



ASTRID PETTERSSON

Bedömning – vad, varför och vart är vi på väg?

FENOMENET BEDÖMNING har förekommit länge även om inte själva ordet har använts. Ordet bedömning kom ganska sent in i det svenska skriftliga språket. I *Nationalencyklopedins ordbok*¹ står att ordet bedömning kom in cirka 1900. Som jämförelse kan nämnas att förhör finns i skriven text sedan 1571, prov sedan 1630 och betyg sedan 1767.

Bedömning är ett mycket komplext begrepp och används i olika sammanhang och i olika betydelser. Ofta jämföras bedömning i skolan med betyg, vilket är en form av bedömning. Men vad har betygsatts? I skolordningen från 1820 var elevens hela personlighet i fokus. Eleven fick betyg i skalan A–D i flit, förstånd, minne och lättjefullhet. Men läskunnigheten bedömdes också med betyg.²

Läskunnigheten var redan i mitten av 1700-talet hög i Sverige, då mellan 70 och 90 procent av den vuxna befolkningen kunde läsa. I början av 1800-talet var andelen över 90 procent.³ Anledningen till det var att Sverige under 1500-talet hade övergett katolicismen och övergått till protestantismen, som fordrade att varje person skulle kunna »läsa i bok«, dvs.

1. Nationalencyklopedin, 1995, s. 109.

2. Andersson, 2002, s 153.

3. Samuelsson, 2022, s. 27.

Bibeln. Det var husbondens plikt att lära sitt husfolk läsa och det var prästens plikt att kontrollera att församlingsmedlemmarna kunde läsa. Detta skedde genom husförhören, som lagstadgades 1686. Läskunnigheten var förutsättningen för att få konfirmeras och konfirmation var ett krav för att få gifta sig. Resultaten av husförhören skrev prästen in i husförhörlängder, där förutom omdömen om läskunnigheten också fanns omdömen om personlighet och leverne. Många präster bedömde läskunnigheten utifrån en femgradig skala: har börjat läsa, läser svagt, läser någorlunda, läser försvarligt, läser med färdighet. I husförhörlängderna står också om församlingsmedlemmarnas personliga egenskaper exempelvis »har svagt förstånd, trögt minne men är flitig« eller »har gott förstånd och minne men är lättjefull« eller »är döv och dum«, eller »är snäll mot barnen«.⁴

Betyg i personliga egenskaper levde kvar länge, framför allt i ordning, uppförande och flit. Under 1950-talet fanns inte flit med, men däremot ordning och uppförande i folkskolan. Betygen blev mer och mer inriktade mot bedömning av elevernas kunskaper i olika ämnen och var en form av målrelaterad bedömning. Lärare skulle på egen hand bedöma vad som fordrades för respektive betyg. Betygssystemet har kallats för absolut betygsättning⁵ och det fanns inga krav på att betygen skulle vara jämförbara.

Från ett absolut betygssystem till ett relativt

En elev genomgick inträdesprov för att pröva in till läroverk. Inträdesproven var både muntliga och skriftliga och det var bara resultaten av elevernas prestationer på dessa prov som avgjorde om eleven fick börja på läroverket eller ej. Någon hänsyn till vad eleven hade för betyg i folkskolan togs inte. När antalet sökande till läroverket blev för stort blev detta system ohanterligt. Dessutom ansåg man att betygen som grundar sig på flera års

4. Pettersson, 2010, s. 7–8.

5. Andersson, 2002, s. 154.

skolarbete borde vara tillförlitligare än den bedömning som man kunde få genom några timmars prov i en för den sökande ny och ovan omgivning. Ytterligare ett skäl att överge inträdesproven var att dessa skapade stress och drill hos eleverna innan proven genomfördes. Frits Wigforss fick i slutet av 1930-talet i uppgift att undersöka om intagningsförfarandet skulle kunna ske på något annat sätt. Efter att ha konstaterat att prognosförmågan för studieframgång var minst lika bra för skolbetygen som för inträdesproven föreslog Wigforss att skolbetyg skulle användas för urvalet till läroverket.⁶

Men det fanns i det absoluta betygssystemet inget krav på att lärarnas betyg skulle vara jämförbara i riket och det var naturligtvis viktigt om betygen skulle ligga till grund för urval. Wigforss kunde i sina undersökningar konstatera att lärare satt olika betyg på samma elevprestationer. Man måste då göra något för att lärarnas betyg skulle kunna bli jämförbara. Man antog att elevernas skolprestationer i ett visst ämne och i en viss årskurs var normalfördelade eftersom de flesta fysiska egenskaper var det. Denna utgångspunkt kunde inte bevisas och det fanns de som var kritiska till att använda samma modell som fungerade på fysiska egenskaper som längd och vikt också på kunskapsprestationer. Men så blev det. För att detta skulle få genomslag vid lärares betygsättning krävdes något instrument som gav lärarna möjlighet att jämföra sina elevers prestationer med rikets alla elevers. Därför infördes på 1940-talet standardprov i folkskolan i vissa ämnen och i vissa årskurser. Standardprov användes till och med våren 1997 i årskurs 9 i ämnena engelska, matematik och svenska. Standardprovet skulle vara styrande för lärares betygsättning och syftet var att skapa största möjliga likformighet i betygsättningen i landet. Standardproven konstruerades, med undantag för några år, vid forskningsinstitutioner vid några av våra lärosäten. Eleverna genomförde proven och därefter skickade lärarna in resultaten för ett slumpmässigt urval av sina elevers prestationer till forskningsinstitutionerna, som överförde elevernas resultat till normtabeller

6. Andersson, 2002, s. 154–155.

som gav resultatfördelningar som överensstämde med normalfördelning. Gränserna för de olika provbetygen sattes i efterhand och skickades ut till skolorna, så att lärarna kunde se hur många elever (inte vilka elever) som skulle få respektive betyg. Även om provresultaten var styrande för lärarna fick betygsättningen skilja sig från provbetyget. Betygsskalan var under de sista åren femgradig med betygen 1–5, där 5 var högsta betyg. För gymnasieskolan fanns liknande nationellt utarbetade prov, centrala prov. För gymnasieskolan var det tillåtet att klassens medelbetyg fick avvika med 0,2 enheter från klassens medelbetyg på det centrala provet. Medelbetygen fick också skilja mer från provbetygens medelvärde om rektor gav tillstånd (efter lärares motivering) till större skillnad. Någon motsvarande begränsning fanns inte för grundskolan. Eftersom normalfördelning skulle tillämpas ställde det krav på de uppgifter som ingick i proven. De skulle ha olika svårighetsgrad och varje uppgift skulle ha förmågan att skilja eleverna åt, de högrepresterande från de lägrepresterande.⁷

Från ett relativt betygssystem till ett mål- och kunskapsrelaterat

Det relativa betygssystemet ifrågasattes av flera skäl. Ett var att betyget inte informerade om vad eleven hade för kunskaper utan bara informerade om elevens kunskaper i relation till andra elevers kunskaper. I Sverige var en vanlig missuppfattning att normalfördelningen gällde en enskild skola eller en enskild klass. Eleverna kunde få höra att »femmorna är slut men ettor finns det gott om«. Att normalfördelningen skulle gälla en årskurs i respektive ämne i hela landet fick inte alltid genomslag. Den statliga styrningen över skolan förändrades under 1990-talet. Under den tid då det relativa systemet tillämpades var skolan regelstyrd men ersattes av ett mål- och resultatstyrt system. Regering och riksdag skulle styra skolverksamheten genom att formulera nationella mål, men huvudmännen skulle ha stor fri-

7. Pettersson m.fl., 2019, s. 73.

het att bestämma över verksamheten. Betygssystemet förändrades också från och med 1994/95. I grundskolan blev det tre betygssteg, Godkänt, Väl godkänt och Mycket väl godkänt. I gymnasieskolan kompletterades den tregradiga skalan med betyget Icke godkänt. Det utarbetades nationella betygskriterier, till att börja med bara för betygen Godkänt och Väl godkänt, men efter några år formulerades betygskriterier också för betyget Mycket väl godkänt. Senare ersattes betygskriterier med kunskapskrav och betygsstegen utökades till sex och betecknades F–A, där F är ett icke godkänt betyg, som också infördes i grundskolan. A är det högsta betyget. Från och med den 1 juli 2022 återgick man till begreppet betygskriterier i stället för kunskapskrav.

I det relativa systemet är utgångspunkten att man på ett ganska säkert sätt ska kunna mäta skillnader i kunskaper och rangordna individer. I ett mål- och kunskapsrelaterat system är utgångspunkten att det går att formulera mål som går att förstå och tillämpa. För att detta ska kunna uppfyllas måste lärarna få utbildning.⁸ Men lärarna fick så gott som ingen utbildning för hur bedömning skulle ske och betyg sättas.

Till det relativa betygssystemet fanns standardprov för grundskolan och centrala prov för gymnasieskolan i vissa ämnen och dessa prov skulle styra lärarens betygsättning på gruppnivå. För det mål- och kunskapsrelaterade systemet infördes nationella prov i vissa ämnen för att stödja lärarens bedömning av elevernas prestationer på individnivå. Tanken var att dessa prov bara skulle förekomma under några år – sedan skulle lärarna vara så förtrogna med betygskriterierna att nationella prov inte skulle behövas. Provens syfte var att stödja, inte styra, lärarnas betygsättning. De nationella provens syften har förändrats över tid. I grundskolan var syftena till att börja med

- * Stödja läraren i dennes bedömning av om och hur väl den enskilda eleven nått målen i kursplanen

8. Andersson, 2002, s. 159.

- * Ge stöd för betygsättningen
- * Bidra till att betygen blir jämförbara över landet.

I gymnasieskolan fanns ytterligare ett syfte med att proven skulle ge en så god grund som möjligt för en rättvis betygsättning i landet, då betygen var viktiga som urvalsinstrument.

Från och med 2018/19 är syftet för såväl grundskolan som gymnasieskolan att stödja en likvärdig och rättvis betygsättning.⁹

Om vi jämför de olika syftena kan vi konstatera att standardproven och de centrala proven endast hade ett syfte, att åstadkomma största möjliga likformighet i betygsättningen i landet. Då de nationella proven infördes formulerades flera olika syften, där det till att börja med var att stödja läraren i bedömningen av om och hur väl den enskilda eleven nått målen i kursplanen. Nu står likvärdigheten och rättvisan i bedömning och betygsättning i fokus.

Att styra eller stödja

Ett sätt att styra respektive stödja lärares betygsättning är att använda centralt utarbetade prov – standardprov och centrala prov för att styra och nationella prov för att stödja. I och med reformen på 1990-talet blev statens inflytande på skolverksamheten betydligt mindre och huvudmännens betydligt större. Staten kunde exempelvis inte bestämma om och i vilken utsträckning lärarna skulle få utbildning i bedömning och att sätta betyg, det var huvudmännens ansvar. Till att börja med märktes förändringen på provnivå på det sättet att antalet prov utökades. För matematikens del utökades antalet prov både för grundskolan och gymnasieskolan. Tidigare fanns bara ett standardprov i årskurs 9, i det nya systemet tillkom ett nationellt prov för årskurs 5 utöver ett för årskurs 9. För gymnasieskolans del utökades antalet prov ännu mer. Från två centrala prov (ett för NT-lin-

9. Pettersson, m.fl. 2019, s. 75-76.

jerna och ett för SE-linjerna) blev det ett prov per matematikkurs. Eftersom staten inte skulle lägga sig i skolornas planering, måste det utvecklas och konstrueras två prov per år för respektive kurs. Men det infördes också provbanker, diagnostiska prov, olika typer av bedömningsstöd och kartläggningmaterial för att stödja lärare i deras bedömningsarbete.

De nationella proven skulle spegla läroplanens kunskapssyn och kursplanens ämnessyn. För matematikens del skulle poäng på elevernas lösningar undvikas. Poängen skulle nämligen kunna locka till att beräkna medelvärden och det var alltför lika det som användes i det relativa systemet. I stället skulle elevernas kvaliteter vad gällde deras lösningar bedömas med olika beskrivningar. Elevens prestationer på hela provet skulle sammanfattas i en kunskapsprofil, där elevens styrkor och svagheter framgick. Detta system genomfördes i stort sett bara för de prov som inte skulle vara betygsgrundande. Men för alla prov i matematik användes i det nya systemet många fler öppna uppgifter, uppgifter som både kunde ge flera korrekta svar men också kunde lösas på flera olika sätt och som gjorde att olika kvaliteter i lösningarna kunde synliggöras. Ett exempel på en uppgift som skulle kunna användas i det mål- och kunskapsrelaterade systemet är: »Hur många pepparkakor kan man baka av en halv liter pepparkaksdeg?«

Nu finns nationella prov i matematik i grundskolan i årskurs 3, 6 och 9, där de två sistnämnda ska vara betygstödjande. Nationella prov i svenska/svenska som andraspråk finns i årskurs 3, 6 och 9 och i engelska i årskurs 6 och 9. I årskurs 9 förekommer också nationella prov i de naturorienterade och samhällsorienterade ämnena. I matematik infördes också muntliga delprov, som numera bara är obligatoriska för grundskolan. I matematik används för de betygstödjande proven kvalitetspoäng, dvs. kvaliteterna i lösningarna bedöms med E-poäng, C-poäng och A-poäng beroende av hur relaterade lösningarnas kvaliteter är till respektive kunskapskrav/betygskriterier. Resultaten sammanfattas som en totalpoäng och ett antal olika kvalitetspoäng. För att exempelvis få provbetyget C behövs ett minsta antal totalpoäng och ett minsta antal poäng på C- eller A-nivå. En väsentlig skillnad till de prov som utarbetades i det relativa systemet är att

numera får lärarna ut bedömningsanvisningar och betygsgränser i anslutning till att proven genomförs på skolorna. Bedömningsanvisningarna i det relativa systemet bestod bara av ett par sidor, medan det i det nuvarande systemet är mångdubbelt fler sidor eftersom dessa också ska innehålla bedömda elevlösningar, så att lärarna ska kunna se kvalitetsskillnader i lösningarna.

Det första nationella provet i matematik som gavs i det nya systemet var ett prov för kurs A, vårterminen 1995 för gymnasieskolan. Det bestod av två delar, en tidsbunden del som skulle genomföras på 120 minuter och en breddningsdel som skulle ge underlag för att bedöma andra aspekter av kunskap än de som avsågs att prövas i den tidsbundna delen. Den tidsbundna delen skulle genomföras någon gång under en vecka i maj, med den första provdagen som rekommenderad provdag. Breddningsdelen skulle genomföras under ordinarie matematiklektioner under drygt en månad i april och maj. Eleverna fick i breddningsdelen välja en av tre uppgifter inom tre olika kunskapsområden (aritmetik, geometri och statistik) och provtiden var högst 180 minuter. Eftersom det bara fanns nationellt utarbetade betygskriterier för Godkänt och Väl godkänt gav det första nationella provet bara stöd för betygssättningen för betygen Godkänt och Väl godkänt. I informationen till lärarna fördes en diskussion om betygsgränser och det gavs riktmärken för provbetyget Godkänt respektive Väl godkänt. Men det påpekades att det bara var riktmärken och att läraren naturligtvis kunde väga in andra aspekter av elevernas prestationer för att bestämma provbetyget. Mycket överläts alltså till lärarnas profession.¹⁰ Dessutom lyftes fram att »Resultatet på ett nationellt kursprov är alltså bara en del i lärarnas totala bedömning av elevernas kunskaper.«¹¹

Mycket har hänt sedan det första kursprovet gavs för gymnasieskolan men också för grundskolans nationella prov. Elevernas val av uppgifter i breddningsdelen har tagits bort och ersatts av en mindre omfattande

10. Pettersson m.fl., 2019, s. 78–79.

11. Skolverket/PRIM-gruppen, 1995, s. 2.

uppgift. Diskussionen om spridningsrisken har lett till att fastställda provdagar har införts för hela landet. Utförliga bestämmelser och föreskrifter har getts ut av Skolverket om hur de nationella proven ska hanteras och genomföras och när lärarna ska få tillgång till bedömningsanvisningarna. För vissa nationella prov har också exakt klockslag på provdagen angetts av Skolverket när eleverna ska börja arbeta med proven. Hanteringen av de nationella proven och genomförandet är, med den utveckling som varit, mycket styrande.

Jämfört med proven i det relativa systemet är det nuvarande systemet mycket mer styrande. Det fanns en frihet för grundskolan och lärarna när det gällde provdag och tidpunkt på provdagen då eleverna skulle arbeta med dessa. En rekommendation var att alla klasser i samma skola skulle göra proven på samma dag. Läraren fick själv bestämma hur elevens formella behandling i sina lösningar skulle bedömas. Den formella behandlingen skulle bedömas på det sätt som eleverna var vana vid i klassen. Lärarna fick också bestämma om invandrarelever/nyanlända elever skulle delta i provet eller ej. Hanteringen och genomförandet var inte alls lika styrande som de är i det nuvarande systemet.

Men fokus är inte på hanteringen och genomförandet utan på provbetygen och vilken vikt som dessa ska ha för slutbetyg/terminsbetyg. I det relativa systemet styrdes det på gruppnivå med hjälp av normalfördelningen. I det nuvarande systemet finns det än ingen fördelning att följa. Skolverket har kompletterat och stramat upp sina anvisningar vad gäller relationen provbetyg/slutbetyg genom att skriva fram att provbetyget ska särskilt beaktas vid betygssättning, vilket innebär att provresultaten ska ha en särskild betydelse vid betygssättningen och ha större betydelse än andra enskilda underlag. Men provbetyget ska inte helt styra betyget och det nationella provet får inte vara lärarens enda underlag vid betygssättningen.

Men diskussioner om hur man ska hantera betygsinflation och att många (alltför många) elever får ett högre betyg än vad provbetyget visar har medfört att man diskuterar om man kan skapa någon modell som lärare ska använda för att fördelningen av slutbetyg bättre ska överensstämma

med fördelningen av provbetyg. Men frågan är om och hur ett system vars syfte är att stödja lärares betygsättning också ska tillåtas styra den.

Internationella kunskapsmätningar

De internationella storskaliga kunskapsmätningarna blir fler och fler och antalet länder som deltar ökar. Resultaten från de internationella kunskapsmätningarna blir mer och mer uppmärksammade både medialt och politiskt. Syftena med de internationella kunskapsmätningarna är att de ska ge ett generellt mått på svenska elevers prestationer i ett internationellt perspektiv, ge tillgång till trendresultat i relation till andra länder, ge ett mått på effektiviteten i olika länders utbildningssystem samt ge underlag för politiska beslut.¹² Men varför började man med internationella kunskapsmätningar?

Utbildningsforskare började redan på 1950-talet diskutera behovet av internationella studier av olika utbildningssystem. Många länder inklusive Sverige hade vid den tiden påbörjat en intensiv reformverksamhet av skolan och många forskare ansåg att det var viktigt att göra jämförande studier. Det skulle ha ett värde för att ta reda på vilka faktorer som hade betydelse för hur och vad eleverna presterade. De olika länderna skulle kunna lära av varandra. En samarbetsorganisation bildades 1967, med förkortningen IEA (International association for the Evaluation of educational Achievement). En av initiativtagarna var professor Torsten Husén vid Lärarhögskolan i Stockholm. Han var också under flera år IEA:s ordförande. Många utmaningar väntade, bland annat vad gäller hur undersökningarna skulle läggas upp. Det blev stickprovsundersökningar där ett slumpmässigt urval elever valdes ut. Det var det som var gemensamt i de olika ländernas kursplaner som skulle ligga till grund för innehållet i testerna och de ämnen/ämnesområden som skulle undersökas var matematik, naturvetenskap

12. Pettersson, 2018, s. 150.

och läsning. Ett antal undersökningar har det blivit under åren med oftast fyra års mellanrum. Undersökningarna i matematik och naturvetenskap benämns numera TIMSS (Trends in Mathematics and Science Study) och i läsning PIRLS (Progress in International Reading Literacy Study). Enkäter till elever och lärare är en del av IEA-undersökningarna.¹³

Under 1990-talet tog OECD (Organisation for economic co-operation and development) initiativ till ytterligare internationella kunskapsmätningar. OECD har ett stort intresse för utbildningspolitiska frågor eftersom ekonomisk utveckling är relaterad till invånarnas utbildningsnivå och OECD var intresserade av det livslånga lärandet. OECD-undersökningarna fick förkortningen PISA (Programme for international student assessment) och var också stickprovsundersökningar men var till viss del bredare än IEA-undersökningarna. Bland annat undersöktes elevernas kunskaper och förmågor även i andra ämnesområden än matematik, naturvetenskap och läsförståelse, som exempelvis problemlösning. IEA utgår från de deltagande ländernas kursplaner och vänder sig till elever i årskurserna 4 och 8. PISA-undersökningarna utgår däremot från de kunskaper och färdigheter som anses vara betydelsefulla för det vuxna livet och vill också fånga in andra förmågor som att fatta beslut och att kunna analysera och designa system, förmågor som är viktiga i det vuxna livet. PISA vänder sig därför till elever som redan har genomfört en stor del av den grundläggande utbildningen, dvs. 15-åriga elever. Det innebär att vissa elever har gått i skolan ett år längre än andra eftersom skolstarten inte är lika i alla länder. I OECD-undersökningarna förekommer också enkäter till elever och skolledare. De utförs vanligtvis vart tredje år och den första undersökningen startade 2000.¹⁴

PISA och IEA-undersökningarna fokuserar på olika aspekter av kunskap. Den ena undersöker elevernas förmåga att omsätta kunskap i olika situationer, den andra i vilken utsträckning eleverna har specifika ämnes-

13. Pettersson, 2018, s. 150–152.

14. Pettersson, 2018, s. 151.

kunskaper. På så sätt kan man säga att de olika undersökningarna kompletterar varandra.¹⁵

Men är PISA och TIMSS relevanta för att mäta elevernas kunskaper och förmågor i den svenska skolan? Här behandlas ämnet matematik och PISA. PISA har formulerat ett ramverk som kan liknas vid den svenska skolans kursplan. En jämförelse mellan PISA:s ramverk och kursplanen och kunskapskraven i matematik har gjorts men även de konkreta provuppgifterna har jämförts för PISA-uppgifterna och de uppgifter som ingår i det nationella provet för årskurs 9.

Innehållsmässigt är det inte så mycket som skiljer PISA:s ramverk från de svenska kursplanerna. Men det finns andra tydliga skillnader. I Sverige ordnas uppgifterna i svårighetsordning, vilket inte PISA gör. I PISA-proven kan den svåraste uppgiften komma först. Ytterligare en skillnad är att eleverna gör PISA-proven endast vid ett tillfälle medan de nationella proven är fördelade på flera tillfällen. Det finns också skillnader i hur elevernas lösningar bedöms. På de nationella proven finns krav på redovisning i cirka hälften av uppgifterna och där kan elevernas lösningar få någon poäng om det bara finns en lösning som påbörjats korrekt. I PISA-proven i matematik är en klar majoritet flervals- och kortsvarsuppgifter och det finns endast ett fåtal uppgifter där mer än svaret ska bedömas. PISA-proven har en betydligt större andel flervalsuppgifter och de nationella proven har en betydligt större andel uppgifter där eleverna ska redovisa sina lösningar. De nationella proven har också en större variation av uppgifter, från huvudräkningsuppgifter till stora problemlösningssuppgifter och muntliga uppgifter. En stor skillnad är att PISA-uppgifterna är betydligt mer texttunga än uppgifterna i de nationella proven.¹⁶

Förklaringen till att det finns skillnader är att de nationella proven och PISA-proven har olika syften. Syftet med de nationella proven är att vara ett stöd för att likvärdigt och rättvist göra bedömningar och sätta betyg på

15. Sollerman, 2019, s. 112.

16. Pettersson, 2018, s. 161–162.

elevernas prestationer. PISA syftar till att undersöka i vilken utsträckning 15-åringar är förberedda för vuxenlivet och rustade för att möta framtidens behov.

PRIM-gruppen, som utvecklar de nationella proven i matematik, har gjort en undersökning, på uppdrag av Skolverket, om relevansen av PISA ur ett svenskt perspektiv. I rapporten till undersökningen konstateras att

PISA är ett relevant instrument, med avseende på ramverk och uppgifter, för att mäta svenska 15-åringars kunskaper i matematik utifrån ramverk och matematiskt innehåll. Alla förmågor och aspekter av matematiken i de svenska styrdokumenterna provas dock inte i PISA.¹⁷

PRIM-gruppen har även gjort motsvarande undersökning för relevansen hos TIMSS-studien där även TIMSS-studien, med avseende på ramverk och uppgifter, bedöms vara ett relevant instrument för att mäta matematik-kunskaper ur ett svenskt perspektiv.¹⁸

Bedömningen av elevernas lösningar på PISA- och TIMSS-uppgifter är mycket detaljerad. De som ska bedöma lösningarna får utbildning om hur bedömningen ska gå till, de får öva sig i att bedöma elevlösningar och man kontrollerar också hur stor överensstämmelsen är mellan bedömare som bedömer samma elevlösningar. Överensstämmelsen måste vara mycket stor. Bedömningsanvisningarna måste vara skrivna på ett sådant sätt att det finns få möjligheter att tolka dessa på olika sätt. Men dessa stränga krav på bedömningen innebär att alla typer av uppgifter inte passar in i proven. Uppgifterna får inte vara för öppna så att flera korrekta svar och lösningar är möjliga. Jag tar PISA som exempel. Varje land får möjlighet att skicka in förslag på uppgifter. Så gjorde också Sverige inför huvudundersökningen 2003. Ett tiotal uppgifter skickades in och några kom med i PISA:s matematikprov. Men en uppgift, som Sverige skickat in, användes som exempel på

17. Skolverket, 2015, s. 35.

18. Sollerman, 2019, s. 111.

en uppgift som inte kan få vara med. Den handlade om att eleverna skulle föreslå hur långa kantlinjerna kan vara i ett akvarium, som har formen av ett rätblock, och som rymmer ett visst antal liter. En liknande uppgift hade varit med i det nationella provet. Men PISA ansåg att till den uppgiften var det svårt att skriva tillräckligt utförliga bedömningsanvisningar eftersom uppgiften kunde ha oändligt många korrekta respektive felaktiga lösningar och svar. Detta är begränsande för matematikinnehållet i PISA och kan därför inte på ett helt trovärdigt sätt representera matematikinnehållet i den svenska kursplanen.

De internationella kunskapsmätningarna har blivit kritiserade. Huvuddelen av kritiken kommer från forskarhåll. Kritiken har handlat om att kunskaper inte kan mätas och att resultaten har fått ett alltför stort inflytande på politiska beslut. PISA har också fått kritik för att den styrs av OECD och inte av forskare som IEA gör. En farhåga som lyfts fram är att om resultaten försämras finns risk att politiker fattar snabba beslut i stället för mer genomarbetade beslut om skolan. Kritiken är viktig så att undersökningarnas styrkor och svagheter synliggörs. Alla undersökningar har sina begränsningar och man måste vara medveten om vad undersökningarna mäter och inte mäter.¹⁹

Vart är vi på väg?

Lärare i folkskolan fick för hundra år sedan själva göra sina bedömningsmaterial och var inte styrda av någon extern bedömning. Lärarna hade i stort sett makten i bedömningsarbetet och eleverna blev föremål för bedömningen, som till att börja med gällde både personliga egenskaper och kunskaper och färdigheter. Efter några decennier blev urvalsfrågan mer akut och man måste ha ett instrument som skulle kunna göra betygen mer jämförbara över landet och som skulle kunna rangordna eleverna och styra lärarens betygsättning på gruppnivå. Då infördes standardprov för

19. Pettersson, 2018, 157.

grundskolan och centrala prov för gymnasieskolan. För trettio år sedan infördes nationella prov som inte skulle styra utan stödja lärarens betygsättning av elevernas prestationer på individnivå. Men undersökningar har visat att betygsättningen inte är likvärdig över landet och Skolverket har därför rekommenderat lärarna att särskilt beakta provbetyget vid betygsättningen. I ett kommentarmaterial²⁰ uppmanas lärarna att vara observanta på avvikelser mellan provbetyg och betyg, särskilt om det är systematiska avvikelser på gruppnivå eftersom dessa avvikelser kan vara tecken på att resultat från de nationella proven inte har särskilt beaktas. Skolverket skriver vidare att det är rimligt att lärare på en skola inför betygsättningen träffas och kalibrerar sina betygskrav utifrån provbetygen. Skolverket menar att bristande likvärdighet i betygen är ett allvarligt problem för skolsystemet och att ansvaret för den bristande nationella likvärdigheten ligger på systemnivå och inte hos lärarkollektivet.²¹ Skolverket har utarbetat förslag på olika modeller för hur betygen ska kunna normeras och alla modeller utnyttjar på något sätt resultaten från de nationella proven.²² Ingen av dessa modeller utgår från normalfördelningen, som var fallet i det relativa systemet.

En fråga som också lyfts är vem som ska bedöma elevernas prestationer på de nationella proven. För närvarande är det oftast elevens lärare som bedömer sina egna elevers prestationer även om sambedömning sker, vilket också rekommenderas av Skolverket. Men diskussioner har förts om huruvida detta förfaringsätt är tillräckligt rättssäkert. Skolinspektionen har under flera år genomfört omrättningar av elevernas prestationer på de nationella proven och har konstaterat att lärarens rättningar och den externa bedömarens i vissa ämnen kan skilja sig mycket åt. Det har föranlett en diskussion om att elevens lärare inte ska rätta sina elevers nationella prov utan att någon annan, oftast en lärare, ska göra det. Men en viktig

20. Skolverket, 2022, s. 41.

21. Skolverket, 2020, s. 3.

22. Skolverket, 2020, s. 6.

fråga är hur elevens egen lärare kan dra nytta av en extern bedömning utan att vara insatt i denna bedömning.²³

Den externa rättningen har sina fördelar, en annan lärare än elevens bedömer elevens visade kunskaper, så att eleven får en bedömning av ytterligare en person av sina visade kunskaper. Men det är också viktigt att elevens lärare får ta del av bedömningen på uppgiftsnivå, för att få kompletterande information om elevens styrkor och utvecklingsområden i sina lösningar. Bedömningen av elevens visade kunskaper på ett nyanserat sätt är svår att göra om endast en poängsumma eller ett provbetyg kommuniceras med läraren. Det är därför viktigt att få en rutin som innebär att läraren får ett så omfattande och varierande underlag som möjligt för sin bedömning samtidigt som eleven inte är utlämnad till en lärares bedömning. Det talar för en rutin som innebär medbedömning, det vill säga där det både är den undervisande läraren och ytterligare en lärare som gör bedömningen. Ännu vet vi inte vad som kommer att beslutas när det gäller extern bedömning.

En diskussion som är ganska intensiv är den om lärarnas arbetsbörda och regering och riksdag försöker minska denna. När det gäller de nationella proven ska minskningen av lärares arbetsbörda göras på tre sätt, förutom extern bedömning är det digitalisering av proven och att göra de nationella proven mindre omfattande som är aktuellt.²⁴

Digitalisering innebär många och stora utmaningar. En grundläggande fråga är hur man kan säkerställa att det viktigaste blir bedömbart och inte motsatsen, det enkelt mätbara.²⁵ En annan fråga är hur de digitala proven ska utformas så att eleverna får största möjliga förutsättningar att visa vad de kan. Det finns en risk att de digitala proven har svårt att hantera uppgifter och lösningar som, för matematikens del, innehåller symboler och figurer och längre beräkningar, och därför kan en digitalisering innebära

en begränsning av såväl innehåll som form. Det som är lätt att digitalisera är uppgifter som bara fordrar korta svar, men dessa uppgifter brukar vara mycket lättbedömda, så att arbetsbördan för lärare kommer bara att minska marginellt. Däremot kan digitalisering innebära att mer dynamiska och interaktiva delar kan införas, vilket fordrar datorkapacitet och gedigen datorkompetens såväl hos elever som lärare. En grundläggande fråga är vilka kompetenser som krävs hos elever för att göra sig själva och sina prestationer rättvisa när de arbetar digitalt respektive på papper. Som vanligt har varje förändring sina möjligheter och begränsningar, detta gäller även när de nationella proven går från att vara analoga till att bli digitala. En förutsättning för införandet av digitala prov är att datorer fungerar väl vid provtillfället och att eleverna har fått tillräcklig övning i att genomföra prov digitalt.

Att göra de nationella proven mindre omfattande är ett risktagande. Om en mindre del av betygskriterierna, förmågorna och det centrala innehållet blir föremål för bedömning finns risken att bedömningen inte blir tillräckligt mätsäker. Det är redan nu svårt att få en tillräckligt mätsäker bedömning då kursplaner och ämnesplaner är omfattande och tolkningsbara vad gäller förmågor, centralt innehåll och kunskapskrav/betygskriterier.

Huruvida dessa tre åtgärder, enskilt eller sammantaget, minskar lärarens arbetsbörda återstår att se. Det som är viktigt är att säkerställa att de olika förändringarna inte äventyrar elevernas möjligheter att visa sina kunskaper på ett så varierat sätt som möjligt.

Detta kapitel har bara tagit upp kunskapsmätningar för betygsättning och för internationell jämförbarhet. Nationella prov och internationella kunskapsmätningar har summativa syften dvs. de används för att bedöma elevens visade prestationer vid några tillfällen och är en bedömning av resultatet av elevens samlade kunskaper. En annan typ av bedömning handlar om att förbättra elevens och lärares lärande, en så kallad bedömning som har formativa syften och är en bedömning som inte ska resultera i ett betyg eller ett annat slutligt omdöme. I den formativa bedömningen är återkopplingen av elevernas visade kunskaper central, en återkoppling

23. Erickson m.fl., 2021, s. 220.

24. Pettersson m.fl., 2019, s. 83–85.

25. Erickson, m.fl., 2021, s. 220.

som visar på elevers styrkor och svagheter och som kan ges av elevens lärare eller av en klasskamrat. En annan viktig aspekt är att eleven själv får bedöma sina kunskaper genom att göra så kallade självvärderingar. Eleven är alltså mycket involverad i den formativa bedömningen och får själv göra sina bedömningar. Det är något som står i stark kontrast till den summativa bedömningen där eleven inte alls är involverad när bedömningen görs. Elevens förmåga att själv bedöma sina resultat framhålls såväl i grundskolans som i gymnasieskolans läroplan. Formuleringarna har varit likalydande under flera år.

I läroplanen för grundskolan²⁶ står som ett av skolans mål att varje elev

utvecklar förmågan att själv bedöma sina resultat och ställa egen och andras bedömning i relation till de egna arbetsprestationerna och förutsättningarna

För gymnasieskolan²⁷ lyder motsvarande formulering

Skolans mål är att varje elev tar ansvar för sitt lärande och sina studieresultat, och kan bedöma sina studieresultat och utvecklingsbehov i förhållande till kraven för utbildningen.

Bedömning i formativt syfte är oerhört viktigt för att utveckla såväl elevernas som lärarnas kompetens i bedömning och behöver stort utrymme i kompetensutvecklingen för lärare och för elever. En grupp lärare som mycket systematiskt använt sig av bedömning i formativ mening kunde konstatera flera positiva konsekvenser av ett formativt arbetssätt för elevers kunskapsutveckling och en slutsats var att det till och med spar tid för läraren.²⁸

26. Skolverket, 2022, s. 18.

27. Skolverket, 2011, s. 10.

28. Pettersson, 2011, s. 14.

Några avslutande reflektioner

De internationella undersökningarna startades med en grupp forskares intresse för att kunna göra jämförelser mellan olika utbildningssystem, men har på senare år haft ett politiskt inflytande som bland annat har influerat kursplanerna. Det är mycket troligt att de internationella kunskapsmätningarna kommer att fortsätta under en lång tid. De kommer att använda slumpmässiga urval, frågor med dominans för frågor som ska besvaras med ett kort svar eller att ange ett alternativ till svar och med strikta bedömningar av elevlösningar. Däremot kommer troligen en förskjutning att ske i innehåll och förmågor efter att de olika ämnena/ämnesområdena utvecklas.

Det som händer med de nationella proven är svårare att sia om, mer än att de troligtvis kommer att finnas kvar. När de nationella proven infördes var tanken att de bara skulle behövas i några år, sedan skulle lärarna vara så förtrogna med bedömningen av elevernas kunskaper att nationella prov inte skulle behövas. Men sett i backspegeln blev utvecklingen en helt annan. De nationella proven har fått större och större betydelse. Lärare och elever är i stort sett nöjda med dem. Ett orosmoln som jag ser det är i vilken utsträckning lärare får ta del av sina elevers lösningar om man övergår till bara extern rättning av proven. Digitaliseringen kan minska lärarnas arbetsbörda och är säkert oproblematiskt för uppgifter där eleverna ska lämna ett alternativ av flera eller bara ge ett kort svar. Det är i och för sig uppgiftstyper som är mycket lätträttade och inte tar särskilt mycket tid för läraren. Det blir troligtvis bara en marginell minskning av lärarens arbetsbörda. Men vad händer med mer öppna uppgifter som exempelvis den tidigare uppgiften om hur många pepparkakor man kan baka av en halv liter pepparkaksdeg? Den fordrar många olika val och beräkningar som eleven måste göra och redovisa. En sådan uppgift ställer stora krav på datorkapacitet och programmering och på elevers kunskaper och färdigheter i datahantering. En fördel med digitaliseringen är att den som ska bedöma

lösningar inte är beroende av elevens handstil. Vad händer med elevernas möjligheter att visa vad de kan om konsekvensen med digitaliseringen blir att allt som enligt kursplanerna ska prövas inte blir prövat. Det finns stor risk att bedömningen av elevernas kunskaper i matematik inte blir trovärdig om inte proven kan representera det innehåll som finns i kursplanerna. En viktig fråga som alltid måste ställas är vad syftet är med digitaliseringen. Det måste väl vara något mer än att minska lärares arbetsbörda? En viktig fråga är att fråga sig vad som enligt kursplanerna ska bedömas. Vilka uppgifter ska konstrueras så att vad-frågan kan besvaras. Därefter kommer frågan hur eleverna ska få visa sina kunskaper, är det bästa sättet digitalt eller med papper och penna? Kanske ska eleverna själva få välja i vilken form de ska redovisa sina uppgifter? På de första nationella proven fick de välja uppgifter i en begränsad skala. Men den tiden är nog förbi. Nationella prov är här för att stanna och det är troligt att de flesta proven blir digitaliserade. Men det måste forskning till för att kunna följa vilka konsekvenser det får för elevernas möjligheter att visa vad de kan.

ASTRID PETTERSSON är professor emerita i pedagogik. Hon har även varit verksam som högstadielärare i matematik, projektledare/vetenskaplig ledare för PRIM-gruppen 1992–2019, föreståndare för den nationella forskarskolan i pedagogisk bedömning 2007–2013 samt ledare för forskningscirkel i bedömning för verksamma lärare. Hon tilldelades 2016 Lärarstiftelsens stipendium för »sina insatser inom forskning ... som på ett framgångsrikt sätt varit till gagn för lärarkåren«.



Foto: Eva Dalin.

KÄLLOR

- Andersson, Håkan, 2002, »Betygen i backspeglarna«, Skolverket (red.), *Att bedöma eller döma. Tio artiklar om bedömning och betygsättning*, s. 153–168.
- Erickson, Gudrun, Palmér, Anne & Pettersson, Astrid, 2021, »Hållbar utformning och användning av de nationella proven«, Åsa Hirsh & Christian Lundahl (red.), *Hållbar bedömning. Bildning, Välbefinnande och utveckling i skolans bedömningsarbete*, s. 201–226.
- Nationalencyklopedin, 1995, *Ordbok A–HZ*. Höganäs: Bra Böcker.
- Pettersson, Astrid, 2010, *Bedömning av kunskap för lärande och undervisning*, Solweig Eklund (red.), i *Bedömning för lärande – en grund för ökat kunnande. Forskning om undervisning och lärande 2010:3*, s. 7–18.
- Pettersson, Astrid, 2011, *Inledning*, Utbildningsförvaltningen, Stockholms stad (red.) *Gensvar för lärande. Rapport med utvecklingsartiklar från en forskningscirkel*, s. 8–15.
- Pettersson, Astrid, 2018, *Vad kan lärare ha för nytta av internationella studier?* Viveca Lindberg, Inger Eriksson & Astrid Pettersson (red.) *Lärares bedömningsarbete. Förutsättningar – Villkor – Agens*, s. 149–171.
- Pettersson, Astrid, Enoksson, Margareta, Evén, Niclas, Kristiansson, Katarina, Nydahl, Anette & Sandström, Heléne, 2019, *Centralt utarbetade prov*, Maria Nordlund & Astrid Pettersson (red.), *Bedömning i matematik – i lärandets och undervisningens tjänst*, s. 73–91.
- Samuelsson, Marielouise, 2022, »Kyrkan lade grunden till svenskarnas läskunnighet«, i *Universitetsläraren*, 2022:3, s. 27–28.
- Skolverket/PRIM-gruppen, 1995, *Matematik A. Nationellt prov. Information till lärare om det nationella provet för kurs A våren 1995*. Stockholm: Skolverket.
- Skolverket, 2011, *Läroplan för gymnasieskolan*, Gy11. Stockholm: Skolverket.
- Skolverket, 2015, *Med fokus på matematik: Analys av samstämmigheten mellan svenska styrdokument och den internationella studien PISA*. Stockholm: Skolverket.
- Skolverket, 2020, *Likvärdiga betyg och meritvärden. Kortversion. Ett kunskapsunderlag om modeller för att främja betygs och meritvärdenas likvärdighet*. Stockholm: Skolverket.
- Skolverket, 2022, *Betyg och prövning. Kommentarer till Skolverkets allmänna råd om betyg och prövning*. Stockholm: Skolverket.
- Skolverket, 2022, *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2022*. Stockholm: Skolverket.
- Sollerman, Samuel, 2018, *Internationella studier*, Maria Nordlund & Astrid Pettersson (red.), *Bedömning i matematik – i lärandets och undervisningens tjänst*, s. 92–115.