



UTBILDNINGSFÖRVALTNINGEN

Ombedömning av skrivdelen i nationella provet i svenska i årskurs 9 och Sv1 vårterminen 2012

Rapportförfattare: Pernilla Lundgren

Projektgrupp: Christer Blomkvist, Inger Willner, Robert Lundh, Sören Bernhardt, Anna-Carin Dalborg, Elisabeth Forsberg Uvemo och Anders Müller.

Innehåll

Inledning	3
Utbildningsförvaltningens tidigare arbete med likvärdig bedömning	3
Tidigare ombedömningsstudier	3
Likvärdighet	5
Genomförande	6
Upplägg och omfattning	6
Praktiskt genomförande	8
Resultat och Diskussion	9
Grundskola	9
Hela urvalet	9
Skolnivå	10
Pojkar/flickor	11
Kommunala/fristående grundskolor	11
Högt och lågt index	12
Gymnasieskola	13
Hela urvalet	13
Pojkar/flickor	14
Högskoleförberedande/yrkesförberedande program	15
Skolnivå	16
Betygsändringar	17
Om avvikelserna	18
Slutsats	19
Rekommendationer för fortsatt arbete för en mer likvärdig bedömning ..	20
Källor	22

Bilaga 1. Bedömningsanvisningar grundskola.

Bilaga 2. Bedömningsanvisningar gymnasieskola.

INLEDNING

Betyg i årskurs 9 och i gymnasieskolan bör självklart vara så likvärdiga som möjligt då de används vid urval till gymnasiestudier och till högre utbildning. Huvudmannen bör se till att skolorna fortlöpande arbetar för att bedömningen och betygssättningen ska vara likvärdig med andra skolor.^{1 2} I Stockholms stads skolplan understryks även att alla elever har rätt till betyg som på ett rättvist sätt speglar elevens kunskapsnivå. I stadens skolplan och i verksamhetsplanen för 2013 rekommenderas att de nationella ämnesproven ska (i normalfallet) rättas av en annan lärare än den som undervisat eleverna.

De ombedömningar av nationella ämnesprov i svenska som genomförts under våren 2012 på uppdrag av utbildningsnämnden har också varit en del i förvaltningens arbete inom området bedömning och betyg. Bakgrunden till försöket var att Skolinspektionens nyligen genomförda ombedömningar på nationell nivå av över 30 000 nationella ämnesprov visade på avvikelser mellan skolornas ursprungliga rättning av elevernas prov och omrättningens resultat, framförallt på uppgifter som krävde längre skriftliga texter. Syftet med försöket är att göra en nulägesanalys och resultatet ska användas för att främja en mer likvärdig bedömning och betygssättning i Stockholms skolor.

Utbildningsförvaltningens tidigare arbete med likvärdig bedömning

Utbildningsförvaltningen har arbetat med olika kompetensutvecklingsinsatser för att utveckla bedömarkompetensen, främst riktat mot grundskolan (i satsningen ”Bedömning och betyg” som pågick 2009-2012). Under flera år har förvaltningen genomfört en undersökning av sambandet mellan grundskolebetyg i årskurs 9 och gymnasiebetyget tre år efter avslutad grundskola. Det så kallade stockholmsprovet i matematik har införts i gymnasieskolan i syfte att utvärdera grundskolans betygsättning i ämnet.

Under åren 2005-2006 pågick ett stort arbete i staden – ”Likvärdig bedömning och betygsättning” med medel från Kompetensfonden. En stor del av både gymnasieskolans och grundskolans lärare och skolledare deltog. Arbetet organiserades via nätverk och nätverksledarseminarier.

Tidigare ombedömningsstudier

Skolinspektionen har under tre års tid ombedömt ett antal nationella prov och redovisar det i rapporten ”Lika för alla”³. Överensstämmelsen är hög för

¹ Skolverket (2011). *Allmänna råd för planering och genomförande av undervisningen*.

² Skolverket (2012). *Allmänna råd för bedömning och betygsättning i gymnasieskolan*.

³ Skolinspektionen (2012). *Lika för alla?*

majoriteten av alla prov som granskats och elevernas lärare och Skolinspektionens ombedömare gör för flertalet prov i hög grad samma bedömning.

För vissa delprov finns dock större avvikelser. Det gäller framförallt de delprov där elevens svar ska vara i form av en längre redogörelse, och särskilt för uppsatsskrivning.

På uppsatsdelen, som denna studie undersöker, var bedömaröverensstämmelsen i Skolinspektionens studie (2012) 56 procent för grundskolan och 55 procent för gymnasieskolan (Sv B). I Skolinspektionens studie var de tränade ”ombedömarena” oeniga i 30-40 procent av de fall när de bedömde samma prov (opublicerat). Alltså var skillnaden mellan ursprungs rättarna och ombedömarena inte så mycket större än inom gruppen av ombedömare, vilket visar på svårigheten att bedöma mer komplexa uppgifter.

Även Skolverket har låtit göra egna studier av bedömaröverensstämmelsen som visat att överensstämmelsen när det gäller skrivuppgifter i svenska var 54 procent.

⁴ Beroende på hur tränade bedömare var och uppgiftens komplexitet varierade överensstämmelsen kraftigt.

I en annan studie på just uppsatser i svenska från åk 9 som bedömts av fyra lärare (inklusive ursprungs rättaren) var majoriteten eniga i 35 procent av fallen (fyra gradig skala).^{5,6} För de fall där en eller flera bedömare avvek var det vanligare att en extern bedömare avviker från majoriteten än att den egna läraren gör det

”Resultatet skulle kunna vara en indikation på att felfrekvensen vid bedömning inte nödvändigtvis behöver vara större för en undervisande lärare än vad den skulle kunna vara för en extern bedömare. Den undervisande lärarens bedömning som grundar sig inte enbart på en fristående text utan även på kunskap om skribenten kan eventuellt vara minst lika rättvisande som en extern bedömares förmåga att bedöma en elevtext.”

I Skolinspektionens omrättning var de negativa avvikelserna vanligare än positiva. Några orsaksförklaringar som ges är dels att flera av de anställda omrättarna uppgav att de följer bedömninganvisningarna striktare som omrättare än som lärare. Dessutom kan lärare omedvetet väga in andra faktorer vid bedömning av de

⁴ Skolverket (2012). *Nationella prov – Skolverkets synpunkter i anledning av Skolinspektionens omrättning.*

⁵ Skolverket (2008). *Central rättning av nationella prov.*

⁶ Ciolek Laerum, Beatrice (2009). *Elever skriver och lärare bedömer - en studie av elevtexter i åk 9.* (FUMS Rapport nr 226.)

egna elevernas prov^{7, 8}. En strängare bedömning behöver inte betyda att omrättarna gör en mer korrekt bedömning än lärarna. I ett antal uppsatsdelprov i ämnesprovet i svenska för årskurs 9 gjorde Skolinspektionen en tredje rättning av prov där det varit stora skillnader mellan lärarens ursprungliga bedömning och omrättarens. Den tredje rättningen, som omfattade ca 20 procent av alla prov, fick till resultat att en lika stor andel styrkte den ursprungliga bedömningen som Skolinspektionens första omrättning.⁹

Likvärdighet

Syftet med de nationella proven är enligt Skolverket att:

- stödja en likvärdig och rättvis bedömning och betygssättning.
- ge underlag för en analys av i vilken utsträckning kunskapskraven uppfylls på skolnivå, på huvudmannanivå och på nationell nivå.

I en undersökning som Riksrevisionen gjort uppgav drygt 40 procent av rektorerna att de önskar mer stöd från huvudmannen i arbetet med att nå en likvärdig betygssättning.¹⁰ Det stöd från huvudmannen som önskas gäller framförallt mer samverkan mellan skolor i samma kommun eller koncern för att diskutera betygssättning och bedömning. Ett annat vanligt förekommande svar är att rektorerna önskar mer kompetensutveckling kring dessa frågor. Knappt 60 procent av lärarna har i Riksrevisionens enkät svarat att de önskar mer lokalt stöd från huvudmannen, rektor eller andra kollegor. Det lokala stödet som efterfrågas är framförallt mer av diskussioner och samverkan mellan lärarna på den egna skolan och med lärare på andra skolor. Många av lärarna menar dock att det inte finns tid för sådana ”professionella samtal”.

Skolinspektionen har nyligen genomfört en granskning och tillsynen har visat att skolorna inte vet om deras betygssättning är likvärdig eller inte. I de allra flesta skolorna kritiseras rektorns uppföljning av likvärdigheten i betygssättningen. Ofta omfattar analysen av betygssättningen endast den egna skolan, och ibland sker den endast på elevnivå¹¹.

All forskning, svensk och internationell, visar att man inte uppnår samma överensstämmelse i bedömning av längre produktiva uppgifter som av slutna

⁷ Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics*.

⁸ Klenowski, V. and Wyatt-Smith, C. (2010). *Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform*. Assessment Matters, 2. pp. 107-131.

⁹ Skolinspektionen (2012). *Lika för alla?*

¹⁰ Riksrevisionen (2011). *Lika betyg, lika kunskap? En uppföljning av statens styrning mot en likvärdig betygssättning i grundskolan*.

¹¹ Skolinspektionen (2013). *Tillsyn av bedömning och betygssättning*.

uppgifter, där eleven ska välja ett alternativ eller svara med ett ord eller några meningar. En överensstämmelse på 70-80 procent för uppsatser är mycket hög. I studier av bedömning av uppsatser skrivna på modersmålet ligger oftast samstämmigheten mellan 40 och 70 procent.¹²

Bedömningsanvisningarna för delprov C (skriva) gav enligt 84 procent av lärarna i utprovningen tillräckligt god hjälp för bedömning av elevlösningarna.¹³ En god eller ökad likvärdighet vid bedömning kan fås genom (Jönsson)¹⁴:

1. Välkonstruerade uppgifter.
2. Tydliga bedömningsanvisningar med bedömda elevexempel.
3. Att använda ”måttstocken” för bedömning på ett korrekt sätt.
4. Samedömning.

GENOMFÖRANDE

Utbildningsförvaltningens ledningsgrupp gav uppföljningsenheten inom tillhandahållaravdelningen i uppdrag att leda och samordna försöket under våren och sommaren 2012. I uppdraget angavs att försöket skulle genomföras genom att några utvalda fristående och kommunala grundskolor och gymnasieskolor skulle byta prov med varandra inom respektive skolform.

En projektgrupp tillsattes med representanter för uppföljningsenheten samt för ledningsstaben och grundskoleavdelningen. Sweco Eurofutures AB fick i uppdrag att organisera försöket utifrån de ramar och förutsättningar som projektgruppen angav samt sammanställa resultaten.

En grupp av lärare som haft rollen som ombedömare träffades tillsammans med tre medarbetare på central förvaltning i januari 2013 för att samla in synpunkter och reflektioner.

Upplägg och omfattning

Nationella prov i svenska, delprov C, uppsatsdelen, i grundskolans årskurs 9 (Lpo94) och i gymnasieskolans årskurs 1 (Lgy11), togs ut för försöket.

För grundskolan fanns möjlighet till bedömning av delprovet i tio steg (som kan förstås som: G, G+, G-, VG, VG+, VG-, MVG, MVG+, MVG- samt EUM (ej uppnått målen)). Ombedömarna gjorde däremot sin bedömning i fyra steg: G, VG,

¹² Brennan, L. Robert (red) (2006). *Educational Measurement*, 4th ed., s. 412.

¹³ Skolverket (2012). *Nationella prov – Skolverkets synpunkter i anledning av Skolinspektionens omvärterning*.

¹⁴ Jönsson, A. *Lärande bedömning*. www.larandebedomning.com.

MVG och EUM. En bedömningsmatris fanns med i instruktionen med fem bedömningsaspekter (Bilaga 1). Därutöver fanns bedömda elevexempel som stöd.

De nationella proven för gymnasieskolan relaterade till de nya ämnesplanerna, där betyg ges i sex steg (A-F). I stödet för gymnasieskolan fanns bedömda exempeltexter och en analytisk matris med delkriterier, medan sammanvägningen sedan skedde till ett betyg i relation till en annan mer holistisk matris (Bilaga 2). För gymnasieskolan utelämnades proven i svenska som andraspråk och för grundskolan bedömdes dessa (men ibland på en tredje skola för att ha en behörig lärare).

De nationella proven genomfördes i februari och april 2012 och var alltså genomförda när försöket inleddes.

Varje bedömt prov ombedömdes av lärare från två andra skolor. De ombedömande skolorna var skolor med motsatta förutsättningar jämfört med "ursprungsskolan". Syftet var att få ett bredare underlag för analys. Deltagande skolor valdes därför ut med hänsyn tagen till skolornas förutsättningar avseende innerstad/ytterstad, kommunala/fristående huvudmän, samt i grundskolan även elevsammansättningen utifrån socioekonomiskt index¹⁵ och inom gymnasieskolan högskoleförberedande program/yrkesprogram. Fyra kommunala och två fristående grundskolor och fyra kommunala och tre fristående gymnasieskolor, som representerade de olika förutsättningarna omfattades av försöket (Tabell 1 och 2).

Grundskola	Huvudman	Socioekonomiskt index	Ursprungs rättning
A	Fristående	Högt	Annan lärare på samma skola.
B	Fristående	Lågt	Sambedömning.
C	Kommunal	Högt	Två lärare bytte prov med varandra och man bedömde sedan även själv proven.
D	Kommunal	Lågt	Sambedömning.
E	Kommunal	Lågt	Alla bedömer allas prov översiktligt och sedan har man en gemensam diskussion och sambedömning.
F	Kommunal	Högt	i.u.

Tabell 1. Grundskolor som deltog i försöket .

¹⁵ Indexvärdena bygger på en statistisk sambandsanalys mellan skolors kunskapsresultat och olika socioekonomiska faktorer. Faktorena är: kort tid i Sverige (upp till två år), två till sex år i Sverige och föräldrarnas utbildningsnivå. Ett lågt indexvärde indikerar gynnsamma socioekonomiska bakgrundsfaktorer och ett högt värde indikerar icke gynnsamma bakgrundsfaktorer.

Gymnasieskola	Huvudman	Program	Ursprungs rättning
A	Kommunal	Högskoleförberedande	Lärarna bedömde sina egna och en kollegas uppsatser. En tredje part tillfrågades i några tveksamma fall.
B	Fristående	Yrkesförberedande	i.u.
C	Fristående	Högskoleförberedande	Avidentifierat och undervisande lärare har bedömt. Dessutom har utbyte skett mellan fyra skolor, och sedan har lärarna på den egna skolan diskuterat fall när det avvek mycket.
D	Kommunal	Högskoleförberedande	Avidentifierat och sambedömning i mindre grupper.
E	Kommunal	Yrkesförberedande	i.u.
F	Fristående	Högskoleförberedande	i.u.
G	Kommunal	Yrkesförberedande	i.u.

Tabell 2. Gymnasieskolor som deltog i försöket .

Inom grundskolan deltog en klass, 9 A, per skola i försöket och i gymnasieskolan var femte elev i alla klasser på skolan med undervisning i Sv1. I gymnasieskolan ingick inte proven för elever med undervisning i svenska som andraspråk. Det slutliga antalet prov inom grundskolan var 196 prov och inom gymnasieskolan 157. Under försöket föll några ombedömningar bort bland annat på grund av oklarheter i tolkningen av resultatet, logistiska problem vid insamling och lärare som lämnade sin tjänst.

Praktiskt genomförande

Skolorna skickade i första skedet in sina egna prov märkta med betyg och elevens namn och personnummer samt när det gäller grundskolan även märkta med om eleven hade undervisning i svenska som andraspråk. Alla inkomna prov avidentifierades och kodades genom Swecos försorg. Inom grundskolan ingick en klass, klass 9A, per skola i försöket och i gymnasieskolan var femte elev i alla klasser som skrev nationella provet.

Därefter skickades proven för omdömning till två andra skolor med motsatta förutsättningar. Lärarna antogs använda de av Skolverket framtagna bedömningsanvisningarna¹⁶.

Med anledning av att några av de mottagande grundskolorna saknade lärare med behörighet att bedöma prov från elev med svenska som andraspråk gjordes en av omdömningarna av dessa elevers prov av lärare med sådan kompetens vid en annan skola som inte deltog i försöket.

Efter två omdömningar samlades alla prov in av Sweco för bearbetning och sammanställning.

RESULTAT OCH DISKUSSION

Grundskola

Hela urvalet

För hela urvalet var bedömaröverensstämmelsen 49 procent (Diagram 1), negativa avvikelser (35 procent) var vanligare än positiva (16 procent) och de relativt vanligaste skillnaderna var att ett prov vid omdömning antingen omdömdes från MVG till VG eller från ej uppnått målen till G (data ej visade). I Skolinspektionens omrättning bedömdes 56 procent med samma betyg, 32 procent avvek negativt och 12 procent av proven hade en positiv avvikelse. Mönstret var alltså likartat i vår undersökning, trots ett mycket mer begränsat material.

Möjligen kan en delförklaring vara att upplägget med att proven omdömdes på skolor med motsatta förutsättningar gav förutsättningar för större skillnader, där en negativ avvikelse i princip alltid är mer frekvent. Det begränsade urvalet av prov jämfört med Skolinspektionens omrättningar gör att det är svårt att påvisa någon signifikant skillnad mellan dessa.

¹⁶ Grundskola C har uppgett att de inte använde bedömningsmatrisen från nationella provet utan enbart relaterade till betygskriterierna.

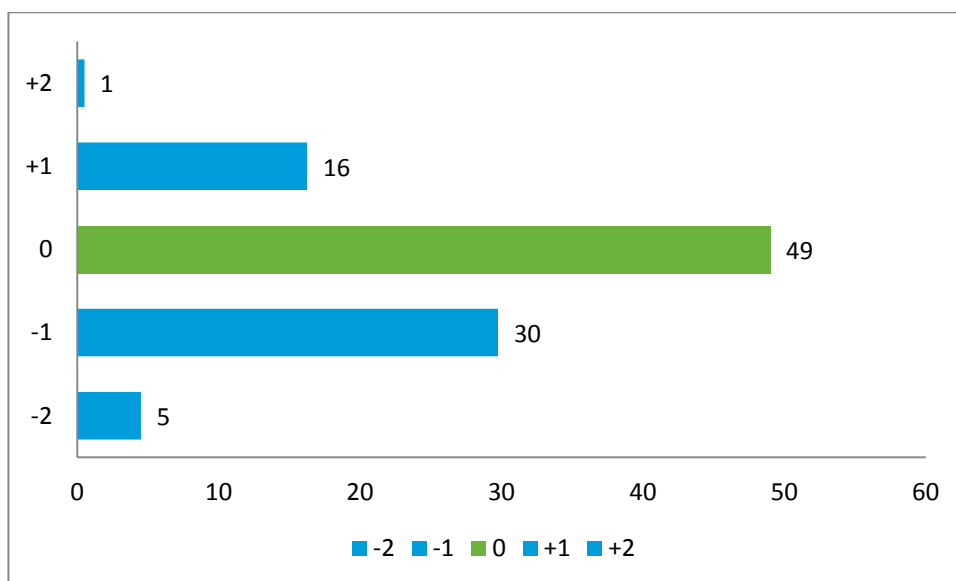


Diagram 1. Grundskolan. Andel i procent av de omrättade proven som avviker olika antal betygssteg .

Skolnivå

Tittar vi på skolnivå så ser vi att bedömaröverensstämmelsen varierar mellan 39 och 60 procent (Diagram 2). I Skolinspektionens omrättning var denna mellan 24 och 77 procent (de mest avvikande var dock skolor med ganska få prov). Negativa avvikelser är mest vanliga, förutom för skola D som uppvisar en större positiv avvikelse (37 procent).

För grundskolan ingick bara en klass per skola i studien och resultatet handlar därmed mer om skillnader mellan olika lärares bedömningar, inte skolors.

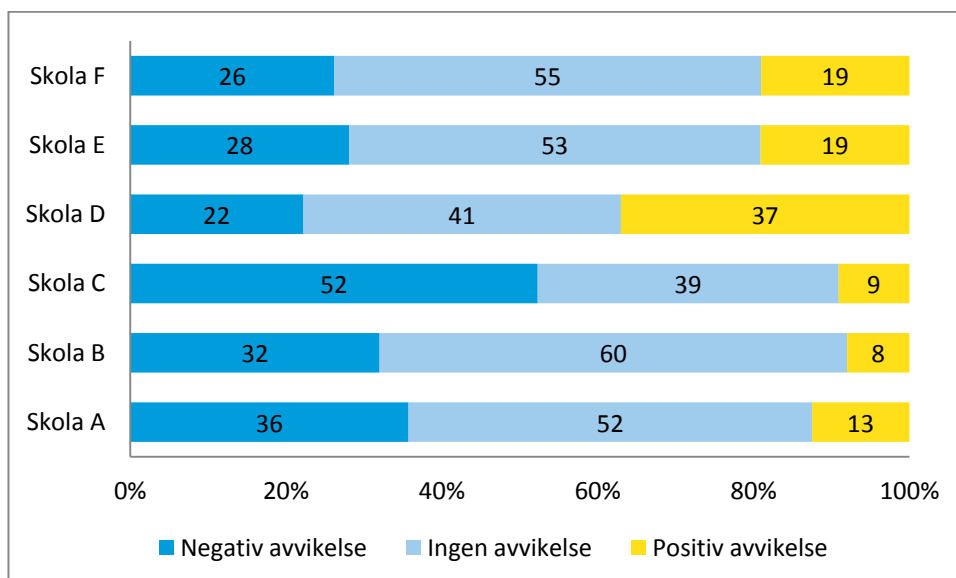


Diagram 2. Grundskolan. Andel i procent av de omrättade proven, per skola, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Pojkar/flickor

Även skillnader mellan pojkar och flickor studerades, och för grundskolan hittades inga signifikanta skillnader (Diagram 3).

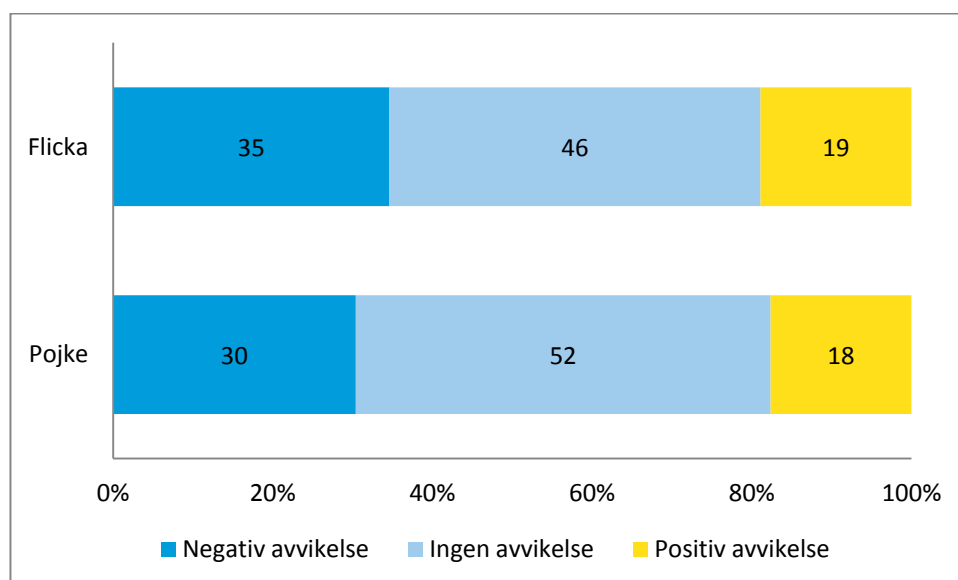


Diagram 3. Grundskolan. Andel i procent av de omrättade proven, uppdelat per kön, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Kommunala/fristående grundskolor

Vid en jämförelse mellan kommunala och fristående grundskolor var ingen eller en negativ avvikelse vanligare bland fristående skolor, och den största skillnaden fanns för positiva avvikelser där 10 procentenheter fler av de kommunala än de fristående gav ett högre betyg vid omrättningen. Detta beror till stor del på att en kommunal grundskola (Skola D) hade många positiva avvikelser och att endast två fristående skolor ingick i studien.

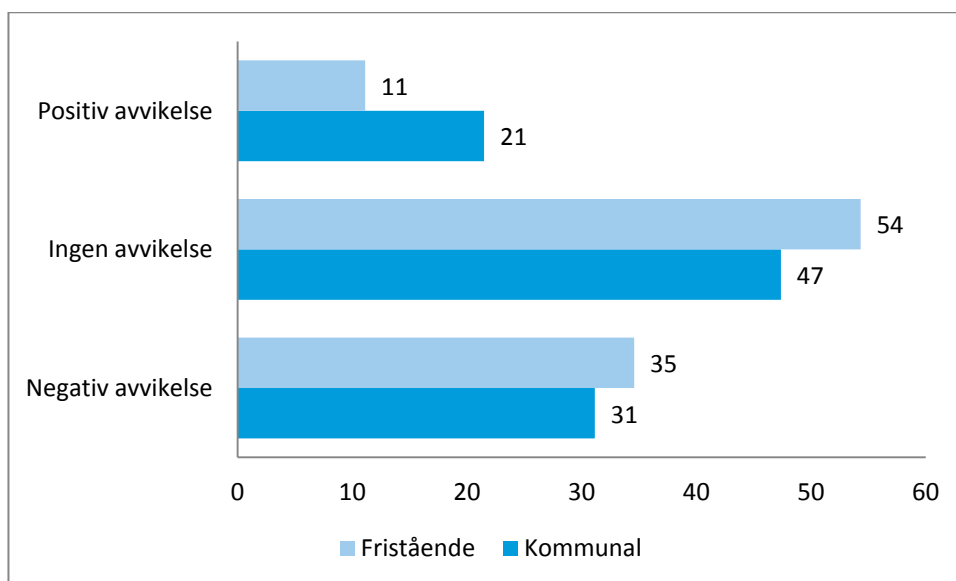


Diagram 4. Grundskolan. Andel i procent av de omrättade proven, uppdelat på kommunala respektive fristående grundskolor, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Högt och lågt index

Inga större skillnader kunde ses om man studerade hur ombedömare från grundskolor med olika socioekonomiskt index bedömde proven (Diagram 5).

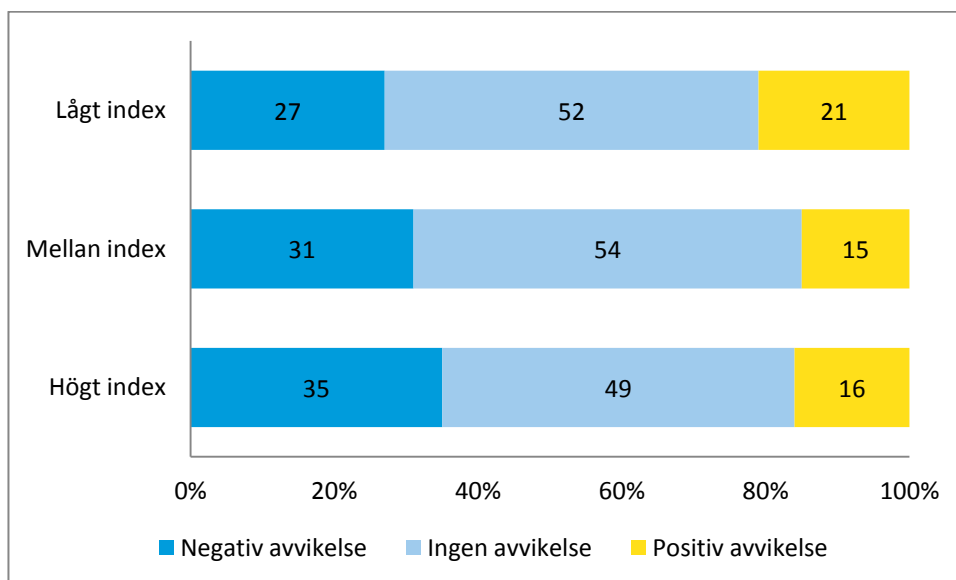


Diagram 5. Grundskolan. Andel i procent av de omrättade proven, uppdelat på skolor med olika socioekonomiskt index, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Gymnasieskola

För gymnasieskolan ingick fler klasser per skola i studien och för skolor med flera parallella klasser. Därmed visas för många skolor ett resultat på skolnivå inte bara på lärarnivå. För de skolor med få (för några skolor enbart en) klasser studeras mer av enskilda lärares bedömningar.

Hela urvalet

Den vanligaste omrättningen var här ett negativt betygsstegs skillnad (29 procent), och endast 21 procent av proven ombedömdes med samma betyg (Diagram 6). 71 procent av proven avvek därmed från ursprungsbedömningen. I Skolinspektionens ombedömning var överensstämmelsen 55 procent (gjord i relation till Lgy 94 och med fyra betygssteg). Inspektionens granskning visade att 34 procent av proven hade en negativ avvikelse, och i vår studie ombedömdes 56 procent av proven med ett lägre betygssteg. Även positiva avvikelser var vanligare i vår studie - 25 procent av proven - jämfört med 11 procent i Skolinspektionens omrättningar.

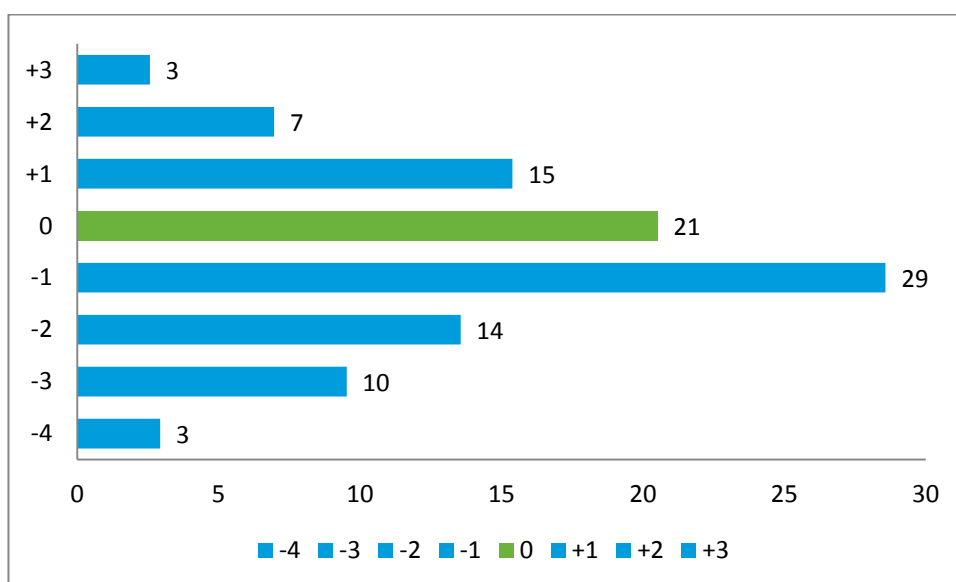


Diagram 6. Gymnasieskolan. Andel i procent av de omrättade proven som avviker olika antal betygssteg från ursprungsbedömningen.

I Diagram 7 kan vi tydligare se hur många prov som vid ombedömningen avvek mycket från ursprungsbedömningen. Hela 26 procent av proven ombedömdes 2-4 betygssteg lägre och 10 procent av proven ombedömdes 2-3 betygssteg högre än vid ursprungsbedömningen. Två betygssteg är då exempelvis mellan A och C eller D och F, och 4 betygssteg kan vara mellan A och E eller B och F. Inom intervallet inget eller max ett betygsstegs avvikelse återfinns 64 procent av proven.

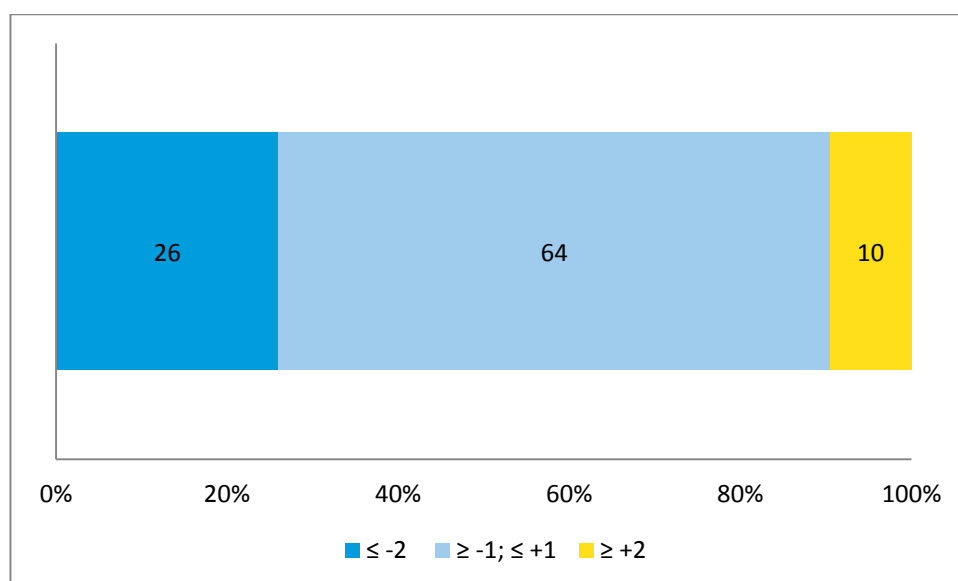


Diagram 7. Gymnasieskolan. Andel i procent av de omrättade proven som avviker: (blått) två betygssteg eller mer negativt, (ljusblått) inget eller max ett betygssteg, (gult) 2 betygssteg positivt eller mer.

Pojkar/flickor

Uppdelat per kön ser vi att negativa avvikelser var mer vanligt förekommande vid omdömning av flickors prov och det omvända för positiva avvikelser (Diagram 8). I denna studie finns ett begränsat urval och att pojkars prov i stor utsträckning visar en positiv avvikelse kan delvis förklaras av att en gymnasieskola (Skola C) med många pojkar uppvisade stora positiva avvikelser vid omdömningen, och en skola (F) med en större andel flickor uppvisade stora negativa avvikelser (Diagram 9). I Skolinspektionens omrättning fanns en signifikant skillnad där pojkars prov ursprungligen (medelvärde, alla ämnen) bedömdes hårdare än flickors (7,2 procent av en standardavvikelse).¹⁷ Detta har även observerats i andra studier.¹⁸ Mer studier krävs för att se om det finns någon större skillnad mellan hur lärare på gymnasieskolan väger in andra faktorer relaterade till kön vid bedömning av de egna elevernas prov.

¹⁷ Skolinspektionen (2012). *Lika för alla?*

¹⁸ Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics.*

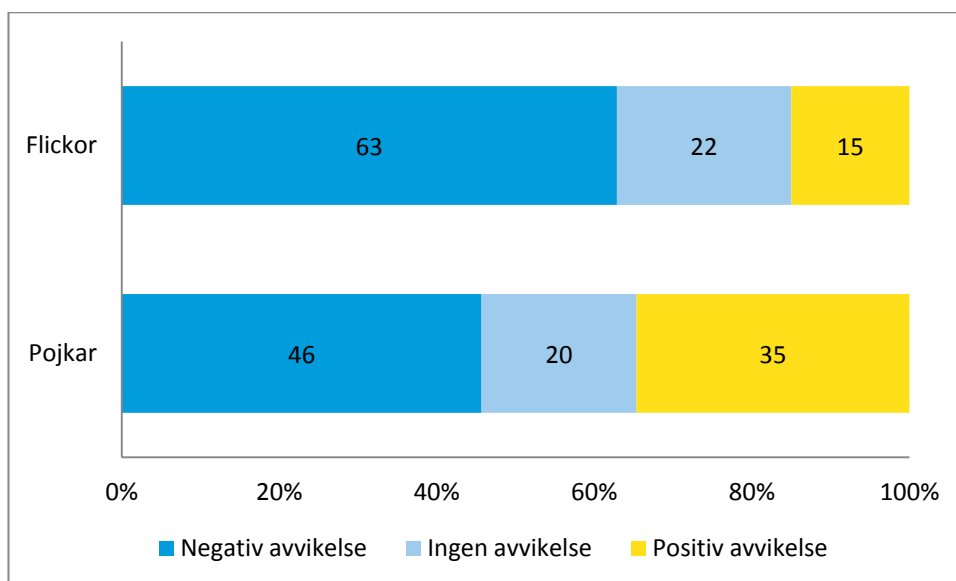


Diagram 8. Gymnasieskolan. Andel i procent av de omrättade proven, uppdelat per kön, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Högskoleförberedande/yrkesförberedande program

Omrättade prov från yrkesförberedande program var en liten andel jämfört med de från högskoleförberedande. För de yrkesförberedande programmen var negativa avvikelser mer vanligt förekommande än positiva, och för de högskoleförberedande var det vanligare med ett högre betyg vid omdömning (25 procent) än var det för prov från yrkesförberedande program (6 procent) (Diagram 9).

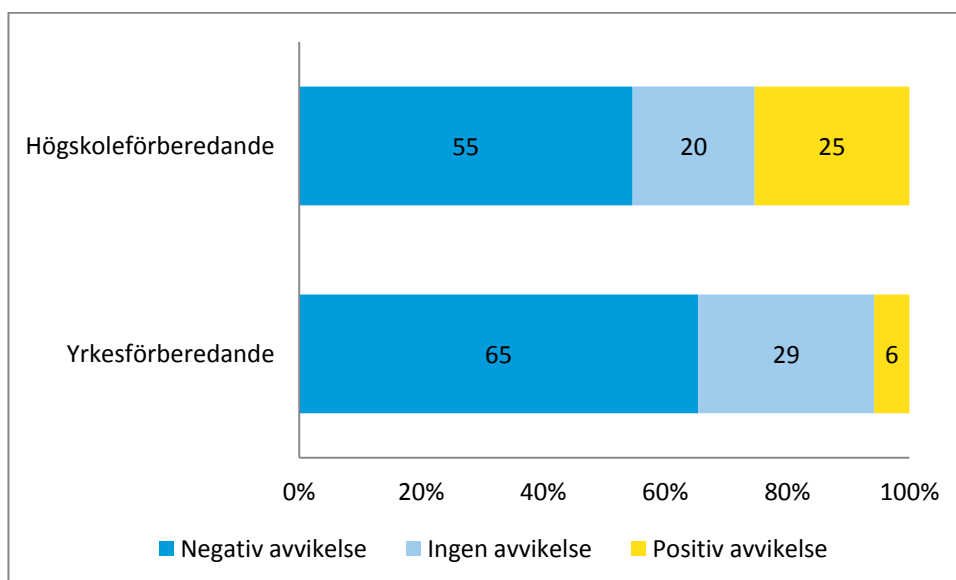


Diagram 9. Gymnasieskolan. Andel i procent av de omrättade proven, uppdelat på yrkesförberedande respektive högskoleförberedande program, som avviker positivt, negativt eller inte alls från ursprungs rättningen.

Skolnivå

Uppdelat per gymnasieskola syns stora avvikelser. I genomsnitt ombedömdes 21 procent av proven med samma betyg, och på skolnivå ser vi att mellan 10 och 43 procent av proven ombedömdes likadant (diagram 10). I Skolinspektionens omrättning bedömdes 55 procent av proven likadant som vid ursprungsrättningen. Dock har den enbart gjorts i relation till Lgy94 med 4 betygssteg, och inte som här med 6 betygssteg, vilket är en av förklaringarna till den stora diskrepansen.

Prov från skola F avvek avsevärt mer negativt vid ombedömningen; 33 procent av proven med hela 3-4 betygsstegs skillnad. På skola C var det inget av proven som avvek mindre än ett betygssteg lägre.

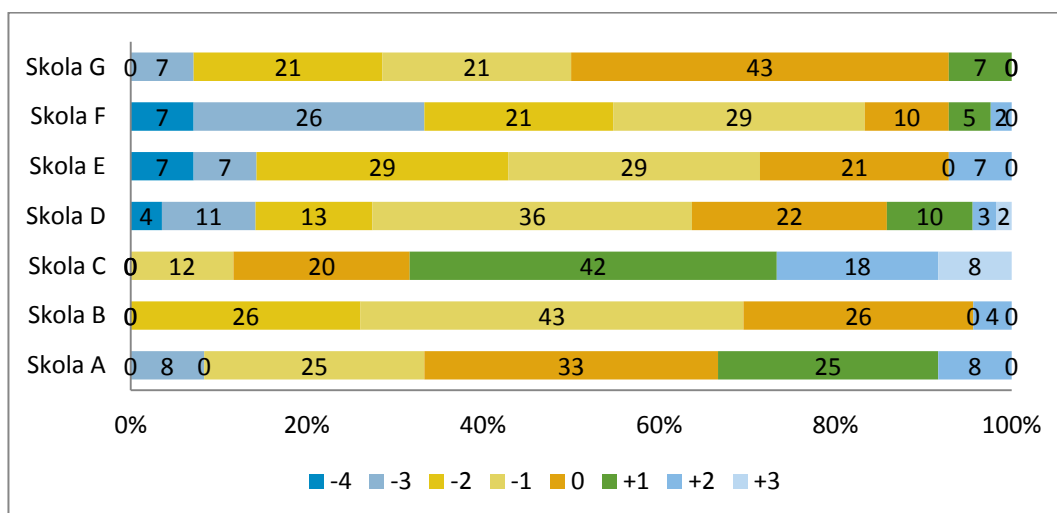


Diagram 10. Gymnasieskolan. Andel i procent av de omrättade proven som avviker olika antal betygssteg från ursprungsrättningen.

I diagram 11 kan vi tydligare se avvikelser på mellan två till fyra betygssteg. Andelen prov med en negativ avvikelse på 2 eller fler betygssteg varierar mellan 0 (Skola C) och 55 procent (Skola F) (Diagram 11). Prov med inget eller max ett betygsstegs avvikelse varierar mellan 43 och 83 procent (Diagram 11).

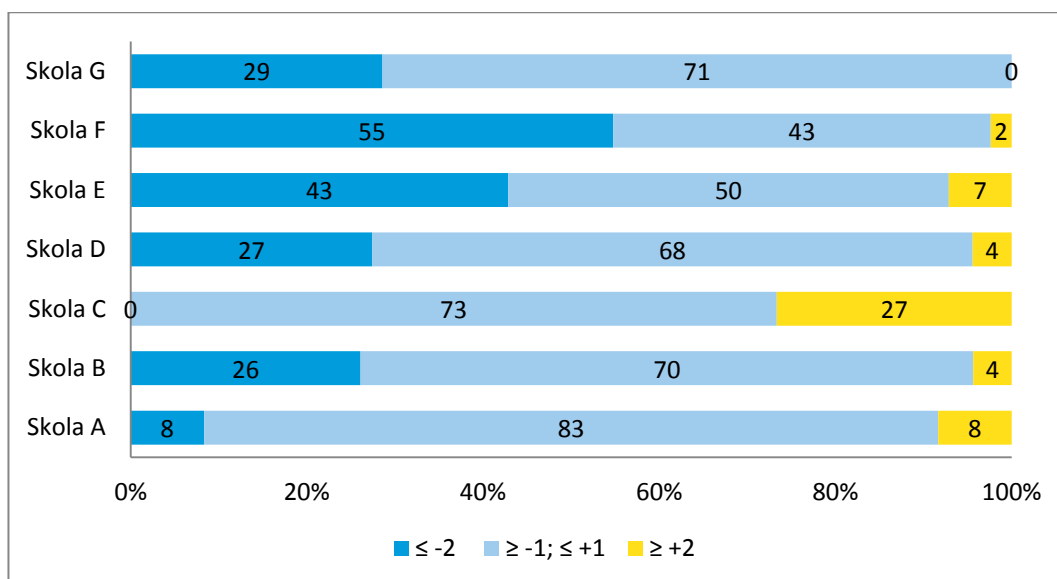


Diagram 11. Gymnasieskolan. Andel i procent av de omrättade proven, per gymnasieskola, som avviker: (bått) två betygssteg eller mer negativt, (ljusblått) inget eller max ett betygssteg, (gult) 2 betygssteg positivt eller mer.

Betygsändringar

I diagram 12 kan vi se vilka avvikelser som var vanligast. Framförallt var det prov med ursprungsbetygen A eller B som bedömdes lägre (93 respektive 80 procent av proven), och prov som ursprungligen bedömts som E som i 72 procent av fallen bedömdes högre i ombedömningen.

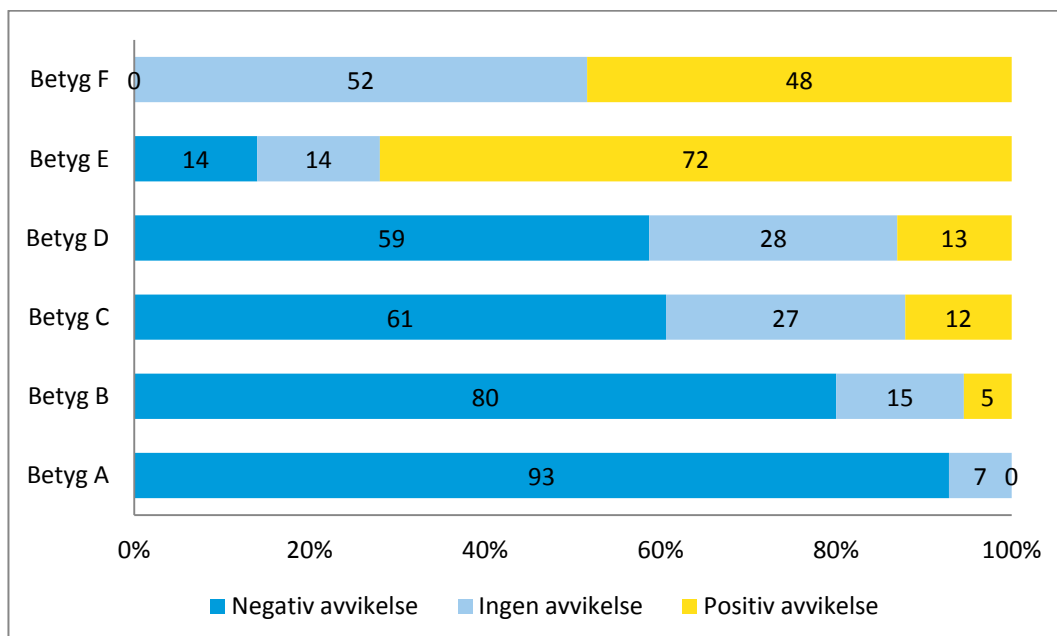


Diagram 12. Gymnasieskolan. Andel i procent av de omrättade proven som har ingen avvikelse, negativ eller positiv avvikelse från ursprungsbetyget.

Om avvikelserna

Proven från gymnasieskolan kan förväntas att ha en mindre bedömaröverensstämmelse då dessa bedöms i sex steg och inte fyra som för grundskolan. Begreppet ”Likvärdig bedömning” måste självklart sättas i relation till hur många betygssteg som finns – ju fler steg desto lägre likvärdighet fås per automatik, om vi med likvärdighet menar att två lärare ger samma betyg på en uppgift. Från och med våren 2012 bedöms även nationella prov för grundskolans årskurs 9 i sex steg. Skolinspektionen har ett fortsatt uppdrag att granska dessa, vilket publiceras i en rapport 2013. Utifrån vår studie kan vi anta att man även i Skolinspektionens framtida omrättningar kommer att se bedömaröverensstämmelsen minska från de runt 55 procent man sett i systemet med fyra steg.

Skolinspektionens granskning ska inte heller ses som ett exakt mått på hur stora skillnader i bedömaröverensstämmelse det finns nationellt, flera av Skolinspektionens omrättare uppger att de upplever att de följer bedömningsanvisningarna striktare när de omrättar på Skolinspektionens uppdrag än när de bedömer de elever de själva har undervisat. Huruvida det gäller även de lärare som medverkat i denna studie har inte undersökts. En strängare bedömning betyder inte att omrättarna gör en mer korrekt bedömning än lärarna. I ett antal uppsatsdelprov i ämnesprovet i svenska för årskurs 9 gjorde Skolinspektionen en tredje rättning av prov där det varit stora skillnader mellan lärarens ursprungliga bedömning och omrättarens. Den tredje rättningen, som omfattade ca 20 procent av alla prov, fick till resultat att en lika stor andel styrkte den ursprungliga bedömningen som Skolinspektionens första omrättning.

Bedömningsanvisningarna skiljer sig åt mellan grundskolan och gymnasieskolan. För grundskolan fås en matris med ett kvantitativt stöd för sammanvägning till betygsresultat (Bilaga 1). På gymnasieskolan, där detta genomfördes för första gången i relation till den nya ämnesplanen, ges ett omfattande stöd som 84 procent av lärarna i en utprövningsomgång upplevde som bra. I stödet finns bedömda exempeltexter och en analytisk matris med delkriterier, däremot är det inte helt klart hur sammanvägningen sedan sker till ett delprovsbetyg i relation till en annan matris (Bilaga 2). Bedömningen sker för gymnasieskolan i relation till fem bedömningsaspekter. Utgångspunkten är att för ett visst provbetyg ska alla kraven för betyget vara uppfyllda. För betygen B och D gäller att betygssteget under måste vara helt uppfyllt och det övre ”till övervägande del”. Det betyder till exempel att om fyra aspekter bedöms som A och en med E skulle delprovsbetyget bli D. Om läraren istället utgår från den holistiska matrisen (med bara en aspekt) kanske bedömningen blir annorlunda. Utan tydliga anvisningar för hur sammanvägningen till ett provbetyg går till kan särskilt bedömningarna av prov med spretiga kunskapsprofiler ha bedömts olika.

En hundra procentig överensstämmelse är omöjligt att åstadkomma i ett kvalitativt system, men likvärdigheten kan säkerligen ökas genom beprövade insatser. Exempelvis har man i Queensland i Australien nått en mycket hög överensstämmelse (minst 85 procent) genom medvetet arbete med att involvera lärare i att ta fram kriterier och genom moderering.¹⁹

Slutsats

Bedömaröverensstämmelsen på grundskolan är 49 procent, och de skolor som avviker mest har 39 respektive 60 procents överensstämmelse. Detta kan jämföras med Skolinspektionens stora omrättning där överensstämmelsen var 55 procent. På grundskolan kan dock inget generellt sägas om skillnader mellan skolor, då det endast var en lärare per skola som medverkat.

För gymnasieskolans del finns stora skillnader mellan ursprungsomrättning och vid omrättningen. Endast 21 procent av proven ombedömdes med samma betyg, och det var även stora skillnader mellan skolor. Avvikelser med två eller flera steg var vanligt förekommande.

I gymnasieskolan togs vart femte prov ut, vilket innebär att för de flesta skolor är det fler lärare som medverkat och resultatet kan i större grad återspegla en bedömningskultur på skolan. Visserligen kan avvikelserna antas vara större då det är en sexgradig skala som tillämpas, till skillnad från grundskolan där det då fanns fyra steg.

En viktig utgångspunkt för en kvalitativ bedömning är att det finns tydliga bedömningsanvisningar. Bedömningsanvisningarna skiljer sig åt mellan grundskolan och gymnasieskolan. För grundskolan fås en matris med ett kvantitativt stöd för sammanvägning till betygsresultat (Bilaga 1). I gymnasieskolan, där detta genomfördes för första gången i relation till den nya ämnesplanen, ges ett omfattande stöd som 84 procent av lärarna i en utprövningsomgång upplevde som bra. Trots det ser vi att man tolkar bedömningsanvisningarna olika. I stödet finns bedömda exempeltexter och en analytisk matris med delkriterier, däremot är det inte helt klart hur sammanvägningen sedan sker till ett betyg i relation till en annan mer holistisk matris. För gymnasieskolan var det även första gången ett nationellt prov genomfördes utifrån de nya ämnesplanerna med sex betygssteg, vilket kan ha påverkat undersökningen.

¹⁹ QSA (2010). *School-based assessment. The Queensland system.*

Vid samtal med lärarna efteråt framkom mer information om hur de hade arbetat även ursprungligen med bedömningen. Några av lärarna hade positiva erfarenheter av att avidentifiera proven (genom att använda ett kodsysteem). En gymnasieskola berättade om hur de parvis hade bedömt sina egna och varandras elevers prov och diskuterat gemensamt och i tveksamma fall rådfrågat en tredje lärare. Den skolan hade minst avvikelser från ursprungsbedömningen vid ombedömningen (Gymnasieskola A). Av de lärare som vi träffade uppgav alla att de föredrog sambedömning framför att bara lämna proven till annan lärare (och då även själv få en till klassuppsättning). Att lämna prov till andra skolor, som i detta fall är även juridiskt problematiskt då rektor enligt skollagen ansvarar för betygsättning. De nationella proven ska beaktas vid betygsättning, men inte vara "slutprov" och några lärare pratade om att de inte såg det som ett stort problem om de var oense i bedömningen av en del i ett nationellt prov, men att det var väldigt ovanligt att de var oense om ett slutbetyg.

I verksamhetsplanen för 2013 står att "De nationella ämnesproven kommer i normalfallet att bedömas av en annan lärare än den som sätter elevens betyg". Utifrån denna studie och andra ser vi att för att få en ökad bedömaröverensstämmelse är det viktigare att arbeta med sambedömning inom skolor, och kanske även mellan skolor. Dessutom är det ju viktigt för undervisande lärare att analysera de egna elevernas resultat, i både summativt och formativt syfte.

Rekommendationer för fortsatt arbete för en mer likvärdig bedömning

För en ökad likvärdighet och med utgångspunkt i bestämmelser runt bedömning och betygsättning ges följande rekommendationer.

Lärare

- Sätta sig in i de nya kunskapskraven och det stödmaterial som finns.
- Analysera likvärdigheten i den egna betygssättningen (intra-bedömarreliabiliteten).²⁰
- Arbeta kollegialt med analys och sambedömning av elevtexter kontinuerligt för att stärka interreliabiliteten.²¹

Rektor

- Aktivt främja sambedömning

²⁰ Skolverket (2012). *Allmänna råd för Bedömning och betygsättning i gymnasieskolan*.

²¹ Skolverket (2011). *Allmänna råd för planering och genomförande av undervisning*.

- Stödja lärarna i deras arbete med att bedöma elevernas kunskaper genom att tillsammans med dem utveckla gemensamma rutiner och former för att utvärdera elevernas kunskaper.²²
- Följa upp resultat och betygssättning på skolan så att ändamålsenliga insatser kan sättas in för att betygssättningen ska vara likvärdig.²³
 - Särskilt följa upp och analysera könsskillnader (särskilt på gymnasiet).
- Organisera modereringstillfällen mellan skolor där elevexempel diskuteras i relation till kunskapskraven.

Huvudmannen

- Aktivt främja sambedömning
- Se till att skolorna kontinuerligt arbetar för att bedömning och betygssättning ska vara likvärdig mellan skolorna och följa upp detta.²⁴
- Särskilt följa upp arbetet med likvärdig bedömning i resultatdialogerna.
- Revidera/lägga till frågor i medarbetarenkät/elevenkät?
- Stödja skolor i att organisera modereringstillfällen mellan skolor där elevexempel diskuteras i relation till kunskapskraven.

²² Skolverket (2011). *Allmänna råd för planering och genomförande av undervisning.*

²³ Skolverket (2012). *Allmänna råd för Bedömning och betygsättning i gymnasieskolan.*

²⁴ Skolverket (2011). *Allmänna råd för planering och genomförande av undervisning.*

KÄLLOR

- Brennan, L. Robert (red) (2006). *Educational Measurement*, 4th ed., s. 412.
- Ciolek Laerum, Beatrice (2009). *Elever skriver och lärare bedömer - en studie av elevtexter i åk 9*. (FUMS Rapport nr 226.)
- Jönsson, A. *Lärande bedömning*. www.larandebedomning.com. (2013-03-10)
- Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics*.
- Klenowski, V. and Wyatt-Smith, C. (2010). *Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform*. Assessment Matters.
- Riksrevisionen (2011). *Lika betyg, lika kunskap? En uppföljning av statens styrning mot en likvärdig betygssättning i grundskolan*.
- Skolverket (2008). *Central rättning av nationella prov*.
- Skolverket (2009). *Bedömaröverensstämmelse vid bedömning av nationella prov*.
- Skolverket (2011). *Allmänna råd för planering och genomförande av undervisningen*.
- Skolverket (2012). *Allmänna råd för bedömning och betygsättning i gymnasieskolan*.
- Skolverket (2012). *Nationella prov –Skolverkets synpunkter i anledning av Skolinspektionens omrättning*.
- Skolinspektionen (2012). *Lika för alla?*
- Skolinspektionen (2013). *Tillsyn av bedömning och betygsättning*.
- QSA (2010). *School-based assessment. The Queensland system*.

BILAGA 1

Bedömningsmatris för delprov C – ämnet svenska

Texten	Godkänt	Väl godkänt	Mycket väl godkänt
1. Kommunikativ kvalitet Syfte och begriplighet Medvetenhet om mottagare	• fungerar i det tänkta sammanhanget efter viss bearbetning* • visar strävan efter anpassning till läsaren	• fungerar i det tänkta sammanhanget • är anpassad till läsaren	• fungerar väl i det tänkta sammanhanget • är väl anpassad till läsaren
2. Innehåll Årskursadekvat komplexitet Genremedvetenhet	• innehåller i huvudsak relevanta uppslag; de flesta uppslag utvecklas något • redovisar händelser mer än förmedlar upplevelser (berättande text) • den egna tankegången framgår och är delvis underbyggd (utredande/argumenterande text) • innehåller exempel/detaljer med viss betydelse för helheten	• innehåller relevanta uppslag/egna synvinklar som utvecklas • förmedlar upplevelser och reflekterar kring dem (berättande text) • den egna tankegången framgår och är underbyggd (utr./arg. text) • ansatser till kritiskt resonemang finns (utr./arg. text) • innehåller träffande exempel/detaljer med betydelse för helheten	• genrens möjligheter utnyttjas • förmedlar, tolkar och värderar upplevelser (berättande text) • den egna tankegången är väl underbyggd (utr./arg. text) • ett kritiskt resonemang förs (utr./arg. text) • god balans mellan: detalj-översikt, personligt-allmänt
3. Struktur Fokus Textbindning	• uppbyggnaden är i huvudsak tydlig, t.ex. kronologisk • enkel textbindning, t.ex. genom sambandsord	• uppbyggnaden är tydlig • varierad textbindning, t.ex. genom sambandsord	• uppbyggnaden är välvald med tanke på ämnet • varierad och ändamålsenlig textbindning, t.ex. genom sambandsord
4. Språk Ordval Tempusbruk Meningsbyggnad Stil	• ordvalet uppfyller i huvudsak uppgiftens krav • tempusbruket är i huvudsak korrekt • meningsbyggnaden är i huvudsak korrekt	• ordvalet är passande och varierat • tempusbruket är anpassat efter olika tidsperspektiv • meningsbyggnaden är varierad och i huvudsak korrekt • viss stilistisk anpassning efter ämne/situation	• ordvalet är träffande, varierat och specifikt • tempusbruket är anpassat efter olika tidsperspektiv och används på ett medvetet sätt • meningsbyggnaden är varierad och träffsäker • god stilistisk anpassning efter ämne/situation
5. Skrivregler Skiljetecken Styckemarkering Stavning	• relativt gott bruk av skiljetecken • vissa stavfel som inte stör förståelsen	• relativt gott bruk av skiljetecken, styckemarkering och rubriker • få stavfel	• gott bruk av skiljetecken, styckemarkering och rubriker • ytterst få stavfel

* "efter viss bearbetning" innebär lättare bearbetning som kan utföras i det tänkta sammanhanget av till exempel en redaktör.

KOPIERINGSUNDERLAG DELPROV C – TILL LÄRAREN
SVENSKA 1 – UPPGIFT 4

Elevens namn:

Bedömningsmatrill till uppgift 4: Min passion för ...

Aspekter	F	E	C	A		
Innehåll och kritisk föreläsning Aktiviteter Bearbetning av textunderlag Källhantering		Eleven redogör för ett fritidsintresse och förklarar varför det är viktigt Eleven jämför översiktligt sitt intresse med Marcus Birros kärlek till fotbollen. I sin jämförelse skiljer eleven på egna och andras tankar, även om käll- känvisningarna inte är fullständiga.	Redogörelsen är fyllig och förklaringen är övertygande. Eleven jämför utförligt sitt intresse med Marcus Birros kärlek till fotbollen.	Elevens inlägg håller hög kvalitet och bidrar med nya perspektiv förhållande till Birros text.		
Disposition och sammenhang Referensbündning Inledning, avslutning Sammanhangssignaler Styckeindelning, styckemarkering Omfång		Texten är sammanhäng- ande och begriplig, t.ex. genom referensbündning och enkla sammanhangs- signaler. Omfånget är till viss del anpassat till situationen, dvs. texten är inte alltför kort (se s. 23).	Texten är välstrukturerad, t.ex. genom styckeindelning markerad med konsekventa styckemarkeringar, inledning och avslutning samt sammanhangssignaler. Omfånget är anpassat till situationen, dvs. texten är inte alltför lång (se s. 23).	Texten är väl disponerad, t.ex. genom god balans mellan olika delar, tydlig styckeindelning som är korrekt markerad och effektiv inledning och avslutning.		
Språk Skriftspråksnormer för språkriktighet Variation Välformulerat språk Situationanpassning		Eleven följer i huvudsak skriftspråksnormer för språkriktighet, t.ex. normer för skiljetecken och stor/ lite bokstav.	Språket är varierat och delvis välformulerat. Stilen är anpassad till situationen, t.ex. genom ordformer, ordval eller läsar- tilltal.	Språket tar till största del välformulerat, dvs. texten är som helhet lättläst, varierad och har övervägande goda formuleringar, även om någon mindre lyckad formulering kan före- komma.		
Betyg på delprovet	F	E	D	C	B	A
Betyg på delprovet Kommunikativ kvalitet Situationanpassning, t.ex. mottagaranpass- ning Självbärande text		Eleven följer instruk- tionen. Textens innehåll och form är till viss del anpassade efter syfte, mottagare och kommunikations- situation. Den fungerar efter viss bearbetning som ett läsarinlägg i en lokal tidning. Bearbetning kan tex. krävas för att texten ska bli helt själv- bärande.		Textens innehåll och form är anpassade efter syfte, mottagare och kommunikations- situation. Texten är självbärande, t.ex. genom mottagar- anpassat innehåll och språk. Den fungerar som ett läsarinlägg i en lokal tidning.		Textens innehåll och form är väl anpassade till syfte, mottagare och kommunikations- situation. Den fungerar väl som ett läsarinlägg i en lokal tidning.