

Betyg – deras funktioner och vad de mäter

Alli Klapp, Göteborgs universitet

De allra flesta människor har erfarenhet av betyg. Betyg kan av den enskilda individen ha upplevts som ett rättvist och bra mått på dennes kunskaper medan andra individer kan ha upplevt betyg som ett orättvist och ett dåligt mått på deras kunskaper. Om betyg i ett utbildningssystem används för urval till nästa nivå inom utbildningssystemet som i Sverige, blir betygen av *high-stakes* karaktär. Det betyder att betygen har stor betydelse för den enskilda elevens möjligheter att komma in på vidare utbildning. Ett exempel på detta är att betygen i årskurs nio används för urval till gymnasieskolan. Betygen kan även användas som behörighetsgivande för tillträde till utbildning, vilket betyder att de flesta eftergymnasiala utbildningar kräver ett godkänt betyg i vissa ämnen för att anta en sökande. Betygens jämförbarhet eller likvärdighet är därför en central aspekt eftersom vi använder dem för urval och behörighet. Betygen ska vara jämförbara så att samma nivå på prestationer och kunskaper hos eleverna ska ge samma nivå på betyg över tid, mellan lärare och skolor.

Validitet och reliabilitet är två begrepp som ofta används i diskussionen om betyg och likvärdighet. *Validitet* handlar om hur väl betygen mäter det som betygen är tänkta att mäta, dvs. elevens ämneskunskaper i relation till kunskapskraven i det målrelaterade betygssystemet. *Reliabilitet* handlar om jämnheten i kvaliteten på själva mätningen. Ett större antal mätningar över tid, t.ex. ett större antal uppgifter, ökar möjligheten för läraren att göra en mer korrekt och pålitlig bedömning. Vid urval blir betygens validitet och reliabilitet speciellt viktigt, eftersom betygen påverkar elevernas fortsatta möjligheter till studier och deras livschanser. Den här artikeln diskuterar betygens funktioner, betygens validitet och reliabilitet och vad betygen mäter utifrån forskning och styrdokument.

Varför har vi betyg?

Den grundläggande förutsättningen för att sätta betyg är att elevers kunskaper och prestationer varierar, och att det är möjligt att identifiera och knyta dessa variationer till individer. Elever presterar på olika nivåer, har olika djupa eller olika mycket kunskaper, vilket leder till olika betyg. Sverige har en lång tradition av att använda betyg för att garantera att eleven har en viss nivå av kunskaper och för att fördela platser i högre utbildning (Andersson, 1999). Betygen kan förstås som ett sammanfattande mått, en summativ bedömning, på elevens kunskaper i relation till en viss referens t.ex. kunskapskrav eller i relation till andra elevers prestationer såsom i det tidigare normrelativa betygssystemet som följde en normalfördelningsprincip. Alla typer av betygssystem skapas utifrån den funktion samhället och beslutsfattarna vill att betygen ska ha, och vissa antaganden om nödvändiga tekniska mätkaraktäristika. Betygssystemet speglar samhällets politiska, ideologiska och ekonomiska förutsättningar och olika aktörers uppfattningar om vad som är ett rättvist betygssystem, vilka kunskaper som anses vara viktiga och vad som får elever att prestera.

De målrelaterade betygens funktioner – urval, information och motivation

Betygen i den svenska skolan har flera uttalade funktioner:

- för att göra ett urval
- för att förmedla information
- för att motivera eleverna att prestera.

Att göra ett urval är den viktigaste funktionen. Urvalsfunktionen innebär att betygen används som ett selektionsverktyg till nästa nivå inom utbildningssystemet genom att de elever som söker samma utbildningsplatser rangordnas efter det meritvärde de har fått. Betygen i årskurs 9 är det dominerande urvalsinstrumentet för fördelning av platser på de nationella gymnasieprogrammen, och i begränsad omfattning förekommer antagning via fri kvot och färdighetsprov (Utbildningsdepartementet, 2010: 2039). Betygens urvalsfunktion har i tidigare studier visats fungera relativt väl (Cliffordson, 2008). Cliffordson (2008) undersökte hur det gick för studenter på högskola- och universitetsutbildningar om de antagits på sina betyg från gymnasieskolan eller på resultat från högskoleprovet. Resultatet visade att de studenter som antagits på sina gymnasiebetyg tog ut examen i större utsträckning, gjorde färre avbrott från utbildningen och tog fler poäng, jämfört med studenter som antagits på sitt resultat från högskoleprovet. Detta förklaras troligen av att betyget är ett bredare och mer allsidigt mått på både kognitiva färdigheter (ämneskunskaper) och socioemotionella kompetenser som uthållighet, flit och motivation. Men det saknas studier som undersökt hur väl betyg satta med betygsskalan som infördes 2011 fungerar som prognos på senare studieresultat.

Att förmedla information om hur det gått för eleverna i skolan innebär att betygen ger en sammanfattande information i efterhand om hur det har gått för eleven. Informationen som betyget ger kommer i slutet av en termin eller ett läsår. I dagens betygssystem ges betygen i form av en bokstav (A–F), vilket i sig ger ganska lite information om vad som har gått bra eller dåligt och vad eleven kan göra för att utveckla sina kunskaper och förbättra sina prestationer. Den information som betygen ger är främst summativ och är i ett retrospektiv - det vill säga att den summerar hur det har gått för eleven i skolan. Betyg är även offentlig information och kan användas för att följa upp elevernas resultat på skol-, kommunal och nationell nivå och kan användas av forskare för att utvärdera skolan och ligga till grund för förändringar och reformer av utbildnings- och betygssystemet.

Att motivera eleverna att prestera utgår från antaganden om att belöningar och bestraffningar är viktiga för att motivera människor att prestera, hämtat från ekonomiska teorier (Klapp, Abrams & Levin, 2017; Prendergast, 1999). En del forskare anser att de har betydelse för de lärandesituationer som uppstår i skolan och menar att elever motiveras av tävling, om eleverna får information om kamraternas prestationer och det skapas en rangordning (Becker & Rosen, 1992; Azmat & Iriberry, 2009). Forskarna inom denna teoribildning menar därför att elever bör tävla om betygen, och betygen ses som en viktig

sporre för elevernas lärande. Men detta antagande är inte i linje med pedagogiska och psykologiska teorier om vad det är som motiverar elever att lära sig mer och att prestera. Inom motivationsforskning undersöks och diskuteras olika kvaliteter av motivation, t.ex. inre, yttre, autonom och kontrollerad motivation (se t.ex. Deci, Koestner & Ryan, 1999; Klapp m.fl. 2017; Vansteenkiste m.fl., 2005). För att lättare kunna förstå och diskutera motivationsbegreppet kan man skilja på inre och yttre motivation. Inre motivation är individens egen drivkraft som utvecklas utifrån att individen vill utveckla sin kompetens för sin egen skull, medan yttre motivation innebär att individen gör något för att få en belöning eller undvika en bestraffning. Undersökningar definierar motivation olika utifrån dessa olika kvaliteter (inre, yttre, autonom och kontrollerad), vilket påverkar resultaten.

I utvärderingen av betyg i årskurs 6 (2017a) har lärare svarat på frågan *Hur upplever du elevernas motivation?* En svårighet med den här typen av fråga är att det inte framkommer vilken aspekt av motivation det handlar om. En viktig fråga är hur elevernas inre motivation påverkas om betygen används som incitament för att motivera eleverna, och i vilken utsträckning det görs. I ett bedömningssystem som använder yttre incitament i stor utsträckning, finns det en risk för att elever utvecklar vissa förhållningssätt till sitt lärande som kallas *performance* och *avoidance* inom motivationsteorierna (se t.ex. Vansteenkiste m.fl., 2005). I korthet innebär *performance* att elever presterar för att visa andra att de har lyckats och *avoidance* att elever presterar för att undvika en bestraffning eller ett misslyckande.

I Skolverkets utvärdering (2017a) framkommer att många elever uppfattar betygen som *high stakes* redan i årskurs 6, trots att de formellt sett inte är det. I utvärderingen framkommer å ena sidan att vissa elever menar att de arbetar hårdare med vetenskapen om att de får betyg, men å andra sidan konstaterar Skolverket att ”den ökade motivationen inte alltid kan likställas med ett ökat kunskapsintresse hos eleverna, utan ibland istället kan vara tecken på en jakt på ett högre betyg, utan att intresset för ämnet eller förståelsen för de kunskaper som betyget ska stå för har ökat.” (s. 79).

Sammanfattningsvis behöver motivationsfunktionen diskuteras mer utifrån forskning om elevernas lärandeprocesser och skolans verksamhet.

Betyg i årskurs 6

Betygen i årskurs 6 har främst informations- och motivationsfunktionerna i fokus. Det finns ingen urvalsfunktion med betygen i årskurs 6, de används inte för att rangordna elevernas resultat i en antagningsprocess, vilket innebär att de är av *low-stakes* karaktär. Betygen i årskurs 6 är inte heller behörighetsgivande, dvs. eleverna behöver inte ha ett visst antal godkända betyg för att få fortsätta till årskurs 7. De har inte heller någon avgörande betydelse för elevernas fortsatta möjlighet att välja utbildningsväg, men på grund av att Sverige har en lång tradition av att använda betyg för urval till högre studier, kan de ändå uppfattas som *high-stakes* av eleverna.

Lärarens förhållningssätt till betyg påverkar rimligtvis hur eleverna reagerar och förhåller sig till betyg. Om läraren pratar om att ”detta måste ni kunna för det kommer på provet och

kommer att påverka era betyg” signalerar läraren att lärandet handlar om att bli bedömd och att få betyg. Om läraren istället pratar om att ”detta är bra om ni lär er för då kommer ni att förstå mer” signalerar läraren att det är bra att lära sig för lärandets egen skull och för att bli kompetent (Klapp, 2015; Vansteenkiste m.fl., 2005). En studie visade att när läraren pratade om lärandet som ett sätt att ”lära för livet” och inte i relation till betyg, lärde sig eleverna mer och presterade bättre. Det gällde även elever som hade utvecklat ett instrumentellt förhållningssätt till sitt lärande dvs. att lära sig för att få en yttre belöning t.ex. betyg (Vansteenkiste m.fl., 2005).

Betygen i årskurs 6 används ofta av huvudmännen för att utvärdera hur det går för eleverna i skolan och indirekt om skolan och lärarna gör ett bra jobb. Betygen i årskurs 6 kan därför bli av *high-stakes* karaktär även för *lärarna* och *rektorerna*. Men för *eleverna* är det slutbetygen i grundskolan (vårterminen i årskurs 9 eller när undervisningen i ett ämne tar slut) som är av stor betydelse, eftersom de används för ansökan till gymnasieskolan.

Vad ska betygssättas och hur: en fråga om validitet och reliabilitet

Betygssättning ska ske i relation till kunskapskraven och läraren ska beakta alla typer av underlag, skriftliga såväl som muntliga och andra vid betygssättning. Ett insamlat underlag som används vid betygssättning ska vara ”brett” och läraren ska titta på det i syfte att göra en allsidig bedömning av elevens kunskaper. I läroplanen står det att:

Lärare ska vid betygssättningen utnyttja all tillgänglig information om elevens kunskaper i förhållande till de nationella kunskapskraven och göra en allsidig bedömning av dessa kunskaper (Skolverket, 2017b).

Vidare står det om lärarens betygsunderlag:

Genom sin utbildning och erfarenhet får lärare en fördjupad kunskap om bedömning och betygssättning. Precis som när det gäller annan myndighetsutövning handlar betygssättning om att samla in relevanta underlag för att därefter fatta beslut. (Skolverket, 2016).

Det är alltså läraren som bestämmer vad som är ett relevant underlag. Det står inte i styrdokumentet att underlaget måste vara av en viss typ, mängd eller skriftligt, bara att elevernas kunskaper ska utvärderas allsidigt. För årskurs 1–5 ska skriftliga individuella utvecklingsplaner (IUP) upprättas men i och med att betygen i årskurs 6 infördes, togs kravet på IUP efterhand bort för årskurs 6 och senare årskurser.

Eftersom det är läraren som avgör vad som är ett relevant underlag, är det viktigt att underlaget är av hög kvalitet och visar på en variation av elevernas prestationer i relation till kunskapskraven. Underlaget utgörs av flera observationer och ”bevis” på elevens kunskaper som i så hög grad som möjligt ska vara relevant och representativt. Dessa bevis kan komma från samtal mellan elever och mellan elever och lärare, olika muntliga praktiker, klassrumsarbete, presentationer och skriftliga arbeten. Underlagen kan skifta för olika elever och det kan vara så att läraren sett vissa kunskaper vid andra tillfällen än när ett visst arbete redovisades eller vid en specifik bedömningssituation. Det är alltså inte heller nödvändigt

att eleven deltagit i en viss bedömningsituation för att kunna få betyg, utan betyget sätts utifrån om eleven har de samlade kunskaper som krävs för ett visst betyg relaterat till kunskapskraven.

Betygen ska vara ett mått på elevens kunskaper ställda i relation till kunskapskraven. Det innebär att personliga karaktärsdrag eller socioemotionella kompetenser såsom elevens engagemang, motivation eller intresse inte ska vägas in i betyget. Ett exempel på hur annat än ämneskunskaper riskerar att vägas in i betyget är om lärare väger in elevers förmåga att lämna in arbeten i tid. Att ett sent inlämnat arbete inte räknas som ett fullgott underlag inför betygssättningen, är en princip som inte har stöd i styrdokumentet. Att läraren har fysisk tillgång till ett visst material innebär dock inte att det är självklart att materialet är praktiskt möjligt att ta i beaktande vid betygssättning. Läraren måste även ha tid att bedöma materialet och väga ihop detta med övrigt betygsunderlag (Skolverket, 2017c). Att elever utvecklar en förmåga att hålla deadlines är rimligt och förmodligen viktigt för elevens studier, men detta är främst en pedagogisk uppgift för läraren att träna eleverna i t.ex. hur de kan utveckla strategier för att klara deadlines.

”När läraren bedömer ett sent inlämnat arbete ska detta ske utifrån samma principer som när ett arbete som lämnas in i tid bedöms” (Skolverket, 2017c).

Ett annat exempel på hur begränsningar i underlagen kan påverka betyget, och leda till att betyget får en låg validitet och reliabilitet, är om läraren i förväg bestämmer vilka nivåer av kunskaper ett prov ska mäta. Det kan till exempel vara ett prov med enbart faktafrågor som relateras till ett visst betyg t.ex. E-frågor. Den kunskapssyn som finns i läroplanen innebär att lärandet är en process där olika aspekter av kunnande löper parallellt och förutsätter varandra, t.ex. att faktakunskaper kräver förståelse, och att analysförmåga kräver samspel mellan förståelse och faktakunskaper (se t.ex. Skolverket, 2017b).

Sammanfattningsvis har detta avsnitt diskuterat några exempel på vad som kan hota betygens validitet och reliabilitet. Det är problematiskt för validiteten om betygen riskerar att mäta annat än det som betygen är tänkta att mäta t.ex. att betygen mäter elevernas personliga karaktärsdrag. Det är problematiskt för reliabiliteten om betyget riskerar att sättas utifrån ett underlag som brister i kvalitet, variation och omfång.

Vad mäter betygen?

Även om betygen i årskurs 6 inte är *high-stakes* på så sätt att de används för urval, är betygens förmåga att spegla elevernas kunskaper på ett rättvist sätt centralt för att betygen ska uppfattas som legitima av elever, föräldrar och lärare. Att betygen mäter det som de är tänkta att mäta är centralt för att betygen ska upplevas rättvisa av elever och av andra aktörer i skolan och samhället. Så - även i en *low-stakes* kontext, såsom betyg i årskurs 6, är betygens validitet central.

Olika problem kan uppstå om elever inte upplever att betygen mäter rätt saker. Om elever exempelvis upplever att betyg sätts utifrån deras flit eller ordningssamhet innebär det en risk för att de satsar sina resurser på att bli bättre på just dessa kompetenser, istället för att

lära sig mer i de olika skolämnena (McMillan, 2003; Murphy, 2000; Pintrich, 2002). Dessutom, om betygen brister i validitet innebär det att likvärdigheten mellan lärare riskerar att brista eftersom olika lärare väger in olika mycket av ”annat”, såsom flit och intresse i betygen. Det finns en mängd forskningsstudier, både svenska och internationella, som visar på brister i betygens validitet, det vill säga att de mäter annat än det som de är avsedda att mäta (Klapp Lekholm, 2008; Thorsén, 2014). Det har beskrivits som att betyg kan vara en ”hodge-podge” dvs. en blandning av elevernas ämneskunskaper och deras personliga egenskaper och socioemotionella kompetenser (Brookhart, 1994; Klapp Lekholm, 2008; Pilcher, 1994). Det kan förstås samtidigt vara svårt att som lärare veta om man väger in sådant som flit eller intresse i betyget, eftersom det är något man ofta gör omedvetet. För lärare är det viktigt att fundera på i vilken utsträckning elevernas ämneskunskaper och elevernas personliga karaktärsdrag påverkar betygssättningen. Att synliggöra för sig själv som lärare vad betygen mäter i termer av ämneskunskaper och socioemotionella kompetenser är angelägna frågor för att lärare ska kunna utveckla den egna betygssättningspraktiken.

Bengt Selghed (2004) skrev en avhandling där han intervjuade ett trettiotal lärare om hur de satte betyg i det dåvarande systemet. Flera av dem beskrev där att de ”sköt från höften” och gick på magkänsla framför allt för de elever som låg på gränsen mellan två olika betygssteg. Internationella forskningsstudier har visat på liknande resultat (se t. ex. Brookhart, 1994 och Rinne, 2014). I dessa studier har det framkommit att lärare inom olika ämnesområden har olika rationaliteter när de betygssätter elevernas kunskaper (Pilcher, 1994). Lärarna i studien som undervisade i matematik och naturorienterade ämnen visade sig kompensera elevernas lägre ämneskunskaper med elevernas personlighet och socioemotionella kompetenser när eleven låg på gränsen till ett högre betyg. Lärare i språk tog istället in elevernas socioemotionella kompetenser under hela lärandeprocessen, inte enbart när det gällde att ta beslut om betyg för elever som ligger på gränsen till ett högre betyg. Dessa resultat kan vara intressanta att fundera över i relation till en undersökning genomförd av Skolverket (2016) som visade att 88 procent av de tillfrågade lärarna på grundskolan menade att de lyckats ”bra” eller ”mycket bra” med att sätta rättvisa betyg. Samtidigt är det förstås intressant att fundera över hur man kan veta i vilken grad de betyg man sätter är mer eller mindre rättvisa?

Det har även visats sig att skolornas sammansättning av elever kan påverka lärarnas betygssättningspraktik. I en studie genomförd av Klapp (2011) undersöktes hur skolans storlek, elevsammansättning på skolan och lärarnas undervisningserfarenhet påverkade lärarnas betygssättning i årskurs 9. Resultatet visade att på skolor med stor andel elever med lågutbildade föräldrar satte lärare betyg i större grad utifrån elevernas socioemotionella kompetenser såsom intresse, medan lärare på skolor med stor andel elever med högutbildade föräldrar, i större grad satte betyg utifrån elevernas ämneskunskaper. Denna kompensatoriska betygssättningspraktik som förekom på skolor med stor andel lågpresterande elever, kan bero på att lärarna använder elevernas socioemotionella kompetenser, som exempelvis deras intresse och motivation, för att kompensera för

elevernas brister i ämneskunskaper. Detta kan förstås som att lärare befinner sig i en etiskt svår situation när det gäller betygssättning i skolor som har elever med sämre förutsättningar. Lärare kan fundera på i vilken utsträckning en kompensatorisk betygssättningspraktik förekommer på den egna skolan och vilka konsekvenser som kan uppkomma, t.ex. för elevernas motivation och lärande och för lärarnas arbete. Förekommer ”snällbetyg” eller bedrivs en mer strikt betygssättning på skolan?

I två studier undersökte jag och Christina Cliffordson betygens validitet dvs. vad betygen mäter i termer av elevers ämneskunskaper och deras socioemotionella kompetenser (Klapp Lekholm & Cliffordson, 2008; 2009). Socioemotionella kompetenser kan vara elevernas visade intresse för att lära sig mer, motivation och samarbetsförmåga. Genom att använda elevers resultat på de nationella proven och betyg i årskurs 9 i ämnena svenska, engelska och matematik, genomfördes två studier där det var möjligt att skilja ut vad betygen mätte för typ av kunskaper och kompetenser. De nationella proven förutsätts till största delen mäta ämneskunskaper, medan betygen förutsätts vara ett mått på både ämneskunskaper och socioemotionella kompetenser.

Resultaten för den första av dessa två studier (Klapp Lekholm & Cliffordson, 2008) visade att betygen till största delen mäter elevernas ämneskunskaper. Men det finns en gemensam dimension i betygen i svenska, engelska och matematik, vilket tyder på att betygen innehåller något som är gemensamt över ämnen och som är något annat än ämneskunskaper, t.ex. olika typer av socioemotionella kompetenser och personliga karaktärdrag. Studien visade även på könsskillnader främst för svenskämnet och i den gemensamma dimensionen. Det innebär att flickor har mer av den gemensamma dimensionen i sina betyg, jämfört med pojkarna (Klapp Lekholm, 2008).

I den andra studien (Klapp Lekholm & Cliffordson, 2009), lades enkätdata till i analysen för att undersöka vad den gemensamma dimensionen kunde innehålla i termer av personliga karaktärdrag eller socioemotionella kompetenser. Resultatet visade att flickornas fördel i den gemensamma dimensionen förklaras av att flickorna visar ett större intresse för att lära sig mer, än pojkarna. Även om de presterar lika på de nationella proven tenderar flickor att få högre betyg än pojkar. Dessa resultat är i linje med internationella forskningsresultat (se t.ex. Murphy, 2000; Voyer & Voyer, 2014). Flickor (i många länder) får högre betyg jämfört med pojkar, vilket kan vara ett tecken på att flickors betyg i större utsträckning sätts utifrån socioemotionella kompetenser, än på deras ämneskunskaper (Klapp Lekholm, 2008; Murphy, 2000).

Gruppskillnader i betyg

Det finns en stor mängd forskning som visar på att olika bakgrundsfaktorer påverkar elevernas betyg, t.ex. social bakgrund och kön. Elevernas socioekonomiska bakgrund har ett stort inflytande på elevernas resultat i skolan och på betygen. Elever med hög socioekonomisk status presterar oftare bättre i skolan och får högre betyg, jämfört med elever med låg socioekonomisk status. Framför allt är det föräldrarnas utbildningsnivå som visats ha betydelse för betygen (Yang, 2003; för översikt se t.ex. Klapp Lekholm, 2008).

Dessa gruppskillnader i föräldrars utbildningsnivå har i forskning visats vara stabila över tid och i olika utbildningssystem. När vi tittar på skolnivå, visar forskning att på skolor med stor andel lågpresterande elever sätts betyg mer utifrån socioemotionella kompetenser (Klapp Lekholm, 2011). Det finns alltså betygssättningspraktiker på skolor som skiljer sig åt beroende på elevernas bakgrund och elevsammansättningen på skolnivå (Klapp Lekholm & Cliffordson, 2008; Klapp Lekholm, 2011). Lärare kan fundera på hur sammansättningen av elever på den egna skolan kan påverka den bedömningspraktik och betygssättningspraktik som råder.

Att det finns ett samband mellan elever som har föräldrar med låg utbildningsnivå och betygssättning av socioemotionella kompetenser tycks delvis kunna förklaras av segregation. Segregationen innebär att elever med liknande förutsättningar samlas på vissa skolor, vilket gör att lärares förväntningar på elevernas prestationer kan göra att lärarens betygssättning kan komma att utgå från just den gruppens prestationsnivåer. Det kan innebära att lärare kommer att använda olika referenspunkter för olika nivåer av prestationer, t.ex. vad lägsta nivån ska vara för ett E.

Skillnader mellan könen återfinns framför allt i betygen och i mindre grad i de nationella proven, vilket kan tyda på att flickors och pojkars ämneskunskaper och socioemotionella kompetenser tillmäts olika stark betydelse för betygen. De existerande könsskillnaderna i betyg kan förklaras av att pojkar och flickor utvecklat olika förhållningssätt till lärande. Utifrån förväntningar från t.ex. föräldrar, syskon och skolan samt de erfarenheter som individen gör under uppväxten, utvecklar individen lärandestrategier som kan påverka de socioemotionella kompetenserna, t.ex. intresset för att lära sig (Murphy, 2000). Detta kan innebära att flickor och pojkar har utvecklat olika uppfattningar om kvinnliga och manliga ämnesdomäner och förhållningssätt till lärande och att deras kompetenser inriktas mot det som samhället förväntar sig att de ska bli duktiga på, t.ex. att flickor förväntas prestera bättre i de språkliga domänerna medan pojkarna förväntas prestera bättre i matematik och teknik. Som lärare är det viktigt att tänka på om man har olika förväntningar på pojkar och flickor, och vilka konsekvenserna kan bli.

Att elevernas bakgrunder och förutsättningar påverkar betygen tyder på att lärare i viss utsträckning använder sig av en kompensatorisk betygssättningspraktik. Att låta betyget mäta annat än ämneskunskaper kan vara ett sätt för lärare att (medvetet eller omedvetet) förhålla sig till det etiska dilemma där grupper av elever inte har så goda förutsättningar att nå upp till målen och kunskapskraven. Ämneskunskaper och socioemotionella kompetenser sammanvägs då på olika sätt beroende på elevernas förutsättningar. Detta kan vara problematiskt för betygens validitet och likvärdighet. Ett sätt att arbeta med att höja betygens likvärdighet och validitet är att lärare diskuterar aspekter av validitet, vad som kan hota betygens validitet och hur kvaliteten på lärares bedömningar (reliabilitet) kan utvecklas, t.ex. i fråga om vad som bildar ett betygunderlag.

Olika betygssättningsmodeller

I forskningen diskuteras tre olika övergripande betygssättningsmodeller som visar på tre olika sätt som lärare använder när de sammanväger elevernas kunskaper vid betygssättning (Korp, 2006). Dessa modeller skiljer sig åt på fundamentala sätt och är förenklingar, men de är användbara när man vill diskutera grunderna för betygssättning. De är

- den aritmetiska modellen
- den blandade modellen
- den analytiska modellen.

Den *aritmetiska modellen* innebär att läraren räknar samman elevens resultat på enskilda uppgifter t.ex. läxförhör, inlämningsuppgifter och prov. Ofta använder läraren poäng eller betyg på de enskilda uppgifterna och använder sedan dessa poäng eller betyg för att beräkna en genomsnittspoäng eller ett genomsnittsbetyg. Att lärare tillämpar denna typ av aritmetisk sammanvägning framkommer i olika studier bl. a. i studien *Lärares förtrogenhet med betygssättning* (2016) där de intervjuade lärarna i historia sammanväger betyget på de olika skriftliga betygssatta uppgifter som eleverna gjort (Vallberg Roth m.fl., 2016). Denna modell är problematisk utifrån flera aspekter. Exempelvis undviks med fördel betygsbeteckningar eller betygsliknande omdömen i uppgifter som inte på ett relevant sätt kan knytas till kunskapskraven, t. ex. på grund av att de är mindre omfattande och enskilda. Dessutom riskerar denna modell att bilda ett alltför snävt och ensidigt underlag, eftersom det främst inkluderar skriftliga läxförhör och prov. Vallberg Roth m.fl. (2016) sammanfattar i sin studie att de intervjuade lärarna uppvisade en bedömningspraktik som bestod av traditionella skriftliga prov som ledde till ett alltför ensidigt underlag.

I den *blandade modellen* använder läraren olika typer av underlag, kunskaper och kompetenser utan någon medvetenhet om vilken typ av kunskaper och kompetenser som ligger till grund för betyget. Läraren går på magkänsla och tar ofta in elevernas närvaro, attityder och aktivitet i klassrummet som grund för betyget. Risken med denna typ av sammanvägning till ett betyg är att elevernas personliga karaktärsdrag påverkar betyget, vilket kan leda till att elever inte fokuserar på att utveckla ämneskunskaper. Ett stort problem med denna modell är att betygets validitet kan bli låg och att likvärdigheten brister mellan lärare, eftersom de lägger betydelse vid olika kompetenser och karaktärsdrag hos eleverna.

I den *analytiska modellen* analyserar läraren elevernas kunskaper i detalj i relation till kunskapskraven. Detta innebär en närgående analys av vad eleven har för kunskaper i relation till specifika kunskapskrav. Detta är ett sätt att sätta vissa moment i fokus för att förstå vad eleven kan och vad som brister kunskapsmässigt hos eleven. Det som kan bli problematiskt med denna modell är om små detaljer styr ett ämnesbetyg i alltför stor grad, t.ex. att eleven har missat eller misslyckats med något moment, vilket då riskerar att leda till ett lägre betyg. Eleven kan då uppfatta det som att det inte är lönt att fortsätta lära sig, eftersom eleven har missat en viss detalj och därför inte uppnår ett visst betyg.

Avslutande reflektioner

Det är centralt för betygssättande lärare att känna till betygens funktioner och att betygens validitet och reliabilitet är av stor vikt för att betygen ska uppfattas rättvisa och anses vara legitima mått på elevernas kunskaper. Betygen ska vara ett mått på det som beskrivs i kunskapskraven och inte ett mått på elevernas intresse eller motivation. Lärare kan däremot i sin verksamhet och undervisning arbeta med ett innehåll som intresserar eleverna och även arbeta med att elever utvecklar socioemotionella kompetenser, eftersom forskning visar att dessa kompetenser är av stor vikt för lärandet (Durlak m.fl., 2011).

Olika sätt för läraren att arbeta medvetet med att öka reliabiliteten i bedömningar och höja betygens validitet är att utvärdera och analysera

- de kvaliteter som är synliga i underlagen läraren använder vid betygssättningen
- grunderna för betygssättningen (ämneskunskaper kontra socioemotionella kompetenser)
- skillnader i grunderna för betygssättningen mellan olika elever i klassrummet.

Referenser

- Andersson, H. (1999). *Varför betyg?* Lund: Studentlitteratur.
- Becker, W.E., & Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, 11(2), 107-118.
- Brookhart, S. M. (1994). Teachers' grading: practice and theory. *Applied Measurement in Education*, 7(4), 279-301.
- Cliffordson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment*, 13(1), 56-75.
- Deci, E.L., Koestner, R., & Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 126(6), 627-668.
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., & Schellinger, K.B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82, 405-432.
- Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics*. Academic dissertation. Göteborg: Göteborg studies in educational science 269, Acta Universitatis Gothoburgensis.
- Klapp Lekholm, A., & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: effects of gender and family background. *Educational Research and Evaluation*, 14(2), 181-199.
- Klapp Lekholm, A., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15(1), 1-23.
- Klapp Lekholm, A. (2011). Effects of school characteristics on grades in compulsory school. *Scandinavian Journal of Educational Research*, (6)55, 587-608.
- Klapp, A., Cliffordson, C., & Gustafsson, J-E. (2014). The effect of being graded on later achievement: evidence from 13-year olds in Swedish compulsory school. *Educational Psychology: An international Journal of Experimental Educational Psychology*. DOI: 10.1080/01443410.2014.933176.
- Klapp, A. (2015). Does grading affect educational attainment? A longitudinal study. *Assessment in Education: Principles, Policy and Practice*, 22(3), 302-323.
- Klapp, A. (2015). *Bedömning, betyg och lärande*. Lund: Studentlitteratur.
- Klapp, A. (2015). Betygen i grundskolan – relationen mellan bedömning och elevers senare prestationer. *Pedagogisk Forskning i Sverige*, 19(4-5), s. 367-385.

Klapp, A., Abrams, S., & Levin, H. (2017). *Utbildningsekonomi – om lärandets värde*. Stockholm: Natur & Kultur.

Koch, A., Nafziger, J., & Nielsen Skyt, H. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115, 3-17. Lundahl, C., Hultén, M., Klapp, A., & Mickwitz, L. (2015). Betygens geografi – forskning om betyg och summativa bedömningar i Sverige och internationellt. Stockholm: Vetenskapsrådet.

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decisions: implications for theory and practice. *Educational Measurement: Issues and Practice*, winter, 34-43.

Murphy, P. (2000). Equity, assessment and gender. I Salisbury, J., & Ridell, S. (red.) *Gender, policy and educational change* (s. 134-152). London: Routledge.

Pilcher, J.K. (1994). The value-driven meaning of grades. *Educational Assessment*, 2(1), 69-88.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching and assessing. *Theory into Practice*, 41, 219-225.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economics Literature*. 37(1), 7-63.

Rinne, I. (2014). Pedagogisk takt i betygssamtal. Gothenburg Studies in Educational Sciences 364. Göteborg: Acta Universitatis Gothoburgensis.

Selghed, B. (2004). *Ännu icke godkänt. Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Doktorsavhandling, Malmö Högskola.

Skolverket (2016). Utvärdering av den nya betygsskalan samt kunskapskravens utformning. Stockholm: Skolverket.

Skolverket (2016). Betygsskalan och betygen B och D. Stockholm: Fritzes

Skolverket (2017a). Utvärdering av betyg från årskurs 6. Rapport 451. Stockholm: Skolverket.

Skolverket (2017b). Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011. Hämtad 2017-06-20 på <http://www.skolverket.se/laroplaner-amnen-och-kurser/grundskoleutbildning/grundskola/laroplan>.

Skolverket (2017c). Fråga oss. Hämtad 2017-09-22 på: <https://kundo.se/org/skolverket/d/sent-inlamnat-arbete/>

Skolverket (2017d). PM-Betyg och stöd från årskurs 6 till 9. Hämtad 2017-07-24 på: <https://www.skolverket.se/om-skolverket/publikationer/betyg->

bedomning?_url_=http%3A%2F%2Fwww5.skolverket.se%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FBlob%2Fpdf3782.pdf%3Fk%3D3782

Thorsén, C. (2014). *Dimensionality and predictive validity of school grades: the relative influence of cognitive and social-behavioral aspects*. Academic dissertation. University of Gothenburg: Acta Universitatis Gothoburgensis.

Utbildningsdepartementet 2010:2039. Gymnasieförordningen, hämtad 2017-07-24 på https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/gymnasieforordning-20102039_sfs-2010-2039

Utbildningsdepartementet 2017:175. Om försöksverksamhet med betyg från årskurs 4 i grundskolan, grundsärskolan och sameskolan och från årskurs 5 i specialskolan, hämtad 2017-07-24 på https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/forordning-2017175-om-forsoksverksamhet-med_sfs-2017-175 .

Utbildningsdepartementet (2017). Nationella prov – rättvisa, likvärdiga, digitala, hämtad 2017-08-15 på <http://www.regeringen.se/rattsdokument/lagratsremiss/2017/06/nationella-prov--rattvisa-likvardiga-digitala/>

Vallberg Roth, A-C., Gunnemyr, P., Londos, M., & Lundahl, B. (2016). Lärares förtrogenhet med betygssättning. Malmö: Malmö Högskola.

Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., & Matos, L. (2005). Examining the impact of extrinsic versus intrinsic goal framing and internally controlling versus autonomy-supportive communication style upon early adolescents' academic achievement. *Child Development*, 76, 483–501.

Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*. Advanced online publication. <http://dx.doi.org/10.1037/a0036620>.