

Samverkan för en likvärdig bedömning

Ombedömning av nationella prov 2015

Redovisning av regeringsuppdrag Dnr U2014/7535/GV

Innehåll

FÖRORD	4
SAMMANFATTNING	5
UPPDRAGET	7
BAKGRUND, SYFTE OCH GENOMFÖRANDE.....	7
DENNA RAPPORT	7
PUBLICERING AV RESULTAT PÅ SKOL- OCH HUVUDMANNANIVÅ.....	8
RESULTAT	9
OMBEDÖMDA DELPROV FÖR OMGÅNG 7	9
REPRESENTATIVITET	9
BEDÖMNINGSPROCESSEN	9
<i>Vanligast att elevens ordinarie lärare bedömer provet.....</i>	<i>10</i>
<i>Läraren känner vanligen till elevens identitet vid bedömning</i>	<i>11</i>
<i>Sambedömning tillämpas i fler än hälften av fallen.....</i>	<i>12</i>
<i>Sambedömning är vanligare på kommunala skolor</i>	<i>13</i>
<i>Avidentifiering vanligare i lägre årskurser på skolor med kommunal huvudman</i>	<i>14</i>
<i>Olika mönster vid avidentifierade prov</i>	<i>15</i>
<i>Bedömningsprocessen kan ha stor betydelse för utfallet</i>	<i>15</i>
<i>Sambedömning och extern bedömning minskar avvikelsegraden</i>	<i>17</i>
EXEMPEL FRÅN TVÅ SKOLOR SOM ARBETAR MED SINA BEDÖMNINGSPROCESSER	20
SKOLENHETER MED OLIKA RESULTAT I OMBEDÖMNINGEN VALDES UT	20
<i>Likvärdighet inom bedömning förutsätter att lärare samarbetar</i>	<i>20</i>
<i>En levande diskussion kring bedömningsfrågor är en framgångsfaktor</i>	<i>22</i>
<i>Hur går bedömningsprocessen till?</i>	<i>22</i>
<i>Fortbildning och uppföljning av bedömningsarbetet</i>	<i>23</i>
HUR MYCKET SKILJER SIG BEDÖMNINGARNA AV SAMMA ELEV LöSNING?.....	24
<i>Bedömningarna skiljer sig åt för mer än varannan elevlösning</i>	<i>25</i>
<i>Små uppmätta skillnader för språkproven mellan omgång 6 och 7</i>	<i>25</i>
<i>Inga skillnader totalt sett för NO-prov mellan åren</i>	<i>25</i>
<i>Störst skillnad mellan åren för fysik</i>	<i>27</i>
<i>SO-prov bedöms i hög grad likvärdigt.....</i>	<i>27</i>
STOR VARIATION MELLAN OLIKA SKOLOR FÖR SAMMA ELEV LöSNING.....	28
<i>Kan skolan förklara avvikelsegraden?.....</i>	<i>30</i>
<i>Små avvikelser mellan kommunala och fristående skolor</i>	<i>31</i>
<i>Avvikelsena kan ha en avgörande betydelse för det slutgiltiga provbetyget.....</i>	<i>32</i>
UTVECKLINGSOMRÅDEN.....	33
<i>Sambedömning ökar likvärdigheten.....</i>	<i>33</i>
<i>Avidentifiering minskar avvikelsegraden</i>	<i>33</i>
<i>Digitalisering av de nationella proven främjar likvärdighet</i>	<i>33</i>
<i>Samverkan mellan skolor ökar samstämmigheten i bedömningen.....</i>	<i>34</i>
<i>Rektorns betydelse för uppnåendet av en likvärdig bedömning.....</i>	<i>34</i>

<i>Provens och bedömningsanvisningarnas utformning är avgörande</i>	34
BILAGA 1: GENOMFÖRANDET AV OMBEDÖMNINGEN	36
URVAL AV PROV OCH SKOLOR	36
<i>Ämnen, delprov och årskurser som ingår i omdömmningen</i>	36
<i>Hela provet omdöms inte, endast längre skrivuppgifter.....</i>	37
<i>Prov i matematik omdömdes inte</i>	37
URVALETS SAMMANSÄTTNING	37
TERMINOLOGISK ÖVERSIKT	38
FÖRDELNING AV PROV MELLAN OMBEDÖMARE	38
OMFATTNING OCH BORTFALL	39
<i>Bortfall</i>	41
<i>Vad beror bortfallet på?</i>	41
<i>Frågor om bedömningen på skolan</i>	41
BILAGA 2: ETT HARMONISERAT MÅTT FÖR ATT MÄTA SKILLNADER I BEDÖMNING	43
BERÄKNING AV DET HARMONISERADE RESULTATMÅTTET	43
BILAGA 3: METODOLOGISK NOT AVSEENDE BERÄKNING AV VARIATIONEN MELLAN SKOLOR.....	45
BILAGA 4: ANVÄNDNING AV RESULTATEN I TILLSYNNEN	47
<i>Betydande avvikelser på skolnivå: en bedömningsfråga</i>	47
REFERENSFÖRTECKNING	49

Förord

Statens Skolinspektion har sedan 2009 haft i uppdrag av regeringen att årligen genomföra viss central rättning av nationella prov. Syftet med ombedömningen är att stödja en likvärdig bedömning och betygssättning av proven över landet.

Skolinspektionen redovisar här till regeringen resultatet av sjunde årets centrala rättning. Denna avser nationella prov i grundskolan och gymnasiet som genomfördes under 2015, och nationella prov inom den kommunala vuxenutbildningen på gymnasienivå som genomfördes under 2014.

Vid Skolinspektionen har följande personer arbetat med att genomföra uppdraget: Klara Staffansdotter Klingspor (projektledare), Jonna Andersson (tf. avdelningschef), Gunnar Persson (tf. enhetschef), Arjann Akbari, Björn Ljung, Jacob Kurkiala, Johan Granath och Peter Gustafsson (utredare/statistiker).

Sammanfattning

Den här rapporten redovisar resultaten från 2015 års omdömning av nationella prov, den sjunde i ordningen sedan 2009. Syftet med omdömningen är att stödja en likvärdig bedömning och betygssättning av proven över landet och därmed befästa elevens rätt till en likvärdig betygssättning.

Slutsatserna nedan bekräftas av 2015 års omdömning (samt tidigare års omdömningar):

- Skolor som *sambedömer* sina prov är generellt sett mer restriktiva i sin bedömning än skolor där elevens ordinarie lärare ensam bedömer proven.
- Det är vanligare att omdömaren sätter ett *lägre* provbetyg än läraren vid skolan.
- Det är vanligt att omdömaren sätter ett annat delprovsbetyg än läraren vid skolan. För vissa delprov och ämnen skiljer sig bedömningarna åt för mer än varannan elevlösning.
- De största och mest frekventa avvikelserna återfinns för de delprov där eleven ska avge svar i form av en längre text.

Fokus i årets analys har även varit att ge praktiska exempel på hur två olika skolor arbetat med sina bedömningsprocesser.

Bedömningsprocessen har stor betydelse för betyget

Skolor tillämpar olika processer vid bedömningen av proven. För ungefär vartannat insänt prov tillämpade skolan en *organiserad process för sambedömning*, det vill säga en process där flera lärare på något sätt samverkat kring bedömningen av provet. *Extern bedömning* är mer ovanligt och tillämpades för cirka vart femte insänt prov.

Både sambedömning och extern bedömning verkar ha en återhållande effekt på betygssättningen av delprovet, vilket resulterar i ett något lägre betyg för en given elevlösning i snitt än om elevens ordinarie lärare bedömer provet ensam.

Det går att identifiera framgångsfaktorer för bedömningsprocesser

För att ge tydligare praktiska exempel redovisas i denna rapport även hur två olika skolor arbetat med sina bedömningsprocesser. En framgångsfaktor för ena skolenheten verkar vara att på ett organiserat sätt arbeta med betygssättning på huvudmannanivå. För den andra skolenheten verkar framgångsfaktorerna handla om att tillämpa en rad olika metoder för att trygga samstämmigheten kring betygssättning.

Bedömningarna skiljer sig åt för mer än varannan elevlösning

I knappt en femtedel av bedömningarna (18 procent) har skolans lärare satt ett delprovsbetyg som är *minst* ett betygssteg lägre än ombedömaren och i 41 procent av fallen har skolans lärare satt ett delprovsbetyg som är *minst* ett betygssteg högre än ombedömaren.

Skillnader i bedömning mellan olika skolor uppgår till ett betygssteg i snitt

För de delprov som ombedömts visar resultaten att en elev på en skola som sätter relativt låga delprovsbetyg typiskt sett får ett delprovsbetyg som ligger cirka ett halvt till ett betygssteg lägre på skalan A-F än om eleven lämnat in samma lösning på en skola som sätter relativt höga delprovsbetyg. Sett till de fem procent mest utstickande skolorna växer skillnaden till mellan ett till två betygsstegs skillnad i snitt i bedömningen av likvärdiga elevlösningar.

Vid betygssättningen i ämnet av elevens kunskaper har resultatet från de nationella proven ofta en avgörande betydelse. Det är därför rimligt att anta att variation i bedömningen av likvärdiga elevlösningar, som ombedömningen visar, även kan slå igenom på elevers ämnes- och kursbetyg.

Utvecklingsområden för att bidra till likvärdighet i bedömningarna

Sammanfattningsvis understryker resultaten att det finns ett flertal utvecklingsområden, vilka kan bidra till att öka likvärdigheten i bedömningarna av nationella prov, samtidigt som proven behåller möjligheten för elever att visa sådana kunskaper som inte låter sig prövas i frågor med fasta svarsalternativ. Till exempel har olika former av sambedömning en utjämnande effekt på bedömningsavvikelsena. Aidentifiering av elevens identitet förfaller också ha en viss återhållande effekt på lärarens bedömning.

Uppdraget

Bakgrund, syfte och genomförande

Sedan 2009 har Skolinspektionen årligen fått i uppdrag av regeringen att genomföra viss central rättning av nationella prov. Syftet med den centrala rättningen är att stödja en likvärdig bedömning och betygssättning av proven över landet och således befästa elevens rätt till en likvärdig och rättvis betygssättning.

Skolinspektionen samlar in kopior av genomförda nationella prov från ett urval av skolor. Proven avidentifieras och bedöms därefter på nytt av erfarna lärare som rekryterats av Skolinspektionen. Resultatet används sedan som ett av underlagen som Skolinspektionens inspektörer använder sig av inför tillsyn av skolor.

Bakgrunden till uppdraget är risken för skillnader i hur olika lärare och skolor bedömer proven, framför allt avseende de delprov där elevens svar ges i form av en eller ett mindre antal längre texter, vilka inte kan bedömas utifrån ett entydigt facit. Tidigare års omdömmingar har visat att det i dessa fall inte är ovanligt att det skiljer ett betygssteg eller mer mellan två olika lärares bedömning av samma elevsvar. Provbetygen på nationella prov påverkar i nästa led betyget på ämnet eller kursen. Med tanke på betygets många gånger avgörande roll för urval till högre studier eller vid ansökan till jobb kan det få stora konsekvenser för enskilda elever om till exempel prestationer som anses motivera ett A på en skola endast ges betyget C på en annan.

De prov som omdöms har inga fasta svarsalternativ, och därför heller inget facit. Att Skolinspektionens omdömare gör en annan bedömning av en elevlösning än skolan betyder inte att skolan gjort en felaktig bedömning. En stor variation i bedömningen av likvärdiga elevlösningar mellan *olika skolor* är dock problematisk, eftersom ett av provens huvudsyften är stödja en likvärdig och rättvis bedömning och betygssättning. Därför är det av stor vikt att skolorna får ett underlag som ger dem möjlighet att jämföra sina bedömningar av de nationella proven med andra skolors, och att Skolinspektionen följer upp dessa indikatorer i sin tillsynsverksamhet.

Denna rapport

Skolinspektionen redovisar med denna rapport resultatet av sjunde årets centrala rättning av nationella prov. Rapporten presenterar både ett antal statistiska analyser samt deskriptiva resultat med bäring på frågan om likvärdig bedömning av proven.

Den detaljerade redogörelsen för projektets genomförande, urvalet av skolor och delprov, beräkningstekniska detaljer samt hur resultatet används i tillsynen, har sammanställts i bilagor. Den läsare som vill skaffa sig en djupare förståelse av resultaten uppmanas att läsa bilagorna.

Publicering av resultat på skol- och huvudmannanivå

Rapporter på skol- och huvudmannanivå för 2015 års ombedömning finns publicerade i skoldatabasen Siris.¹ Materialet på Siris, som är det som utgör underlag i tillsynen, redovisas terminsvis.

Resultaten som redovisas i föreliggande rapport avser de *sammanslagna* resultaten från hela 2015 års ombedömning. De totaler som redovisas i den här rapporten skiljer sig därför från de som publicerats i de terminsvisa skol- och huvudmannarapporter som används i samband med tillsyn av skolenheter.

¹ <http://siris.skolverket.se/siris/ris.omrattning.searchForm>.

Resultat

Ombedömda delprov för omgång 7

2015 års ombedömning innefattar prov i svenska och engelska för årskurs 6 och 9, gymnasieskolans grundläggande kurser, samt motsvarande kurser inom den kommunala vuxenutbildningen på gymnasienivå. För årskurs 9 omfattas även proven i biologi, fysik, kemi (NO), geografi, historia, religionskunskap och samhällskunskap (SO).

För de flesta ämnen ombedöms endast *ett delprov*, förutom för svenska årskurs 6 och SO-ämnena för årskurs 9, där två delprov valts ut till ombedömning. Delprov som valts är de där risken för skillnader i bedömning mellan olika lärare bedömts vara störst. Urvalet av prov och delprov för ombedömning beskrivs mera utförligt i Bilaga 1.

Representativitet

Skolurvalet för ombedömningen följer från och med 2014 Skolinspektionens tillsynscykel och kravet på riksrepresentation har därmed utgått. Syftet med urvalet är alltså inte längre att ge en representativ bild av avvikelser mellan skolor på nationell nivå. Resultatet kan därför inte med säkerhet generaliseras till riket i stort. Skolinspektionen bedömer däremot att urvalets sammansättning och stora omfattning gör det möjligt att använda resultaten som en indikation på den allmänna situationen i landet.

Det totala bortfallet på provnivå är 14 procent, men bortfallet varierar kraftigt mellan olika delprov. Framför allt är bortfallet högt för vuxenutbildningen samt för SO-proven i årskurs 9. Resultaten för nämnda prov får därför anses mindre stabila än för språkproven. Den kompletta bortfallstabellen redovisas i Bilaga 1.

Bedömningsprocessen

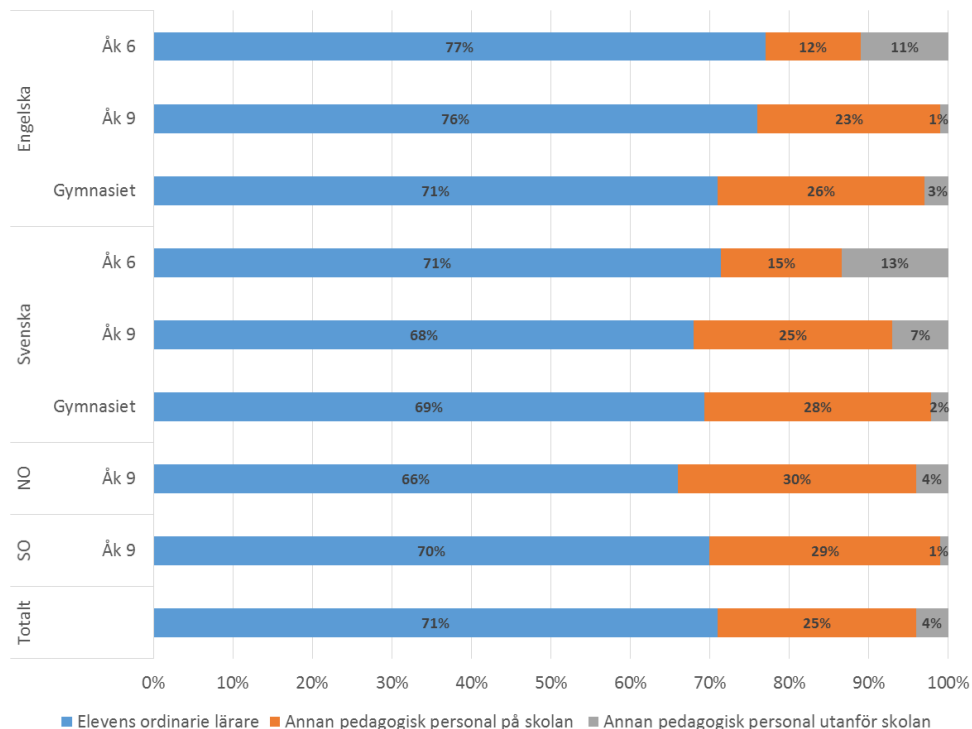
I samband med provinsamlingen gavs skolorna möjlighet att besvara frågor kring hur bedömningen av de nationella proven genomförts. Nedan redovisas svar från skolor som medgett att uppgifterna från deras skola får användas i analyssyfte.

Hur bedöms proven på skolorna?

Eftersom det är frivilligt för skolorna att fylla i uppgifterna kring bedömningens genomförande är svarsfrekvensen lägre än för det totala antalet skolor som ingick i ombedömningen. För prov skrivna inom den kommunala vuxenut-

bildningen samlas ingen information om provens genomförande in. Dessa prov är därför exkluderade ur analysen.

Analysen visar att sambedömning har en återhållande effekt på betygssättningen, vilket leder till ett lägre betyg i snitt. Det är även tydligt att organiserad sambedömning generellt sett har en större återhållande effekt på betygssättningen än sambedömningar av mer informell karaktär.



Figur 1: Skolornas svar på frågan om vem som bedömt det insända provet. Frågan besvaras för varje enskild elevlösning och är frivillig för skolan att besvara.

Vanligast att elevens ordinarie lärare bedömer provet

Tidigare ombedömningsomgångar har visat att elevens ordinarie lärare i de flesta fall bedömt provet. Även denna ombedömningsomgång pekar på samma trend. Detta är fallet för cirka 70 procent av alla ombedömda prov där skolan besvarat frågorna om hur bedömningsprocessen gått till. Andelen prov som bedömts av annan än elevens ordinarie lärare är ungefär lika stor för samtliga stadier, dock något lägre för årskurs 6. Att pedagogisk personal² utanför skolan

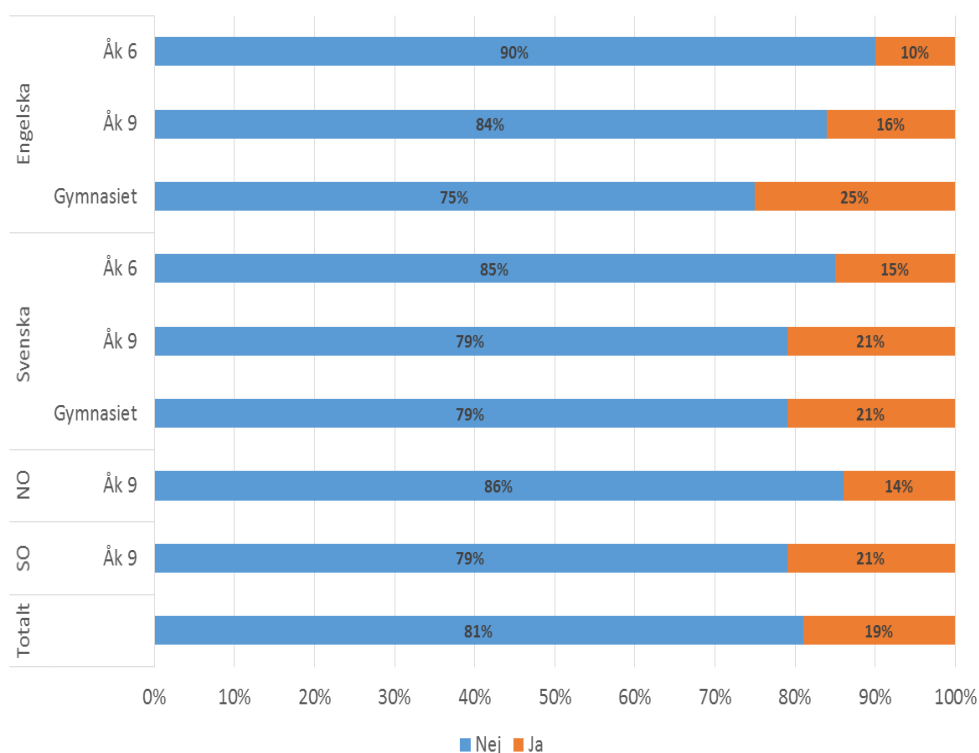
² Avser andra än elevens ordinarie lärare.

står för bedömningen är däremot betydligt vanligare för prov i årskurs 6 än för övriga årskurser.

Läraren känner vanligen till elevens identitet vid bedömning

En knapp femtedel av alla prov (19 procent) som ingick i ombedömningen var avidentifierade då ursprungsbedömningen på skolan gjordes. Andelen var högst för gymnasiet. Orsaken bakom detta kan vara att allt fler gymnasieelever avger sina provsvar på dator, vilket underlättar avidentifiering.

Andelen avidentifierade prov är högst inom ämnen där eleverna avger sina svar i form av essäer eller längre textsvar. För de flesta av dessa prov medföljer inget provhäfte i vilket eleven ska avge in sina svar, utan uppgiften skrivs på externa papper, alternativt på dator. Andelen elever som avger sina svar elektroniskt kan tänkas öka om inget svarsblad erbjuds, vilket även medför att en högre andel av dessa prov avidentifieras inför bedömning.

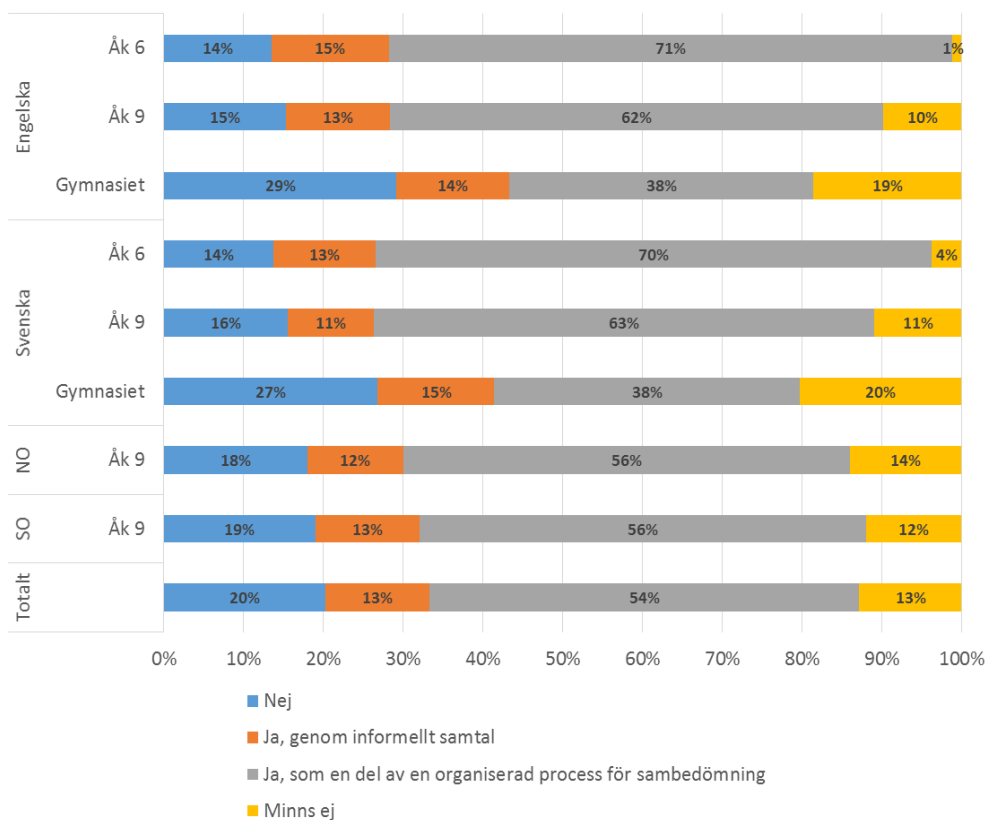


Figur 2: Skolornas svar på frågan om det insända provet var avidentifierat vid bedömningen. Frågan besvaras för varje enskild elevlösning och är frivillig för skolan att besvara.

Sambedömning tillämpas i fler än hälften av fallen

Drygt hälften av de inrapporterade proven, 54 procent, bedömdes inom ramen för en organiserad process för sambedömning. Vart femte prov bedömdes helt utan någon form av sambedömning. För 13 procent av proven tillämpades en sambedömning av en mera informell karaktär, och i en lika stor andel av fallen uppgav skolan att de inte kan minnas hur bedömningsprocessen gick till.

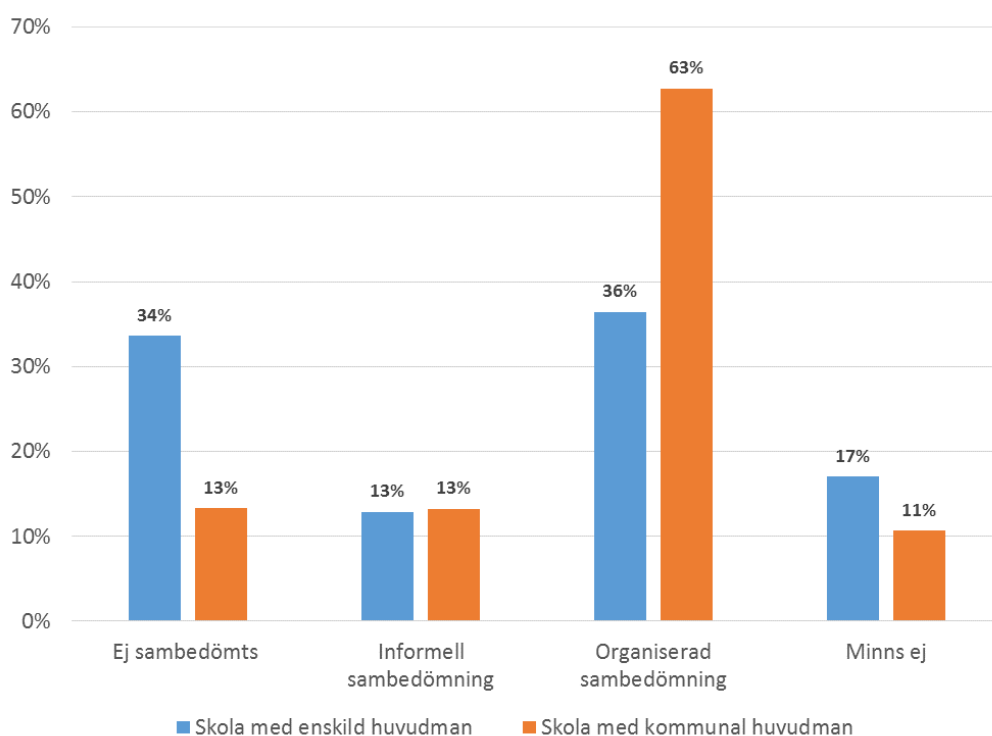
Enligt figur 3 nedan är skillnaden mellan grund- och gymnasieskolan påtaglig, sett till tillämpning av organiserad sambedömning. I årskurs 6 tillämpades sambedömning i någon form för 84 procent av proven. För prov i årskurs 9 var andelen 72 procent. Motsvarande siffra i gymnasieskolan var endast 53 procent.



Figur 3: Skolornas svar på frågan om sambedömning skett vid bedömningen av det insända provet. Frågan besvaras för varje enskild elevlösning och är frivillig för skolan att besvara.

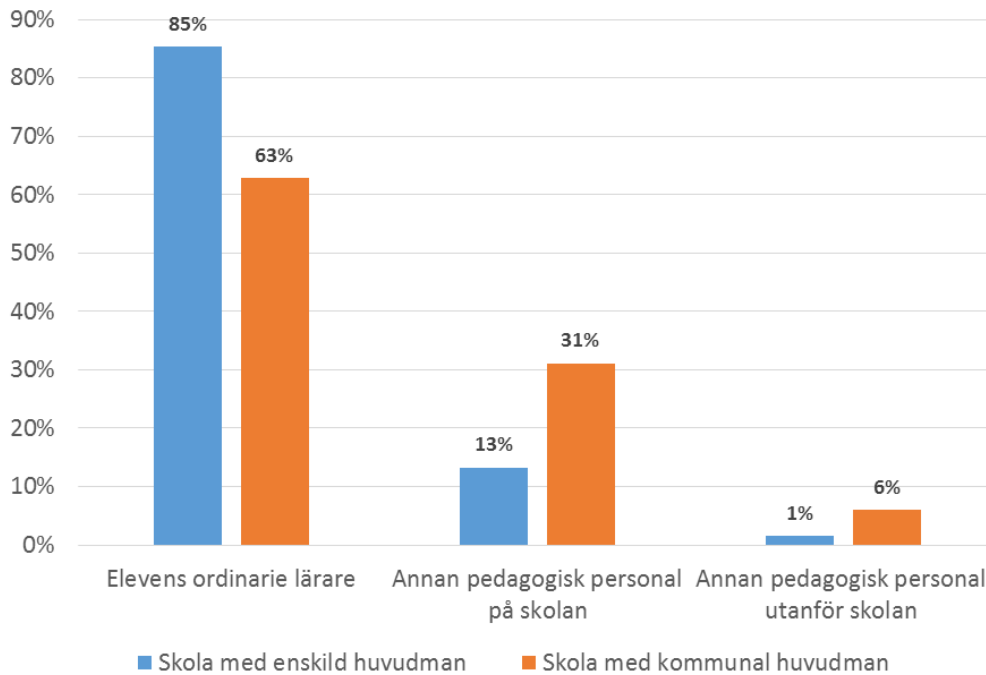
Sambedömning är vanligare på kommunala skolor

Organiserad sambedömning är betydligt vanligare på kommunala skolor än i skolor med enskilda huvudman, något även tidigare års ombedömningsresultat visat. 63 procent av proven genomförda på skolor med kommunal huvudman bedömdes genom en organiserad process för sambedömning. Motsvarande siffra i skolor med enskilda huvudmän var betydligt lägre, endast 37 procent. För båda huvudmannakategorierna var andelen sambedömda prov lägst i gymnasieskolan.



Figur 4: Andelen prov som inte sambedömts, bedömts genom organiserad sambedömning eller genom sambedömning av mera informell karaktär, utifrån skolans huvudmannatyp; enskild eller kommunal.

Andelen prov som bedömts av pedagogisk personal utanför skolan är mer än fyra gånger så stor för skolor med kommunal huvudman än för skolor med enskild huvudman. Däremot rör det sig om en tämligen liten mängd prov för denna grupp, varför siffrorna bör tolkas med försiktigt. Prover som bedömts av annan pedagogisk personal på skolan är betydligt lägre inom fristående skolor, med 13 procent, mot 31 procent på kommunala skolor. Andelen prov som bedömts av annan pedagogisk personal på skolan är för de båda huvudmannatyperna större i de högre årskurserna.



Figur 5: Andelen prov som bedömts antingen av elevens ordinarie lärare eller av annan pedagogisk personal, utifrån skolans huvudmannatyp; enskild eller kommunal.

Orsaken till varför sambedömning av olika slag tillämpades i betydligt större utsträckning på skolor med kommunal huvudman än på skolor med enskild huvudman kan bero på att de flesta kommunala huvudmännen har flera skolor, varpå skolöverskridande sambedömning och organiserade sambedömningsprocesser är enklare att samordna. Skolorna med kommunal huvudman har därutöver i regel haft längre tid på sig att upprätta strukturer och rutiner kring bedömningsprocesserna. Enskilda skolor som nyligen startat sin verksamhet har därför inte ännu hunnit koordinera processen kring bedömning av prov i lika hög utsträckning.³ Storlek på skolan kan även vara av avgörande betydelse för förekomsten av sambedömning. På mindre skolor försvåras till exempel förutsättningarna för sambedömning om det finns få undervisande lärare inom ett ämne.

Aidentifiering vanligare i lägre årskurser på skolor med kommunal huvudman

Andelen aidentifierade prov för årskurs 6 är mer än tre gånger så hög på skolenheter med kommunal huvudman som på skolor med enskild huvudman.

³ Friskolereformen genomfördes 1992, men det var först i och med millennieskiftet som friskoleverksamheten i Sverige tog fart på allvar. De flesta skolor med enskild huvudman i landet har därför endast varit verksamma i 10–15 år.

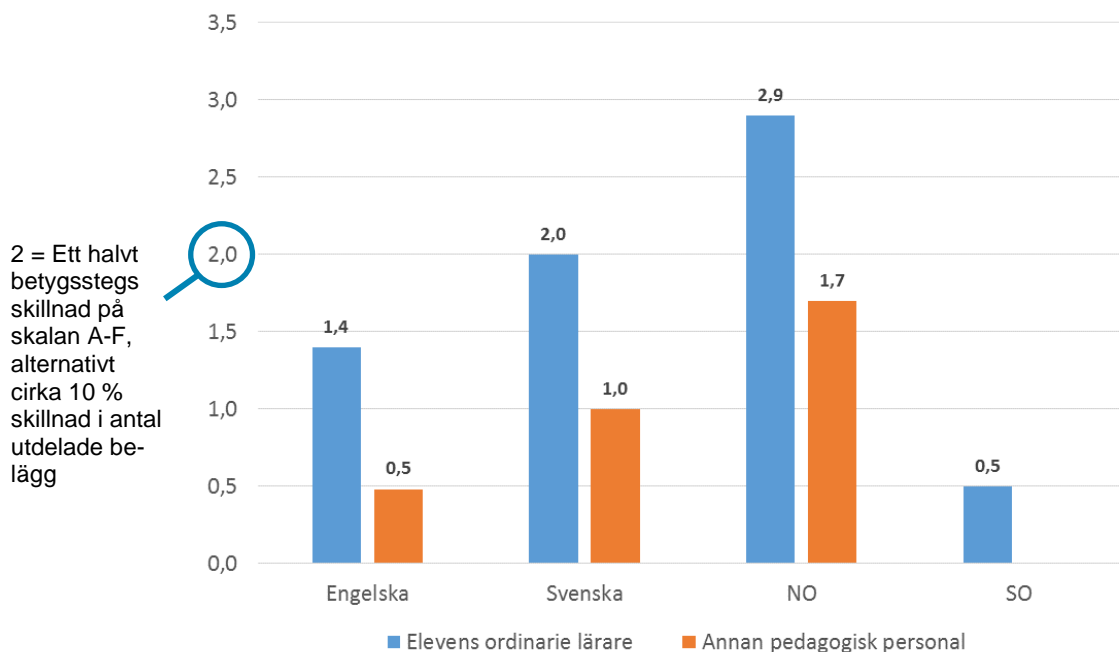
För årskurs 9 har skillnaden i graden av avidentifiering minskat något mellan de båda huvudmannatyperna, men är ändå dubbelt så vanligt på skolor med kommunal huvudman. För gymnasieskolan är andelen avidentifierade prov å andra sidan något högre för skolor med enskild huvudman, jämfört med skolor med kommunal huvudman.

Olika mönster vid avidentifierade prov

För de båda huvudmannatyperna är andelen avidentifierade prov störst för språk- och SO-proven. Allra störst är avvikelsen mellan de båda grupperna för svenska i årskurs 9, där 28 procent av proven skrivna på skolor med kommunal huvudman var avidentifierade, medan endast 10 procent på en skola med enskild huvudman. Däremot var andelen avidentifierade prov i engelska 10 procentenheter högre bland enskilda skolor än kommunala. Det bör poängteras att andelen prov från skolor med enskilda huvudmän är betydligt lägre än den för kommunala skolor, vilket gör att resultatet får tolkas med viss reservation, speciellt i de ämnen där totalt endast ett tusental prov omdömts (NO- och SO-ämnena).

Bedömningsprocessen kan ha stor betydelse för utfallet

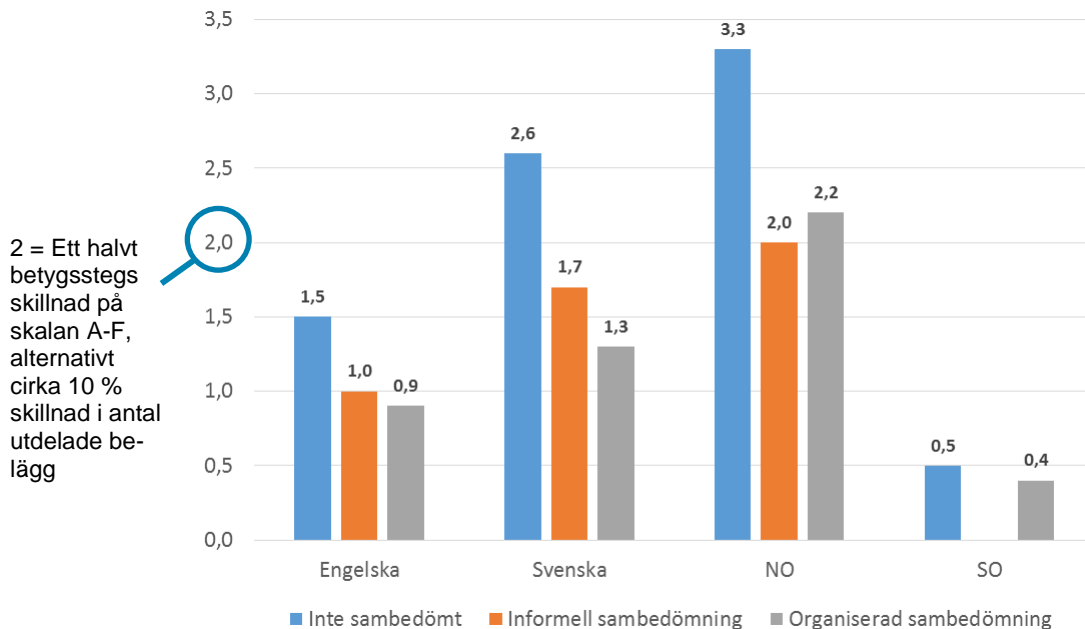
Den genomsnittliga skillnaden mellan skolans bedömning och Skolinspektionens omdömning är signifikant mindre för prov som bedömts av någon annan än elevens ordinarie lärare, vilket figur 6 visar.



Figur 6: Genomsnittlig skillnad mellan skolans bedömning och omdömningen för prov som bedömts av elevens ordinarie lärare eller av annan pedagogisk personal.

Den generella tendensen till att skolan gör en något mer generös bedömning än omdömningen minskar generellt sett kraftigt när någon annan än den ordinarie läraren står för bedömningen. För vissa ämnen och årskurser försvinner den helt.

Språk- och NO-prov är de prov där skolor oftare gör mer generösa bedömningar än Skolinspektionens lärare. För de enskilda NO- och SO-proven är resultatet däremot spretigt och avviker i varierande grad, i vissa fall med signifikanta skillnader i *motsatt* riktning.

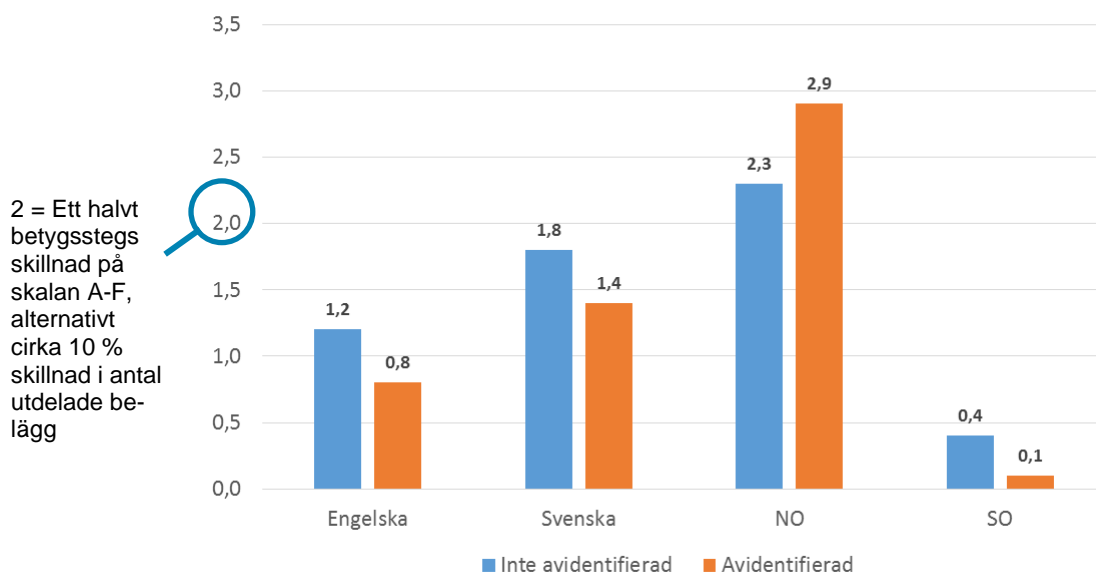


Figur 7: Genomsnittlig skillnad mellan skolans bedömning och omdömningen för prov där bedömaren inte diskuterat bedömningen med någon annan, för prov där bedömningen diskuteras genom informellt samtal, respektive för prov som bedömts i organiserad sambedömning.

Avidentifiering verkar inte ha samma betydande effekt på betygssättningen som sambedömning har. I alla ämnesgrupper utom en minskade avidentifieringen avvikelsegraden något. För NO-prov hade avidentifieringen intressant nog motsatt effekt, det vill säga ledde till en ökad avvikelsegrad.

Att avidentifiering generellt sett inte har desto större effekt på avvikelsegraden tyder möjligen på att ordinarie ämneslärare via till exempel elevens handstil

eller uttrycksätt i många fall kan avgöra vilken elev som skrivit provet, även om elevsvaren aidentifierats på namnuppgifter.⁴

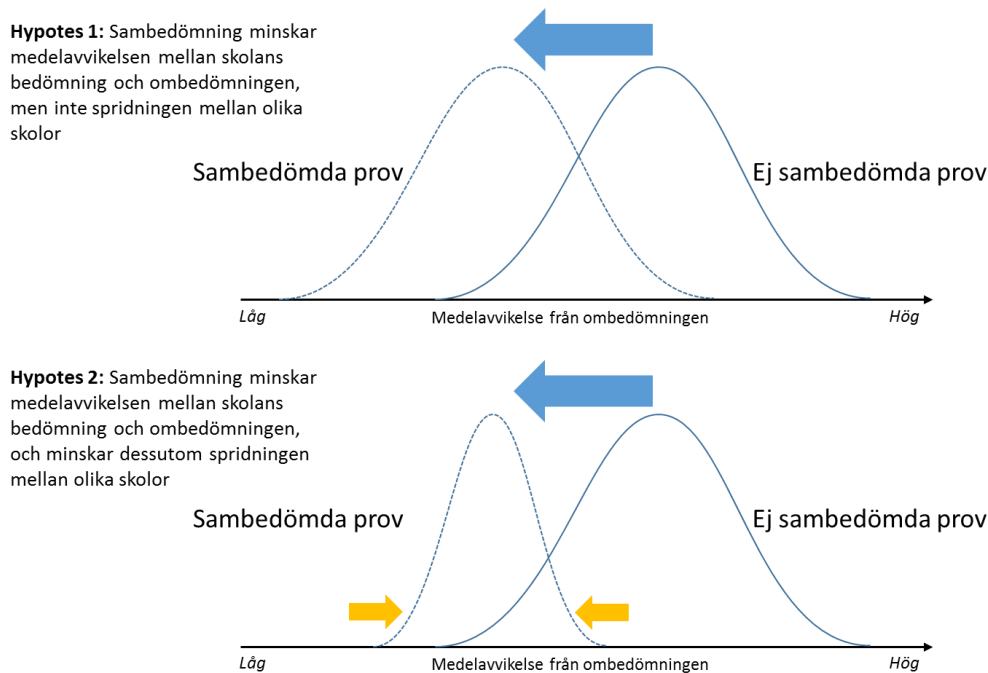


Figur 8: Genomsnittlig skillnad mellan skolans bedömning och omdömningen för prov som var aidentifierade vid bedömningen, respektive för prov där elevens identitet var känd för läraren.

Sambedömning och extern bedömning minskar avvikelsegraden

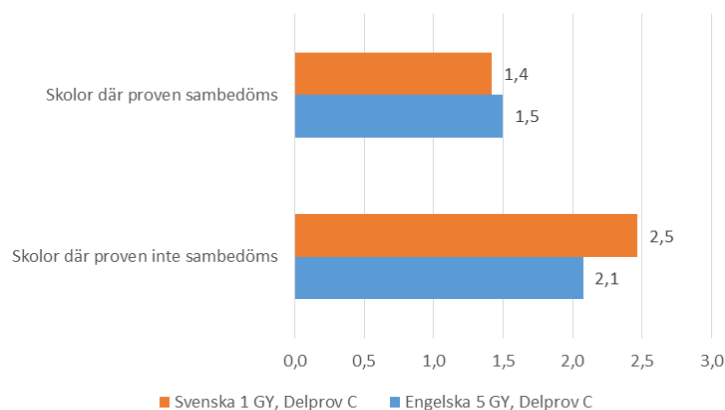
Föregående avsnitt visar att både extern bedömning och sambedömning generellt sett har en återhållande effekt på provbetygen. Detta resulterar i ett något lägre betyg i snitt för motsvarande elevlösning än om elevens ordinarie ämnesläraren gör bedömningen själv. Frågan är däremot om det även minskar *spridningen* i bedömning mellan skolor, eller om det bara är bedömningarnas medelpunkt som flyttas. Ur likvärdighetssynpunkt är nämligen spridningen mellan skolor det intressanta. Var "mittpunkten" i bedömningarna ligger är däremot relativt ointressant ur likvärdighetssynpunkt, givet att alla skolor lägger ribban på samma nivå.

⁴ Skolinspektionen (2015). Omdömning av nationella prov 2014: "Processerna spelar roll".



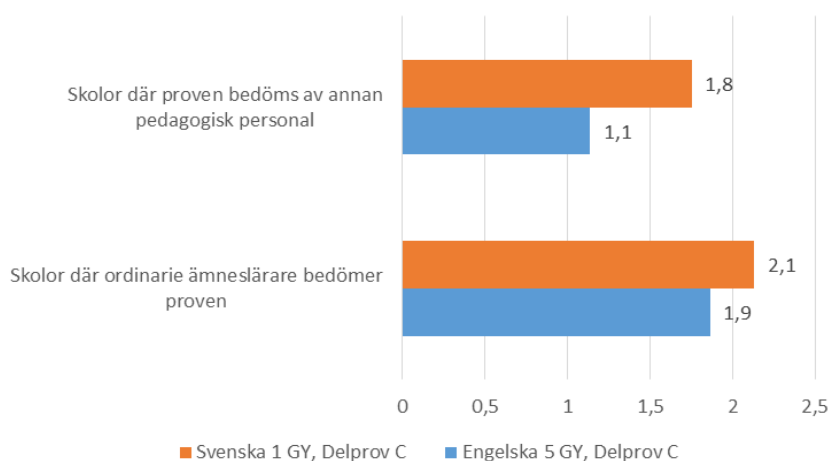
Figur 9: Två alternativa hypoteser kring effekten av sambedömning och extern bedömning

Denna analys är endast möjlig att företa för de ämnen där de allra flesta skolor inkluderats i ombedömningen, vilket denna ombedömningsomgång endast gäller Svenska 1 respektive Engelska 5 på gymnasiet. I övriga fall är antalet skolor i urvalet för litet för att några tydliga slutsatser ska kunna dras. Resultatet redovisas i figur 10 och 11 nedan.



Figur 10: Spridning i medelavvikelse från ombedömningen bland skolor som sambedömer sina prov respektive för skolor där sambedömning inte tillämpas. p75-p25 = skillnaden mellan den 75:e percentilens och den 25:e percentilens skolor, med avseende på genomsnittlig avvikelse från ombedömningen. 4 på skalan motsvarar ett betygsstegs skillnad på skalan A-F.

Resultaten ger en viss indikation på att extern bedömning möjligen kan minska, inte bara medelavvikelsen från ombedömningen, utan även spridningen mellan olika skolor. Skillnaden är däremot så pass liten att en slumpmässig tillfällighet inte kan uteslutas. För sambedömning kan å andra sidan en desto större spridning uppmätas, speciellt för proven i svenska.



Figur 11: Spridning i medelavvikelse från ombedömningen bland skolor där ordinarie ämneslärare bedömer proven respektive skolor där annan pedagogisk personal på eller utanför skolan bedömer proven. $p_{75}-p_{25}$ = skillnaden mellan den 75:e percentilens och den 25:e percentilens skolor, med avseende på genomsnittlig avvikelse från ombedömningen. 4 på skalan motsvarar ett betygsstegs skillnad på skalan A-F.

Att vare sig sambedömning eller extern bedömning löser problemet med spridning mellan skolor är utifrån omständigheterna inte särskilt förvånande. En sambedömargrupp kan nämligen ha goda förutsättningar att *sinsemellan* enas kring uttolkningen av bedömningsanvisningarna. Däremot hindrar det inte att en annan sambedömargrupp på annat håll i landet enas om en helt annan tolkning.

Resultatet från ombedömningen visar att både sambedömning som extern bedömning har en viss återhållande effekt på den genomsnittliga bedömningen av en given elevlösning (tenderar att leda till ett något lägre betyg i snitt). Skolornas olika tillvägagångssätt vid bedömningen kan därför mycket väl undergräva elevens rätt till en likvärdig bedömning.

Exempel från två skolor som arbetar med sina bedömningsprocesser

För att ge det statistiska dataunderlaget en kvalitativ dimension har Skolinspektionen inför redovisningen av denna ombedömningsomgång valt att exemplifiera hur två skolenheter arbetar med sin bedömningsprocess. Telefonintervjuer har genomförts med rektorer för två skolor som deltagit i ombedömningen. I detta kapitel görs nedslag i hur rektorerna beskriver arbetet med likvärdighet, bedömning och betygssättning på sina skolenheter. Finns det några framgångsfaktorer eller exempel att ta lärdom av?

I detta stycke utreds även i vilken utsträckning Skolverkets allmänna råd införlivats i skolornas arbete med bedömning och betygssättning. Skolverkets allmänna råd anger hur regleringen runt olika skolområden bör tolkas och implementeras i praktiken. Syftet med de allmänna råden är att ge lärare, rektorer och huvudmän det stöd de behöver för att ge eleven goda förutsättningar för lärande.

Skolenheter med olika resultat i ombedömningen valdes ut

De skolor som valts ut till intervju har ingått i ombedömningen som genomfördes 2014 och 2015, det vill säga den sjätte och sjunde ombedömningsomgången.

Skola 1: En gymnasieskola med enskild huvudman som ligger i en svensk storstad. Skolan har i båda ombedömningarna haft en mittenplacering sett till ombedömningsresultatet. Ombedömningen indikerar att skolan lyckats upprätta rutiner kring bedömningsprocessen och betygssättning, vilka gör att skolan erhållit en mittenplacering bland skolorna som deltog i bedömningen.

Skola 2: En kommunal mellanstadieskola på landsbygden som 2014 gjorde en generösare bedömning sett till genomsnittet för ombedömningen. I den senaste ombedömningsomgången har skolan däremot gjort en bedömning mera i linje med genomsnittet. För skola 2 är det en fråga om hur skolan arbetat med sina bedömningsprocesser sedan förra ombedömningen, och i sådana fall vilka åtgärder de vidtagit för att öka samstämmigheten inom lärarkollegiet.

Likvärdighet inom bedömning förutsätter att lärare samarbetar

Skolverkets allmänna råd kring bedömning och betygssättning betonar vikten av sambedömning och diskussion lärarkollegor emellan, som en grundförutsättning för uppnåendet av likvärdig bedömning. Enligt Skolverket bidrar diskussioner och analyser kring variationer i elevarbeten till att läraren utvecklar sin bedömarförmåga och känner sig förtrogen med kunskapskraven. Detta främjar i sin tur en likvärdig och rättvis bedömning av elevens kunskaper. Rektorn bär ansvaret för att organisera skolverksamheten så att lärare har möjlighet och tid att gemensamt diskutera bedömning och betygssättning.⁵ Rektorn bör i samråd med lärarna se till så att gemensamma rutiner och former för att utvärdera elevens kunskaper upprättats.⁶

⁵ Skolverket (2012). Skolverkets allmänna råd: Bedömning och betygssättning i gymnasieskolan. s. 17.

⁶ Skolverket (2011). Skolverkets allmänna råd: Planering och genomförande av undervisningen – för grundskolan, särskolan, specialskolan och sameskolan. s. 22.

Rektorerna på de båda skolorna påpekade att deras lärare generellt sett upplever perioden då nationella proven genomförs som särskilt stressig. Lärarna upplever att de inte hinner utföra alla arbetsuppgifter, då de nationella proven är så pass tidskrävande. Rektorn för skola 2 betonade vikten av att avvara tid för samordning av bedömningsarbetet vid just denna hektiska tid, vilket kan vara utmanande.

Rektorn för skola 2 understryker vikten av att stötta lärarna i samband med bedömningen av prov. De flesta lärare är vana vid de nationella proven, men många lärare, speciellt i de lägre årskurserna, kanske aldrig har satt betyg tidigare. Rektorn uppger att de flesta mellanstadielärare behöver mycket stöd vid betygssättning och att rektorns roll som stöd i samband med bedömningsprocessen då blir av desto större vikt.

Under 2015 har en kvalitetssamordnare inom kommunen där skola 2 ligger initierat en jämförelse mellan elevernas betyg och resultaten i de nationella proven bland kommunens skolor. Jämförelsen har öppnat upp för en diskussion kring bedömningsprocesserna på skolorna. Rektorn uppger vidare att rektorskollena inom kommunen har ett gott samarbete och att de sinsemellan träffas och diskuterar bedömningsprocesser, betygssättning eller andra områden de upplever att det finns behov av att diskutera.

Enligt Skolverkets råd är huvudmannens insats av särskild vikt för utvecklandet av bedömningsarbetet och betygssättning, som ska vara likvärdig med andra skolor. Exempel på sådana insatser kan vara upprättandet av nätverk och tekniska lösningar för att underlätta kommunikationen.⁷

Tillsättandet av en kvalitetssamordnare inom kommunen där skola 2 ligger är ett exempel på hur man på huvudmannanivå koordinerat arbetet för att öka samstämmigheten inom kommunens skolor gällande bedömning och betygssättning. Huruvida man på kommunal nivå upplevt att det funnits ett behov av att arbeta med betygssättning, och därmed tillsättandet av en kvalitetssamordnare, förblir däremot osagt.

För att avlasta de lärare som bär huvudansvaret för de nationella proven ser rektorn till att alla lärare på skola 2 är delaktiga i bedömningen av proven. Skolan byter även prov med en annan skola i kommunen, för att på så vis få proven bedömda av en extern bedömare, utan anknytning till eleven. Skola 1 tillämpar inte denna typ av skolöverskridande samedömning, utan samtliga prov bedöms av lärare på skolan.

Skolverket betonar, förutom samarbete inom skolan, även skolöverskridande samarbete för uppnåendet av en likvärdig bedömning och betygssättning. Huvudmannen ansvarar för att rektorer och lärare har möjlighet att diskutera bedömning och betygssättning, inte bara inom skolan utan även mellan olika skolor. Diskussioner kring bedömning och betygssättning med kollegor på andra skolor bidrar till att bedömningen blir mera tillförlitlig och minskar påverkan av lokala bedömningskulturer. Förutsättningarna för detta

⁷ Skolverket (2012). Skolverkets allmänna råd: Bedömning och betygssättning i gymnasieskolan. s. 17., Skolverket (2011). Skolverkets allmänna råd: Planering och genomförande av undervisningen – för grundskolan, särskolan, specialskolan och sameskolan. s. 27.

skolöverskridande samarbete är delvis beroende av huvudman och ort. På mindre skolor kanske det endast finns en lärare per undervisningsämne, vilket försvårar denna typ av samarbete.⁸

En levande diskussion kring bedömningsfrågor är en framgångsfaktor

Skolenheterna lyfter i samband med intervjun upp ett antal framgångsfaktorer för att uppnå samstämmighet kring bedömningsfrågor. Dessa summeras i korthet nedan:

- Avvara tid för samordning av bedömningsarbetet
- Diskutera resultat och göra jämförelser på kommunnivå mellan elevernas betyg och resultat från nationella prov
- Nätverk där rektorer och lärare kan diskutera bedömningsfrågor
- Rektorer stöttar lärarna i samband med betygssättning
- Att byta prov med en annan skola i sin kommun, för att på så vis få en extern bedömning av sina elever
- Att byta prov mellan lärare på skolan
- Arbeta i ämneslag med bedömning
- Diskussion kring bedömningsunderlaget och provfrågorna inför provtillfällena

Hur går bedömningsprocessen till?

I samband med bedömning av de nationella proven på skola 1 väljer läraren ut ett antal prov från sin klass som läraren anser representerar en god prestation. Dessa prov avidentifieras och bedöms därefter av andra lärare på skolan. Utöver det görs slumpvisa stickprov på prov som bedöms genom en organiserad process för sambedömning.

På skola 1 arbetar lärarna i ämneslag kring bedömning av elevens prestationer. Ämnesansvarig lärare för varje ämne kallar till möte en gång i månaden eller en gång var annan månad. Syftet med dessa arbetslag är att alla lärare ska få en uppfattning av var eleven ligger, även i andra ämnen än de läraren själv undervisar eleven i. Det ses som ett effektivt sätt att följa elevens utveckling.

Inför bedömningen av de nationella proven brukar provfrågorna och bedömningsunderlagen diskuteras på båda skolorna, för att säkerställa att samtliga lärare tolkar underlagen korrekt.

Enligt Skolverkets råd för bedömning och betygssättning bör rektorn för skolan se till så att rutiner för kunskapsbedömning vid arbetsplatsförlagt lärande finns etablerade. Likaså ska rektorn skapa förutsättningar för arbete med likvärdig bedömning och betygssättning.⁹ När diskussionen kring kvaliteten i elevens prestationer hålls levande bland lärarna, ökar bedömaröverensstämmelsen och likvärdigheten i bedömningen.¹⁰

⁸ Skolverket (2011). Skolverkets allmänna råd: Planering och genomförande av undervisningen – för grundskolan, särskolan, specialskolan och sameskolan. s. 27. Skolverket (2012). Skolverkets allmänna råd: Bedömning och betygssättning i gymnasieskolan. s.17, 24.

⁹ Skolverket (2012). Skolverkets allmänna råd: Bedömning och betygssättning i gymnasieskolan. s.15, 26.

¹⁰ Skolverket (2011). Kunskapsbedömning i skolan – praxis, begrepp, problem och möjligheter. s.35.

Fortbildning och uppföljning av bedömningsarbetet

I enlighet med skollagen ska huvudmannen se till så att personalen på skolan ges möjlighet till den kompetensutveckling som krävs för att de ska kunna utföra sina arbetsuppgifter efter rådande föreskrifter.¹¹

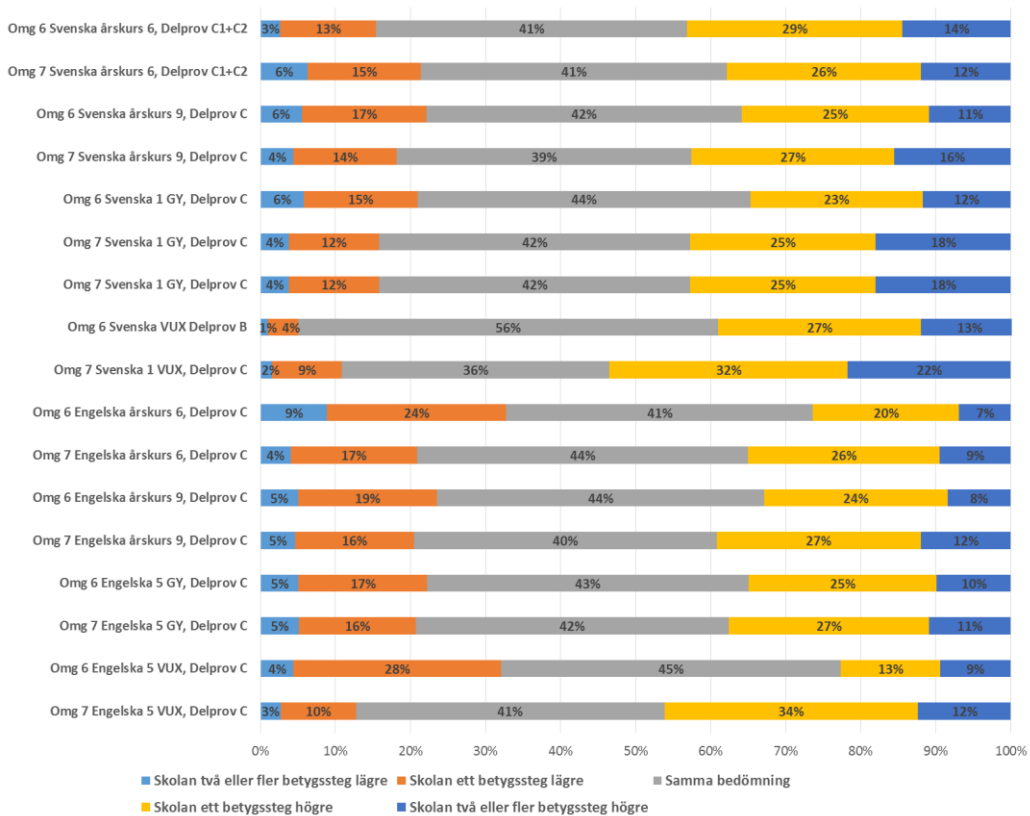
För att öka kompetensutvecklingen på skola 1 deltar lärarna på skolan i ämnesvisa fortbildningar, vilka ordnas för skolor med enskilda huvudmän inom kommunen. I samband med fortbildningarna diskuteras frågor som upplevs vara svåra, bland annat bedömning och betygssättning. Därutöver arbetar förstelärarna vid skola 1 aktivt med fortbildning och alternativa sätt att arbeta med formativ bedömning.

Rektorn på skola 1 betonar vikten av formativ bedömning på skolan, där kontinuerlig uppföljning av elevens kunskaper ger både lärare och eleven en god uppfattning av den faktiska kunskapsnivån eleven ligger på. Möjligen kan det formativa inslaget ha en positiv effekt för bedömningssamstämmigheten, då läraren har möjlighet att kontinuerligt följa elevens kunskapsutveckling över tid. Måhända minskar effekten av lärarens förutfattade uppfattning om elevens kunskaper, den nivå läraren *tror* att eleven ligger på, då kunskapsnivån kontinuerligt följs upp.

¹¹ 2 kap. 34 § skollagen (2010:800).

Hur mycket skiljer sig bedömningarna av samma elevlösning?

För proven i engelska och svenska ges delprovsbetyg, som i sin tur vägs samman till ett provbetyg. De omdömda språkproven består av uppsatser där eleven har skrivit fritextsvar. I cirka 41 procent av fallen har skolans lärare satt *samma* provbetyg som omdömdaren.¹²



Figur 12: Andel avvikelser mellan bedömningen från skolans lärare och omdömdarens bedömningar, för delprov i svenska och engelska, för omgång 6 och 7.

¹² På betygsskalan A-F. För proven i engelska och svenska ges delprovsbetyg som i sin tur vägs samman till ett provbetyg enligt givna modeller.

Bedömningarna skiljer sig åt för mer än varannan elevlösning

I 18 procent av fallen har skolans lärare satt ett delprovsbetyg som är *minst* ett betygssteg lägre än omdedömarens och i 41 procent av fallen har skolans lärare satt ett delprovsbetyg som är *minst* ett betygssteg högre än omdedömarens. Andelen avvikelser är ungefär den samma för samtliga årskurser.

Små uppmätta skillnader för språkproven mellan omgång 6 och 7

I figur 12 ovan visas även avvikelserna från respektive språkprov från omgång 6, för att tydliggöra hur avvikelserna ser ut vid de två senaste omgångarna.¹³ Förutom för gymnasieskolan kunde inga större skillnader för ämnet svenska konstateras mellan de två omdömningsomgångarna. Tydligast blir skillnaden i de fall där skolans lärare satt ett delprovsbetyg som är *minst* två betygssteg högre än omdedömarens, där har det skett en ökning med 6 procentenheter jämfört med föregående omgång. I absoluta termer kan skillnaden betraktas som liten, men vid en jämförelse med övriga årskurser blir förändringen påtaglig.

Inga större skillnader kunde uppmätas för delproven i engelska. Den enda skillnaden som sticker ut är för årskurs 6, där andelen delprovsbetyg som är *minst* ett betygssteg lägre än omdedömarens har sjunkit med totalt 12 procentenheter. Samtidigt har andelen provbetyg som är *minst* ett betygssteg högre än Skolinspektionens omdömnings betyg stigit med totalt 8 procentenheter för årskurs 6.

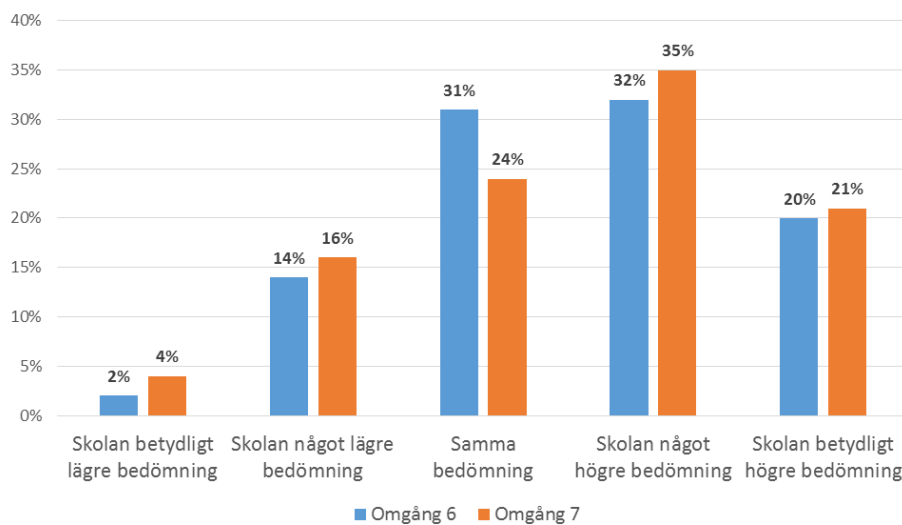
Inga skillnader totalt sett för NO-prov mellan åren

För NO- och SO-prov sätts inga delprovsbetyg utan istället delas det ut *belägg*. I de fall där bokstavsprov inte förekommer har beläggen räknats om till en gemensam skala som sträcker sig från 0-20 för att ge en mer lättbegriplig bild över hur avvikelserna ser ut.¹⁴

¹³ Notera att proven i svenska inom den kommunala vuxenutbildningen avser två olika delprov, varför jämförelser mellan åren inte är möjliga.

¹⁴ I bilaga 2 redovisas i detalj hur resultatmättet beräknats.

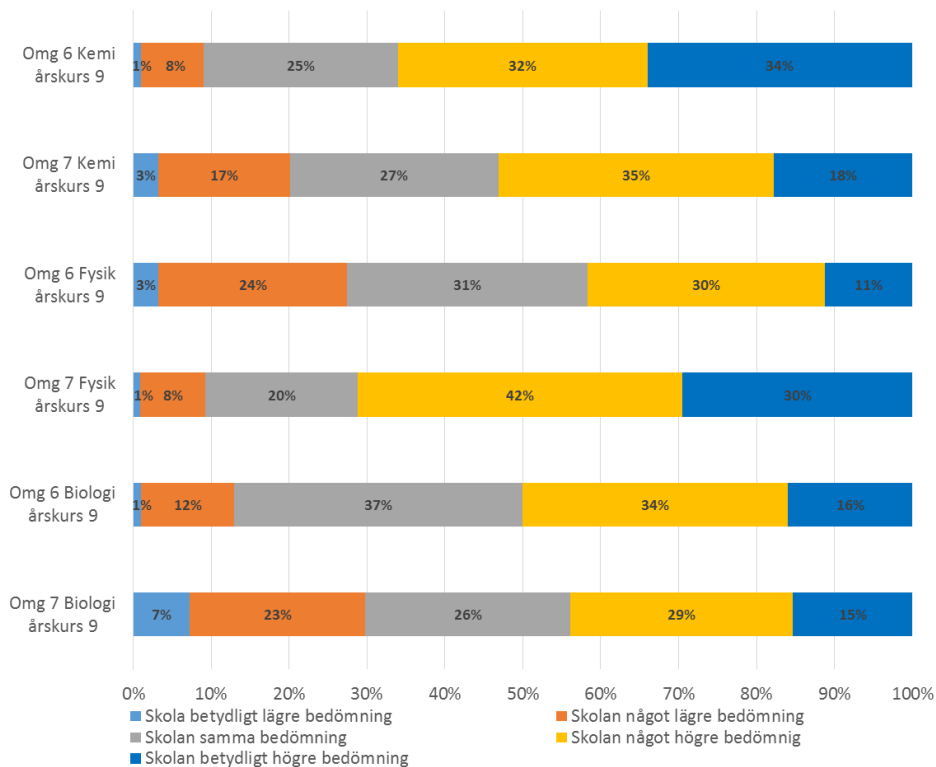
NO-ämnena har totalt sett inga betydande avvikelser mellan de två senaste omgångarna. Däremot kan man notera att avvikelserna ökat något för samtliga kategorier, förutom för de fall där skolans lärare och omdömare gjort samma bedömning. Omdömare har i en knapp fjärdedel av bedömningarna (24 procent) gjort samma bedömning som skolans lärare. I drygt hälften av fallen (56 procent) har skolans lärare gjort en bedömning som är något högre/betydligt högre än vad omdömare gjort. I 20 procent av fallen har skolans lärare gjort en bedömning som är något lägre/betydligt lägre.



Figur 13: Andel avvikelser mellan bedömningen från skolans lärare och omdömarens bedömningar, för delprov i NO-ämnena i årskurs 9, för omgång 6 och omgång 7.

Störst skillnad mellan åren för fysik

En jämförelse mellan de två senaste omgångarna visar att avvikelserna skiljer sig mellan de olika ämnena. Den mest påtagliga avvikelserna är för fysik, där det skiljer 19 procentenheter mellan omgångarna för andelen delprovsbetyg som av skolans lärare bedömts som betydligt högre.



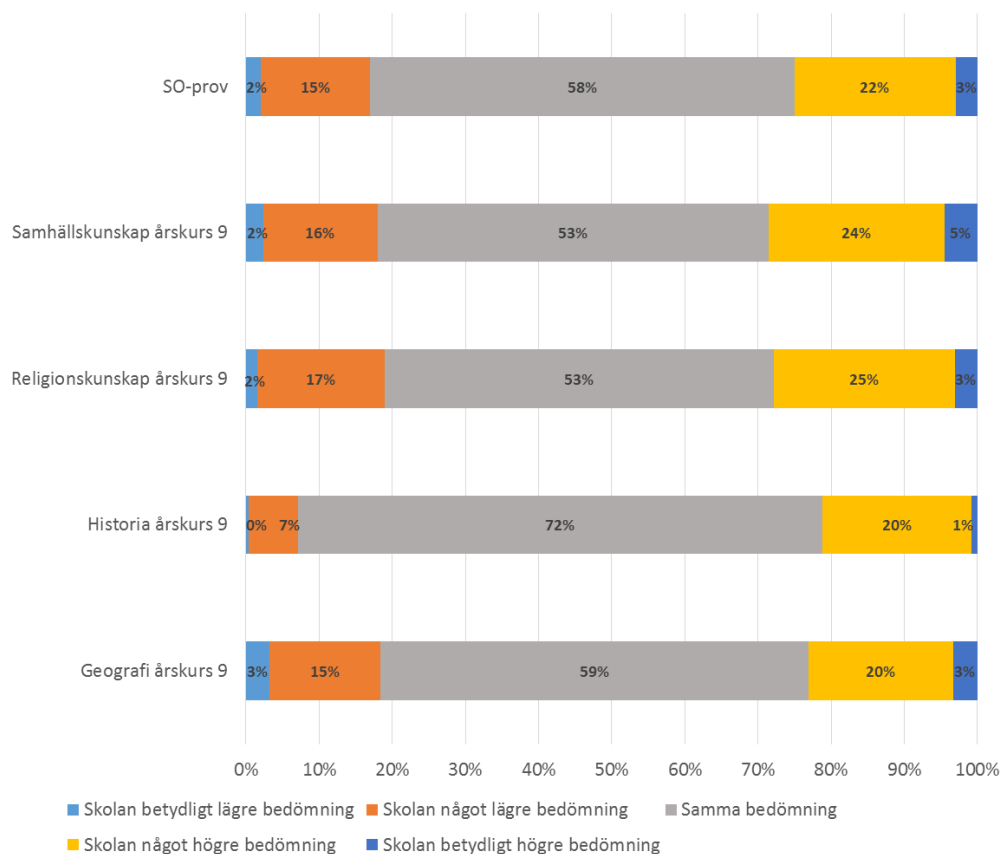
Figur 14: Andel avvikelser mellan bedömningen från skolans lärare och omdömarens bedömningar, för delprov i NO-ämnena i årskurs 9, för omgång 6 och 7.

SO-prov bedöms i hög grad likvärdigt

För SO-proven totalt sett har skolans lärare i 58 procent av fallen gjort samma bedömning som omdömarens.¹⁵ I 25 procent av fallen har skolans lärare gjort en bedömning som är något högre/betydligt högre än den omdömarens gjort och i 17 procent av fallen har skolans lärare gjort en bedömning som är något

¹⁵ För SO-ämnena har inte samma delprov omdömts den sjätte och sjunde omdömningsomgången, vilket gör att jämförelser mellan åren inte är möjligt att göra.

lägre/betydligt lägre än den omdömdaren gjort. De mest extrema avvikelserna är i princip obefintliga för SO-ämnet. För de omdömda proven i historia finns det inga prov alls där skolans lärare gjort en betydligt lägre bedömning. För övriga SO-ämnena är andelen mycket låg. För den här omgången är det vid SO-ämnena som skolans lärares bedömningar i störst utsträckning överensstämmer med omdömdarens.



Figur 15: Andel avvikelser mellan bedömningen från skolans lärare och omdömdarens bedömningar, för delprov och totalt sett för SO-ämnena i årskurs 9, för omgång 7.

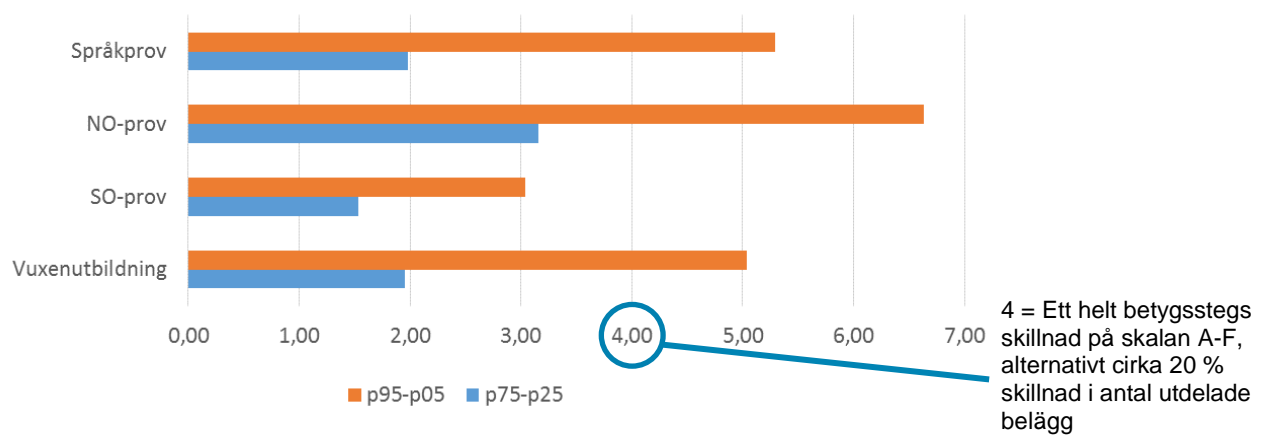
Stor variation mellan olika skolor för samma elevlösning

Utifrån ett likvärdighetsperspektiv är det mer intressant att jämföra spridningen *mellan olika skolor än* mellan en viss skola och omdömdarens.

Eftersom de flesta skolorna bedömer sina egna prov är det inte möjligt att utifrån omdömdarens analys analysera andelen avvikelser mellan olika skolor. Resultaten från omdömdarens analys kan däremot användas som en gemensam referenspunkt för att undersöka om olika skolor *i genomsnitt* sätter lika höga eller

låga provbetyg för likvärdiga elevlösningar.¹⁶ I Bilaga 3 beskrivs tillvägagångssättet i beräkningarna närmare.

I figur 16 nedan redovisas den beräknade genomsnittliga skillnaden av likvärdiga elevlösningar mellan olika skolor, redovisat per ämnesgrupp. Resultaten visar att en elev på en skola som sätter relativt låga delprovsbetyg kan förväntas få ett delprovsbetyg som ligger cirka ett halvt till ett betygssteg lägre än om eleven lämnat in samma svar på en skola som sätter relativt höga delprovsbetyg.



Figur 16: Estimat av spridning i bedömning av likvärdiga elevlösningar mellan olika skolor. p75-p25 betyder skillnaden mellan den 75:e percentilens och den 25:e percentilens skolor, med avseende på genomsnittlig avvikelse från ombedömningen.

Sett till de skolor som placerar sig längst ute på kanterna så växer skillnaderna ytterligare. En elev på en skola som tillhör de fem procent mest strikta skolorna i landet sett till betygssättning kommer typiskt sett att få ett provbetyg som ligger runt ett till två betygssteg lägre på skalan A-F, än om eleven lämnat in samma svar på en av de fem procent mest generösa skolorna.

Spannet mellan de skolor som är relativt strikta respektive relativt generösa i betygssättningen vad gäller delprovsbetyg är i princip lika stort för de ombedömda språkproven samt för delprov skrivna inom den kommunala vuxenutbildningen. Däremot kan en tydlig skillnad mellan NO- och SO-ämnena¹⁷ uppmätas, där NO-provens spridning är klart högre.

¹⁶ "Likvärdiga elevlösningar" syftar här på en hypotetisk situation där olika skolor bedömt samma elevlösningar.

¹⁷ Uttryckt på den harmoniserade bedömningsskalan 0–20. För delproven i NO- och SO- ämnen ges inga bokstavs-betyg.

Kan skolan förklara avvikelsegraden?

Att det förekommer skillnader i hur en elevlösning bedöms av olika sinsemellan oberoende omdömare säger inget i sig om varför så är fallet. En hypotes, som kan testas utifrån omdömningsresultaten, är ifall det förekommer systematiska skoleffekter.

Systematiska skoleffekter. Uppmätta skillnaderna mellan skolor beror på att vissa skolor har *skolövergripande processer* eller en informell *skolkultur* av generös alternativt strikt betygssättning, som färgar av sig på merparten av lärarna. Det finns sedan tidigare indikationer på att det är lättare att nå ett högre slutbetyg på skolor där snittnivån är låg, än på skolor där snittnivån är hög.¹⁸ De tillgängliga register som används medger inte att svaren analyseras på lärarnivå, utan endast på skolnivå. För flera skolor omdöms prov från två olika ämnen. Om systematiska skoleffekter spelar en stor roll bör det finnas en korrelation mellan avvikelsegraden i förhållande till omdömningen för de båda ämnena.

Svag positiv korrelation för de systematiska skoleffekterna

För 433 skolenheter finns omdömningsresultat för två ämnen. För dessa 433 skolor finns en statistiskt signifikant skoleffekt, definierad som korrelationen mellan genomsnittlig avvikelsegrad för det ena ämnet, det vill säga "ämne 1" och genomsnittlig avvikelsegrad för det andra ämnet, det vill säga "ämne 2". Den uppmätta korrelationen mellan de två ämnena är 0,198.¹⁹ Ett samband förekommer, men är inte särskilt starkt.

Ett annat sätt att uttrycka sambandets styrka på är följande: Ifall en skola ligger över den genomsnittliga avvikelser i "ämne 1" så är det 38 % sannolikhet att samma skola hamnar under den genomsnittliga avvikelser för "ämne 2". Det innebär även att samma skola har 62 % sannolikhet att hamna ovanför den genomsnittliga avvikelser för "ämne 2" ifall man hamnat ovanför den genomsnittliga avvikelser för "ämne 1".

Om sambandsanalysen begränsas till språkproven i svenska och engelska, vilket kan göras i gymnasiet, är sambandet aningen starkare: korrelationen är här 0,221 mellan de två avvikelsegraderna. Skoleffekten förekommer men den är inte särskilt påtaglig, eftersom korrelationen mellan de genomsnittliga avvikelserna är svag.

¹⁸ Se till exempel Fredriksson, Peter & Vlachos, Jonas (2011). Reformen och resultat: kommer regeringens utbildningsreformer att ha någon betydelse? s. 29.

¹⁹ Single Measure Intraclass Correlation Coefficient.

Det innebär att ifall en skola ligger över den genomsnittliga avvikelsen i "ämne 1" så är det 43 % sannolikhet att samma skola hamnar under den genomsnittliga avvikelsen för "ämne 2". Det innebär även att samma skola har 57 % sannolikhet att hamna ovanför den genomsnittliga avvikelsen för "ämne 2" ifall man hamnat ovanför den genomsnittliga avvikelsen för "ämne 1".

En möjlig förklaring till de uppmätta skoleffekterna är skolövergripande skillnader i bedömningsprocesserna, eftersom dessa har en effekt på betygssättningen för delproven.²⁰ Till exempel kan det vara så att vissa skolor konsekvent tillämpar sambedömning för alla ämnen, och vissa inte gör det.

Antalet skolor som ingår i ombedömningen är för litet för att kunna utreda denna fråga grundligt. Korrelationen mellan de två ämnesresultaten är signifikant, och i samma storleksordning som tidigare, vilket antyder att de uppmätta skoleffekterna inte i någon större utsträckning kan förklaras av systematiska skillnader i bedömningsprocessen på skolnivå.

Sammanfattningsvis bör det poängteras att datamaterialet inte är tillräckligt stort för att med säkerhet kunna svara på frågan om det finns systematiska skoleffekter. Eventuellt kan skillnader i bedömningsprocesser mellan skolorna förklara en del av den uppmätta skoleffekten, men de kan inte förklara hela effekten.

Resultaten från ombedömningen indikerar att en så kallad skoleffekt (systematiska effekter på skolnivå), förekommer men att skoleffekten bara förklarar en viss del av bedömaravvikelserna. Det betyder inte att återstående variation är osystematisk – flera andra typer av systematiska effekter är tänkbara, till exempel systematiska effekter med avseende på lärar- eller elevgenskaper (ålder, kön, bakgrund och så vidare).²¹ Däremot kan dessa effekter inte kontrolleras inom ramen för detta uppdrag.

Små avvikelser mellan kommunala och fristående skolor

Skillnaden i andelen betydande avvikelser mellan de olika huvudmannatyperna är liten. För skolor med enskilda huvudmän har 80 procent av proven fått en likartad ombedömning och för skolor med kommunal huvudman är motsva-

²⁰ Skolinspektionen (2015). Ombedömning av nationella prov 2014: "Processerna spelar roll".

²¹ Vissa av dessa aspekter har dock belysts i tidigare omgångar av ombedömningen, se t ex Skolinspektionen (2012): "Lika för alla: Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år. Resultatbilaga" s. 63.

rande andel 84 procent. Däremot är andelen prov som bedömts generösare något högre bland de enskilda skolorna (17 procent) jämfört med de kommunala (11 procent) skolorna.

Avvikelserna *kan* ha en avgörande betydelse för det slutgiltiga provbetyget

Eftersom endast ett delprov per prov ombedöms är det inte givet att en uppmätt skillnad för ett enskilt delprov får genomslag på det slutliga provbetyget. Ofta räcker det däremot med ett betygsstegs skillnad för ett delprov för att det totala provbetyget ska påverkas. Ett exempel är NO-proven i årskurs 9, där en skillnad på 2–3 belägg, vilket är vanligt förekommande för de ombedömda delproven, inte sällan påverkar det totala provbetyget.

Det står alltså klart att de avvikelser som uppmäts i ombedömningen på delprovsnivå är av den storleksordningen att det *kan* leda till att det totala provbetyget påverkas, även om det långt ifrån alltid är fallet.

Nationella prov är inga examensprov, utan bör betraktas som summativa. Detta innebär att elevens slutbetyg/kursbetyg inte behöver överensstämma med resultatet på det nationella provet. Läraren ska utgå ifrån elevens prestationer på de nationella proven, men förväntas även utnyttja annan tillgänglig information om elevens kunskaper i förhållande till kunskapskraven.²²

²² Skolverket (2009). Likvärdig betygssättning i gymnasieskolan? En analys av samband mellan nationella prov och kursbetyg.

Utvecklingsområden

Alla elever har rätt att bli bedömda utifrån samma grunder. Betygsättning i skolan innebär en myndighetsutövning vilket kan ha en avgörande betydelse för elevens framtida förutsättningar att lyckas i livet. Betyg har en direkt koppling till behörighet och urval för vidare studier och kan även vara av avgörande betydelse för att lyckas på en konkurrenspräglad arbetsmarknad. Därför är det viktigt att alla elever, oavsett vilken lärare de haft eller vilken skola de gått i, har blivit bedömda utifrån samma kriterier.

I ljuset av resultaten från föreliggande rapport vill Skolinspektionen belysa ett antal utvecklingsområden, vilka kan bidra till att öka samstämmigheten mellan bedömningarna av nationella prov.

Sambedömning ökar likvärdigheten

Olika former av sambedömning har en utjämnande effekt på bedömningsavvikelserna, vilket tydligt framgår i resultatet från ombedömningarna. Ombedömningen indikerar att prov som bedömts av pedagogisk personal utanför skolan uppvisar färre avvikelser än prov som bedömts av lärare på elevens skola.

Aidentifiering minskar avvikelsegraden

Aidentifiering av elevens identitet förfaller ha en viss återhållande effekt på lärarens bedömning. Resultatet från ombedömningen visar att många lärare har svårt att förhålla sig objektiva vid bedömning av sina egna elevers provsvar. Aidentifiering av proven har visat sig minska avvikelsegraden något, och därmed främja en likvärdig bedömning.

Digitalisering av de nationella proven främjar likvärdighet

Digitaliseringen innebär att prov genomförs på dator eller med hjälp av andra digitala verktyg. Mycket pekar på att digitalt genomförda prov kan bidra till att utjämna ojämlikheter i skolan, då en rätt genomförd digitalisering kan bidra till främjandet av likvärdiga bedömningsgrunder för landets elever.

I betänkandet *”Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning”* från 2016 lyfts för- respektive nackdelarna med digitalisering av de nationella proven upp.²³ Elever med skrivsvårigheter kan få stöd i

²³ Betänkandet *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning* (SOU 2016:25) s. 160-179.

stavning med hjälp av standardiserade stavningskorrigeringsfunktioner. Även elever med funktionsnedsättningar kan med hjälp av olika digitala verktyg dra fördel av digitaliseringen.

Om proven skrivs på dator underlättar det avidentifieringen avsevärt, jämfört med om proven skrivs för hand. Handstil är nämligen en betydande faktor som gör att specifika elevlösningar kan kopplas till enskilda elever, trots att namnet maskerats. Även möjligheten för skolor att bedöma varandras prov kan främjas av digitaliseringen, då elektroniska elevlösningar är betydligt enklare att skicka mellan skolenheter än prov skrivna på papper.

Samverkan mellan skolor ökar samstämmigheten i bedömningen

Sambedömning *inom* skolor har endast ett begränsat potential att öka likvärdigheten i bedömningen, medan sambedömning *mellan* olika skolor har desto större genomslagskraft i termer av ökad likvärdighet.

Bedömningsprocesserna har en systematisk effekt på betygssättningen. Resultaten från ombedömningen indikerar att inslaget av sambedömning varierar kraftigt både mellan skolor och mellan de olika årskurserna, liksom mellan skolor med kommunal respektive enskild huvudman. Denna "spretighet" på systemnivå riskerar att utgöra ett hot mot likvärdigheten i bedömningen.

Rektorns betydelse för uppnåendet av en likvärdig bedömning

Många rektorer överlämnar helt betygssättningen till sina lärare, utan att själva vara delaktiga i processen. Betygssättningen i svenska skolor bör vara likvärdig och fungera som ett rättssäkert behörighets- och urvalsinstrument inför framtida studier och arbetsmöjligheter.

Det är den enskilde läraren som är ansvarig för att sätta betyget, men det är rektorn som ansvarar för att säkerställa att lärarna är förtrodda med det regelverk som gäller. Rektorn ska även leda arbetet för att garantera att läraren får det stöd den behöver för att kunna sätta rättvisa och likvärdiga betyg. Rektorn ansvarar för att den myndighetsutövning läraren anförtros inte missbrukas. För oerfarna lärare är rektorns stöd vid betygssättning av extra stor betydelse.

Provens och bedömningsanvisningarnas utformning är avgörande

Bedömningsanvisningarnas utformning är en avgörande faktor för uppnåendet av en likvärdig bedömning av elevsvar.

Delprov där eleven ska ge sitt svar i form av en längre sammanhängande text tenderar generellt sett att påvisa en hög avvikelsegrad. Det är viktigt att ta det i

beaktande när proven och principerna för sammanvägning till ett samlat provbetyg konstrueras.

Skolinspektionen välkomnar det arbete som för närvarande bedrivs för att samordna struktur och terminologi mellan de olika ämnesproven. Långsiktigt torde det ge bättre förutsättningar för systematiskt utvärderings- och utvecklingsarbete på systemnivå än vad som idag är fallet.

Bilaga 1: Genomförandet av ombedömningen

Urval av prov och skolor

2015 års ombedömning innefattar nationella prov som under året skrevs i årskurs 6, årskurs 9, gymnasiets grundläggande kurser, samt motsvarande kurser inom den kommunala vuxenutbildningen på gymnasial nivå. För den kommunala vuxenutbildningen ombedömdes prov skrivna 2014.

Skolor som ingår i ombedömningen väljs ut genom ett slumpvist urval bland skolor vars huvudmän följande år ingår i Skolinspektionens regelbundna tillsyn. Urvalet av skolor för ombedömning år 2015 gjordes bland skolor vars huvudmän ingår i tillsynsplanen för 2016. Detta för att ett så aktuellt resultat som möjligt ska kunna användas som underlag vid tillsyn.

När ursprungs rättningen av proven genomfördes kände skolorna ännu inte till att de skulle ingå i urvalet för ombedömningen av nationella prov. Skolinspektionen bedömer därför att risken för att skolorna i förväg skulle kunnat ha manipulerat proven, med avsikt att uppnå ett fördelaktigt resultat, vara i princip obefintlig.

Ämnen, delprov och årskurser som ingår i ombedömningen

För att kunna genomföra ombedömningen vid så många skolor som möjligt ombedöms endast ett eller två ämnen per skola, i regel ett språkämne och ett NO- eller SO-ämne. En slumpvis tilldelning av ämnen görs bland de skolor som valts ut till ombedömningen. Därefter görs ett slumpvist urval bland de elever som skrivit provet den aktuella terminen. Urvalet görs tillräckligt stort för att resultaten ska vara representativa på skolnivå. Det maximala antalet prov som ombedöms har begränsats till i regel 100 för grundskolor och 200 för gymnasieskolor. Om fler prov krävs för att uppnå representativitet för två ämnen väljs endast ett ämne för den aktuella skolan.

Gymnasieskolorna har möjlighet att genomföra antingen en höstterminsversion eller en vårterminsversion av varje prov. Skolinspektionen har valt att endast ombedöma vårterminsversionen, då denna provversion skrivs av cirka tre fjärdedelar av landets elever och skolor och därmed ger ett bredare provunderlag att utgå ifrån.

Hela provet omdöms inte, endast längre skrivuppgifter

Alla nationella prov är uppdelade i delprov. Vissa delprov innehåller laborativa eller muntliga element, eller uppgifter som genomförs i grupp. Denna typ av delprov är svåra, om än inte omöjliga att omdöma. Tidigare erfarenhet har även visat att större avvikelser är mycket sällsynt för denna typ av prov.

Av ovanstående skäl omdöms inte hela prov, utan endast ett eller två delprov. De delprov som väljs ut för omdömning är sådana som är möjliga att omdöma (som saknar laborativa eller muntliga inslag), men där det inte finns tydliga "rätt eller fel svar" på frågorna. I praktiken betyder det att de prov som väljs till stor del består av uppsatsfrågor, där eleven fritt ska formulera sig utifrån en öppen frågeställning.

Prov i matematik omdömdes inte

Tidigare omdömningsomgångar har matematikprov omdömts. Däremot har ämnet av tradition saknat längre skrivuppgifter, varpå endast små avvikelser mellan ursprungsbedömningen och omdömningen kunnat påvisas inom ämnet. Av ovanstående orsak har inte matematikprov omdömts denna omgång. Skolinspektionen har valt att fokusera på andra ämnen, där större utrymme för subjektivitet i tolkningen av provsvaren ges.

Urvalets sammansättning

Andelen omdömda prov från skolor med enskild huvudman som ingår i urvalet är högre än riksgenomsnittet. Andelen är högre för gymnasieskolan, där 55 procent av omdömda skolor och 53 procent av omdömda prov är skrivna på skolor med enskild huvudman. Vid tiden då urvalet drogs uppgick andelen gymnasieskolor med enskild huvudman till 34 procent av den totala andelen gymnasieskolor i landet.

Urvalet exkluderar vidare de skolor där färre än tio prov för ett visst ämne skrivits, eftersom de inte redovisas av sekretesskäl. För små insamlingsgrupper medför nämligen att en röjningsrisk uppstår.

Terminologisk översikt

Skolinspektionen kan inte fastslå att ombedömarnas bedömningar är mer korrekta än de bedömningar som ursprungs rättarna ute på skolorna gjort. I denna rapport talas det därför i termer av *avvikelser* mellan skolans ursprungsrättning och Skolinspektionens ombedömning, snarare än korrekta eller felaktiga bedömningar. Med avvikelse avses en skillnad i bedömningen av en elevprestation med minst ett betygssteg på en standardiserad skala.²⁴

Generös bedömning: Ursprungsrättaren på skolan har bedömt att elevens svar motsvarar ett högre betyg än Skolinspektionens ombedömning.

Strikt bedömning: Ursprungsrättaren på skolan har bedömt att elevens svar motsvarar ett lägre betyg än Skolinspektionens ombedömning.

Betydande avvikelse: En avvikelse på två betygssteg eller mer mellan Skolinspektionens och skolans bedömning.

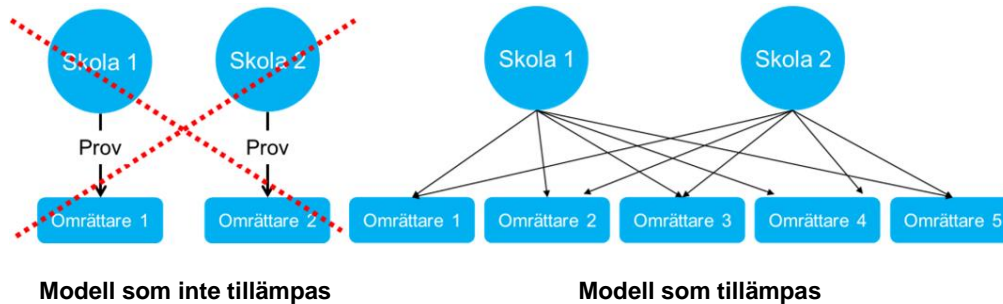
Då Skolinspektionen inte kan säga vilken av de båda bedömningarna som är mera "korrekt", läggs inte heller någon värdering vid begreppen. Det är snarare ett sätt att åskådliggöra i vilken riktning skolorna placerat sig i förhållande till Skolinspektionens ombedömning.

Fördelning av prov mellan ombedömning

Skolor som ingår i ombedömningen ombeds skicka in kopior av bedömda prov till Skolinspektionen. Proven maskeras och avidentifieras därefter på personuppgifter, bedömningsanteckningar och andra uppgifter från ursprungsrättningen. Därefter bedöms proven på nytt av yrkesverksamma och erfarna lärare som rekryterats av Skolinspektionen. Ursprungsbedömningen jämförs sedan med ombedömningen, och resultatet sammanställs i skolvisa rapporter.

Skolinspektionens lärare ombedömer sina prov enskilt. Samedömning tillämpas inte. För att minimera risken att en enskild ombedömares bedömningar får stort genomslag i de sammantagna resultaten sprids elevlösningar för varje enskild skola ut bland flera olika ombedömning. Jämförelsen av resultatet mellan skolor blir på så vis så rättvist som möjligt. Beroende på ämne och årskurs fördelas proven mellan 5–26 olika ombedömning.

²⁴ Det harmoniserade måttet för mätandet av skillnader i bedömningen beskrivs i bilaga 2.



Figur 17: Schematisk bild över fördelningen av prov mellan omdömare. Prov från varje enskild skola fördelas så långt som möjligt ut till flera olika omdömare. Detta för att minimera risken för att den enskilda omdömarens bedömningar får för stort genomslag i de sammantagna resultaten.

För prov i svenska och engelska där det skiljer mer än två betygssteg mellan bedömning och omdömning, alternativt om ursprungsbedömaren ger godkänt betyg (A–E) men omdömare ger ett F, eller vice versa, går provet vidare till en andra omdömning. Slutbetyget från omdömningen blir då resultatet från den av de två omdömningarna som ligger närmast ursprungsbedömningen. Syftet med andra omdömningen är att kvalitetssäkra prov med betydande avvikelser, för att garantera att Skolinspektionens lärare inte bedömt proven felaktigt.

Omfattning och bortfall

Totalt 43 881 delprov från 627 skolor har omdömts i 2015 års omgång. Omdömningen genomfördes av totalt 179 lärare som anställdes av Skolinspektionen. Fördelningen över ämne och delprov framgår av tabell 1 nedan.

Tabell 1: Externt och internt bortfall av prov. Antal begärda prov är totala antalet prov som begärdes in från skolorna.

	Antal begärda prov	Från antal skolor	Antal inkomna prov	Från antal skolor	Externt bortfall (prov)	Antal kompletta omedömda prov	Från antal skolor	Internt bortfall (prov)	Totalt bortfall	Antal omedömdare
Totalt:	50 783	656	47 515	641	6 %	43 881	627	8 %	14 %	179
<i>Ämne och delprov:</i>										
Engelska åk 6 delprov C	3 718	126	3 669	126	1 %	3 456	123	6 %	7 %	10
Engelska åk 9 delprov C	6 726	145	6 371	141	5 %	6 055	140	5 %	10 %	20
Engelska Gy delprov C	8 341	137	7 685	132	8 %	7 305	129	5 %	12 %	25
Engelska Vux delprov C	853	28	710	26	17 %	598	23	16 %	30 %	12
Svenska åk 6 delprov C1+C2	3 744	126	3 646	126	3 %	3 362	121	8 %	10 %	15
Svenska åk 9 delprov C	7 091	175	6 904	173	3 %	6 553	171	5 %	8 %	23
Svenska Gy delprov C	8 708	140	7 885	134	9 %	7 626	131	3 %	12 %	26
Svenska Vux delprov C	1 560	41	1 156	37	26 %	916	28	21 %	41 %	13
Biologi åk 9 delprov A2	1 241	29	1 228	29	1 %	1 141	27	7 %	8 %	5
Fysik åk 9 delprov A2	1 180	34	1 130	33	4 %	1 043	33	8 %	12 %	5
Kemi åk 9 delprov A2	1 451	30	1 399	30	4 %	1 189	25	15 %	18 %	5
Geografi åk 9 delprov A+B	1 718	29	1 645	29	4 %	1 457	28	11 %	15 %	9
Samhällskunskap åk 9 delprov A+B	1 648	29	1 591	29	3 %	1 101	24	31 %	33 %	9
Historia åk 9 delprov A+B	997	29	952	28	5 %	755	24	21 %	24 %	7
Religion åk 9 delprov A+B	1 807	41	1 544	39	15 %	1 324	35	14 %	27 %	11

Bortfall

Det *externa bortfallet* utgörs av prov som begärts in, men som inte inkommit från skolorna. Elevlösningar begärs in från skolorna utifrån föregående års elevlistor, varför till exempel elever som bytt skola under året utgör en del av det externa bortfallet. Till detta tillkommer elever som inte skrivit provet på grund av sjukdom eller frånvaro av annan orsak. Det händer även att skolor inte kan återfinna proven eftersom ansvarig lärare slutat.

Internt bortfall utgörs av prov som inkommit från skolorna, men som av olika orsaker inte kunnat analyseras. De vanligaste orsakerna är att provkopiorna håller för dålig kvalitet (kopieringen har inte tillräcklig svärta), sidor saknas, eller att resultatet inte rapporterats in på korrekt sätt. Till det interna bortfallet räknas även prov från skolenheter där mindre än 10 prov kunnat ombedömas. Av sekretesskäl exkluderas i sådana fall hela skolans resultat från analysen.

Vad beror bortfallet på?

Det sammanvägda bortfallet varierar kraftigt från delprov till delprov. Bortfallet är särskilt högt för den kommunala vuxenutbildningen samt för prov i grundskolan där resultatet rapporteras in i form av summering av antal belägg, där problem med felaktig summering av belägg lett till ett stort internt bortfall.

Bortfallsanalyser som gjorts i samband med tidigare ombedömningsomgångar har inte kunnat påvisa några systematiska bortfall. Då insamlingens genomförande inte förändrats över åren kan det även antas att inte heller detta förändrats.

Orsaken till varför bortfallet av prov är särskilt stort för prov i den kommunala vuxenutbildningen kan delvis härledas till att det är frivilligt för de kommunala huvudmännen att skicka in kopior på bedömda nationella prov. Eftersom det totala antalet utvalda skolenheter och antalet insamlade prov är lågt, bör resultatet avseende proven som rör den kommunala vuxenutbildningen tolkas med försiktighet.

Frågor om bedömningen på skolan

I samband med insamlingen av prov fick skolorna även möjlighet att svara på ett antal frågor om hur bedömningen av provet gick till. Frågorna besvarades för varje enskild insänd elevlösning. Det är frivilligt för skolorna att besvara dessa frågor, vilket medför att antal svar som kunnat analyseras är lägre än antalet inkomna prov.

I tabell 2 redovisas resultaten från de skolor som besvarat frågorna om bedömningens genomförande på skolan och som gett sitt samtycke till att svaren får användas i analysyfte. Bortfallet utgörs av skolor som inkom med provkopior som gick att ombedöma, men som inte svarat på frågorna om bedömningens genomförande eller godkänt att svaren får användas i analysyfte.

Frågorna kring hur bedömningsprocessen på skolan gick till samlades inte in för proven inom den kommunala vuxenutbildningen. Därför är de inte inkluderade i analysen.

Tabell 2: Bortfall vid frågor om bedömningens genomförande på skolorna.

Frågor om bedömningen:	1. Vem bedömde provet?			2. Var provet avidentifierat vid bedömningen?			3. Diskuterades bedömningen med annan pedagogisk personal?			
	Totalt:	An- tal svar	Från antal sko- lor	Bort- fall	Antal svar	Från antal skolor	Bort- fall	An- tal svar	Från antal skolor	Bort- fall
<i>Ämne och delprov:</i>										
Engelska åk 6 delprov C	2 542	96	26 %	2 967	113	14 %	2 916	112	16 %	
Engelska åk 9 delprov C	4 169	111	31 %	5 104	130	16 %	5 076	129	16 %	
Engelska Gy delprov C	5 883	115	19 %	6 054	118	17 %	6 164	117	16 %	
Svenska åk 6 delprov C1+C2	2 321	96	31 %	2 860	112	15 %	2 837	110	16 %	
Svenska åk 9 delprov C	4 974	147	24 %	5 821	161	11 %	5 746	162	12 %	
Svenska Gy delprov C	6 292	113	17 %	6 445	118	15 %	6 521	117	14 %	
Biologi åk 9 delprov A2	850	21	26 %	939	25	18 %	955	25	16 %	
Fysik åk 9 delprov A2	679	25	35 %	822	29	21 %	846	29	19 %	
Kemi åk 9 delprov A2	928	20	22 %	1 034	23	13 %	1 042	23	12 %	
Geografi åk 9 delprov A+B	1 287	27	12 %	1 396	28	4 %	1 366	27	6 %	
Samhällskunskap åk 9 delprov A+B	699	16	37 %	913	20	17 %	901	20	18 %	
Historia åk 9 delprov A+B	464	16	39 %	631	20	16 %	628	20	17 %	
Religion åk 9 delprov A+B	858	27	35 %	1 016	30	23 %	1 040	30	21 %	

Bilaga 2: Ett harmoniserat mått för att mäta skillnader i bedömning

Ett nationellt prov består av flera *delprov*, typiskt sett 2–5 stycken. Enligt vad som beskrivs i avsnittet *Urval av prov och skolor* i Bilaga 1 så ombedöms inte alla delprov, utan endast de skriftliga prov där risken för avvikelser kan antas vara störst. Det gör att det samlade provbetyget inte kan användas som resultatmått, utan jämförelsen måste endast avse det aktuella delprovet.

För språkproven sätts delprovsbetyg på skalan A–F, men för övriga prov (NO, SO) är det endast det samlade provbetyget som anges på skalan A–F. För de olika delproven erhålls istället ett antal *poäng* på E-, C- respektive A-nivå (för NO- och SO-proven är *belägg* det ord som används). Hur många poäng/belägg som kan uppnås på respektive nivå varierar mellan olika ämnen och delprov.

För att kunna göra en översiktlig sammanställning av ombedömningens resultat för de vitt skilda proven, behöver de olika bedömningarna därför konverteras till en gemensam skala.

Denna har konstruerats på så sätt att bedömningarna av samtliga delprov kan uttryckas på skalan 0–20, där 0 representerar lägsta möjligt bedömning och 20 representerar högsta möjliga bedömning. Skillnaden mellan bedömning och ombedömning kan då på motsvarande sätt uttryckas på skalan -20 till 20. Ett positivt värde indikerar att skolan satt ett högre betyg (språkprov) eller gett fler belägg (SO/NO) än ombedömaren för en viss elevlösning. Ett negativt värde indikerar motsatsen.

Beräkning av det harmoniserade resultatmålet

För de delprov där bokstavsbedömning A–F sätts är konverteringen mellan den faktiska bedömningen och den harmoniserade skalan: A=20, B=16, C=12, D=8, E=4, F=0. En genomsnittsskillnad på 2 mellan bedömning och ombedömning för en viss skola betyder alltså i dessa fall att det i snitt skiljer "ett halvt betygssteg" mellan skolans och ombedömarens bedömning, där skolan i snitt sätter högre betyg.

För de delprov där delprovsbetyg inte sätts omvandlas i stället antalet belägg/poäng på E-, C- respektive A-nivå till skalan 0–20 enligt följande modell:
$$((E*10 + C*15 + A*20) / (\max(E)*10 + \max(C)*15 + \max(A)*20)) * 20$$

Där E, C respektive A står för antalet belägg som skolan samt ombedömaren angivit för delprovet, och max (E), max (C) respektive max (A) står för det maximala antal belägg som kan uppnås för det aktuella delprovet.

Resultatet av denna beräkning är att även bedömningar som uttrycks i belägg (NO- och SO-prov) kan standardiseras till ett enskilt indexvärde på skalan 0–20, där A-belägg väger dubbelt så tungt som E-belägg i sammanräkningen. Om det maximala antalet belägg som eleven kan erhålla för ett delprov är lågt, väger varje enskilt belägg i framräkningen av det harmoniserade måttet desto tyngre.

En snabb approximation som underlättar den intuitiva förståelsen är att 2 på den harmoniserade skalan motsvarar 10 procent av antalet möjliga belägg för det aktuella delprovet, 10 motsvarar 50 procent av antal möjliga belägg och så vidare. Denna approximation tar dock inte hänsyn till att A-, C- och E-belägg viktas olika i beräkningen.

Det är viktigt att komma ihåg att det harmoniserade resultatmåttet avseende de prov där det bygger på en sammanräkning av belägg endast är ett grovt sätt att på en gemensam skala ange om en skillnad i bedömning är "stor" eller "liten". Det finns dock ingen enkel korrespondens mellan detta mått och ett faktiskt betyg på skalan A-F, eftersom reglerna för hur belägg summeras till ett samlat provbetyg ser olika ut för olika prov, och dessutom involverar summering av belägg från flera olika delprov.

Bilaga 3: Metodologisk not avseende beräkning av variationen mellan skolor

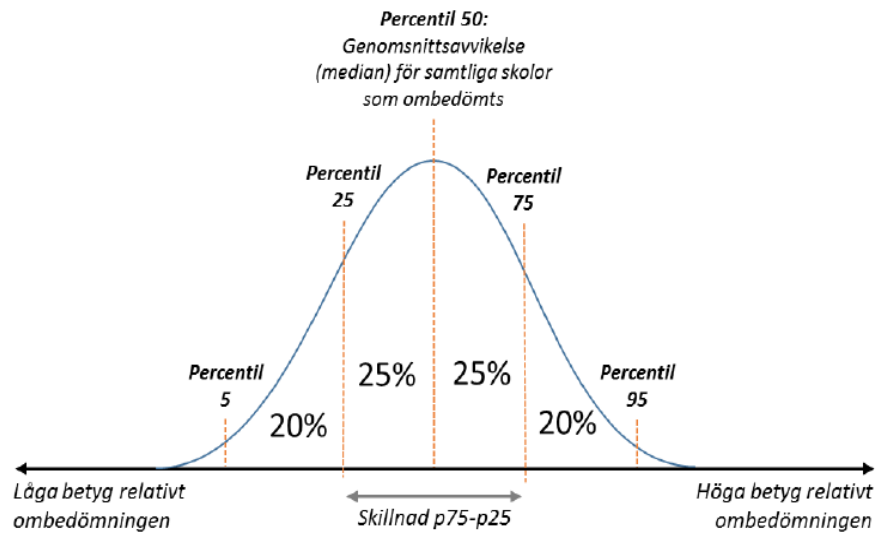
Spridningen mellan olika skolor i bedömningen av samma elevlösning kan inte observeras direkt utifrån ombedömningen, eftersom de flesta skolor endast bedömer sina egna elevers prov. Resultatet från ombedömningen kan dock användas som en gemensam referenspunkt för att göra en indirekt approximation av skillnaderna mellan skolor. Eftersom proven från varje skola slumpas ut till samma grupp av ombedömare, så borde skolorna om de bedömde likvärdiga elevlösningar på exakt samma sätt också ha ungefär samma medelavvikelse från ombedömningen. Spridningen med avseende på genomsnittsavvikelsen från ombedömningen ger därmed en indikation på hur stora skillnaderna mellan olika skolor skulle vara *om de hade bedömt samma prov*.

Måttet är endast approximativt eftersom sammansättningen av gruppen ombedömare inte ser exakt likadan ut för alla skolor, vilket kan påverka jämförelsen.

Skolinspektionen bedömer däremot att det ger en tillförlitlig bild av *storleksordningen* på skillnaderna mellan olika skolor. Mindre skillnader mellan olika delprov bör dock tolkas med försiktighet. Resultaten redovisas därför endast sammanslaget per ämnesgrupp.

För denna analys har även betygen på språkproven räknats om till skalan 0–20, enligt principen F=0, E=4, D=8, C=12, B=16, A=20. För språkproven är alltså konverteringen mellan den faktiska bedömningen och den harmoniserade skalan rättfram. En skillnad på 4 mellan två bedömningar motsvarar ett betygsstegs skillnad, 8 motsvarar två betygssteg och så vidare.

I analysen jämförs olika percentiler av skolor, med avseende på genomsnittlig avvikelse från ombedömningen. Den hundradel av skolorna som sätter de lägsta betygen i förhållande till ombedömningen utgör percentil 1. På motsvarande sätt utgörs percentil 25 av de skolor som relativt ombedömningen sätter lägre betyg än 75 procent av skolorna, och högre betyg än 25 procent. Figur 18 ger en schematisk illustration av jämförelsen mellan två percentilgrupper.



Figur 18: Schematisk illustration av jämförelsen mellan olika percentiler av skolor.

Bilaga 4: Användning av resultaten i tillsynen

I enlighet med regeringsuppdraget ska Skolinspektionen använda resultatet från omdömningen som ett underlag i tillsynen av skolor och huvudmän. Myndigheten ska "... utifrån indikationer på betydande avvikelser när det gäller dels rättning av nationella prov, dels betygssättning i ämnen och kurser, inom ramen för sin tillsynsverksamhet vidta lämpliga åtgärder."²⁵

Det övergripande syftet med analysen är att främja en rättvis och likvärdig betygssättning. Eftersom det inte finns några lagstadgade gränser för hur stora avvikelserna bör eller får vara, är det alltid processen som leder fram till resultatet som granskas i Skolinspektionens tillsyn, inte provbetygen i sig. En hög avvikelsegrad i förhållande till andra skolor kan däremot vara en indikation på att skolan behöver utveckla sina processer för bedömning och betygssättning, och att området bör ägnas extra uppmärksamhet i samband med tillsynen.

Betydande avvikelser på skolnivå: en bedömningsfråga

Vad som klassas som en "betydande avvikelse" är i grunden en bedömningsfråga. Det går inte att fastslå någon entydig definition av vad som är rätt svar för de frisvarsbaserade delproven. Skolinspektionen har därför valt att sätta ribban för vad som ska anses vara en betydande avvikelse relativt högt.

I ett första steg gäller det att för varje prov avgöra vad som är en avvikelse mellan skolans bedömning och Skolinspektionens omdömning. För språkproven innebär en *betydande avvikelse* en skillnad på *två betygsteg eller mer* mellan skolans bedömning och Skolinspektionens omdömning. För de delprov där bedömningen sker i form av belägg är gränsen satt till en skillnad på approximativt 30 procent eller mer i antal utdelade belägg mellan de två bedömningarna.²⁶

I ett andra steg gäller det att avgöra vad som är en avvikelse mellan skolans andel avvikelser från omdömningen och den genomsnittliga avvikelserna för samtliga skolor som ingår i omdömningen. De betydande avvikelserna i den ena eller andra riktningen ska vara mer vanligt förekommande vid den aktuella skolan än för genomsnittet av de omdömda skolorna. Eftersom omdömningens syfte är likvärdighet i bedömningen *mellan skolor* är det den enskilda

²⁵ Utbildningsdepartementet (2015). Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion. U2013/5175/GV. U2014/7521/SAM (delvis). U2014/7535/GV.

²⁶ Motsvarar +/- 6 i skillnad på den harmoniserade skalan 0-20. I sammanvägningen väger dock A-belägg tyngre än C-belägg, som i sin tur väger tyngre än E-belägg. Beräkningen av den harmoniserade betygsskalan beskrivs i Bilaga 2.

skolans avvikelse från andra skolor, inte från omdömningen i sig, som bör minimeras. En skolas resultat uppmärksammas ifall andelen betydande avvikelser, antingen i generösare eller striktare riktning, skiljer sig med mer än 10 procentenheter från genomsnittet för de omdömda skolorna för det aktuella delprovet. Avvikelsen fungerar som en signal till tillsynsverksamheten att särskilt granska skolans processer för bedömning och betygssättning vid nästkommande tillsynsbesök.

Resultaten från omdömningen är ett av flera underlag som Skolinspektionens inspektörer granskar inför tillsyn av skolor. Om Skolinspektionen bedömer att skolans arbetssätt inte uppfyller de i lag och förordning fastställda kraven, föreläggs skolan att åtgärda bristerna.

Referensförteckning

Betänkandet Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning (SOU 2016:25).

Fredriksson, Peter & Vlachos, Jonas (2011). Reformen och resultat: Kommer regeringens utbildningsreformer att ha någon betydelse? Finanspolitiska rådet. Hämtad den 21 oktober 2016, från:
<http://www.finanspolitiskaradet.se/download/18.34d8785812fd178ec94800022394/110608+Fredriksson+och+Vlachos.pdf>.

Skolinspektionen (2012). Lika för alla? Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år.

Skolinspektionen (2015). Ombedömning av nationella prov 2014: "Processerna spelar roll".

Skollagen (2010:800).

Skolverket (2009). Likvärdig betygssättning i gymnasieskolan? En analys av samband mellan nationella prov och kursbetyg. Hämtad den 21 oktober 2016, från: <http://www.skolverket.se/publikationer?id=2286>.

Skolverket (2011). Kunskapsbedömning i skolan – praxis, begrepp, problem och möjligheter.

Skolverket (2011). Skolverkets allmänna råd: Planering och genomförande av undervisningen – för grundskolan, särskolan, specialskolan och sameskolan.

Skolverket (2012). Skolverkets allmänna råd: Bedömning och betygssättning i gymnasieskolan.

Utbildningsdepartementet (2015). Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion. U2013/5175/GV. U2014/7521/SAM (delvis).

U2014/7535/GV. Hämtad den 25 oktober 2016, från:
<http://www.esv.se/statsliggaren/regleringsbrev/?RBID=16509>.