

Validering av den högre utbildningens antagningssystem (VALUTA)

Inledning och syfte

Då antalet sökande till högre utbildning överstiger antalet platser får tillträdesystemets regelverk och formerna för antagning stor betydelse. När högskoleplatser är en bristvara blir tillträdesystemet, genom vilket behörighet definieras, kriterier för urval bland behöriga sökande fastställs och mätinstrument och metoder utformas, ett mycket värdeladdat inslag i hela utbildnings systemet. Dess utformning och funktion får återverkningar både inom och utanför den högre utbildningen.

Trots en radikal utbyggnad av antalet platser i grundläggande högskoleutbildning i Sverige (en fördubbling mellan läsåret 1988/89 och budgetåret 1997) har antalet sökande hela tiden överstigit antalet tillgängliga platser. Endast ca 50 % av de sökande har kunnat antas. Den svenska högskoleutbildningen har i sin helhet platsbegränsningar, och bortsett från de konstnärliga utbildningarna används ett fåtal enhetliga instrument för rangordning av sökande till samtliga utbildningar. Denna enhetlighet utgör ett intressant drag i utformningen av antagningssystemet, vilket, bland många andra, bör göras till föremål för studium (jmf Brandell & Kim, 2000).

Inom forskningen som studerat metoder och instrument för urval av individer till utbildningar har successivt allt bredare perspektiv utvecklats. Under första halvan av 1900-talet var forskningen i huvudsak fokuserad på instrumentutveckling och på att fastställa instrumentens prognostiska validitet gentemot olika kriterier. Runt 1950-talet kom många forskare dock att ifrågasätta fokuseringen på prognostisk validitet, och menade att frågor om instruments tolkning i termer av begrepp förankrade i psykologisk teori är mer fundamentala ur såväl vetenskaplig som ur praktisk synpunkt. Cronbach och Meehl (1955) lanserade termen begreppsvaliditet för detta utvidgade validitetsbegrepp, och det är nu en allmänt accepterad syn att begreppsvaliditet är ett överordnat begrepp, under vilket den prognostiska validiteten kan inordnas som ett specialfall. Under 1990-talet har emellertid validitetsbegreppet, framförallt genom Messicks (1989) arbete, kommit att få en betydligt vidgad innebörd, och det är denna vidgade innebörd som utgör den gemensamma teoretiska ramen för det föreslagna forskningsprogrammet.

Medan begreppsvaliditet primärt refererar till empiriskt baserade tolkningar av mätinstruments innebörd, menar Messick (1989) att det därutöver är nödvändigt att avgränsa tre ytterligare aspekter av validitetsbegreppet. Instrument för urval till utbildning är sålunda utvecklade för praktiskt bruk, och den faktiska användningen av mätinstrument, och effekterna av detta, är, menar Messick, en andra validitetsaspekt som kan och bör studeras med empiriska metoder. Men konstruktioner som exempelvis betygssystem och prov för att mäta olika förmågor har också sociala konsekvenser vilka måste uppmärksammas. Dessa påverkar tänkande och värderingar på samhälls nivå, och de har konkreta konsekvenser för olika grupper av människor. De värdemässiga implikationerna identifierar Messick som en tredje validitetsaspekt, och de sociala konsekvenserna som den fjärde aspekten. Dessa två aspekter kan sammanföras under termen konsekventiell validitet.

Studium av systemet för tillträde och rekrytering till högre utbildning enligt Messicks modell kräver sålunda att alla fyra validitetsaspekterna uppmärksammas, vilket är syftet med det integrerade forskningsprogram som presenteras i denna ansökan. Förslaget innebär samverkan mellan två institutioner (Institutionen för pedagogik och didaktik vid Göteborgs universitet och Enheten för pedagogiska mätningar vid Umeå universitet) vilka båda har mångårig erfarenhet av forskning inom området.

Det övergripande syftet med forskningsprogrammet är att undersöka hur antagningssystemet fungerar i relation till samhällets mål och intentioner och vilka effekterna av detta system är för enskilda individer,

för olika grupper av individer, för högskolan samt för samhället som helhet, dvs för samtliga avnämare till antagningssystemet.

Det är möjligt att identifiera olika komponenter i antagningssystemet:

I	Urvalsinstrumenten
II	Kriteriet/Studieframgång
III	Urvalsinstrumenten i relation till kriteriet
IV	Avnämarna
V	Regelsystemet

Inom programmet kommer samtliga komponenter att studeras liksom även interaktionen mellan dem. Forskningsprogrammet är uppbyggt kring 13 delprojekt, tillsammans med en integrerande del. På grundval av de två medverkande institutionernas erfarenhets- och kompetensprofil har en arbets- och ansvarsfördelning gjorts så att Göteborgsinstitutionen i första hand fokuserar på empiriskt baserade studier av urvalsinstrumentens innebörd (dvs begreppsvaliditet) och på studier av användningseffekter, medan enheten i Umeå i första hand fokuserar på den konsekventiella validiteten. Här finns dock av naturliga skäl en betydande överlappning, och avsikten är att de delprojekt som är förlagda till de två institutionerna skall bedrivas i nära samverkan. Planeringsarbetet för detta forskningsprogram har genomförts under det senaste året, men grunden för förslaget är en mångårig samverkan kring pedagogiska mätfrågor mellan den pedagogiska institutionen i Göteborg och nuvarande Enheten för pedagogiska mätningar i Umeå.

Tidigare forskning

Sedan 1960-talet har forsknings- och utvecklingsarbete avseende urvals- och tillträdesfrågor intagit en central plats vid Umeå universitet. Sten Henrysson fick då i uppdrag från Kompetensutredningen att utveckla ett studielämplighetsprov för urval till högre studier och efter ett omfattande forsknings- och utvecklingsarbete under 1970-talet kunde det första högskoleprovet administreras 1977. Alltsedan dess har konstruktions-, forsknings- och utvecklingsarbete avseende Högskoleprovet bedrivits i Umeå. Som ett exempel kan nämnas att Sten Henrysson tillsammans med medarbetare i Umeågruppen, på uppdrag av Tillträdesutredningen, deltog i utredningsarbetet avseende prov för urval till högskolan. (SOU 1985: 59). Denna utredning baserades på kunskap och erfarenheter som fanns och fortfarande finns i Umeå.

Frågor i anslutning till högskoleprovet, samt dess konstruktion och resultat, har varit föremål för intensiva studier med olika inriktning. En nyligen sammanställd rapportförteckning visar att ungefär 200 rapporter hittills har skrivits av Umeå-gruppen (Eriksson, 2000). Som framgår av Erikssons sammanställning (2000) har högskoleprovets mättekniska aspekter beskrivits i ett 40-tal rapporter, där såväl provets faktoriella struktur som mätspecifika egenskaper har studerats. Ett annat område som studerats ingående är könsskillnader i provresultat, som har analyserats på såväl totalprovs-, delprovs som uppgiftsnivå på ett flertal olika sätt, vilket finns beskrivet i ett 40-tal rapporter. Ett annat relevant område är upprepat provdeltagande, testträning och testvana och inom detta område finns ett 20-tal rapporter skrivna. Även prognosstudier har förekommit tämligen frekvent och även på det området finns närmare 20 rapporter. Relationen mellan betyg och provpoäng har beskrivits i 7 rapporter. Därutöver har studier genomförts avseende utprövningsrutiner, provinnehåll, uppgifts- och provformat, provtider, nya provtyper etc. Den forskningstradition som byggts upp under en 30-årsperiod har således både djup och bredd. Djup i den bemärkelsen att en ingående kunskap har etablerats om såväl konstruktionsteoretiska som provteoretiska frågor. Bredd i den meningen att det alltid finns en strävan att enskilda instrument, som prov och betyg, ska analyseras och värderas i ett större sammanhang. Sammanfattningsvis kan konstateras att en unik och synnerligen omfattande kunskap och erfarenhet om framför allt högskoleprovet, ett av de två viktigaste urvalsinstrumenten i dagens antagningssystem, har utvecklats och finns i Umeå.

I högskoleprovets databank finns, för samtliga provdeltagare sedan 1977 resultat på uppgifts-, delprovs- och totalprovsnivå. För samtliga provdeltagare finns även uppgift om ålder, kön och utbildningsbakgrund, den senare angiven både som högsta avslutade utbildning och pågående utbildning uppdelad på gymnasieprogram. Denna databank har redan tidigare använts i ett flertal studier av såväl Umeå- som Göteborgsgrupperna och denna databank kommer att användas även i detta forskningsprogram.

Flera av de delprojekt som beskrivs i denna ansökan knyter samman olika forskningstraditioner vilka under lång tid utvecklats inom den pedagogiska forskningen vid Göteborgs universitet. En sådan tradition utgörs av empiriska studier med i huvudsak longitudinell uppläggning, med fokus dels på frågor av utbildningssociologisk art (t ex den sociala bakgrundens betydelse för val och prestationer i utbildningssammanhang), dels på frågor kring individuella differenser i förutsättningar för utbildning och i utbildningsresultat. I anslutning till dessa empiriska och teoretiska inriktningar har också intresse riktats mot metodfrågor och metodutveckling, och då särskilt kring mät- och analysfrågor. Nedan beskrivs denna tidigare forskning kortfattat.

År 1961 etablerades den första kohorten inom det s k Individualstatistikprojektet (Härnqvist & Svensson, 1973) genom att ett slumpmässigt urval om ca 12 000 individer födda 1948 i åk 6 bjöds en uppsättning enkäter och begåvningsstest. Denna kohort har sedan följts genom utbildningssystemet och upp i vuxen ålder, och har givit underlag för ett mycket stort antal studier. Flera nya kohorter har därefter förts in i projektet, vilket med kohorten född 1972 fick namnet UGU (Utvärdering genom uppföljning). De longitudinella UGU-materialen har utgjort grund för ett mycket stort antal studier av utbildningssociologiska frågeställningar, kring bl a den sociala snedrekryteringen till högre utbildning (t ex Härnqvist, 1994; Härnqvist & Svensson, 1980), och av betydelsen av individens egenskaper i olika hänseenden för prestationer och för val av utbildning (t ex Härnqvist, 1998, Reuterberg, 1996, Svensson & Stahl, 1996). Under de senaste åren har dessa stickprovsundersökningar kompletterats med totalundersökningar baserade på samkörningar av registerdata. En databas (UGU-R), vilken omfattar samtliga födda 1972 – 1979 har sålunda skapats. Detta material har nyligen blivit tillgängligt för analys, och har givit prov på intressanta möjligheter att med totalmaterial studera tendenser och skeenden som eljest skulle vara svåra att upptäcka (Gustafsson, Andersson & Hansen, 2000).

En annan forskningstradition inom Göteborgsinstitutionen har fokuserat på individuella differenser i förutsättningar för utbildning, och då i synnerhet på kognitiva förmågor. Tidiga insatser gjordes av Härnqvist (1960) och under de två senaste årtiondena har Gustafsson (t ex 1984, i tryck-a, i tryck-b; se också Gustafsson & Undheim, 1996) genomfört en serie studier kring de kognitiva förmågornas struktur. Ett huvudresultat från denna forskning är en hierarkisk modell som beskriver kognitiva förmågor av olika generalitetsgrad, från en generell kognitiv förmåga, med tyngdpunkt på icke-verbal problemlösningsförmåga, över ett fåtal breda kognitiva förmågor (t ex en bred verbal förmåga och en bred spatial förmåga) till ett stort antal specifika förmågor. Denna modell har bl a visat sig vara användbar som en referensram för studier av olika instruments begreppsvaliditet, som t ex Högskoleprovet (Gustafsson, Wedman & Westerlund, 1992) och grundskolebetygen (Andersson, 1998).

Forskningen om de kognitiva förmågornas struktur har baserats på de metodmässiga framsteg som gjorts inom den multivariata statistiska modelleringen, och då framförallt de tekniker som går under samlingsnamnet strukturell ekvationsmodellering. Inom detta fält utgör den svenske statistikern Karl Jöreskog och hans elever (t ex Bengt Muthén och Dag Sörbom) internationella förgrundsgestalter. Informationsteknologins snabba utveckling har gjort att utvecklingen inom området accelererat ytterligare, och de metoder som nu finns tillgängliga tillåter hantering av data med mer komplex struktur än vad som tidigare var fallet. Inte minst inger framstegen vad gäller tekniker för att hantera inkompleta data (s k "missing-data"-modellering, Arbuckle, 1996; Gustafsson & Stahl, 1999; Muthén & Muthén, 1998) stora förhoppningar att några av de fundamentala metodproblemen vad gäller studier av urvals instruments egenskaper nu skall kunna lösas.

Uppläggnig och genomförande

Programmet hålls samman av den ovan beskrivna gemensamma övergripande Messickska validitetsmodellen. De olika projekt, som planeras ingå i programmet, skall anläggas ett eller flera av de perspektiv som angetts ovan.

Flera delprojekt har också en empirisk bas i den longitudinella databas som etablerats i Göteborg. I databasen ingår, som redan nämnts, samtliga individer födda 1972-1979 (N=842 800). Bland variablerna finns uppgifter om elevernas betyg i enskilda ämnen från grundskolans åk 9, tillsammans med vissa demografiska uppgifter. Från FoB finns uppgifter om föräldrarnas utbildning och yrke. Därtill finns från högskoleprovsregistret i Umeå resultat på uppgiftsnivå på Högskoleprovet för de som valt att delta i ett eller flera prov. För pojkar finns resultat från Värnpliktsverkets register på de delprov som ingår i de militära inskrivningsproven. Från SCBs register har inhämtats uppgifter om val av gymnasielinje och gymnasiebetyg, liksom sökalternativ och utfall vid ansökan till högskoleutbildning. Vidare finns uppgifter om studieresultat från högskolans LADOK-system, liksom uppgifter om sysselsättning från det s.k. ÅRSYS-registret. Två av de kohorter som ingår i UGU-R sammanfaller med UGU-projektets stickprovsundersökningar.

Dessa mycket rikliga data kan omedelbart göras till föremål för analys, vilket också kommer att ske inom programmet. Icke desto mindre framstår det som angeläget att göra kompletteringar av de nu tillgängliga data. Under början av 2001 finns det sålunda möjlighet att lägga till ytterligare 4 kohorter, och det är också angeläget att komplettera data med ytterligare variabler (t ex mer detaljerade uppgifter om antagningsgrupp, och uppgifter om KOMVUX och utlandsstudier). Ett ytterligare skäl varför databasen bör utvecklas vidare är att en lagändring är att förvänta, vilken kommer att tillåta SCB att bevara identifikationsnyckeln för de avidentifierade forskningsdatabaserna. Detta gör det möjligt för SCB att i efterhand lägga till ytterligare uppgifter. Denna omständighet gör det angeläget att SCB kan ges i uppdrag att rekonstruera hela den nuvarande databasen för att erhålla denna kompletteringsmöjlighet. Eftersom kodnyckeln inte finns tillgänglig måste hela proceduren för sammanslagningen upprepas. Detta kan emellertid ske utan alltför omfattande arbete och alltför stora kostnader, eftersom de rutiner som utvecklats kan användas på nytt. Arbetet med komplettering och vidareutveckling av databasen kommer att förläggas till projektets första år, men arbete med underhåll, dokumentation och datauttag kommer att ske under programmets hela löptid. Databasen kommer härigenom att utgöra en viktig grund även för framtida studier utanför programmets ram.

Nedan presenteras i korthet de i programmet ingående delprojekten insorterade under antagningssystemets komponenter. Avslutningsvis planeras en integrering av samtliga delprojekt inom den gemensamma teoriramen.

I. Urvalsinstrumenten

Egenskaperna hos de urvalsinstrument som används i ett antagningssystem är av avgörande betydelse för vilka individer som väljs ut till olika utbildningar, och vilka som tvingas stå utanför de eftertraktade utbildningarna. I forskningsprogrammets första, och i termer av antalet delprojekt, största del, studeras urvalsinstrumentens mätegenskaper, med särskilt fokus på begreppsvaliditet. Det urvalsinstrument som särskilt fokuseras är högskoleprovet, men även de nya gymnasiebetygen och särskilda urvalsprov ägnas uppmärksamhet.

Högskoleprovet

Tre delprojekt fokuserar på Högskoleprovets mätegenskaper.

Delprojekt 1: Högskoleprovets interna struktur

Ett test eller urvalsinstrument är inte homogent i det avseendet att det kan förväntas mäta endast en generell kognitiv förmåga. Inte heller sönderfaller det alltid i *a priori* avsedda/förutsedda faktorer. Tidigare arbeten med modellering av Högskoleprovet visar sålunda att provet mäter minst två faktorer (Gustafsson, Wedman & Westerlund, 1992). Detta gör det angeläget att i större detalj med hjälp av flerdimensionella modeller kartlägga Högskoleprovets inre struktur.

Medan tidigare studier i huvudsak analyserat Högskoleprovet struktur på delprovsnivå kommer de planerade analyserna framförallt att fokusera på uppgiftsnivån. En anledning till detta är att sådana analyser är betydligt mer informativa än analyser på aggregerad nivå. Detta illustreras av en studie som tidigare genomförts av Åberg-Bengtsson (1999). Ett huvudresultat i denna studie var att vissa av uppgifterna i delprovet DTK influeras av en numerisk förmåga vid sidan av den problemlösningsförmåga som dominerar huvuddelen av uppgifterna. Denna numeriska dimension uppvisade också betydligt större könsskillnader än den andra dimensionen. Studier av detta slag kan endast genomföras på uppgiftsnivå. Ett ytterligare skäl för varför det nu är aktuellt att göra analyser av Högskoleprovets interna struktur på denna nivå är att effektiva algoritmer utvecklats och implementerats vilka gör det möjligt att bygga modeller som omfattar både kontinuerliga variabler och kategorivariabler (Muthén & Muthén, 1998).

I de planerade studierna kommer samtliga delprov att göras till föremål för analyser av den interna strukturen på uppgiftsnivå. Fortsatt arbete kommer sålunda att göras kring den numeriska dimensionen i DTK-provet, varvid flera provomgångar kommer att analyseras och även det mer utpräglade numeriska NOG-provet kommer att inkluderas i modellerna.

LÄS-provet är intressant ur två synpunkter. En fråga av stort intresse är vilken grad av inflytande på delprovets dimensionalitet de texter som ingår i provet har. Om det är möjligt att identifiera separata dimensioner för varje text, vilket preliminära studier visar är fallet, är det intressant att studera om andra variabler som ålder, kön, utbildningsbakgrund och social bakgrund har samband med dessa textdimensioner. En annan intressant fråga är om det är möjligt att identifiera en snabbhetsdimension av liknande slag som identifierats i andra läspröv (Gustafsson, 1997, 1998). Det faktum att provdeltagarna rapporterar en självupplevd tidsbrist i LÄS-provet antyder att så kan vara fallet. Även för det engelska läsförståelseprovet ELF är dessa två frågor av intresse att studera, men eftersom detta delprov uppvisar en större variation i uppgiftsmaterialets utformning finns här även möjlighet att studera betydelsen av faktorer som lästexternas omfattning och olika svarsmodeller.

ORD-provet utgör en stor del av uppgifterna i Högskoleprovet. Som visats av Gustafsson och Holmberg (1992) har urvalet av ord som ingår i provet betydelse för vilka egenskaper som provet mäter. I pågående arbete i annat provsammanhang har Ullstadius, Gustafsson & Carlstedt (2000) utnyttjat Muthén och Muthén's (1998) nyutvecklade tekniker för analys av dikotoma uppgifter, och därvid funnit ett mycket högt samband mellan orduppgifters lösningsfrekvens och deras relation till den generella begåvningsdimensionen. Denna typ av analyser kommer att genomföras även för Högskoleprovets ORD-prov, och resultaten kan förväntas vara såväl av teoretiskt som praktiskt intresse.

Avsikten är även att i ett nästa steg anknyta de olika modellerna till variabler hämtade från tillgängliga register för att närmare belysa identifierade dimensioners karaktär och egenskaper.

Delprojekt 2: Högskoleprovet och andra instrument för att mäta kognitiva förmågor

Ytterligare ett sätt att vinna insikt om vilka förmågor och färdigheter som Högskoleprovet mäter är att relatera det till andra instrument för mätning av kognitiva förmågor, vars mätegenskaper är kända och möjliga att förstå inom ramen för en teoretisk modell. En sådan möjlighet erbjuds av de militära inskrivningsproven, vilka bjudits till huvuddelen av de män som tagit Högskoleprovet. Fram till 1994 omfattade inskrivningsproven 4 delprov (ett verbalt, ett induktivt och två tekniska/spatiala prov). Från 1994 omfattar inskrivningsproven 10 delprov vilka utvecklats för att mäta tre hierarkiskt ordnade dimensioner

(generell förmåga, verbal förmåga och spatial förmåga) enligt den tidigare omnämnda hierarkiska modellen (Mårdberg & Carlstedt, 1998).

Genom att med hjälp av multivariata modeller relatera Högskoleprovet till dessa dimensioner kan Högskoleprovets mätegenskaper bestämmas i termer av den begreppsapparat som utvecklats inom den omfattande forskningen om intellektets struktur (Gustafsson, 1988, i tryck-a, i tryck-b; Gustafsson & Undheim, 1996). Sådana studier kan sålunda besvara frågan om den analytiska dimensionen som Gustafsson, et al. (1992) identifierade i Högskoleprovet sammanfaller med den icke-verbala problemlösningsförmåga som benämns "fluid intelligence" och som visats vara identisk med den allmänna intelligensfaktorn (Gustafsson, 1984, i tryck-c). Högskoleprovet har också visats mäta en verbal förmågedimension, och även vad gäller denna är det av intresse att undersöka samband med den verbala dimension som mäts av inskrivningsproven. De senare omfattar också spatiala prov av ett slag som inte ingår i Högskoleprovet. Detta gör det möjligt att studera i vilken utsträckning spatiala prov mäter förmågor som är fristående från vad som mäts med Högskoleprovet, och om dessa har samband med studieframgång inom vissa utbildningar (t ex arkitektutbildning, tekniska utbildningar),

Delprojekt 3: Gruppkillnader i Högskoleprovresultat

Utfallet i resultat på högskoleprovet för män respektive kvinnor har undersökts i en mängd studier på såväl totalprovs- som delprovs- och uppgiftsnivå (se t ex Stage, 1985, 1992c, 1993, 1997) Genom dessa studier har även metoder för att studera gruppkillnader prövats, utvecklats och förfinats – metoder som kan och bör användas även för andra gruppjämförelser. Kunskapen är tämligen bristfällig om skillnader mellan provdeltagare med olika social bakgrund. Det har konstaterats att delprovet ELF ger upphov till de största skillnaderna följt av delprovet NOG (Hansen, 1997). På uppgiftsnivå är kunskapen om socialgruppskillnader närmast obefintlig. Det vore därför angeläget att, på samma sätt som har gjorts avseende könsskillnader, genomföra systematiska analyser även av socialgruppskillnader på uppgiftsnivå.

De metoder som avses användas är DIF-analyser (differential item functioning) med såväl Mantel-Haenszel-metoden (se t ex Stage, 1997) som IRT (item response theory)-teknik. Det som är av betydelse ur testkonstruktionssynpunkt är nämligen inte om det föreligger skillnader mellan grupper i allmänhet utan om skillnaderna kvarstår efter kontroll för skillnader i förmåga. En fördel med Mantel-Haenszel-metoden i detta sammanhang är att den erbjuder både mått på effektstorlek och statistisk signifikans och dessutom är tämligen lätt att använda. IRT-metoden har dock kompletterande fördelar, nämligen att den ger ytterligare information om konstaterade skillnader. Det är nämligen möjligt att på uppgiftsnivå konstatera om skillnaderna föreligger mellan låg-, medel- eller högpresterande eller om skillnaderna är jämnt fördelade på alla prestationsnivåer.

En hypotes, som har väckts i samband med granskning av provuppgifter och som också avses prövas i detta sammanhang är om det föreligger en interaktion mellan kön och socialgrupp, så att uppgifter som "gynnar" kvinnor samtidigt gynnar högre socialgrupper. Det skulle innebära att den ansträngning som har gjorts att minska könsskillnaderna i resultat på högskoleprovet har orsakat ökade skillnader mellan sociala grupper. Det finns även anledning till fortsatta studier av könsskillnader i provresultat. Dessa studier bör dock genomföras avseende de skilda dimensioner som mäts i provet. Vid analys av könsskillnader på uppgiftsnivå har med hjälp av IRT-teknik konstaterats att en uppgift kan t ex vara lättare för kvinnor än män på låg prestationsnivå, medan den är svårare för kvinnor än för män på hög prestationsnivå. (Stage, 1999). Detta skulle kunna tolkas så att en uppgift kan innehålla olika dimensioner och att det leder till att uppgiften fungerar olika på olika förmågenivåer. Detta är ett område som är mycket angeläget att klarlägga.

Gymnasiebetygen

De relativa gymnasiebetygen som användes fram till 1996 var utvecklade för att fungera som urvalsinstrument. Det kriterierelaterade betygssystem som infördes 1997 ger emellertid anledning att ställa frågor om de nya betygens funktion som urvalsinstrument. De nya betygen är målrelaterade till skillnad från de gamla som var normrelaterade; de nya betygen anges inte som de gamla i siffrvärden och vidare är kriterierna för olika betygssteg tämligen inexakta och för det högsta betygssteget finns ännu inga centralt utformade kriterier, vilket innebär att kraven för detta betyg kan variera mellan olika skolor. Införandet av de nya gymnasiebetygen väcker sålunda en lång rad frågor om vilka mätegenskaper som kännetecknar dessa i förhållande till de relativa betygen. Syftet med detta delprojekt är att ur ett begreppsvaliditetsperspektiv undersöka mätteoretiska aspekter på gymnasiebetygen.

Delprojekt 4: De kriterierelaterade gymnasiebetygens mätegenskaper

En serie analyser kommer att ta sin utgångspunkt i det faktum att införandet av det nya betygssystemet i grundskolan skett senare än i gymnasieskolan. De första kriterierelaterade grundskolebetygen gavs sålunda vt 1998, vilket innebär att tre kohorter (dvs de födda 1978 – 1981) erhållit kriterierelaterade gymnasiebetyg och relativa grundskolebetyg. Detta faktum gör det möjligt att utnyttja multivariata metoder för att bestämma mätegenskaper hos såväl de nya som de gamla gymnasiebetygen genom analys av deras samband med de relativa grundskolebetygen. I dessa analyser kommer också variabler som kön och social bakgrund att inkluderas, och även skolskillnader kommer att undersökas. Det torde också vara möjligt att använda de relativa grundskolebetygen för en kalibrering av de nya gymnasiebetygen gentemot de relativa gymnasiebetygen.

En annan serie analyser kommer att genomföras där såväl de gamla gymnasiebetygen som de nya gymnasiebetygen kommer att relateras till resultat på högskoleprovet. Eventuella skillnader i dessa sambandsmönster är intressanta såväl ur teoretisk som ur praktisk synvinkel.

Särskilt urval

Förutom betyg och högskoleprov kan högskolorna använda särskilt prov som urvalsgrund vid antagning till utbildningar som förbereder för yrken som ställer krav på vissa personliga egenskaper eller särskilda kompetenser. Avsikten är att egenskaper eller kompetenser som den sökande inte alls eller i otillräcklig omfattning kunnat visa i gymnasieskolan och som inte heller mäts via högskoleprov skall kunna beaktas. Såväl utformandet som genomförandet av en alternativ urvalsprocedur är behäftad med en komplex validitets- och reliabilitetsproblematik. Vilka kvalifikationer hos den sökande ska provet mäta? Hur ska dessa egenskaper eller kompetenser mätas? Då kunskapen kring såväl effekter av alternativa urvalsformer som av de enskilda instrumentens mätegenskaper är otillräcklig är forskning inom området angelägen utifrån både system- och individperspektiv.

Delprojekt 5: Egenskaper hos vissa särskilda prov

Ett syfte med delprojektet är att studera effekter av antagning via särskilt prov med avseende på dels sådana faktorer som de rekryterades utbildningsbakgrund, kön, sociala bakgrund och etniska härkomst, dels studieframgång. Ytterligare ett syfte är att studera de enskilda instrumentens mättekniska egenskaper med särskild fokus på intervju och de övriga metoder som används i samband med antagning till läkarutbildning, och på de särskilda prov som används vid antagning till arkitektutbildningen. Läkar- och arkitektutbildningarna har använt sig av alternativt urval för en del av studieplatserna under större delen av 1990-talet, vilket innebär att ett tillräckligt stort material nu kumulerats för att göra analyser möjliga. Undersökningarna kommer i huvudsak att genomföras på grundval av registerdata där utgångspunkten är antagningsgrund, och där analyserna fokuseras på skillnader i olika bakgrunds- och prestationsvariabler mellan de grupper som antagits till utbildningarna på grundval av särskilda prov, och de grupper som antagits med de reguljära instrumenten. För arkitektutbildningarna används samma instrument för alla tre högskolor och varje år antas ett relativt blygsamt antal elever. Vad gäller läkarutbildningarna föreligger dock

vissa skillnader mellan de urvalsystem som används vid olika orter, och här är antalet antagna också större, vilket gör det intressant att göra jämförelser mellan olika orter.

II. Kriteriet/Studieframgång

Ett av de stora problemen i studier av olika instruments förmåga att predicera studieframgång är att finna goda mått på studieframgång (Henriksson & Wedman, 1992; Henrysson, 1984). I avsaknad av optimala mått har studier inom området använt tämligen grova mått på studieframgång, t ex erhållna poäng efter ett års studier, godkänd eller icke godkänd examen (SOU 1968:25). Med denna typ av kriterier tenderar man dock att erhålla låga samband med olika prediktorer.

Studier på området, har visat att studieframgång kan anta många olika former (jfr Lenning et al., 1974; Taber & Hackman, 1976; Hackman & Taber, 1979; Willingham, 1985). Även studier som har behandlat genomströmningsproblematiken har pekat på behovet att problematisera studieframgångsbegreppet (se exempelvis Dahllöf, 1968; Elgqvist-Saltzman, 1979; Henrysson, 1984). Trots detta har innebörden av studieframgång inte varit föremål för något större intresse. Detta är anmärkningsvärt med tanke på de skillnader som existerar mellan olika högskoleutbildningar såväl avseende innehåll, utformning, examination som betygssystem (Andersson & Grysell, 1995). Därutöver skiljer sig även utbildningarna på andra sätt, t ex avseende status och koppling till offentlig respektive privat sfär (se exempelvis Broady, 1987). De studerande uppvisar även mycket skilda motiv till valet att påbörja högre studier, och olika studiemönster där de s.k. icke-traditionella studenterna utgör en växande grupp (se t ex Brandell, 1998; Wästerfors, 1998).

Delprojekt 6: Traditionella och icke traditionella kriterier för studieframgång

Det finns därför anledning att problematisera innebörden av begreppet studieframgång. Ett sätt att närma sig begreppet är att dela in kriterierna i traditionella respektive icke traditionella kriterier. Med traditionella kriterier avses här de kriterier som hittills vanligtvis använts för att mäta studieframgång, vilka i huvudsak utgjorts av kvantitativa mått, t ex uppnådda högskolepoäng. Med icke traditionella kriterier avses här t ex personlig utveckling, förmåga till kritisk reflektion och samarbetsförmåga. Denna typ av kriterier är mer svårångade men viktiga, och då inte minst med tanke på att den högre utbildningens mål ofta beskrivs i sådana termer. Avsikten är att genom en litteraturgenomgång kartlägga båda dessa kategorier av kriterier.

Henrysson (1984) ställde två övergripande krav på ett kriterium, där det första relaterades till kriteriets mättekniska kvaliteter och det andra till kriteriets relevans. Även om det inte föreligger någon absolut skiljelinje mellan de mättekniska kraven och kravet på relevans är de förra lättare att bedöma. Till de mättekniska kraven kan exempelvis reliabiliteten för kriteriet och dess skalnivå hänföras. Inom högskolan tillämpas olika examinationsformer vilket innebär att studieframgången mäts på olika sätt. Det faktum att lärare ofta betygssätter studenterna på olika sätt medför brister i kriteriets reliabilitet. Även den skala kriteriet uttrycks i kan vara problematisk utifrån ett mättekniskt perspektiv. Dikotoma kriterier (exempelvis underkänd - godkänd) ger mindre information än en mer fingraderad kriterieskala då de medför att variationen mellan studenternas prestationer blir låg. Utifrån ett mättekniskt perspektiv är således en reliabel och mer fingraderad skala att föredra.

Att bedöma om en kriterievariabel är relevant eller ej är däremot betydligt mer problematiskt. Olika avnämare bedömer detta från olika perspektiv, med påföljd att vad som för en grupp är ett kriterium med en hög grad av relevans, kan sakna relevans för en annan grupp. Frågan om olika kriteriers relevans kommer därför att bedömas i relation till de resultat som framkommer i avnämarestudien i delprojekt 10.

Inom ramen för det här aktuella delprojektet kommer även de traditionellt använda kriterierna för studieprestationer att närmare granskas. I de lokalt hanterade LADOK-registren (Lokalt ADB-baserad

dokumentationssystem) finns en avsevärd mängd information om de antagna studenterna. Informationen gäller bland annat deras antagning, resultat på högskoleprov och gymnasiebetyg samt information om deras studieprestationer, vanligtvis i form av antal högskolepoäng. En fråga som är relevant är vad dessa yttre mått på studieprestation egentligen står för? Tanken är här att närmare granska dessa mått utifrån det bakomliggande innehållet, dvs vad representerar dessa kriterier sett utifrån ett utbildningsperspektiv? Vilka är de kriterier (examinationsformer och deras fokus) som ligger till grund för den prestation som finns uttryckt i LADOK? I detta sammanhang blir det således relevant att undersöka om det finns andra kriterier som inte kommer till uttryck i LADOK? Ytterligare en fråga är om den bakomliggande innebörden av de yttre måtten skiljer sig åt beroende på vilken utbildning som granskas? Två utbildningslinjer vid högskolan (ekonomutbildning och socialt arbete) kommer att väljas ut och studeras närmare utifrån denna aspekt.

Syftet med föreliggande delprojektet är sålunda att upprätta en modell som består av de ovan beskrivna komponenterna (relevans, mättekniska krav och överensstämmelse med antagningssystemets målsättningar) och utifrån denna analysera de kriterier som framkommer i litteraturgranskningen. Därutöver kommer en empirisk studie att genomföras, dels i form av en genomgång av befintliga kriteriemått i LADOK, dels i form av en analys av de kriterier som olika utbildningar använder sig av för att bedöma innebörden av studieframgång.

III. Urvalsinstrumenten i relation till kriteriet

Det resursfördelningssystem som infördes med 1993 års högskolereform har medfört ett ökat intresse för betygens och Högskoleprovets förmåga att förutsäga studieframgång. Att bestämma ett urvalsinstruments prognostiska värde möter emellertid stora teoretiska och metodiska problem.. Ett sådant är att empiriskt bestämma graden av samband mellan prognosinstrumentet och kriterierna, eftersom kriteriemåtten är kända enbart för dem som blivit uttagna till yrket eller utbildningen. Denna grupp är självfallet positivt selekterad och därmed är variationen beskuren, vilket leder till en underskattning av det faktiska sambandet mellan prognosinstrumentet och kriterierna (Lord & Novick, 1968). Om urvalet sker på grundval av ett enda prognosinstrument (enkelt urvalsförfarande) finns kända metoder med vars hjälp man kan korrigera för den beskurna variationen och därigenom få en skattning av det samband som skulle ha erhållits om samtliga sökande blivit uttagna.

Den svenska urvalsmodellen till högre utbildning är emellertid mera komplicerad därigenom att den sker utifrån två urvalsinstrument, där man kan bli antagen antingen på grundval av gymnasiebetyg eller på grundval av högskoleprovresultat (kompensatoriskt urvalsförfarande). I ett sådant fall har vi inte enbart att räkna med den beskurna variationen utan här får vi också en inverkan av graden av samband mellan de två urvalsinstrumenten samt deras samband med kriteriet och dessutom av de lägsta poänggränser som krävs på de två urvalsinstrumenten för antagning. Under vissa betingelser kan detta leda till att man finner tämligen starka negativa samband mellan kriteriet och ett urvalsinstrumenten inom den antagna gruppen trots ett relativt starkt positivt mellan urvalsinstrumentet ifråga och kriteriet bland samtliga sökande.

Genom de metodmässiga framsteg som gjorts med den så kallade "missing-data"-modelleringen (Arbuckle, 1996; Gustafsson & Stahl, 1999; Muthén, Kaplan & Hollis, 1987) är det numera möjligt att från den information som finns tillgänglig för gruppen antagna och den information som finns tillgänglig om de ej antagna vad gäller sambanden mellan urvalsinstrumenten och kriteriet skatta de faktiska sambanden mellan urvalsinstrumenten och kriterierna. Med de faktiska sambanden menas då de samband som skulle ha erhållits om samtliga sökande blivit antagna. Eftersom "missing-data"-metoden hittills är oprövad i de sammanhang som gäller här har den prövats i en serie simuleringsstudier. Dessa har visat att den ger synnerligen säkra skattningar av de faktiska sambanden (Gustafsson & Reuterberg, 2000; Reuterberg, 2000).

Med hjälp av de här beskrivna metoderna kommer följande problemställningar att belysas inom tre delprojektet.

Delprojekt 7: Prognosvärdet för betyg respektive högskoleprov, samt genomströmning och examination inom olika utbildningar

Enligt det nuvarande urvalssystemet sker urvalet utifrån medelbetyget från gymnasieskolan och utifrån det resultat som uppnås på högskoleprovet i sin helhet. Detta enhetliga urvalsförfarande föranleder en rad frågor. Hur skiljer sig prognosinstrumentens prediktiva förmåga mellan olika typer av högskoleutbildningar, t ex utbildningar inom de olika utbildningssektorerna? Finns det skillnader i prediktiv förmåga mellan olika kategorier av studerande, t ex mellan män och kvinnor, mellan studerande med olika social eller etnisk bakgrund, mellan studerande på olika åldersnivåer? Varierar den prediktiva förmågan mellan olika lärosäten inom en och samma utbildning? Har den prediktiva förmågan förändrats över tid, dvs över de årskullar som ingår i databasen? Hur förändras den prediktiva förmågan över studietiderna?

En annan intressant fråga är om man kan förbättra prognosvärdet genom att utnyttja urvalsinstrumenten på ett mera flexibelt sätt. Varken medelbetyget eller totalpoängen på högskoleprovet är homogena i den bemärkelsen att de mäter en enda generell förmåga. Bl a har Gustafsson, Wedman och Westerlund (1992) visat att provet mäter minst två faktorer. När analyserna i delprojekt 1 ovan är redovisade blir det naturligt att studera vilket prognostiskt värde de olika dimensionerna har för olika typer av högskoleutbildningar. Om det visar sig att de olika dimensionerna har olika prognostiskt värde för olika utbildningar har man fått ett underlag för att bedöma i vad mån den prediktiva förmågan kan höjas genom att låta olika betyg eller olika delar av högskoleprovet få en större vikt än andra för urvalet till en viss utbildning.

En annan metodisk utveckling som skett under senare år, nämligen den som gäller analys av tillväxtkurvor (Muthén & Muthén, 1998) har skapat goda förutsättningar för att beskriva och förklara skillnader i tillväxt. Exempel på sådana tillväxtkurvor är t ex den enskilde studerandes ackumulerade poängproduktion över studietid och avsikten är i denna del av projektet att tillämpa den nya statistiska metodiken på den ackumulerade poängproduktionen.

Metodiken innebär att man som utgångspunkt tar den studerandes samlade poäng efter t ex en termins studier (A) och därefter bestäms den tillväxtfaktor som bäst beskriver förändringen i ackumulerad poäng över studietid (B). A och B behandlas sedan som latenta variabler i en regressionsmodell varvid de kan relateras till andra faktorer i modellen. Som exempel på sådana andra faktorer kan nämnas bakgrundsvariablerna kön, social bakgrund och etnisk tillhörighet. Det kan även gälla sådana faktorer som resultaten på urvalsvariablerna eller egenskaper hos den utbildning som studeras. De relationer man därvid finner beskriver således hur A och B påverkas av de andra faktorerna i modellen.

Modellen kan även byggas ut till sk flergruppsmodeller och till modeller med missing-data varvid man kan göra direkta jämförelser mellan t ex olika utbildningar vad gäller den roll som de övriga faktorerna spelar för tillväxten även om ett visst bortfall förekommer. Denna teknik gör det möjligt att studera vilken roll kön, social bakgrund och etnisk bakgrund spelar för de studerandes poängtillväxt och om det eventuellt föreligger skillnader mellan olika utbildningar i detta avseende. Även här är det givetvis av stort intresse att undersöka vilken roll resultaten på urvalsinstrumenten spelar i kombination med bakgrundsvariablerna för poängtillväxten, och om det föreligger några skillnader mellan olika utbildningar. I detta sammanhang är det också av intresse att studera vilka skillnader som finns mellan olika utbildningar och universitet/högskolor i poängtillväxt.

Med tillgång till data som sträcker sig över ett stort antal födelsekohorter kommer de ovan angivna frågeställningarna att studeras i tidsperspektiv. Vi får härigenom också veta huruvida de olika faktorernas

betydelse för poängtillväxten förändrats över en period på ca 15 år. Med hjälp av flernivåteknik kommer vi också att granska i vad mån speciella egenskaper i utbildningarna inverkar på tillväxten i poängproduktion. Exempel på sådana egenskaper är antagningskrav, undervisnings- och examinationsformer.

Delprojekt 8: Studier av högskoleutbildningarna inom teknik- respektive undervisningssektorn

Vid sidan av de mer översiktliga studierna av urvalsinstrumentens prognosförmåga inom flertalet utbildningar i högskolan är det också angeläget att mer fördjupade studier görs av enskilda utbildningar. Inom ett delprojekt planerar vi att undersöka vissa utbildningar inom teknik- och undervisningssektorn. I detta sammanhang kommer vi också att granska rekryteringen till dessa utbildningar och hur rekryteringsläget påverkar genomströmningen.

Vad gäller den tekniska sektorn kommer vi att syna både civilingenjörs- och högskoleingenjörsutbildningarna. Bland de frågor som här skall besvaras märks: Vilka har konsekvenserna blivit av det starkt växande antalet nybörjare? Har rekryteringsmönstret förändrats vad gäller sådana faktorer som tidigare genomgången utbildning, studieförutsättningar, kön, social bakgrund och etnisk härkomst? Hur har det ökande antalet nybörjare påverkat studieframgången? Varierar studieframgången för olika elevkategorier, för olika studieinriktningar samt mellan olika högskolor? Har olika prediktionsvariabler varierande prognosvärde för skilda grupper?

Inom undervisningssektorn kommer vi att koncentrera oss på utbildningarna till grundskollärare. Även här skall vi studera hur olika faktorer påverkar rekryterings- och genomströmningmönstren. Ett speciellt problem inom lärarutbildningarna är svårigheterna att rekrytera studerande som vill satsa på en inriktning mot matematik och naturvetenskap. Vilka konsekvenser har detta för det svenska urvalssystemet och prediktionsvariablernas prognosvärde?

Delprojekt 9: Studier av två högskoleutbildningar inom samhällsvetenskaplig utbildningssektor

Även detta delprojekt fokuserar på enskilda utbildningar, nämligen de två utbildningar som valts ut i delprojekt 6 (ekonomutbildning och socialt arbete). Information om de studenter som är antagna till dessa utbildningar kommer att inhämtas från respektive lokalt LADOK-register men även från institutioner kopplade till respektive utbildning. Informationen kommer även att kompletteras med information avseende etniskt ursprung och social bakgrund.

I den första fasen av denna delstudie kommer de studerande att jämföras utifrån hur de har antagits. Föreligger det någon skillnad avseende kön, utbildningsbakgrund eller social bakgrund mellan de grupper, som antagits genom gamla respektive nya betyg och högskoleprov med och utan tillägg för arbetslivserfarenhet (ALE)? Samtidigt som det finns en målsättning om social rättvisa, finns även krav på att antagningssystemet ska vara effektivt, vilket kan tolkas så att det skall förmå välja ut de sökande som presterar väl i högre utbildning. Andersson, Lexelius & Ramstedt (1998) visade att den grupp som antas till högskolan utifrån högskoleprovet består av en betydligt större andel män än kvinnor, medan det motsatta har gällt för de gamla gymnasiebetygen. Utifrån ett rättviseperspektiv blir det därmed motiverat att jämföra antagningsgrupperna.

Nästa fas blir att granska de prestationer dessa urvalsgrupper uppvisar (se exempelvis Henriksson & Wolming, 1998). Genom att jämföra urvalsgruppernas prestationer erhålls en bild av hur väl respektive urvalsgrupp (eller urvalsinstrument) förmår välja ut studenter som kan uppvisa goda prestationer. I detta sammanhang blir definitionen av studenternas prestationer synnerligen central. Förutom ”positiva” prestationskriterier som exempelvis antal högskolepoäng och betyg på kurser kommer även ”negativa” kriterier som studieavbrott att närmare granskas.

Avslutningsvis kommer dessa två utbildningar att studeras även utifrån ett konventionellt prediktivt perspektiv. Urvalsinstrumentens prediktiva validitet kommer att undersökas med avseende på skillnader mellan studenter med olika utbildningsbakgrund, etnisk härkomst, social bakgrund och könstillhörighet, liksom mellan de två utbildningarna.

IV. Avnämarna

Att antagningssystemets avnämare upplever att systemet är relevant och användbart och att det förmår uppfylla sina syften är av grundläggande betydelse. Frånsett att antagningssystemets syften kan vara flerfaldiga kan det övergripande syftet sägas vara att välja ut studerande som kan förväntas bedriva studier med framgång. Emellanåt höjs dock röster som gör gällande att rådande antagningssystem inte förmår uppfylla de krav som kan och bör ställas på ett sådant, t ex med avseende på olika grupper möjlighet att påbörja högre studier (Kim, 1998). Antagningsinstrumenten har kritiserats för att vara otillräckliga och för att inte tillvarata den mångfald av talanger eller kompetenser som olika sökande besitter och som kan vara av betydelse för deras förmåga att framgångsrikt bedriva studier vid olika utbildningar (Brandell & Kim, 2000). I sammanhanget kan även aspekter på rättvisa anföras (se exempelvis Wolming, 1999b). Frågan är dock om antagningssystemet kan uppfylla flera syften samtidigt samt hur väl det förmår uppfylla sina syften.

Ett sätt att besvara denna fråga är att vända sig till antagningssystemets avnämare eller intressenter vilka utgörs av de personer eller grupper som på olika sätt berörs av systemet, dvs. studerande som antagits till olika utbildningar, sökande som inte antagits, företrädare för universitet och högskola samt därtill knutna myndigheter (Högskoleverket/ Skolverket), politiker, och arbetsliv. Har exempelvis studerande som antagits samma uppfattning om hur antagningssystemet fungerar som studerande som inte antagits? Har politiker samma uppfattning om antagningssystemets funktion som företrädare för arbetslivet?

Delprojekt 10: Olika avnämargrupperns syn på antagningssystemet

Det är lätt att inse att det förmodligen föreligger en skillnad mellan olika avnämarnas uppfattning av hur antagningssystemet fungerar, vilket sannolikt hänger samman med att även deras kontakter med antagningssystemet skiljer sig åt. De olika avnämarkategorierna utgör inte heller en enhetlig grupp och deras uppfattningar kan även skilja sig åt inbördes. Det utesluter dock inte att samtliga intressenter i princip kan vara överens om att det övergripande målet med den högre utbildningens antagningssystem är att anta studenter som kommer att bedriva sina studier med framgång. Detta innebär dock inte att frågan om innebörden av studieframgång är besvarad, vilket är en fråga som kommer att få en särskild belysning i delprojekt 6.

I tidigare studier av det svenska antagningssystemet lyser avnämarperspektiv i huvudsak med sin frånvaro, även om företrädare för arbetslivet tillfrågats om sin syn på vilka egenskaper de efterfrågar hos sökande vid nyrekryteringar (Skolverket/Högskoleverket, 1999). En rad amerikanska studier kan dock sägas tangera ett avnämarperspektiv (se exempelvis Lenning m.fl., 1974; Pascarella & Terenzini, 1991; Willingham, 1985). Delvis kan detta vara en effekt av att det amerikanska utbildningssystemet, till skillnad från det svenska, har en decentraliserad antagning där flertalet amerikanska college i hög grad är beroende av ett gott förhållande till avnämarna inte minst ur ekonomisk synvinkel (Pascarella & Terenzini, 1991). Avnämarnas koppling till utbildningsinstitutionerna vid amerikanska college kommer även till uttryck i att hänsyn tas till andra meriter hos de sökande än vad som omfattas av betyg och urvalsprov (Shepard, 1993).

För en studie som gör anspråk på att validera ett antagningssystem är emellertid avnämarnas uppfattning av central betydelse. Mot bakgrund av att någon avnämarestudie av antagningssystemet inte tidigare genomförts i Sverige kommer en validering av antagningssystemet, ur detta perspektiv, att kunna ge ett

värdefullt kunskapsstillskott genom att ge en samlad bild av olika avnämares uppfattning om antagningssystemet.

Studiens övergripande syfte är att undersöka olika avnämares syn på antagningssystemet med avseende på komponenterna urvalsinstrumenten, kriteriet/studieframgång och regelsystemet med utgångspunkt i Messicks modell. Som exempel på frågeställningar som anknyter till urvalsinstrumenten kan nämnas: Vilka förmågor mäts via gymnasiebetyg, högskoleprovsresultat eller högskoleprovsresultat i kombination med poäng för arbetslivserfarenhet? Hur fungerar respektive instrument? Hur borde respektive instrument fungera (vad borde mätas?), och vilka blir konsekvenserna i olika avseenden om man använder olika urvalsinstrument. På motsvarande sätt kan också frågeställningar formuleras som är relaterade till antagningssystemets komponenter kriteriet och regelsystemet.

I ett första steg kommer information att samlas in på en övergripande nivå. Detta innebär att studier kommer att genomföras i syfte att erhålla en bred beskrivning, exempelvis via enkäter, riktade till ett urval av olika avnämarkategorier. Tanken är här att erhålla en bred bild av avnämnarnas uppfattning och hur de olika avnämarkategoriernas uppfattningar förhåller sig till varandra.

Det andra steget består av en fördjupad undersökning, i form av intervjuer, där intressanta resultat på den övergripande nivån ges en mer genomgripande och fördjupad belysning. Resultaten från den övergripande beskrivningen, som erhållits i det första steget, kommer således att styra såväl intervjuernas inriktning och uppläggning som val av intervjupersoner.

Eftersom de olika avnämarkategorierna i sig är tämligen omfattande måste ett urval göras inom kategorierna. Nedan anges vilka avnämarkategorier som delprojektet omfattar samt hur urvalet kommer att göras inom respektive avnämarkategori.

Sökande till universitet och högskola: Urvalet i sökandegruppen kommer att omfatta de utbildningsinriktningar och orter som omfattas av projektet i sin helhet. De utbildningar som i första hand kommer att studeras är socialt arbete och ekonomiprogrammet

Företrädare för universitet och högskola: Det är angeläget att välja personer som har en god kännedom om hur antagningssystemet fungerar och som även möter de studerande, vilket innebär att de även kan förväntas ha en god kännedom om och överblick av antagningssystemets effekter. Av denna anledning faller valet på studierektorerna vid de i projektet valda utbildningarna.

Företrädare för verk och myndigheter: Här avses tongivande personer inom verk och myndigheter kopplade till den högre utbildningen vilket framför allt utgörs av Högskoleverket men även Skolverket som ansvarar för de utbildningsstadier som föregår högskola.

Företrädare för det politiska livet: Företrädare för det politiska livet kommer att utgöras av de politiker som på riksnivå, vid tidpunkt för studiens genomförande, arbetar med frågor kopplade till den för den högre utbildningen.

Företrädare för arbets- och näringsliv: Även valet av företrädare för arbets- och näringsliv är en grannliga uppgift. I likhet med de principer som ligger till grund för val av företrädare för universitet och högskola torde personer verksamma med rekrytering och urval vara den mest angelägna gruppen. Valet faller därför på personalchefer vid företag och myndigheter som anställer personer med en utbildningsbakgrund som motsvarar de utbildningar som valts inom projektet.

V. Urvalsregler

Vid urval till högre utbildning i Sverige används ett kvotgruppssystem som i huvudsak baseras på tre urvalsgrupper: sökande med betyg, sökande med högskoleprov och sökande med högskoleprov + ALE. Till detta system finns även specifika regler som berör högskoleprovet och betygen, nämligen s k upprepat provtagande och s k konkurrenskomplettering. Dessa regler har effekter på urvalet men relativt lite är känt om dessa reglers verkningar.

Delprojekt 11: Effekter av upprepat provtagning på urval med högskoleprovet

Enligt de regler som gäller för högskoleprovet kan en provdeltagare genomföra provet hur många gånger som helst och, inom en 5-årsperiod, få räkna det bästa resultatet vid urval till universitet/högskola. En närmare analys och granskning av de motiv som kan tänkas ligga till grund för beslutet om införandet av denna regel förefaller angelägen, liksom frågan om vilken effekt som denna regel har på urvalet.

Sett utifrån ett strikt mätteoretiskt perspektiv medför denna regel att individens faktiska förmåga överskattas. Denna slutsats baseras på antagandet att en provdeltagares erhållna poäng varierar slumpmässigt kring ett sant (teoretiskt) värde. Detta kan balanseras mot det faktum att, sett utifrån ”testwiseness”-perspektiv, medför regeln att man erhåller en bättre skattning av provdeltagarens faktiska förmåga för de individer som har bristande testförtrogenhet. En granskning av regelns effekter dels vad avser vilka som utnyttjar möjligheten till upprepat provdeltagande, dels vilka poängförändringar som kan iaktas förefaller således angeläget.

Sett i ett övergripande teoretiskt perspektiv utgör upprepat provdeltagande ett specialfall av övning enligt den begreppsstruktur som Henriksson (1981) etablerade utifrån en teoretisk granskning av studier inom området effekter av övning och instruktion på testprestation. Upprepat provdeltagande har också en koppling till det teoretiska begreppet test-wisness (se t ex Nilsson, 1979) och detta skapar en teoretisk bas för deduktioner angående förväntade effekter av upprepat provdeltagande för standardiserade prov av typ högskoleprovet. Ett flertal studier har genomförts avseende effekter av upprepat provdeltagande på högskoleprovet, såväl på totalpoängsnivå som på grupp- och delprovsnivå (se t ex Henriksson, 1981, 1991, 1993, 1994a; Henriksson et al, 1985; Henriksson & Bränberg, 1994) och den sammanfattade nationella resultatbilden överensstämmer väl med den internationella resultatbilden för test av samma karaktär som högskoleprovet. Såväl generella effekter av övning, som effekter av övning i relation till olika kategorier av provdeltagare och till olika typer av prov, har också visat sig vara förhållandevis konstanta och stabila (se t ex Henriksson, 1981, 1994a; Becker, 1990).

Undersökningen kommer att inriktas på att komplettera de resultat som erhållits i tidigare studier av upprepat provdeltagande av högskoleprovet. Speciellt intresse kommer därvid att ägnas frågan att separera effekten av utbildning och effekten av upprepat provdeltagande per se. Detta är motiverat av att populationen provdeltagare nu till stor del består av gymnasieelever som genomför provet under sin utbildning. Vidare kommer också ansatsen att inriktas på att utforma en analys och beskrivning på delprovsnivå.

Delprojekt 12: Effekter av konkurrenskomplettering på urval med målrelaterade gymnasiebetyg

Enligt de regler som gäller för betygssättning kan en student åstadkomma en förhöjning av sitt betyg via s.k. konkurrenskomplettering. Detta innebär att man tillskapar speciella omständigheter som tillåter såväl ytterligare skolning som en förnyad bedömning. Dessutom har kommunerna lokalt beslutanderätt angående principer och villkor (exempelvis kostnad) för hur konkurrenskompletteringen ska genomföras. Det förefaller också som om villkoren varierar avsevärt mellan olika kommuner. Sammantaget medför detta att populationen elever på en viss nivå inte bedöms utifrån samma villkor. En granskning av denna

regels effekter vad avser såväl vilka som utnyttjar möjligheten till konkurrenskomplettering av sina betyg som vilka betygsförändringar som kan iakttas förefaller därför angelägen.

Undersökningen kommer i en första fas att i ett urval av kommuner inriktas på att beskriva olika betingelser och villkor för konkurrenskomplettering. I en andra fas kommer utfallet av fas ett att ligga till grund för slutsatser angående effekter av konkurrenskomplettering såväl i ett individ- som ett kvotgrupps- och systemperspektiv.

Delprojekt 13: Kvotgruppssystemet

Ett problem som kommer att belysas inom detta delprojekt är effekterna av att urvalet till högre utbildning sker i olika grupper. För närvarande placeras således sökande med gamla respektive nya gymnasiebetyg i olika urvalsgrupper, sökande med högskoleprovsresultat ingår i ytterligare en annan urvalsgrupp som i sin tur är indelad i a) Högskoleprovsresultat och b) Högskoleprovsresultat med tilläggspoäng för arbetslivserfarenhet (ALE). Det övergripande syftet är således att jämföra de grupper som väljs ut med gamla respektive nya gymnasiebetyg, med enbart högskoleprovsresultat samt med högskoleprovsresultat i kombination med ALE.

Det faktum att sökande med gamla och nya gymnasiebetyg placeras i olika urvalsgrupper, innebär i praktiken att sökande med gamla gymnasiebetyg successivt blir allt mer gynnade. Detta har sin grund i att de mest meriterade sökande med gamla gymnasiebetyg successivt försvinner från sökandegruppen genom att de antas till olika utbildningar. I gruppen med nya betyg sker hela tiden en påfyllning av nya personer, medan de enda nytillkommande till gruppen med gamla gymnasiebetyg är sökande med utländska betyg. Fördelningen av antalet studieplatser mellan urvalsgrupperna sker dock enbart i proportion till antalet sökande.

Ett ytterligare problem i detta sammanhang är att kunskapen är tämligen dålig om vad som egentligen mäts med de nya betygen (se delprojekt 4). Det finns därför anledning att närmare analysera innebörden i och konsekvenserna av de nya betygen genom att jämföra utfallet i urvalssammanhang för dem som har betyg enligt det gamla och det nya systemet och även att jämföra utfallet med det nya betygssystemet och högskoleprovsresultat.

En annan viktig komponent i det nuvarande urvalssystemet är, som redan påpekats, begreppet ALE. Motiven för att vissa sökande ska få tillgodoräkna sig poäng för ALE har i första hand varit att det ska bredda rekryteringen till högskolan (Wolming, 2000). Men, och detta förefaller viktigt att uppmärksamma, det fastslås också att man knappast kan förvänta sig att generella arbetslivsmeriter har ett prognosvärde vad avser studieframgång i högskolan (Högskoleverket, 1997). Beaktat att man också konstaterat ett negativt samband mellan ALE och studieframgång (se Henriksson & Wolming, 1998) förefaller en närmare analys och granskning av de motiv som kan tänkas ligga till grund för denna regel angelägen, liksom frågan om vilken effekt som regeln om ALE har på urvalet.

VI. Integrering

I den integrerande delen av projektet kommer resultaten från samtliga delstudier, vilka avser samtliga komponenter i antagningssystemet att ställas i relation till de fyra aspekterna i Messicks modell. I detta arbete kommer särskilt två frågor att fokuseras. Den ena avser målkonflikter i systemet, vilket särskilt berör den konsekventiella validiteten, och den andra avser den enhetliga utformningen av det svenska antagningssystemet, vilket särskilt avser instrumentens begreppsvaliditet och deras användning.

En av de övergripande målsättningarna med 1977 års högskolereform, vilken i huvudsak fortfarande bestämmer antagningssystemets utformning, var att skapa möjligheter att prioritera mellan olika utbildningsområden och att anpassa utbildningsutbudet till arbetsmarknadens behov. En annan målsättning kan relateras till en politisk strävan efter socialt utjämning åtgärder vilket skulle leda till nya

studerandegrupper i den högre utbildningen. Man ville även tona ner betydelsen av studenternas betygsmärker i samband med antagningen till högskolan.

Kim (1998) menade att argumenten för tillträdesreformen år 1977 kan delas in i tre olika kategorier: (1) sociala rättviseargument, (2) rättsäkerhetsargument, och (3) effektivitetsargument. De sociala rättviseargumenterna grundades i en strävan mot att erhålla en mer alls idig sammansättning av studerandegruppen. Rättsäkerhetsargumenterna fokuserade på att regelsystemet både skulle fungera och upplevas som rättvist och enkelt. Effektivitetsargumenterna kunde bland annat relateras till antagningssystemets förmåga och effektivitet att välja ut studenter som skulle uppvisa goda studieprestationer. Dessa målsättningar kan ses som ett försök att skapa en balans mellan politikernas visioner, högskolornas krav och individernas önskemål (Högskoleverket, 1997).

De resultat som framkommit i de 13 delprojekten kommer att nyttjas för undersöka dels i vilken utsträckning de olika målen har nåtts, dels i vilken utsträckning det föreligger konflikter mellan de olika målen.

Delprojektens resultat kommer också att användas för att granska den enhetliga utformningen av det svenska antagningssystemet, där samma instrument används på samma sätt för samtliga utbildningar. Ett genomgående drag i många av delprojekten är att de fokuserar på instrumentens och kriteriernas flerdimensionalitet, liksom på skillnader mellan olika utbildningar och utbildningsområden. Detta innebär att den enhetliga utformningen kan diskuteras med utgångspunkt i en rik mängd av aktuella empiriska resultat. I den mån det finns anledning av utveckla ett mer differentierat antagningssystem finns goda förutsättningar att nyttja forskningsprogrammets resultat i detta arbete.

Som ett exempel kan nämnas att studier i Israel, där ett urvalstest liknande högskoleprovet har använts i många år, visat på betydande olikhet i funktionssätt för provets tre delar: en verbal, en kvantitativ och en engelsk. Den kvantitativa delen har sålunda hög validitet för kvantitativa samhällsvetenskaper och naturvetenskap, tämligen hög validitet för medicin och ingenjörsutbildning medan den har tämligen låg validitet för juridik och verbala samhällsvetenskaper. Den verbala delen har speciellt hög validitet för humaniora och juridik, tämligen hög för samhällsvetenskap, naturvetenskap och ingenjörsutbildning, men låg validitet för medicin. Den engelska delen hade hög validitet enbart för humaniora (Kennet-Cohen et.al. 1995). Preliminära studier i Sverige har även visat att högskoleprovets delprov fungerar olika för olika utbildningar (Svensson, 2000).

Det finns skäl misstänka att liknande resultat kommer att erhållas i de planerade analyserna av högskoleprovet och gymnasiebetygen. I den integrerande delen kommer resultaten att bedömas med avseende på om de stödjer ett fortsatt enhetligt bruk av instrumenten, eller om ett mer differentierat förfarande är aktuellt. Om det senare är fallet kommer också arbetet i denna del att inriktas på att belysa konsekvenserna av detta för utformningen av antagningssystemet.

Forskningsinformation

Projektresultaten kommer att presenteras i svenska och internationella tidskrifter samt i Enhetens och Göteborgsinstitutionens rapportserier. Den integrerande delen, som sammanfattar och analyserar resultaten på en övergripande nivå, kommer att redovisas i en mer omfattande publikation.

Resultaten kommer också att presenteras på sådana sätt att de blir tillgängliga för grupper som inte läser vetenskapliga primärpublikationer. Intresset för antagningsfrågor är stort i massmedia, liksom i den populärvetenskapliga litteraturen, vilket ger goda förutsättningar för att forskningsresultaten kommer att kunna redovisas i sådana former. Flera av de medverkande i forskningsprogrammet medverkar också i diskussioner kring antagningsfrågor på både lokal och nationell nivå, vilket också ger goda förutsättningar för praktisk tillämpning av den vunna kunskapen.

Forskargruppens kompetens och meriter för projektet

Huvudansvarig för programmet är Jan-Eric Gustafsson (se nedan), som inom ramen för sin professur kommer att bedriva projektledningsarbetet.

Göteborgsgruppen

Delprojekten kommer att ledas av flera seniora forskare, vilka kortfattat presenteras nedan:

Berit Askling, professor i pedagogik. Har varit projektledare för utformningen av de alternativa antagningsförsöket vid Hälsouniversitetet, Linköping. Är f.n. ordförande i Högskoleverkets rådgivande organ för frågor som rör antagning till högre utbildning, det s.k. Provrådet. Har visst utrymme för forskning inom ramen för tjänsten.

Jan-Eric Gustafsson, professor i pedagogik. Han har bland annat forskat kring modeller över intellektets struktur, där en hierarkisk modell som integrerar flera tidigare modeller har utvecklats. En stor del av forskningen har varit inriktad mot mättekniska och metodologiska problem. Gustafsson har varit ordförande i UHÄs vetenskapliga råd för antagningsfrågor och är för närvarande ledamot i Högskoleverkets rådgivande organ för frågor som rör antagning till högre utbildning, det s.k. Provrådet. Han är också ledamot i det rådgivande organet Visting Panel on Research vid Educational Testing Service, Princeton, USA. Har visst utrymme för forskning inom ramen för tjänsten.

Reuterberg, Sven-Eric, professor i pedagogik, har sedan 1990 varit huvudansvarig för UGU -projektet och har under de senaste åren till stor del ägnat sig åt att studera rekryteringen till Högskoleprovet samt gruppskillnader i provresultat. Ingår i den nationella expertgrupp som granskar NOG- och DTK-proven samt arbetar deltid med statistisk analys av ELF-provet. Har ej utrymme för forskning inom ramen för tjänsten.

Allan Svensson, professor i pedagogik ägnar sig åt forskning om studieval och studieframgång på olika nivåer inom utbildningsväsendet. Tidigare ledamot av UHÄs vetenskapliga råd för antagningsfrågor, nu ordförande i SCBs programråd för arbetsmarknads- och utbildningsstatistik. Har ej utrymme för forskning inom ramen för tjänsten.

Dessutom skall följande personer engageras i programmet:

Åberg-Bengtsson, Lisbeth, fil dr i pedagogik. Åberg-Bengtsson disputerade 1998 på en avhandling om barns tolkning av grafisk information. Hon har därefter bland annat arbetat med multivariata analyser av uppgifterna i DTK- och NOG-proven.

Andersson, Anette, doktorand i pedagogik. Andersson kommer att disputerat vt 2001 på en avhandling som behandlar grundskolebetygens flerdimensionalitet, och det informationsinnehåll som går att utvinna genom att ta hänsyn till mer än en generell dimension.

Cliffordson, Christina, doktorand i pedagogik. Cliffordson kommer att disputerat vt 2001 på en avhandling som behandlar empatibegreppet och hur det kan mätas med olika metoder, däribland intervju i en antagningssituation.

Berndtsson, Åsa, databasansvarig

Gustafson, Doris, administratör

En nyrekryterad doktorand

Umeågruppen

Umeåverksamheten bedrivs vid Enheten för pedagogiska mätningar (EPM). EPM, som fram till 1995 var en avdelning vid pedagogiska institutionen, blev detta år en självständig enhet under samhällsvetenskapliga fakulteten. Med undantag av en professorstjänst, till vilken föreståndarskapet för EPM är knutet, är verksamheten helt externfinansierad. Detta innebär ett starkt beroende av externa finansiärer för att kunna vidmakthålla och utveckla den kompetens inom mätområdet som EPM för närvarande har.

Umeådelen av projektet kommer att ledas av professor Christina Stage. Dessutom medverkar Ewa Andersson, Mats Hamrén, Widar Henriksson, Lena Konradsson, Birgitta Törnkvist och Simon Wolming. Dessa presenteras kortfattat nedan:

Christina Stage, professor i pedagogik, har sedan 1995 varit vetenskaplig ledare för högskoleprovet och var under åren 1992 till 1996 projektledare. Hennes forskning har huvudsakligen behandlat könsskillnader i provresultat och under senare år även den nya testteorin "Item Response Theory". Hon är ledamot i Högskoleverkets Rådgivande organ för provfrågor. Har ej utrymme för forskning inom ramen för tjänsten

Widar Henriksson, docent och tf professor i pedagogik, har lång erfarenhet av konstruktions- och FoU-arbete i anknytning till högskoleprovet. Han var projektledare för högskoleprovet åren 1981 - 1985. Hans forskning har främst inriktats på testträning och upprepat provdeltagande. Han är ordförande i Umeå universitets antagningsråd. Har ej utrymme för forskning inom ramen för tjänsten.

Ewa Andersson, doktorand i pedagogik, har flerårig erfarenhet av uppgiftskonstruktion i anslutning till högskoleprovet. Hon kommer under innevarande kalenderår att presentera sin avhandling som behandlar begreppet studieframgång ur ett studentperspektiv, vilket även är ett centralt begrepp inom detta projekt. *Simon Wolming*, försvarade sin doktorsavhandling i pedagogik 2000-05-26. Avhandlingen som har titeln "Validering av urval" har (som framgår av titeln) direkt koppling och relevans för detta projekt och handlar om problematiken med att tillämpa valideringsbegreppet i samband med urval till högre utbildning.

Birgitta Törnkvist, FD i statistik, har multivariat statistisk teknik och metodik som specialitet. Inom ramen för samarbetet mellan EPM och statistiska institutionen har hon tidigare medverkat vid analys och bearbetning av TIMSS-data och i övrigt fungerat som konsult i statistiska frågor.

Mats Hamrén, systemchef, har sedan början av 1970-talet arbetat med datorfrågor i anslutning till högskoleprovet och har ansvaret för registret över provresultat. Han har bl.a. utarbetat det program för uppgiftsanalys som används för högskoleprovet.

Lena Konradsson, assistent, arbetar med administrativa uppgifter i anslutning till högskoleprovet.

Inför år tre av den planerade fyraåriga projektiden är avsikten att anställa en doktorand, som kommer att beredas möjlighet att påbörja sitt avhandlingsarbete inom projektets ram och samtidigt inskolas i EPMs forskningstradition.

Referenser

Andersson, A. (1998). The dimensionality of the leaving certificate in Swedish compulsory school. *Scandinavian Journal of Educational Research*, 42(1), 1998, 25-40.

- Andersson, E. & Grysell, T. (1992). *Umeå sommaruniversitet 1992. En utvärdering baserad på kursvärderingar och intervjuer med lärare och kursansvariga* (Arbetsrapporter från pedagogiska institutionen. Nr 97). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Andersson, E. & Grysell, T. (1995). *Studenternas syn på studieframgång. En pilotstudie inför studier av studieframgångskriteriet inom ramen för prognosprojektet vid Högskoleprovet* (Pm nr 67). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- 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). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Andersson, E. (1996). Kriterier för studieframgång. I E. Johansson, (red.) *Högskoleprovet genom elva forskares ögon* (Högskoleverkets rapportserie 1996:22R). Stockholm: Högskole-verket.
- Andersson, E., Lexelius, A. & Ramstedt, K. (1998). *Kvinnor och män på ekonom- och läkarlinjen. En studie av antagna och deras studieprestation på några utbildningsorter* (Pm nr 146). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.) *Advanced structural equation modeling. Issues and techniques*. Mahwah, New Jersey: Erlbaum, pp. 243-277.
- Becker, B. J. (1990). Coaching for the scholastic aptitude test: further synthesis and appraisal. *Review of Educational Research, 41*, 373-417.
- Bond, L. (1989). The effects of special preparation om measures of scholastic ability. In R.L. Linn (Ed), *Educational Measurement*, Third Edition. New York: Macmillan.
- Brandell, L. (1998). *Nittiolets studenter. Bakgrund och studiemönster* (StudS Arbetsrapport nr 2). Stockholm: Högskoleverket.
- Brandell, L., & Kim, L. (2000). Privilegium eller rättighet - en ESO-rapport om antagningen till högskolan (Ds 2000:24). Stockholm: Finansdepartementet
- Broady, D. (1987). Att utforska högskolan? (UHÄFoU skriftserie 1987:2). Varför Eva men inte Nils -Erik? Perspektiv på studerande i högskolan. Stockholm: Universitets- och högskoleämbetet.
- Cronbach, L., J., & Meehl, P., E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.
- Dahllöf, U. (1968). *Examensfrekvens, studentkategori och lärosäte. En omprövning av de fria fakulteternas effektivitet* (Pedagogiska rapporter, Nr 31). Göteborg: Göteborgs universitet, Pedagogiska institutionen.
- Elgqvist-Saltzman, I. (1976). Vägen genom universitetet. En forskningsöversikt och en empirisk analys av några studievägar inom filosofisk fakultet i Umeå 1968-1975 (avhandling för doktorsexamen). Uppsala Studies in Education. Acta Universitatis Upsaliensis. Uppsala: Uppsala universitet.
- Eriksson S. (2000). En översikt av forskningsrapporter om Högskoleprovet. Pm nr 163, Enheten för pedagogiska mätningar, Umeå universitet.
- Gambetta, D. (1987). *Were they pushed or did they jump? Individual decision mechanisms in education*. Cambridge: Cambridge University Press.
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 1984, 8*, 179-203.
- Gustafsson, J.-E. (1988). Hierarchical models of individual differences in cognitive abilities. In R.J. Sternberg (Ed.) *Advances in the psychology of human intelligence* (Vol 4). Hillsdale, N.J.: Erlbaum, 1988.
- Gustafsson, J.-E. (1997). Measurement characteristics of the IEA Reading Literacy scales for 9-10 year-olds at country and individual levels. *Journal of Educational Measurement, 34*(3), 233-251.
- Gustafsson, J.-E. (1998). Social background and teaching factors as determinants of reading achievement at class and individual levels. *Journal of Nordic Educational Research, 18*(4), 176-185
- Gustafsson, J.-E. (i tryck-a). On the hierarchical structure of personality and ability. In J. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates
- Gustafsson, J.-E. (i tryck-b). Measurement from a hierarchical point of view. In Braun, H., Jackson, D. N., & Wiley, D. E. (Eds.), *Under construction: The role of constructs in psychological and educational measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gustafsson, J.-E. (i tryck-c). On the Relation between Fluid and General Intelligence: A Reanalysis of the Holzinger and Swineford (1939) Study. *Intelligence*.
- Gustafsson, J.-E., Andersson, A., & Hansen, M. (2000). Prestationer och prestationsskillnader i 1990-talets skola. Underlagskapitel för Kommittén Valfärdsbokslut .

- Gustafsson, J.-E. & Benjegård, M. (1996). Högskoleprovet, social bakgrund och upprepad provtagning. I: *Högskoleprovet. Genom elva forskares ögon*. (Högskoleverkets rapportserie 1996:22 R) Stockholm: Högskoleverket.
- Gustafsson, J.-E., & Holmberg, L. (1992). Psychometric properties of vocabulary test items as a function of word characteristics. *Scandinavian Journal of Educational Research*, 36(3), 191-210..
- Gustafsson, J.-E., & Reuterberg, S.-E. (2000). Metodproblem vid studier av urvalsinstruments prognosförmåga – och deras lösning! Manus för eventuell publicering i *Forskning och Framsteg*.
- Gustafsson, J.-E., & Stahl, P. A. (1999). *STREAMS User's Guide. Version 2.0 for Windows*. Mölndal, Sweden: MultivariateWare.
- Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. Berliner, & R. Calfee (Eds.), *Handbook of Educational Psychology*, New York: Macmillan, pp. 186-242.
- Gustafsson, J.-E., Wedman, I., & Westerlund, A. (1992). The dimensionality of the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 36, 21-39.
- Gustafsson, J.-E. & Westerlund A. (1994). Socialgruppskillnader i prestationer på Högskoleprovet. I: R. Erikson & J. O. Jonsson (Red.). *Sorteringen i skolan. Studier av snedrekrytering och utbildningens konsekvenser*. Stockholm: Carlsson Bokförlag.
- Hackman, J. D., Taber, T. D. (1979). Patterns of Undergraduate Performance Related to Success of College. *American Educational Research Journal*, 16, 2, 117-138.
- Hansen, M. (1997). Social bakgrund och resultat på högskoleprovet (Rapport 1997:08). Göteborg: Göteborgs universitet, institutionen för pedagogik.
- Henriksson, W. (1981). Effekter av övning och instruktion på testprestation. En litteraturgranskning. *Pedagogiska rapporter, Umeå, Nr 93*. Pedagogiska institutionen, Umeå universitet.
- Henriksson, W. (1991) Effekter av upprepat provtagande. *Pedagogiska Mätningar, Nr 40*. Pedagogiska institutionen, Avdelningen för pedagogiska mätningar, Umeå universitet.
- Henriksson, W. (1993). The problem of repeated test taking and the SweSAT. *Educational Measurement, No 5*. Division of Educational Measurement, Department of Education, University of Umeå.
- Henriksson, W. (1994a). Meta-analysis as a method for integrating results of studies about effects of practice and coaching on test scores. *British Journal of Educational Psychology*, 64, 319-329.
- Henriksson, W. (1994b). Effects of Repeated Test Taking on Swedish Scholastic Aptitude Test. In, Mauritius Examinations Syndicate (Eds.). *School-Based and External Assessments*. Reudit, Mauritius: Mauritius Printing Specialists Ltd.
- Henriksson, W. (1995). Repeated test taking and the SweSAT. (Educational measurement, no. 13). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Henriksson, W., Henrysson, S., Stage, C., & Wedman, I. (1985). Prov för urval till högskolan. *SOU 1985:59*.
- Henriksson, W., & Bränberg, K. (1994). The effects of practice on the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 38(2), 129-148.
- Henriksson, W., & Wedman, I. (1992). Prediction of Academic Success in a Perspective of Criterion-related and Construct Validity (Em Nr 2). Umeå: University of Umeå. Department of Education. Division of Educational Measurement.
- Henriksson, W. & Wolming S. (1998). Academic Performance in Four Study Programmes: a Comparison of Students Admitted on the basis of GPA and SweSAT Scores, with and without credits for Work Experience. *Scandinavian Journal of Educational Research* 4 (2), 135-150.
- Henrysson, S. (1984). Prognos av framgång i högre studier, metoder och problem. En rapport inom projektet "Antagningssystemets effektivitet och prognosförmåga" (Arbetsrapporter från pedagogiska institutionen Nr 12). Umeå: Umeå universitet, Pedagogiska institutionen.
- Härnqvist, K. (1960). *Individuella differenser och skoldifferentiering*. SOU 1960:3. Stockholm: Ecklesiastikdepartementet.
- Härnqvist, K. (1994). Social selektion till gymnasium och högskola. I: R. Erikson & J. O. Jonsson (Red.). *Sorteringen i skolan. Studier av snedrekrytering och utbildningens konsekvenser*. Stockholm: Carlsson Bokförlag.
- Härnqvist, K. (1998). Explaining the recruitment to post-secondary education: the case of Sweden. *Journal of Nordic Educational Research*, 18(4), 207-216.
- Härnqvist, K., & Svensson, A. (1973). A Swedish data bank for studies of educational development. *Sociological Microjournal*, 7, 35-42.
- Härnqvist, K., & Svensson, A. (1980). *Den sociala selektionen till gymnasiet. En jämförelse mellan fyra årskullar*. Betänkade av Gymnasiutredningen. SOU 1980:30.

- Högskoleverket. (1997). *Tillträde till högre utbildning – en evighetsfråga* (Högskoleverkets skriftserie 1997:13 S). Stockholm: Högskoleverket.
- Kennet-Cohen, T., Bronner, S. & Oren, C. (1995). A meta-Analysis of the Predictive Validity of the Selection Process to Universities in Israel. Jerusalem, National Institute for Testing and Evaluation.
- Kim, L. (1998). *Val och urval till högre utbildning. En studie baserad på erfarenheterna av 1977 års tillträdesreform*. Akademisk avhandling, Uppsala: Uppsala Universitet.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading: Addison-Wesley.
- Lenning, O. T., Munday, L. A., Johnson, O. B., Vander Well, A. R., & Brue, E. J. (1974) The many faces of college success and their nonintellectual correlates: the published literature through the decade of the sixties (Monograph No 14). Iowa City: American College Testing Program.
- Messick, S. (1989). Validity. In: R.L. Linn (Ed.), *Educational measurement* (Vol. 3, pp. 13-103). New York: American Council on Education/MacMillan.
- Messick, S. & Jungeblut, A. (1981). Time and method in coaching for the SAT. *Psychological Bulletin*, 89, 191-216.
- Muthén, B. O., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52, 431-462.
- Muthén, B., & Muthén, L. (1998). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Mårdberg, B., & Carlstedt, B. (1998). Swedish Enlistment Battery (SEB): Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment*, 6(2), 107-113.
- Nilsson, I. (1979). *Test-wisness och provkonstruktion* (avhandling för doktorsexamen, Umeå universitet).
- Pascarella, E. P., & Terenzini, P. T. (1991). *How College Affects Students?*. San Francisco: Jossey-Bass.
- Powers, D. E. (1993). Coaching for the SAT: a summary of the summaries and an update. *Educational Measurement: Issues and Practice*, 12, 24-30.
- Reuterberg, S.-E. (1996). *Matematik i grundskolan*. (Rapport 1996:6). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Reuterberg, S.-E. (2000). Correcting validity coefficients for restriction of range and compensatory selection rules with missing-data modeling: A simulation study. Submitted for publication.
- Shepard, L.A. (1993). Evaluating test validity. In: L. Darling-Hammond (Ed.), *Review of research in education* (pp. 405-450). Washington, DC: American Educational Research Association.
- Shepard, L. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16(2), 5-8.
- Similä, M. (1994). Andra generationens invandrare i den svenska skolan. I: R. Erikson & J. O. Jonsson (Red.). *Sorteringen i skolan. Studier av snedrekrytering och utbildningens konsekvenser*. Stockholm: Carlsson Bokförlag.
- Skolverket/Högskoleverket. (1999). *Hur gick det sedan? Avnämarna av gymnasieskolan och högskolans utbildningar* (Skolverkets rapport nr 168/Högskoleverkets rapportserie 1999:7). Stockholm: Liber.
- SOU 1968:25. *Studieprognos och studieframgång*. Stockholm: Utbildningsdepartementet.
- Stage, C. (1985). *Gruppskillnader i provresultat. Uppgiftsinnehållets betydelse för resultat-skillnader mellan män och kvinnor på prov i ordkunskap och allmänorientering*. Akademisk doktorsavhandling. Umeå: Umeå universitet, Pedagogiska institutionen.
- Stage, C. (1990). Gender Differences in Test Results. *Scandinavian Journal of Educational Research*, 32(3), 101-111.
- Stage, C. (1992a). *Betyg och högskoleprov* (Pm nr 53). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Stage, C. (1992b). *Skillnader mellan betyg och högskoleprovresultat* (Pm nr 62). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Stage, C. (1992c). How Important are Age and Education for Gender Differences in Test Results? *Scandinavian Journal of Educational Research*, 36(3), 223-235.
- Stage, C. (1993). *Gender Differences on the SweSAT. A Review of Studies since 1975*: (EM No 7), Umeå: Umeå University, Division of Educational Measurement.
- Stage, C. (1995). Gender Differences on Admission Instruments for Higher Education. In T., D. Kellaghan (Ed.) *Admission to Higher Education: Issues and Practice*. Dublin: Educational Research Centre.
- Stage, C. (1996). Experiences with the Swedish Scholastic Aptitude Test. In Hänsen, K., Hofer, R. & Ruefli, D. *Eignungsdiagnostik und Medizinstudium. Bericht 2. Zentrum für Testentwicklung und Diagnostik am Psychologischen Institut der Universität Fribourg*.
- Stage, C. (1996). Kvinnor och män. I E. Johansson (red.) *Högskoleprovet genom elva forskares ögon*. Högskoleverkets rapportserie 1996:22R, 9-17.

- Stage, C. (1997). *Do Males and Females with Identical Test Scores Solve Test Items in the Same Way?* Paper presented at the 23rd Annual Conference of International Association for Educational Assessment in Durban, South Africa. (EM No 23). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1998). A Comparison between Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest ERC. (EM No 30). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1999). *Predicting Gender Differences in Word Items. A Comparison of Item Response Theory and Classical Test Theory* (EM No 34). Umeå: Umeå University, Department of Educational Measurement.
- Svensson, A. (2000). Kan högskoleprovet förutsäga studieframgång inom civilingenjörsutbildningarna? Manus.
- Svensson, A., & Stahl, P. A. (1996). Framgång och misslyckanden i gymnasieskolan. (Rapport 1996:17). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Taber, T. D., Hackman, J. D. (1976) Dimensions of Undergraduate College Performance. *Journal of Applied Psychology*, 61, 5, 546-558.
- Ullstadius, E., Gustafsson, J.-E., Carlstedt, B. (2000). Separating the influence of general and crystallized intelligence on vocabulary test items. Submitted for publication in *Educational and Psychological Measurement*.
- Willingham, W. W. (1985). *Success in College. The Role of Personal Qualities and Academic Ability*. New York: College Entrance Examination Board.
- Wolming, S. (1998). Validitet. Ett traditionellt begrepp i modern tillämpning. *Pedagogisk forskning i Sverige*. 3(2), 81-103.
- Wolming, S. (1999a). Validity issues in Higher Education Selection: A Swedish Example. *Studies in Educational Evaluation*, 25(4), 335-351.
- Wolming, S. (1999b). Ett rättvist urval? *Pedagogisk forskning i Sverige*, 4(3), 245-258.
- Wolming, S. (2000). Validering av urval. (avhandling för doktorsexamen, Umeå universitet)
- Wästerfors, D. (1998). Studiestarten som ett äventyr. Rationalitet, slump och nödvändighet inför högskolestudierna (Studs Arbetsrapport nt 5). Stockholm: Högskoleverket.
- Åberg-Bengtsson, L. (1999). Dimensions of performance in the interpretation of diagrams, tables and maps: Some gender differences in the Swedish Scholastic Aptitude Test. *Journal of Research in Science Teaching*, 36, 565-582.

Bilaga 1: Vetenskapliga skrifter

Berit Askling

- Askling, B. (1990). *Att utforska högskolans utbildningar*. Rapport. Linköping: Institutionen för pedagogik och psykologi
- Askling, B., Ottosson, J.-E., & Ulfberg, S. (1991). *Utveckling av särskilda urvalsinstrument. Några erfarenheter från försöket med lokal antagning till läkarlinjen och sjukgymnastutbildningen vid Hälsouniversitetet i Linköping höstterminen 1991*. Rapport från Universitetet i Linköping, Institutionen för pedagogik och psykologi. LiU-PEK-R-155.
- Askling, B., & El-Khawas, E. (1997). *The Academic Profession: Evolving Roles in Diverse Contexts*. Paper presented at the Association for the Study of Higher Education Annual Conference. November 5-9, 1997.
- Askling, B., & Bauer, M. (1997). *The Role, Functions, and Impact of a National Agency in the Evaluation of a Decentralised Higher Education System*. Paper presented at the 10th Annual CHER conference 18-20 September 1997
- Askling, B. (1998). *Quality Work in Swedish Universities in a Period of Transformation* (Report 1998:04). Göteborg: Göteborg University, Department of Education and Educational Research.
- Askling, B. (1998). Främja och granska — motto för lokalt kvalitetsarbete. I M. Bauer (Red.), *Kraften ligger i det okända*. En festskrift till Stig Hagström, universitetskansler 1992 — 1998. Högskoleverket. Stockholm.
- Askling, B. (1998). 'Sweden: Professional Diversity in an Egalitarian System.' In D. Farmham (Ed.), *Managing the Academic Profession*. London: Open University Press.
- Askling, B., Bauer, M., & Marton, S. (1998). *Swedish Universities towards Self-regulation — a new look at institutional autonomy*. Paper presented at the 20th EAIR Forum in San Sebastian, Spain, 9-12 September, 1998.
- Bauer, M., Marton, S., Askling, B., & Marton, F. (in press). *Transforming Universities. Patterns of Governance, Structure and Learning in Swedish Higher Education at the Millennial Turn*. Jessica Kingsley Publ. London.

Jan-Eric Gustafsson

- Gustafsson, J.-E. (1992). The relevance of factor analysis for the study of group differences. *Multivariate Behavioral Research*, 27(2), 239-247.
- Gustafsson, J.-E., Wedman, I., & Westerlund, A. (1992). The dimensionality of the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 36, 21-39.
- Gustafsson, J.-E. (1992). The "Spearman hypothesis" is false. *Multivariate Behavioral Research*, 27(2), 265-267.
- Gustafsson, J.-E., & Undheim, J. O. (1992). Stability and change in broad and narrow factors of intelligence from ages 12 to 15 years. *Journal of Educational Psychology*, 84(2), 141-149.
- Gustafsson, J.-E., & Holmberg, L. (1992). Psychometric properties of vocabulary test items as a function of word characteristics. *Scandinavian Journal of Educational Research*, 36(3), 191-210.
- Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28(4), 407-434.
- Gustafsson, J.-E. (1994). Hierarchical models of intelligence and educational achievement. In A. Demetriou, & A. Efklides (Eds.), *Intelligence, mind, and reasoning. Structure and development*, (pp. 45-73). Amsterdam: North-Holland.
- Gustafsson, J.-E. (1995). General intelligence. In R. E. Sternberg (Ed.), *Encyclopedia of Intelligence*. New York: Macmillan.
- Gustafsson, J.-E. (1995). Hierarchical theories. In R. E. Sternberg (Ed.), *Encyclopedia of Intelligence*. New York: Macmillan.
- Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. Berliner, & R. Calfee (Eds.), *Handbook of Educational Psychology*, New York: Macmillan, 1996, 186-242.
- Gustafsson, J.-E., & Snow, R. E. (1997). Ability profiles. In R. F. Dillon (Ed.), *Handbook of Testing*. Westport, Connecticut: Greenwood Press, 107-135.
- Gustafsson, J.-E. (1997). Measurement characteristics of the IEA Reading Literacy scales for 9-10 year-olds at country and individual levels. *Journal of Educational Measurement*, 34(3), 233-251.
- Gustafsson, J.-E., & Stahl, P. A. (1997). *STREAMS User's Guide*. Mölndal, Sweden: MultivariateWare.
- Crawford, J. R., Deary, I. J., Allan, K. M., & Gustafsson, J.-E. (1998). Evaluating competing models of the relationship between inspection time and psychometric intelligence. *Intelligence*, 26(1), 27-42.
- Gustafsson, J.-E. (1998). Social background and teaching factors as determinants of reading achievement at class and individual levels. *Journal of Nordic Educational Research*, 18(4), 176-185.

- Gustafsson, J.-E. (1998). Editorial for a special issue on Modern Quantitative Methods in Educational Research. *Journal of Nordic Educational Research*, 18(4), 129-132.
- Gustafsson, J.-E. (i tryck-a). On the hierarchical structure of personality and ability. In J. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates
- Gustafsson, J.-E. (i tryck-b). Measurement from a hierarchical point of view. In Braun, H., Jackson, D. N., & Wiley, D. E. (Eds.), *Under construction: The role of constructs in psychological and educational measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gustafsson, J.-E. (i tryck-c). On the Relation between Fluid and General Intelligence: A Reanalysis of the Holzinger and Swineford (1939) Study. *Intelligence*.
- Gustafsson, J.-E., Andersson, A., & Hansen, M. (2000). Prestationer och prestationsskillnader i 1990-talets skola. Underlagskapitel för Kommittén Välfärdsbokslut .
- Gustafsson, J.-E., & Reuterberg, S.-E. (2000). Metodproblem vid studier av urvalsinstrumenters prognosförmåga – och deras lösning! Manus för eventuell publicering i Forskning och Framsteg.

Sven-Eric Reuterberg

- Reuterberg, S-E., & Gustafsson, J-E. (1992). Confirmatory Factor Analysis and Reliability: Testing Measurement Model Assumptions. *Educational and Psychological Measurement, Vol 52*(4), 795-811.
- Reuterberg, S-E. (1994). *Selection of Swedish Scholastic Aptitude Test takers as a function of social background and ability* (Rapport 1994:05). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Reuterberg, S-E. (1996). Selektionsmekanismernas betydelse för gruppskillnader på Högskoleprovet. I *Högskoleprovet. Genom elva forskares ögon*. Högskoleverket. Rapport 1996:22R.
- Reuterberg, S-E. (1996). Den sociala bakgrundens betydelse för prestationer på Högskoleprovet. I *Högskoleprovet. Genom elva forskares ögon*. Högskoleverket. Rapport 1996:22R.
- Mäkitalo, Å., & Reuterberg, S-E. (1996). *Who takes the Swedish Scholastic Aptitude Test? A study of differential selection to the SweSAT in relation to gender and ability*(Report 1996:03). Göteborg: Göteborg University, Department of Education and Educational Research.
- Reuterberg, S-E. (1996). *Matematik i grundskolan* (Rapport 1996:06). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Reuterberg, S-E. (1997). *Gender differences on the Swedish Scholastic Aptitude Test* (Report 1997:02). Göteborg: Göteborg University, Department of Education and Educational Research.
- Reuterberg, S-E. (1997). Educational Careers for Gifted Students in Sweden. In J. Li, R. Chan, & J. Spinks (Eds.), *Maximizing Potential: Lengthening and Strengthening Our Stride*. Hong Kong: Social Sciences Research Centre.
- Reuterberg, S-E. (1998). On Differential Selection in the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 42(1), 81-97.
- Reuterberg, S-E. (1998). *Hur var det på högstadiet? Grundskoleelever värderar högstadietiden beträffande upplevda problem, uppnådda kunskaper samt nyttan av kunskaperna* (Rapport 1998:08). Göteborg: Göteborgs universitet, Institutionen för pedagogik.

Allan Svensson

- Reuterberg, S-E., & Svensson, A. (1992). *Social bakgrund, studiestöd och övergång till högre studier*. Statens offentliga utredningar 1992:122.
- Emanuelsson, I., Reuterberg, SE, & Svensson, A. (1993). Changing differences in Intelligence? *Scandinavian Journal of Educational Research*, 37, 259-277.
- Reuterberg, S-E., & Svensson, A. (1994). Financial aid and recruitment to higher education in Sweden: changes between 1970 and 1990. *Studies in Higher Education*, 19, 33-45.
- Svensson, A. (1995). *Att välja eller välja bort naturvetenskap och teknik*. Skolverket & VHS: Not-häfte 1995:3.
- Svensson, A. (1996). Kulturfria instrument - finns sådana hybrider? I *Högskoleprovet. Genom elva forskares ögon*. Högskoleverket. Rapport 1996:22R.
- Svensson, A. (1996). *NT-resan. Så får högskolan fler studenter till naturvetenskap och teknik*. Skolverket och högskoleverket: Not-häfte 1996:6.
- Svensson, A. (1997). Why do not gifted students choose technical and science programmes in upper secondary school? In J. Li, R. Chan, & J. Spinks (Eds.), *Maximizing Potential: Lengthening and Strengthening Our Stride*. Hong Kong: Social Sciences Research Centre.

- Svensson, A., & Reuterberg, S-E. (1998). *How to get more students to science and technical programmes in higher education*. Paper presented at the International Conference on Science, Technology and Society. March 16-22, Tokyo.
- Svensson, A. (1998). *Hur lyckas eleverna i den nya gymnasieskolan?* (Rapport 1998:07). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Svensson, A. (1999). *Socialgruppsbegreppet - sett ur den pedagogiska forskningens synvinkel*. Göteborgs universitet, Institutionen för pedagogik och didaktik. Under tryckning.

Lisbeth Åberg-Bengtsson

- Gerber, R., Ottosson, T., Boulton-Lewis, G., Bruce, C., & Åberg-Bengtsson, L. (1993). *Towards an understanding of how children understand graphic representations of quantitative data*. Paper presenterat vid The 5th EARLI Conference, Aix-en-Provence, Frankrike.
- Åberg-Bengtsson, L. (1994). Elevinteraktion i en åldersblandad mellanstadieklass. *Didaktisk Tidskrift*, 4(4), 4-23.
- Ottosson, T., & Åberg-Bengtsson, L. (1995). *Children's understanding of graphically represented quantitative information*. Paper presenterat vid The 6th EARLI Conference, Nijmegen, Nederländerna.
- Åberg-Bengtsson, L. (1996). *En sammanfattning av James J. Gibsons ekologiska ansats: An "ecological approach to visual perception"*. Rapport nr 1996:02. Göteborg: Göteborgs Universitet, Institutionen för pedagogik.
- Åberg-Bengtsson, L. (1996). *Observationer av undervisning i kvantitativa metoder: En rapport inom projektet Statistisk Undervisning Med Datorkraft (SUMDA)*. Nr 24 i skriftserie utgiven av Enheten för kvalitetsutveckling och kvalitetssäkring. Göteborg: Göteborgs Universitet.
- Åberg-Bengtsson, L. (1996). *Education in small rural Swedish schools: An initial overview of the field*. Paper presenterat vid The European Conference on Educational Research, ECER96, Sevilla, Spanien, 1996.
- Åberg-Bengtsson, L. (1997). *Does being a student from a small rural school make a difference? An analytic model of reading achievement among Swedish nine-year-olds*. Paper presenterat vid The European Conference on Educational Research, ECER97, Frankfurt-am-Main, Tyskland, 1997.
- Åberg-Bengtsson, L. (1997). *"Hur väl lever vi upp till läroplanens målsättning?" Utvärdering vid en gymnasieskola där några grupper arbetat PBL-inriktat*. Rapport nr 1997:07. Göteborg: Göteborgs Universitet, Institutionen för pedagogik.
- Åberg-Bengtsson, L. (1998). *Entering a graphic society: Young children learning graphs and charts*. Göteborg Studies in Educational Sciences, 127. Göteborg: Acta Universitatis Gothoburgensis.
- Åberg-Bengtsson, L. (in press). Dimensions of performance in the interpretation of diagrams, tables, and maps: Some gender differences in the Swedish Scholastic Aptitude Test. *Journal of Research in Science Teaching*.

Ewa Andersson

- Wedman, I., Karp, S., Andersson, E. & Åström, T. (1989). Rückmeldung von Erfolgen Überregionaler Lernerfolgsmessungen an Lehrer, Schüler und Eltern. (Betydelsen av ett indikatorperspektiv för återkoppling av resultat från nationell utvärdering till lärare, elever och föräldrar) I K. Ingenkamp. & W. H. Schreiber (Hrsg). *Was wissen unsere Schüler? Überregionale Lehrerfolgsmessung aus internationaler Sicht*. Weinheim: Deutscher Studien Verlag.
- Andersson, E. & Gysell, T. (1992). *Umeå sommaruniversitet 1992. En utvärdering baserad på kursvärderingar och intervjuer med lärare och kursansvariga* (Arbetsrapporter från pedagogiska institutionen. Nr 97). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Andersson, E. & Gysell, T. (1992). *Studenternas syn på studieframgång. En pilotstudie inför studier av studieframgångskriteriet inom ramen för prognosprojektet vid Högskoleprovet* (Pm nr 67). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Andersson, E. & Gysell, 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). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.
- Andersson, E. (1996). Kriterier för studieframgång. I E. Johansson, (red.) *Högskoleprovet genom elva forskares ögon* (Högskoleverkets rapportserie 1996:22R). Stockholm: Högskole-verket.
- Andersson, E., Lexelius, A. & Ramstedt, K. (1998). *Kvinnor och män på ekonom- och läkarlinjen. En studie av antagna och deras studieprestation på några utbildningsorter* (Pm nr 146). Umeå: Umeå universitet, Enheten för pedagogiska mätningar.

Widar Henriksson:

- Henriksson, W. (1981). *Effekter av övning och instruktion på testprestation. Några empiriska studier och analyser avseende övningens och instruktionens betydelse för testprestationen*. Akademisk doktorsavhandling. Umeå: Umeå universitet, pedagogiska institutionen.
- Henriksson, W., Henrysson, S., Stage, C., & Wedman, I. (1985). Provför urval till högskolan. *SOU 1985:59*.
- Henriksson, W., Bränberg, K., Wedman, I., & Nyquist, H. (1990). The influence of sex, education and age on test scores on the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 34(3), 189-203.
- Henriksson, W. (1991). *Effekter av upprepat provtagande* (Pm, Nr 40). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Henriksson, W., & Wedman, I. (1992). *Prediction of academic success in a perspective of criterion-related and construct validity* (EM, No 2). Umeå: Umeå University, Division of Educational Measurement.
- Henriksson, W. (1994). Meta-analysis as a method for integrating results of studies about effects of practice and coaching on test scores. *British Journal of Educational Psychology*, 64, 319-329.
- Henriksson, W. (1994). Effects of Repeated Test Taking on the Swedish Scholastic Aptitude Test. In: Mauritius Examinations Syndicate (Eds.) *School-Based and External Assessments*. Reduit, Mauritius: Mauritius Printing Specialists Ltd.
- Henriksson, W., & Bränberg, K. (1994). The effects of practice on the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 38(2), 129-148.
- Wedman, I., & Henriksson, W. (1996). The Swedish Scholastic Aptitude Test. Research and main findings. In: K-D. Hänsgen & N. Ischi. (Eds.). *Eignungsprüfung für das Medizinstudium. Kriterien und Testverfahren. Kriterien und Testverfahren*. Bern: Zentrum für Testentwicklung und Diagnostik.
- Adolfsson, L., & Henriksson, W. (1999). Different Methods Different Results. How the Swedish Results in Mathematics vary Depending on Methodological Approach. *Educational Research and Evaluation. An International Journal on Theory and Practice*, 5(2), 127-138.

Christina Stage

- Stage, C. (1985). *Gruppskillnader i provresultat. Uppgiftsinnehållets betydelse för resultat-skillnader mellan män och kvinnor på prov i ordkunskap och allmänorientering*. Akademisk doktorsavhandling. Umeå: Umeå universitet, Pedagogiska institutionen.
- Stage, C. (1990). Gender Differences in Test Results. *Scandinavian Journal of Educational Research*, 32(3), 101-111.
- Stage, C. (1992a). *Betyg och högskoleprov* (Pm nr 53). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Stage, C. (1992b). *Skillnader mellan betyg och högskoleprovresultat* (Pm nr 62). Umeå: Umeå universitet, Avdelningen för pedagogiska mätningar.
- Stage, C. (1992c). How Important are Age and Education for Gender Differences in Test Results? *Scandinavian Journal of Educational Research*, 36(3), 223-235.
- Stage, C. (1993). *Gender Differences on the SweSAT. A Review of Studies since 1975*: (EM No 7), Umeå: Umeå University, Division of Educational Measurement.
- Stage, C. (1995). Gender Differences on Admission Instruments for Higher Education. In T., D. Ke llaghan (Ed.) *Admission to Higher Education: Issues and Practice*. Dublin: Educational Research Centre.
- Stage, C. (1996). Experiences with the Swedish Scholastic Aptitude Test. In Hänsgen, K., Hofer, R. & Ruefli, D. *Eignungsdiagnostik und Medizinstudium. Bericht 2*. Zentrum für Testentwicklung und Diagnostik am Psychologischen Institut der Universität Fribourg.
- Stage, C. (1996). Kvinnor och män. I E. Johansson (red.) *Högskoleprovet genom elva forskares ögon*. Högskoleverkets rapportserie 1996:22R, 9-17.
- Stage, C. (1997). *Do Males and Females with Identical Test Scores Solve Test Items in the Same Way?* Paper presented at the 23rd Annual Conference of International Association for Educational Assessment in Durban, South Africa. (EM No 23). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1998). A Comparison between Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest ERC. (EM No 30). Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. (1999). *Predicting Gender Differences in Word Items. A Comparison of Item Response Theory and Classical Test Theory* (EM No 34). Umeå: Umeå University, Department of Educational Measurement.

Simon Wolming:

- Henriksson, W. & Wolming S. (1998). Academic Performance in Four Study Programmes: a Comparison of Students Admitted on the basis of GPA and SweSAT Scores, with and without credits for Work Experience. *Scandinavian Journal of Educational Research* 4 (2), 135-150.
- Wolming, S. (1998). Validitet. Ett traditionellt begrepp i modern tillämpning. *Pedagogisk forskning i Sverige*. 3(2), 81-103.
- Wolming, S. (1999a). Validity issues in Higher Education Selection: A Swedish Example. *Studies in Educational Evaluation*, 25(4), 335-351.
- Wolming, S. (1999b). Ett rättvist urval? *Pedagogisk forskning i Sverige*, 4(3), 245-258.
- Wolming, S. (2000). Validering av urval. (Avhandling för doktorsexamen, Umeå universitet)