

Om statens verktyg för skoljämförelser Vem vill dansa SALSA?

MICHAEL HANSEN
ROLF LANDER

Institutionen för pedagogik och didaktik, Göteborgs universitet

***Sammanfattning:** SALSA är Skolverkets webbtjänst där skolors genomsnittliga meritvärde i Årskurs 9 redovisas justerade för elevunderlagets sociala sammansättning. Tanken är att det skall ge pålitligare besked om skolans kvalitet än betyg i rådata, eftersom betygsnivån i hög grad speglar elevernas sociala bakgrund. Artikeln diskuterar några av de problem som är förenade med att göra sådana mått. Den redovisar även ett försök att komplettera SALSA:s mått på social bakgrund (föräldrautbildning) med ett mått på kulturellt kapital samt en modell där elevers skolmotivation och prestationstillit är resultatmått istället för betyg. I försöket används strukturell ekvationsmodellering istället för Skolverkets multipel regressionsanalys och innebörden av det diskuteras. Till slut sätts frågan in i ett styrningsperspektiv med hänsyn till utbildningsmarknadens expansion och en ny skolinspektion.*

For an English abstract see p. 18

I diskussioner om skolors kvalitet och effektivitet brukar begreppet *value added* dyka upp. Vi saknar ett bra svenskt ord för det, men innebörden är ungefär: Faktorer i elevers sociala bakgrund påverkar deras förmåga att hävda sig i skolan och skolors resultat bör därför jämföras med hänsyn till det. »Value added» är resultatmått som statistiskt justerats för social bakgrund och är ett försök att skatta det skolan »lägger ovanpå» hemmets bidrag till elevens resultat. I Sverige publicerar Skolverket på nätet sådana justerade betygsmaått för Årskurs 9 i alla landets grundskolor. Tjänsten kallas SALSA¹. Syftet är att instrumentet skall ge underlag för diskussioner i kommuner och skolor om måluppfyllelsen och dess betingelser.

När value added-analyser först gjordes i Storbritannien såg man dem som en korrektion av marknadsmekanismerna inom skolektorn. Föräldrar skulle med dessa mått få en bättre vägledning om skolors verkliga kvalitet än de listor som pressen publicerar över skolornas examinationsresultat (Saunders 1999). SALSA-verktyget vänder sig i mindre grad till föräldrar eller massmedia, även om alla har tillgång till det, utan är tänkt som ett stöd i skolors utvecklingsarbete. Med SALSA kan skolor sätta sina resultat i relation till de förutsättningar som elevunderlaget ger. Utredningen om en ny svensk skolinspektion, (*Tydlig och öppen. Förslag till ny skolinspektion*, SOU 2007:101 s 134) har i samma anda föreslagit att SALSA skall användas för att skapa ett bättre underlag för inspektionen. Här skymtar alltså en användare utöver skolan och kommunen, nämligen den statliga inspektionen.

Det finns en gråzon mellan faktorer, som kan sägas höra till skolans ansvar, och sådana faktorer, som den inte kan påverka. Föräldrarnas utbildningsnivå och elevernas invandrarbakgrund räknas till sådant som skolor inte kan påverka och alltså inte kan hållas ansvariga för eller ta åt sig äran av. Man kan visserligen säga att i marknadskonkurrensens tid är det just vad en hel del skolor försöker göra, det vill säga påverka rekryteringen av vilka elever man tar emot. Det ingår till och med i strategin för kvalitetsmedvetna skolor, enligt en del förespråkare för Total Quality Management (Bradley 1993). Men i SALSA-data tillåter man dem i princip inte att dra nytta av detta. Resultatens rensä i görligaste mån från elevernas sociala status, så att bara det som skolan kan påverka träder fram och det är det senare som lägger grunden för bedömningen av kvalitet. I teorin.

Här presenteras en kritisk granskning av metoden i SALSA. Verktuget förefaller vara ett viktigt bidrag till värderingen av skolors kvalitet, men vi menar att det inte är tillräckligt bra för sin uppgift. Olika problem redovisas i fyra avsnitt. Frågan är sedan vilka möjligheter det har att bli användbart. Vi återkommer till den frågan efter ett eget försök att med hjälp av enkätmått från eleverna dels förbättra instrumentet som det ser ut nu, dels att vidga dess användning. Att ta in enkätmått i value added-analyser är inte helt vanligt, men det förekommer inom forskningen om effektiva skolor (Thomas, Smees, MacBeath, Robertson & Boyd 2000; Thomas 2001). Avslutningsvis tar vi upp frågan vem som skall »dansa SALSA» och vilken funktion den dansen kan få i skolans styrning.

PROBLEMBESKRIVNINGAR

Bestämning av skolans resultat

SALSA har betygen som det enda resultatmåttet för skolans arbete. Det reser två frågor: Är betygen ett bra resultatmått? Är det skolans enda valida resultatmått? Omedelbart bör sägas att SALSA och value added-analyser handlar om relativa mått, skolor jämförs med varandra. Däremot säger det inget om i vilken mån skolorna uppfyller absoluta kriterier på kvalitet. Om vi tar exemplet sex- och samlevnadsundervisning, som skolor regelbundet får kritik för, så skulle en value added-analys kunna peka på skolor som är bättre än andra, men som till exempel Folkhälsoinstitutet kanske inte skulle tycka höll måttet. På andra områden kan det hända att även de sämsta skolorna gör tillräckligt bra ifrån sig.

Gustafsson (2006 s 56) avslutar en genomgång av forskning kring det svenska betygssystemet med slutsatsen att det inte ger bra kvalitetsmått för kommunala skolsystem. Det tidigare normbaserade systemet var inte avsedd för detta och det nuvarande har vare sig den logiska uppbyggnaden eller den rätta lokala hanteringen för att duga. Det synes också som om införandet av det senare systemet förvärrade likvärdighetsproblemet på grundskolan (Björklund, Edin, Fredriksson & Krueger 2003) och på gymnasieskolan (Cliffordson 2004). Det finns systematiska problem i själva konstruktionen av de mål- och kriteriebaserade betygen, som låter skolors olika bedömningar få stort inflytande. Till detta kommer att vare sig implementeringen av det nya systemet, eller det politiska ansvarstagandet efteråt, varit tillräckliga för att ge lärare och skolor en bra grund för jämförelser och lärande om praxis (Tholin 2003).

Till de systematiska problem som Gustafsson lyfter fram hör den underliggande föreställningen att det finns generella kunskapsnivåer, som kan beskrivas med giltighet för alla elever. Ett annat problem är att beskrivningen av mål och betygskriterier ligger så nära varandra i abstraktionsnivå att kriterierna uppfattas som mål. Vidare har olika ämnen så olika antal mål att det är mycket svårt att förstå om samma betyg i två ämnen verkligen betyder samma kunskapsnivå. Vare sig utredningar eller Skolverket har tagit »ett gemensamt grepp över hela kursplanekonstruktionen», citerar Gustafsson (2006) från Kroksmark (2002 s 67 f).

I sin genomgång refererar Gustafsson inte till några studier av betygens reliabilitet eller validitet i det nuvarande systemet. Betygen används, trots kritiken, ändå i allmän debatt och i politiska överväganden. Ett tecken på det politiska och administrativa intresset för betygen och SALSA är den nämnda utredningen om en ny inspektion. Om betygen är mycket otillförlitliga är det närmast oetiskt att låta elever, lärare och föräldrar leva i okunnighet om det.

Å andra sidan drar Cliffordson (2008) slutsatsen, att på gymnasienivå är de målrelaterade betygen den bästa tillgängliga prediktorn för framtida studieframgång, bättre än högskoleprov och normbaserade betyg. Detta trots att de principiella problem, som Gustafsson (2006) refererar, lär vara större för gymnasiebetygen än för grundskolebetygen. I en annan studie skattar Klapp Lekholm och Cliffordson (2008) att skolskillnader i betyg utgör ca 6 procent av den totala betygsvariationen. Det är i linje med, och till och med något lägre än, skolskillnadernas storlek på de nationella proven.

Gustafsson diskuterar inte de nationella proven i Årskurs 9 som alternativ till betygen. Proven torde ha bättre specifik validitet och reliabilitet för just vad de mäter, men är å andra sidan begränsade till särskilda aspekter i endast tre ämnen, och bör dessutom inte, till skillnad från betygen, ta hänsyn till den lokala tolkningen av läroplanen, vilket innebär en insnävning relativt de faktiska kunskapsmålen för eleverna. Samtidigt är det just den lokala tolkningens brister som är ett huvudproblem för betygen, eftersom de nationella styrmedlen inte anses hålla denna inom tillräckligt valida och reliabla ramar.

Som indikator på en skolas resultat använder SALSA dels andelen elever med fullständiga betyg på skolan, dels elevernas genomsnittliga betyg (meritvärdet). Thomas (2001) menar att sådana summerande mått på effekter kan dölja skillnader mellan olika kön, etniska grupper, hög- och lågpresterande elever med flera. Skolans bidrag behöver inte vara likformigt relativt dessa kategorier och deras prestationer i olika ämnen. Analyser av gruppskillnader skulle alltså kunna fördjupa instrumentets användbarhet. Men summerande betyg saknar av detta skäl inte validitet, särskilt inte så länge urvalet till högre studier baseras på dem.

Vi har i vår egen studie inte kunnat göra något åt betygens brister som kvalitetsmått. Däremot har vi genom enkäter till eleverna försökt bredda den nuvarande SALSA-modellen med mått på vad vi kallar skolanpassning. Med detta begrepp vill vi visa att det går att hitta andra mått på skolors resultat än betygsnivån. I grundskolans läroplan anges att skolan skall sträva efter att varje elev utvecklar »nyfikenhet och lust att lära» och »tillit till sin egen förmåga» (Lpo 94 s 9). Även om det finns ett samband mellan betygen och

dessa förhållningssätt så kan skolor rimligtvis vara bättre på att stimulera det ena och sämre på det andra.

Hur mäts social bakgrund?

I SALSA representeras tre aspekter av elevunderlagets sociala sammansättning: socioekonomiska resurser, nationellt ursprung och kön. Elevunderlagets socioekonomiska resurser baseras på utbildningsnivån hos biologiska föräldrar till elever i Årskurs 9. Indexet utgörs av skolans medelvärde på detta mått. Nationellt ursprung indexeras på två olika sätt; andelen elever som är födda utomlands respektive andelen elever vars föräldrar är födda utomlands. Kön indexeras som andelen pojkar på skolan. Betygsresultaten sätts i relation till elevunderlaget genom att de tre indexen tillsammans får predicera en betygsnivå, som sedan jämförs med skolans faktiska resultat.

Man kan hävda att SALSA:s indikatorer på elevunderlagets sammansättning döljer viktiga aspekter av relevans för skolans förutsättningar att fullgöra sitt uppdrag. Det finns rimligen betydande skillnader i hemmens inställning till skola och utbildning inom de sociala kategorier som ett utbildningsindex arbetar med. Om denna inomgruppsvariation innebär att elever med mer resurser tenderar att välja vissa skolor så gör det att indexet på elevunderlaget blir missvisande.

Nash (2003) illustrerar sådana selektionseffekter med data från Storbritannien. Han fann att på skolor med låg genomsnittlig SES hade cirka 60 procent av eleverna med låg SES få böcker hemma (0–50 böcker). Detta ska jämföras med situationen bland motsvarande elever på skolor med hög genomsnittlig SES. Där var det bara nära hälften så många av eleverna med låg SES, eller cirka 35 procent, som hade få böcker hemma. Trots att eleverna hade samma SES var de med andra ord inte jämförbara på en indikator på skolrelaterade resurser i hemmet.

När man kontrollerar för skolornas förutsättningar med grova mått som SES eller utbildningsnivå så missar man sådana skillnader i mer skolnära faktorer. Det innebär i sin tur att man överskattar förutsättningarna för skolor med lägre genomsnittlig SES eller utbildningsnivå. Dessa skolor framstår därför som mindre effektiva än de kanske är.

Samma effekt får andra faktorer, som kan associeras med socialgruppsskillnader, men som inte enkelt kan subsumeras under detta begrepp. I England visar studier av ett par lokala områden att den allmänna sociala miljön inverkar starkt på undervisningskvaliteten, så som den statliga inspektionen mäter den. Ju lägre inkomster, ju mera arbetslöshet, ju sämre bostadskvalitet och ju sämre hälsa i skolornas närområde, desto sämre undervisning fick barnen (Bell 2003). Rimligen beror det i viss grad på att miljön påverkar vilka lärare skolorna kan dra till sig. Engelska myndigheter satsar nu på att få fram vad man kallar kontextualiserade mått på »value added» (Schagen 2006).

Vem är ansvarig för vad?

Vem är ansvarig för vad? Denna frågeställning kan delas upp i flera: Hur mycket ansvarar skolan för när det gäller betygsskillnader? Går det säkert att urskilja skolans ansvar från andra aktörers? Är »hela skolan» den mest adekvata analysenheten i en analys av SALSA-typ?

Teddlie och Reynolds (2001 s 53, 62) sammanfattar flera studier med att 12–18 procent av individuella resultat kan tillskrivas skolan, en effekt de karaktäriserar som »relatively small.» Men de är optimistiska när det gäller betydelsen av dessa procent, därför att man faktiskt hittar skolor med låg social bakgrund hos eleverna, som presterar bra skolresultat och därför att det verkar finnas en tendens att skolor med lägre social bakgrund visar större skillnader i effektivitet. I synnerhet visar en del studier att effektiva skolor framför allt är effektiva i att påverka förändringen av prestationsresultat över tid och att skolfaktorn (som i detta fall kan beräknas till närmare 25%) då är viktigare än bakgrundsfaktorerna.

Den så kallade school effectiveness-forskningen uppmärksammar numera klassrummens betydelse utöver skolans. Creemer och de Jong (2002) menar att om man bara kan undersöka en nivå ovanför individen, så bör man välja klassrumsnivån, eftersom den vanligen har störst betydelse. Författarna tvingas inte till det valet eftersom de lutar på hierarkisk linjär modellering (HLM), som tillåter tre eller fler nivåer. För forskare med större tillit till strukturell ekvationsmodellering är det ett mera allvarligt problem, eftersom denna bara tillåter tvånivåanalys.

Mot framhållandet av klassrummets betydelse kan man hävda att skolan är den styrningsmässigt och formellt mest ansvariga nivån nära eleverna. Eleverna kan välja skola, men inte vilken klass de ska gå i eller vem som ska vara deras lärare. Skolan är den jämförelsevis mest stabila nivån medan ansvaret för klassrummet även i enskilda ämnen över tid kan variera mellan många lärare eller lärarlag. Man kan säga att skolan har enskilda lärare och lag som sina instrument för att påverka kvaliteten i undervisningen och att därför skolor bör jämföras.

PISA-studien av matematik för 15-åringar visade att man kunde förklara 71 procent av den skolvisa variationen i resultat med hjälp av socio-ekonomiska data om elevernas bakgrund och den sociala kompositionen i skolorna, skolornas resurser, vissa drag i styrningen av skolorna samt data om elev- och rektorsattityder till undervisningen och skolans klimat förmedlade via enkäter (OECD 2004 s 256).

Här kommer alltså styrning och resurser in, vilket i Sverige dels har en statlig komponent tämligen lika för alla, dels en kommunal komponent som kan variera rätt betydligt. Det påminner om en kritik mot value added-studier överlag, nämligen att de tenderar att lägga all skuld på skolorna, därför att man inte tillräckligt adekvat kan mäta det omgivande samhällets betydelse. Kritiken ansluter sig till den ovan framförda synpunkten att indikatorn föräldrars utbildning inte är ett tillräckligt mått på sociala skillnader.

Kommunen kan visserligen påverka närsamhället, men inte tillräckligt. Vilket ansvar har då lokal administration och politik? Problemet illustreras av frågan om vilken effekt den fysiska miljön (luftkvalitet, oväsen, temperatur, ljusförhållanden etcetera) har på elevernas skolprestationer och anpassning (Woolner, Hall, Higgins, McCaughey & Wall 2007). Den som sitter på pengarna (oftast huvudmannen) kan tyckas vara främst ansvarig för den fysiska miljön. Å andra sidan kan en skola med fantasi och arbete få även usla miljöer drägliga. Det kan vara en tolkningsfråga vad man gör av sina yttre villkor.

Den statistiska metoden

SALSA:s justerade mått tas fram via multipel regressionsanalys. Den så kallade förklarade variansen är blott cirka 40 procent, det vill säga mindre än hälften av variationen i skolornas betyg kan hänföras till variationen i elevunderlagets sociala sammansättning. Det skulle betyda att cirka 60 procent beror på att skolorna har olika kvalitet, men riktigt så är det inte. Restposten, det som inte »förklaras», kallas residual och i den ingår inflytandet från alla faktorer, som påverkar resultatet för en enskild skola, utom just dem man tagit hänsyn till i regressionsanalysen.

Skolverket säger försiktigtvis att residualen inte duger för att rangordna skolor, men att det kan ge en utgångspunkt för den lokala diskussionen om förutsättningar och resultat. Det är lätt att underskatta betydelsen av detta problem för kvaliteten i den »lokala diskussionen». Faran med alla analyser, som har sin bas i en regressionsmodell, är att de fäster blicken på det som modellen inkluderar. I själva verket måste de slutsatser man drar om skolverkligheten också väga in det som *inte* finns med i modellen, det som inte mäts. Men det kräver ibland ett betydande mått av fantasi och kunskap för att se vad som saknas i en regressionsmodell. Risken är därför att systemets användare slår sig till ro med den enklaste tolkningen; att man med SALSA gör skolor med olika förutsättningar jämförbara.

Både den praktiska användningen och den teoretiska giltigheten i SALSA skulle vinna på att den förklarade variansen från sådant som skolan inte kan hållas ansvarig för, till exempel elevunderlagets sociala sammansättning, närmar sig ett tänkt »sant» värde. Ovan har vi sett siffror från svensk och internationell forskning, som tillmäter annat än skolan den största betydelsen. Vi menar att den nuvarande förklarade variansen i SALSA bör kunna höjas. Huvudvägen för det är naturligtvis urvalet av faktorer man kontrollerar för, men den statistiska metoden är också viktig.

Vi menar att SALSA-analyser bör ha strukturell ekvationsmodellering (SEM) som statistiskt redskap. Strukturell ekvationsmodellering är en betydligt kraftfullare metod än multipel regressionsanalys. Kraftfullheten ligger i flera egenskaper. Genom konfirmatorisk faktoranalys skapas i den så kallade mätmodellen latent variabler ur manifesta rådata (se t ex Loehlin 2000). Latenta variabler är rensade från osystematiska mätfel och sådant som inte hör till måttens gemensamma kärna. Därmed höjs precisionen betydligt jämfört med multipel regressionsanalys. Genom »nested models» kan precisionen höjas ytterligare. Det innebär att alla delfrågor i ett mått får lämna variation till en generell faktor medan vissa delfrågor därutöver ger variation till en eller flera specifika faktorer. I den så kallade strukturmodellen kan man med SEM dessutom relatera de oberoende variablerna till varandra och därigenom betydligt bättre studera dynamiken i de olika faktorernas bidrag till den beroende variabeln.

Med SEM kan man via tvånnivåanalys (Muthén 1994) hålla isär variation som systematiskt hör till individerna resp. skolorna (ifall man väljer skola som den andra nivån). En del av variationen i individernas svar beror alltså på att de tillhör en viss social enhet och det är denna som tvånnivå-analysen förmår

skilja ut från variationen mellan individerna som sådana (se t ex Snijders & Bosker 1999). I multipel regressionsanalys kan inga nivåer urskiljas.

Fördelen med SEM illustreras av en analys av betygens dimensionalitet som Klapp Lekholm och Cliffordson (2008) genomförde. I konfirmatoriska faktoranalyser kunde man, både på individnivå och skolnivå, etablera dels kunskapsfokuserade latent variabler med hjälp av betyg och nationella prov i svenska, engelska och matematik, dels en latent betygsfaktor, som hade rensats från den ämnes- och kunskapsspecifika variationen som är gemensam för dessa betyg och nationella prov. På individnivå stod denna andra faktor för cirka 4 procent av betygsvariationen medan den på skolnivå uppgick till cirka 7 procent av variationen i betygen i engelska och svenska. Ur SALSA-perspektiv är det intressanta här att man kan dela upp skolornas kunskapsmått på olika sätt och se hur de prediceras av bakgrundsfaktorerna. En bättre validitet och mera nyanserade resultat väntar.

Ytterligare en fördel med SEM är möjligheten till så kallad tillväxtmodellering (Muthén & Khoo 1998), som kan hantera det inom value added-forskningen påtalade problemet med osäker stabilitet i data, dvs. att skolors residualer sjunker och stiger oväntat mycket på kort tid (Thomas m fl 1997, Schagen & Hutchinson 2003). Särskilt instabila data har små skolor, där ett mindre antal elever med avvikande värden kan ha tydlig påverkan på resultatet i enskilda kullar. SALSA-resultaten skulle redan nu kunna redovisas som glidande medelvärden i tre- eller femårsperioder, men ett bättre sätt vore att rensa data från tillfälliga fluktuationer med tillväxtmodellering. Då kan man också bättre hålla reda på förändrade resultat i olika grupper av skolor, vilket påtalades av Teddlie och Reynolds (2001) ovan. Metoden kräver dock minst tre mätpunkter över tid, så den lösningen har vi inte kunnat använda i vår studie.

DEN EMPIRISKA STUDIEN

Det vi vill demonstrera med den här aktuella studien är (i) vad kompletterande mått på social bakgrund kan åstadkomma, samt (ii) hur en motsvarande SALSA-analys kan göras med attitydmått som indikatorer på skolresultat istället för betyg. Studien hör till ett uppdrag från Skolverket (Lander & Hansen 2006) och dess data har möjliggjorts av ett tidigare uppdrag från Skolverket, nämligen att genomföra enkätstudier i samband med de kvalitetsgranskningar som verket gjorde med skolbesök åren 2001–2002.

Detta läsår granskades tre teman på totalt 83 grundskolor, vilket gav enkäterna delvis olika innehåll (Skolverket 2003)². Några av frågorna användes också i en utvärdering av timplanedelegationens försök där grundskolor fick arbeta utan den nationella timplanen (Hansen & Lander 2005), vilket gav ett tillskott av 12 skolor. Antalet skolor, som kan användas i vår analys, är dock som mest 70, eftersom alla skolorna inte hade aktuella underlagsdata för SALSA. Antalet elever är då 5.137. Enkät svar har vi från eleverna på 52 skolor (3.776 elever).

Utan att vara strikt representativa har båda studierna i sina urval eftersträvat en variation av skolor med hänsyn till social bakgrund och geografiskt läge. De tre storstäderna är dock inte med. Till utvärderingen av timplane-försöket valdes skolor med högre ambitionsnivå för detta försök. Svarsfre-

kvensen per skola varierar mellan 76 och 86 procent. Alla enkäter är insamlade i klassrummen av lärare eller forskare.

Det kulturella kapitalets betydelse

Skolverkets metod är, som nämnts, multipel regressionsanalys medan vi använt strukturell ekvationsmodellering (LISREL 8 och Mplus 3 genom STRE-AMS 3.03; Gustafsson & Stahl 2001, 2005). Skolverkets metod på vårt urval skolor ger en förklarad varians för meritvärdet på 38 procent. Resultatet med SEM visar mycket stor likhet med den multipla regressionsanalysen på samma urval skolor, men den förklarade variansen blir 44 procent.

Vi finner att föräldrautbildningens nivå är starkt positivt relaterad till skolornas meritvärde (den standardiserade regressionskoefficienten är 0,54). Relationen mellan andelen elever födda utomlands och meritvärdet är negativ, men något mindre (-0,34). Det finns en svagt negativ korrelation mellan utbildningsnivå och utländsk härkomst (-0,07). Högre andel pojkar på skolan tenderar att sänka det genomsnittliga meritvärdet, men värdet är långt ifrån signifikant och ingår därför inte i vår modell. Skolverket har en alternativ operationalisering av utländsk härkomst, som räknar andelen elever födda i Sverige, men med föräldrar födda utomlands. Denna variabel tillför inte modellen något i konkurrens med de övriga.

Detta är grundmodellen, som vi söker förbättra med hjälp av faktorn *kulturellt kapital* via enkätfrågor till eleverna. Vi mäter inte kulturellt kapital specifikt i Bourdieus (t ex 1986) mening, som handlar om av andra erkända kulturellt betydelsefulla egenskaper. Men vi behåller namnet, eftersom vi ser måttet som en indikator som träffar ungefär samma räjong. Vi vill fånga sådant i hemmiljön som gynnar eleven i skolmiljön. Enkätfrågorna och mätmodellen i den konfirmatoriska faktoranalysen framgår av Tabell 1.

Vi skapade en »nested model» för individnivån, som fick den bästa modellanpassningen om föräldrastöd fick vara en specifik faktor. Till denna fördes frågorna om stöd i skolarbetet (e), om framtidsdiskussioner (d) och om man diskuterar böcker, filmer eller tv-program hemma (c). Den sista frågan hade vi tänkt skulle höra till enbart den generella faktorn kulturellt kapital, men nu fick den dela sin variation mellan den specifika och den generella faktorn.

Detta betyder att den generella faktorn i hög grad bestäms av frågorna a och b, men att det »samhällstillvända» inom variationen i de övriga frågorna bidrar. Till detta skall nämnas att det i enkäten fanns ännu ett påstående att ta ställning till: »hjälp dig med läxor eller annat skolarbete.» Det fungerar även med denna fråga i både generell och specifik faktor, men anpassningen blir betydligt bättre om den utesluts.

Tendensen att vårt mått fångar det »samhällstillvända» i uppväxtmiljön stärktes när vi utformade grupp-nivån i nästa steg. Det kunde endast ske med de tre första frågorna. Politik, samhälls- och miljöfrågor och relationen till media utgör alltså basen för det kulturella kapital, som systematiskt skiljer elever i olika skolor åt. Föräldrastödet i skolarbetet skiljer inte dessa skolors elever åt. Anpassningen, modellens förmåga att återspegla variationen i data, är god (se Tabell 1).

Vi kan tillägga att modellen prövats särskilt på urvalet av 12 skolor i timplañeförsöket, men med skolklassensom grupp-nivån (46 klasser med 721 elever)

Tabell 1. Mätmodell för kulturellt kapital i två nivåer. Individnivån har den generella faktorn kulturellt kapital (Kkap) och den specifika föräldrastöd (Fstöd). Skolnivån har bara en faktor (KkapB). Termen β är standardiserade regressionskoefficienter och & är residualer. I tabellen är decimalnoteringen amerikansk. Modellanpassning i LISREL: $\chi^2 = 41,8$; $df = 9$; RMSEA = 0,039; CFI = 0,99

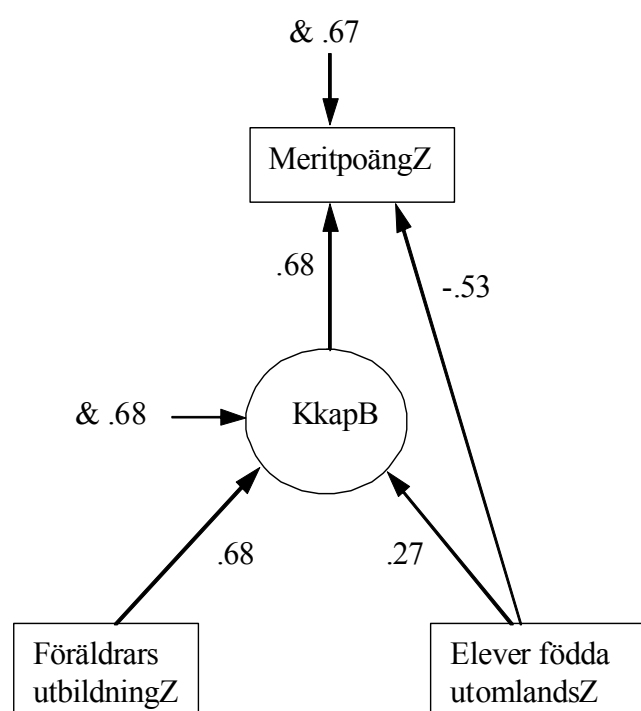
<i>Hur ofta gör någon av dina föräldrar detta? (5-gradig skala från aldrig till ofta)</i>	<i>Kkap</i>	<i>Fstöd</i>		<i>KkapB</i>	
	β	β	&	β	&
a. diskuterar politik med dig	.67		.68	.29	.06
b. diskuterar samhälls- eller miljöfrågor med dig	.86		.46	.21	.06
c. diskuterar böcker, filmer eller tv-program med dig	.40	.30	.83	.21	.11
d. pratar med dig om vad du kan göra i framtiden	.29	.54	.79		
e. uppmuntrar dig i skolarbetet/hjälper dig att förstå saker du tycker är svårt i skolarbetet*	.25	.57	.79		

* Den första formuleringen kommer från Skolverkets studie, den andra från timplane-studien. De två måtten uppförde sig likadant i förberedande faktoranalytiska studier. Därför ansågs det säkert att låta dem representeras av samma variabel.

och då också fått tillfredsställande anpassning. Även skolklasser skiljer sig systematiskt åt i denna variabel.

Är kulturellt kapital ett mått som skolan inte kan påverka? Idealt borde skolan via sina föräldrakontakter kunna tillföra hemmens samtal och intresse för samhällsfrågor något, men i realiteten är det nog knappast så. Karlson Bäck (2005) räknar kulturellt kapital i Bourdieus mening som föräldradeterminerat inte bara genom att det förser eleverna med en viss utgångsnivå för sin förståelse av skolan och motivation för skolarbete, utan också genom att föräldrarnas egna kontakter med skolan styrs av hemmets kulturella kapital. Empiriska norska data stöder hennes tes. Därmed till vår modell, vars skolnivå presenteras i Figur 1.

Den förklarade variansen är 55 procent (100 minus residualen i kvadrat). Förbättringen från Skolverkets modell är 17 procentenheter. I skolverksmodellen är inflytandet från andelen elever födda utomlands på meritpoängen enbart negativ. Men genom att vi använt strukturell ekvationsmodellering ser vi att det också finns en positiv inverkan via kulturellt kapital. Alltså, skolor med elever från utlandet, vars föräldrar kan förse sina barn med kulturellt kapital i den här angivna meningen, har ett högre genomsnittligt meritvärde. Skolor, vars föräldrar i mindre utsträckning förmedlar detta kulturella kapital, har elever med lägre betyg. Det finns naturligtvis andra faktorer förknippade med utländsk bakgrund, som påverkar elevernas betyg, men dessa saknar vi mått på. Summan av dessa icke-observerade faktorer repre-



B-tecknet står för "between" och är en konvention för att beteckna en latent variabel på gruppnivån. Z står för data, som endast förekommer på gruppnivån, här manifesta variabler om skolorna.

Figur 1. Strukturmodell för reanalys av Skolverkets SALSA-modell med strukturell ekvationsmodellering kompletterad med data om kulturellt kapital. Standardiserad modell. & = residualer. Övriga siffror = regressionskoefficienter. Modellanpassning i LISREL: $\chi^2 = 80,4$; $df = 23$; RMSEA = 0,029; CFI = 0,99

senteras av den direkta och negativa relationen från andel elever med utländsk bakgrund till skolans genomsnittliga meritpoäng. Den direkta relationen är betydligt starkare än den indirekta. Dess negativa effekt på meritpoängen är som synes 0,53 medan den totala positiva effekten via kulturellt kapital ($0,27 \times 0,68$) är 0,18.

Kulturellt kapital är en kraftfull mediator, som i Figur 1 tar hand om hela effekten av föräldrarnas utbildningsnivå. I Skolverkets ursprungliga modell är relationen utbildningsnivå – betyg 0,54. Men här är den 0,46 ($0,68 \times 0,68$). Den totala standardiserade effekten av kulturellt kapital är alltså något starkare än den från formell utbildning.

Vad skulle då detta betyda konkret för enskilda skolor? För att få en referensnivå skattade vi först varje skolas residual när föräldrarnas genomsnittliga utbildningsnivå på skolan och andelen elever med utländsk bakgrund fick predicera skolans genomsnittliga meritvärde. Här användes enbart manifesta variabler i en multipel regressionsanalys, vilket innebär att referensnivån motsvarar Skolverkets nuvarande modell. Därefter skattades skolornas residualer enligt vår modell (se Figur 1) med kulturellt kapital som prediktor tillsammans med föräldrarnas utbildningsnivå och andelen elever födda utomlands.

Kulturellt kapital mäts med enkätfrågor och det finns flera sätt att använda informationen i enkätsvaren. Ett sätt är att summera varje individs svar på frågorna om aktiviteterna i hemmet och sedan beräkna varje skolas medelvärde på summan. Det finns emellertid skäl att anta att bland annat bristande reliabilitet i enkätdata gör att denna metod underskattar det kulturella kapitalets betydelse och att man istället bör arbeta med sk faktorpoäng (Yang-Hansen, Rosén & Gustafsson 2006). Faktorpoäng innebär att den samlade informationen i SEM-modellen används för att skatta varje skolas värde på den latent variabeln kulturellt kapital. På detta sätt åstadkommer vi ett mått på kulturellt kapital med mindre mätfel, som är mindre känsligt för bortfall i enkätdata och som tar hänsyn till att olika frågor mäter kulturellt kapital olika bra. Vi skattade faktorpoängen med hjälp av programmet Mplus 3 (Muthén & Muthén 2004).

På detta sätt fick varje skola två residualer, en baserad på Skolverkets nuvarande modell och en baserad på en modell som även inkluderade kulturellt kapital. Om man sätter 10 poängs skillnad mellan de två residualerna, som något att bry sig om, så förändras läget för 10 procent av skolorna. Om 10 poängs skillnad är meningsfull kan inte vi avgöra, men om den är det och vårt urval skolor är någorlunda representativt, så skulle var tionde skola i landet ligga bättre eller sämre till i dessa nya SALSA-data än vad de gör idag. En förändring i en skolas residual på 10 poäng kan exempelvis sättas i relation till den variation som finns över tid för en enskild skolas meritvärde. Utifrån sju års meritvärden för 829 skolor skattades den genomsnittliga standardavvikelsen till 8,0 poäng, med större mellanårsvariation för mindre skolor och mindre för större skolor.

Skolanpassning som skolans resultat

Vi har nu gett exempel på alternativ till SALSA:s mått på skolornas elevunderlag och påvisat vinsterna med strukturell ekvationsmodellering. I det följande kommer vi att visa hur alternativ till betygen som skolans resultat kan utformas. Vi har möjlighet att undersöka skolmotivation med våra data. Skolmotivation är ett bredare begrepp än lust att lära. Vi har även försökt ta in aspekten att vara uthållig i svårigheter. Prestationstillit är en egen översättning av det internationella begreppet »academic self-concept» (Marsh, Walker & Debus 1991), som är ett delvis överlappande, delvis konkurrerande begrepp till »self-efficacy» (Bandura 1997), översatt till *kapacitetsupplevelse*. Det överlappande ligger i att båda avser att mäta självkänsla inför prestationer.

»Academic self concept» nöjer sig med att specificera ett skolämne med till exempel 3–4 frågor. Man kan också låta det gälla skolarbete i allmänhet, men även ett område utan nödvändig relation till »academic», såsom hälsa (t ex hälsotillit; se Lander 2002). »Self-efficacy» kräver en noggrann specifikation av särskilda kunskaper och färdigheter i en viss kompetens eller uppgift och ett antal frågor som matchar den. Konkurrensen gäller vilket mått som predicerar prestationer bäst.

Det har gjorts jämförande undersökningar med olika resultat. Pajares (1996) drar slutsatsen att kapacitetsupplevelse nog är bättre när den svarande är väl förtrogen med kontexten och uppgifterna medan *prestationstillit* kan vara bättre när man känner till området på ett allmänt sätt och när det inte är helt klart vad som specifikt åsyftas. I de två studier, som bildar basen för denna

undersökning, används båda, men på grund av den mera specifika inriktningen i mätningen av kapacitetsupplevelse, så är de olika enkätvarianterna inte jämförbara. Återstår alltså prestationstilliten, som allmänt gäller skolprestationer utan ämnesspecificering. För vårt syfte kan vi anta att mått på prestationstillit är mera rättvisande för en jämförelse, eftersom det inte, till skillnad från mått på kapacitetsupplevelse, behöver anpassa sig till särskilda uppläggningar av kraven, lokalt eller i försök.

Motivation och självkänsla har gemensamma teoretiska rötter. Motivation kan sägas vara en funktion av intentioner, känslor och föreställningar man har om sig själv som agent. Det sista, »personal agency beliefs», är det som avses med begrepp som »self-efficacy» och »academic self concept» (Giota 2001 s 26). Man kan då förvänta sig att det går att skapa ett gemensamt mått, vilket vi här kallar skolanpassning. Ja, det visar sig dessutom vara nödvändigt, eftersom två mått skulle innebära kollinearitet i relation till meritvärdet, dvs deras samband med det senare ligger så nära att de stör varandra.

Analysen bjöd dock på vissa svårigheter (se Tabell 2). Vi försökte återigen skapa en »nested model» för den individuella nivån och hade, som nämnts, förväntat oss att Fråga d om uthållighet skulle höra till den specifika faktorn motivation. Men det gick inte, utan den vände sig mot den generella faktorn prestationstillit. Vi tror att orsaken dels kan vara den särskilda formuleringen

Tabell 2. Mätmodell för skolanpassning i två nivåer. Individnivån har den generella faktorn prestationstillit (Ptill) och den specifika skolmotivation (Mot). Skolnivån har bara en faktor (PtillB). Standardiserad modell. Modellanpassning i LISREL: $\chi^2 = 66,6$; $df = 15$; RMSEA = 0,036; CFI = 0,99. COV står för kovarians mellan residualer. Förklaringar se tabell 1.

<i>Hur fungerar skolarbetet för dig? (5-gradig skala från aldrig till ofta)</i>	<i>Ptill</i>	<i>Mot</i>		<i>PtillB</i>	
	β	β	&	β	&
a. jag är bra på att lägga upp skolarbetet så att jag hinner med det	.65		.75	.12	.04
b. jag är bra på att skriftligt redovisa ett arbete jag gjort*	.57		.80	.18	.04
c. jag har lätt för att lära mig i skolan	.55		.81	.18	.06
d. jag kämpar på även när det är svårt**	.65		.75	.08	.08
e. innehållet i olika ämnen är ofta spännande	.53	.35	.76	.14	.09
f. jag tycker det är roligt att lära sig olika saker i skolan	-.05	.25	.86	.14	.42

COV (b&,c&) = .22; COV (e&,f&) = -.09

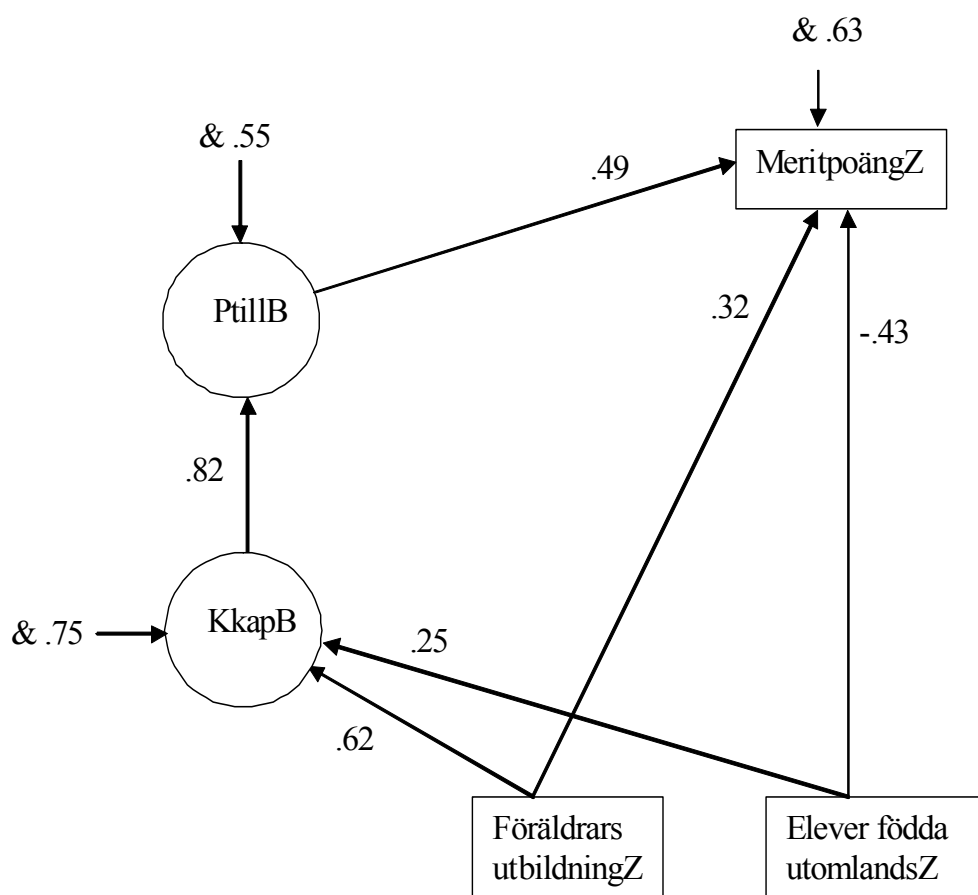
* = Formuleringen i Skolverkets studie. Formulering i Timplanestudien: »skolarbetet brukar gå bra för mig».

** = Formuleringen från skolverksstudien, i timplanestudien löd den: »jag jobbar hårt för att lära mig viktiga saker».

som användes i timplanestudien, dels att en enda fråga har svårt att balansera den lustfyllda aspekten i de två andra motivationsfrågornas formuleringar.

Lägg också märke till det svagt negativa sambandet för frågan om det roliga i skolarbetet relativt prestationstilliten. Tilliten på individnivå hör samman med att innehållet kan vara spännande, men inte med att lärandet är roligt. En kovarians visar också att det som blir kvar i variablerna spännande och roligt när det gemensamma för faktorerna dragits ur, inte alls är synonymt. På skolnivå upphävs dock dessa nyanser mellan de två motivationsfrågorna. En konfirmatorisk faktoranalys med enbart motivationsfrågorna får inte en signifikant grupp-nivå, men plockas de in i en modell tillsammans med prestationstillit går det alltså. Det syns inte i Tabell 2, men det är Fråga f om det roliga i att lära sig i skolan, som står för det större bidraget till variationen på skolnivå.

Figur 2 visar modellens grupp-nivå när skolanpassning får mediera mellan kulturellt kapital och meritpoäng. Anpassningen i modellen är god. Jämfört



Figur 2. Strukturmodell där Skolverkets modell kompletterats med kulturellt kapital (KkapB) och skolanpassning (PtillB). Enbart grupp-nivå.

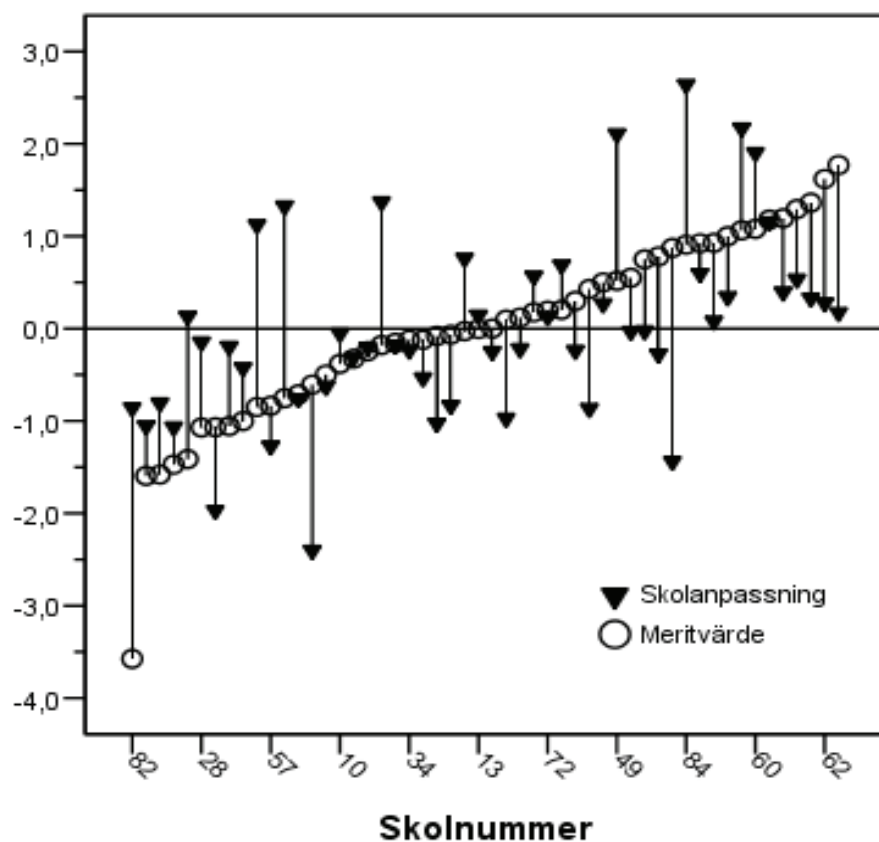
Modellanpassning i LISREL: $\chi^2 = 1299$; $df = 298$; RMSEA = 0,025; CFI = 0,94. Förklaringar, se figur 1.

med i Figur 1 finns inte längre någon direkt relation mellan kulturellt kapital och betygen, utan kapitalet arbetar helt genom skolanpassningen. Vi ser att skolanpassningen är starkt beroende av kulturellt kapital.

Residualen för skolanpassning i Figur 2 visar att måttets förklarade varians är 68 procent.³ Modellens förmåga att förklara skolanpassning är alltså bättre än dess förmåga att förklara meritbetygen. Skillnaden är 68–55 procent, dvs 13 procentenheter.

Vilken skillnad skulle det göra för enskilda skolor om skolanpassning vore ett mått på skolresultat i SALSA? Vi utgår från modellen i Figur 1, men ersätter meritvärdet med skolanpassning som enda resultatmått. Vi räknade ut skillnaderna mellan varje skolas faktiska och justerade värde på skolanpassning respektive meritvärde. Residualerna är standardiserade denna gång, eftersom enkät- och betygsmått inte har samma skalor. Resultatet visas i Figur 3.

Om vi sätter en standardiserad residual av någon betydelse som 1,0, så får 40 procent av skolorna en annorlunda position på de två måtten. Ungefär lika många har bättre meritvärde än skolanpassning (21%, trianglar nedanför cirklar i figuren) respektive bättre skolanpassning än meritpoäng (19%, trianglar ovanför cirklar). Resultatet för de senare skolorna kan tolkas som att man förmår att hålla modet uppe hos eleverna mer än vad deras relativa



Figur 3. Standardiserade residualer för skolanpassning och meritvärde för 52 skolor.

prestationer egentligen borde tillåta dem. Men skillnaden kan också ses som ett uttryck för att dessa skolor har potential att nå högre resultat, eftersom eleverna är förhållandevis positivt inställda till skolans verksamhet.

Innan vi lämnar modellen i Figur 2 bör vi uppmärksamma att där även finns en pil från skolanpassning till meritvärdet. Den påverkar inte de analyser av skolskillnader som nyligen gjorts, utan visar den logiska konsistensen i modellen. Men om vi skulle dra in denna relation i beräkningen av skolskillnader, vilken effekt skulle det få och hur kan det motiveras? Den förklarade variansen skulle med detta öka till 60 procent. Argumentet för att acceptera detta skulle kunna vara, att det kulturella kapitalet ändå inte fångar upp tillräckligt av hemmens inflytande över aspirationer, självbilden i relation till akademiska studier och vad Nash (2003 s 446) kallar »a willingness to accept the regime of schooling». Denna lösning skulle föra begreppet till hemmets ansvar.

Skolanpassning i vår modell ovan är egentligen teoretiskt ett »gråzonsbegrepp» med delat ansvar för skolan och hemmen. Motivation kan antas vara mera skolberoende än prestationstillit. Det hade varit mera användbart om en »nested model» på skolnivå hade förmått skilja på prestationstillit och skolmotivation. Det kan nog åstadkommas med mättekniskt utvecklingsarbete.

Det är möjligt att istället låta meritvärdet predicera skolanpassningen. Därmed skulle man åstadkomma att skolanpassningsresidualen inte speglar skolskillnader som är relaterade till betygsnivån. När vi genomför en sådan analys visar sig det inte ge så stora effekter på skolnivå annat än för några skolor med extrema residualer. Vi kommer alltså inte avsevärt längre med en sådan metodik, skolor som klart håller isär självkänsla och prestationer är sällsynta.

DISKUSSION

Vi ska strax diskutera i vilken mån som vår empiriska studie sagt något väsentligt i frågan om SALSA:s vara eller icke vara. Vi ska också diskutera de problembilder vi tog upp inledningsvis. Men vi vill börja diskussionen med ännu en problembild, vilken bara antydde i början, men som ger frågeställningen en viss angelägenhet. Idag har vi en kraftig expansion av marknaden för val av skolor, både privata och offentliga. Det finns dock politiska röster från olika håll, som framhåller behovet av att moderera marknadsmekanismerna. SALSA kan ha ett intresse i detta sammanhang. Vi menar att modererande krafter bör arbeta både med lärarna, som ansvarar för betygssättningen, och med föräldrarna och eleverna, som använder betygen som ett sätt att värdera skolors kvalitet.

En kraftig inverkan på lärarnas relation till marknaden finns redan i bruk, nämligen de nationella proven. De tas fram med vetenskapliga metoder, men baseras även på utbildningsideologiska antaganden, vilka kan stämma mer eller mindre med enskilda lärares inställning. Proven utgör en syntes mellan akademiska förhållningssätt på högskole- och universitetsnivå och lärarnivå, eftersom människor från dessa miljöer samarbetar i utprovandet av dem.

Proven främjar också en gemensam professionalism genom att lärarna måste förhålla sig till gemensamma kunskapskriterier. Man kan säga att proven, tillsammans med en gemensam lärarutbildning, bidrar till att upprätthålla en professionell marknad till skillnad från marknader där lärarna delar upp sig efter andra kriterier (Erickson & Lander 2007).

Det har påpekats att skolor med särskilda pedagogiska eller religiösa ideologier hellre konkurrerar med det unika i sina egna kunskapsdefinitioner för att locka till sig speciella grupper och därmed sluter sig mot extern utvärdering och jämförelser. Marknaden tenderar då mot oligopol. Om man, som vi, positivt värderar en gemensam professionalism med utbyte mellan forskning och skola, så bör man fråga sig om SALSA kan spela en motsvarande roll som de nationella proven.

Betyg är mera utsatta för marknadspress än nationella prov. Betygsnivåerna på gymnasiet har stigit på senare år, men de motsvaras inte av bättre resultat på universitet och högskolor (Cliffordson 2004). Som påpekas i utredningen om en ny skolinspektion finns det tendenser att betygssättningen på grundskolan avviker mer än rimligt uppåt jämfört med de nationella provens resultat. En fungerande SALSA-modell skulle kunna bidra till att motverka sådant.

Betygsinflation med social snedvikt skulle sannolikt öka den varians som förklaras av föräldrautbildning och kulturellt kapital. En SALSA-modell med enkätbaserad skolanpassning som predicerad faktor torde inte påverkas på samma sätt, trots att ett visst samband mellan betygsgivning och skolanpassningsattityder är trolig. Men skillnaden mellan de två måtten borde tendera att öka. Det skulle kunna ge underlag för självprovning inom skolan, eventuellt hjälpt på traven av kommunal utvärdering eller skolinspektionen. Det finns en del som talar för att det är i sådana sammanhang, i diskussionen mellan skolor och extern utvärdering, som måtten har sin relevans.

En nationell inspektion bör, liksom nationella prov, inverka på lärarprofessionalismen, så att den bättre främjar en vakthållning kring gemensamma kriterier för kvalitet i utbildning⁴. Som påpekats har staten försummat implementeringen av och fortbildningen om betygsreformen. Här ges en chans att återvinna mark och stärka lärarnas medvetenhet om problemens karaktär.

Det bör öka medvetenheten även hos föräldrarna, men man bör inte förvänta sig några enkla sådana effekter. Inledningsvis nämndes att value added-mått kunde ses som ett sätt att understödja föräldrars och elevers valfrihet med information om vad skolor egentligen presterar. Vissa föräldrar vill dock egentligen inte hitta den »bästa» skolan, utan söker sig till sådana skolor där deras barn kan vidmakthålla eller utveckla rätt attityder för en karriär inom vissa sociala skikt, eller för hemmahörighet i en viss kultur, just därför att barn i dessa skikt eller kulturer redan går i dessa skolor. Hur »bra» skolor är rent undervisningsmässigt är då underordnat. Vi får alltså inte tro att marknaden följer någon enkel rationalitet. Den behöver inte påverkas av tekniska metoders säkerhet, vilket dock inte minskar behovet av forskning och utvärdering för att balansera sådana attityder.

Pekar man på vådan av marknadens tryck bör man även peka på vådan av för stark styrning. Det som händer då är vanligen att den styrda parten utvecklar manipulativa motmedel, vilket har kunnat iakttas i engelska skolor inför deras nationella inspektion (Lander & Granström 2000). Visscher (2001) pekar också på sådana frestelser då »value added»-data ses som farliga eller lockande: Skolor fokuserar sina åtgärder på ur »value added»-perspektiv profitabla elever. Man eftersträvar urval i elevantagningen, förflyttar besvärliga elever, lägger tonvikten vid snävt betygsmätta kunskaper och ignorerar andra mål, uppfinnar kreativa rapporteringstekniker etcetera.

Vasquez Heilig och Darling-Hammond (2008) rapporterar likartade konsekvenser från ett testbaserat »accountability»-system i Texas, där många på test lågpresterande elever tvingades ur skolorna för att inte dra ner testresultaten. Det syntes inte i delstatsdata, varav politiker och administratörer kunde slå sig för bröstet och hävda att systemet var effektivt i att höja resultaten. Sådana »gaming strategies» anses av vissa vara ett argument för att endast informell rapportering av »value added»-analyser direkt till skolan skall tillåtas.

Skall man överhuvudtaget fortsätta med verktyg av SALSA:s slag är det vår mening, att det bör sättas in i ett sammanhang med kraftigt förstärkt stöd för betygs- och kursplaneutveckling samt en diskussions- och fortbildningsapparat, gärna knuten till den nya skolinspektionen. Därtill måste det tekniska arbetet med själva modellen fortsätta, och därmed kommer vi till vår studie och våra övriga problembeskrivningar.

Vårt bidrag består alltså av ett metodskifte, som dels innefattar användandet av strukturell ekvationsmodellering, dels elevenkäter som kompletterar social bakgrund och ger ett nytt effektmått. Vi har med detta förbättrat den förklarade variansen i meritbetyget med 17 procentenheter (till 55 procent) jämfört med SALSA och skapat resultatmåttet skolanpassning med en förklarad varians på 68 procent. Vi har demonstrerat att detta i det första fallet kan göra skillnad i värderingen för 10 procent av skolorna och att 40 procent av skolorna får olika bedömningar när det gäller residualerna för meritvärdet och skolanpassningen. Särskilt det senare verkar ganska lovande för en lokal diskussion. Men vi kan inte hävda att detta har placerat SALSA på säker mark som ett verktyg att hålla social bakgrund under kontroll när man jämför skolors resultat.

Troligen skulle det gå att komma vidare med de här modellerna redan genom att skapa individuellt baserade manifesta variabler för de ursprungliga SALSA-måtten. Som sades ovan är dessa genomsnittssiffror för respektive skola fördelade på alla skolans elever. Våra enkäter ger individuell variation dess rättmätiga utrymme och det är denna som gruppnivåns latent variabler baseras på. På samma sätt skulle man kunna göra med föräldrarnas utbildningsbakgrund, kön och utländsk härkomst, givet att man kan föra samman dem med enkätsvaren i samma datafil via individuella koder. Vi har pekat på möjligheten att utnyttja dimensionaliteten i kunskapsmått för mera nyanserade och valida mått, vilket ställer samma krav på individmätning.

Det reser förstås frågor om etik i datalagens mening. Utvägar som kan prövas är att ta de tre bakgrundsfrågorna via elevenkäten eller att föräldrarna ger sitt samtycke till sammanförande av individuella data för kvalitetsmätande syfte. En sådan datainsamling och diskussion kring resultaten kan få en adekvat form i samband med den självvärdering skolan tänks göra inför den nya skolinspektionen. Med inspektioner var tredje år skapas snabbt en betydande databas. Enligt vår mening har det väsentliga fördelar att inspektörerna genomför datainsamlingen. Det bör innebära möten med klasserna i nian på respektive skola, vilket lämpar sig utmärkt för inledande information om inspektionens syften och genomförande. Att inspektörerna genomför enkätinsamlingen förstärker klart enkätdatas reliabilitet. Omfattningen av de använda frågeformulären behöver inte vara stor. De här använda frågorna för eleverna ryms väl på en A4-sida.

Om ett verktyg för value added-analyser ska vara tillförlitligt och meningsfullt, så förutsätter det att merparten av de variabler, som inverkar på elevers och skolors betygsnivå eller skolanpassning, kan mätas och specificeras i en analytisk modell (Raudenbusch & Willms 1995). I SALSA är det framför allt mått på elevernas tidigare kunskaper och motivation, till exempel i Årskurs 5, som saknas, förutom mer nyanserade mått på hemmiljön. Gör man sådana studier kan man finna att förkunskaper står för nära 80 procent av variationen i skolors genomsnittliga betygsnivå (Gustafsson & Myrberg 2002 s 22). Denna andel kommer förstås både från den sociala bakgrunden och personalens insatser i skolans tidigare årskurser. Vi är alltså återigen ute i gråzonen mellan skolornas och hemmens ansvar, men högstadiets roll blir i alla fall klarare. Dessutom har vi kvar den i stort okända interaktionen mellan skolornas ansvar och deras huvudmäns. Debatten kunde hyfsas om detta erkändes och den ovan nämnda PISA-studien kan ge vägledning åt utvecklingsarbete.

Det måste dock stå klart att en modell, som tar hänsyn till alla viktiga faktorer, inte kan åstadkommas. Av det skälet bör SALSA tillämpas med flera utfallsvariabler. Vi har här pekat på betyg, nationella prov och skolanpassning. Enligt Visscher (2001) är »value added» en nödvändig förutsättning för valida jämförelser mellan skolor, men inte en tillräcklig. Rätt sammanhang för användningen av instrumentet är i en analytisk diskussion och absolut inte i endimensionella rangordningar över skolor.

Själva definitionen av skoleffektivitet som restprodukten i en statistisk analys verkar som en för enkel representation av komplexiteten i verkliga kausala relationer. Inom internationell forskning är man ivrigt sysselsatt med att komplettera sina statistiska analyser med kvalitativa fallstudier samtidigt som länderna engagerar sig i uppbyggnaden av nationella databaser på vilka den här typen av analyser är möjliga (Schagen & Hutchinson 2003). När väl data började bli tillgängliga kunde man också studera skillnader inom skolor, till exempel mellan ämnesgrupper. Det är ju inte givet att en skola, som är effektiv i matematik också är det i språk. I jämförelse med detta har SALSA stått stilla.

Vi menar att kunskapsläget på detta område förändras ganska snabbt och att svensk forskning borde uppdatera sig och på allvar pröva ett value added instruments möjligheter och risker. I synnerhet är detta viktigt om en ny skolinspektion skulle börja intressera sig för verktyget. Utredningens formuleringar om SALSA och value added är inte särskilt nyanserade. Det borde vara så att regeringar inte lanserade starka styrmedel inom någon samhällssektor, utan att låta den åtföljas av forskning och utvärdering, det vill säga av analys grundad i adekvat empiri. En samverkan mellan till exempel inspektion och forskning borde kunna åstadkomma realistiska bedömningar av huruvida det kommer att löna sig att dansa SALSA eller ej.

ABSTRACT

Hansen, Michael & Lander, Rolf: A state instrument for comparing schools: Who wants to dance salsa. *Pedagogisk Forskning i Sverige*, Vol. 14, No. 1, pp. 1–00.

Sweden has an indicator system for grade 9 in comprehensive schools building on value added measures. The aim is to stimulate improvement, and not to rank schools in quality. Data are easily available at the webb. A government committe suggested

recently that the system should be used by a new, more powerful state inspection. Whatever purpose, the problem with the statistical model is its weakness in taking account of the variance in marks that cannot be related to school composition in terms of parents' education, ethnicity, and gender distribution among pupils. Explained variance is about 40 per cent, and the rest (the residual) is taken as the indicator of a school effect. The article shows some ways to improve this: Structural equation modelling, and pupil questionnaire data about the transfer of cultural capital in families. By these means the explained variance is increased to 55 per cent. By using school adjustment as another value added measure explained variance was 68 per cent. Variation between schools is markedly higher with school adjustment than with marks. Still, there are heavy problems with the model. The article concludes by discussing what functions this indicator system may have in the steering system as a whole.

Keywords: Evaluation, assessment, value added, governing, school indicators, cultural capital

NOTER

1. Skolverkets arbetsverktyg för lokala sambandsanalyser; <http://salsa.artisan.se/>
2. Författad av Rolf Lander och Joanna Giota.
3. En parallell körning med programmet AMOS gav 70 procent.
4. Vakhållningen bör dock inte bli för snävt avgränsad. Inspektionsutredningen (SOU 2007: 101) tenderar att gå för långt i det avseendet genom att vara för närvarande i värderingen av undervisningskvalitet.

LITTERATUR

- Bandura, A. 1997: *Self-efficacy: The exercise of control*. New York: Freeman.
- Bell, J. 2003: Beyond the school gates: The influence of school neighbourhood on the relative progress of pupils. *Oxford Review of Education*, 29(4), 485–502.
- Björklund, A., Edin, P.-A., Fredriksson, P. & Krueger, A. 2003: *Den svenska skolan: Effektiv och jämlik?* Stockholm: SNS Förlag.
- Bourdieu, P. (1986): The forms of capital. I G. Richardson (red): *Handbook of theory and research for the sociology of education*. New York: Greenwood.
- Bradley, L.H. 1993: *Total quality management for schools*. Lancaster: Technomic.
- Cliffordson, C. 2004: Betygsinflation i de målrelaterade gymnasiebetygen. *Pedagogisk forskning i Sverige*, 9(1), 1–14.
- Cliffordson, C. 2008: Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment*, 13(1), 56–75.
- Creemers, B. & De Jong, R. 2002: Explaining differences in student outcomes. *Journal of Classroom Interaction*, 37(2), 16–26.
- Erickson, G. & Lander, R. 2007: Der Kitt, der ein wachsendes System zusammenhält? Nationale Test als Kern der Qualitätssicherung in Schweden. *Pädagogik*, 59(3), 32–35.
- Giota, J. (2001): *Adolescents' perceptions of school and reasons for learning*. (Göteborg Studies in Educational Sciences 147) Göteborg: Acta Universitatis Gothoburgensis.
- Gustafsson, J.-E. 2006: *Barns utbildningssituation: Bidrag till ett kommunalt barnindex*. Stockholm: Rädda Barnen.
- Gustafsson, J.-E. & Myrberg, E. 2002: *Ekonomiska resursers betydelse för pedagogiska resultat: en kunskapsöversikt*. Stockholm: Liber.
- Gustafsson, J.-E. & Stahl, P.A. 2001: *STREAMS User's guide. Version 2.5 for Windows*. Mölndal: MultivariateWare.
- Gustafsson, J.-E. & Stahl, P.A. 2005: *Using Mplus with STREAMS 3.0*. Mölndal: MultivariateWare.

- Hansen, M. & Lander, R. 2005: *Effekter i försök med timplanen på grundskolan: Slutrapport till regeringens timplanedelegation*. (Rapport 2005:14) Mölndal: Göteborgs universitet, Institutionen för pedagogik och didaktik.
- Karlsen Bæck, U-D. 2005: School as an arena for activating cultural capital. *Nordisk Pedagogik*, 25(3), 217–228.
- Klapp Lekholm, A. & Cliffordson, C. 2008: Discrepancies between school grades and test scores at individual and school level: effects of gender and family background. *Educational Research and Evaluation*, 14(2), 181–199.
- Krokmark, T. 2002: En tankes fall i praktiken – då den målrationala styrningen möter skolan. I *Att bedöma eller döma. Tio artiklar om bedömning och betygsättning*. Stockholm: Skolverket.
- Lander, R. 2002: Professional cooperation around self-related measures in school indicator instruments. *Journal of Classroom Interaction*, 37(2), 27–36.
- Lander, R. & Hansen, M. 2006: *Bättre och sämre skolor? Analyser av Skolverkets SALSA-data tillsammans med enkätdata*. Göteborg: Göteborgs universitet, Institutionen för pedagogik och didaktik.
- Lander, R. & Granström, K. 2000: Skolinspektion i England och Sverige. Hjälptill självhjälp eller självstyrning med blets i munnen? *Pedagogisk forskning i Sverige*, 5(3), 215–234.
- Loehlin, J.C. 2004: *Latent variable models: an introduction to factor, path, and structural equation analysis*. Mahwah, N.J.: Lawrence Erlbaum.
- Lpo 94: *Läroplan för det obligatoriska skolväsendet, förskoleklassen och fritidshemmet Lpo 94*. Stockholm: Skolverket.
- Marsh, H.W., Walker, R. & Debus, R. 1991: Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology*, 16(4), 331–345.
- Muthén, B.O. 1994: Multilevel covariance structure analysis. *Sociological Methods Research*, 22(3), 376–398.
- Muthén, B.O. & Khoo, S-T. 1998: Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, 10(2), 73–101.
- Muthén, L.K. & Muthén, B.O. 2004: *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Nash, R. 2003: Is the school composition effect real? A discussion with evidence from UK PISA data. *School Effectiveness and School Improvement*, 14(4), 441–457.
- OECD 2004: *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.
- Pajares, F. 1996: Self-efficacy in academic settings. *Review of Educational Research*, 66(4), 543–578.
- Raudenbush, S.W. & Willms, J.D. 1995: The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307–335.
- Saunders, L. 1999: A brief history of educational 'Value Added': How did we get to where we are? *School Effectiveness and School Improvement*, 10(2), 233–256.
- Schagen, I. 2006: The use of standardized residuals to derive value-added measures of school performance. *Educational Studies*, 32(2), 119–132.
- Schagen, I. & Hutchinson, D. 2003: Adding value in educational research. The analysis of data and analytical power. *British Educational Research Journal*, 29(5), 749–765.
- Skolverket (2003): *Självkänslan och skolans vardag. En enkätstudie av elever och lärares attityder till information och kommunikation, lusten att lära, tid för lärande. Nationella kvalitetsgranskningar 2001–2002*. Stockholm: Skolverket.
- Snijders, T.A.B. & Bosker, R.J. 1999: *Multilevel analysis, an introduction to basic and advanced multilevel modeling*. London: SAGE Publications.
- SOU 2007:101: *Tydlig och öppen. Förslag till ny skolinspektion* (Dnr U2007/7907/G) Stockholm: Utbildningsdepartementet
- Teddlie, C. & Reynolds, D. 2001: Countering the critics: Responses to recent criticism of school effectiveness research. *School Effectiveness and School Improvement*, 12(1), 41–82.

- Tholin, J. 2003: *En roligare dans? Svenska skolors första tolkning av innebörden i lokala betygskriterier i tre ämnen för skolår åtta*. Borås: Högskolan i Borås, Institutionen för pedagogik.
- Thomas, S. 2001: Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement*, 12(3), 285–322.
- Thomas, S., Sammons, P., Mortimore, P. & Smees, R. 1997: Stability and consistency in secondary schools' effects on students' GCSE outcomes over three years. *School Effectiveness and School Improvement*, 8(2), 169–197.
- Thomas, S., Smees, R., MacBeath, J., Robertson, P. & Boyd, B. 2000: Valuing pupils views in Scottish schools. *Educational Research and Evaluation*, 6(4), 281–316.
- Vasquez Heilig, J. & Darling-Hammond, L. 2008: Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110.
- Visscher, A.J. 2001: Public school performance indicators: Problems and recommendations. *Studies in Educational Evaluation* 27(3), 199–214.
- Woolner, P., Hall, E., Higgins, S., McCaughey, C. & Wall, K. 2007: A sound foundation? What we know about the impact of environments on learning and the implications for Building Schools for the Future. *Oxford Review of Education*, 33(1), 47–70.
- Yang-Hansen, K., Rosén, M. & Gustafsson, J-E. 2006: Measures of self-reported reading resources, attitudes and activities based on latent variable modeling. *International Journal of Research & Method in Education*, 29(2), 221(2)237.