

PEDAGOGISK FORSKNING I UPPSALA 157
AUGUSTI 2010

CHRISTIAN LUNDAHL
HENRIK ROMÁN
ULLA RIIS

Tidigt ute med sena betyg – sent ute med tidiga!

Svensk betygspolitik i ljuset av internationell betygsforskning och betygsättningen i Europa

UPPSALA UNIVERSITET
UNIVERSITETSTRYCKERIET
UPPSALA 2010-08-13

ISBN 978-91-506-2143-3
ISSN 0348-3630

Innehåll

<i>Förord</i>	5
<i>Sammanfattning</i>	6
<i>1. Inledning och bakgrund</i>	8
<i>2. Betygssystemet i ett antal europeiska länder</i>	8
2.1 Faktaöversikt	8
2.2 Eurydices indelning: Fyra typer av övergångskrav inom grundskolan	11
2.3 Betygsättningens betydelse för elevernas kunskapsutveckling	13
2.4 Sammanfattning och slutsatser utifrån jämförelsen med andra betygssystem	15
<i>3. Vad säger bedömningsforskningen?</i>	17
3.1 Betyg och lärande – inledning	17
3.2 Betyg som feedback och relationen mellan feedback och lärande	18
3.3 Betyg tydliggör – vad och för vem?	23
3.4 Effekten av lärares pedagogiska bedömningar är stor	23
3.5 Sammanfattning och slutsatser om feedback, betyg och lärande	24
<i>Referenser</i>	25

Förord

Våren 2010 presenterade regeringen betänkandet *Betyg från årskurs 6 i grundskolan* (Ds 2010:15) i vilket två förslag till viktiga förändringar lanseras när det gäller grundskolans betygssystem: Dels införandet av en sexgradig betygsskala, dels införandet av betyg fr.o.m. höstterminen i årskurs 6.

Betänkandet remitterades till bland andra vårt universitet och för Pedagogiska institutionens räkning utarbetade vi ett yttrande. Vi tror att texten är av intresse utanför institutionen och har därför valt att ge ut den i rapportform.

Uppsala fredagen den 13 augusti 2010

Christian Lundahl

Henrik Román

Ulla Riis

Sammanfattning

Under våren 2010 presenterade regeringen betänkandet *Betyg från årskurs 6 i grundskolan* (Ds 2010:15) i vilket två förslag till viktiga förändringar lanseras när det gäller grundskolans betygssystem: Dels införandet av en sexgradig betygsskala, dels införandet av betyg fr.o.m. höstterminen i årskurs 6.

I den här rapporten har vi valt att fokusera på två aspekter: Dels vad vi vet om tidig betygsättning från andra länder, dels vad vi vet om effekterna av betygsättning för lärandet. En viktig och nödvändig funktion hos ett betygssystem är urval, men den lämnas åt sidan i vårt yttrande.¹ Inte heller diskuterar vi frågan om övergång till en sexgradig betygsskala.

Sverige sätter i ett europeiskt perspektiv relativt sena betyg, men det är viktigt att påpeka att den europeiska trenden under 2000-talet är att senarelägga betygsättningen och att komplettera eller ersätta betyg med mer deskriptiva omdömen. I flera länder utöver de skandinaviska sätts inte betyg förrän i grundskolans sekundärstadium (motsvarande det tidigare svenska högstadiet). Regeringens förslag om betyg från årskurs 6 innebär därför å ena sidan en anpassning till en europeisk de facto norm, men är å andra sidan ett brott mot trenden att senarelägga betygen. Denna trend bottnar delvis i internationella forskningsrön, där betygen och dessas pedagogiska effekter ställs mot andra former av bedömning.

Det är svårt att finna forskning som tydligt visar att betyg *i sig* främjar lärandet. Däremot finns det en betydande mängd forskning som pekar mot att betyg är en typ av bedömning som hämmar lärandet och som dessutom håller tillbaka duktiga och kreativa lärare i deras undervisning.

Det finns knappast några belägg för att eleverna inte vet hur de ligger till. Snarare tyder mycket på att de flesta elever som riskerar bli underkända är medvetna om detta.

Vi argumenterar för att läraren har en mycket stor, sannolikt den största, betydelsen för elevernas resultat och kunskapsutveckling. Vi visar också att *feedback* är ett effektivt pedagogiskt redskap för lärare att påverka elevers kunskapsutveckling, men att just *betyg*, enligt de evidensbaserade studier som finns, kan ha en avsevärd negativ effekt för lärandet. Feedback som främjar lärande bör vara frekvent och informationsrik, den bör ge eleven information om inlärningsuppgiften som sådan, om hur lärprocessen fortskridit och om möjliga meta-kognitiva strategier. Betyg innefattar endast indirekt detta.

Ett annat problem i sammanhanget är att de kunskaper samhället är i behov av i hög grad handlar om förmåga och vilja att lösa problem, att tolka komplicerade

¹ Se Lundahl (2006) för en historisk genomgång av de negativa effekter inträdesprövningar ansetts ha på det pedagogiska arbetet.

sammanhang och att kommunicera och ställa nya frågor. Vi visar att utvecklandet av just denna typ av kunskaper antagligen hämmas mest, relativt sett, av betygsättning.

Självklart finns det i det svenska betygssystemet inslag som är centrala för elevernas kunskapsutveckling. Skolans mål och kriterier för framgång måste vara tydliga och delas av elever och föräldrar, om elevernas kunskapsutveckling ska kunna optimeras. Det är viktigt att läroplan och kursplaner är tydliga med vad som förväntas av lärarna och eleverna. Mål och kvalitetskriterier förutsätter dock inte betygsättning och givet de risker för kunskapsutveckling som betygsättning medför föreslår vi att dessa sidor hos betygen hålls isär från varandra. Mycket av argumentationen för tidiga betyg har egentligen varit argumentation för tydliga mål. Det är, menar vi, en helt annan sak.

Regeringen genomför och planerar för närvarande en serie reformer, i vilken förändringar i betygssystemet är ett led. Detta sker i syfte att höja den svenska skolans standard. Mot denna bakgrund vore det, ur ett vetenskapligt perspektiv, mer förnuftigt att i ett första skede satsa på fortbildning av lärares bedömningskompetens. Om denna kompetens blir avsevärt högre än idag är det rimligt att anta att några av de kända negativa effekterna med betyg, särskilt tidiga betyg, kan motverkas. Då skulle det om ett antal år kunna bli aktuellt att se över betygssystemet. Förutsatt att det fortfarande skulle finnas en opinion för det.

1. Inledning och bakgrund

Under våren 2010 presenterade regeringen betänkandet *Betyg från årskurs 6 i grundskolan* (Ds 2010:15) i vilket två förslag till viktiga förändringar lanseras när det gäller grundskolans betygssystem: Dels införandet av en sexgradig betygsskala, dels införandet av betyg fr.o.m. höstterminen i årskurs 6.

I denna rapport diskuterar vi dels vad vi vet om tidig betygsättning från andra länder, dels vad vi vet om effekterna av betygsättning för lärandet. Den första aspekten diskuterar vi utifrån behovet av att nyansera jämförelserna mellan Sverige och andra länder i fråga om tidig betygsättning. Den andra aspekten avser relevanta forskningsresultat rörande betyg och elevkunskaper, särskilt krav på det slags kunskaper som den moderna skolan förväntas utveckla hos eleverna, t.ex. problemlösning och kommunikation, och det vi vet om vad som främjar elevers och lärares arbete med den typen av kunskaper.

En viktig funktion hos ett betygssystem är att ge underlag för urval, i första hand till vidare studier. Denna funktion menar vi är viktig och nödvändig, men eftersom urvalsproblematiken idag i hög grad handlar om övergången från gymnasieskolan till högre utbildning lämnar vi denna funktion åt sidan i vårt yttrande. Inte heller diskuterar vi frågan om övergång till en sexgradig betygsskala, eftersom vi menar att denna fråga är underordnad frågan om betygens betydelse för lärandet.

2. Betygssystemet i ett antal europeiska länder

2.1 Faktaöversikt

I regeringens betänkande finns en översikt över betygssystemen i Norden, som visar att betyg sätts relativt sent – först i åttonde klass – i de nordiska länderna (Sverige, Norge och Danmark), eller åtminstone att den möjligheten finns (Finland, Island). Vidgas blicken till det övriga Europa kan konstateras att Norden hör till dem som sätter betyg senast i grundskolan, men också att tidpunkt och former för betygsättning varierar stort. I ett historiskt perspektiv är betyg den bedömningsform som de flesta länder regelmässigt använt genom hela skolgången. Över tid har dock många europeiska länder successivt ersatt betygen, framför allt i primärskolan, med andra bedömningsformer: muntliga och skriftliga omdömen, utvecklingssamtal och elevrapporter motsvarande de individuella studieplanerna. Frågan är hur den svenska regeringens förslag om betyg från årskurs 6 stämmer överens med denna generella rörelse i Europa att senarelägga betygsättningen? Förslaget innebär ett slags harmonisering med betygspraxis i merparten av Europas länder, men det innebär också att Sverige går emot en allmän europeisk trend att senarelägga betygsättningen, som i många fall grundas på de internationella forskningsrön (se kap. 3). Vi befarar att Sverige genom en ideologisk låsning i betygsfrågan inte tagit tillvara denna forskning tillräckligt.

Nedan följer en översikt över 26 europeiska länder, som för varje land anger när de första betygen sätts, grundskolans längd, samt vid vilka tidpunkter skolan differentieras i skilda skolformer. Länderna är ordnade efter hur tidigt betyg sätts.

	Land:	Betygs sätts från årskurs (elevålder):	Grundskolans längd (antal år och elevernas ålder):	Struktur	Primärskolan slutar efter årskurs (elevålder)	Organisatorisk differentiering inför årskurs (elevålder)
1	Turkiet	åk 1 (6 år)	8 år (6-14)	G		åk 9 (14 år)
2	Italien	åk 1 (6 år)	8-10 år (6-14/16)	<u>PS</u>	åk 5 (11 år)	åk 9 (14 år)
3	Rumänien	åk 1 (6 år)	8-10 år (6-14/16)	<u>PS</u>	åk 4 (10 år)	åk 9 (14 år)
4	Österrike	åk 1 (6 år)	8-9 år (6-14/15)	D	åk 4 (10 år)	åk 5 (10 år), åk 9 (15 år)
5	Ungern	åk 1 (6 år)	8 år (6-14)	G/D	<i>åk 4 (10 år)</i>	åk 5 (10 år), åk 9 (14 år)
6	Lettland	åk 2 (8 år)	9 år (7-16)	G		åk 10 (16 år)
7	Portugal	åk 2 (7 år)	9 år (6-15)	<u>PS</u>	åk 6 (12 år)	åk 9 (15 år)
8	Tyskland	åk 2 (7 år)	10 år (6-16)	D	åk 4 alt 6 (10 år alt 12 år)	åk 5/7 (10/12 år), åk 11 (16 år)
9	Grekland	åk 2 (7 år)	9 år (6-15)	<u>PS</u>	åk 6 (12 år)	åk 10 (16 år)
10	Bulgarien	åk 2 (8 år)	9 år (7-16)	<u>PS</u>	åk 4 (11 år)	åk 8 (14 år)
11	Estland	åk 3 (9 år)	9 år (7-16)	G		åk 10 (16 år)
12	Slovenien	åk 4 (9 år)	9 år (6-15)	G		åk 10 (15 år)
13	Polen	åk 4 (10 år)	9 år (7-16)	<u>PS</u>	åk 6 (13 år)	åk 11 (16 år)
14	Slovakien	åk 5 (10 år)	9-10 år (6-15/16)	G/D	<i>åk 4 (10 år)</i>	åk 5 (10 år), åk 10 (15 år)
15	Litauen	åk 5 (11 år)	9 år (7-16)	<u>PS</u>	åk 4 (11 år)	åk 8 (14 år), åk 9 (15 år)
16	Frankrike	åk 6 (11 år)	9 år (6-15)	<u>PS</u>	åk 5 (11 år)	åk 9 (14 år), åk 10 (15 år)
17	Spanien	åk 7 (12 år)	10 år (6-16)	<u>PS</u>	åk 7 (12 år)	åk 11 (16 år)
18	Irland	åk 7 (12 år)	9 år (6-15)	<u>PS</u>	åk 6 (12 år)	åk 10 (15 år)
19	England	åk 7 (11 år)	11 år (5-16)	<u>PS</u>	åk 6 (11 år)	åk 10 (16 år)
20	Sverige	åk 8 (14 år)	9 år (7-16)	G		åk 10 (16 år)
21	Danmark	åk 8 (14 år)	9 år (7-16)	G		åk 10 (16 år)
22	Norge	åk 8 (13 år)	10 år (6-16)	G		åk 11 (16 år)
23	Finland	valbart 1-7, obl. från åk 8 (14 år)	9 år (7-16)	G		åk 10 (16 år)
24	Tjeckien	Valbart	9 år (6-15)	G/D	<i>åk 5 (11 år)</i>	åk 6 (11 år), åk 10 (15 år)
25	Nederländerna	Valbart	10 år (5-15)	D	åk 7 (11 år)	åk 8 (12 år), åk 12 (16 år)
26	Island	Valbart	10 år (6-16)	G		åk 11 (16 år)
27	Belgien, (Flandern)	Valbart	8-9 år (6-14/15)	<u>PS</u>	åk 6 (12 år)	åk 9 (14 år)

Tabell: Betygsättning och grundskolestruktur i 27 europeiska länder. Tabellen är bearbetad utifrån nationsöversikter hämtade från www.eurydice.org. Förklaring till förkortningarna i femte kolumnen (struktur): **G** står för sammanhållen grundskola, **PS** för uppdelning i primär och sekundärskola och **D** för organisatorisk differentiering efter primärskolan.

Med utgångspunkt i svensk grundskolas historia och den tidigare svenska stadiindelningen som fram till 1982 ledde fram till stadietyg i åk 3 och 6, kan vi se att tio av länderna (från Turkiet till Estland) föreskriver betyg från åk 1-3, att fem länder (från Slovenien till Frankrike) föreskriver betyg från åk 4-6, och att sju länder föreskriver betyg från åk 7-8 (där barnens åldrar dock varierar från 11-14 år). Dessutom har vi fyra länder (från Tjeckien till Belgien) där skolor själva väljer mellan att sätta betyg eller att ge verbala omdömen, och där de verbala omdömena, skriftliga och/eller muntliga, dominerar under grundskolans första del. Till den gruppen kan också Finland föras, eftersom betyg där är valbart under skolåren 1-7.

Översikten visar med andra ord att majoriteten europeiska länder börjar sätta betyg någon gång under de första sex skolåren, och så till vida framgår att Sverige och Norden i ett internationellt perspektiv sätter relativt sena betyg. Detta konstaterande förtjänar dock en närmare analys och nyansering.

En bakomliggande förklaring till variationen i betygsättning är skolsystemens skilda strukturer. En markerad avgränsning mellan primär- och sekundärskola har haft en viss betydelse för ett lands krav på betygsättning. För det första bör man skilja mellan länder med sammanhållna grundskola och de som organisatoriskt delar upp grundskolan i primärskola och lägre sekundärskola (som i flera fall utgör en del av en högre sekundärskola). Norden tillsammans med Estland, Lettland, Slovenien och Turkiet saknar en sådan uppdelning. I de länder där primärskolan utgjort en särskild enhet har detta genererat krav på särskilda övergångsregler. Betyg har då i en del fall fungerat som ett urvalsinstrument, men notabelt är att kraven i flera andra fall riktats mot att skolorna ska upprätta mer utförligt beskrivande bedömningsunderlag, muntliga och skriftliga. Nio av sjutton länder med uppdelade grundskolor sätter skriftliga betyg i primärskolan, åtta gör det inte. I sekundärskolan sätts dock betyg i princip i alla de sjutton länderna. Undantagen är Tjeckien och Nederländerna, där betygen fortsätter att vara valbara.

För det andra kan man skilja mellan de länder som delar upp grundskolan i två delar utan att differentiera den lägre sekundärskolan och de länder som har en sådan organisatorisk differentiering redan efter primärskolan. Som framgår av översikten ovan, så sker differentieringen av eleverna i skilda skolformer (studieförberedande, yrkesförberedande eller andra praktiskt-estetiskt inriktade skolor) först efter grundskolan i de flesta europeiska länderna, även i många av de länder som har en tydlig organisatorisk avgränsning mellan primär- och sekundärskola. Några av de länderna (Litauen, Frankrike, Belgien) har dock en organisatorisk differentiering under de sista två åren i den lägre sekundärskolan. Men därutöver finns de länder där differentiering sker redan efter primärskolan. I vissa av dem (Ungern, Tjeckien och Slovakien) utgör den sammanhållna grundskolan huvudalternativet, medan organisatorisk differentiering utgör huvudalternativet i Tyskland, Österrike och Nederländerna.

2.2 Eurydices indelning: Fyra typer av övergångskrav inom grundskolan

Eurydice heter EU:s informationsnätverk i fråga om de europeiska ländernas utbildningssystem och utbildningspolitik. Det är Eurydices nationsöversikter som huvudsakligen ligger till grund för avsnitt 2 (Eurybase – Descriptions of National Education Systems and Policies). Av dessa informativa översikter framgår att olika länder lägger olika vikt vid att förankra bedömningarnas pedagogiska funktion, och att många av länderna under åren gjort förändringar i riktning mot att utveckla alternativ till betygsättning, i första hand under de primära skolåren, men i vissa fall – Tjeckien och Island – under hela grundskolan. Detta är en trend som också är tydlig bland flera andra europeiska länder. I flera länder ställs exempelvis krav på deskriptiva elevplaner eller deskriptiva avgångscertifikat från grundskolan eller grundskolans olika stadier. Samtidigt finns det många europeiska länder som förefaller hålla fast vid en traditionell bedömningsordning, uppbyggd kring betyg, prov och examina, utan att fästa särskild vikt vid elevens kunskapsprogression eller föräldrarnas delaktighet i barnens skolgång. Skillnaderna är stora mellan hur t.ex. Nederländerna, Tjeckien och Finland motiverar och resonerar kring elevbedömningar jämfört med t.ex. Italien, Rumänien och Turkiet.

Eurydice gör en indelning i fyra typer av skolsystem, utifrån hur olika europeiska länder slussar eleverna vidare i den obligatoriska skolan (European Commission 2009, p. 233f). I princip sammanfaller Eurydices indelning med hur olika länder strukturerar grundskolan. *Den första typen av skolsystem* omfattar tolv länder där den obligatoriska skolan utgör en enhet och där i princip alla elever går vidare ogallrat till slutklassen (klass nio eller tio). Hit räknas alla de nordiska länderna, flera länder med östeuropeisk tradition, såsom Estland, Lettland, Tjeckien, Slovakien, Slovenien och Ungern, samt Turkiet, trots att några av dessa länder har tidig differentiering som alternativ.

Länderna med en sammanhållen skola kan utifrån tidpunkt för betygsättning delas in i olika undergrupper. Sverige, Danmark och Norge utgör som sagt en undergrupp, eftersom betygsättningen där inträder sent: från åttonde skolåret, vilket innebär från 14 års ålder i Sverige och Danmark och från 13 år i Norge. Dessa tre länder har istället föreskrivet andra bedömningsinstrument såsom utvecklingsamtal och individuella studieplaner (förutom i Sverige också i Danmark sedan 2006). Återkommande nationella prov förekommer i alla tre länderna (löpande år 2-8 i Danmark och år 5 och 8 i Norge), och i Norge och Danmark genomförs examensprov i slutet av grundskolan.

Finland, Tjeckien och Island bildar en annan undergrupp eftersom de långt upp i grundskolan har skriftliga eller muntliga (verbala) omdömen som ett valbart alternativ till betyg, vilket innebär att deras skolor från första skolåret (Island) sätter antingen betyg eller ger verbala omdömen. I Finland är betyg obligatoriska från åttonde skolåret, men vanliga från fjärde skolåret. I Tjeckien är betyg i kombination med verbala omdömen vanligt från första klass. Varken Finland eller Tjeckien genomför nationella prov och har inte heller några slutexamensprov. På Island varierar det mellan skolor under hela grundskolan, om man sätter betyg och skriftliga omdömen. Island genomför nationella prov åk 4, 7 och 10, men har inga examensprov.

En tredje undergrupp bildas av systemen i Turkiet, Ungern, Estland, Lettland, Slovenien och Slovakien där eleverna betygsätts tidigt: i Turkiet och Ungern från första klass, i Slovakien får eleverna muntliga graderade betyg fram till femte skolåret och därefter skriftliga betyg, i Slovenien får eleverna skriftliga omdömen fram till fjärde skolåret och därefter betyg, i Lettland betygsätts eleverna från skolår 2 och i Estland från skolår 3. I Turkiet får eleverna ett avgångsbetyg baserat på medelvärdet av betygen år 4-8, och inga nationella prov eller examensprov genomförs. Ungern har inga examensprov, men nationella prov år 4, 6 och 8. I Estland och Lettland förekommer examensprov och nationella prov (år 3, 6 och 9). I Slovakien har nationella prov nyligen införts under år 9 och i Slovenien genomförs under slutåret ett prov som är ett mellanting mellan nationella prov och examensprov.

Den andra typen av skolsystem omfattar länder som har en uppdelning mellan primärskola och lägre sekundärskola, och där kravet för att gå vidare är att man gått igenom primärskolan med tillräckliga kunskaper, manifesterade i någon form av dokumentation – betyg, skriftligt omdöme, utvärderingsrapport och/eller individuell elevplan. Irland, Spanien, Portugal, Italien och Rumänien har alla den typen av system. Hit räknas också Frankrike och Storbritannien även om man i dessa två länder i princip låter alla elever gå vidare till sekundärskolan när de nått den tillämpliga åldern. Gemensamt för Irland, England (som har det största av flera skolsystem i Storbritannien), Frankrike och Spanien är att primärskolan i princip är betygsfri, dvs. betyg sätts först i den lägre sekundärutbildningen som i dessa länder inträder när eleverna är 11 eller 12 år, beroende på när primärskolan börjar. I alla dessa länder börjar barnen skolan när de är 6 år eller yngre.

Irland tillämpar utvecklingssamtal och skriftliga omdömen i primärskolan, som omfattar barn i åldrarna 6-12 år. Dessutom upprättas elevrapporter, ett slags svarighet till individuella utvecklingsplaner. I England gör man skriftliga elevutvärderingar i primärskolan (5-11 år). I Frankrike och Spanien upprättas individuella elevrapporter från första året som följer eleverna genom hela grundskolan. I Spanien (primärskola 6-12 år) ligger dessa rapporter till grund för bedömningen av elevernas övergång till lägre sekundärskola, medan den bedömningen i Frankrike (primärskola 6-11 år) sker på underlag från nationella prov. Betygen i den franska lägre sekundärskolan är för övrigt inte centralt reglerade och kan bestå både av siffror och/eller skriftliga kommentarer. Frankrike har nyligen infört nationella prov i primärskolan, och ett examensprov i slutet av den sekundära primärskolan. Spanien har vissa nationella kunskapsutvärderingar, men inga direkta nationella prov eller examinationer i grundskolan.

I Portugal, Italien och Rumänien däremot betygsätts eleverna redan tidigt i primärskolan. I Portugals nioåriga, stadieindelade grundskola (6-15 år) varvas betyg och skriftliga omdömen från andra skolåret och detta ligger till grund för bedömningen av en elevs möjligheter att gå vidare till nästa stadium. I Italien och Rumänien med 8-årig respektive 9-årig grundskola sätts betyg från första klass och slutbetygen i primärskolan (efter femte respektive fjärde skolåret) ligger till grund för bedömningen av elevens vidare studier. Examensprov under sista skolåret i grundskolan genomförs i Italien och Portugal, medan Rumänien sedan 2006 har nationella prov år 7 och 8.

Den tredje typen av skolsystem omfattar länder där övergången mellan primärskola och lägre sekundärutbildning både bygger på att eleven blivit godkänd i primärskolan och rekommendationer från ett skolråd/klassråd, ofta i samråd med föräldrar och i vissa fall baserat på särskilda prov. Gemensamt för dessa länder är att den lägre sekundärutbildningen är differentierad i olika skolformer. Resultaten i primärskolan blir därför avgörande för vilken skolform eleven kan nå. Den här typen av skolsystem finns i Tyskland, Österrike, Nederländerna och i småländerna Liechtenstein och Luxemburg. I de länder som har tidig differentiering kan vi se två varianter. Antingen föreskrifter om tidiga betyg, som i Tyskland och Österrike, eller valbarhet som i Nederländerna, liksom i tidigare nämnda Ungern, Tjeckien, Slovakien, där tidig differentiering finns som ett mindre vanligt alternativ. Valbarheten kan förklaras med hänvisning till att den organisatoriska differentieringen i dessa länder bottenar i decentraliseringssträvanden och grundläggande principer om skolautonomi – att varje skola ges möjlighet att välja metoder och utvärderingsinstrument anpassade både till lokala eller etniska förhållanden och till utbildningens art. I Tyskland varierar reglerna för de yngre elevernas val av skolform mellan delstaterna, men i princip är det fråga om en förhandling mellan föräldrarna och skolan, utifrån ett bedömningsunderlag om varje enskild elev, där betygen väger tungt. I Österrike är betygen i primärskolan direkt avgörande för elevens fortsatta valmöjligheter. För att få gå vidare till den skolform som leder till högre studier, måste eleven ha mer än godkänt i flera ämnen. Också i Nederländerna sker skolvalet efter primärskolan i samråd mellan föräldrar och skola, baserat på ett individuellt bedömningsunderlag. Nationella prov under sjätte primärskoleåret ska tjäna som stöd för dessa val. Inga av dessa tre länder har några examensprov i den lägre sekundärutbildningen.

Den fjärde typen av skolsystem omfattar länder där villkoret för att gå vidare från primärskola till sekundärskola är att eleven får ett slutcertifikat från primärskolan, oftast grundat på skolarbetet under åren, i något fall i kombination med ett slutprov eller utifrån en bedömning av elevens målpuppfyllelse under sista året i primärskolan. Förutom Belgien hör Bulgarien, Grekland, Polen och Litauen till de länder som har ett sådant uppdelat system. I Polen övergår skriftliga omdömen till betyg från fjärde året och primärskolan avslutas under sjätte året med examensprov som grund för vidare studier. Också Bulgarien avslutar sin fyraåriga primärskola med examensprov. Grekland utfärdar avgångscertifikat grundat på sjätte årets slutbetyg. I Belgien utfärdas primärskolans avgångscertifikat också efter sjätte året, grundat på antingen betyg eller skriftliga omdömen. I Litauen grundas avgångscertifikatet efter fjärde året på skriftliga omdömen.

2.3 Betygsättningens betydelse för elevernas kunskapsutveckling

Ett underliggande antagande bakom regeringens förslag är att tidigare betyg ska påverka de svenska elevernas kunskapsutveckling positivt. Tidigare betyg ger enligt förslaget ”ökade möjligheter till en systematisk uppföljning av elevernas kunskapsutveckling” (Ds 2010:15, s. 9). Vilken betydelse tidpunkt och former för betygsättning har för kunskapsutvecklingen i skolan är emellertid en svårbesvarad fråga, eftersom många andra faktorer spelar in, och det svårt att isolera betygens betydelse. Ett indicium på betygens betydelse skulle dock kunna vara hur länder med skilda betygssystem klarat sig i PISA 2006 (Skolverket 2007). Här följer en sådan jämförelse utifrån Eurydices indelning i fyra typer av skolsystem.

Jämför vi medelvärdet i Sverige med de övriga elva länderna med sammanhållen grundskola i förhållande till de tre områdena läsning, matematik och naturvetenskap, så kan vi konstatera att:

- Finland har ett signifikant bättre resultat i läsning, Estland befinner sig på samma nivå som Sverige, medan övriga nio länder har ett signifikant sämre resultat.
- tre länder (Finland, Estland, Danmark) har signifikant bättre resultat i matematik, två (Island, Slovenien) ligger på samma nivå, medan sex länder har ett signifikant sämre resultat.
- fyra länder (Finland, Estland, Slovenien, Tjeckien) har signifikant bättre resultat i naturvetenskap, ett land (Ungern) ligger på samma nivå medan sex länder har ett signifikant sämre resultat.

Jämför vi sedan Sverige i PISA 2006 med de sju länderna med en markerad, men relativt mjuk övergång mellan primär- och sekundärskola, så kan konstateras att:

- Storbritannien hade signifikant bättre resultat än Sverige i naturvetenskap och att Irland låg på samma nivå, medan de övriga fem länderna hade signifikant sämre resultat.
- Sverige låg i nivå med Irland och Frankrike i matematik, medan övriga fem länder hade signifikant sämre resultat.
- Irland signifikant bättre resultat än Sverige i läsning, medan övriga sex länder hade signifikant sämre resultat.

En PISA-jämförelse mellan Sverige och de tre länderna som tidigt differentierar skolan ger följande resultat:

- Nederländerna har signifikant bättre resultat än Sverige i matematik, medan Tyskland och Österrike ligger på samma nivå som Sverige.
- Sverige ligger på samma nivå som Nederländerna i läsning, men har signifikant bättre resultat än Tyskland och Österrike.
- Nederländerna och Tyskland har signifikant bättre resultat än Sverige i naturvetenskap, medan Österrike ligger på samma nivå som Sverige.

Och slutligen kan vi jämföra Sverige med de fem länder där primärskolans elever får avgångscertifikat, och konstatera att:

- Belgien hade signifikant bättre resultat än Sverige i naturvetenskap, Polen var på samma nivå som Sverige, medan övriga länder hade signifikant sämre resultat.
- Belgien hade signifikant bättre resultat än Sverige i matematik, medan övriga länder hade signifikant sämre resultat.
- Polen och Belgien hade resultat på samma nivå som Sverige i läsning, medan övriga länder har signifikant sämre resultat.

För att sammanfatta PISA: Sverige befinner sig på den övre halvan i Europa, sett till de tre kunskapstesterna. Sju länder placerar sig sammantaget framför Sverige, varav två marginellt. Nitton länder placerar sig efter Sverige, varav fem marginellt. Finland och Estland är de europeiska länder som presterat klart bäst i PISA. I dessa länder sätts betyg relativt tidigt (obligatoriskt i Estland, valbart i Finland). Relativt tidig betygsättning gäller även i fyra av de fem andra länder som sammantaget placerar sig klart eller marginellt framför Sverige – Nederländerna, Belgien (valbart), Slovenien och Tyskland (obligatoriskt), medan det femte landet Irland sätter betyg relativt sent. Bland de fem länder som placerar sig marginellt efter Sverige så sätts tidiga betyg i Österrike, Polen (obligatoriskt) och Tjeckien (valbart), medan England och framför allt Danmark sätter betyg relativt sent.

Tilläggs bör att de länder som bland de 27 jämförda presterat klart sämst i PISA (Italien, Portugal, Grekland, Bulgarien, Turkiet och Rumänien) alla tillämpar tidig betygsättning. Att döma av Eurydices beskrivningar av skolorsystemen är detta en grupp länder som i liten grad uppmärksammat relationen mellan elevbedömningar och elevernas kunskapsprogression. I många av de andra ländernas (t.ex. Nederländerna, Tjeckien, Danmark och Finland) beskrivningar av hur eleverna bedöms, läggs stor vikt vid att motivera olika typer av bedömningar och vad de betyder för elevernas kunskapsutveckling och för kommunikationen mellan skola och hem. I dessa skrivningar betonas återkommande vikten av att lärarnas och skolornas elevbedömningar inte stannar vid prov och sifferbetyg, den arbetsordning som däremot förefaller vara normen i svagpresterande länder som Italien, Turkiet och Rumänien.

2.4 Sammanfattning och slutsatser utifrån jämförelsen med andra betygssystem

Betygsättningen i Sverige sker i ett europeiskt perspektiv relativt sent i grundskolan. I sex av 27 jämförda länder får eleverna betyg först från sjunde skolåret, däribland Sverige. I elva av de 27 länderna får eleverna sina första betyg någon gång under de tre första skolåren, i fem länder får eleverna sina första betyg under åren 4-6 och i fem länder får skolorna från första eller andra året välja mellan betyg och verbala omdömen. Utmärkande för många europeiska länder är dock en strävan att minska betygens betydelse i grundskolan, dels genom att erbjuda alternativ i form av mer deskriptiva, verbala omdömen eller genom att senarelägga betygsättningen, dels genom att komplettera betygsättningen med mer deskriptiva bedömningar.

Regeringens utredare antar att tidigare betyg kommer att gynna elevernas kunskapsutveckling. Vi tror att det är svårt att dra några bestämda slutsatser om betygsättningens betydelse för kunskapsutvecklingen. Vår i och för sig ganska grunda granskning av resultaten i PISA 2006 ger sådana indikationer. Svaren går nämligen i flera riktningar. Det förefaller inte så självklart att förklara Finlands goda resultat med hänvisning till den finska skolans utrymme för tidig betygsättning, med tanke på att många andra länder med tidiga betyg uppvisar anmärkningsvärt svaga resultat. I den mån den nationella bedömningskulturen kan förklara resultaten i PISA, så ligger förmodligen nyckeln till framgång i hur skolorna

i Finland förankrar elevbedömningarna, vare sig skolan sätter betyg eller avger verbala omdömen.

Sammantaget anser vi att regeringens utredare har ett tunt underlag för jämförelse med betygsutvecklingen i andra länder, eftersom endast Nordens länder tas upp i rapporten. Följaktligen tar rapporten heller inte upp det faktum att variationerna i betygsättningen mellan olika länder hänger samman med skolsystemens uppbyggnad och inte heller att förändringarna i Europa går mot senarelagd betygsättning. Regeringens förslag om betyg från årskurs 6 kan visserligen motiveras som en anpassning till en europeisk de facto norm, men innebär på samma gång ett brott mot trenden att senarelägga betygen. Sverige sätter visserligen i ett europeiskt perspektiv relativt sena betyg, men noteras kan att flera länder utöver de skandinaviska inte sätter betyg förrän i grundskolans sekundärstadium (motsvarande det tidigare svenska högstadiet). Den allmänna europeiska trenden under senare decennier är att senarelägga betygsättningen och utveckla alternativ och komplement till betygen. Detta är en trend som bottnar i de internationella forskningsrön vi nu går över till: en forskning som systematiskt ställer betygen och dessas pedagogiska effekter mot andra former av bedömning.

3. Vad säger bedömningsforskningen?

3.1 Betyg och lärande – inledning

Vi stöder oss i det här avsnittet på internationell forskning från flera olika discipliner såsom pedagogik, psykologi, sociologi, företagsekonomi och antropologi. Vi har särskilt utgått från betygsutredningens (Ds 2010:15) formuleringar kring betygs effekter för elevernas resultat. I utredningen hävdas att forskningen kring sådana betygseffekter är mager:

”Det finns få större nationella och internationella studier om betygs effekter på elevernas resultat. Förutom att de är få vilar de också på ett litet empiriskt underlag. Det är därför inte möjligt att dra några entydiga slutsatser utifrån dem.

Att det finns så få studier om betyg och dess effekter på elevernas resultat ser arbetsgruppen som en brist. Ett införande av betyg tidigare i grundskolan och motsvarande skolformer bör därför följas upp. Arbetsgruppen anser att regeringen bör ge Skolverket ett sådant uppdrag.” (Ds 2010:15, s. 40).

Att en förändring i betygssystemet måste följas upp med studier är något vi håller vi med om. Men utredningen gör en alltför snäv avgränsning av den relevanta forskningen. Vi menar att det redan finns gott om relevanta internationella studier att grunda ett politiskt beslut om tidigarelagd betygsättning på. Inte minst finns det forskning som undersöker och jämför sambanden mellan olika slags elevbedömningar och elevers kunskapsutveckling. Låt oss därför se vad den forskningen kan säga om betygs effekter på elevernas resultat.

Eftersom betygs vara eller icke vara är en ganska unik svensk fråga finns inte särskilt mycket forskning om betyg i detta specifika avseende. Däremot är betyg en typ av *feedback*, och *feedback* finns det mycket skrivet om. Betyg är teoretiskt sett att betrakta som en form av summativa bedömningar. *Feedback* i skolan är den information som någon agent (eleven själv, klasskamrater, föräldrar, lärare, skolboken) ger på ett utfört arbete. *Feedback* är konsekvensen av ett handlande som i sin tur får konsekvenser för det fortsatta handlandet (Hattie & Timperley 2007, s. 81). Konsekvenserna betingas av vem som ger vilken typ av *feedback* till vem och i vilken situation. Betyg och andra s.k. summativa omdömen är en av två huvudkategorier av bedömningar, där syftet främst är att sammanfatta ett kunnande i ett meritvärde. Den andra huvudkategorin är formativa bedömningar, vilket är löpande omdömen i syfte att förändra elevens lärstrategier eller lärarens undervisning. Även om summativa omdömen kan användas formativt är det inte där deras huvudstyrka generellt anses ligga (Andrade & Cicek 2009). Det finns faktiskt ganska gott om indikationer på att betyg och betygsliknande omdömen har liten eller nästan ingen positiv effekt på lärandet eller t.o.m. negativ effekt. *Feedback* bör vara frekvent och informationsrik om den ska gagna lärandet. Betyg ges sällan och är en abstraktion som vilar på iakttagelser av många händelser och olika typer av prestationer. Betyg innehåller ingen konkret information om vad eleven behöver arbeta vidare med. Elever som exempelvis har problem med läsning i svenska behöver i första hand inte träna på att skriva. Betyget är en sammanvägning av elevens förmåga att såväl läsa, skriva som tala och upplyser inte om ens så pass enkla skillnader som mellan språkämnenas tre färdigheter. Det

finns heller inga belägg för att betyg skulle kunna vara en feedback som i sig leder till ökad ansträngning och motivation, något som ofta anförs i den svenska debatten (jfr Kluger & DeNisi 1996, Hattie & Timperley 2007)

I en aktuell, uppmärksamman rapport från IFAU (Sjögren 2010) har författaren använt sig av det faktum att Sveriges kommuner under perioden 1969-1982 skulle avskaffa betygen i grundskolans årskurs 3 och 6. Kommunerna avgjorde själva när under perioden som förändringen skulle träda i kraft och detta genererade ett naturligt experiment. De undersökta var individer födda 1954-1974. Beroendevariabler var utbildningslängd efter grundskola och årsinkomst som vuxen. Sjögren noterar ett samband mellan tidiga betyg och sannolikheten att genomgå (delar av eller fullständig) gymnasieutbildning. Såväl söner som döttrar till lågutbildade föräldrar missgynnades av att inte få betyg i den meningen att sannolikheten för att genomgå gymnasiet minskade. Söner till högutbildade föräldrar gynnades av att slippa betyg både vad gäller utbildningslängd och inkomst. Sjögren gör inte anspråk på att kunna förklara vad som ligger bakom dessa samband men hon pekar ut några riktningar för fortsatt sökande. Hon diskuterar inledningsvis utifrån befintlig (pedagogisk) forskning betygens effekter i termer av motivation och information. Hon menar att betygen kan ha haft motiverande effekter särskilt för ungdomar från utbildningssvaga grupper. De kvartssamtal som ersatte betygen kan ha inneburit tolkningssvårigheter för föräldrar med kort utbildning men knappast för dem med en längre utbildning, enligt Sjögren. Hennes resultat är vederhäftiga men frågan är vad de säger om situationen idag, vilket Sjögren också själv påpekar. Under 1970-talet var övergångsfrekvensen från grundskola till gymnasieskola lägre än idag, när praktiskt taget alla bereds plats, och på 70-talet gällde fortfarande att en del unga inte antogs till gymnasiet beroende på för låga betyg.

Vi återkommer till frågor om motivation och information. Innan så sker är det emellertid också värt att notera skillnaden mellan ett relativt betygssystem och ett målrelaterat. Det senare gäller idag. De kohorter som Sjögren studerat hade betygssatts enligt ett relativt system. I det relativa betygssystemet kunde studieframgång till del bero på att andra misslyckades. I relation till modern bedömningsforskning är betyg, som vi strax ska visa, en typ av feedback som visat sig särskilt negativa för högpresterande elevers lärande. Sjögrens resultat, att vissa svaga grupper gynnades av tidiga betyg, kan alltså vara en effekt av ett relativt betygssystem som missgynnade de starkare elevgrupperna.

3.2 Betyg som feedback och relationen mellan feedback och lärande

Betyg är en tacksam politisk fråga som ofta vädjar till människors personliga erfarenheter och tyckanden. Åsikterna kan därtill ganska lätt dikotomiseras – man är antingen för eller emot. Forskningen har dock visat att verkligheten är mer komplex än så. Ett inom samhällsvetenskap vedertaget sätt att bevisa något utanför de personliga erfarenheterna är att arrangera jämförelser mellan en insats med en annan insats eller ingen insats. Två grupper, experimentgruppen och kontrollgruppen, ska idealt vara matchade så att de liknar varandra i största möjliga mån. Skillnaden i utfall kan då till huvudsak förklaras av att grupperna utsatts för olika slag av påverkan. Utifrån den här typen av experiment går det att dra så kallade evidensbaserade slutsatser, men även en sådan slutsats kan i de flesta fall

inte utesluta att verkligheten är beskaffad på något annat och eller något mer komplicerat vis.

Det finns ett femtontal stora internationella meta-studier där evidensbaserad forskning om summativ och formativ feedback analyseras (Fuchs & Fuchs 1986; Natriello 1987; Crooks 1988; Bangert-Drowns, Kulik, Kulik & Morgan 1991; Dempster 1991, 1992; Elshout-Mohr 1994; Kluger & DeNisi 1996; Black & Wiliam 1998; Nyquist 2003; Brookhart 2004; Allal & Lopez 2005; Köller 2005; Brookhart 2007; Wiliam 2007; Hattie & Timperley 2007; Shute 2008, Wiliam 2009). I den mån det finns bevis för och emot tidig betygsättning är det från studier som dessa och liknande. Sedan finns det givetvis enskilda studier och personliga erfarenheter som lett fram till avvikande uppfattning i frågan men det förändrar inte bevisläget i stort. Slutsatserna från de olika meta-studierna är förhållandevis entydiga. Feedback är generellt sett positivt för lärandet. I runt 40 procent av de observerade fallen har dock feedback negativa effekter. Gemensamt för dessa fall var att den feedback individerna fått antingen var riktad mot individens ego eller mot hennes motivation. För att feedback ska fungera effektivt för lärandet behöver den riktas mot den uppgift som ska utföras och innehålla information om hur uppgiften kan utföras mer effektivt.

Vi kan exemplifiera med några resultat från två av de mest frekvent citerade artiklarna om feedback, Kluger och DeNisi (1996) och Hattie och Timperley (2007). Kluger och DeNisi (1996) analyserade resultaten av 3000 forskningsartiklar och rapporter om effekter av feedback i både utbildning och i arbetslivet. Bland dessa publikationer fanns 131 undersökningar som uppfyllde krav på evidensbaserad forskning, i betydelsen av effektstudier, dvs. studier där effekten av feedback isolerats från effekten av andra faktorer av relevans för resultaten. Totalt genomfördes i dessa undersökningar 607 olika effektstudier baserat på 12 652 deltagare och 23 663 observationer (s. 258).

Kluger och DeNisi visade att feedback oftast hade positiva om än ibland svaga effekter för lärandet, men att i hela 38 procent av observationerna var effekterna negativa. I sin genomgång såg de att feedback kan fokuseras mot tre hierarkiskt ordnade kontrollnivåer i förhållande till en uppgift: 1) lärande, 2) motivation, 3) metanivå (personlighet, extern vinning etc). Ett sätt att uttrycka skillnaden är: 1) Gör så här så utförs *uppgiften* bättre, 2) gör så här så *får* du det bättre, 3) gör så här så *är* du bättre. Det generella mönstret var att när feedback rör sig mot punkt 2 och 3 avtar den positiva effekten. För att feedback i sig ska vara verkningsfull behöver den fokusera på uppgiften och inrymma information om: (a) hur uppgiften utförts, (b) vilket målet är och (c) hur uppgiften kan utföras bättre. När feedback utformas eller uppfattas som kopplat till den egna personen eller i syfte att motivera är effekten av feedback generellt sett negativ, dvs. den kan bidra till att resultaten försämras.

Hattie och Timperley (2007) studerade hur olika typer av feedback kan främja *olika* typer av lärande/kunskap/uppgifter. Utifrån en metaanalys av över 7 000 studier och 13 370 effektmått såg de att feedback är mest effektiv när den ger information om hur ett mål ska nås. Minst effektiv är feedback i form av beröm, bestraffning och belöningar. Den typen av feedback innehåller ingen lärorelaterad information. Feedback som utgår från externt driven motivation, som att ”plugga

för betygen” tycks ha viss effekt på enklare uppgifter, men motverkar lösningen av kognitivt mer komplexa problem. Hattie och Timperley framhåller utifrån sina resultat att feedback som gagnar lärandet bör ske i förhållande till frågor om vad som är målet och var individen befinner sig i relation till målet samt vad som är nästa steg (t.ex. ”gör om uppgiften” eller ”gå vidare”). I relation till dessa tre frågor kan feedback riktas mot fyra olika nivåer: uppgiften, processen, regler/metakognitiva begrepp samt individens ego.

Uppgiftsrelaterad feedback är mest effektiv när den används för att identifiera felaktiga hypoteser, idéer eller fakta, dvs. när den ger information som kan användas också utanför den specifika uppgiften (Holm 1975, 1976). Feedback är mer effektiv i relation till uppgiften när det handlar om att korrigera felaktiga svar utifrån bristande tolkningar men inte när det finns brister i förståelse. Processrelaterad feedback är mest effektiv när den hjälper eleverna att avfärda felaktiga antaganden samt när den ger signaler om var eleverna ska leta efter alternativa svar och strategier. Idealt sett syftar processrelaterad feedback till att åstadkomma en förflyttning av elevens uppmärksamhet från uppgiften till processerna bakom dess lösning. Feedback som syftar till regelkunnande och självreglering handlar om att utveckla studietekniker, ett s.k. metakognitivt tänkande, och då inte minst att själv kunna söka feedback av rätt slag i förhållande till olika uppgifter. Feedback kopplat ego, slutligen, är inte effektiv. Det har t.ex. visat sig att det kan vara bättre att låta bli att berömma en elev än att berömma. När feedback kan kopplas till jaget visar det sig att eleverna undviker att ge sig på krävande problem i rädsla för att inte bli berömda eller att bli kritiserade. Rädsla för att misslyckas är genuint kontraproduktivt för lärandet (Dweck 2006; Black & Wiliam 1998). Feedback på uppgifter eller processer som kombineras med en feedback riktad mot egot (eller som av eleven tolkas som riktad mot egot) är alltså ofta kontraproduktiv. Betyg har just den problematiska effekten och i synnerhet när komplexiteten i kunskaperna ökar. Läroplanernas kunskapsmässiga komplexitet har ökat för varje läroplan sedan grundskolan infördes 1962, varför det också blir allt mer centralt att bedöma elevernas kunskaper med nya metoder.

Bedömning och mätning är situationer barn lär sig av. Så som de oftast blir bedömda, så kommer de också att uppfatta lärande och kunskap, samt sig själva. Bedömning leder till ett meta-lärande (Lundahl 2006). Carol S. Dweck (ex. 1986) frågade ett stort antal barn vad det beror på när det går bra för dem i skolan och när det går dåligt för dem. De svar hon fick kategoriserades i två dimensioner. Den första skillnaden barnen gav uttryck för var om de förklarade sina prestationer utifrån interna eller externa faktorer: jag jobbade hårt vs. läraren tycker om mig. Den andra skillnaden handlade om stabiliteten i prestationen: jag lyckades bra eftersom jag är smart vs. jag lyckades bra just den här gången eftersom jag lade ner mycket arbete. Avgörande för elevers framgång över tid är om de förstår sina resultat utifrån interna och ostabila förhållanden – det vill säga att det är upp till mig själv som elev hur resultaten blir och jag kan göra någonting åt dem. Dweck men också flera andra forskare har bland annat sett att pojkar gärna förklarar sin framgång med interna stabila faktorer som ”begåvning” och motgångar med externa ostabila faktorer som ”otur”, medan flickor betraktar framgång som internt ostabilt, ”hårt arbete” och motgångar som internt stabilt ”bristande begåvning”. Huvudpoängen här är att återkoppling som inte skiljer mellan person och prestation och som möjliggör ranking, tenderar att förstärka känslan av

att både framgång och motgång är relaterad till interna stabila faktorer (förmåga, begåvning).² Det är särskilt betydelsefullt undvika fokus vid prestationer och ranking med yngre barn (Dweck 2006). Genom att undvika den typen av återkoppling och genom att visa på kunskapernas komplexitet, t.ex. genom dialogiska frågor, kan eleverna lättare se och förklara sina resultat i relation till hur de har arbetat och hur de har tänkt.

Hur resultaten redovisas för eleverna (och därmed för föräldrarna) har alltså en stor betydelse för vilka effekterna av en bedömning blir. Problemet är att många av skolans bedömningar tjänar två funktioner samtidigt: Att fungera som stöd för elever och lärare samt att utgöra underlag för styrning och administration (kring t.ex. urval, som vi alltså i allt väsentligt lämnat utanför vårt resonemang). Den förra funktionen bör undvika siffror och rangordning medan den senare förutsätter det. När det gäller stöd är den enklaste varianten att redovisa kvantitativa resultat för eleverna först när de förstått vad de kan göra bättre och hur. Så snart de kvantitativa resultaten finns med är det på dem fokus hamnar och feedbacken används för jämförelser snarare än för vidare lärande (Kluger & DeNisi 1996). Samtalsforskaren Johan Hofvendahl (2006) studerade hur lärare i ett utvecklingssamtal kommenterar obestridligt goda elevprestationer på nationella prov, dvs. resultat där eleven antingen har alla eller nästan rätt. I 80 samtal hittade han 222 sådana instanser. Studiens resultat visade att i åtta fall av tio (eller i 181 fall) utgick läraren från ett bristperspektiv, dvs. det är de obefintliga eller den helt betydelselösa proportionen fel som hamnar i fokus istället för elevens framgång. Denna *red pen mentality*, (även vid riktigt goda resultat) är unik för skolan och strider mot vad man vet om hur feedback får goda effekter. Ett ökat inslag av betyg och nationella prov i grundskolan riskerar givetvis att förstärka den här mentaliteten när den, med beaktande av forskningsrön, borde utmanas. Utvecklingssamtalen präglas negativt av avrapporteringskraven av provresultat och betyg. Istället för att fokusera på elevens utveckling riskerar de förstärka elevernas rädsla för att göra fel i skolan. Utvecklingssamtalen blir lätt allt för ensidigt inriktade på om eleverna klarar sig eller inte. Det finns rättmätiga förväntningar från föräldrarnas sida om tydliga signaler men den tydligheten bör inte innebära en förenkling av skolans uppdrag. Samhällets krav på skolan som dessa uttrycks i nuvarande läroplaner är avsevärt svårare att kommunicera än tidigare. Tidiga betyg, stramare utvecklingssamtal och fler nationella prov kan komma att vilseleda såväl lärare som föräldrar kring vad som faktiskt förväntas av eleverna; framtidens arbetstagare (jfr OECD 2009).

En av de starkaste effektundersökningarna kring skillnaden mellan mer utförlig feedback och betyg är Ruth Butlers (1988) studie av 11-12-åringar som jämfördes utifrån om de fått betyg på ett prov, betyg och kommentarer eller enbart kommentarer. Butler valde ut grupper med både högpresterande och lågpresterande individer. Hon gjorde för- och eftermätningar samt såg till att grupperna var matchade. Eleverna studerades utifrån två olika uppgifter. Den första handlade om ordförståelse och var faktabaserad. Den andra handlade om problemlösning. Den generella effekten efter att elevgrupperna fått två omgångar feedback enligt en av de tre betingelserna var tydlig: I den svagare gruppen nådde eleverna samma resultat kring den svårare uppgiften såväl utifrån kommentar som med betyg. Den negati-

² Se även Säljö 2000, kap. 1 för en genomgång av liknande forskningsresultat.

va effekten hos de duktigare eleverna var dock överväldigande. Då kommentarer enbart hade positiva effekter på båda elevgrupper och betyg kraftigt negativa effekter på de starkare eleverna är den generella effekten av kommentarer större. Kring den enklare uppgiften hade betyg negativ effekt också för de svagare eleverna. Butler såg även att när eleverna både fick betyg och kommentarer tenderade de framför allt att komma ihåg betygen, dvs. den feedback som innehöll minst information om vad de gjort för fel och hur de kunde korrigera för dem. Betygen eliminerade sålunda värdet av den formativa bedömningen. Butlers resultat bekräftas också i en stor metastudie som visar att motivation, lärande och intresse för uppgifter ökar om feedbacken går ut på att skapa tillfredsställelse kring arbetet i sig självt snarare än i externa belöningar (Rawsthorn & Elliot 1999).³

Ett sätt att motverka den negativa effekt som betyg kan ha för elevernas förståelse av varför det gått som det gått är att dela på informationen, dvs. att först arbeta med elevernas felaktiga svar på olika vis innan man ger dem ett slutomdöme (William et al. 2004). Forskare har noterat att lärarna inledningsvis kan ha svårt att få med sig elever och föräldrar på detta, även om dessa brukar ändra uppfattning när de väl ser de positiva effekterna (Black et al. 2003). I ett pågående utvecklingsprojekt i Borås, där man särskilt arbetat med kommentarerna på elevarbeten så att man undviker poäng och betyg, skrev eleverna ned sina reflektioner till lärare efter att ha fått tillbaka proven (se vidare Holmgren 2010). En lärare arbetade med prov med bara kommentarer i ämnet svenska och med både betyg och kommentarer i filosofi. En pojke i gymnasieskolans första år skrev så här efter att ha fått tillbaka sitt filosofiprov:

”Jag har kollat igenom kommentarerna men eftersom betyget står på blir man smått blind av det och fokuserar lite för mycket på det. Så personligen (även om jag mycket möjligt skulle klaga om jag inte får betyget skrivet) skulle vilja att du inte gör det för jag märker att jag lägger mer fokus på kommentarerna och utvecklas bättre om betyget inte står på pappret.”

Typiskt för betyg är, som sagt, att de summerar upp en serie prestationer i slutet av ett moment. Betygen sätts i förhållande till mål att uppnå och när betyget är satt kan det inte revideras förrän tidigast nästa termin. Det kan därför vara svårt att engagera eleverna i en diskussion om vad de konkret behöver arbeta med och hur de ska arbeta med detta. Särskilt gäller det när målen kan anses ”färdigundervisade”. Eleverna kommer inte att behöva visa att de behärskar samma sak en gång till. Beträktade utifrån detta perspektiv fyller betyg dåligt en pedagogisk funktion.

En intressant iakttagelse i Kluger och DeNisis (1996) studie är att även om det inte gått att bevisa att all feedback ger önskade effekter på lärandet så är feedback något attraktivt för de allra flesta människor. Feedback tycks sålunda tillfredsställa vissa psykologiska behov oavsett om den är positiv eller negativ och oavsett dess konsekvenser. Det är med andra ord begripligt att många vill tro att exempelvis betyg och nationella prov är bra för lärandet, när det egentligen kanske är en mer existentiell bekräftelse man efterfrågar, eller en ångest av att inte veta som behöver lindras.

³ I fackspråk talar man om skillnaden mellan *intrinsic* och *extrinsic motivation*.

3.3 Betyg tydliggör – vad och för vem?

I debatten förekommer ofta argumenten att betyg tydliggör skolans förväntningar. Det är egentligen målen och betygskriterierna som åsyftas och vi delar uppfattningen att skolans kriterier för framgång bör vara tydliga och delade av eleverna. Detta är dock inte samma sak som att sätta betyg och som vi pekat på är betygs problem i förhållande till lärandet kopplade till deras knapphet på information, deras otydliga koppling till elevernas ego, och slutligen deras uppsummerande karaktär.

Det hävdas ibland att betygen kommer som en överraskning för eleverna. Det stämmer inte sett till Skolverkets senaste nationella utvärdering av grundskolan. Sju av tio elever, med viss ämnesvariation, anger att de fått samma eller högre betyg än förväntat medan tre av tio anger att de fått lägre än förväntat. Av de elever som inte uppnått målen, och detta är den i sammanhanget viktigaste gruppen, är det tre av 100 (3 procent) som förväntat sig betyget godkänt men inte fått det. Det är givetvis inte bra men kan knappast betraktas som ett stort problem. Det förefaller sålunda som om lärarna är relativt bra på att kommunicera till eleven hur hon eller han ligger till.

3.4 Effekten av lärares pedagogiska bedömningar är stor

Analys av resultaten i PISA-studierna visar att lärarnas betydelse för elevernas resultat är mycket stor. I PISA 2006 är variationen mellan skolor i Sverige relativt sett liten, medan variationen mellan klasser i samma skolor är förhållandevis stor. Detta understryker lärarfaktorns stora tyngd. Det finns, inte oväntat, ett mycket stort antal undersökningar av lärarfaktorns generella betydelse för elevers prestationer (t.ex. Sanders och Rivers 1996; Hanushek, Kain, O'Brien & Rivkin 2005; Rivkin, Hanushek & Kain 2005; Kane, Rockoff & Staiger 2006; Barber & Mourshed 2007). Det är svårt att förklara vad det är hos olika lärare som orsakar skillnader i utfallet för eleverna, men modern pedagogisk forskning lutar alltmer åt att svaret till stor del ligger i frågor om feedback och bedömning, framför allt lärares förmåga att använda formativa bedömningar (Kluger & DeNisi 1996; Black & Wiliam 1998; Brookhart 2004; Allal & Lopez 2005; Köller 2005; Brookhart 2007; Wiliam 2007; Hattie & Timperley 2007; Wiliam 2009; Andrade & Cicek 2009).

I flera experimentella studier har man också konstaterat radikala resultatförbättringar, upp mot 100 procent, när lärare givits möjligheten att *systematiskt och under en längre tid*, arbeta med formativ bedömning (Black et al. 2003; Wiliam et al. 2004; Wiliam & Leahy 2009). Vad dessa studier visar är att lärares, elevers och föräldrarnas hela förhållningssätt till bedömning kan behöva förändras om resultatförbättring ska kunna ske.

Om lärarens kompetens och engagemang är av stor vikt för elevernas prestationer och lärande är det naturligtvis viktigt, som arbetsgruppen föreslår, att satsa på kompetensutveckling avseende bedömning och betygsättning. Självklart måste även sådana frågor behandlas i lärarutbildningen och det bör ske i långt högre grad än idag.

Innan vi lämnar lärarkompetensen och lärarnas arbete bör frågan ställas om det inte går att arbeta med formativ bedömning även när man sätter betyg. Det är klart att det är möjligt men erfarenheterna visar, som vi nämnde ovan, att betyg och betygsliknande omdömen lätt skymmer sikten för det som är värt att lära och som läroplanerna betonar. Andra studier visar också på en bedömningarnas ekonomi. Nationella prov och betyg är tidskrävande verktyg och bidrar till att lärare drar sig för att lägga ytterligare tid till bedömning och då särskilt den formativa bedömningen (Torrance & Pryor 1995; Cowie & Bell 1999; Black 2001). Implementeringsstudier har visat att det kan vara effektivt för lärare att starta i formativa bedömningar för att sedan successivt inkludera bedömningar av summativt slag (se vidare Brookhart 2009).

3.5 Sammanfattning och slutsatser om feedback, betyg och lärande

Det finns en betydande mängd internationell forskning som pekar mot att betyg är en typ av bedömning som hämmar lärandet och som dessutom håller tillbaka duktiga och kreativa lärare i deras undervisning. Det är däremot svårt att finna forskning som tydligt visar att betyg *i sig* främjar lärandet.

Det finns knappast heller några belägg för att betyg behövs av det skälet att eleverna inte vet hur de ligger till. Snarare tyder mycket på att de flesta elever som riskerar att bli underkända är medvetna om detta.

Vi har argumenterat för att läraren har en mycket stor, sannolikt den största, betydelsen för elevernas resultat och kunskapsutveckling. Vi har också visat att *feedback* är ett effektivt pedagogiskt redskap för lärare att påverka elevers kunskapsutveckling, men att just *betyg*, enligt de evidensbaserade studier som finns, kan ha en avsevärd negativ effekt för lärandet. Feedback som främjar lärande bör vara frekvent och informationsrik; den bör ge eleven information om inlärningsuppgiften som sådan, om hur lärprocessen fortskridit och om möjliga meta-kognitiva strategier. Betyg innefattar endast indirekt detta.

Ett annat problem i sammanhanget är att de kunskaper samhället är i behov av i hög grad handlar om förmåga och vilja att lösa problem, att tolka komplicerade sammanhang och att kommunicera och ställa nya frågor. Vi visar att utvecklandet av just denna typ av kunskaper antagligen hämmas mest, relativt sett, av betyg-sättning.

Syftet med de reformer som regeringen nu planerar är att höja den svenska skolans standard. Mot denna bakgrund vore det det, ur ett vetenskapligt perspektiv, mer förnuftigt att i ett första skede satsa på fortbildning av lärares bedömningskompetens. Om denna kompetens blir avsevärt högre är det rimligt att anta att några av de kända negativa effekterna med betyg kan motverkas. Då skulle det om ett antal år kunna bli aktuellt att se över dagens betygssystem.

Referenser

- Allal, L., & Lopez, L. M. (2005). Formative assessment of learning: a review of publications in French. In J. Looney (Ed.), *Formative assessment: improving learning in secondary classrooms* (pp. 241-264). Paris, France: Organisation for Economic Cooperation and Development.
- Andrade, H. L. & Cicek, G. J. (Eds.) (2009). *Handbook of formative assessment*. NY: Routledge.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, **61**(2), 213-238.
- Barber, M. & Mourshed, M. (2007). *How the world's best performing school systems come out on top*. McKinsey & Company.
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, **5**(1), 7-73.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003): *Assessment for learning: putting it into practice*. Maidenhead: Open University Press.
- Brookhart, S. M. (2004). Classroom assessment: tensions and intersections in theory and practice. *Teachers College Record*, **106**(3), 429-458.
- Brookhart, S. M. (2007). Expanding views about formative classroom assessment: a review of the literature. In J. H. McMillan (Ed.), *Formative classroom assessment: theory into practice* (pp. 43-62). New York, NY: Teachers College Press.
- Brookhart, S. M. (2009). Mixing it up. Combining Sources of Classroom Achievement Information for Formative and Summative Purposes. In H. L. Andrade & G. J. Cicek. (Eds.) *Handbook of formative assessment* (pp. 279 – 296). NY: Routledge.
- Butler, R. (1988). Enhancing and Undermining Intrinsic Motivation. *British Journal of Educational Psychology* **58** (1988): 1-14.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, **58**(4), 438-481.
- Dempster, F. N. (1991). Synthesis of research on reviews and tests. *Educational Leadership*, **48**(7), 71-76.
- Dempster, F. N. (1992). Using tests to promote learning: a neglected classroom resource. *Journal of Research and Development in Education*, **25**(4), 213-217.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, **41**, pp. 1040–1048.
- Dweck, C. S. (2006). *Mindset. The New Psychology of Success*. New York: Ballentine Books.
- Elshout-Mohr, M. (1994). Feedback in self-instruction. *European Education*, **26**(2), 58-73.
- European Commission (2009). *Key Data on Education in Europe 2009*.

- Eurybase – Descriptions of National Education Systems and Policies. http://eacea.ec.europa.eu/education/errydice/eurybase_en.php
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation - a meta-analysis. *Exceptional children*, **53**(3), 199-208.
- Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality*. (NBER Working Paper 11154). Cambridge, MA: National Bureau of Economic Research. Retrieved December 6, 2007, from <http://edpro.stanford.edu/hanushek/admin/pages/files/uploads/w11154.pdf>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, **77**(1), 81-112.
- Holm, U. (1975). Criterion information and principle information. Studies of verbal learning. *Linköping Studies in Education, No 4*. Linköping: Department of Education, University of Linköping.
- Holm, U. (1976). Criterion Information and Principle Information. Studies of the Concept Formation Process. *Linköping Studies in Education, Reports, No 3*. Linköping: Department of Education, University of Linköping.
- Holmgren, A. (2010). Lärargrupperns arbete med bedömning för lärande. C. Lundahl & M. Folke-Fichtelius (red.), *Bedömning i och av skolan – praktik, politik, principer*, s. 165–181. Lund: Studentlitteratur.
- Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City*. (NBER Working Paper 12155). Cambridge, MA: National Bureau of Economic Research. Retrieved December 6, 2007, from http://rsss.anu.edu.au/themes/TQConf_Rockoff.pdf
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, **119**(2), 254-284.
- Köller, O. (2005). Formative assessment in classrooms: a review of the empirical German literature. In J. Looney (Ed.), *Formative assessment: improving learning in secondary classrooms* (pp. 265-279). Paris: Organisation for Economic Cooperation and Development.
- Lundahl, C. (2006). *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Arbetsliv i omvandling 2006:8. Stockholm: Arbetslivsinstitutet och Uppsala universitet.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, **22**(2), 155-175.
- Nyquist, J. B. (2003). *The benefits of reconstruing feedback as a larger system of formative assessment: a meta-analysis*. Unpublished Master of Science, Vanderbilt University.
- OECD (2009). *Creating Effective Teaching and Learning Environments*. Paris: Organisation for Economic Cooperation and Development.
- Rawsthorne, L. J., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review*, **3**, 326-344.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, **73**(2), 417–458.

- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, **78**(1), 153-189.
- Sjögren, A. (2010). Betygsatta barn – spelar det någon roll i längden? *IFAU Rapport 2010:8*. Uppsala: Institutet för arbetsmarknadspolitisk utvärdering.
- Skolverket (2007). PISA 2006. *15-åringars förmåga att förstå, tolka och reflektera – naturvetenskap, matematik och läsförståelse*. Rapport 306.
www.skolverket.se/publikationer?id=1760
- Säljö, R. (2000). *Lärande i praktiken. Ett sociokulturellt perspektiv*. Stockholm: Prisma.
- William, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles Policy and Practice*, **11**(1), 49-65.
- William, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning. In F. K. Lester Jr (Ed.), *Second handbook of mathematics teaching and learning* (pp. 1053-1098). Greenwich, CT: Information Age Publishing.
- William, D. (2009). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment*. New York, NY: Taylor & Francis.