



Likvärdiga betyg och meritvärden

KORTVERSION

Ett kunskapsunderlag om modeller för att främja
betygens och meritvärdenas likvärdighet

Denna kortversion av rapporten Likvärdiga betyg och meritvärden sammanfattar övergripande problembilden, de olika modeller som utvärderats och vilka effekter modellerna har. Kortversionen innehåller också rekommendationer och medskick till en fortsatt diskussion.

Publikationen finns att ladda ner som kostnadsfri PDF från Skolverkets webbplats: skolverket.se/publikationer

ISBN: 978-91-7559-358-6
Grafisk form: AB Typoform
Skolverket, Stockholm 2020

Varför behöver betygen normeras?

Bristande likvärdighet i betygen är ett allvarligt problem för skolsystemet. Om tilliten till betygen urholkas påverkas hela utbildnings-systemets legitimitet. Ansvaret för den bristande nationella likvärdigheten vid betygssättning ligger på systemnivå och kan därför inte tillskrivas lärarkollektivet. Om betygen ska kunna bli mer rättvisa krävs förändringar i själva betygssystemet.

Syftet med den här undersökningen är att utvärdera och bidra med kunskap om hur de nationella proven skulle kunna användas för att uppnå mer likvärdiga betyg och meritvärden. Skolverket vill med rapporten bidra till kunskapsläget och till vidare diskussioner.

Bakgrund

I början av 1990-talet infördes målrelaterade betyg i den svenska skolan, det vill säga betyg som sätts utifrån nationella betygskriterier eller kunskapskrav som enda referenspunkt. Målrelaterade betyg ansågs vara bättre än så kallade normrelativa betyg i alla avseenden – förutom när det gäller urval. Men eftersom gymnasieskolan och högskolan höll på att byggas ut så att det skulle finnas en plats åt alla elever menade regeringen att urval inte längre var en lika viktig uppgift för betygen. Man tänkte att andra urvalsförfaranden, särskilt högskoleprovet, skulle ersätta betyg vid antagning till högskolan. Därför infördes de målrelaterade betygen trots medvetenhet om att de inte skulle bli nationellt likvärdiga.

Betyg och betygssättning har fått större betydelse för urvalet

Men tvärtom vad som förutspåddes fick betygen och betygssättningen en större betydelse för urvalet än tidigare. Skolvalsreformen, friskolereformerna och betygsintaget till gymnasieskolan gjorde att

betygen blev *ännu* viktigare för antagningen till en viss utbildning på en viss skola – inte bara för antagningen till ett nationellt program i gymnasieskolan. När behörighetskrav infördes till gymnasieskolan blev betygssättningen av godkänt (senare betyget E) dessutom helt avgörande för om en elev över huvud taget skulle komma in på ett nationellt program. Även betygen i grundskolan blev ett viktigt mått på skolans resultat och ett sätt att konkurrera om elever. Betygen kom att bli det mest centrala måttet för staten och huvudmännen att utvärdera skolors resultat.

Problembild

Dagens målrelaterade betyg går inte att få nationellt likvärdiga

Stämmer det då att de målrelaterade betygen inte går att få nationellt likvärdiga och att de därför inte är anpassade för urval? Ja, det visar både forskares analyser och Skolverkets undersökningar. Våra senaste rapporter från åren 2019 och 2020 visar att det finns stora skillnader i betygssättning mellan skolor och att skillnaderna är lika stora som tidigare. Den mest sannolika förklaringen till skillnaderna är att lärare gör olika tolkningar av mål, kriterier och kunskaper som eleverna ska uppvisa för att få ett visst betyg.

Problemen med bristande likvärdighet i betygssättningen är väl kända och har varit föremål för många olika insatser. Skolverket har bidragit med nationella underlag och stöd för att hjälpa lärarna att utveckla en mer likvärdig bedömning och betygssättning. År 2011 ändrades betygsskalan och betygskriterier ändrades till kunskapskrav för att öka kursplanernas tydlighet. Skolinspektionen har också ökat sitt fokus på att granska betygssättningen. Sedan

läsåret 2018/19 ska lärarna ”särskilt beakta” de nationella proven.¹ Sammantaget har dock inga av de insatser som gjorts gett någon påvisbar ökad likvärdighet i betygssättningen. Det är tänkbart att likvärdigheten i betygssättningen skulle ha minskat ytterligare om inte dessa insatser hade gjorts men faktum kvarstår: skillnaderna mellan skolors betygssättning är stora och de minskar inte. Efter 27 år måste vi konstatera att det är stora svårigheter för skolmyndigheterna att genom ändringar i styrdokument, stödmaterial och policyer påverka betygssättningen mot högre likvärdighet.

Lärare gör ett bra jobb, men systemet skapar ett relativt inslag

Skolverket har kunnat visa en tydlig relativ betygssättning, det vill säga att betygssättningen tenderar att anpassas till prestationsnivån på skolan. Konsekvensen blir att givet samma prestation, är det svårare för en elev att få ett visst betyg om hen går på en högpresterande skola jämfört med om hen går på en lågpresterande skola. Det innebär att elever som går på högpresterande skolor tenderar att missgynnas av den relativa betygssättningen och elever på lågpresterande skolor gynnas. I grundskolan kan den relativa betygssättningen förklara 25–40 procent av de totala skillnaderna mellan olika skolors betygssättning. I gymnasieskolan är mönstret liknande. Betygsskillnaderna mellan skolorna skulle alltså ha varit betydligt större utan den relativa betygssättningen.

Skolverkets analyser visar också att fristående skolor i genomsnitt sätter mer generösa betyg. Skillnaderna mellan huvudmännens betygssättning är inte den huvudsakliga förklaringen till skillnader i betygssättning mellan skolor, men är ändå ett likvärdighetsproblem som minskar legitimiteten i betygssystemet.

Ett viktigt och positivt resultat är samtidigt att lärarna på det stora hela förmår sätta betyg utan att gynna eller missgynna elever utifrån kön, migrationsbakgrund eller socioekonomisk bakgrund. Det finns vissa skillnader men de bedöms inte som stora.

Systemfel bör åtgärdas på systemnivå

Det finns en inbyggd problematik i betygssystemets utformning som orsakar bristande likvärdighet i betygssättningen. För att nå en markant ökad likvärdighet i betygssättningen krävs mer grundläggande systemförändringar i betygssystemet.

Lärarna kan i hög grad betygssätta sina egna elevers kunskaper på ett rättvist sätt utan att gynna eller missgynna någon elevgrupp. Det är ett krav som är rimligt att ställa och det klarar lärarna väl. Det är däremot orimligt att ställa krav på att en lärare, utifrån kunskapskraven, ska kunna tolka olika elevprestationer som på ett likvärdigt sätt stämmer överens med alla andra lärares tolkningar i alla andra skolor i hela landet. Den bedömningen gjordes av de utredare och beslutsfattare som införde det målrelaterade betygssystemet. Det stämmer också överens med hur lärare själva uppfattar möjligheterna att nå en nationell likvärdighet i dagens betygssystem.

En konsekvens av rådande betygssystem är att lärare sätter betyg utifrån olika förståelser, tolkningar och referensramar. Ansvaret för den bristande nationella likvärdigheten vid betygssättning kan därför inte tillskrivas lärarna utan faller på själva betygssystemet. Det är bakgrunden till varför denna rapport har tagits fram.

Vad leder olikvärdiga betyg till i praktiken?

Skillnaderna mellan skolors betygssättning är inte abstrakta räkneövningar utan har stor och direkt betydelse för alla landets elever i både grundskolan och gymnasieskolan. För att visa hur skillnaderna påverkar i praktiken kan vi använda följande exempel från år 2016.

Två gymnasieskolor (A och B i figuren nedan) med snarlika resultat på de nationella proven kan avvika olika mycket i betygssättningen. När elever ska välja mellan dessa skolor framstår de därför som olika attraktiva utifrån kursbetygen. Skillnaden kan tydlig-

görs genom skolornas rankingplacering baserat på deras kursbetyg. Trots snarlika nationella provresultat hamnar skola A på plats 291 av de totalt 440 skolorna där provet genomfördes medan skola B rankas som den 24:e bästa skolan i riket.

De elever som går på skola A eller B får mycket olika förutsättningar när de gått ut och ska konkurrera med varandra om utbildningsplatser. I skola A fick 14 procent av eleverna ett steg högre kursbetyg än provbetyg. I skola B däremot fick 53 procent av

eleverna ett steg högre kursbetyg än provbetyg. Dessutom fick 26 procent av eleverna två steg högre och 3 procent (1 elev) ett kursbetyg som var tre steg högre än elevens provbetyg. Eleverna i skola B har i genomsnitt fått ett högre kursbetyg än eleverna i 201 av de 224 skolor som presterade bättre resultat på det nationella kursprovet jämfört med skola B.

Tabell 1. Exempel på skillnaden mellan provbetyg och kursbetyg, Matematik 1b vårterminen 2016.

	Provbetyg	Kursbetyg	Rankingplacering provbetyg	Rankingplacering kursbetyg	Skillnad kursbetyg-provbetyg (steg)			
					+0	+1	+2	+3
Skola A	10,54	11,0	224/440*	291/440*	86 %	14 %	0 %	0 %
Skola B	10,53	14,1	225/440*	24/440*	18 %	53 %	26 %	3 %

* Totalt 440 skolor med minst 15 elever per skola ingår i exemplet.

Modeller och resultat

I undersökningen har vi kvantitativt utvärderat hur likvärdigt betygsutfallet blir utifrån ett antal olika betygsmodeller som alla på något sätt utnyttjar resultaten från ett nationellt prov, nationellt slutprov eller examensprov.

Modeller



Dagens betygssystem/Nollmodell: I dagens betygssystem sätter läraren elevens betyg utifrån kunskapskraven och väger in elevens resultat från det nationella provet utifrån att det särskilt ska beaktas. Provet bedöms av läraren själv. Betyget sätts på en målrelaterad AF-skala.



Examensprovmodell: Elevens betyg (i det specifika ämnet) avgörs helt och hållet av elevens resultat på ett nationellt slutprov som bedöms centralt av oberoende bedömarlag. Resultatet redovisas som en intagningspoäng på en årsvis standardiserad 1-100-skala.




Direktstyrningsmodell: Läraren betygsätter varje elevs kunskaper utifrån såväl egna prov och bedömningstillfällen som resultatet från det nationella provet (som i nollmodellen). Men läraren eller skolan måste sedan anpassa betygen så att genomsnittet av skolans betyg överensstämmer med skolans genomsnittliga resultat på det nationella provet.² Provresultat och slutbetyg redovisas som ett målrelaterat betyg på AF-skalan.





Sammanvägningsmodell: Läraren betygsätter elevens kunskaper helt oberoende av resultatet på det nationella slutprovet. Eleven genomför sedan det nationella slutprovet som bedöms centralt (på samma sätt som examensprovsmodellen). Provresultatet redovisas på en årsvis standardiserad 1-100-skala. Elevens lärarsatta betyg vägs sedan samman med resultatet på det nationella slutprovet till en intagningspoäng (i det specifika ämnet). Intagningspoängen redovisas på en standardiserad 1-100-skala.³



Modereringsmodell: Lärarna på skolan bedömer och betygsätter tillsammans elevernas kunskaper. Eleverna skriver sedan ett nationellt slutprov som bedöms centralt av ett oberoende bedömarlag (på samma sätt som sammanvägnings- och examensprovsmodellen). De lokala betygen på varje skola modereras eller justeras sedan till ett betygsvärde så att skolans genomsnittliga modererade betygsvärde och spridning motsvarar skolans genomsnittliga resultat och spridning på det nationella slutprovet.⁴ I ett sista steg vägs elevens modererade betygsvärde ihop med elevens individuella resultat på det nationella slutprovet till en intagningspoäng i det specifika ämnet (på samma sätt som i sammanvägningsmodellen). Intagningspoängen redovisas på en årsvis standardiserad 1-100-skala.

 Lärarsatt betyg

 Nationellt prov/slutprov

 Modererat lärarbetyg



Nationella provets betydelse

Likheter och skillnader i modellerna

Det finns viktiga egenskaper som både skiljer sig åt och är gemensamma för de olika modellerna:

- I nollmodellen och direktstyrningsmodellen redovisas både betygen och resultatet på det nationella provet på en diskret målrelaterad A-F-skala (sex steg), det vill säga en skala med enbart hela betygssteg där betygen sätts utifrån kunskapskraven. I sammanvägningsmodellen, modereringsmodellen och examensprovsmodellen redovisas resultaten på det nationella slutprovet på en årsvis standardiserad 1–100-skala, det vill säga en mer fingradig skala där varje årskull kommer att ha samma medelvärde och spridning från år till år. Därmed redovisas även intagningspoängen i det specifika ämnet på en standardiserad 1–100-skala.
- Det enda som skiljer modereringsmodellen från sammanvägningsmodellen är att det nationella slutprovet även används för att normera betygen på skolnivå (förutom att ge eleven ett individuellt provresultat). I sammanvägningsmodellen används resultatet på det nationella slutprovet endast för att ge eleven ett individuellt provresultat. Sammanvägningsmodellen påverkar således inte skolornas betyg över huvud taget.
- I nollmodellen och direktstyrningsmodellen väger läraren in resultatet på det nationella provet i betyget och avgör därmed själv hur stor betydelse det nationella provet ska ha för betyget för den enskilde eleven. I sammanvägningsmodellen och modereringsmodellen bestäms det centralt och på förhand hur stor betydelse den enskilde elevens resultat på slutprovet ska ha för intagningspoängen. Sammanvägningen sker centralt, det vill säga sammanvägningen är transparent och lika för alla elever.

- I nollmodellen och direktstyrningsmodellen har det nationella provet betygsgränser och provbetyg, vilket kräver ekvivalering med ankaruppgifter⁵ för att proven inte ska variera i svårighetsgrad mellan olika år. I sammanvägningsmodellen och modereringsmodellen standardiseras provresultaten årsvis till en 1–100-skala utan provbetygsgränser. Det innebär att olika svåra prov inte påverkar likvärdigheten i betygssättningen mellan olika årskullar. Därmed behövs inga ankaruppgifter och proven kan offentliggöras efter provtillfället.

Hur utvärderar vi?

Utvärderingsmättet som används i den kvantitativa analysen är den Genomsnittliga Avvikelsen i Percentilranking (GAP). GAP-indikatorn mäter (utifrån datasimuleringar) avvikelser mellan betygsutfall och ett perfekt rättvist betyg. Avvikelse kan dels uppstå på grund av orättvis betygssättning (till exempel att betygen sätts mer generöst på vissa skolor), dels på grund av slumpvariation då prov och olika bedömningstillfällen är förknippade med mätfel (till exempel att eleven kan ha en bra eller dålig dag på det nationella slutprovet).

GAP-indikatorn fångar upp båda dessa typer av avvikelser. GAP-indikatorn visar därmed hur väl elevrankingen (respektive skolrankingen) enligt en viss betygsmodell stämmer överens med motsvarande elevranking (skolranking) om betygen hade satts perfekt rättvist och utan mätfel. Ett lägre värde på GAP-indikatorn innebär ett bättre likvärdighetsutfall.

Vilken modell ger bäst effekt på betygens likvärdighet?

I diagram 1 visas utfallet från den kvantitativa analysen utifrån GAP-indikatorn för respektive betygsmodell.⁶

Modereringsmodellen är den modell som ger allra bäst effekt i de olika kvantitativa utvärderingarna (lägst värden i diagram 1). Sammanvägningsmodellen ger nästan samma effekt som modereringsmodellen för likvärdigheten på elevnivå (blågrå stapel i figuren). Modellen kommer däremot inte åt olikvärdig betygsättning mellan skolor, eftersom den inte styr betygsättningen på gruppnivå (orange stapel). Direktstyrningsmodellen ger bättre resultat än sammanvägningsmodellen för likvärdigheten på skolnivå, men däremot sämre på elevnivån.

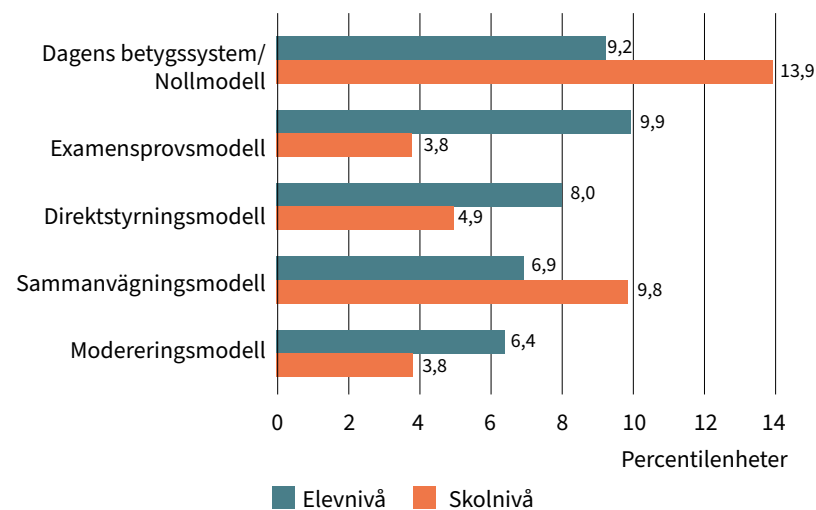
Examensprovsmodellen, som innebär att en elevs betyg helt och hållet avgörs vid ett enda provtillfälle (per ämne), leder däremot till sämre effekt på likvärdigheten på elevnivå än dagens betygssystem. Det beror på att även om ett centralt bedömt examensprov minimerar systematiska orättvisor är mätfelet av den enskilda elevens kunskaper mycket stort eftersom betyget baseras på ett enda provtillfälle. På skolnivå fungerar däremot examensprovsmodellen bättre än dagens modell.

I datasimuleringarna har vi antagit en hundraprocentig regel-efterlevnad av direktstyrningsmodellen. Dessutom har vi bortsett från en potentiell olikvärdig bedömning av de nationella proven i dessa datasimuleringar (lärarna bedömer själva proven i nollmodellen). Därmed kan de positiva effekterna på likvärdigheten vara något överskattade både när det gäller direktstyrningsmodellen och nollmodellen (GAP-värdet underskattat).

Det är inte enbart själva normeringen på skolnivå som bidrar till ett mer likvärdigt totalt betygsutfall jämfört med nollmodellen (för 3 av 4 betygsmodeller). En annan viktig faktor är att slutprovet utnyttjas som en oberoende mätning av elevernas kunskaper och där informationsförlosterna minimeras genom att resultatet

rapporteras på en mer fingeradig skala. Dessutom vägs det oberoende provresultatet ihop med det lokala lärarsatta betyget på ett för alla elever likvärdigt sätt. Direktstyrningsmodellen använder endast normeringen medan sammanvägningsmodellen endast använder ett nationellt oberoende slutprov. Modereringsmodellen utnyttjar däremot både normering och ett oberoende slutprov, vilket sannolikt är förklaringen till att modereringsmodellen sammantaget är den modell som ger bäst effekt på betygens likvärdighet enligt denna analys.

Diagram 1. Genomsnittlig avvikelse i elevers respektive skolors percentil-ranking för olika betygsmodeller (GAP), matematik.*



* Ett lågt värde visar på en mer likvärdig betygsmodell för hela årskullen sammantaget. GAP på elevnivå anger hur väl elevernas ranking stämmer överens med den ranking de skulle fått om de rankats efter perfekt rättvisa betyg utan mätfel. GAP på skolnivå anger hur väl skolornas ranking stämmer överens med den ranking de skulle fått om de rankats efter genomsnittet av sina elevers perfekt rättvisa betyg och utan mätfel.

GAP-värdet på elevnivå är det mest betydelsefulla måttet och indikerar det totala likvärdighetsutfallet. Om en elev tex går på en skola med en relativt restriktiv betygsättning fångas detta upp av GAP-indikatorn på elevnivå. GAP-indikatorn på skolnivå ger dock en mer direkt uppfattning om hur olikvärdigheten på skolnivå faller ut i de olika modellerna.

Slutsatser

Det går att förbättra betygens likvärdighet med nationella slutprov

Undersökningens huvudsakliga resultat är att det går att förbättra betygens likvärdighet avsevärt genom att använda centralt bedömda nationella slutprov för antingen moderering, sammanvägning eller direkt styrning av betygen. En fördel med alla modeller är att det framstår som både välavvägt och effektivt att utnyttja en kombination av lärarnas unika information om elevernas kunskaper med centralt bedömda prov på ett sätt som tydliggör de nationella provens roll. Däremot hanteras systemproblemet med att skolor tolkar kunskapskraven på olika sätt enbart i modereringsmodellen och direktstyrningsmodellen.

Modereringsmodellen har störst positiv effekt

Modereringsmodellen har störst positiv effekt på betygens likvärdighet, sammanvägningsmodellen något lägre följt av direktstyrningsmodellen. Modereringsmodellen ger lärarna och rektorn hela ansvaret för betygsättningen inom den egna skolan. Det är centralt för modellen att betygssättningen verkligen samordnas på skolan med till exempel gemensamma bedömningsituationer. Den potentiella effekten på likvärdigheten är stor jämfört med dagens system, runt 30 procent i de undersökta ämnena utifrån den likvärdighetsindikator som använts i analysen.⁷ Det beror på att staten genom proven tar ansvar för likvärdigheten mellan skolor samt att det nationella slutprovet fungerar som en oberoende mätning av elevens


kunskaper. Provens syfte kan på det viset renodlas vilket kan ge prov som både är enklare och billigare att utveckla, mer varierade och lättare att bedöma centralt. Med modereringsmodellen försvinner dessutom alla incitament för skolorna att inflatera betygen.

Examensprov leder till sämre likvärdighet jämfört med dagens system

Ett viktigt resultat är att en modell med renodlade examensprov visserligen löser den systematiska olikvärdigheten mellan skolor, men ger sämre effekt på den totala likvärdigheten än nuvarande betygssystem. Det beror på att elevens betyg avgörs vid ett enda mättillfälle i respektive ämne.

Viktigt med transparens om provens uppgift och betydelse

Ju större vikt proven får, desto större anledning att förbereda eleverna till proven. Risken är att undervisningen inriktas mot proven och på så vis smalnar av läroplanen. Därför är det viktigt att en modell inte leder till att proven får omotiverat stor betydelse och att provens uppgift i systemet tydligt framgår. I modereringsmodellen och sammanvägningsmodellen behöver visserligen provens vikt bestämmas på förhand, men det behöver inte betyda att vikten blir större än vad den är i dag.



Ju färre nya regler och arbetsmetoder desto bättre för lärarna

Det är en fördel ju färre nya regler, förhållningssätt och arbetsmetoder en modell skulle föra med sig. Modererings- och sammanvägningsmodellerna löser detta genom att inte lägga ansvaret på lärarna själva att väga ihop provresultaten. Modereringsmodellen ställer däremot krav på att lärare arbetar kollegialt för samsyn inom skolorna. Direktstyrningsmodellen förutsätter att lärare ändrar sina arbetssätt så att de kan hålla sig till de tillåtna medelvärdena.

Rekommendationer

Modellerna behöver utredas mer

Förmodligen finns det inget system som bokstavligen kan garantera alla elever perfekt rättvisa betyg. Snarare handlar det om att utforma system så att de eliminerar så mycket systematisk orättvisa som möjligt samtidigt som informationen om elevens kunskaper omhändertas med minsta möjliga mätfel. Förutom examensprovmodellen, ger samtliga modeller bättre likvärdighet jämfört med nuvarande system men det behöver utredas ytterligare hur de olika modellerna påverkar i förhållande till fler elevgrupper. I modererings- och direktstyrningsmodellerna behöver det undersökas hur elever som avviker mycket från sin skolas genomsnittliga prestationsnivå påverkas, och vid behov utveckla lösningar för att se till att de eleverna får en rättvis behandling.

Behörighetskraven och åtgärder för att dämpa oönskade effekter behöver ses över

Att betygen är trovärdiga är en central legitimitetsfråga för hela utbildningssystemet. Om betygen inte går att lita på kommer viktiga beslut om skolan att fattas på felaktiga grunder. Men mer likvärdiga betyg innebär också att nu dolda skillnader synliggörs i högre grad. Med modererings- och direktstyrningsmodellerna kommer den relativa betygssättningen att försvinna. Därmed behöver behörighetskraven och eventuella åtgärder för att dämpa oönskade effekter ingå i en översyn. Risken är annars att ännu fler elever på skolor med dåliga socioekonomiska förutsättningar blir obehöriga till gymnasieskolan och inte klarar av det som krävs för en gymnasieexamen.

Med en direktstyrningsmodell skulle andelen elever som får betyget F sannolikt öka. Modererings- och sammanvägningsmodellerna kan däremot (om önskvärt) användas enbart för att kalibrera meritvärdet för urval och inte behörigheten, och påverkar då inte andelen F. Likvärdighet är naturligtvis viktigt även vid godkändgränsen, men samtidigt kan det finnas fördelar med att hantera olikvärdighet där den kanske spelar allra störst roll: för urval i konkurrensen om attraktiva utbildningsplatser. Detta är viktiga frågor att beakta.

Hur kan likvärdigheten stärkas i flera ämnen?

De nuvarande nationella proven ger endast begränsad möjlighet att påverka likvärdigheten. Fem av cirka 16 ämnen i grundskolans årskurs 9 och 300 av 2 500 poäng i en gymnasieexamen är för lite för att göra en större skillnad. Fler provämnen kan vara en lösning, normering av närliggande ämnen mot provämnen en annan och att begränsa urvalsgrundande ämnen till provämnen en tredje. På grundskolan skulle meritpoängen kunna viktas efter timplanen för att öka likvärdighetseffekten. Det kan också diskuteras om samtliga gymnasiekurser ska ligga till grund för urvalet till högskolan. Det skulle också kunna gå att kombinera någon av de tre provbaserade modellerna med en annan slags modell för ämnen där prov inte fungerar.

Slutord

Skolverket vill med den här undersökningen synliggöra systemproblematiken i dagens betygssystem och visa hur de nationella proven kan användas för att uppnå mer likvärdiga betyg och meritvärden. Undersökningen visar att någon form av normeringsmodell är nödvändig för att åstadkomma en märkbar effekt på betygens likvärdighet. Om en modell inte tillämpas behöver frågan ställas huruvida dagens betyg är ändamålsenliga för urval till fortsatta studier. Innan några förändringar görs i dagens betygssystem återstår att undersöka, ta hänsyn till och diskutera flera viktiga aspekter. Rapporten ska därför ses som ett bidrag till kunskapsläget och en katalysator för en bred och saklig diskussion om alternativa betygssystem.

Slutnoter

1. Utifrån analys av data från det enda läsåret då bestämmelsen gick att tillämpa (2018/19) syns inga positiva effekter av bestämmelsen. Men bestämmelsen är relativt ny och flera års data kan behöva analyseras innan mer robusta slutsatser kan dras.
2. Modellen påminner om hur betygen ankrades under det relativa betygssystemet, men med den viktiga skillnaden att betygen här sätts i relation till mål.
3. Sammanvägningen behöver inte nödvändigtvis beräknas som ett enkelt genomsnitt där betyg och provresultat ges samma betydelse. Det lärarsatta betyget eller resultatet från det nationella slutprovet kan ges större eller mindre betydelse. Detta gäller även modereringsmodellen.
4. Modereringsprocessen innebär att betygen inte nödvändigtvis behöver justeras ned eller upp i diskreta steg utan kan ske i mindre kontinuerliga steg. Den interna rangordningen påverkas inte av modereringen, dvs den elev som fått det högsta lokala betyget på skolan kommer också att få det högsta modererade betygsvärdet på skolan.
5. Justeringar av provens kravgränser i efterhand utifrån uppgifter som återkommer i de olika proven över tid.
6. Figuren gäller matematik men mönstret är detsamma även för svenska och engelska som också analyserats (se huvudrapport).
7. I datasimuleringen har ingen gemensam betygssättning antagits för modereringsmodellen, vilket innebär att det finns en viss olikvärdighet mellan olika klasser inom samma skola. Det är därmed troligt att det finns ytterligare förbättringspotential vilket innebär att den positiva effekten på likvärdigheten med modereringsmodellen kan vara något underskattad (GAP-indikatorn överskattad).