totalpoängen) förhållandevis direkt av specifik undervisning i ämnet. De provdeltagare som t.ex. genomför högskoleprovet på hösten i årskurs 3, och upprepar sitt provtagande på våren (ett mycket vanligt beteende), har således erhållit ett halvt års undervisning i engelska mellan dessa provtillfällen. Det är en rimlig slutsats att denna skolning medfört en kunskaps- och färdighetsökning som sedan återspeglar sig i en högre poäng på ELF-provet (Henriksson, 1995). En ökning av provpoängen kan således vara en effekt av att den faktiska förmågan förändras från ett provtillfälle till ett annat d.v.s. de kunskaper och färdigheter som provet mäter har förbättrats t.ex. genom utbildning. Det finns ett starkt samband mellan utbildningsnivå och provpoäng, ju högre utbildning ju högre provpoäng. Studierna visar också att när tiden mellan de två provtillfallen ökar, ökar också provpoängen förutsatt att den mellanliggande perioden har använts till att åstadkomma en höjning av utbildningsnivån (Henriksson, 1990).

På individnivå föreligger stora variationer avseende poängförändring mellan det första och det andra provtillfället. I de flesta fall utgörs de av förändringar

från en mycket låg nivå på skalan till en genomsnittlig nivå och i många fall är det provdeltagare med utländsk utbildning som, via sina studier i svenska, erhållit en högre poäng (Henriksson & Bränberg, 1994).

En annan anledning till poängförändring kan vara att det resultat som en provdeltagare erhöll vid det första provtillfället är en underskattning av provdeltagarens faktiska förmåga. Detta kan exempelvis bero på att personen saknar s.k. testförtrogenhet. Med testförtrogenhet avses en persons samlade erfarenhet av prov och provsituationer, vilken är av stor betydelse för att den erhållna provpoängen skall överensstämma med faktisk förmåga.

Ännu en förklaring till att det kan ske en poängökning kan vara att provdeltagaren vid genomgång och analys av sina erfarenheter från det första provtillfället inser sina bristande kunskaper och försöker att åtgärda dessa eller finner onödiga slarvfel som kan undvikas vid nästa provtillfälle. Erfarenheter från det första provtillfället kan också medvetandegöra att den tid som fanns till förfogande inte disponerades på ett rationellt sätt och denna erfarenhet kan utnyttjas vid det andra provtillfället.

Ytterligare en förklaring till att det uppkommer en poängökning mellan första och andra provtillfället är att det faktiskt finns provdeltagare som ser det första provtillfället som ett tillfälle ”att se vad det går ut på”. Med denna kunskap som bas gör de sedan sitt bästa för att erhålla en bra poäng vid det andra tillfället.

Den största poängförändringen på delprovsnivå sker för NOG- och DTK-proven medan poängförändringen för ORD-provet är liten. Orsaken till denna skillnad mellan delproven skulle kunna vara att provdeltagaren inte var förtrogen med själva testet och uppgiftsformatet (t.ex. NOG-provet) och därmed inte heller förtrogen med innebörden i ett rationellt och fungerande testbeteende. Som tidigare nämnts krävs en viss testförtrogenhet för att den erhållna testpoängen skall överensstämma med individens faktiska förmåga. Förklaringen till att man förbättrar sig på DTK kan vara att man går igenom provet efter det att man fått facit och därigenom finner egna strategier för lösande av uppgifterna. Att poängförändringen är minst på ORD kan förklaras av att det är ett renodlat kunskapsprov och påverkas inte av ökad testförtrogenhet.

Sammanfattningsvis kan sägas att det bästa sättet att höja sin provpoäng är att genom utbildning öka sina kunskaper och färdigheter och att i förväg

bekanta sig med provet och provsituationen för att uppnå den tidigare nämnda testförtrogenheten. Detta kan exempelvis ske genom att ta del av den informationsbroschyr som ges ut av Högskoleverket och den handbok som Enheten för pedagogiska mätningar ger ut samt att träna på gamla prov.

På senare år har ett antal privata företag erbjudit s.k. förberedelsekurser för högskoleprovet. Kurserna, som ofta är dyrbara, är inriktade på att kursdeltagarna ska höja sin provpoäng på kort sikt. Kursdeltagarna erbjuds bl.a. att lära sig tekniker som gynnar den egna provprestationen genom att träna på uppgifter som är så lika högskoleprovets uppgifter som möjligt samt att genomföra ett simulerat högskoleprov. Mot bakgrund av den forskning som gjorts av upprepat provdeltagande, och som mycket kort sammanfattats i föreliggande artikel, kan det sägas att den höjning av provpoäng som deltagarna utlovas, i de fall den sker, troligen beror på att man tränar på gamla prov och därigenom uppnår s.k. testförtrogenhet. En annan förklaring kan vara att man blir medveten om sina kunskapsbrister och därför ägnar sig åt kunskapsinhämtande på området. I stället för att betala dyrbara kursavgifter kan provdeltagarna uppnå samma effekt genom att själva öva på gamla prov eller att beställa den tidigare nämnda handboken och att sätta sig in i provsituationen och därmed bli testförtrogna.

Referenser

Henriksson, W., Henrysson, S., Stage, C. & Wedman, I., (1985). Prov för urval till högskolan. Stockholm: Utbildningsdepartementet, *SOU 1985:59*.

Henriksson, W. (1990). Effekter av upprepat provdeltagande. *Pm nr 40*. Pedagogiska institutionen, Umeå universitet.

Henriksson, W. & Bränberg, K. (1992). Effekter av upprepat provtagande. En studie av poängförändringar från första till andra provgenomförandet. *Pm nr 70*. Avdelningen för pedagogiska mätningar, Umeå universitet.

Henriksson, W., & Bränberg, K. (1994). The effects of practice on the Swedish Scholastic Aptitude Test (SweSAT). *Scandinavian Journal of Educational Research*, 38(2), 129-148.

Henriksson, W. (1995). Repeated Test Taking and the Swe SAT. *Em nr 13*. Division of Educational Measurement, Department of Education, Umeå university.

Författarpresentation

- Andersson Ewa** Doktorand vid Enheten för pedagogiska mätningar vid Umeå universitet.
Forskningsområde: Kriteriefrågor i samband med urval.
- Benjegård Maria** Biträdande forskare vid Pedagogiska institutionen vid Göteborgs universitet.
- Erickson Gudrun** Universitetsadjunkt vid institutionen för pedagogik, Avdelningen för språkpedagogik vid Göteborgs universitet.
Forskningsområde: Bedömning av språkfärdighet.
- Gustavsson Jan-Eric** Professor vid institutionen för pedagogik vid Göteborgs universitet.
Forskningsområde: Pedagogisk och psykologisk mätning, individuella differenser, utvärdering och multivariat analys.
- Hellekant Jan** Universitetsadjunkt vid institutionen för pedagogik, Avdelningen för språkpedagogik vid Göteborgs universitet.
Forskningsområde: Språkpedagogiska frågor, bl. a. könsskillnader i språkprov.
- Ohlander Sölve** Docent vid Engelska institutionen vid Göteborgs universitet.
Forskningsområde: Engelsk lexikologi, dvs. ordförrådet i engelska.
- Reuterberg Sven-Eric** Bitr. professor vid institutionen för pedagogik vid Göteborgs universitet.
Forskningsområde: Bakgrundsfaktors betydelse för val och framgång i utbildning utifrån longitudinella material.
- Stage Christina** Docent vid Enheten för pedagogiska mätningar vid Umeå universitet.
Forskningsområde: Könsskillnader i provresultat.
- Svensson Allan** Bitr. professor vid institutionen för pedagogik vid Göteborgs universitet.
Forskningsområde: Bakgrundsfaktors betydelse för val av och framgång i olika gymnasie- och högskolestudier utifrån longitudinella material.
- Wolming Simon** Doktorand vid Enheten för pedagogiska mätningar vid Umeå universitet.
Forskningsområde: Prognosforskning.
- Ögren Gunilla** Doktorand vid Enheten för pedagogiska mätningar vid Umeå universitet.
Forskningsområde: Kvantitativa prov.