

# Att värdera urvalsinstrument

## Några reflektioner över begränsningar och möjligheter

SIMON WOLMING

Enheten för pedagogiska mätningar, Umeå universitet

**Sammanfattning:** *Urvalsinstrument har traditionellt blivit värderade utifrån sin förmåga att förutsäga studieframgång, det som med andra ord kallas för prognostisk validitet. Denna prognosförmåga uttrycks med ett mått som anger styrkan på sambandet mellan resultatet på urvalsinstrumentet och den senare studieprestationen. I föreliggande artikel problematiseras användandet av prestationsvariabeln utifrån den traditionella korrelationsbaserade utvärderingsansatsen. I artikeln presenteras även ett antal alternativa ansatser för utvärderingen av ett urvalsinstrument.*

De instrument som används vid urval till högre studier har nästan uteslutande och tämligen oproblematiskt blivit värderade utifrån hur väl de kan förutsäga de sökande studenternas senare studieprestationer. Att belägga denna egenskap hos ett urvalsinstrument innebär att sambandet mellan instrument och senare studieprestationer på olika sätt studeras. De metodologiska och teoretiska problemen vid den här typen av studier är dock många.

Ett av dessa problem är kopplat till det prestationskriterium som används vid studier av urvalsinstrumentets prognosförmåga. Att det finns en spridning i prestationsvariabeln är ett grundläggande metodologiskt krav för att prognosförmågan för ett urvalsinstrument ska kunna beläggas. De prestationer som studenterna uppvisar påverkas dock av en mängd olika faktorer, vilket i sin tur kan leda till att spridningen begränsas. En begränsad spridning i någon av de två variablerna (urvalsinstrumentet eller studieprestationen) innebär att prognosförmågan underskattas.

Förutom de metodologiska problem som är förknippade med denna ansats kan även den ensidiga fokuseringen på prognosvärde diskuteras. I många studier rapporteras att prognosvärdet, uttryckt i en korrelationskoefficient, sällan överstiger en viss nivå. Enligt Bloom och Peters (1961) kan man som bäst räkna med en korrelation mellan urvalsinstrument och de senare studieprestationerna på ungefär 0,50. En liknande sammanfattning av sambandets nivå har även framförts av Willingham, Lewis, Morgan och Ramist (1990). Ett prognosvärde med denna styrka innebär i statistiska termer att 25 procent av den variation som förekommer i prestationsvariabeln förklaras av resultatet på urvalsinstrumentet, vilket innebär att de resterande 75 procenten förklaras av andra orsaker.

I ljuset av detta faktum framstår det som något märkligt att urvalsinstrument till stor del endast värderas utifrån dessa ramar. Wedman (2000) framhåller att andra aspekter än den traditionella urvalsmodellen, där

korrelationskoefficienten mellan instrument och senare prestation står i centrum, förmodligen kommer att erhålla en mer framträdande roll i framtiden. Avsikten med denna artikel är dels att diskutera de problem som är kopplade till prestationsvariabeln utifrån den traditionella korrelationsbaserade ansatsen, dels att visa på alternativa och kompletterande ansatser vid utvärdering av urvalsinstrument.

## DEN GODA PRESTATIONENS PARADOX

Den högre utbildningens antagnings- och urvalssystem har en mängd olika krav ställda på sig. Ett av dessa krav är enligt Högskoleverket (1997a) att antagningssystemet på olika sätt ska vara rättvist. De sökande ska bland annat behandlas likvärdigt och rättvist oberoende av kön och socialgrupp. Ett annat krav som ställs på antagningssystemet är att man med hjälp av systemet ska kunna välja ut de sökande, som med hänsyn till deras meriter, har den bästa studieprognosen. Trost (1995) hävdade att dessa krav är motiverade både ur ett ekonomiskt och ett rättviseperspektiv, så länge en begränsning av antalet utbildningsplatser vid universitets- och högskoleutbildningar föreligger.

Den prognostiska förmågan hos ett urvalsinstrument utvärderas genom att studera sambandet mellan urvalsinstrumentets resultat och de antagna studenternas studieprestationer. Detta samband uttrycks ofta med en korrelationskoefficient och en enkel tolkningsregel är: ju högre samband desto bättre prognosförmåga för instrumentet. Ett av de största problemen när det gäller att undersöka detta samband har varit, och är fortfarande, att finna användbara och meningsfulla mått på studieprestationer. För att dessa mått ska vara meningsfulla att använda vid studier av ett urvalsinstrumentets prognosvärde, bör de uppfylla en mängd olika krav. Messick (1989 s 72) preciserade dessa krav genom att ange att måtten som används bör vara:

relevant to the applied purpose, reliable, free from bias or distortion, and practical, in the sense of being plausible and acceptable to decision makers as well as available or attainable.

De mått på studieprestation som vanligtvis används vid studier av ett instruments prognosvärde är olika former av betygsomdömen eller antal avlagda högskolepoäng. Den här typen av mått kan mycket väl vara relevanta och funktionella utifrån de perspektiv och målsättningar som utbildningar och studenter har. Däremot är de inte speciellt användbara när de ska användas för att värdera styrkan på sambandet mellan urvalsinstrumentets resultat och den senare studieprestationen. Ett grundläggande statistiskt problem är att ett högt samband mellan urvalsinstrumentets resultat och de antagna studenternas studieprestationer ytterst är beroende av en variation i de två variablerna. En hög variation i materialet behöver dock inte innebära att sambandet blir högt, men däremot innebär en obefintlig variation alltid att sambandet blir noll.

En konkretisering av detta problem är följande: Antag att samtliga studenter som antagits till en utbildning slutför sin utbildning med maximala resultat och inom ramen för utsatt tid. Urvalet har därmed till fullo lyckats uppfylla en

av målsättningarna, nämligen att välja ut de studenter som hade den bästa studieprognosen. Att samtliga studenter presterar maximalt är inget problem, snarare en positiv bedrift, sett ur utbildningens eller studenternas perspektiv. Den obefintliga variationen utgör dock ett problem när prestationen används för att utvärdera urvalsinstrumentets prognosförmåga. Den obefintliga variationen i prestationsvariabeln innebär att korrelationen blir noll. Utfallet skulle därmed kunna tolkas så att det aktuella urvalsinstrumentet har en begränsad (eller obefintlig) förmåga att förutsäga studieframgång.

Om det däremot finns en stor variation i studenternas prestationer, det vill säga från mycket låga prestationer till mycket höga prestationer, blir korrelationen högre. Problemet är dock att det bland de antagna studenterna i så fall skulle finnas studenter som inte klarat sina studier, vilket står i motsats till urvalets målsättningar, det vill säga att välja ut de studenter som kommer att prestera väl i sina studier. Att värdera ett urvalsinstruments prognosvärde enbart utifrån ett korrelationsmått kan därmed vara behäftat med grova feltolkningar.

Sett i detta perspektiv är det relevant att nämna att Högskoleverket (2000 s 6) anger följande: »En ökad vikt fästs numera vid genomströmning och examination. Till följd av detta får prognoskraften hos provet ökad betydelse.» Om målsättningen med urvalet är att anta studenter som uppfyller kraven på genomströmning och examination, och om prognosvärdet även likställs med en hög korrelation, finns risk för att felaktiga tolkningar kommer att göras. En hög korrelation behöver, som redan beskrivits, inte innebära en hög genomströmning av studenter med en hög examinationsfrekvens.

## STUDIEPRESTATIONENS MÅTT OCH PROBLEM

En students prestationer i samband med högre utbildning kan, som redan nämnts, uttryckas på en mängd olika sätt. I samband med studier av ett urvalsinstruments prognosvärde hävdade exempelvis Lavin (1965) och Linn (1990) att den typiska studien använder sig av olika former av betygsomdömen som prestationskriterium. Vissa av dessa studier har inriktat sig mot olika former av medelbetyg som mått på studenternas prestationer (Bridgeman 1991, Paskus 1997) medan andra mer fokuserat på betyget i enskilda ämnen (Royer, Marchant, Sinatra & Lovejoy 1990).

Ett annat prestationsmått som tillämpas inom ramen för svenska högskolestudier är antal högskolepoäng (t ex Henriksson & Wolming 1998, Wolming 1999a). En högskolepoäng ska motsvara resultatet av en veckas heltidsstudier, vilket innebär att ett läsår normalt omfattar 40 poäng. Att använda dessa två prestationsmått för att värdera ett urvalsinstruments prognostiska värde är behäftat med ett antal problem, vilka kan påverka måttens betydelse samt begränsa dess variation, vilket i sin tur påverkar utfallet och tolkningen av ett korrelationsmått.

En faktor som kan ha betydelse i detta sammanhang är den högre utbildningens betygssättning. Denna baseras inte på några enhetliga examinationsformer eftersom dessa kan markant skilja sig åt beroende på utbildning och utbildningsort. Dagens studenter möter ofta flera olika former av examina-

tionsformer och den vanligaste formen är salskrivning (Högskoleverket 1997b). Andra vanliga examinationsformer är; individuella PM, grupparbeten, seminarier och eventuellt någon form av självständigt arbete i slutet av utbildningen (Wiiand 1998).

Vilka betygsomdömen som ges kan variera stort beroende på utbildning. Andersson och Grysell (1995) fann, vid en beskrivning av fem olika utbildningar, att antalet betygssteg vid dessa utbildningar varierade från två till fyra betygssteg. Tre av de fem utbildningarna använde sig enbart av två betygsomdömen. Generellt kan man säga att en fingradig betygsskala visserligen är mer användbar, då syftet är att använda omdömet i samband med beräkandet av en korrelation, men det kan diskuteras om den mer graderade skalan har någon större relevans sett ur utbildningarnas perspektiv. I Högskoleverket definieras examinationen som »en bedömning för godkännande eller underkännande av studenters kunskaper och färdigheter i samt förståelse av ett ämne eller område efter en inläring alternativt inlärningsinsats» (Högskoleverket 1997 s 27). Mot bakgrund av denna definition kan hävdas att det som egentligen har betydelse vid examinationen är huruvida studenterna uppnått de olika kunskaps- och färdighetsmål som är kopplade till det område examinationen omfattar, inte att placera in prestationen på en mer graderad skala.

En annan faktor som påverkar antalet högskolepoäng är förekomsten av studieavbrott. Enligt Högskoleverket (1999) har ungefär 25 procent av kvinnorna och 17 procent av männen gjort ett långt uppehåll i studierna (minst tre terminer sammanhängande). Studieavbrott innebär att antalet högskolepoäng som redovisas minskar. Orsakerna till dessa studieuppehåll kan dock vara flera. Enligt Hindbeck, Hagenfeldt och Åberg (1994) kunde 41 procent av studieuppehållen för en grupp läkarstuderande förklaras av bristande studieprestationer (resttentamina) medan övriga studieuppehåll kunde förklaras av sjukdom, studieortsbyte, forskning, arbete, barnafödande eller byte av studieinriktning. Ytterligare en förklaring till studieuppehåll, kan vara att dagens studenter i större utsträckning än tidigare förlägger delar av sin utbildning utomlands.

Studieuppehållen påverkar visserligen prestationsmättet negativt men samtliga studieuppehåll bör, som redovisats, inte nödvändigtvis kopplas samman med dåliga studieprestationer. Problemet uppstår först när dessa prestationsmätt används för att utvärdera urvalsinstrumenten. Prestationsmättet blir då skevt och missvisande.

Ytterligare en faktor som påverkar antalet högskolepoäng är de prestationskrav som ställs av Centrala Studiestödsnämnden. I allmänhet krävs att en student ska klara 30 högskolepoäng under ett år för att utan särskild prövning erhålla nya studiemedel (CSN 2001). Konsekvensen av denna prestationsgräns kan därmed vara att studenter som inte uppfyller Centrala Studiestödsnämndens prestationskrav faller bort och avbryter sina studier. Prestationskravet resulterar därmed i att endast studenter med goda resultat, dvs 30 högskolepoäng eller mer, fortsätter sina studier, vilket i sin tur innebär att variationen i studenternas prestationer minskar avsevärt.

En annan faktor som problematiserar tolkningen av ett korrelationsmått är studieprestationernas jämförbarhet. Olika typer av utbildningar skiljer sig åt,

både avseende innehåll och svårighetsgrad, vilket innebär att jämförelser av studieprestationer från olika utbildningar kan vara svåra att göra. En generell frågeställning i detta sammanhang är huruvida 40 högskolepoäng från utbildning A kan likställas med samma antal högskolepoäng från utbildning B?

De prestationsmått som nämnts kan således vara behäftade med ett antal problem då de används i ett utvärderingssammanhang. Bristande variation i studenternas prestationer, samt olika faktorer som stör innebörden av prestationsmått, innebär betydande svårigheter att på ett meningsfullt sätt tolka sambandet mellan urvalsinstrument och studieprestationer. Korrelationsbaserade utvärderingsmetoder, där urvalets kvalitet bedöms utifrån sambandets styrka, är därmed oerhört svårtolkade och underskattar vanligtvis urvalsinstrumentens förmåga att förutsäga den senare studieprestationen.<sup>1</sup>

Ett sätt att förbättra möjligheten till en mer relevant tolkning av sambandets styrka är att tillämpa olika statistiska korrigeringsmetoder. Gustafsson och Reuterberg (2000) visade på möjligheterna att på statistisk väg göra en uppskattning av hur mycket korrelationen sänks av den beskurna variationen och därmed göra en mer realistisk skattning av det »sanna» värdet. Metoden bygger på att de flesta sökande har resultat på gymnasiebetyg och högskoleprov. Dessa resultat kan användas för att skatta hur väl den grupp som inte antas skulle ha presterat på utbildningen om de blivit antagna. Enligt Svensson, Gustafsson och Reuterberg (2001) kan även fel som uppkommer genom beskuren variation i urvalsinstrumenten elimineras genom att tillämpa denna metod.

Även om den här typen av komplexa analysmetoder används för att tydliggöra sambandet mellan urvalsinstrument och studieprestation, kommer det dock alltid att finnas en övre gräns för prognosvärdet. Som redan nämnts finns en mängd andra mekanismer och faktorer som kan påverka och förklara studieprestationens utfall, vilket ytterligare styrker argumentet att en värdering av ett urvalsinstrument inte enbart ska inriktas mot instrumentets prognosvärde. Att urvalsinstrument används för att prognostisera de senare prestationerna är endast en av många funktioner ett urvalsinstrument har (t ex Wolming 1999b). Det finns därmed grund för Wedmans (2000) antagande att även andra aspekter kommer att erhålla en mer framträdande roll i framtiden än den traditionella urvalsmodellen där korrelationskoefficienten mellan instrument och senare prestation står i centrum.

## ALTERNATIV TILL EN PROGNOsinRIKTAD ANSATs

Vilka alternativa ansatser är då möjliga? Förutom frågan om ett urvals-instruments prognosförmåga pekade Willingham, Lewis, Morgan och Ramist (1990) och Shepard (1993) exempelvis på nödvändigheten av att inrikta studier mot urvalsinstrumentens innehåll, samt att undersöka vilken relevans ett visst innehåll har för den senare utbildningen. Vilken typ av kunskaper, färdigheter och förmågor mäter egentligen instrumentet och hur stämmer dessa överens med den senare utbildningens innehåll?

Shepard (1993) och Willingham m fl (1990) pekade även på vikten av att studera huruvida urvalet kan ses som rättvist för olika grupper av sökande. Ett

uttalat mål är att de urvalsinstrument som används inte ska diskriminera olika grupper. Problemet är att kunna bestämma om eventuella skillnader mellan olika grupper av studenter beror på att det är något fel på instrumentet, eller om det faktiskt föreligger skillnader i deras förmåga. Ett sätt att studera detta är att jämföra dessa skillnader med gruppernas senare studieprestationer. Om de senare studieprestationerna uppvisar samma skillnader kan det tolkas som om gruppernas förmåga faktiskt skiljer sig åt. Däremot om studieprestationerna inte skiljer sig nämnvärt finns det anledning att misstänka att urvalsinstrumentet missgynnar någon grupp. Frågan om huruvida urvalsinstrumenten missgynnar någon grupp eller ej kan därmed enbart avgöras genom att studera eventuella skillnader i gruppernas resultat på urvalsinstrumentet. Ett sätt att belysa denna fråga mer konkret skulle kunna vara att studera antagningsinstrumentens effekter för olika grupper, exempelvis olika studenter med olika urvalsgrund, socialgrupp, etnicitet och kön.

Även urvalsinstrumentens förmåga att välja ut mest lämpade sökande kan studeras med en liknande ansats. Detta kan exemplifieras av det försök att utvärdera olika urvalsinstrumenters funktion som genomfördes av Henriksson och Wolming (1998). Tre olika urvalsgrunder, gymnasiebetyg, högskoleprov samt högskoleprov med tillägg för arbetslivserfarenhet, jämfördes med avseende på de antagna studenternas prestationer. Ett sådant förfaringsätt innebär att jämförelsen av urvalsinstrumenten baseras utifrån gruppernas prestationsskillnader och inte utifrån olika sambands styrka. Även Hindbeck, Hagenfeldt och Åberg (1994) använde en liknande ansats och jämförde prestationer för studenter som antagits på skilda sätt.

Man kan i det här sammanhanget även uppmärksamma de fördelar en longitudinell ansats har. Vid en korrelationsbaserad ansats väljer man vanligtvis att använda studenternas prestationer efter det första årets studier som kriterievariabel. En begränsning med detta val av prestationsvariabel är att det första årets studier vanligtvis skiljer sig i karaktär från resterande studietid. Av denna anledning blir det därför angeläget att vidga perspektivet och även använda sig av prestationsmått som är representativa för den övriga utbildningen.

En annan mycket viktig infallsvinkel är olika avnämares uppfattning huruvida urvalsinstrumenten uppfyller de uppsatta målsättningarna. Två centrala avnämare i detta sammanhang är de sökande studenterna samt de utbildningar som i slutänden ska utbilda studenterna. Uppfattningen bland de sökande och antagna har beskrivits i några studier. Exempelvis har Kolmodin (1993), Lexelius och Wedman (1985) liksom Wester-Wedman (1989) studerat hur högskoleprovets deltagare uppfattar högskoleprovet med avseende på dess innehåll och relevans. Andersson och Grysell (1992) och Wester (1995) belyste de antagna studenternas syn på innebörden av deras studieprestationer. Även Kim (1998) presenterade resultat som gav indikationer på hur de sökande upplevt antagningsprocessen. Även om studenternas uppfattningar därmed kan sägas vara väl kartlagda sker kontinuerliga förändringar i utbildningssystem och samhälle. Studier av studenters attityder och uppfattningar bör därför upprepas.

Framför allt saknas studier där representanter för utbildningarna fått möjlighet att få ge sin uppfattning om antagningssystemet. Olika utbildningar kan exempelvis ha skilda uppfattningar om antagningssystemet. Hur ser de på kraven om rättvisa och prognosförmåga hos olika urvalsinstrument? Är de tillfreds med det rådande systemet, exempelvis antagningsregler, behörighetskrav, urvalsinstrument?

## SAMMANFATTNING

Att studera värdet av ett antagningsinstrument enbart ur en korrelationsbaserad ansats kan vara missvisande. Som beskrivits, finns mekanismer som kraftigt kan påverka variationen i studenternas prestationer. Enbart ett mått på sambandet mellan dessa två variabler utgör inte en adekvat indikator på ett urvalsinstrumentets värde. Om variationen är obefintlig (på grund av studenternas maximala prestationer) innebär detta att korrelationen blir noll. Finns det däremot en variation i studenternas prestationer ökar förutsättningarna för att nivån på korrelationskoefficienten stiger. En oreflekterad tolkning av ett nollsamband är att urvalsinstrumentet inte besitter någon som helst prognostisk förmåga. I det senare fallet, när det finns ett samband, tolkas detta däremot som om instrumentet besitter en viss prognostisk förmåga. Samtidigt ska det nämnas att korrelationen blir noll även när studenternas prestationer inte uppvisar någon variation på grund av minimala prestationer. Att instrumentet inte har någon prognostisk förmåga är då självklart.

De olika mekanismer som påverkar variationen i studieprestationerna, och de olika tolkningar som kan göras av en korrelationskoefficient, illustrerar nödvändigheten av alternativa och kompletterande ansatser som belyser det faktiska utfallet för olika grupper av studenter vid analysen av ett urvalsinstrumentets värde. Den prognostiska styrkan i form av en korrelationskoefficient behöver alltid kompletteras med information om studenternas faktiska prestationer för att kunna tolkas på ett mer adekvat sätt.

## NOTER

Denna artikel är skriven inom ramarna för VALUTA-projektet (Validering av den högre utbildningens antagningssystem), finansierat av Riksbankens Jubileumsfond.

1. För en vidare diskussion om tolkningen av korrelationer i detta sammanhang se exempelvis Henriksson och Wedman (1992) och Henrysson (1984).

## LITTERATUR

- Andersson, E. & Grysell, T. 1992: *Studenters syn på studieframgång. En pilotstudie inför studier av studieframgångskriteriet inom ramen för prognosprojektet vid Högskoleprovet* (PM Nr 67). Umeå: Umeå universitet, Pedagogiska institutionen, 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 17). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för

- pedagogiska mätningar.
- Bloom, B.S. & Peters, F.R. 1961: *The use of academic prediction scales for counseling and selecting college entrants*. New York: The Free Press of Glencoe.
- Bridgeman, B. 1991: Essays and multiple-choice tests as predictors of college freshman GPA. *Research in Higher Education*, 32(3), 319–331.
- CSN, 2001: *Nya regler för studiemedel*. Tillgänglig: www.csn.se [2001-02-01].
- Gustafsson, J-E. & Reuterberg, S-E. 2000: Metodproblem vid studier av Högskoleprovets prognosförmåga – och deras lösning! *Pedagogisk Forskning i Sverige*, 5(4), 273–283.
- Henriksson, W. & Wedman, I. 1992: *Prediction of academic success in a perspective of criterion-related and construct validity* (EM No 2). Umeå: Umeå University, Department of educational measurement.
- Henriksson, W. & Wolming, S. 1998: Academic performance in four study programmes: a comparison of students admitted on the basis of the SweSAT-scores, with and without credits for work experience. *Scandinavian Journal of Educational Research*, 42(2), 131–150.
- Henrysson, S. 1984: *Prognos av framgång i högre studier. Metoder och problem*. (Arbetsrapport Nr 12). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Hindbeck, H., Hagenfeldt, K. & Åberg, H. 1994: *Lokal antagning till läkarutbildning vid Karolinska institutet. Genomströmning och studieresultat för studerande antagna vårterminen 1992* (Rapport 59). Stockholm: Karolinska institutet, Institutionen för medicinsk informatik och pedagogik.
- Hynes, K. & Givner, N. 1981: Restriction of range effects on the new MCAT's predictive validity. *Journal of Medical Education*, 56, 352–353.
- Högskoleverket, 1997a: *Tillträde till högre utbildning: En evighetsfråga* (HSV 1997:13 S). Stockholm: Högskoleverket.
- Högskoleverket, 1997b: *Studenters upplevelser av examinationen: Om hur högskolestuderande retrospektivt ser på examinationen vid högskolan*. (HSV 1997:5 S). Stockholm: Högskoleverket.
- Högskoleverket, 1997c: *Examinationen i högskolan. Slutrapport från Högskoleverkets examinationsprojekt*. (HSV 1997:39 R) Stockholm: Högskoleverket.
- Högskoleverket, 1999: *Årsrapport för universitet och högskolor 1999*. Stockholm: Högskoleverket.
- Högskoleverket, 2000: *Högskoleprovet. Gårdagens mål och framtida inriktning*. (HSV 2000:12 R) Stockholm: Högskoleverket
- Kim, L. 1998: *Val och urval till högre utbildning: En studie baserad på erfarenheterna av 1997 års tillträdesreform*. Uppsala: Acta Universitatis Upsaliensis.
- Kolmodin, H. 1993: *Vad tycker provdeltagarna om högskoleprovet 1993-04-17?* (PM Nr 77). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Lavin, D.E. 1965: *The prediction of academic performance*. New York: John Wiley.
- Lexelius, A. & Wedman, I. 1985: *Vad tycker provdeltagarna om högskoleprovet 1985-05-04?* (PM Nr 3). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Linn, R.L. 1990: Admissions testing: recommended uses, validity, differential prediction, and coaching. *Applied Measurement in Education*, 3(4), 297–318.
- Messick, S. 1989: Validity. I R.L. Linn (red): *Educational measurement*. New York: American Council on Education.
- Paskus, T.S. 1997: *Alternative prediction, selection, and utility models for identify-*



- ing academically successful college student-athletes*. Charlottesville: University of Virginia, Department of Psychology.
- Royer, J.M., Marchant, H.G., Sinatra, G.M. & Lovejoy, D.A. 1990: The prediction of college course performance from reading comprehension performance: Evidence for general and specific factors. *American Education Research Journal*, 27(1), 158–179.
- Shepard, L. 1993: Evaluating test validity. *Review of Research in Education*, 19, 405–450.
- Svensson, A., Gustafsson, J-E. & Reuterberg, S-E. 2001: *Högskoleprovets prognosvärde: Samband mellan provresultat och framgång första studieåret vid civilingenjör-, jurist- och grundskollärautbildningarna*. (Högskoleverkets rapportserie 2001:19 R) Högskoleverket: Stockholm.
- Trost, G. 1995: Principles and practice in selection for admission to higher education. I T. Kellaghan (red): *Admission to higher education: Issues and practice*. Dublin: Educational Research Centre.
- Wedman, I. 2000: *Behörighet, rekrytering och urval. Om övergången från gymnasieskola till högskola*. (HSV 2000:6 AR) Stockholm: Högskoleverket.
- Wester, A. 1995: *Ambitions and attitudes toward studies and study results: Interviews with students of the business administration study program in Umeå* (EM No 14). Umeå: Umeå University, Department of Educational Measurement.
- Wester-Wedman, A. 1989: *Vad tycker prövdeltagarna om högskoleprovet 1988-05-07* (PM Nr 23). Umeå: Umeå universitet, Pedagogiska institutionen, Avdelningen för pedagogiska mätningar.
- Wiiand, T. 1998: *Examinationen i fokus: Högskolestudenters lärande och examination: En litteraturöversikt* (Rapport nr 14). Uppsala: Uppsala universitet: Enheten för utveckling och utvärdering.
- Willingham, W.W., Lewis, C., Morgan, R. & Ramist, L. 1990: *Predicting college grades: An analysis of institutional trends over two decades*. Princeton: Educational Testing Service.
- 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.