

## **Att mäta skolans bidrag till elevernas kunskapsutveckling**

Utbildningsnämnden gav i verksamhetsplanen för 2018 utbildningsförvaltningen i uppdrag att skapa en modell av skolors resultat genom en value added-modell. Denna modell ska visa hur skolorna bidrar till elevernas kunskapsutveckling, i syfte att förbättra förutsättningar för uppföljning av kunskapsresultat och måluppfyllelse. Uppdraget kom ursprungligen från kommunstyrelsen som specificerar att framtagandet av måttet inte får innebära ytterligare administration för den pedagogiska personalen.

Denna rapport redogör för ovanstående uppdrag.

Rapporten kommer inledningsvis att diskutera olika resultatmått i skolan. Skolbidragsmått – som mäter skolors bidrag till elevernas kunskapsutveckling – kommer särskilt att belysas. Ett value added-mått är en variant av skolbidragsmått. En annan variant är tvärsnittsmått, som bygger på elevernas bakgrundsfaktorer. Metodöverbäganden är ständigt närvarande för skolbidragsmått, för att skapa mått som på ett bra sätt visar skolans bidrag till elevernas lärande. De metodutmaningar som finns kommer att belysas, och främst kommer vi att koncentrera oss på dessa frågor för value added-mått.

Därefter kommer vi att rikta blicken mot skolbidragsmått i praktiken. Vi tittar på några intressanta mått som finns i andra länder. Sedan redogör vi för vilka diskussioner som förs i Sverige och vilka olika mått som används i Stockholms stad och nationellt.

Avslutningsvis diskuteras, med bakgrund av diskussionen i rapporten, hur man kan och bör använda skolbidragsmått, och då framförallt value added-mått.

De rekommendationer som vi ger är följande:

- Att fortsätta använda de skolbidragsmått som redan används i Stockholms stad. För närvarande finns två value added-mått. För gymnasieskolan kallas måttet förädlingsvärde och detta mått utgörs av den uträknade differensen mellan den faktiska andel elever som tar gymnasieexamen och den andel elever som kan förväntas göra det utifrån avgångsbetygen i årskurs 9. För grundskolan finns ett mått som mäter betygsutvecklingen på en skola mellan årskurs 7 och 9 för de elever som gått på skolan hela perioden. För grundskolan finns även ett mått som kallas Stockholmssalsa, vilket är ett tvärsnittsmått där betygen i årskurs 9 värderas utifrån vilka socioekonomiska bakgrundsfaktorer eleverna på skolan har.
- De value added-mått som idag används i Stockholms stad bedöms vara av den kvaliteten att de på ett godtagbart sätt belyser skolornas bidrag till elevernas kunskapsutveckling. De är därför ett användbart instrument för att bedöma och diskutera skolutveckling och fungerar därför som stöd i skolornas kvalitets- och utvecklingsarbete.
- Att det för närvarande saknas förutsättningar för fram fler value added-mått. De nationella proven är utifrån ett metodperspektiv idag inte lämpliga för att skapa ett value added-mått. Förändringar i utformningen av nationella prov krävs för att de skulle fungera som ett tillförlitligt value added-mått. De andra alternativ som övervägts skulle innebära mer administration för den pedagogiska personalen och bedöms därför inte vara aktuellt. Däremot finns det anledning att se över hur de existerande måtten konstrueras för att eventuellt förbättra träffsäkerheten.
- Att det ständigt sker metodöverväganden i framtagandet i value added-mått, vilket påverkar tillförlitligheten i måttet, innebär att de bör diskuteras och analyseras tillsammans med andra tillgängliga resultatmått och andra mer kvalitativa bedömningar av skolkvalitet, i syfte att ge en mer heltäckande bild.
- Att nuvarande skolbidragsmått inte är lämpliga att använda som ett underlag för skolval. De är metodkänsliga och kan vara svårtolkade, och de riskerar att driva på en betygsinflation samtidigt som det är osäkert i vilken utsträckning de beaktas i en skolvalssituation.

De existerande skolbidragsmåtten är inriktade mot grundskolans högre årskurser och gymnasieskolan. För de yngre årskurserna kan det nationella bedömningsstödet i årskurs 1–3 vara en källa som man kan konstruera ett skolbidragsmått. Det kräver dock att det

digitaliseras. Även de nationella proven i lägre årskurser skulle kunna användas i syftet om de rättas centralt och med vissa förändringar i upplägget av frågor. Men det behöver i så fall ytterligare utredning om ovanstående förutsättningar föreligger.

### **Skolbidragsmått – att mäta skolans bidrag till elevernas kunskapsutveckling**

Ett skolbidragsmått är ett mått på hur utbildningen på en skola bidrar till elevernas kunskapsutveckling och resultat. För att konstruera ett skolbidragsmått behövs någon form av bedömning av elevernas kunskapsnivåer, och om man ska jämföra skolor behövs ett enhetligt mått. I Sverige har vi två sådana mått som vi kan utgå ifrån: betygen och de nationella proven.

Vanligt är att resultaten på skolnivå redovisar alla elevers resultat. Hädanefter benämner vi dessa totalresultat, och beskriver de faktiska resultaten och tar inte någon hänsyn till elevernas förutsättningar och delas inte upp utifrån olika elevgrupper på skolan.

Skolornas bidrag till elevernas kunskapsutveckling fångas inte nödvändigtvis bäst genom att titta på totalresultaten. Det är inte alls säkert att de skolor som har de högsta totalresultaten bidrar i störst utsträckning till elevernas lärande, och omvänt att skolor med lägst totalresultat bidrar i minst utsträckning till elevernas lärande. En anledning är att skolorna inte har identiska elevsammansättningar utifrån elevernas förutsättningar och förkunskaper.

En variant av totalresultaten är att redovisa skolresultat där man exkluderar vissa elevgruppers resultat eller redovisar resultaten nedbrutet på olika kategorier av elever. Skolverket redovisar till exempel skolresultat där man exkluderat nyinvandrade elever (elever som varit mellan 0 och 4 år i Sverige) och elever med okänd bakgrund.<sup>1</sup> Anledningen till att använda resultatet exklusive nyanlända elever och elever med okänd bakgrund är att den gruppen av elever normalt sett har låga betygsresultat, vilket i mycket stor utsträckning förklaras av att de har bott en kort tid i Sverige och därmed inte behärskar undervisningsspråket svenska samt att elevgruppen är ojämnt fördelade på olika skolor. Skolverket redovisar också skolresultaten för vissa typer av elevgrupper: utifrån kön, vårdnadshavarnas utbildningsnivå samt svensk och

utländsk bakgrund. Sådana redovisningar kan underlätta jämförelsen mellan skolor och ger en viss men ändå begränsad information om skolans bidrag till elevernas lärande. Dessa mått är därför inte att betrakta som skolbidragsmått.

Om man i högre utsträckning ska bedöma det som kan benämnas *skolbidrag* behöver man i ännu högre grad i beräkningarna minska eller ta bort betydelsen av elevsammansättningen på skolan. Grunden i ett skolbidragsmått är att försöka åstadkomma en hypotetisk situation där alla skolor har en genomsnittlig elevsammansättning, utifrån de bakgrundsfaktorer som eleverna har och som man väger in och kontrollerar för. De bakgrundsfaktorena är vanligtvis socioekonomiska förutsättningar och migrationsbakgrund samt elevförutsättningar som inte kan kopplas till skolans arbete, såsom talang och motivation för studier.<sup>2</sup> Därmed kan man också utifrån skolbidragsmättet bedöma hur mycket skolorna bidrar till elevernas kunskapsutveckling.

Med andra ord skapar man ett mått som möjliggör en direkt jämförelse mellan skolor som har olika elevsammansättningar. Måttet som skapas kan antingen bygga på ett förväntat värde utifrån elevsammansättningen eller på en direkt jämförelse av kunskapsutvecklingen.

#### Tvärsnittsmått och value added-mått

Det finns två varianter av skolbidragsmått. Den första är ett tvärsnittsmått och bygger på ett antal bakgrundsfaktorer (till exempel socioekonomiska och migrationsbakgrund) som statistiskt visat sig ha betydelse för elevernas prestation i skolan. Skolans totalresultat mäts här en gång och relateras till bakgrundsfaktorer hos eleverna. Jämförelsen mellan skolorna görs sedan utifrån det förväntade värdet och hur det förhåller sig till totalresultaten.

Den andra varianten är ett value added-mått (eller förädlingsvärde) där man mäter resultaten vid två olika tidpunkter och ser utvecklingen mellan dessa två mättillfällen. Ett value added-mått kan också innehålla bakgrundsfaktorer. Här gör man en direkt jämförelse mellan skolor av hur utvecklingen har varit mellan de två olika mätningarna.

---

<sup>2</sup> Dessa faktorer kan påverkas av skolans arbete, t.ex. att få eleverna mer motiverade för sina studier, men tanken är här att eleverna kan besitta olika sådana förutsättningar oberoende av vad skolan gör.

Det har i en rad studier konstaterats att på en aggregerad nivå har elevernas socioekonomiska bakgrund en tydlig påverkan på elevernas kunskapsresultat (se till exempel Skolverket 2018). Framförallt har utbildningsnivån på elevernas vårdnadshavare och elevernas migrationsbakgrund en tydlig strukturell påverkan. Elever vars vårdnadshavare har en hög utbildningsnivå har högre kunskapsresultat jämfört med elever vars vårdnadshavare har en låg utbildningsnivå. Likaså har elever med en uppväxtbakgrund i Sverige högre resultat än elever som vuxit upp i ett annat land, eftersom var man växer upp påverkar elevens kunskaper i undervisningsspråket.

När man konstruerar ett mått utifrån de socioekonomiska förutsättningarna och elevernas migrationsbakgrund får man en uppskattning av vad skolan bör prestera utifrån ett övergripande mönster. En skola med en hög andel elever med högutbildade vårdnadshavare förväntas till exempel prestera bättre än en skola med en låg andel högutbildade vårdnadshavare. I vilken omfattning det finns skillnader i det förväntade värdet beror på hur elevsammansättningen ser ut, men också på hur det övergripande mönstret ser ut. Det övergripande mönstret fastställs genom en regressionsanalys där man utifrån ett större datamaterial, nationell nivå eller ett större antal skolor, fastställer vilka bakgrundsfaktorer och i vilken utsträckning de påverkar resultaten.

Fördelen med ett value added-mått jämfört med ett tvärsnittsmått är att genom att använda tidigare elevprestationer så skapas en utgångspunkt som är elevspecifik. I ett tvärsnittsmått är utgångspunkten inte elevspecifik, utan strukturell, det vill säga att den grundar sig på ett genomsnittligt värde för den grupp av elever som eleven tillhör utifrån de strukturella variabler som används (till exempel flicka, högutbildade vårdnadshavare och född i Sverige). Man kan uttrycka det som att ett value added-mått ger en bättre utgångspunkt för att beskriva undervisningens bidrag till elevernas prestationer. Ett value added-mått behöver därför innehålla två mätpunkter av elevernas kunskap (till exempel betyg eller provresultat), och är således en uppskattning av skolans bidrag till elevens lärande under den tid som ligger mellan de två mätpunkterna. Det är också fullt möjligt att i ett value added-mått inkludera elevernas socioekonomiska förutsättningar.

### **Allmänna metodproblem för value added-mått**

Metod och metodproblem är en viktig fråga när det gäller skolbidragsmått, och det vi framförallt kommer att fokusera på är

value added-mått. De tre huvudsakliga metodproblem som behöver hanteras är:

- Vilka mått på kunskap, till exempel betyg och prov, som är lämpliga att använda.
- Måttets tillförlitlighet på skolenheter med få elever.
- Golv- och takeffekter, som kan uppstå för vissa skolenheter utifrån ett lågt eller högt utgångsläge, och som ger olika möjligheter till förändring.

### Vilka kunskapsmått kan användas?

Den första utmaningen för att skapa ett value added-mått är att det finns lämpliga kunskapsmått att mäta utvecklingen ifrån. Ett lämpligt mått behöver mäta kunskaperna på ett liknande sätt över tid och kunskapen behöver också mätas på ett liknande sätt mellan olika skolor.

De huvudsakliga kunskapsmåten som finns i den svenska skolan är nationella prov och betyg. De nationella proven skrivs i årskurs 3, 6 och 9<sup>3</sup>, och betyg ges från årskurs 6 till 9. Det är också dessa mått som står till förfogande i det här sammanhanget. Ett annat alternativ skulle kunna vara att bygga value added-måttet på ett för syftet särskilt utformat prov. I och med att uppdraget specificerat att det inte får innebära ytterligare administration för den pedagogiska personalen så har det alternativet inte varit aktuellt att utreda.

### Nationella prov

När det gäller nationella prov är de som grund för ett value added-mått i dagsläget behäftade med en del problem metodmässigt. Det som lyftes fram i utredningen om nationella prov (SOU 2016:25) är att de inte mäter konsistent över tid då svårighetsgraden på proven kan variera. Det skulle dock kunna hanteras genom att man normaliserar resultaten, vilket innebär att man gör ett genomsnitt för varje provtillfälle som skolornas resultat relateras till. Det blir då ingen exakt bedömning av hur kunskaperna har ökat, men i en jämförelse mellan skolor så kan man se vilka skolor som ligger över eller under den genomsnittliga utvecklingen. Det är också så här de exempel på value added-mått som vi belyser i rapporten mäter, nämligen hur skolorna presterar i förhållande till varandra. Det innebär i praktiken att det kommer att vara en blandning av positiva och negativa värden. Teoretiskt är det möjligt att alla skolor hamnar på samma värde, om alla har samma utveckling.

---

<sup>3</sup> I årskurs 3 är ämnena matematik och svenska/svenska som andraspråk, i årskurs 6 är det ämnena engelska, matematik och svenska/svenska som andraspråk och i årskurs 9 är det ämnena engelska, matematik, svenska/svenska som andraspråk samt ett naturvetenskapligt och ett samhällsvetenskapligt ämne.

Mer problematiskt för de nationella proven är att de i dagsläget är något begränsade i vilka kunskapsområden de omfattar och att innehållet kan variera något mellan olika år. Det kan då slå igenom på resultaten om undervisningen varit inriktad på ett visst område vid någon tidpunkt. För att de nationella proven ska fungera som ett bra mått på kunskap i detta sammanhang behöver de därför troligen omarbetas så att de omfattar tillräckligt mycket av ämnet i samtliga av de årskurser som man gör nationella prov. Eftersom man behöver två mätpunkter i ett value added-mått så kan det bli problematiskt om de nationella proven mäter olika delar av ämnet i olika årskurser. Jämförelsen kommer då att halta något och resultaten kan bli en effekt av skolorna har inriktat sin undervisning olika vid olika tidpunkter.

Ett annat metodproblem med att använda nationella prov som resultatmått i Sverige är att de i dagsläget rättas lokalt på skolan. Om det finns variationer i bedömningen av de nationella proven på olika skolor så påverkar det potentiellt sett tolkningen av resultatet, och framförallt om det finns variationer på skolan mellan olika lärare över tid. Om man till exempel vid ett prov gör en restriktiv bedömning och vid nästa provtillfälle gör en mer generös tolkning så blir skolbidragsfaktorn övervärderad. Eller i den omvända situationen att skolbidragsfaktorn blir undervärderad. För att de nationella proven skulle kunna fungera i det här sammanhanget skulle det behövas någon form av central och gemensam bedömning. Även andra lokala skolskillnader kan påverka resultatet på de nationella proven. Det kan finnas variationer i förberedelser till de nationella proven mellan skolorna, vilket kan påverka resultatet. Och den här problematiken gäller alla former av tester och prov som kan ligga till grund för ett value added-mått.

Sammantaget gör vi därför bedömningen att i dagsläget finns det metodproblem med de nationella proven som gör att de inte är lämplig att använda som ett mått på kunskap i ett value added-mått.

### **Betyg**

Ett alternativ är att använda betygen i de olika ämnena som grund för måttet, antingen för sig eller i ett samlat mått som till exempel för grundskolan genomsnittligt meritvärde och andel behöriga till yrkesprogram, samt för gymnasieskolan genomsnittlig betygspoäng och andel elever med gymnasieexamen. Det ger en bredare bedömning av elevens kunskaper och blir inte påverkbart av skillnader i förberedelser. Betygen är dock precis som de nationella proven bedömda lokalt på skolan, och det är ett problem i det här

sammanhanget, eftersom skillnader i bedömning och betygssättning mellan skolor kan påverka resultatet. Men det är så i något mindre utsträckning jämfört med de nationella proven, eftersom betygen sätts i flera ämnen och det därmed blir ett större antal lärare som är betygssättande. Därmed minskar möjligheten för påverkan av individuella lärares nivå i betygssättningen. Men om det finns skillnader i betygssättningsnivå generellt sett mellan skolor kan det påverka värdet, framförallt om mätpunkterna finns i två olika skolformer eller om bedömningen varierar över tid på en skola. Om man till exempel vid det första mättillfället har en restriktiv betygssättning och vid det andra mättillfället har en generös betygssättning så kommer bidraget till elevernas kunskapsutveckling att övervärderas. Om den omvända situationen råder kommer resultatet att undervärderas. Detta metodproblem förstärks om det är olika lärare som sätter betygen vid de olika mättillfällena, och uppstår främst om mätpunkterna ligger på olika stadier, skolformer eller om det på skolan är en hög läraromsättning.

Men även om det finns metodproblem med betygen så anser vi att betygen är tillräckligt stabila och fungerar därför som ett mått på kunskap i ett value added-mått.

Eftersom ett value added-mått alltid är känsligt för metodproblem i kunskapsmättet så konstaterar flera bedömare att det är viktigt, framförallt om man ska jämföra skolor, att redovisningen sker så att osäkerheten i beräkningarna redovisas (se American Statistical Association 2014, Ds 2001:24, s. 44 och Steffensen m.fl. 2017, s. 85). Det görs vanligen genom att man redovisar resultaten med en konfidensintervall, vilket innebär att man inte ger ett exakt värde utan en uppskattning utifrån en statistik beräkning att utfallet ligger inom ett visst intervall. Det innebär att istället för att säga att en viss skola har värde 5 så redovisas resultat att det ligger inom ett visst intervall, till exempel 4 till 6. Därmed tar man hänsyn till den statistiska osäkerheten i beräkningarna. Att redovisa resultaten på ett sådant sätt gör dock måttet mer komplicerat att tolka, vilket kan tala emot att redovisa med en konfidensintervall. Ett annat sätt att hantera den statistiska osäkerheten är att använda flera års resultat och därmed få större stabilitet i måttet.

### Skolenheter med få elever

En annan utmaning är om skolorna och skolenheterna har få elever, vilket påverkar den statistiska tillförlitligheten. Det kan hanteras genom att man tar bort de enheter där den statistiska tillförlitligheten bedöms som för låg. Nackdelen blir då att man inte kan jämföra alla skolor. Ett annat sätt att hantera det är att slå ihop



flera års resultat och därigenom öka antalet elever som ingår i undersökningen. Nackdelen med att slå ihop flera års data är att modellen i mindre utsträckning kan åskådliggöra förändringar i skolornas mått.

Ett relaterat metodproblem är om skolan har en hög elevomsättning. Eftersom ett value added-mått mäter kunskaperna vid två tillfällen så bör man enbart mäta kunskaperna för de elever som gått på skolan vid båda dessa tillfällen. De elever som lämnat skolan eller kommit till skolan under denna period bör inte inkluderas då resultaten kan bli missvisande, eftersom man då inte fångar skolans bidrag till elevernas kunskapsutveckling. Vid en hög elevomsättning kan det precis som i fallet med få elever på en skola bli en för låg statistisk tillförlitlighet.

### Golv- och takeffekter

Ytterligare ett metodproblem är så kallade golv- och takeffekter, och relaterar till vilket utrymme det finns till förbättring för skolorna. Det innebär att det finns olika, åtminstone teoretiska, möjligheter för skolor som har låga respektive höga skolresultat. Takeffekten innebär att vissa skolor har så pass höga initiala resultat att det finns små möjligheter till en förbättring av det förväntade resultatet. Golfeffekten innebär att skolor med så pass låga initiala resultat att det finns ett begränsat utrymme att försämra resultaten. Dessutom om det finns stora skillnader i möjligheterna till förbättring mellan skolorna så blir det svårare att jämföra skolornas utveckling. Hur stora golv- och takeffekterna blir beror på vilka kunskapsmått man använder, och hur stort spelrum det finns för förbättringar och försämringar.

### Sammanfattning metodproblem

Sammantaget kan konstatera att det inte finns något perfekt mått, utan de påverkas av mätproblem för måtten på kunskap och faktiska omständigheter såsom att det finns skolenheter med få elever. Vi kommer att återkomma till dessa metodproblem nedan när vi redogör för existerande value added-mått och beskriva hur man hanterat dem. Därför bör man alltid vara försiktig med att dra alltför långtgående slutsatser utifrån ett value added-mått. Denna rekommendation gäller för övrigt alla resultatmått i skolan, om man ska använda dessa för att göra en bedömning av skolans prestation. Det är att föredra att använda olika mått, samtliga med måtta, för att ge en mer heltäckande bild. Därför blir rekommendationen att skolbidragsmått bör relateras till en skolas totalresultat och redovisas i anslutning till dessa.

## Value added-mått i andra länder

Det finns i varierande utsträckning value added i andra länder, och nedan kommer vi att redogöra för några exempel.

### Norge

I Norge gjorde den norska statistiska centralbyrån 2017 en nationell studie om olika skolors och kommuner bidrag till elevernas resultat. Anledningen var att man ville komplettera totalresultaten med ett value added-mått som ger bättre information om skolors och kommuners bidrag till elevernas lärande. Nedan redovisas hur måttet är konstruerat.

Som kunskapsmått används de nationella proven i årskurs 5 och 8 samt examensprovet i årskurs 10, i ämnena norska, engelska och matematik.<sup>4</sup> Utifrån dem får alla elever ett genomsnittligt värde baserat på proven i de tre ämnena, så här används ett normaliserat kunskapsmått.

De nationella proven i Norge lider i sammanhanget delvis av samma metodproblematik som de svenska nationella proven och som diskuterades tidigare. De mäter en begränsad del av elevernas färdigheter och kunskaper i de aktuella ämnena. I rapporten lyfts också fram att skillnader mellan skolorna i förberedelser till de nationella proven kan påverka jämförbarheten mellan skolor.

En metodmässig fördel med de norska nationella proven jämfört med de svenska är att de är elektroniska och rättas automatiskt, med undantag av några uppgifter i läsning. Därmed blir de inte sårbara för eventuella variationer i betygsättning mellan skolorna.

I årskurs 10 sätts betyget på examensprovet av en extern bedömare. Alla elever får samma examensprov och eftersom det rättas externt har det fördelen av att man kan bortse från eventuella skillnader i betygsättning mellan skolor. Examensprovet skrivs under en dag och kan då vara känsligt för om elever under- eller överpresterar på provet i förhållande till deras nivå över en längre period. Men man menar att sådana dåliga eller bra dagar borde vara en variation som ser ungefär likadant ut på alla skolor, men att det potentiellt sett får större utslag på små skolor jämfört med stora skolor (Steffensen m.fl. 2017, s. 15).

I den norska studien aktualiseras metodproblematiken att det finns små enheter/få elever. Därför har de också analyserat två års resultat

sammanslaget för att kontrollera för tillfälliga årsvisa variationer. Eftersom resultaten normaliseras så påverkar inte eventuella skillnader mellan svårighetsgraden i provet för olika år och mellan olika ämnen utfallet. Även golv- och takeffekten diskuteras och man kommer till slutsatsen att det inte finns särskilt starka golv- och takeffekter (Steffensen m.fl. 2017, s. 30).

I modellen ingår även bakgrundsfaktorer såsom elevernas socioekonomiska bakgrund och migrationsbakgrund. De har dock i sin analys observerat att socioekonomisk bakgrund och migrationsbakgrund inte bidrar särskilt mycket till förklaringsmodellen när man jämför utvecklingen gentemot tidigare elevresultat.

I årskurs 5 är måttet ett tvärsnittsmått där skolorna får ett förväntat värde utifrån elevernas socioekonomiska bakgrund och migrationsbakgrund. I årskurs 8 jämförs resultaten på de nationella proven med elevernas resultat i årskurs 5 och utvecklingen kan bedömas. Nu blir måttet ett value added-mått eftersom det finns två mätpunkter. I årskurs 10 används examensbetygen och jämförs med de nationella proven i årskurs 8. I tabellen nedan sammanfattas de olika källorna.

	<b>Måttillfälle 1</b>	<b>Måttillfälle 2</b>
<b>Årskurs 5</b> (tvärsnittsmått)	Nationella prov utifrån elevernas bakgrundsfaktorer	Inget
<b>Årskurs 8</b>	Nationella prov i årskurs 5	Nationella prov i årskurs 8
<b>Årskurs 10</b>	Nationella prov i årskurs 8	Examensprov i årskurs 10

### England

I England tas ett value-added mått fram för varje skola. Det finns tre olika mått, ett för primary education (elever från 5 till 11 år), ett för secondary education (elever från 11 till 16 år) och sedan för further education (elever från 16 till 18 år).

Beräkningarna sker utifrån olika källor och graderingssystemet varierar, men för varje nivå i skolsystemet finns ett value added-mått. I tabellen nedan sammanfattas vilka källor som man använder.

	<b>Måttillfälle 1</b>	<b>Måttillfälle 2</b>
<b>Primary education</b>	Standardiserade nationella prov (när eleverna är 7 år), läsning, skrivning och matematik.	Standardiserade nationella prov (när eleverna är 11 år), läsning, skrivning och matematik.
<b>Secondary education</b>	Standardiserade nationella prov (när eleverna är 11 år), läsning, skrivning och matematik.	Progress 8 (ett mått som beräknas på betygen i åtta ämnen, engelska och matematik räknas dubbelt) när secondary education avslutas.
<b>Further education</b>	Progress 8 (ett mått som beräknas på betygen i åtta ämnen, engelska och matematik räknas dubbelt) när secondary education avslutas.	Betygen på prov i de huvudämnena (A-levels) man läst (vanligen tre till fyra stycken) när further education avslutas.

Måtten ser något olika ut för de olika nivåerna, men det centrala är att de är normaliserade, det vill säga de förhåller sig till genomsnittet av alla elever i England. Precis som i det norska exemplet visar måtten hur en skola förhåller sig till genomsnittet, och visar på om de har en bättre eller sämre utveckling än genomsnittsskolan.

Måttet som används i England har två syften. Det första är att ur ett kontroll- och stödperspektiv identifiera skolor som har en svag prestation. De skolor som har ett negativt värde under en viss gräns bevakas av The Office for Standards in Education, Children's Services and Skills (Ofsted). Ofsteds huvudsakliga ansvar är inspektioner av skolor, men de kan också ge särskilda insatser till skolor och deras befogenheter är relativt stora. Ett annat syfte med måtten är att de ska vara till hjälp då elever gör sitt skolval (<https://www.compare-school-performance.service.gov.uk/>).

Metodproblemet små skolenheter diskuteras i instruktionerna till hur man ska tolka. Vid mycket få elever så publiceras inte data, men de nämner också att man måste ta hänsyn till om elevantalet är litet i tolkningen av siffrorna. Även golv- och takeffekterna tas till viss

del upp, eftersom de för varje skola kategoriserar eleverna i tre grupper utifrån viken kunskapsnivå de presterar i utgångsläget, vilket möjliggör att se utvecklingen för de olika grupperna på en skola. De skriver också att man ska ta hänsyn till både totalresultaten och skolbidragsmättet när man ska bedöma en skola.

## USA

Även i USA används i hög grad value added-mått för att mäta skolors bidrag till elevers kunskapsutveckling. En viktig skillnad jämfört med de mått som används i England handlar om analysnivån. Value added-mått användes från allra första början för att uppskatta just skolans bidrag, men metoden kan anpassas så att den mäter effekterna på elevers kunskapsutveckling som andra nivåer inom utbildningssystemet bidrar med. I olika nationella och politiska sammanhang varierar den analysenhet som betonas. I många delstater i USA ses lärarna som den viktigaste analysenheten. Detta har skapat en efterfrågan på value-added-mått som mäter just *lärareffekt*, snarare än den aggregerade skoleffekten (Harris, 2009). Även i amerikansk forskning om value-added-mått läggs relativt liten tonvikt på skolnivåns bidrag till elevernas kunskapsutveckling, eftersom den rådande uppfattningen inom detta forskningsområde är att lärareffekten är avsevärt starkare än skoleffekten (Muijs m.fl., 2014).

## Vad händer i Sverige?

Skolkommisionen föreslog i sitt slutbetänkande, Samling för skolan (SOU 2017: 35) bland annat ett sammanhållet uppföljnings- och utvärderingssystem. Bakgrunden till förslaget var att kommissionen såg att det behövdes mer tillförlitlig resultatinformation om elevernas kunskapsutveckling än den som ges i dag i form av betyg och nationella prov. Dessutom identifierade Skolkommisionen att det idag exempelvis saknas data om resultatutveckling för elever som får någon typ av stöd eller extra anpassning (SOU 2017: 35, s. 311).

Flera offentliga utredningar har adresserat frågan om mått som ger information om undervisningens bidrag till elevernas lärande. Skolkommisionen (SOU 2017:35, s. 88) betonar att det behövs för att följa utvecklingen på nationell nivå och på skolnivå som stöd för skolans utvecklingsarbete. Samtidigt understryker kommissionen att ett sådant system, som utgår från value added-mått, kräver ett omfattande utrednings- och metodutvecklingsarbete som behöver prioriteras (SOU 2017:35,s. 132).

Utgångspunkten i utredningarna har till stor del varit kring de nationella proven som källa till ett value added-mått. I utredningen om nationella prov (SOU 2016:25, 239) beskrivs synpunkter som lyfts fram från tidigare utredningar, vilka dels framhåller att betyg och nationella prov inte kan anses vara tillräckliga för att följa elevernas kunskapsutveckling över tid, samt dels att det finns ett behov av ett sådant instrument. Utredningen om förbättrade resultat i grundskolan (SOU 2014:12), konstaterade att det nationella provsystemet idag inte fungerar till att använda för trendmätning. För att det ska fungera behöver proven uppfylla följande:

- Proven är stabila över tid och de innehåller ett antal uppgifter som har ingått i tidigare prov, så kallade ankaruppgifter.
- Antalet uppgifter räcker för att täcka in tillräckligt mycket av ämnets eller kursens innehåll.
- Innehållet i kurs- och ämnesplaner och utformningen av betygssystemet i huvudsak är stabila över tid, så att det är samma kunskaper och kunskapsnivåer som jämförs.

Skolverket har tidigare haft ett regeringsuppdrag om att utreda möjligheterna att ta fram ett value added-mått. I uppdraget vill man att Skolverket ska presentera statistik så att olika skolors faktiska resultat och kvalitet kan jämföras med varandra (Utbildningsdepartementet 2012). Det står bl.a. att:

Skolverket bör också utveckla mått på skolors kvalitet som underlättar möjligheten att upptäcka vilka skolor som avviker från riksgenomsnittet – positivt eller negativt – så att bra insatser kan spridas och uppmuntras, och dåliga insatser avbrytas. Det kan ske genom att provresultaten på en skola jämförs med de resultat som eleverna uppnådde i de nationella ämnesproven under tidigare årskurser, så kallad value added. Jämförelser kan på så sätt göras av enskilda skolors bidrag till elevernas kunskapsutveckling mellan två olika provtillfällen [...]  
Regeringen anser att det ska vara enkelt att få tillgång till information som underlättar möjligheten för bl.a. föräldrar och elever att välja skola.

I sitt svar på uppdraget svarar Skolverket att de påbörjat ett arbete om att skapa ett value added-mått och att de under 2014 avser att fortsätta utvecklingsarbetet. I ett senare PM beskriver Skolverket (2015) hur arbetet har fortlöpt och synpunkter från forskare har också inhämtats. Där skriver man:

Under arbetet med att ta fram ett value-added mått [sic] har ett flertal problem påtalats, dels från forskarna, dels internt. Problemen är beroende av att det inte finns utfallsmått som är

lämpade för uppgiften. Detta gör det direkt olämpligt att ta fram och publicera ett value-added mått för den svenska skolan [...] Skolverkets bedömning är att avråda användningen av ett value-added mått i dess nuvarande utformning i syfte att ranka skolor för ”välja skola” -siten.

Det bör påpekas att Skolverket inte diskuterar det som lyftes fram inledningsvis i regeringsuppdraget, att ett nationellt value added-mått kan användas för skolutvecklingsfrågor, utan enbart som ett mått i informationen att välja skola. Det skäl till att Skolverket anser att måttet är olämpligt är framförallt att det skulle kunna leda till en betygs- och provinflation så länge det inte finns sådana som rättas centralt, eftersom måttet skulle kunna användas som konkurrensmedel. Därmed skulle också legitimiteten för dessa betyg- och provresultat kunna undergrävas. Andra metodproblem som lyfts är att det inte är ovanligt att skolenheter i Sverige är små och det skulle därmed bli en betydande osäkerhet i beräkningarna. Skolverket oroar sig också för golv- och takeffekter eftersom betyg och prov inte är utformade för att göra value added-mått. Skolor med höga och låga totalresultat har olika förutsättningar för utveckling av resultaten, då skolorna med låga resultat har mindre utrymme för försämring och skolorna med höga resultat har mindre utrymme för förbättring. Dessa metodproblem ska vi diskutera nedan.

### Skolbidragsmått i Sverige

Ett tvärsnittsmått där man ställer skolornas totalresultat i relation till ett antal bakgrundsfaktorer hos eleverna är Skolverkets SALSA (Skolverkets Arbetsverktyg för Lokala SambandsAnalyser), och är således ett tvärsnittsmått. De bakgrundsfaktorer som används är fördelningen flickor/pojkar, vårdnadshavarnas utbildningsnivå samt andel nyinvandrade elever. SALSA är en statistisk modell som tar hänsyn till ovanstående bakgrundsfaktorer och utifrån hur alla skolor i landet presterar så får en specifik skola ett modellberäknat (förväntat) betygsresultat som beskriver på vilken förväntad nivå deras prestation bör vara utifrån de strukturella förutsättningarna för skolans elever. Detta modellberäknade betygsresultat jämförs därefter med totalresultatet, och man kan se om skolan presterar bättre eller sämre än vad som är förväntat.

Den fristående skolkoncernen AcadeMedia har för sina gymnasieskolor ett value added-mått som de benämner förädlingsvärde. Det utgår ifrån elevernas meritvärde i årskurs 9 och utifrån det beräknas en förväntad genomsnittlig betygspoäng i

gymnasieskolan. Differensen mellan det faktiska och förväntade gymnasiebetyget ger förädlingsvärdet på en skola.

### **Skolbidragsmått specifika för Stockholms stad**

#### **Stockholmssalsa**

I Stockholm finns ett motsvarande tvärsnittsmått som Skolverkets SALSA, vanligen benämnd som Stockholmssalsa. Utgångspunkten i Stockholmssalsa är ett socioekonomiskt index, som fördelningen av det socioekonomiska tilläggsanslaget<sup>5</sup> baseras på. Indexet bygger i likhet med Skolverkets Salsa på en statistisk modell om vilka strukturella faktorer som påverkar skolresultatet. De faktorer som finns med i det socioekonomiska indexet är följande, i redovisad styrkeordning.

1. Nyinvandrad till Sverige, graderad efter hur lång tid eleven varit i Sverige (upp till 7 år)
2. Vårdnadshavarnas utbildningsbakgrund
3. Om vårdnadshavarna uppbär ekonomiskt bistånd
4. Om eleven bor med båda, en eller ingen vårdnadshavare

I indexet ingår även två strukturvariabler. Den första är den totala konstellationen av elever på en skola utifrån socioekonomiska faktorer, och benämns skolmiljö. Om det är en hög andel elever med mindre gynnsam socioekonomisk bakgrund så får skolan således en socioekonomisk tyngd som blir större än den varit om man bara utgår de individuella eleverna. Den andra strukturvariabeln är boendemiljö (geografisk) och innebär att den totala socioekonomiska sammansättningen i en boendemiljö för eleverna på en skola påverkar skolans index.

Det socioekonomiska indexet avspeglar därmed sannolikheten för eleverna på en skola att uppnå olika betygsresultat och påminner om Skolverkets SALSA, men tar fler faktorer i beaktande. Precis som för Skolverkets SALSA får varje skola ett förväntat betygsresultat utifrån elevsammansättningen, och det värdet jämförs med totalresultatet.

När man jämför hur Stockholms grundskolor presterar i SALSA och Stockholmssalsa så är det inte särskilt stora skillnader på skolnivå. Vanligen över- och underpresterar en skola i båda måtten,

---

<sup>5</sup> Det socioekonomiska tilläggsanslaget innebär att skolor där eleverna har mindre gynnsamma socioekonomiska förutsättningar får en högre ekonomisk resurstilldelning jämfört med skolor där eleverna har mer gynnsamma socioekonomiska förutsättningar.



men det finns några enstaka undantag där skolor kan överprestera i det ena måttet och underprestera i det andra.

### Jämförelse av betygsresultat i årskurs 7 och 9

Ett annat mått, som är ett value added-mått, som utbildningsförvaltningen i Stockholm för första gången till 2017 har utvecklat är en betygsjämförelse i grundskolan mellan årskurs 7 och årskurs 9. Utgångspunkten är en skolas genomsnittliga meritvärde i årskurs 7, vilket sedan jämförs med det genomsnittliga meritvärdet i årskurs 9. Endast de elever som har betyg från samma skola i årskurs 7, 8 och 9 ingår i underlaget. Anledningen till att studera de som gått på skolan hela högstadieperioden är att man tydligare kan isolera den effekt utbildningen haft på eleverna. Elevomsättningen varierar kraftigt på olika skolor i Stockholms stad och därför blir inte alltid totalresultaten en jämförbar av avspiegling av hur stor påverkan undervisningen haft på elevernas kunskaper. Då måttet är nytt och så här långt bara är framräknat för ett år behövs data för ytterligare år för att testa robustheten i måttet.

Ett metodproblem i detta mått är att det på runt 20 procent av skolorna inte blir statistiskt tillförlitliga resultat eftersom det inte är tillräckligt många elever som går på skolan mellan årskurs 7 och 9. De är också vanligtvis skolor med mindre gynnsamma socioekonomiska förutsättningar. Detta problem skulle delvis kunna lösas med att slå ihop två års data, men det skulle fortfarande finnas skolor där resultatet inte skulle vara statistiskt tillförlitligt. Däremot på övergripande nivå är inte golv- och takeffekter särskilt besvärande. En bidragande orsak är att det är få skolor som har riktigt höga genomsnittliga meritvärden i årskurs 9. Samtidigt är det något man behöver ta hänsyn till i enstaka fall.

En variant av detta mått skulle kunna vara att utgå ifrån årskurs 6-betyget istället för årskurs 7-betyget. Vilket betygsmått man använder kan ge olika för- och nackdelar. I årskurs 7-betyget använder skolorna inte alltid hela betygsskalan, och det kan bli ett problem om skolorna gör på olika sätt. Med årskurs 6-betyget blir det fler skolors betygsättning som blir aktuella (eftersom det finns skolor som inte har högre årskurs än 6), och här kan det också bli problematiskt om det finns variationer i betygssättningen i årskurs 6. Men den avgörande skillnaden till att inte använda årskurs 6-betyget är att det med stor säkerhet skulle bli färre elever på en skola som går hela vägen till årskurs 9, vilket redan är ett metodproblem med detta mått.

### Förädlingsvärde gymnasieskolan

För gymnasieskolan har utbildningsförvaltningen ett liknande mått som det som huvudmannen AcadeMedia använder för sina gymnasieskolor, och som tidigare beskrivits. Stockholms stads mått har samma benämning, förädlingsvärde, och är också ett value added-mått. Varje nationellt gymnasieprogram på en skolenhet får ett uppskattat värde för hur många elever som når gymnasieexamen för de elever som finns på programmet/enheten i början på årskurs 3. Det uppskattade värdet är baserat på de meritvärden eleverna har i årskurs 9 på grundskolan och utifrån en statistik analys finns en beräknad sannolikhet för att en elev tar examen vid ett visst meritvärde. Det kan vara intressant att även titta på elevernas genomsnittliga betygspoäng som resultatvariabel för gymnasieskolan.

Ett metodproblem är att det beräknas på eleverna i början på årskurs 3, så de elever som gått på skolan i årskurs 1 och 2 men slutat ingår inte i materialet. Om det finns skillnader i programmen på skolenheterna i vilken utsträckning man har elever som gått kortare tid på skolan så blir det också en skillnad i den utsträckning utbildningen påverkat elevernas lärande. Det hade varit att föredra att utgå ifrån de elever som påbörjar enheten/programmet i årskurs 1, men problemet är då att det i flera fall blir få elever kvar som går ut årskurs 3 på samma program och enhet, vilket gör att måttet inte är statistiskt tillförlitligt. Även här skulle man kunna överväga att slå ihop två års resultat och låta måttet utgå ifrån årskurs 1. Men även när man använder utgångspunkten i årskurs 3 blir det ett antal program på enheterna där resultatet inte är statistiskt signifikant eftersom det grundas på för få elever. Ovanstående metodproblem medför att det är viktigt att se på genomströmningssiffror, det vill säga hur hög andel av eleverna som tagit examen inom tre eller fyra år från det att de påbörjade sina gymnasiestudier, som komplement till förädlingsvärdet.

I förädlingsvärdet för gymnasieskolorna i Stockholm finns besvärande golv- och takeffekter. Det finns program på gymnasieenheter som har mycket höga förväntade värden med små möjligheter till förbättring, samtidigt som det finns flera program på gymnasieenheter, med låga förväntade värden, som har ett positivt värde som vida överstiger det som är möjligt för skolor med de höga förväntade värdena. Grundorsaken till detta metodproblem är att antagningen till gymnasieskolan är kraftigt segregerad utifrån antagningspoäng, det vill säga att det finns många enheter med höga respektive låga antagningspoäng. Det innebär att det viktigt att alltid titta på förädlingsvärdet i kombination med totalresultaten.

Även kunskapsmåttan som används är en viss metodsvaghet. Förädlingsvärdet för gymnasieskolan grundas det på betygen i grundskolan och den andra mätningen sker utifrån gymnasiebetygen. Om gymnasieskolorna har olika nivåer för sin betygsättning så påverkar det value added-måttet. Till viss del finns det också ett problem i andra ändan, om det finns variationer i betygsättningen på grundskolor kan det påverka förädlingsvärdet om en gymnasieskola har en relativt hög andel av elever från skolor som antingen är generös eller restriktiv i sin betygsättning. Eftersom Stockholms stads gymnasieskolor vanligen har en sammansättning av elever som kommer från ett stort antal olika grundskolor så borde det i hög grad vara slumpmässigt fördelat på de olika gymnasieskolorna.

Sammantaget finns det anledning att se över förädlingsvärdet metodmässigt för att undersöka om det kan förbättras något. Samtidigt handlar det inte om att utveckla något nytt mått.

Sammanfattning tillgängliga resultatmått för Stockholms stad  
 Nedan sammanfattas de resultatmått som är tillgängliga, och de tre sistnämnda är specifika för Stockholms stad. Som redan diskuterats så mäter skolbidragsmåten i högre utsträckning jämfört med totalresultaten och de justerade resultaten skolans bidrag till elevernas lärande.

<b>Mått</b>	<b>Typ av mått</b>	<b>Beskrivning</b>
<b>Betygsresultat, samtliga elever</b>	Totalresultat	Nationell nivå, grund- och gymnasieskola
<b>Nationella prov, samtliga elever</b>	Totalresultat	Nationell nivå, grund- och gymnasieskola
<b>Betygsresultat exklusive vissa elevgrupper</b>	Justerade resultat	Nationell nivå, grundskola
<b>SALSA</b>	Tvårsnittsmått	Nationell nivå, skolbidragsmått, grundskola
<b>Stockholmssalsa</b>	Tvårsnittsmått	Stockholmsnivå, skolbidragsmått, grundskola
<b>Jämförelse betyg åk 7 och 9, grundskola</b>	Value added	Stockholmsnivå, skolbidragsmått, grundskola
<b>Förädlingsvärdet, gymnasieskola</b>	Value added	Stockholmsnivå, skolbidragsmått, gymnasieskola

Inget av de två value added-måtten som används innehåller elevernas bakgrundsfaktorer. I båda fallen skulle det vara möjligt, vilket skulle göra modellen något stabilare, men också mer komplicerad att räkna fram. Dessutom finns bakgrundsfaktorerna till viss del inbyggda i modellen, eftersom betygsresultaten som är utgångspunkten i de två måtten påverkas av bakgrundsfaktorerna eftersom det statistiskt är fastställt att elever med mer gynnsamma socioekonomiska förutsättningar har högre betygsresultat. Sammanfattningsvis, att ta med elevernas bakgrundsfaktorer skulle ge en något mer robust modell, men skulle också göra framräkningen mer komplicerad. Förädlingsvärdet har också tidigare tagit hänsyn till elevernas bakgrundsfaktorer, men på grund av svårigheter att få ut data från SCB så har det inte gjorts de senare

åren. Beräkningar visar också att det inte blir särskilt stora skillnader i modellen om man tar med elevernas bakgrund.

### **Hur kan och bör skolbidragsmått användas?**

En första fråga som behöver besvaras innan man kan ge rekommendationer om hur skolbidragsmått bör och kan användas är i vilken utsträckning ett skolbidragsmått avspeglar skolkvalitet eller skolans prestation på elevernas lärande. Svaret på den frågan kan i sin tur spåras tillbaka till de metodfrågor som diskuterades tidigare.

Ett skolbidragsmått kan användas för nedanstående syften:

- Som ett bidrag till en debatt om skolutveckling på enskilda skolor.
- Indikationer på hur ett skolsystem (till exempel på kommunal och nationell nivå) utvecklas.
- Som information vid skolval.
- Som underlag i det systematiska kvalitetsarbetet på skolan.

När det gäller tvärsnittsmått, så skriver Skolverket följande om SALSA: ”SALSA kan inte ensamt svara på frågan om en skola är bättre än en annan. Hur god kvalitet en skola har är en alltför sammansatt fråga för att fångas i ett enkelt mått, vare sig det är ett faktiskt eller modellberäknat värde.” (Skolverkets hemsida, 2018-04-05).

Istället ser Skolverket att SALSA tillsammans med uppgifter om förutsättningar, processer och resultat kan användas för nyanserade diskussioner om skolors resultat. För ett enstaka mått kan inte säga vilka faktorer som förklarar vad som ligger bakom att en skola presterar högt eller lågt. För att besvara frågan varför krävs ytterligare data och analyser.

När det gäller value added-mått så resonerar man kring detta i den tidigare refererade norska framtagandet av ett value added-mått (Steffensen m.fl. 2017, s. 81f). Författarna menar att måtten inte kan ge ett uttömmande svar på vad som är skolkvalitet, och att man absolut bör undvika måtten för att kategorisera skolor som bra och dåliga. Anledningen är att det som i mer tekniska termer kan beskrivas som att måttet mäter korrelation och inte kausala samband. Med andra ord, de resultat som framkommer och kan kopplas till en skola behöver inte nödvändigtvis förklaras av det som sker på skolan, utan kan vara faktorer som ligger utanför modellen (se också American Statistical Association 2014).

Däremot menar de att ett offentliggörande av måtten kan bidra till erfarenhetsutbyte mellan politiker, skolhuvudmän, skolledare och lärare och också att det leder till mer debatt kring skolutvecklingsfrågor

Dessa resonemang ligger i linje med slutsatserna i denna utredning. Med bakgrund av att det alltid sker metodövervägande i olika resultatmått så går det inte att göra en direkt koppling till skolkvalitet, men skolbidragsmått ger en indikation på skolkvalitet. Använder man sig av en kombination av resultatmått – totalresultat, justerade resultat, tvärsnittsmått och value added-mått – så är det att föredra framför att bara använda ett mått. Men för att bedöma vad som är skolkvalitet så finns det även andra faktorer som behöver vägas in. Dessa aspekter fångas ibland bättre med kvalitativa metoder, till exempel genom intervjuer och att observera kvaliteten på undervisningen, snarare än olika resultatmått.

Sammanfattningsvis kan skolbidragsmått användas som en indikator på skolkvalitet, men för att få en mer sammanvägd bild bör det kompletteras med andra mått, såväl kvalitativa som kvantitativa. Skolbidragsmått är också intressant information när man ska bedöma och diskutera skolutveckling för enskilda skolor och skolsystem. Kvarstår gör frågan om skolbidragsmått, och mer specifikt value added-mått bör användas som information vid skolval.

Som vi redogjorde för tidigare så avråder Skolverket, under nuvarande förhållanden med lokal bedömning av nationella prov och ämnesbetyg, att man använder value added-mått som information vid skolval. Men även de totala provresultaten och betygen är mått som riskerar att driva på betygs- och provbetygsinflation, och dessa mått används idag i stor utsträckning som information vid skolval. Det som kan invändas mot value added-mått är att de är mer metodkänsliga jämfört med totalresultaten, men samtidigt är de en bättre indikator på skolans prestation. Fördelen skulle då vara att med hjälp av ett value added-mått så kan man vid ett skolval göra en bättre bedömning av skolans prestation än vad de totalresultaten ger möjlighet till. Men det är mycket möjligt att introduktionen av ytterligare ett mått, förutsatt att även totalresultaten ska användas, som information vid skolval skulle öka risken ytterligare för betygs- och provbetygsinflation.

Det avråds också allmänt ifrån att rangordna skolor utifrån value added-mått, och att publicera ett sådant mått vid skolval skulle

inbjuda till en rangordning. Alternativet skulle vara att redovisa siffrorna inom ett konfidensintervall, för att uppväga mot metodosäkerheter, men då skulle en del av den information man vill förmedla gå förlorad och det skulle också vara en siffra som kan vara svår att förstå.

En relaterad fråga är i vilken utsträckning value added-mått används vid skolval. En studie från New York om sökmönstret till high school (motsvarande ungefär gymnasieskolan) visar att högpresterande elever i hög utsträckning tenderar att välja skolor med andra högpresterande elever. Däremot finns det inget som visar att man väljer skolor utifrån vilket förädlingsvärde (value added-mått) skolan har (Abdulkadiroglu m.fl. 2017). I studie från Nederländerna bland 12-åriga elever i Amsterdam (där det finns en valsituation som påminner om den i Stockholm) visar inte på något stöd för att skolans förädlingsvärde påverkar valet, utan det är snarare den totala resultatnivån och elevsammansättningen som är avgörande. Även närheten till hemmet och hur elevens tidigare klasskamrater väljer är viktiga faktorer för skolvalet (Ruijs och Oosterbeek 2017). Även svenska studier kring val till gymnasieskolan betonar i stor utsträckning att skolval i stor utsträckning handlar om social tillhörighet och identitet (se Utbildningsförvaltningen, Stockholms stad 2015).

Med den risk för betygs- och provbetygsinflation som ytterligare ett mått innebär tillsammans med att man bör undvika rankinglistor och den metodkänslighet som finns i uträkningen av så är det att föredra att välja den försiktiga vägen och avråda från att använda value added-mått i information vid skolval. Till detta kan läggas att studier i andra länder visar att value added-mått inte är något som i större utsträckning används i skolvalsituationen.

## Källor

- Abdulkadiroglu, A. m.fl. (2017). *Do Parents Value School Effectiveness?* NBER Working Paper No. 23912.
- American Statistical Association (2014). *ASA statement on Using Value-Added Models for Educational Assessment*, april 8 2014.
- DS 2001:24. *Betyg på skolan – en ESO-rapport om gymnasieskolorna*.
- Harris, D. N. (2009) 'Would accountability based on teacher value added be smart policy? An examination of the statistical properties and policy alternatives', *Education*, 4(4), s. 319-350.
- Steffensen K., Ekren R., Zachrisen O.O. og Kirkebøen, L.J. (2017). *Er det forskjeller i skolars og kommuners bidrag til elevenes læring i grunnskolen?* En kvantitativ studie, Statistisk centralbyrå, rapport 2017/2.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H. and Earl, L. (2014) 'State of the art – teacher effectiveness and professional learning', *School Effectiveness and School Improvement*, 25(2), s. 231-256.
- Skolverket (2012). *Redovisning av uppdraget om skol- och resultatinformation* (Dnr 2012:01002).
- Skolverket (2015). *Value added*, PM 2015-02-11.
- Skolverket (2018). *Analys av familjebakgrundens betydelse för skolresultaten och skillnader mellan skolor*, rapport 467.
- Stockholms stad. *Budget 2018 – Ett Stockholm för alla*.
- SOU 2014:12. *Utvärdera för utveckling – om utvärdering av skolpolitiska reformer*, Utbildningsdepartementet.
- SOU 2016:25. *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning*, Utbildningsdepartementet.
- SOU 2017:35. *Samling för skolan – Nationell strategi för kunskap och likvärdighet*. Utbildningsdepartementet.
- Ruijus, Nienk M. and Oosterbeek, H. (2017). "School choice in Amsterdam: Which schools are chosen when school choice is free?" i *Education Finance and Policy*, posted online june 20, 2017.
- Utbildningsdepartementet (2012). *Uppdrag om skol- och resultatinformation* (U2012/4307/S).
- Utbildningsförvaltningen, Stockholms stad (2015). *Utvärdering gymnasierregion*, Dnr 3.4.2.-10587/2015.