



# Ombedömning av nationella prov 2018 – Samma prov, olika bedömningar

Redovisning av regeringsuppdrag Dnr U2017/05038/GV

# Innehållsförteckning

<b>Förord</b> .....	<b>4</b>
<b>Sammanfattning</b> .....	<b>5</b>
Allmänt.....	5
Resultat .....	5
<b>Uppdrag och genomförande</b> .....	<b>7</b>
Uppdraget .....	7
Bakgrund .....	7
Genomförandet.....	8
Ombedömda delprov .....	8
Urvalet av skolor .....	8
Denna rapport.....	9
<b>Resultat</b> .....	<b>10</b>
Hur mycket skiljer sig bedömningarna av samma elevlösning? .....	10
Bedömningarna skiljer sig åt för mer än varannan elevlösning.....	10
Vanligt att skolans lärare gör mer generösa bedömningar i svenska .....	10
Merparten av elevlösningarna i engelska får en annan bedömning .....	11
Hälften av delproven i NO fick samma bedömning .....	12
Överlag små skillnader jämfört med tidigare år .....	12
Små skillnader mellan fristående och kommunala skolor .....	15
Små skillnader i avvikelser mellan delprov skrivna av flickor jämfört med pojkar... ..	16
Liknande resultat för hand- som för datorskrivna prov.....	17
Stor variation mellan skolor .....	19
Hur påverkas skillnaderna av bedömningsprocessen?.....	22
Stora variationer i hur delproven bedöms .....	22
Stora variationer i hur bedömningsmetoderna kombineras .....	25
Bedömningsprocessen har betydelse .....	26
Kombinationer av bedömningsmetoder minskar tendensen att skolan gör en mer generös bedömning .....	29
Vid extern bedömning ges flickor och pojkar i högre grad samma betyg av de olika bedömarna.....	30
Bedömningsprocessen har betydelse för både hand- och datorskrivna delprov ....	32
Sambedömning minskar variationen mellan skolor .....	35
<b>Slutsatser</b> .....	<b>40</b>
Skillnader mellan bedömningarna av de nationella delproven .....	40

Bedömningsprocessen har betydelse .....	40
Svårt att göra samstämmiga bedömningar av delprov med fritextsvar .....	41
Avslutningsvis .....	41
<b>Bilagor.....</b>	<b>43</b>
Bilaga 1 Användning av resultaten i tillsyn och kvalitetsgranskning .....	43
Hur avgörs vad som är betydande avvikelser? .....	43
Att använda resultaten i tillsyn och kvalitetsgranskning .....	44
Skolenhetsrapporter .....	44
Bilaga 2 Genomförandet av ombedömningen .....	45
Urval av prov och skolor.....	45
Hela provet ombedöms inte, endast längre skrivuppgifter .....	45
Urvalets sammansättning .....	46
Fördelning av delprov mellan Skolinspektionens ombedömare .....	46
Hand- respektive datorskrivna prov.....	47
Omfattning och bortfall.....	47
Vad beror bortfallet på?.....	49
Frågor om bedömningen på skolan .....	49
Bilaga 3 Ett harmoniserat mått för att mäta skillnader i bedömning av nationella delprov .....	51
Beräkning av det harmoniserade resultatmättet .....	51
Bilaga 4 Övriga figurer .....	53
<b>Referenslista .....</b>	<b>58</b>

# Förord

Skolinspektionen har sedan 2009 haft i uppdrag av regeringen att varje år genomföra viss omdömning av nationella prov. Syftet med omdömningen är att stödja en likvärdig bedömning och betygssättning av proven över landet.

Skolinspektionen redovisar här till regeringen resultatet av tionde årets omdömning. Denna rapport avser nationella prov i grund- och gymnasieskolan som genomfördes under 2018 och nationella prov inom den kommunala vuxenutbildningen på gymnasial nivå som genomfördes under 2017.

Vid Skolinspektionen har följande personer arbetat med att genomföra uppdraget. Projektledare: Fanny Hemmingsson. Utredare/statistiker: Per Gillström, Peter Gustafsson och Tina Sajjadi. Enhetschef: Emma Uhman.

Helén Ängmo

Generaldirektör

Anna Sellin

Avdelningschef

# Sammanfattning

## Allmänt

Skolinspektionen har sedan 2009 fått i uppdrag att ombedöma vissa nationella prov. Syftet med uppdraget är att, genom att undersöka bedömningarnas likvärdighet samt analysera vilka faktorer som främjar likvärdig bedömning, stödja en likvärdig bedömning och betygssättning av proven över landet. Årets resultat visar, liksom tidigare år, att det är vanligt att ombedömaren gör en annan bedömning än ursprungsbedömaren. Ombedömarnas bedömningar ska dock inte betraktas som ett facit, det vill säga deras bedömningar är inte mer korrekta än de som skolans lärare har gjort. Ombedömningarna kan ses som ett slags second opinion och skillnader ger underlag för att diskutera likvärdighet.

I uppdraget ingår att ombedöma de prov och delprov där risken för avvikelser är störst. Alla delprov i de nationella proven ombedöms därför inte, utan ett urval av de delprov som till stor del består av fritextsvar. I 2018 års ombedömning ingår delprov från de nationella proven i svenska och engelska för årskurs 6, årskurs 9, gymnasiet och den kommunala vuxenutbildningen på gymnasial nivå, samt i NO (biologi, fysik och kemi) för årskurs 9.

Liksom i tidigare rapporter mäts skillnaden mellan skolans och ombedömarnas bedömning samt hur bedömningsprocessen påverkar likvärdigheten i bedömningarna. I årets rapport görs dessutom en jämförelse av bedömaröverensstämmelsen mellan delprov skrivna av flickor och pojkar samt mellan delprov som är skrivna för hand och på dator.

## Resultat

### Bedömningarna skiljer sig för mer än varannan elevlösning

Resultatet visar att det är vanligt att skolans lärare och ombedömarna gör olika bedömningar av samma elevlösning. I 57 procent av fallen har ombedömaren gjort en annan bedömning än skolans lärare. Likt tidigare år är det vanligare att skolorna sätter mer generösa betyg jämfört med ombedömarna än vice versa. I 38 procent av fallen har skolan gjort en mer generös bedömning och i 19 procent av fallen har skolan gjort en mer strikt bedömning än ombedömaren. Denna tendens återfinns i alla ombedömda ämnen, årskurser och kurser, både för delprov skrivna av flickor respektive pojkar samt på fristående respektive kommunala skolor.

### Skillnaderna är störst i svenska och minst i NO-ämnena

Andelen avvikelser mellan ursprungsbedömning och ombedömning varierar mellan olika ämnen, vilket tyder på att utrymmet för individuella tolkningar skiljer sig mellan olika delprov. Skillnaderna är störst i svenska, framför allt inom vuxenutbildningen, och minst i NO-ämnena (omfattar endast årskurs 9). Överlag har resultaten från ombedömningen varit likartade de senaste fem omgångarna. När det gäller delproven i NO-ämnena har samstämmigheten i bedömningarna dock varit högre de två senaste åren. En trolig förklaring till den mer likvärdiga bedömningen är att delproven i NO-ämnena 2017 fick en ny utformning och nya bedömningsanvisningar.

### Liknande resultat för delprov som skrivits för hand och på dator

Drygt hälften av de ombedömda elevlösningarna är skrivna för hand och knappt hälften på dator. Andelen avvikelser mellan skolans och ombedömarnas bedömningar är ungefär lika stora för de hand- och datorskrivna proven. Bland delproven i svenska finns dock en

liten skillnad. Av de handskrivna proven har 45 procent fått en generösare bedömning på skolan, jämfört med 40 procent av de datorskrivna.

### **Bedömningsprocessen påverkar bedömningen**

Vid bedömningen av de nationella proven tillämpades olika processer. För en tredjedel av delproven tillämpades extern bedömning, det vill säga att delprovet bedömdes av någon annan än elevens ordinarie lärare. Sju av tio delprov sambedömdes medan två av tio delprov var avidentifierade vid bedömningen. Olika kombinationer av de tre bedömningsmetoderna förekom också.

Av resultaten framgår att extern bedömning, sambedömning och i viss mån avidentifiering har en återhållande effekt på betygssättningen. Den generella tendensen att skolan gör generösare bedömningar är mindre för delprov som har bedömts av någon annan än elevens ordinarie lärare, sambedömts och avidentifierats. Effekten är störst när dessa bedömningsmetoder har kombinerats, dessutom ökar överensstämmelsen mellan ursprungsbedömarna och ombedömarna något.

### **Små skillnader i ombedömningens resultat mellan delprov skrivna av flickor och pojkar**

Vid en jämförelse av delprov skrivna av flickor och pojkar syns små skillnader. Resultaten visar en svag tendens till att flickor generellt har bedömts generösare än pojkar och att pojkar i något högre utsträckning får samma betyg i de två bedömningarna. För delprov som har bedömts av någon annan än elevens lärare (extern bedömning) får dock flickor och pojkar i lika stor utsträckning samma betyg av de olika bedömarna. I övrigt får tillämpning av extern bedömning, sambedömning eller avidentifiering liknande effekter på delprov som skrivits av flickor respektive pojkar. Resultatet att andelen generösare bedömningar är mindre för delprov där dessa metoder har tillämpats gäller i lika stor utsträckning för båda könen.

# Uppdrag och genomförande

## Uppdraget

Sedan 2009 har Skolinspektionen årligen fått i uppdrag av regeringen att omdedöma nationella prov i sitt regleringsbrev. Syftet med uppdraget är att stödja en likvärdig bedömning och betygssättning över hela landet.<sup>1</sup> I beskrivningen av uppdraget framgår bland annat att de prov och delprov där risken för avvikelser är störst ska ingå i så stor utsträckning som möjligt. Därför genomförs omdedömningen av nationella prov för ett urval av ämnen, årskurser, kurser och delprov.

I likhet med tidigare år framgår i regleringsbrevet avseende 2018 att myndigheten ska använda resultatet från omdedömning av nationella prov i sin tillsyn och kvalitetsgranskning.<sup>2</sup> En utförlig beskrivning av hur resultaten från omdedömningen används i myndighetens tillsyn, regelbundna kvalitetsgranskning och tematiska kvalitetsgranskning finns i bilaga 1.

## Bakgrund

De nationella proven syftar till att stödja en rättvis och likvärdig betygssättning. Sedan 29 juni 2018 ska resultat från de nationella proven särskilt beaktas vid betygssättning i ett ämne eller kurs som det ges nationellt prov i, givet att det inte finns särskilda skäl som gör att resultaten inte bör beaktas särskilt.<sup>3</sup> Det finns särskilda skäl till varför resultat från de nationella proven ska ges särskild tyngd vid betygssättningen. Anledningen är att de nationella proven är välkonstruerade, noga utprovade och har utformats just för att stödja betygssättningen. Därutöver skiljer sig de nationella proven från övriga bedömningsunderlag på så sätt att alla elever runt om i landet gör samma prov och bedömarna utgår från samma bedömningsanvisningar.<sup>4</sup> Betygssättningen ska dock ske utifrån en allsidig utvärdering av all information om en elevs kunskaper.<sup>5</sup>

Skolinspektionens tidigare erfarenheter från omdedömningen av nationella prov är att vissa delprov inte bedöms likvärdigt. Det är vanligt att samma elevlösning får olika bedömningar av skolan och omdedömaren.<sup>6</sup> Skillnaderna mellan bedömare har varit som störst för delprov där eleven ska ange fritextsvar på frågor som inte har ett givet rätt eller fel svar. För sådana delprov är det vanligt att det skiljer ett betygssteg eller mer mellan två olika lärares bedömningar av samma elevlösning.<sup>7</sup>

Med tanke på betygens många gånger avgörande roll för urval till högre studier och arbete kan en olikvärdig bedömning av nationella prov, till exempel att prestationer på ett delprov som anses motivera betyget A på en skola ger C på en annan, få konsekvenser för

<sup>1</sup> Utbildningsdepartementet, *Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion* (2014).

<sup>2</sup> Utbildningsdepartementet, *Regleringsbrev för budgetåret 2018 avseende Statens skolinspektion* (2017).

<sup>3</sup> 10 Kap. 20 a § Skollagen (2010:800), 15 Kap. 25 a § Skollagen

<sup>4</sup> För alla nationella prov i de årskurser/kurser där betyg ges, följer ett delprovsbetyg samt ett sammanvägt provbetyg utifrån bedömningsanvisningarna.

<sup>5</sup> Förordning om läroplan för grundskolan, förskoleklassen och fritidshemmet (SKOLFS 2010:37), 2.7 Bedömning och betyg; Förordning om läroplan för gymnasieskolan 2011 (SKOLFS 2011:144), 2.5 Bedömning och betyg.

<sup>6</sup> Statens skolinspektion, *Omdedömning av nationella prov 2017 – Fortsatt stora skillnader* (2018).

<sup>7</sup> Statens skolinspektion, *Olikheterna är för stora Omrättning av nationella prov i grundskolan och gymnasieskolan, 2013* (2013).

enskilda elever. Det är därför viktigt och angeläget att de nationella proven bedöms likvärdigt.

## Genomförandet

För att genomföra uppdraget samlar Skolinspektionen in kopior på elevlösningar från nationella delprov från ett urval av skolor. De insamlade provkopiorna avidentifieras och maskeras från tidigare bedömares kommentarer, anteckningar och dylikt. Därefter bedöms elevlösningarna på nytt av legitimerade och erfarna ämneslärare som anlitas på uppdrag av Skolinspektionen. Dessa bedömningar görs enskilt. De ursprungliga bedömningarna av elevsvaren jämförs sedan med ombedömarnas bedömningar. Jämförelserna sammanställs både på skolenhetsnivå och nationell nivå.

För att uppfylla uppdragets syfte ska myndigheten även samla in kompletterande information om hur och av vem proven bedömts för att få ytterligare underlag för analys av förutsättningar och faktorer som främjar en likvärdig bedömning. För att Skolinspektionen ska få tillgång till denna information ombes skolorna därför att för varje delprov besvara några frågor kopplade till bedömningen av delprovet.

## Ombedömda delprov

I varje ämne ombedöms endast *ett delprov*.<sup>8</sup> De delprov som ingår är de där risken för skillnader i bedömning mellan olika lärare bedömts som störst. I praktiken innebär det att de delprov som väljs till stor del består av fritextsvar där eleven fritt ska formulera sig utifrån en öppen frågeställning som inte har ett givet rätt eller fel svar. En mer utförlig beskrivning av urvalet av delprov för ombedömningen återfinns i bilaga 2.

Mot bakgrund av ovan ingår i 2018 års ombedömning delprov C i svenska/svenska som andraspråk och engelska för grundskolans årskurs 6 och 9, samt delprov A2 i biologi, fysik och kemi (NO) för årskurs 9. Vidare innefattar ombedömningen delprov C i svenska 1/svenska som andraspråk 1 och engelska 5 för gymnasieskolan och motsvarande kurser inom den kommunala vuxenutbildningen på gymnasial nivå. Totalt har 27 772 delprov från 393 skolor ombedömts.

## Urvalet av skolor

Urvalet av skolor för ombedömning följer från och med 2014 Skolinspektionens inspektionsplanering.<sup>9</sup> Urvalet är därför inte utformat för att ge en representativ bild av likvärdigheten i bedömning och betygssättning på nationell nivå. Resultatet som presenteras i denna rapport kan därför inte med säkerhet generaliseras till riket i stort. På samma sätt går det inte heller att veta i vilken utsträckning resultaten är jämförbara mellan de olika åren som ombedömningen gjorts. Skolinspektionen bedömer dock att urvalets sammansättning och omfattning gör det möjligt att använda resultaten som en indikation på den allmänna situationen i landet, samt på förändringar över tid.

Det totala bortfallet på provnivå är 6 procent. Bortfallet varierar dock mellan de olika delproven. Högst är bortfallet för svenska inom vuxenutbildningen. Bortfallet beskrivs mer ingående i bilaga 2.

<sup>8</sup> I svenska årskurs 6 är delprovet uppdelat i två delar (C1 och C2).

<sup>9</sup> Urvalet regleras i den gällande uppdragsbeskrivning som återfinns i 2015 års regleringsbrev.



## Denna rapport

Med denna rapport redovisar Skolinspektionen resultaten från den tionde omgången av ombedömningen av nationella prov. I rapporten presenteras både deskriptiva resultat och ett antal statistiska analyser med bäring på frågan om likvärdig bedömning av delproven. Liksom i tidigare rapporter redovisas skillnaderna mellan ursprungsbedömarens och ombedömarens bedömning samt hur sambedömning, extern bedömning<sup>10</sup> och avidentifiering påverkar samstämmigheten i bedömningarna. Nytt för denna omgång är att vi i rapporten presenterar resultat och analyser utifrån juridiskt kön och om delprovet varit skrivet för hand eller på dator. Den detaljerade redogörelsen för projektets genomförande, urvalet av skolor och delprov samt beräkningstekniska detaljer har sammanställts i bilagor.

---

<sup>10</sup> Med extern bedömning avses bedömningar som inte är gjorda av elevens ordinarie lärare i ämnet.

# Resultat

I avsnittet presenteras resultaten av 2018 års omdömning. Först presenteras jämförelser mellan skolornas och omdömarens bedömningar av delproven. Därefter analyseras hur bedömningsprocessen påverkar skillnaderna mellan ursprungsbedömning och omdömning.

## Hur mycket skiljer sig bedömningarna av samma elevlösning?

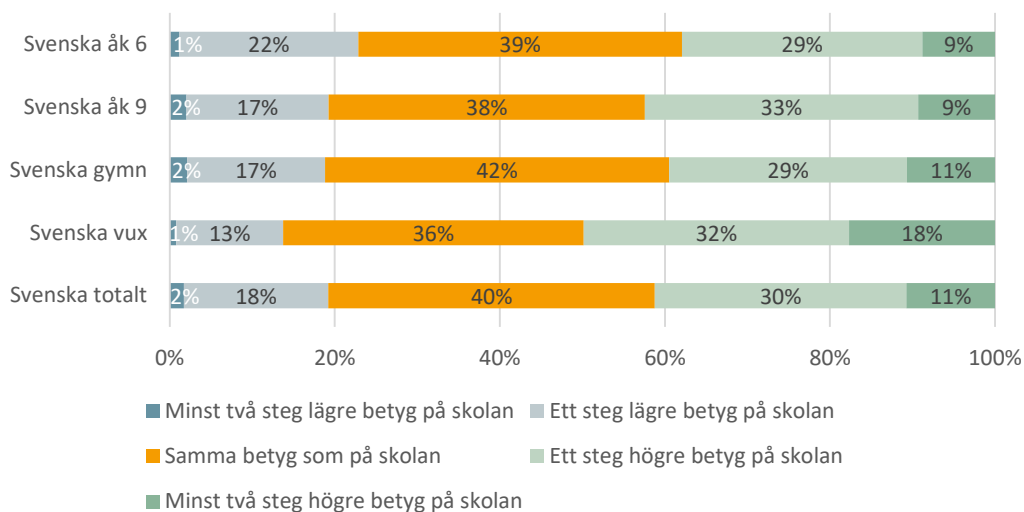
I följande avsnitt presenteras jämförelser mellan skolornas och omdömarens bedömningar av delproven per ämne och årskurs. Därtill jämförs andelen avvikelser mellan fristående och kommunala skolor, mellan delprov skrivna av flickor respektive pojkar samt mellan delprov som är skrivna för hand respektive på dator. Dessutom presenteras spridningen i andel avvikelser mellan olika skolor samt jämförelser med resultatet i tidigare års analyser.

## Bedömningarna skiljer sig åt för mer än varannan elevlösning

Drygt hälften av delproven har fått olika bedömningar av skolornas lärare och omdömare. För 57 procent av elevlösningarna skiljer sig resultatet mellan de olika bedömningarna. Där olika bedömningar gjorts skiljer det oftast ett betygssteg. Det är vanligast att skolan har gjort en mer generös bedömning än omdömare, det gäller för 38 procent av delproven. I 19 procent av fallen har skolan satt ett lägre betyg än omdömare. Andelen avvikelser mellan de två bedömningarna skiljer sig något mellan olika ämnen och årskurser.

## Vanligt att skolans lärare gör mer generösa bedömningar i svenska

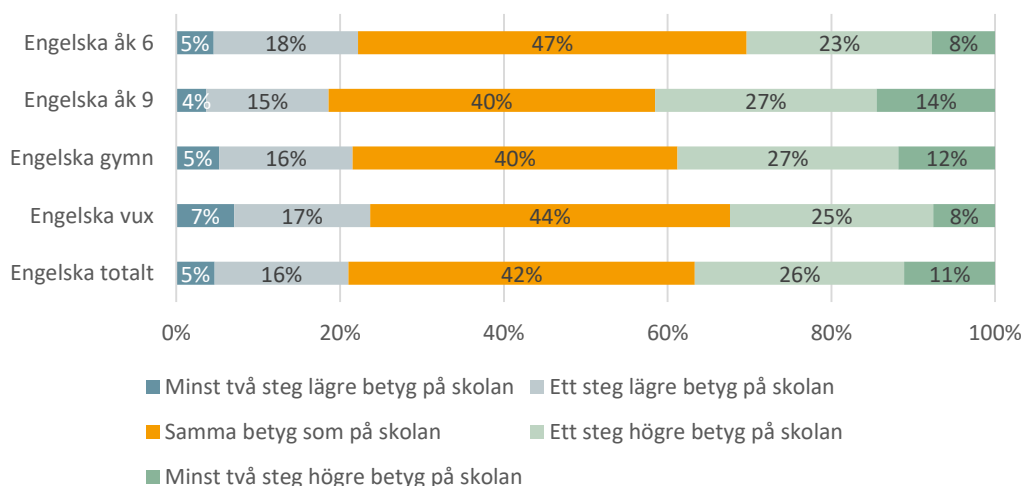
Som framgår av figur 1 har omdömare gjort en annan bedömning än ursprungsbedömare för merparten av de omdömda delproven i svenska. I cirka 40 procent av fallen har skolan gjort en mer generös bedömning, i 40 procent har samma bedömning gjorts och i 20 procent har skolan gjort en mer strikt bedömning. Genomgående för alla årskurser och kurser är att skolan gör en mer generös bedömning. Tendensen är tydligast inom den kommunala vuxenutbildningen där hälften av delproven fått ett högre delprovsbetyg av ursprungsbedömare. Jämfört med övriga årskurser och kurser är det också vanligare inom vuxenutbildningen att ursprungsbedömare ger ett delprovsbetyg som är minst två betygssteg högre än Skolinspektionens omdömare.



Figur 1: Resultat i omdömningen för utvalda delprov i svenska, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

## Merparten av elevlösningarna i engelska får en annan bedömning

Andelen delprov i engelska där skolans lärare och omdömare har gjort olika bedömningar är 58 procent, se nedersta stapeln i figur 2 nedan. Liksom i svenska är det vanligt att skolans lärare har gjort en mer generös bedömning än omdömare. I 37 procent av fallen har skolan satt ett högre betyg, i 42 procent av fallen har samma bedömning gjorts och i 21 procent av fallen har skolan satt ett lägre betyg än omdömare.



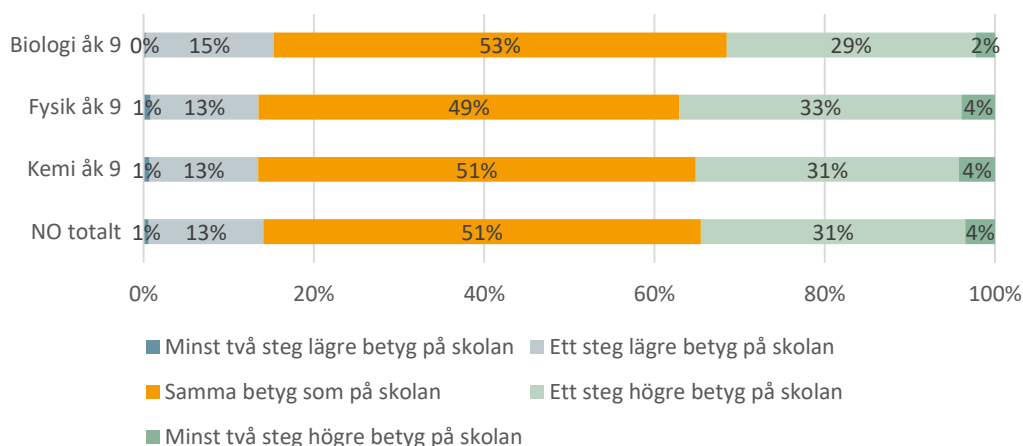
Figur 2: Resultat i omdömningen för utvalda delprov i engelska, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

Samstämmigheten mellan bedömningarna är större i årskurs 6 än i övriga årskurser och kurser. Där har 47 procent av elevlösningarna fått samma bedömning. En orsak till att samstämmigheten är högre kan vara att textmängden i elevlösningarna generellt är mindre i årskurs 6. Andelen delprov där skolan har gjort mer generösa bedömningar än omdömare är högst i årskurs 9, 41 procent, och på gymnasiet, 39 procent.

När det gäller vuxenutbildningen skiljer sig resultaten i engelska från dem i svenska. För delproven i svenska observerades en lägre grad av samstämmighet i vuxenutbildningen än i övriga årskurser och kurser. Dessutom var andelen delprov som bedömts generösare av skolans lärare högre. För delproven i engelska är samstämmigheten istället något högre inom vuxenutbildningen jämfört med genomsnittet för övriga delprov i engelska. Andelen delprov som fått generösare bedömningar av skolans lärare är samtidigt något lägre.

## Hälften av delproven i NO fick samma bedömning

I NO-ämnena (endast årskurs 9) är samstämmigheten kring bedömningarna större än i svenska och engelska, vilket framgår av figur 3 nedan. Drygt hälften av elevlösningarna fick samma bedömning på skolan som av ombedömarna. I de fall där olika bedömningar har gjorts är det till övervägande del skolan som har gjort en generösare bedömning än ombedömarna. I 35 procent av fallen har skolan satt ett högre betyg och i 14 procent av fallen har skolan gjort en striktare bedömning än ombedömarna. Där olika bedömningar gjorts skiljer det oftast ett betygssteg (belägg har räknats om till betygsskalan A-F). Jämfört med svenska och engelska är det mindre vanligt att det skiljer två betygssteg eller mer mellan bedömningarna. Mönstret är ungefär desamma för de olika NO-ämnena.



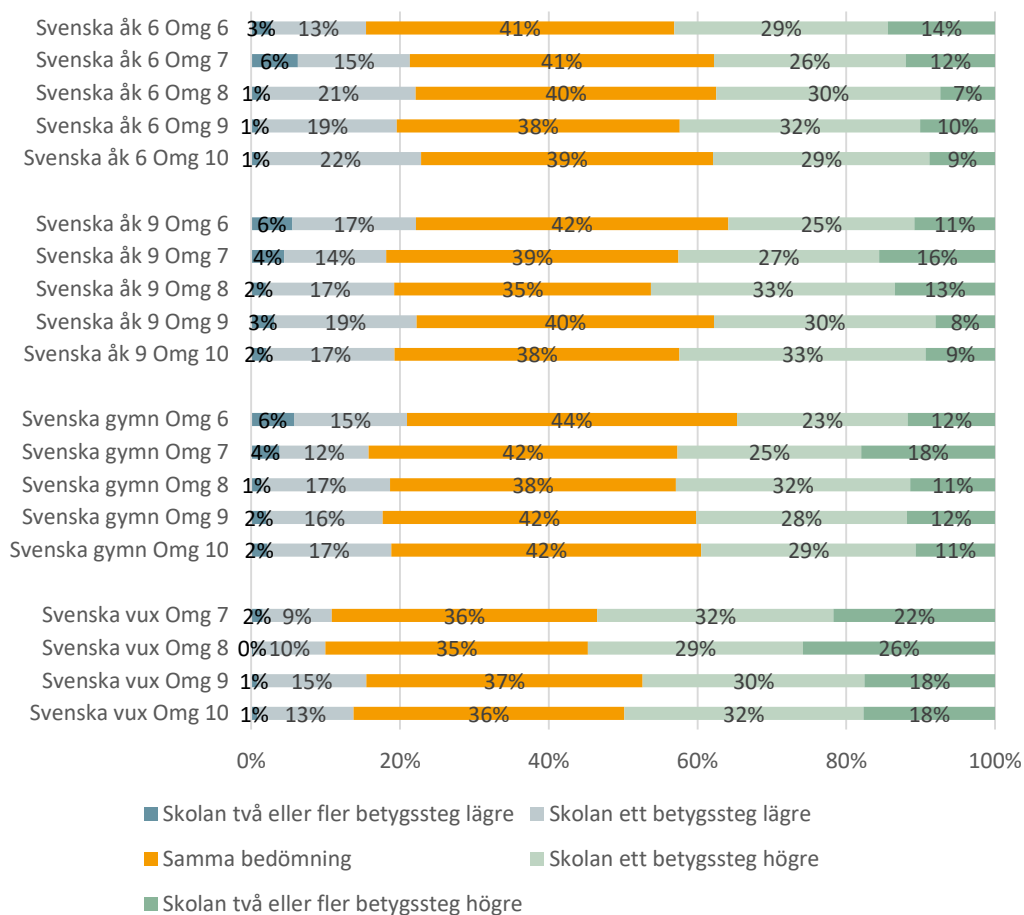
Figur 3: Resultat i ombedömningen för utvalda delprov i kemi, fysik och biologi i årskurs 9. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

## Överlag små skillnader jämfört med tidigare år

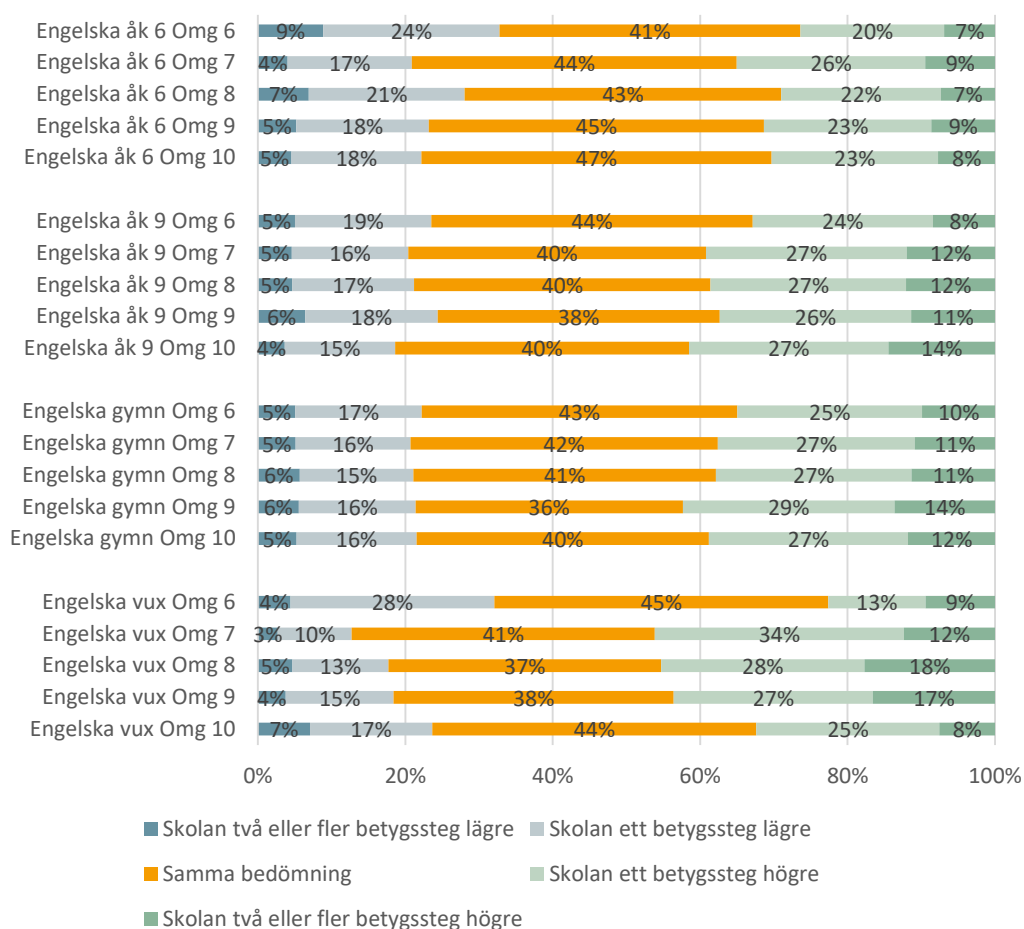
I figurerna 4–6 ställs resultaten från årets omgång av ombedömningen i relation till resultaten från de fyra senaste omgångarna.<sup>11</sup> Det bör noteras att delprov till ombedömningen samlas in från de kommuner och enskilda huvudmän som kan komma att bli föremål för Skolinspektionens tillsyn eller granskning när resultaten från respektive ombedömningsomgång är klara. Detta innebär att resultaten blir mer användbara i inspektionen, men betyder samtidigt att nationella jämförelser mellan åren blir mer osäkra än om ett representativt och slumpmässigt urval av skolor i hela landet hade dragits. Dessutom har fördelningen mellan fristående och kommunala skolor varierat mellan åren.

<sup>11</sup> I svenska, omgång 6, ombedömdes ett annat delprov i den kommunala vuxenutbildningen på gymnasial nivå. Därför redovisas inte resultatet från den kommunala vuxenutbildningen för den omgången.

För ämnena svenska och engelska är resultaten tämligen oförändrade under de senaste fem omgångarna (figur 4 och 5). I de fall där skolans lärare och Skolinspektionens ombedömare gjort olika bedömningar, finns tendensen från föregående omgångar kvar: att skolans lärare oftare gör en mer generös bedömning än ombedömarna. Detta visas av att de gröna delarna av staplarna är större än den blå.

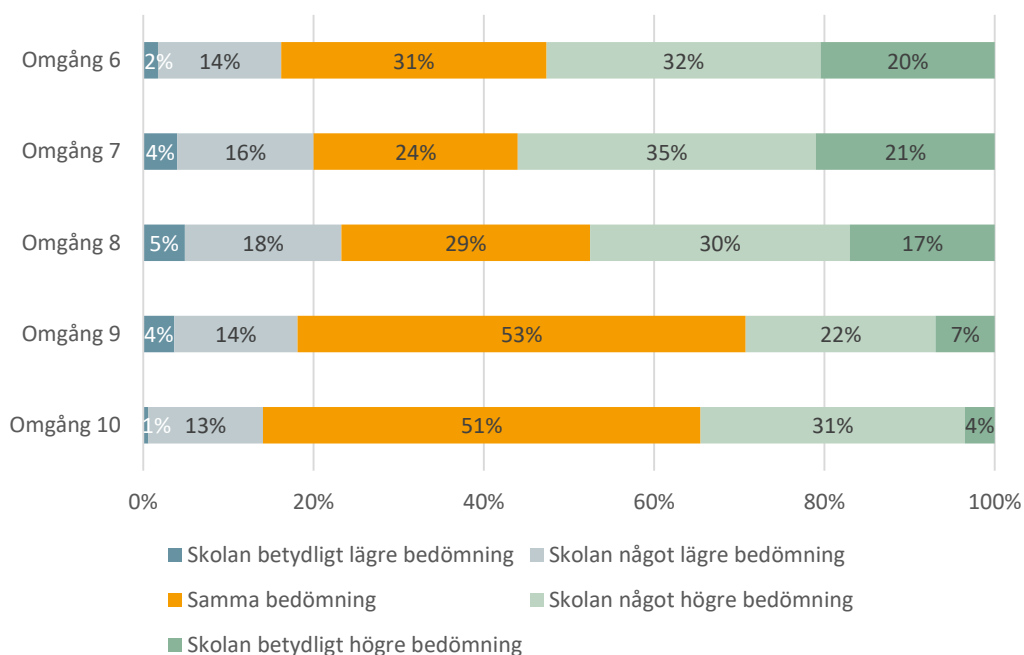


Figur 4: Resultat i ombedömningen för delprov i svenska de senaste fem omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.



Figur 5: Resultat i omdedömningen för delprov i engelska de senaste fem omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdedömare.

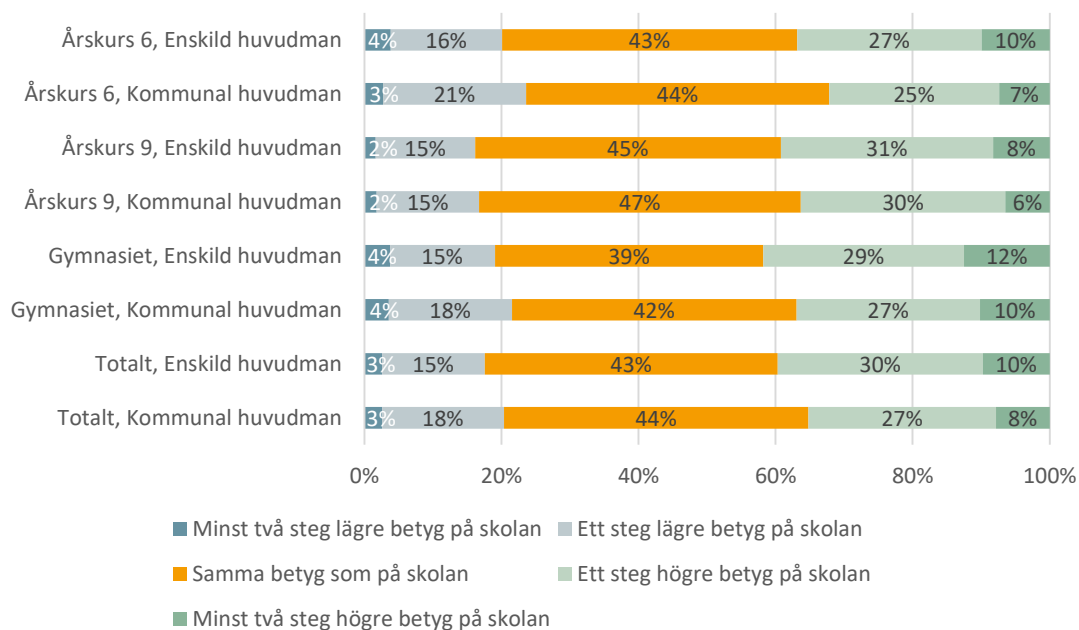
Däremot framträder en tydlig förändring rörande resultatet för NO-ämnena (figur 6). Mellan omgång åtta och nio blev det betydligt vanligare att skolans lärare hade gjort samma bedömning som Skolinspektionens omdedömare. Andelen ökade från 29 procent till 53 procent och uppgår till 51 procent omgång tio. En sannolik förklaring till den mer likvärdiga bedömningen i NO-ämnena är att delproven 2017 fick en ny utformning och nya bedömningsanvisningar. Resultatet för omgång tio visar på små förändringar jämfört med omgång nio. En skillnad är att de fall där skolans lärare gör en högre eller lägre bedömning med två eller flera betygssteg har minskat i omfattning. Nedgången kan troligen förklaras av att i omgång tio fick de elevlösningar i NO-ämnena där det skiljde två betygssteg eller mer (belägg har räknats om till betygsskalan A-F) mellan ursprungsbedömning och omdedömning en andra omdedömning. Den av de två omdedömningarna som låg närmast ursprungsbedömningen användes sedan som resultat från omdedömningen.



Figur 6: Resultat i omedömningen för delprov i NO-ämnena i årskurs 9 de senaste fem omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omedömarna.

## Små skillnader mellan fristående och kommunala skolor

Som framgår av de nedersta staplarna i figur 7 är skillnaderna i resultatet mellan kommunala och fristående skolor relativt små. Det är ungefär lika stor andel av elevlösningarna som fått samma bedömning av skolan och omedömaren. Där olika bedömningar gjorts är det dock något vanligare att de fristående skolorna har gjort mer generösa bedömningar jämfört med de kommunala, 40 jämfört med 35 procent. Samtidigt är det något vanligare att de kommunala skolorna har gjort en mer strikt bedömning jämfört med de fristående, 21 jämfört med 18 procent. Detta mönster återfinns inom samtliga årskurser och kurser.



Figur 7: Resultat i omdömningen, efter årskurs/kurs samt typ av huvudman. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare (belägg i NO-ämnen har räknats om till skalan A–F).

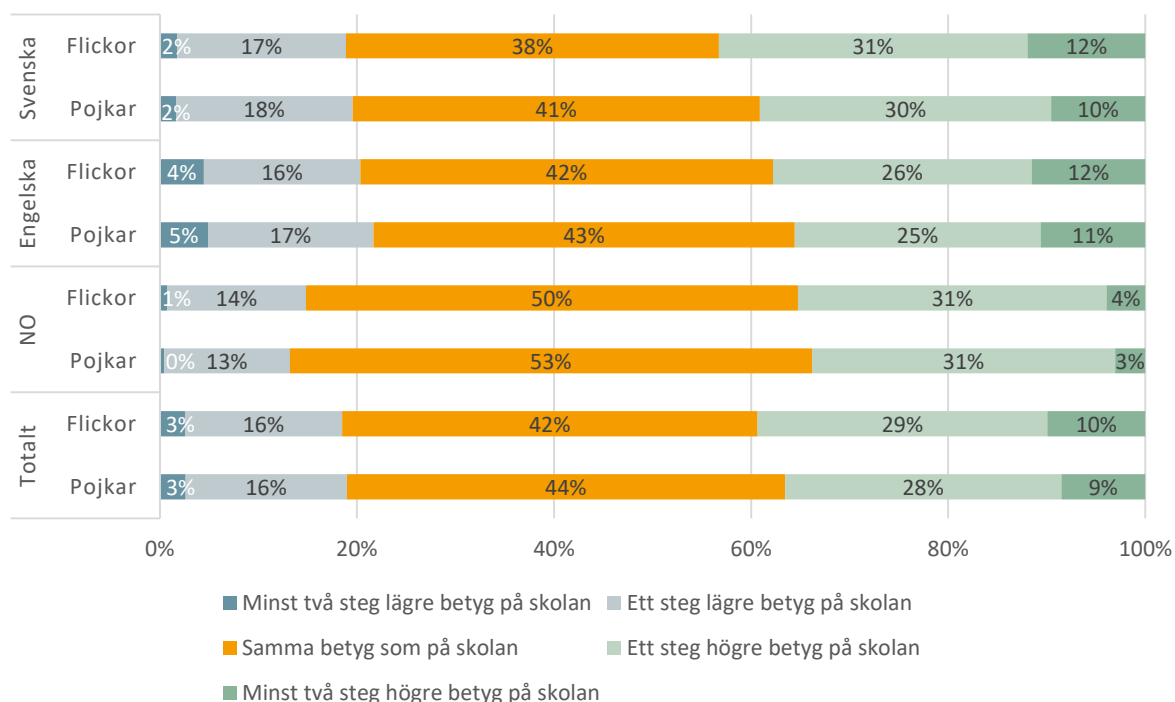
## Små skillnader i avvikelser mellan delprov skrivna av flickor jämfört med pojkar

Nytt för i år är att resultat tagits fram för i vilken mån ursprungsbedömarnas bedömning avviker från omdömarens utifrån elevens kön. Som framgår av figur 8 är skillnaderna i avvikelse mellan delprov som skrivits av flickor respektive pojkar<sup>12</sup> små för samtliga ämnen.

Resultaten visar en svag tendens till att flickor generellt har bedömts generösare än pojkar. Andelen delprov som fått samma betyg i de två bedömningarna är samtidigt aningen högre för delprov skrivna av pojkar. Tendensen är genomgående för alla ämnen och för samtliga årskurser och kurser (se figur 31 i bilaga 4.) Skillnaderna är dock väldigt små och det kan inte uteslutas att de beror på slumpmässiga variationer.

<sup>12</sup> Vi använder begreppen flickor och pojkar för de kvinnor och män som studerar på gymnasial nivå och inom den kommunala vuxenutbildningen.





Figur 8: Resultat i omdömningen, efter ämne samt elevens kön. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare (belägg i NO-ämnena har räknats om till skalan A–F).

## Liknande resultat för hand- som för datorskrivna prov

Av de elevlösningar som samlats in för omdömning är 55 procent skrivna för hand och 45 procent på dator<sup>13</sup>. Provpuppgifterna har dock distribuerats i pappersformat till samtliga elever. I tabell 1 framkommer att andelen datorskrivna elevlösningar varierar stort mellan olika ämnen och årskurser.<sup>14</sup> Det är vanligt att gymnasieelever skrivit delprovet på dator. Det är mindre vanligt inom vuxenutbildningen och inom NO-ämnena.

<sup>13</sup> Från och med höstterminen 2018 ska uppsatser i svenska, svenska som andraspråk och engelska skrivas på dator eller annan digital enhet (Regeringsbeslut U2017/03739/GV), i årskurs 9 i grundskolan, årskurs 10 i specialskolan, gymnasieskolan och vuxenutbildningen (Skolverket.se). De elevlösningar som ingår i denna rapport skrevs vårterminen 2018 och omfattades därmed inte av beslutet.

<sup>14</sup> Resultat saknas för svenska årskurs 6 pga att delprovet är uppdelat i två delar (C1 och C2), där den ena delen kan vara skriven för hand medan den andra skrivits på dator.

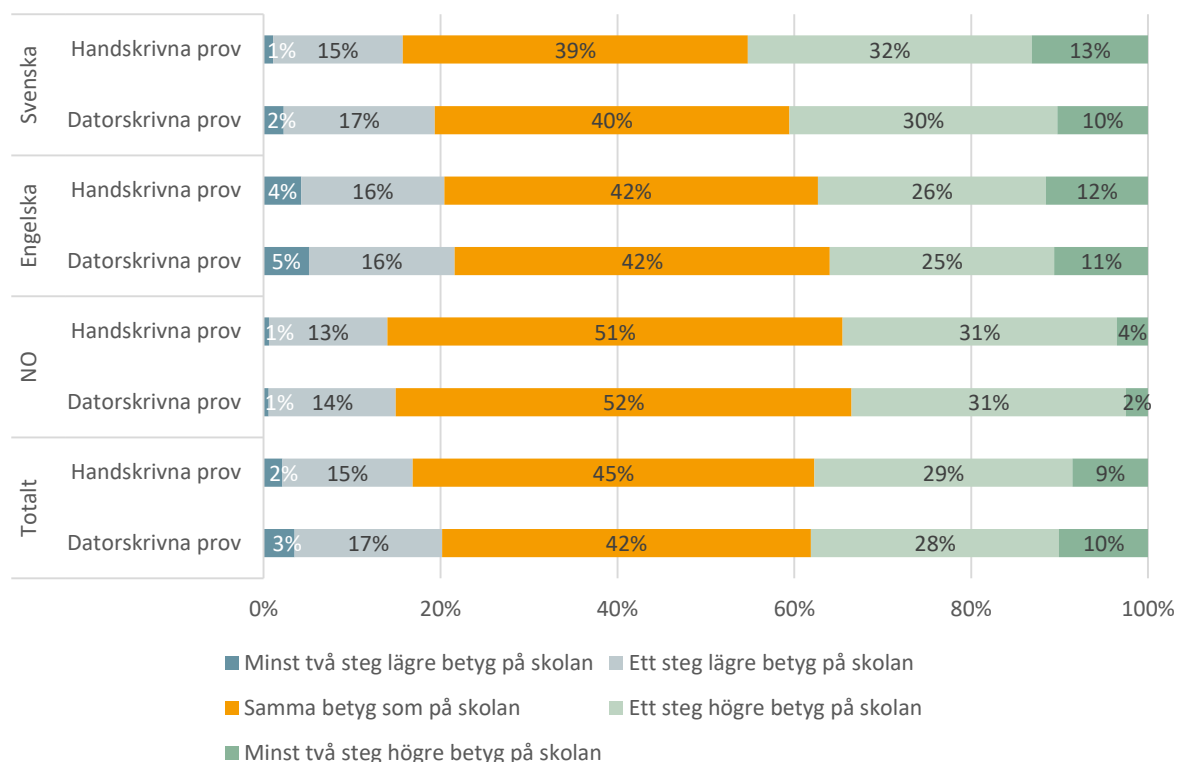
Tabell 1: Andel och antal delprov som skrivits på dator och för hand, efter ämne och årskurs/kurs.

<b>Ämne och årskurs</b>	<b>Delprovet är skrivet på dator</b>		<b>Delprovet är skrivet för hand</b>	
	<i>Andel (%)</i>	<i>Antal</i>	<i>Andel (%)</i>	<i>Antal</i>
Engelska åk 6	32 %	1 036	68 %	2 152
Engelska åk 9	39 %	1 160	61 %	1 835
Engelska gymn	76 %	2 916	24 %	903
Engelska vux	16 %	96	84 %	500
Svenska åk 9	57 %	1 621	43 %	1 230
Svenska gymn	87 %	3 552	13 %	550
Svenska vux	29 %	397	71 %	965
Biologi åk 9	8 %	161	92 %	1 925
Fysik åk 9	6 %	116	94 %	1 829
Kemi åk 9	11 %	245	89 %	2 022
<b>Totalt</b>	<b>45 %</b>	<b>11 300</b>	<b>55 %</b>	<b>13 911</b>

Jämförelser mellan ursprungs- och omdömnings ger överlag liknande resultat för elevlösningar som skrivits för hand som för dator, vilket framgår i figur 9 nedan. Bland delproven i svenska finns dock en liten skillnad. Av de handskrivna delproven har 45 procent fått en generösare bedömning på skolan, jämfört med 40 procent av de datorskrivna. Studier som jämfört hand- och datorskrivna texter har visat att det finns vissa skillnader mellan texterna. Exempelvis är datorskrivna texter oftare längre, har längre meningar och mer korrekt styckesindelning.<sup>15</sup> Utifrån resultaten i omdömnings finns dock inga tydliga tecken på att likvärdigheten i bedömningarna mellan ursprungs- och omdömnings påverkas avsevärt av om elevlösningarna skrivits för hand jämfört med på dator.

---

<sup>15</sup> SOU 2016:25



Figur 9: Resultat i omdömningen, efter ämne samt om provet är skrivet för hand eller på dator. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare (belägg i NO-ämnena har räknats om till skalan A–F).

## Stor variation mellan skolor

I detta avsnitt redovisas resultat på skolenhetsnivå för att visa den genomsnittliga skillnaden mellan skolans bedömning och omdömningen av de utvalda delproven. Skillnader redovisas som absoluta tal och det är därför inte möjligt att dra slutsatser om skolorna överlag bedömer prov striktare eller generösare än Skolinspektionens omdömare.<sup>16</sup>

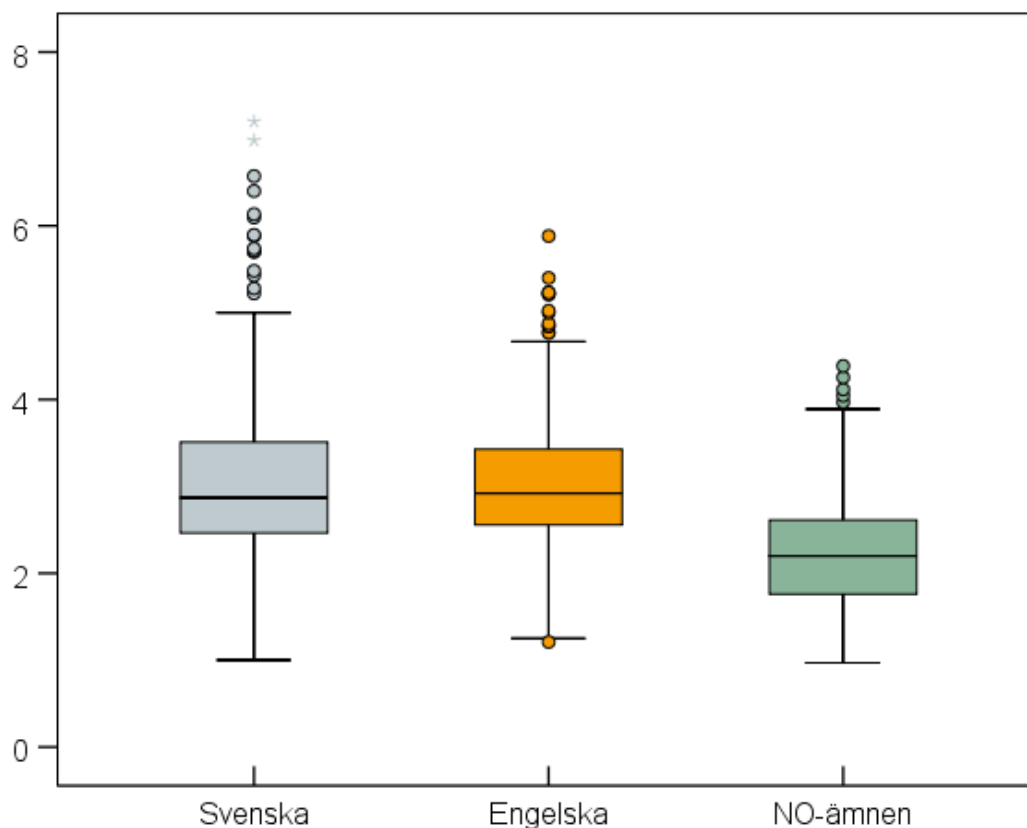
I figur 10 nedan redovisas den genomsnittliga skillnaden mellan olika skolor per ämnesgrupp. Fyra enheter på skalan motsvarar ett steg på betygsskalan A–F.<sup>17</sup> Hälften av skolorna ryms inom boxarna som avgränsas av den 25:e och 75:e percentilen. Den genomsnittliga skillnaden, medianen, anges av strecket i mitten av boxen. Avståndet mellan den 25:e och 75:e percentilen kallas för kvartilavstånd. De T-formade strecken visar det högsta/lägsta värdet som ligger max 1,5 kvartilavstånd från övre/nedre kvartilen. Eventuella värden utanför detta spann (så kallade extremvärden) visas som cirklar eller stjärnor i figuren.

Resultaten visar att medianen för de tre ämnesgrupperna ligger mellan 2,2 och 2,9 enheter på skalan. Det innebär att skillnaden mellan skolorna och omdömare i genomsnitt motsvarar över ett halvt betygssteg. Skillnaderna är som störst i svenska och engelska. Det finns dock en spridning mellan olika skolor. Variationen inom denna grupp närmast medianen, det vill säga kvartilavståndet, motsvarar cirka en enhet på skalan för

<sup>16</sup> Figuren i årets analys presenteras i absoluta tal. Det försvårar jämförelser med motsvarande figur från tidigare omdömningar.

<sup>17</sup> I NO-ämnena har belägg räknats om till skalan A–F, se bilaga 3.

de tre ämnesgrupperna. Skillnaden mellan den skola som har störst respektive minst avvikelse i jämförelse med ombedömningen motsvarar sex enheter på skalan för ämnet svenska. För övriga ämnen är avvikelsen något mindre.



Figur 10: Avvikelse i ombedömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många ombedömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. Totalt ingår därmed 256 skolor i svenska, 244 skolor i engelska och 133 skolor i NO-ämnen. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Värdet på y-axeln motsvarar den genomsnittliga avvikelsen på skolnivå, mätt i absoluta tal, mellan ursprungsbedömaren och ombedömaren. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnen har räknats om till denna skala).

En del av variationen mellan skolor kan bero på att det är svårt att bedöma de utvalda delproven som består av frågor med fritextsvar. I den analys som ingick i förra årets regeringsrapport belyste Skolinspektionen detta. I analysen fick två ombedömare bedöma samma elevlösning var för sig.<sup>18</sup> Resultatet visade att det i många fall blir lika stora skillnader mellan Skolinspektionens båda ombedömare som mellan läraren på skolan och respektive ombedömare.<sup>19</sup> Det är heller inte otänkbart att en lärare skulle kunna bedöma samma elevlösning olika, om det gått lång tid mellan bedömningarna och läraren hunnit glömma sin tidigare bedömning.<sup>20</sup> En annan faktor som bidrar till slumpmässiga variationer är att endast ett begränsat antal delprov från varje skola har ingått i ombedömningen.

<sup>18</sup> Om det skiljde minst två betygssteg mellan ursprungsbedömaren och ombedömaren gick proven i svenska till en andra ombedömning. Den låga samstämmigheten skulle delvis kunna förklaras av att elevlösningarna som ingår i den andra ombedömningen är svårare att bedöma.

<sup>19</sup> Statens skolinspektion, *Ombedömning av nationella prov 2017 – Fortsatt stora skillnader* (2018) s. 32-33.

<sup>20</sup> Statens skolverk, *Sambedömning i skolan – exempel och forskning* (Stockholm: Skolverket, 2013), s. 9.

Sammanfattningsvis visar resultaten på fortsatt bristande samstämmighet i bedömningarna av delprov med fritextsvar. Ombedömarnas bedömningar är inte ett facit och det är fullt möjligt att de överlag är striktare i sin roll som ombedömare. Det finns dock stora variationer mellan olika skolor, vilket tyder på bristande likvärdighet i bedömningen mellan landets skolor. På vissa skolor är skillnaderna mellan skolans och ombedömarnas bedömningar små, medan de på andra är omfattande. I följande avsnitt analyseras vilken betydelse bedömningsprocesserna på skolorna har för likvärdigheten i bedömningarna.

## Hur påverkas skillnaderna av bedömningsprocessen?

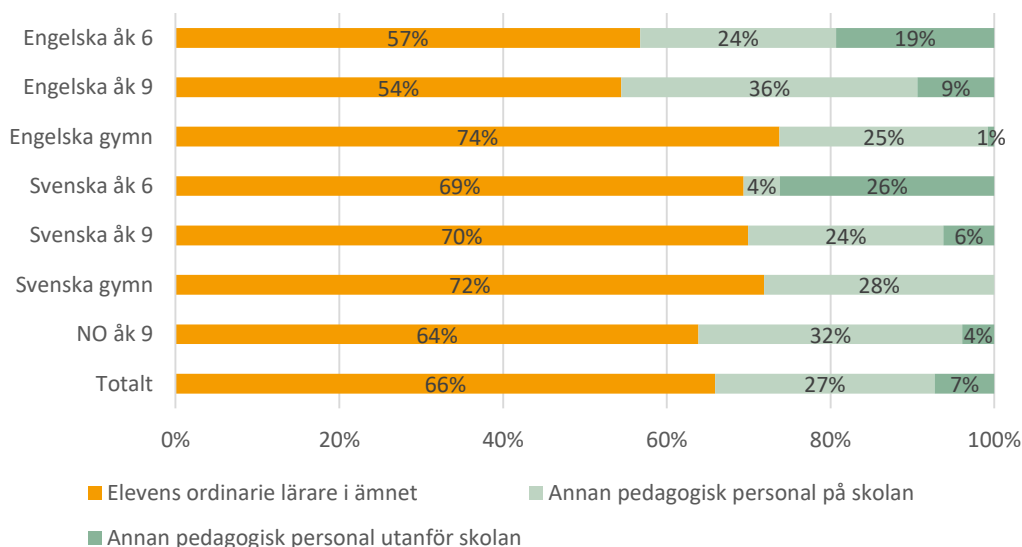
I avsnittet redogörs inledningsvis för hur de omdömda delproven har bedömts på skolorna. Därefter analyseras hur bedömningsprocessen påverkar skillnaderna mellan ursprungs- och omdömning samt om resultatet skiljer sig mellan delprov som är skrivna av flickor respektive pojkar och mellan delprov som skrivits för hand respektive på dator.

### Stora variationer i hur delproven bedöms

För att ta reda på hur bedömningsprocessen sett ut för de elevlösningar som omdömts har Skolinspektionen bitt skolorna att för varje insänd elevlösning besvara en enkät.<sup>21</sup> I enkäten ställdes frågor om vem som bedömt delprovet, om delprovet diskuterades med annan pedagogisk personal samt om delprovet var oidentifierat vid bedömningen. Enkäten var frivillig att besvara och därför finns inte uppgifter för samtliga elevlösningar som omdömts.

### Vanligast att elevens ordinarie lärare bedömt delprovet

Av enkätsvaren framgår att sett till samtliga ämnen, årskurser och kurser har elevens ordinarie lärare bedömt delprovet i 66 procent av fallen, se stapeln för totalen i figur 11 nedan. I övriga fall har extern bedömning tillämpats, 27 procent av delproven har bedömts av annan pedagogisk personal på skolan och 7 procent av annan pedagogisk personal utanför skolan. Det innebär att det likt tidigare år är vanligast att elevens ordinarie lärare bedömt delprovet och att skillnaderna är mycket små jämfört med svaren förra året. Det finns betydande variationer mellan ämnen. Att någon annan än elevens ordinarie lärare bedömt delprovet är vanligast i engelska och minst vanligt i svenska.



Figur 11: Andel delprov i omdömningen som ursprungligen bedömts av elevens ordinarie lärare respektive annan pedagogisk personal, efter ämne samt årskurs/kurs.

<sup>21</sup> Inga uppgifter om bedömningsprocessen har efterfrågats för delprov genomförda inom den kommunala vuxenutbildningen.

I svenska är andelen delprov som bedömts av elevens ordinarie lärare i stort sett densamma för grundskolans årskurs 6 och 9 och gymnasieskolan. I ämnet engelska syns en variation mellan grundskolan och gymnasiet. Andelen delprov som bedömts av elevens ordinarie lärare är som störst för gymnasieskolan, 74 procent, och som minst för årskurs 9, 54 procent.

Det är vanligare med extern bedömning bland kommunala skolor än bland fristående (se figur 32 i bilaga 4). Andelen delprov som bedömts av annan lärare på skolan är 30 procent bland kommunala skolor jämfört med 24 procent bland fristående. Andelen delprov som bedömts av lärare utanför skolan är 13 procent bland kommunala skolor jämfört med 3 procent bland fristående.

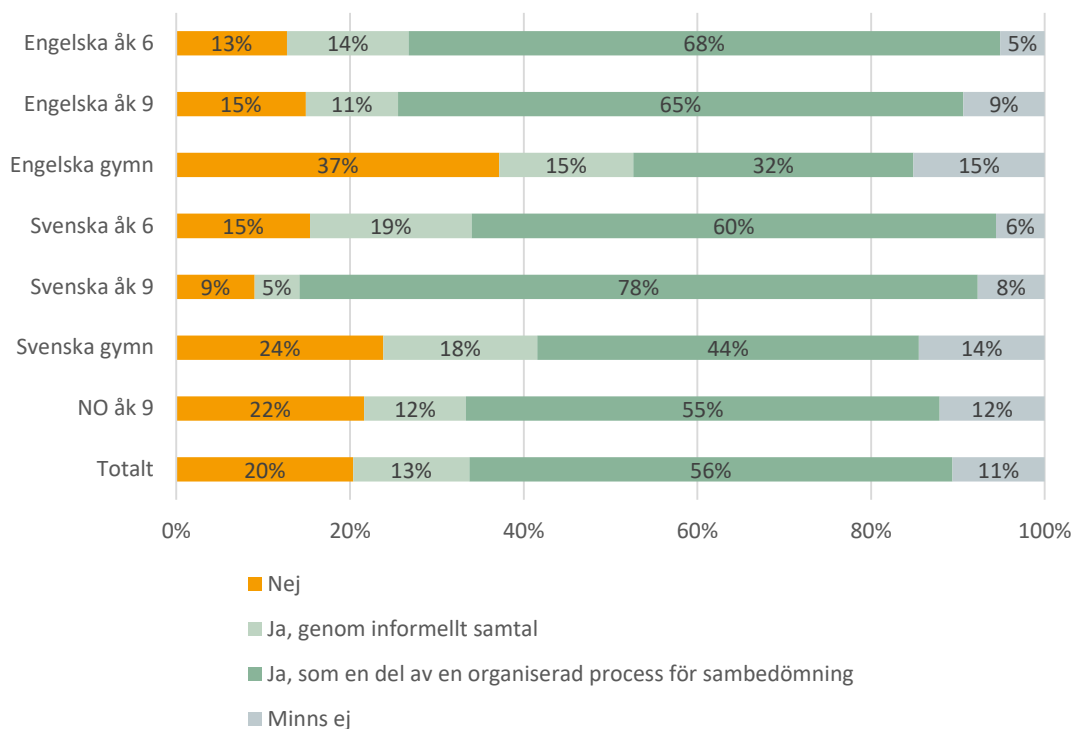
### Sju av tio delprov sambedömdes

Sambedömning innebär att lärare samarbetar vid bedömning och betygssättning, till exempel genom att bedöma elevens prestationer tillsammans eller genom att diskutera bedömningen.<sup>22</sup> Med *organiserad* sambedömning avses här att skolan har strukturer och rutiner för att lärare ska bedöma delprov i samarbete eller i diskussion med kollegor. Detta indikeras av enkätsvaret "Ja, som en del av en organiserad process för sambedömning". Med *informell* sambedömning avses här att lärare bedömt delprov i samråd eller i samarbete med kollegor trots att formella strukturer och rutiner för sambedömning saknas på skolan. Informell sambedömning kan också ske genom att kollegor diskuterar eller samarbetar kring bedömning och betygssättning utanför de eventuella formella strukturer och rutiner för sambedömning som finns på skolan. Informell sambedömning indikeras av svarsalternativet "Ja, genom informellt samtal".

Totalt har det för nära sju av tio delprov tillämpats någon form av sambedömning i samband med bedömningen av elevlösningen (figur 12). Två av tio delprov har bedömts utan sambedömning. Av de delprov som ingår i årets omgång är det fem procentenheter lägre andel som uppgett att delproven sambedömts i jämförelse med de delprov som ingick förra omgången.

---

<sup>22</sup> Statens skolverk, *Sambedömning i skolan – exempel och forskning* (Stockholm: Skolverket, 2013), s. 11.



Figur 12: Andel delprov i omdömningen som sambedömts vid den ursprungliga bedömningen på skolan, efter ämne samt årskurs/kurs.

Den vanligaste formen för sambedömning, oavsett ämne, årskurs och kurs var att den genomfördes som en del av en organiserad process. Totalt har 56 procent av delproven sambedömts på detta sätt. Denna typ av sambedömning var framförallt vanlig i svenska årskurs 9 (75 %). Det var mindre vanligt med organiserad sambedömning på gymnasiet, där 32 procent av delproven i engelska och 44 procent av delproven i svenska sambedömts som en del av en organiserad process.

Av enkätsvaren framgår att det varit vanligare att någon form av sambedömning tillämpats på kommunala skolor jämfört med fristående (se figur 33 i bilaga 4). De kommunala skolorna har uppgett att någon form av sambedömning tillämpats i 76 procent av fallen medan motsvarande andel för fristående skolor är 64 procent. Det är främst den organiserade sambedömningen som är vanligare vid kommunala skolor (61 %) än vid fristående (52 %).

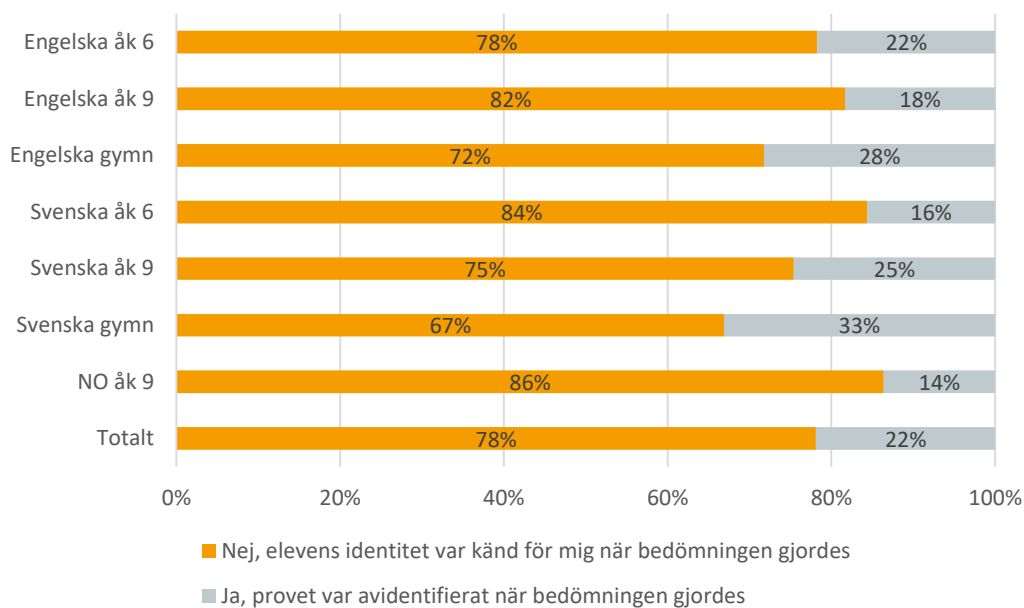
### Två av tio delprov var avidentifierade

Av enkätsvaren framkommer att sett till alla ämnen, årskurser och kurser så var 22 procent av elevlösningarna avidentifierade vid bedömningen, vilket framgår i figur 13 nedan.<sup>23</sup> Andelen delprov som avidentifierats är ungefär lika stor som i förra årets rapport, då var 24 procent av delproven avidentifierade. Andel avidentifierade delprov är högst för svenska på gymnasiet (33 %) och engelska på gymnasiet (28 %) samt i svenska i årskurs 9 (25 %). Lägst är andelen i NO-ämnena (14 %) och svenska i årskurs 6 (16 %). Det är vanligare med avidentifiering i kommunala skolor.

<sup>23</sup> Från och med höstterminen 2018 ska uppsatser i svenska, svenska som andraspråk och engelska skrivas på dator eller annan digital enhet och vara avidentifierade vid bedömningen (Regeringsbeslut U2017/03739/GV), i årskurs 9 i grundskolan, årskurs 10 i specialskolan, gymnasieskolan och vuxenutbildningen (Skolverket.se). De elevlösningar som ingår i denna rapport skrevs vårterminen 2018 och omfattades därmed inte av beslutet.



Bland de kommunala skolorna var 26 procent av delproven avidentifierade jämfört med 19 procent bland de fristående (se figur 34 i bilaga 4).



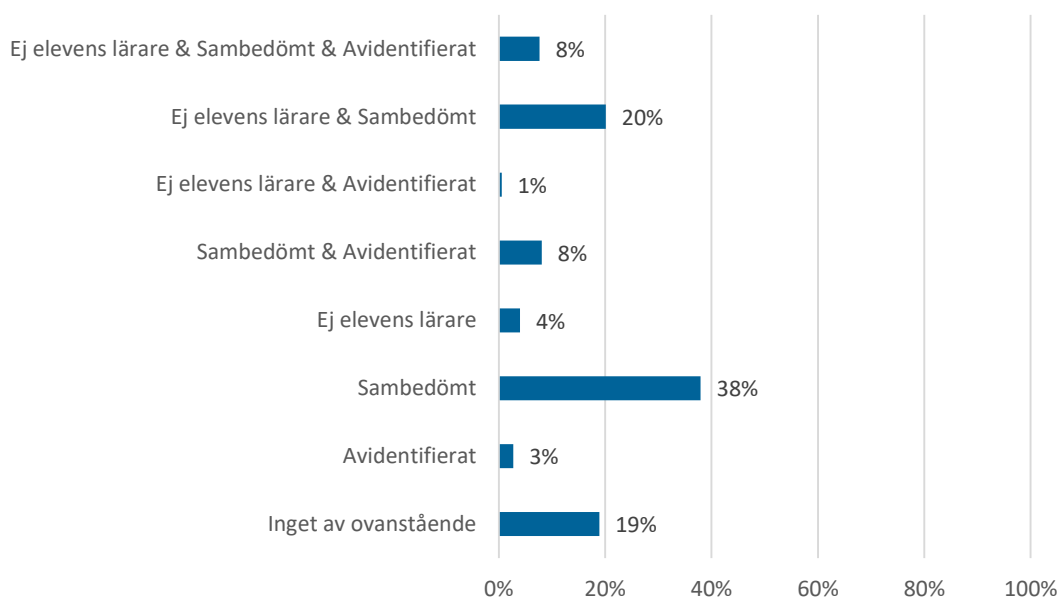
Figur 13: Andel delprov i ombedömningen som var avidentifierade vid den ursprungliga bedömningen på skolan, efter ämne samt årskurs/kurs.

## Stora variationer i hur bedömningsmetoderna kombineras

I figur 14 redovisas resultat för *kombinationer av svar* på de tre frågorna om hur bedömningen av de nationella delproven gått till på skolan. Figuren visar fördelningen mellan kombinationer för de 73 procent av delproven med kompletta svar på frågorna om bedömningsmetoderna.

Det är åtta procent av delproven som bedömts på det sätt som antas *mest* gynnsamt för likvärdig bedömning, det vill säga att någon annan än elevens ordinarie lärare har bedömt delproven, att delproven sambedömts samt att de varit avidentifierade vid bedömningen. Ungefär dubbelt så många delprov (19 %) har bedömts på det sätt som antas *minst* gynnsamt för en likvärdig bedömning, det vill säga att elevens ordinarie lärare har bedömt delproven, att delproven inte har sambedömts och att de inte har varit avidentifierade.

Det är 45 procent av delproven som haft *en* gynnsam faktor för likvärdig bedömning och den absolut vanligaste faktorn är i dessa fall sambedömning. Bland de 29 procent av delproven som haft *två* gynnsamma faktorer för likvärdig bedömning är det vanligast att någon annan än elevens lärare bedömt delproven och att det dessutom skett en sambedömning. Kombinationen att delprovet anonymiserats samt bedömts av någon annan än elevens lärare är däremot mycket ovanlig.



Figur 14: Svar på tre frågor om hur bedömningen av de nationella delproven gått till på skolan. I figuren redovisas i vilken utsträckning extern bedömning, sambedömning och aidentifiering tillämpats i kombination eller var för sig.

Sammanfattningsvis visar resultatet att inslagen av sambedömning, extern bedömning och aidentifiering varierar mellan skolor, ämnen, årskurser och kurser, liksom mellan skolor med kommunal respektive enskild huvudman. Vilken betydelse bedömningsprocessen har för likvärdigheten i bedömningen undersöks i nästkommande avsnitt.

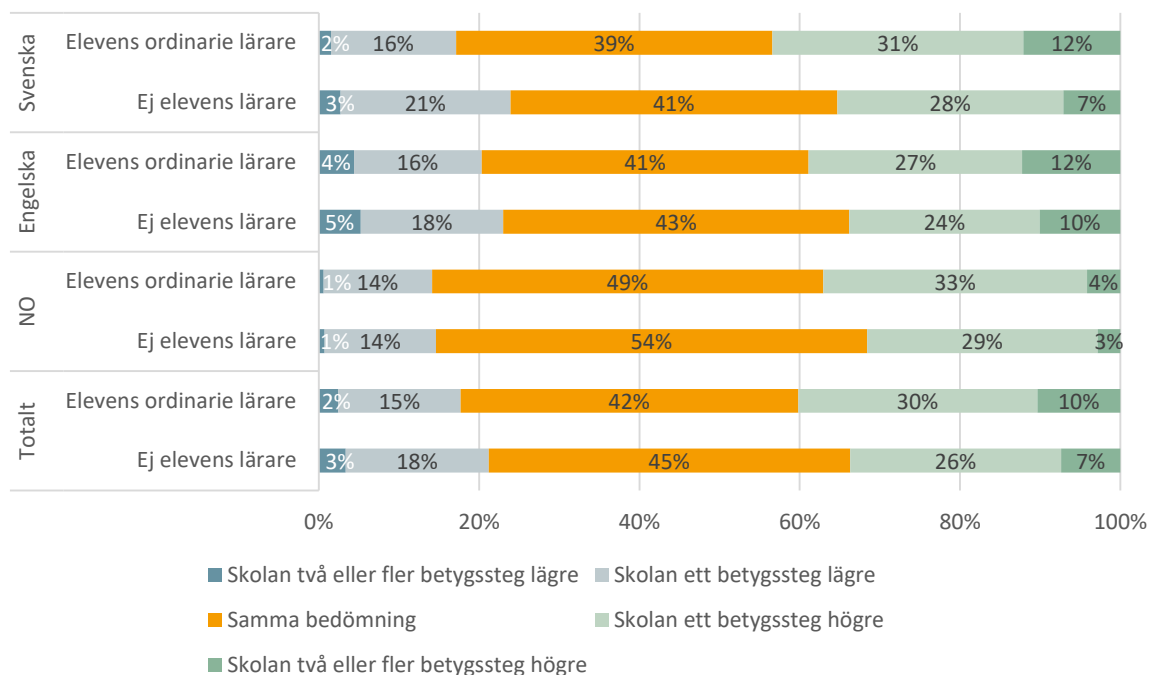
## Bedömningsprocessen har betydelse

I figur 15–17 visas skillnader mellan ursprungs- och ombedömning uppdelat per bedömningsmetod och ämne<sup>24</sup>. I figurerna tas ingen hänsyn till vilka ytterligare bedömningsmetoder förutom den figuren redogör för, som tillämpats vid bedömningen av delprovet. Detta innebär exempelvis att en del av de aidentifierade elevlösningarna även kan ha sambedömts och/eller bedömts av någon annan än elevens ordinarie lärare.

### Extern bedömning minskar andelen delprov som bedömts generösare på skolan

I figur 15 jämförs resultatet från ombedömningen mellan delprov som bedömts av elevens ordinarie lärare respektive av någon annan än elevens ordinarie lärare (så kallad *extern bedömning*). Den tydligaste effekten av extern bedömning är att andelen delprov som bedömts generösare av skolan jämfört med ombedömarna minskar, från 40 procent till 33 procent totalt sett för alla ämnen. I engelska och svenska ökar samtidigt andelen striktare bedömningar. I NO-ämnen ser vi istället att andelen delprov som fått samma bedömning av skolans lärare och ombedömarna ökar, från 49 procent till 54 procent. Totalt, för samtliga tre ämnen, får 45 procent av delproven samma bedömning om extern bedömning tillämpas jämfört med 42 procent om de bedöms av elevens ordinarie lärare.

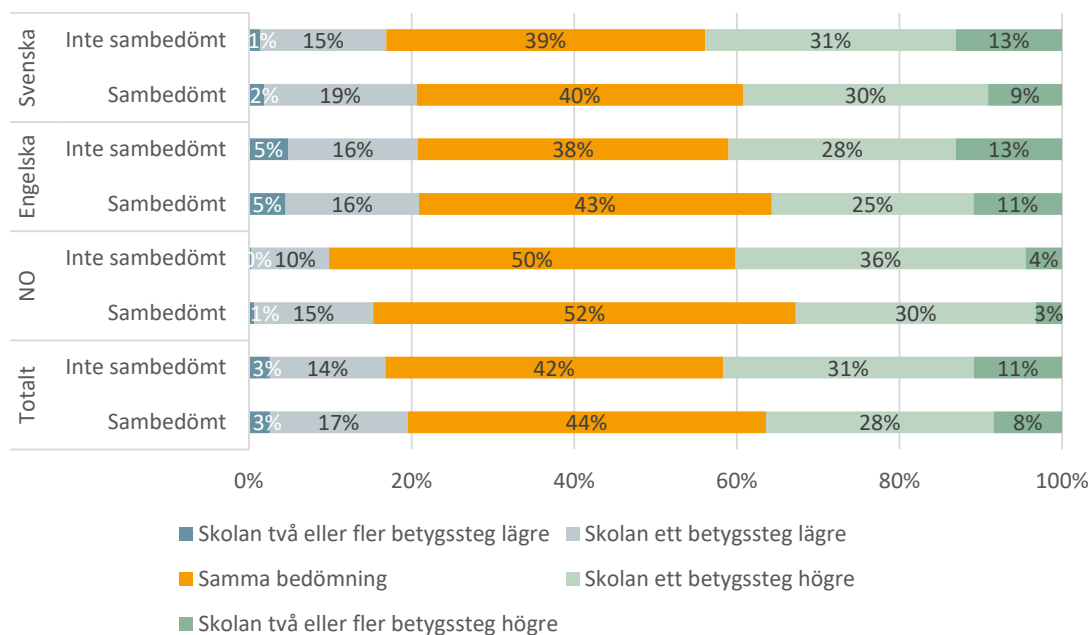
<sup>24</sup> I figur 35-37 i bilaga 4 redovisas skillnader mellan ursprungs- och ombedömning uppdelat per bedömningsmetod och årskurs/kurs.



Figur 15. Resultat i ombedömningen för delprov som ursprungligen bedömts av elevens ordinarie lärare respektive av annan pedagogisk personal, efter ämne. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

### Sambedömning minskar andelen delprov som bedömts generösare

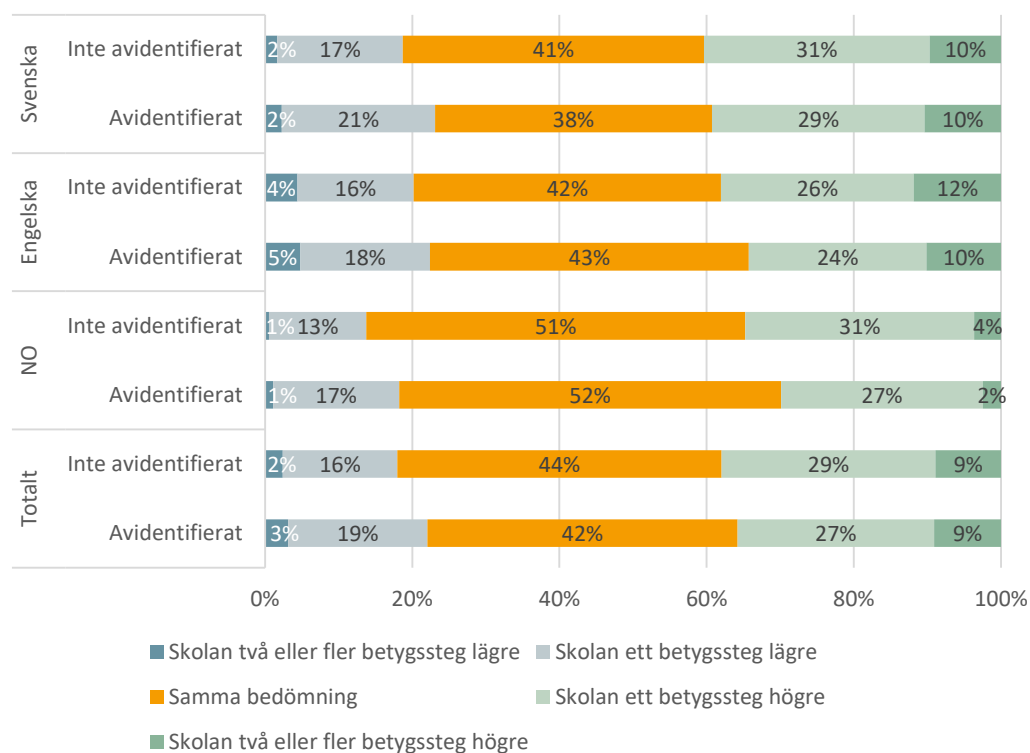
I figur 16 jämförs ursprungs- och ombedömning utifrån om delproven sambedömts eller bedömts enskilt. Likt extern bedömning minskar tillämpningen av sambedömning andelen delprov som bedöms generösare på skolan, från 42 till 36 procent, totalt sett för samtliga ämnen. I svenska och NO-ämnena ökar samtidigt andelen striktare bedömningar. I engelska är andelen delprov som ges en striktare bedömning på skolan densamma för sambedömda och enskilt bedömda delprov. Istället ökar andelen delprov som ges samma bedömning, från 38 till 43 procent. Totalt är det små skillnader, 44 procent av delproven som sambedöms får samma betyg av de olika bedömarna jämfört med 42 procent för delprov som bedöms enskilt.



Figur 16. Resultat i ombedömningen för delprov som sambedömts respektive bedömts enskilt, efter ämne. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

### Andelen striktare bedömningar är större för avidentifierade delprov

I figur 17 redovisas skillnaden mellan ursprungs- och ombedömning av delprov som varit avidentifierade respektive inte avidentifierade. Resultaten är överlag i linje med extern bedömning och sambedömning men inte lika tydliga. Jämfört med ombedömningen så minskar de generösare bedömningarna på skolan medan andelen striktare ökar en aning. Framförallt i NO-ämnena. Tvärtemot resultaten för extern bedömning och sambedömning är andelen delprov som får samma betyg av de olika bedömarna lägre för avidentifierade delprov (42 %) jämfört med delprov som inte varit avidentifierade (44 %).



Figur 17. Resultat i ombedömningen för delprov som var avidentifierade vid ursprungsbedömningen respektive för delprov där elevens identitet var känd, efter ämne. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

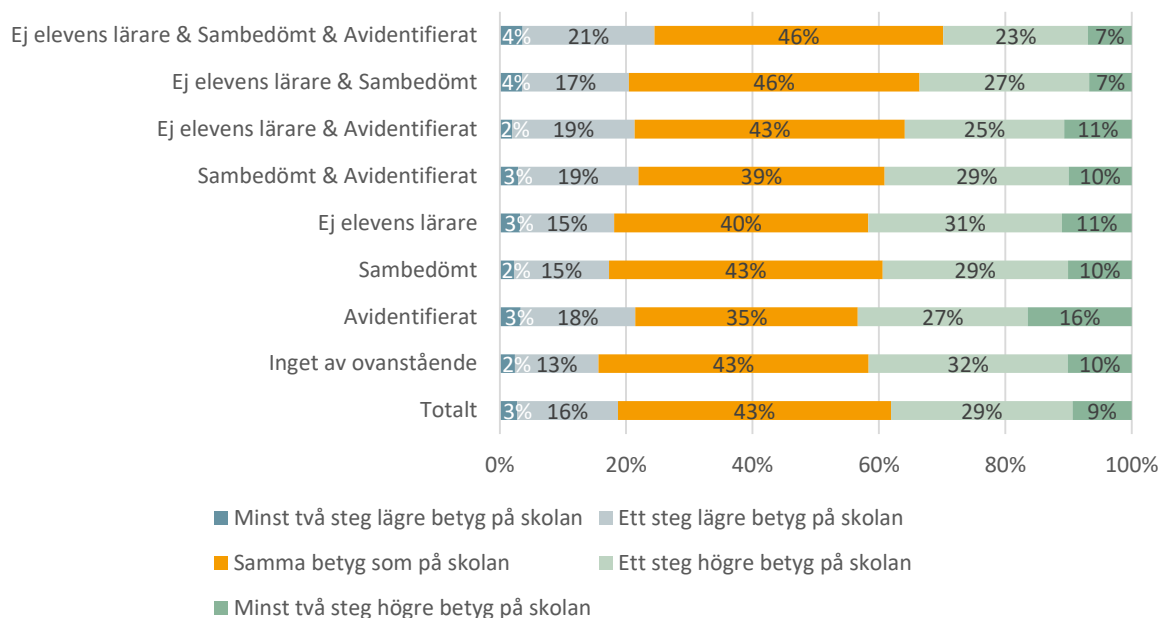
## Kombinationer av bedömningsmetoder minskar tendensen att skolan gör en mer generös bedömning

I detta avsnitt redovisas hur en enskild bedömningsmetod eller *kombinationer* av de tre bedömningsmetoderna påverkar i vilken mån ombedömningen skiljer sig från den ursprungliga bedömningen.<sup>25</sup> Detta till skillnad från resultaten i figur 15–17 där metoderna studerats enskilt men kan innehålla kombinationer. I figur 18 innebär det alltså att kategorin "Ej elevens lärare" innehåller resultat för delprov som inte bedömts av elevens ordinarie lärare, och där varken *sambedömning* eller *avidentifiering* tillämpats. På motsvarande sätt innehåller gruppen "Ej elevens lärare & Sambedömt" resultat för delprov som inte bedömts av elevens ordinarie lärare och som sambedömts, *men inte avidentifierats*.

Som framgår av figur 18 får i genomsnitt 43 procent av de utvalda delproven samma betyg vid Skolinspektionens ombedömning som vid den ursprungliga bedömningen på skolan. Det är dock fler som får ett högre betyg vid den ursprungliga bedömningen på skolan (38 %) än ett lägre betyg (19 %). Resultaten i figur 18 visar att tendensen att skolans lärare gör en mer generös bedömning än Skolinspektionens ombedömare minskar om sambedömning, avidentifiering och extern bedömning tillämpats vid ursprungsbedömningen. Bland delprov utan dessa faktorer i bedömningsprocessen gör skolan en generösare bedömning i 42 procent av fallen och en striktare bedömning i 15 procent, vilket kan jämföras med 30 respektive 25 procent då alla tre faktorerna samtidigt förekommit när delprovet bedömts. Andelen delprov där bedömarna är helt överens ökar däremot endast marginellt när de tre bedömningsmetoderna som ingår i denna rapport kombineras, jämfört

<sup>25</sup> Resultat från den kommunala vuxenutbildningen är ej med i den här analysen eftersom de inte svarar på frågor om bedömningsprocessen.

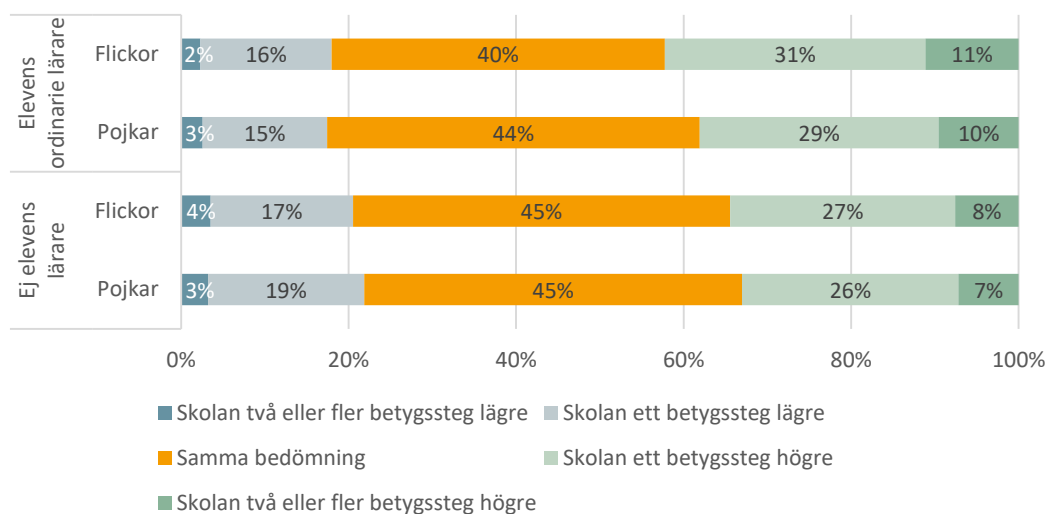
med delprov där ingen av faktorerna tillämpats. Detta kan tolkas som att kombinationerna har en återhållande effekt på betygssättningen men liten effekt på likvärdigheten eftersom andelen generösare bedömningar minskar samtidigt som andelen striktare bedömningar ökar. Denna tendens är som tydligast när alla tre faktorer kombinerats men finns där även när två av faktorerna kombinerats parvis.



Figur 18: Resultat i ombedömningen, efter svar på tre frågor om hur bedömningen av de nationella delproven gått till på skolan (belägg i NO-ämnena har räknats om till skalan A–F). Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

## Vid extern bedömning ges flickor och pojkar i högre grad samma betyg av de olika bedömarna

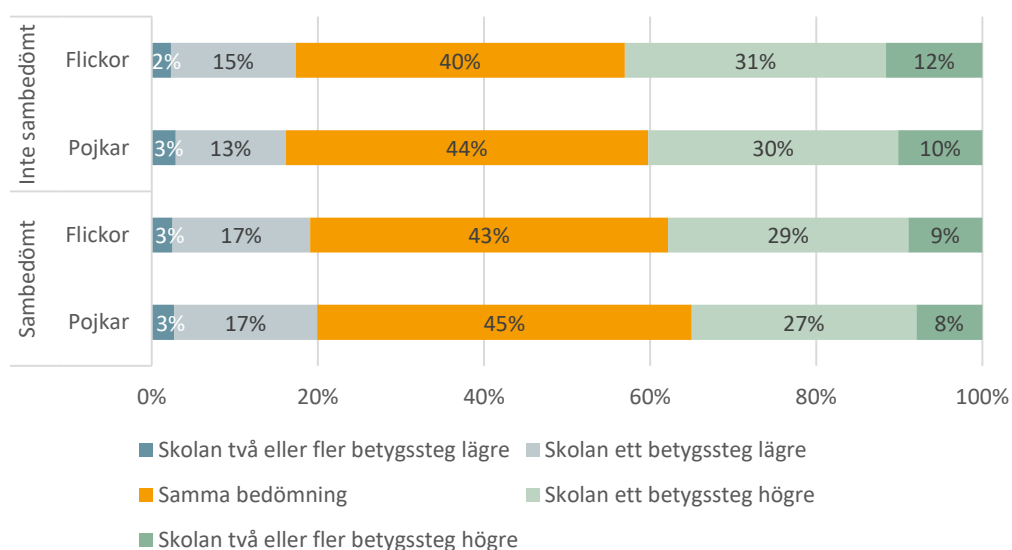
I figur 19 visas resultat från ombedömningen för flickor respektive pojkar uppdelat på om delprovet bedömts externt eller av elevens ordinarie lärare. Resultatet från föregående avsnitt, att extern bedömning minskar andelen delprov som bedömts generösare på skolan, gäller både för delprov skrivna av flickor och pojkar. Andelen delprov som fått samma bedömning är större när extern bedömning tillämpats, i synnerhet för flickor. Detta innebär att skillnaderna mellan flickor och pojkars resultat minskar. För delprov som bedömts av elevens ordinarie lärare är andelen prov som fått samma bedömning högre för pojkar (44 %) än för flickor och (40 %), medan det för delprov som bedömts externt inte finns någon skillnad mellan könen (45 %).



Figur 19: Resultat i omdömningen för delprov som ursprungligen bedömts av elevens ordinarie lärare respektive av annan pedagogisk personal, efter elevens kön. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

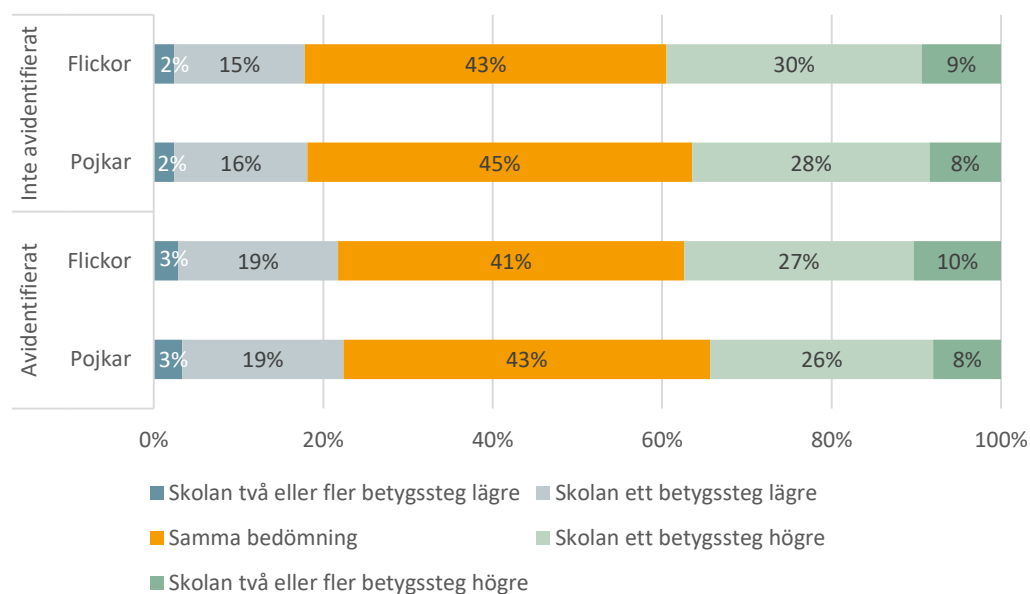
Avseende delprov som sambedömts respektive bedömts enskilt visar resultaten i figur 20 nedan små skillnader mellan flickor och pojkars delprov. Det generella resultatet står sig för båda könen, att andelen delprov där ursprungsbedömarna gjort en mer generös bedömning är något större för delprov som bedömts enskilt jämfört med delprov som sambedömts. Samtidigt är andelen som bedömt striktare något större för delprov som sambedömts jämfört med dem som bedömts enskilt.

Likt extern bedömning visar resultaten för sambedömda delprov att skillnaderna i avvikelser mellan könen är något mindre för delprov som sambedömts jämfört med delprov som bedömts enskilt. Skillnaden är dock inte lika tydlig och mycket liten. För delprov som bedömts enskilt är andelen delprov som fått samma bedömning högre för pojkar (44 %) än för flickor (40 %), en skillnad på fyra procentenheter. För delprov som sambedömts är skillnaden mellan könen två procentenheter, 45 procent för pojkar jämfört med 43 procent för flickor.



Figur 20: Resultat i omdömningen för delprov som sambedömts respektive bedömts enskilt, efter elevens kön. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

Avseende delprov som avidentifierats respektive där elevens identitet varit känd för bedömaren visar resultaten för flickor respektive pojkar ingen nämnvärd skillnad (figur 21). Andelen delprov där skolan gjort en mer generös bedömning minskar något för både flickor och pojkar när delproven avidentifierats. Samtidigt ökar andelen delprov där skolan har gjort en mer strikt bedömning något.

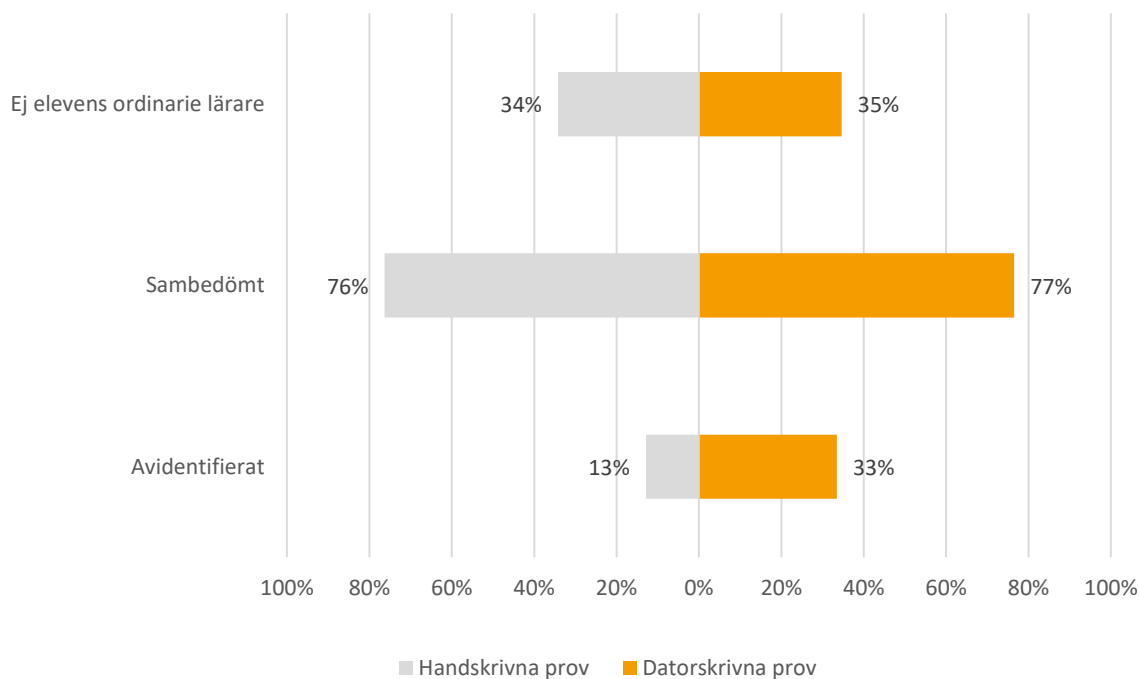


Figur 21: Resultat i ombedömningen för delprov som var avidentifierade vid ursprungsbedömningen respektive för delprov där elevens identitet var känd, efter elevens kön. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

## Bedömningsprocessen har betydelse för både hand- och datorskrivna delprov

Figur 22 visar hur elevlösningarna har bedömts utifrån om de skrivits för hand eller på dator. Ungefär en tredjedel av de hand- och datorskrivna delproven har bedömts av någon annan än elevens lärare. Resterande två tredjedelar har bedömts av elevens ordinarie lärare. Det är inte heller några stora skillnader i andelen elevlösningar som har sambedömts, 76 procent av de handskrivna jämfört med 77 procent av de datorskrivna. Däremot är det vanligare att de datorskrivna delproven har avidentifierats, 33 mot 13 procent av de handskrivna delproven.



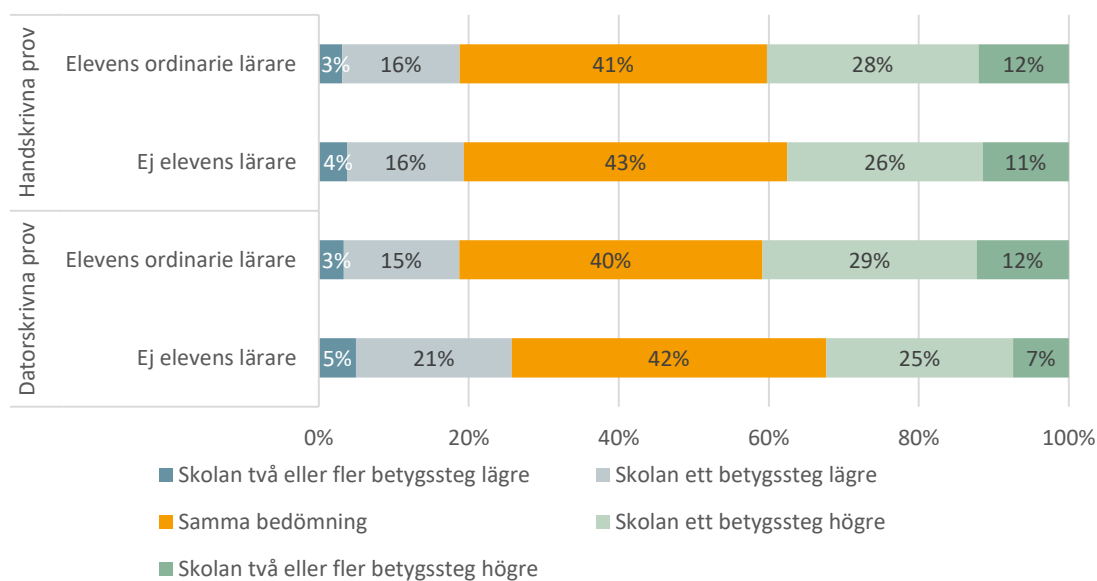


Figur 22: Andel delprov som bedömts av någon annan än elevens lärare, sambedömts respektive aidentificerats, efter om delprovet skrivits för hand eller på dator.

I figurerna 23-25 jämförs hur utfallet av ombedömningen skiljer sig mellan olika bedömningsmetoder och mellan delprov som skrivits för hand respektive på dator. I figurerna ingår delprov i svenska och engelska, men inte i NO-ämnena.<sup>26</sup>

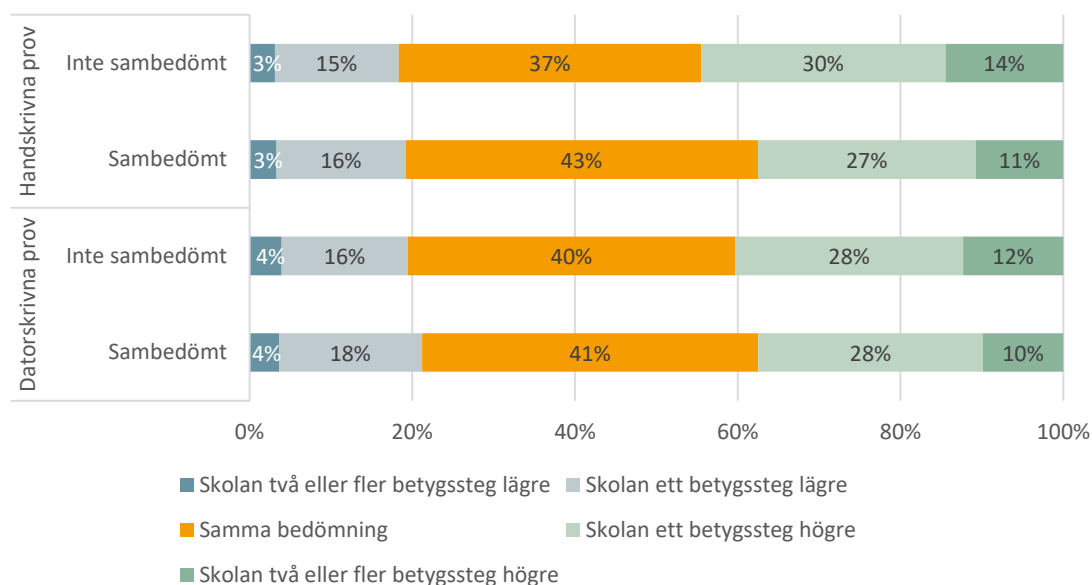
Andelen delprov som bedömts generösare på skolan är något mindre när delproven bedömts av någon annan än elevens lärare. Det gäller främst delprov som skrivits på dator (figur 23). För datorskrivna delprov är andelen generösare bedömningar 41 procent för delprov som bedömts av elevens lärare jämfört med 32 procent för delprov som bedömts av annan lärare. Resultaten visar också att när extern bedömning av de datorskrivna delproven tillämpas är det vanligare att skolan har gjort en striktare bedömning.

<sup>26</sup> NO-ämnena har uteslutits på grund av att skillnaden mellan ursprungsbedömning och ombedömning är markant mindre för delproven i NO jämfört med de övriga ämnena samtidigt som majoriteten av delproven i NO är skrivna för hand. Det innebär att resultaten för NO-ämnena får stort genomslag i kategorierna för handskrivna delprov när jämförelser görs totalt för alla ämnen. En uppdelning av resultaten per ämne var inte möjlig att genomföra med det befintliga dataunderlaget.



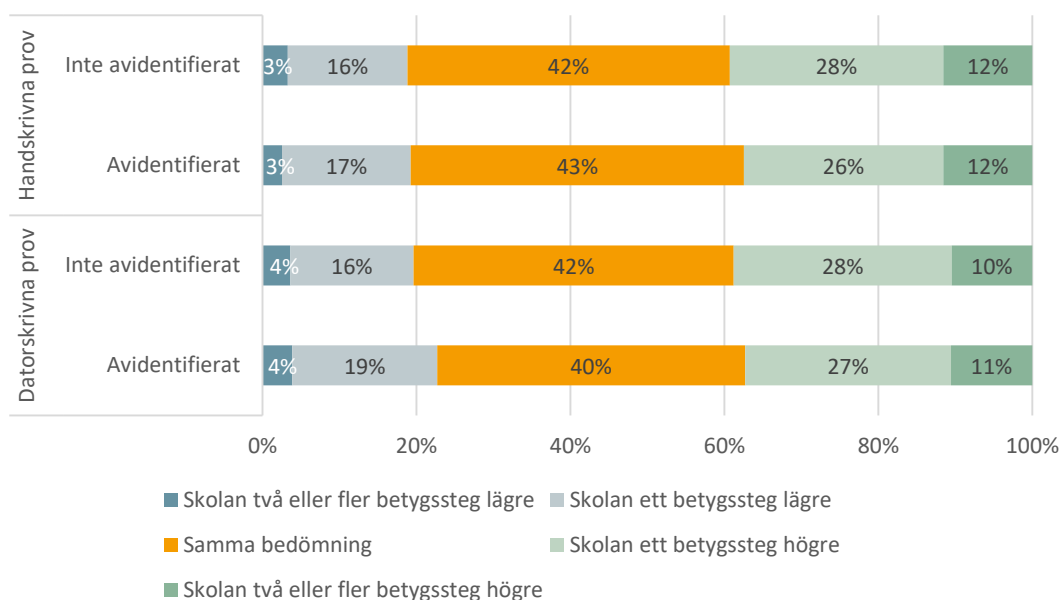
Figur 23: Resultat i omdömningen för delprov i svenska och engelska som bedömts av elevers ordinarie lärare respektive annan pedagogisk personal, efter om delprovet skrivits för hand eller på dator. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

Sambedömning minskar andelen delprov där skolan har satt ett generösare betyg än omdömare, framförallt när delproven skrivits för hand (figur 24). För handskrivna delprov är andelen generösare bedömningar 44 procent för delprov som bedömts enskilt och 38 procent för delprov som sambedömts. För de handskrivna delproven är andelen delprov som fått samma bedömning i ursprungs- och omdömning högre för delprov som sambedömts. För de datorskrivna delproven är skillnaden inte lika stor.



Figur 24: Resultat i omdömningen för delprov i svenska och engelska som sambedömts respektive bedömts enskilt, efter om delprovet skrivits för hand eller på dator. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omdömare.

Skillnaderna mellan ursprungs- och ombedömning skiljer sig mycket lite mellan delprov som avidentifierats och inte, både för delprov som skrivits på dator och för hand (figur 25). Andelen generösare bedömda delprov på skolan är något mindre när delproven är avidentifierade, både för hand- och datorskrivna delprov. Andelen striktare bedömda delprov är något större för delprov som skrivits på dator.



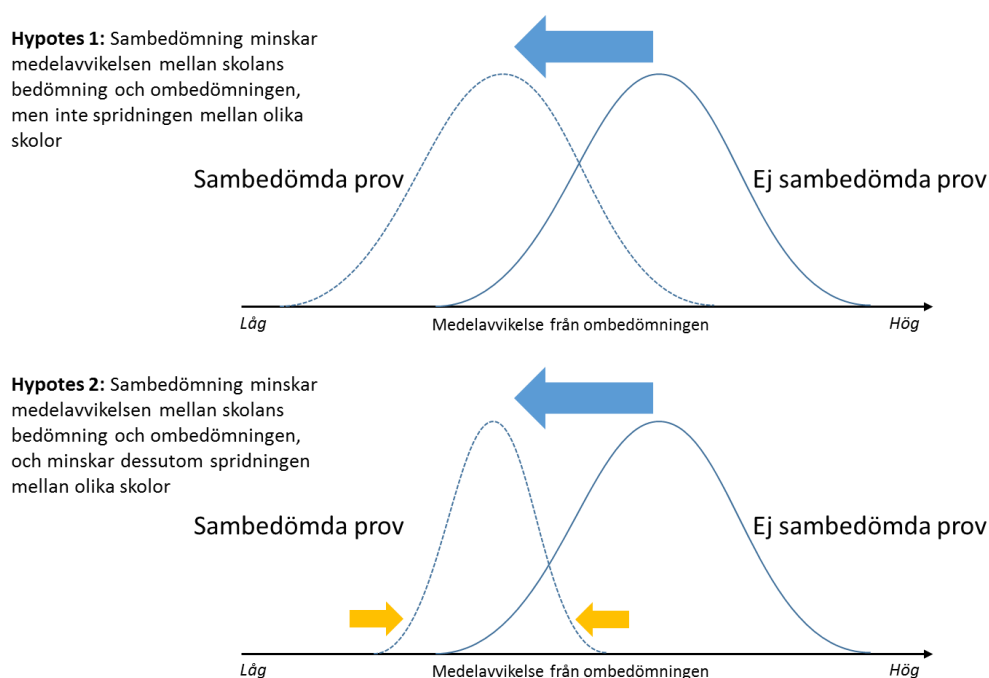
Figur 25: Resultat i ombedömningen för delprov i svenska och engelska som var avidentifierade vid ursprungsbedömningen respektive för delprov där elevens identitet var känd, efter om delprovet skrivits för hand eller på dator. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

Sammantaget indikerar resultaten att bedömningsprocessen har betydelse för likvärdigheten både för delprov som skrivits för hand och på dator, men i lite olika utsträckning beroende på vilken bedömningsmetod som används. Liksom i figurerna 15-17 tas här inte hänsyn till vilka ytterligare bedömningsmetoder, förutom den stapeln redogör för, som tillämpats vid bedömningen av delprovet. Detta innebär exempelvis att en del av de avidentifierade elevlösningarna även kan ha sambedömts och/eller bedömts av någon annan än elevens ordinarie lärare. Med hänsyn till detta samt att resultaten inte delats upp per ämne ska de tolkas med försiktighet. En uppdelning där de olika bedömningsmetoderna delas upp i åtta olika möjliga kombinationer (såsom i figur 18) eller per ämne (såsom i figur 9) var inte möjlig att genomföra med det befintliga dataunderlaget.

## Sambedömning minskar variationen mellan skolor

Föregående avsnitt visar att extern bedömning (bedömning gjord av annan lärare än elevens ordinarie), sambedömning och avidentifiering generellt sett har en återhållande effekt på delprovsbetygen, samt att kombinationer av dessa bedömningsmetoder tenderar att leda till mindre generösa bedömningar. Detta resulterar i ett något lägre genomsnittligt betyg än när elevens ordinarie lärare gör motsvarande bedömningar utifrån en icke avidentifierad elevlösning och utan sambedömning.

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.



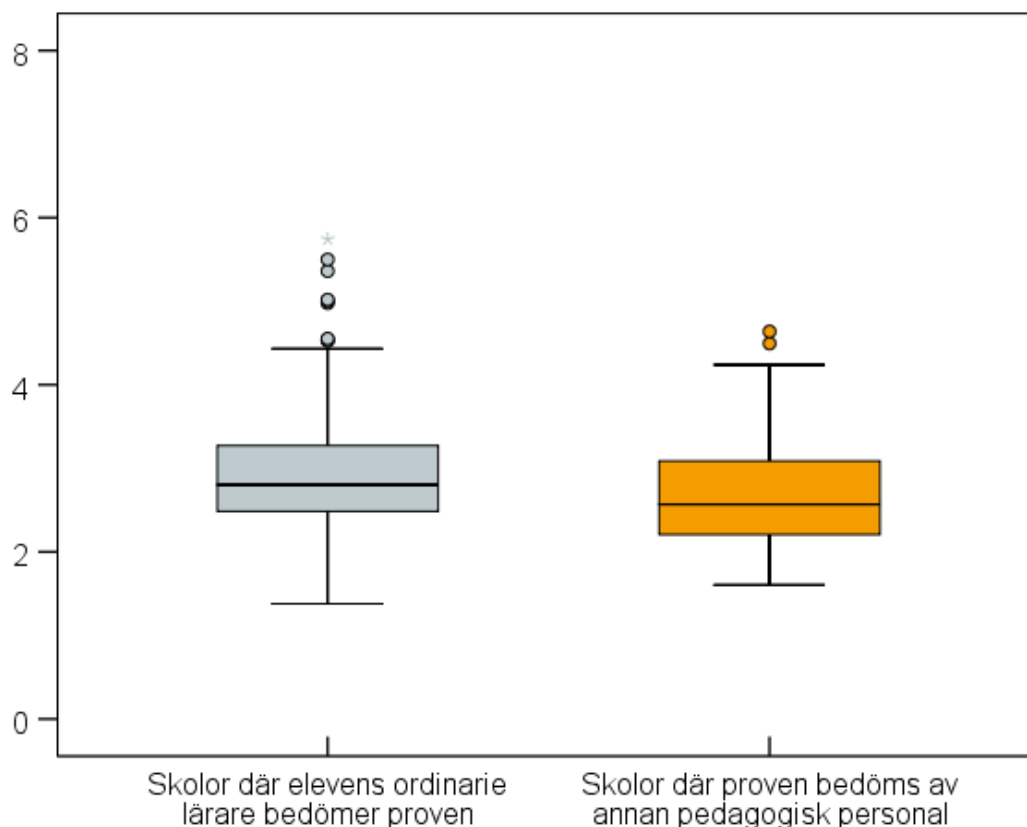
Figur 26: Två alternativa hypoteser kring effekten av sambedömning, extern bedömning och avidentifiering.

I analysen ingår de skolor som uppfyllt Skolinspektionens kriterier för att få en skolenhetsrapport (se kriterier i bilaga 1) och som sambedömer delprov på skolan, tillämpar extern bedömning eller avidentifierar delprov. För att visualisera spridningen redovisas resultaten i boxplotdiagram. Boxarna visar den 25:e och 75:e percentilen, inom detta spann ryms hälften av skolorna (se figur 27). Avståndet mellan den 25:e och 75:e percentilen motsvarar längden på boxen och kallas *kvartilavstånd*. Det heldragna strecket genom boxen visar medianen, det vill säga genomsnittsskillnaden. De T-formade strecken visar det högsta/lägsta värdet som ligger max 1,5 kvartilavstånd från övre/nedre kvartilen. Eventuella värden utanför detta spann (så kallade extremvärden) visas som cirklar/stjärnor i figuren.

Observera att skillnader redovisas som absoluta tal, vi tar här inte hänsyn till om skillnaden är negativ eller positiv. Det är därför inte möjligt att dra slutsatser om skolorna överlag bedömer delprov striktare eller generösare än Skolinspektionens ombedömare.

Figur 27 visar att skolor som tillämpar extern bedömning har en något lägre genomsnittsskillnad i förhållande till Skolinspektionens ombedömningar jämfört med skolor där elevens lärare bedömer delproven. Skillnaden är emellertid liten då medianen för skolor med extern bedömning ligger på 2,6 enheter jämfört med 2,8 enheter för skolor som inte tillämpar extern bedömning. Då fyra enheter motsvarar ett betygssteg utgör 2,8 enheter knappt tre fjärdedels steg på betygsskalan A–F.

Extern bedömning minskar dock inte längden på kvartilavståndet, utan spridningen mellan skolor är ungefär densamma, 0,9 mot 0,8 enheter bland skolor som inte tillämpar extern bedömning. Ser man på avståndet mellan den skola som är mest generös i sina bedömningar med den som är mest strikt, så är detta lägre bland skolor som tillämpar extern bedömning, 3,0 jämfört med 4,4 enheter. Detta eftersom det förekommer klart avvikande resultat för några av de skolor som inte tillämpar extern bedömning, markerade som cirklar/stjärnor i figuren.



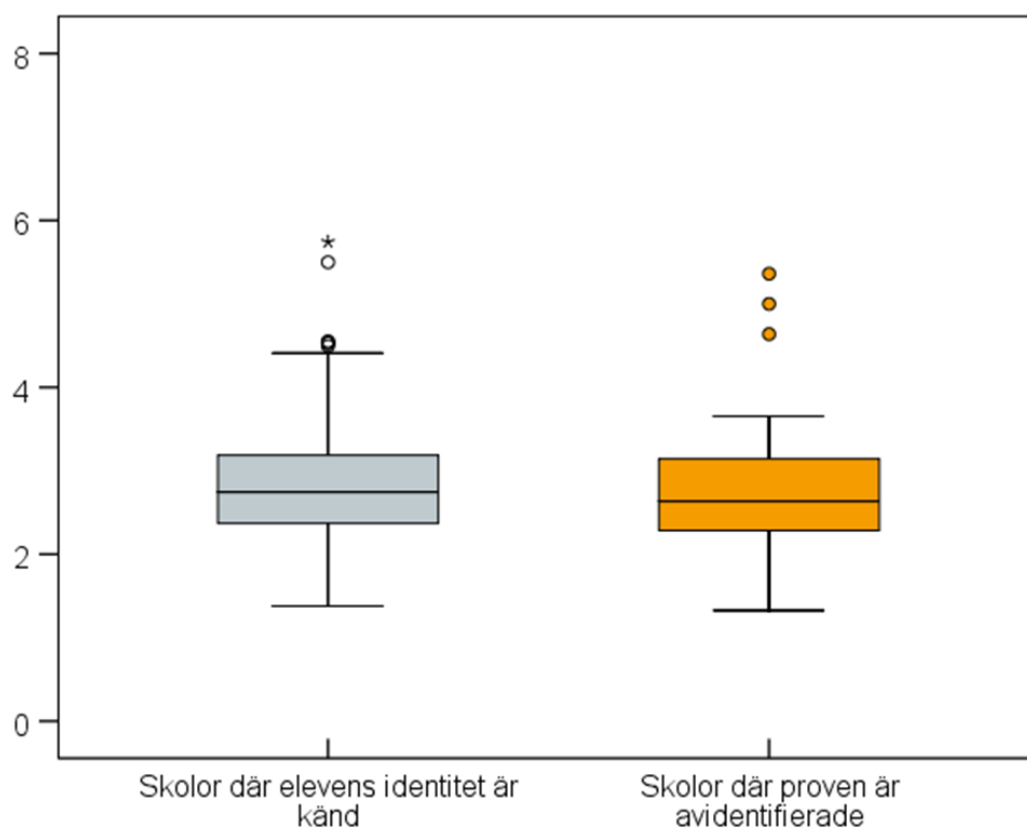
Figur 27: Avvikelse i omdömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många omdömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. I figuren redovisas endast de 205 skolor där under 20 procent av delproven i omdömningen bedöms av annan pedagogisk personal än elevens ordinarie lärare samt de 76 skolor där minst 80 procent av delproven bedöms av annan pedagogisk personal. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Värdet på y-axeln motsvarar den genomsnittliga avvikelsen på skolnivå, mätt i absoluta tal, mellan ursprungsbedömaren och omdömaren. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnena har räknats om till denna skala).

Liksom extern bedömning minskar sambedömning den genomsnittliga skillnaden något. Figur 28 visar att medianen för skolor som tillämpar sambedömning är 2,7 mot 3,1 för skolor som inte sambedömer delprov. Kvartilavståndet för skolor som sambedömer delprov är 0,8 enheter, jämfört med 1,4 enheter för skolor som inte sambedömer delprov. Spridningen mellan skolor är alltså lägre bland skolor som tillämpar sambedömning. Om man däremot ser på skillnaden mellan den skola som är mest generös i bedömningen jämfört med den som är mest strikt, så är skillnaden större bland skolor som tillämpar sambedömning, 4,4 mot 3,8 enheter. Detta förklaras av ett tydligt avvikande resultat för en av skolorna som tillämpar sambedömning.



Figur 28: Avvikelse i omdömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många omdömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. I figuren redovisas endast de 40 skolor där under 20 procent av delproven i omdömningen sambedöms samt de 200 skolor där minst 80 procent av delproven sambedöms. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Värdet på y-axeln motsvarar den genomsnittliga avvikelsen på skolnivå, mätt i absoluta tal, mellan ursprungsbedömaren och omdömaren. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnen har räknats om till denna skala).

Även avidentifiering av delprov leder till en något minskad genomsnittsskillnad mellan skola och omdömaren, där medianen är 2,6 på skolor där delproven är avidentifierade mot 2,8 på skolor där de inte är det (se figur 29). Avidentifiering minskar dock inte längden på kvartilavståndet i någon nämnvärd omfattning, utan de är ungefär lika stora. Skillnaden i bedömning mellan den mest generösa och mest strikta skolan är något större på skolor där delproven inte avidentifieras, då det bland dessa finns några värden som tydligt avviker från medianen. På dessa skolor är skillnaden 4,4 enheter mot 4,0 enheter för skolor som avidentifierar delproven.



Figur 29: Avvikelse i omedömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många omedömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. I figuren redovisas endast de 246 skolor där under 20 procent av delproven i omedömningen är avidentifierade samt de 48 skolor där minst 80 procent av delproven är avidentifierade. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Värdet på y-axeln motsvarar den genomsnittliga avvikelsen på skolnivå, mätt i absoluta tal, mellan ursprungsbedömaren och omedömaren. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnena har räknats om till denna skala).

Analysen på skolnivå visar att extern bedömning, sambedömning, och avidentifiering bidrar alla tre till att skillnaden mellan skola och omedömaren minskar något. Framförallt gäller detta sambedömning som även är den enda metod som tydligt visar på minskad spridning mellan olika skolor.

Bedömningsmetoderna kan dessutom kombineras. Att kombinera olika bedömningsmetoder bidrar till att en lägre andel delprov bedöms generösare av skolorna jämfört med omedömningen (se avsnittet *Kombinationer av bedömningsmetoder minskar tendensen att skolan gör en mer generös bedömning*). Sammantaget visar resultaten att bedömningsprocessen påverkar utfallet i bedömningarna. Skolornas olika tillvägagångssätt vid bedömningen kan därför påverka möjligheterna till en likvärdig bedömning.

# Slutsatser

Alla elever har rätt till en likvärdig och rättvis bedömning och betygssättning. Betyg har en direkt koppling till behörighet och urval för vidare studier och kan därmed ha en avgörande betydelse för elevens framtida förutsättningar att lyckas på en konkurrenspräglad arbetsmarknad. Därför är det viktigt att alla elever, oavsett vilken lärare de har eller vilken skola de går på, i praktiken blir bedömda utifrån samma kriterier.

## Skillnader mellan bedömningarna av de nationella delproven

Resultaten i rapporten visar att mer än hälften av elevlösningarna får olika bedömningar av skolans lärare och ombedömare. Dessutom kan bedömningarna i många fall skilja sig med två betygssteg eller mer. Eftersom endast ett delprov per nationellt prov ombedöms är det inte givet att en uppmätt skillnad för ett enskilt delprov får genomslag på det slutliga provbetyget. Det kan dock räcka med ett betygsstegs skillnad på ett delprov för att det totala provbetyget ska påverkas. Det betyder att de avvikelser som uppmäts i ombedömningen på delprovsnivå är av den storleksordningen som *kan* leda till att det totala provbetyget för en elev påverkas.

## Bedömningsprocessen har betydelse

Resultaten från ombedömningen visar att extern bedömning, sambedömning och i viss mån även avidentifiering leder till att andelen delprov som bedöms generösare på skolan minskar. Effekten verkar öka när flera bedömningsmetoder kombineras. Kombinationen extern bedömning och sambedömning tycks även förbättra samstämmigheten något, det vill säga att skolan och Skolinspektionens ombedömare i högre utsträckning gör samma bedömning.

Vidare visar resultaten på liknande mönster oavsett om det är pojkar eller flickors delprov som bedöms. Tillämpning av extern bedömning, sambedömning och avidentifiering minskar andelen generösare bedömningar på skolan, jämfört med ombedömningen. Vid extern bedömning ser vi, förutom att andelen generösare bedömningarna minskar, att flickor och pojkar i lika stor utsträckning får samma betyg av de olika bedömarna.

Oavsett om delproven skrivits för hand eller på dator leder extern bedömning, sambedömning och avidentifiering till att andelen generösare ursprungsbedömningar minskar, jämfört med ombedömningen. Effekten är ungefär lika stor för hand- respektive datorskrivna delprov.

Resultaten från ombedömningen visar att inslagen av sambedömning, extern bedömning och avidentifiering varierar mellan skolor, ämnen, årskurser och kurser liksom mellan skolor med kommunal respektive enskild huvudman. Den stora variationen i hur delproven bedöms riskerar att i sig minska likvärdigheten, eftersom bedömningsprocessen har en effekt på utfallet av bedömningen. De exempel på bedömningsmetoder som tas upp i den här rapporten kan leda till ökad likvärdighet och stärka rättssäkerheten i bedömningarna, i synnerhet om de sker mer enhetligt mellan landets skolor.



## Svårt att göra samstämmiga bedömningar av delprov med fritextsvar

Även om sambedömning, extern bedömning och till viss del avidentifiering tenderar att minska andelen delprov som bedömts generösare av skolan kvarstår stora skillnader mellan bedömningarna även för delprov där dessa metoder har tillämpats. För delprov som till stor del består av fritextsvar är det svårare att komma fram till samma bedömning och betyg. En slutsats som Skolinspektionen konstaterat i tidigare rapporter<sup>27</sup>. I NO fick de nationella proven en ny utformning och nya bedömningsanvisningar under 2017. I förra årets rapport visade resultaten för dessa delprov att samstämmigheten hade ökat markant. I och med att resultatet står sig för de NO-prov som genomfördes under 2018 tycks utformningen av prov och bedömningsanvisningar, något som skolorna själva inte kan påverka, vara minst lika viktig för att öka samstämmigheten.

## Avslutningsvis

Skolinspektionen har sedan 2009 årligen fått i uppdrag att omdedöma nationella prov och de övergripande slutsatserna av omdedömningsuppdraget har varit desamma över åren. Skolinspektionen vill därför framföra att arbetet med att stödja en likvärdig bedömning och betygssättning kan bedrivas på ett mer verksamt och effektivt sätt genom kvalitetsgranskningar.<sup>28</sup> Mer utförliga redogörelser för skälen till Skolinspektionens förslag finns att läsa i *Ökad kostnadseffektivitet i granskning av likvärdighet i bedömning och betygssättning*<sup>29</sup> och *Budgetunderlag för åren 2017–2019*<sup>30</sup>.

Skolverket har i uppdrag att digitalisera de nationella proven.<sup>31</sup> Om omdedömningsuppdraget till Skolinspektionen ska fortgå under och efter digitaliseringens införande behöver omdedömningsuppdraget anpassas och samverkan inledas med Skolverket under 2020. Digitaliseringen medför nya förutsättningar och därmed behöver Skolinspektionen påbörja ett omställningsarbete. Omställningsarbetet består dels i att utreda IT-tekniska frågor, såsom hur elevlösningar och delprovsresultat ska överföras från den digitala provplattformen till Skolinspektionen, dels i att utreda frågor om sekretess och personuppgiftsbehandling.

Inför år 2022 behöver uppdragsbeskrivningen anpassas i och med att extern bedömning och avidentifiering införs generellt för de digitaliserade proven. För det första behöver kopplingen mellan omdedömningen av nationella prov och Skolinspektionens inspektionsverksamhet på skolenhetsnivå ses över. Om proven bedöms av lärare utanför skolenheten kommer resultatet från omdedömningen inte i samma grad vara relevant att använda i Skolinspektionens tillsyn och granskning på skolenhetsnivå. Resultatet kommer

<sup>27</sup> Statens skolinspektion, *Ombedömning av nationella prov 2017 – Fortsatt stora skillnader* (2018), s. 34. Statens skolinspektion, *Olikheterna är för stora Omrättning av nationella prov i grundskolan och gymnasieskolan, 2013* (2013).

<sup>28</sup> Se exempelvis: Statens Skolinspektion, *Betygsättning på högskoleförberedande program i kursen svenska 3* (2019).

<sup>29</sup> Statens skolinspektion, *PM — Ökad kostnadseffektivitet i granskningen av likvärdighet i bedömning och betygssättning* (2015).

<sup>30</sup> Statens skolinspektion, *Missiv — Statens skolinspektions budgetunderlag 2017–2019 samt regelförteckning* (2016).

<sup>31</sup> Regeringsbeslut, *Uppdrag att digitalisera de nationella proven m.m. U2017/03739/GV* (2017).

inte ge en bild av likvärdigheten i bedömningen bland lärarna på den aktuella skolenheten utan bland de externa bedömarna från andra skolenheter. För det andra behöver behov av och möjligheter att analysera bedömningsprocessen ses över.

# Bilagor

## Bilaga 1

### Användning av resultaten i tillsyn och kvalitetsgranskning

I enlighet med 2018 års regleringsbrev<sup>32</sup>, som hänvisar tillbaka till 2015 års regleringsbrev, ska myndigheten ”... 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.”<sup>33</sup>.

Det övergripande syftet med uppdraget 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 arbetsprocesserna som leder fram till bedömningen och därigenom betygssättningen som granskas i Skolinspektionens inspektion, inte delprovsbetygen i sig. Avviker en skolenhets resultat i hög grad i förhållande till andra skolor kan det vara en indikation på att skolan behöver utveckla sitt arbete med bedömning och betygssättning. Då utreds skolans processer för bedömning och betygssättning antingen i samband med tillsyn eller regelbunden kvalitetsgranskning.

#### Hur avgörs vad som är betydande avvikelser?

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 gränsen för vad som ska anses vara en betydande avvikelse relativt högt.

I ett första steg gäller det att för varje delprov avgöra vad som är en avvikelse mellan skolans bedömning och Skolinspektionens omdömning. För delproven i svenska och engelska innebär en *betydande avvikelse* en skillnad på *två betygssteg eller mer* mellan skolans bedömning och Skolinspektionens omdömning.

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 skolans avvikelse från andra skolor, inte från omdömningen i sig, som bör minimeras. En skolas resultat uppmärksammas om andelen betydande avvikelser, antingen i generösare eller striktare riktning, skiljer sig med mer än tio procentenheter från genomsnittet för de omdömda skolorna för det aktuella delprovet.

<sup>32</sup> Utbildningsdepartementet. *Regleringsbrev för budgetåret 2018 avseende Statens skolinspektion* (2017).

<sup>33</sup> Utbildningsdepartementet. *Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion* (2014).

## Att använda resultaten i tillsyn och kvalitetsgranskning

Skiljer det mer än tio procentenheter mellan andelen delprov där skolenheten gjort en betydligt högre eller lägre bedömning än genomsnittet skickas en signal till den av Skolinspektionens avdelningar som har ansvar för tillsyn och regelbunden kvalitetsgranskning av skolenheten.

Inom tillsyn och regelbunden kvalitetsgranskning utgör resultaten från ombedömningen ett av flera underlag som analyseras inför ett skolbesök. Även Skolverkets statistik, i vilken de jämför provbetyg på det nationella provet med ämnes- eller kursbetyg, utgör en del av inspektörernas underlag. Visar resultaten från ombedömningen att en skolenhet, i större utsträckning än andra skolenheter som medverkat i ombedömningen, gör en betydligt högre eller lägre bedömning är det en risksignal som indikerar att skolenheten kan behöva utveckla sitt arbete med bedömning och betygssättning. En ytterligare risksignal som indikerar att skolan kan behöva utveckla sitt arbete med bedömning och betygssättning är om Skolverkets statistik visar på att skolenheten generellt sätter högre eller lägre ämnes- eller kursbetyg jämfört med det provbetyg eleven fått på det nationella provet, och gjort så under längre tid.

Skolinspektionen genomför också tematiska kvalitetsgranskningar med fokus på bedömning och betygssättning.<sup>34</sup> Urvalet av skolor som ingår i de tematiska kvalitetsgranskningarna dras ibland utifrån resultaten från ombedömningen av nationella prov i kombination med resultatet från Skolverkets jämförelser mellan nationella provbetyg och ämnes- eller kursbetyg.

## Skolenhetsrapporter

Resultatet sammanställs på skolenhetsnivå i skolenhetsrapporter. Inga skolenhetsrapporter görs dock om antalet ombedömda elevlösningar vid en skolenhet understiger 10 stycken eller om antalet ombedömda delprov vid skolenheten uppgår till färre än hälften av det totala antalet delprov som begärts in av Skolinspektionen i det aktuella ämnet från skolenheten den terminen. Resultaten skickas via post till de skolor som deltagit och publiceras på Skolverkets hemsida.

Resultaten från ombedömningen kan användas av rektorer och huvudmän på de skolor som har ingått i ombedömningen i syfte att analysera och utveckla arbetet med bedömning och betygssättning.

---

<sup>34</sup> Se till exempel: Statens skolinspektion, *Rättvis och likvärdig betygssättning i grundskolan?* (2018); Statens skolinspektion, *Uppenbar risk för felaktiga betyg samt Skolinspektionen* (2014); Statens Skolinspektion, *Förstärkt tillsyn av skolors arbete med bedömning och betygssättning* (2014).

## Bilaga 2

# Genomförandet av omdedömningen

### Urval av prov och skolor

2018 års omdedömning innefattar nationella prov som under året genomfördes i årskurs 6, årskurs 9, gymnasiets grundläggande kurser, samt motsvarande kurser inom den kommunala vuxenutbildningen på gymnasial nivå. För den kommunala vuxenutbildningen omdedömdes prov genomförda 2017.

Skolor som ingår i omdedömningen väljs ut genom ett slumpmässigt urval bland skolor vars huvudmän följande år ingår i Skolinspektionens regelbundna tillsyn och granskning. Urvalet av skolor för omdedömning år 2018 gjordes bland skolor vars huvudmän ingår i inspektionsplanen för 2019/2020. Detta för att ett så aktuellt resultat som möjligt ska kunna användas som underlag vid tillsyn och granskning.

När ursprungsbedömningen av proven genomfördes kände skolorna ännu inte till att de skulle ingå i urvalet för omdedömningen av nationella prov. Skolinspektionen bedömer därför att risken för att skolorna i förväg skulle ha kunnat manipulera proven, med avsikt att uppnå ett fördelaktigt resultat, i princip är obefintlig.

För att kunna genomföra omdedömningen vid så många skolor som möjligt omdedöms endast ett eller två ämnen per skola. En slumpvis tilldelning av ämnen görs bland de skolor som valts ut till omdedömningen. Därefter görs ett slumpvist urval bland de elever som genomfört delprovet den aktuella terminen. Urvalet görs tillräckligt stort för att resultaten ska vara representativa på skolenhetsnivå. Det maximala antalet delprov som omdedöms har begränsats till i regel 100 för grundskolor och 200 för gymnasieskolor. Om fler delprov 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 omdedöma vårterminsversionen, då denna provversion skrivs av de flesta av landets elever och skolor och därmed ger ett bredare provunderlag att utgå ifrån.

### Hela provet omdedöms 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 att samla in och fördela ut till omdedömare. Tidigare erfarenheter från omdedömningen har även visat att större avvikelser är sällsynta för delprov med lite utrymme för subjektivitet, till exempel nationella delprov som genomförs i matematik.

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

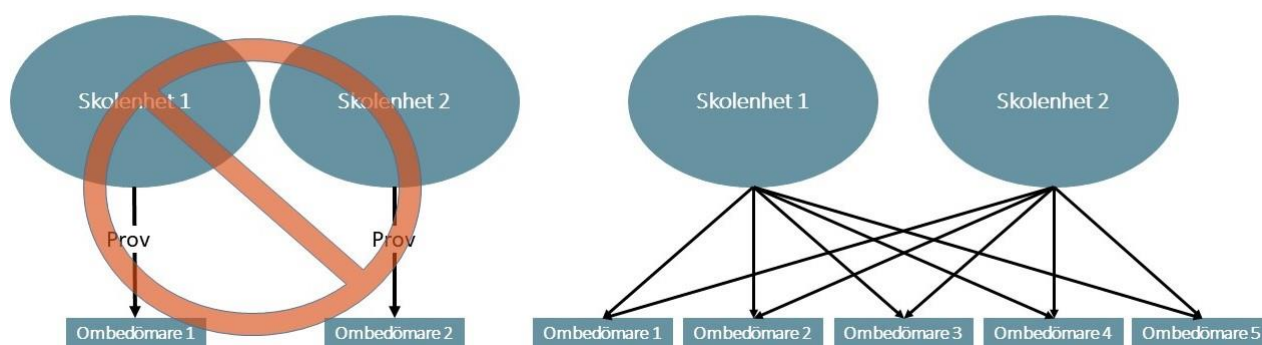
## Urvalets sammansättning

Andelen omdömda delprov från skolor med enskild huvudman som ingår i urvalet är högre än andelen skolor med enskild huvudman i riket vid tidpunkten för urvalet. Andelen skolor med enskild huvudman som ingått i omdömningen är 55 procent för grundskolor och 54 procent för gymnasieskolor. Vid tiden då urvalet gjordes uppgick andelen grundskolor respektive gymnasieskolor med enskild huvudman till 18 procent respektive 33 procent av det totala antalet grund- respektive gymnasieskolor i landet. Anledningen till att fristående skolor är överrepresenterade i urvalet till omdömningen är att urvalet är kopplat till Skolinspektionens tillsyns- och granskningsprocesser.

## Fördelning av delprov mellan Skolinspektionens omdömare

Skolor som ingår i omdömningen ombeds skicka in kopior av bedömda delprov till Skolinspektionen. Delproven maskeras och oidentifieras därefter på personuppgifter, bedömningsanteckningar och andra uppgifter från ursprungsbedömningen. Därefter bedöms delproven på nytt av ämnesbehöriga, yrkesverksamma och erfarna lärare som anlätts på uppdrag av Skolinspektionen. Ursprungsbedömningen jämförs sedan med omdömningen, och resultatet sammanställs i skolenhetsvisa rapporter.

Skolinspektionens lärare omdömer delprov enskilt. Sambedömning tillämpas inte. För att minimera risken att en enskild omdömares bedömningar får stort genomslag i de sammantagna resultaten sprids elevlösningar för varje skola ut bland flera olika omdömare. Jämförelsen av resultatet mellan skolor blir på så vis så rättvist som möjligt. Beroende på ämne, kurs och årskurs fördelas delproven mellan 6 och 21 olika omdömare.



Figur 30: Schematisk bild över fördelningen av delprov mellan omdömare. Delprov 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 delprov i svenska och NO-ämnen, där det skiljer mer än två betygssteg mellan bedömning och omdömning, går delprovet vidare till en andra omdömning. Det slutgiltiga betyget från omdömningen blir då resultatet från den av de två omdömningarna som ligger närmast ursprungsbedömningen. Syftet med den andra omdömningen är att kvalitetssäkra bedömningen av delprov med betydande avvikelser genom att låta ytterligare en lärare bedöma delproven.

## Hand- respektive datorskrivna prov

I samband med att de omdedömande lärarna rapporterar in delprovsbetyg har de även fått svara på frågan huruvida det aktuella delprovet var skrivet för hand eller på dator (eller annan digital enhet). I svenska årskurs 6 genomförs två delprov (C1 och C2) och de omdedömande lärarna rapporterar in det sammanvägda delprovsbetyget. För detta delprov har det inte varit möjligt att sammanställa lärarnas svar på frågan hur eleverna skrev delprov eftersom det ena delprovet kan ha skrivits för hand medan det andra genomförts på dator.

I de fall ett delprov gått till den andra omdedömningen har även den omdedömande läraren fått svara på frågan om delprovet är skrivet för hand eller dator. Om svaren från de två omdedömare inte stämmer överens för ett specifikt delprov exkluderas det från analyser som är uppdelade på hand- respektive datorskrivna delprov.

## Omfattning och bortfall

Totalt 27 772 delprov från 393 skolor har omdedömts i 2018 års omgång. Omdedömningen genomfördes av totalt 127 lärare som anlätts av Skolinspektionen. Fördelningen över ämne och delprov framgår av tabell 2.

Tabell 2: Externt och internt bortfall av delprov. Antal begärda delprov är det totala antalet delprov som begärdes in från skolorna. Externt bortfall har med skillnaden mellan antal begärda delprov och antal inkomna delprov att göra, medan det interna bortfallet beskriver skillnaden mellan antal inkomna delprov och antal komplett omdömda delprov. Totalt bortfall, slutligen, handlar om skillnaden mellan antal begärda delprov och antal komplett omdömda delprov.

<i>Ämne och delprov:</i>	Antal begärda prov	Från antal skolor	Antal inkomna prov	Från antal skolor	Externt bortfall (prov)	Antal kompletta omdömda prov	Från antal skolor	Internt bortfall (prov)	Totalt bortfall	Antal omdömdare
Engelska åk 6 delprov C	3 348	93	3 237	90	3 %	3 188	90	2 %	5 %	11
Engelska åk 9 delprov C	3 166	66	3 034	65	4 %	2 995	65	1 %	5 %	18
Engelska Gy delprov C	4 030	70	3 877	70	4 %	3 819	70	1 %	5 %	17
Engelska Vux delprov C	637	24	619	24	3 %	596	24	4 %	6 %	6
Svenska åk 6 delprov C1+C2	2 687	87	2 576	84	4 %	2 508	84	3 %	7 %	14
Svenska åk 9 delprov C	2 988	62	2 947	62	1 %	2 865	62	3 %	4 %	21
Svenska Gy delprov C	4 351	71	4 185	71	4 %	4 120	71	2 %	5 %	21
Svenska Vux delprov C	1 628	44	1 466	44	10 %	1 368	44	7 %	16 %	8
Biologi åk 9 delprov A2	2 231	45	2 128	44	5 %	2 094	44	2 %	6 %	19
Fysik åk 9 delprov A2	2 055	42	2 009	42	2 %	1 950	42	3 %	5 %	19
Kemi åk 9 delprov A2	2 343	47	2 308	47	1 %	2 269	47	2 %	3 %	20
<b>Totalt:</b>	<b>29 464</b>	<b>397</b>	<b>28 386</b>	<b>393</b>	<b>4 %</b>	<b>27 772</b>	<b>393</b>	<b>2 %</b>	<b>6 %</b>	<b>127</b>

Det externa bortfallet utgörs av delprov 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 kommer elever som inte genomfört delprovet på grund av sjukdom eller frånvaro av annan orsak. Det händer även att skolor inte kan återfinna delproven eftersom de inte har arkiverats enligt gällande regler.

Internt bortfall utgörs av delprov som inkommit från skolorna, men som av olika orsaker inte kunnat analyseras. De vanligaste orsakerna är att delprovskopiorna håller för dålig kvalitet (kopieringen har inte tillräcklig svärta), att sidor saknas eller att resultatet inte rapporterats in på korrekt sätt.



## Vad beror bortfallet på?

Årets sammanvägda bortfall uppmäts till sex procent och är mindre än vad det varit för tidigare omgångar. I tidigare omgångar har bortfallet varit ganska högt för delprov där skolorna själva fyller i belägg (poäng). I NO-ämnen fyller skolorna själva i antal belägg för varje delprov de ska skicka in. I årets omgång har instruktionerna om hur beläggen fylls i förtydligats och kan därmed förklara en del av minskningen.

Det sammanvägda bortfallet varierar från delprov till delprov, mellan 3 till 16 procent. Bortfallet är störst för ämnet svenska inom den kommunala vuxenutbildningen men även här har en minskning skett jämfört med tidigare omgångar. Orsaken till varför den kommunala vuxenutbildningen sticker ut i bortfall kan vara att det är frivilligt för de kommunala huvudmännen att skicka in kopior på bedömda nationella delprov.

## Frågor om bedömningen på skolan

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

I tabell 3 redovisas antal svar per fråga. Bortfallet utgörs av skolor som inkom med delprovkopior som gick att omdöma, men som inte svarat på frågorna om bedömningens genomförande.

För delprov genomförda inom den kommunala vuxenutbildningen ställdes inga frågor om bedömningsprocessen och de är därför inte inkluderade i analysen.

Tabell 3: 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?		
	Antal svar	Från antal skolor	Bortfall*	Antal svar	Från antal skolor	Bortfall*	Antal svar	Från antal skolor	Bortfall*
<b>Ämne och delprov:</b>									
Engelska åk 6 delprov C	2 555	75	20 %	3 061	89	4 %	3 082	89	3 %
Engelska åk 9 delprov C	2 766	61	8 %	2 959	64	1 %	2 826	62	6 %
Engelska Gy delprov C	3 569	66	7 %	3 581	67	6 %	3 583	67	6 %
Svenska åk 6 delprov C1+C2	1 751	64	30 %	2 380	82	5 %	2 427	83	3 %
Svenska åk 9 delprov C	2 041	50	29 %	2 618	59	9 %	2 670	60	7 %
Svenska Gy delprov C	3 778	67	8 %	3 995	70	3 %	4 000	70	3 %
Biologi åk 9 delprov A2	1 539	37	27 %	2 056	44	2 %	1 960	43	6 %
Fysik åk 9 delprov A2	1 732	37	11 %	1 910	42	2 %	1 891	42	3 %
Kemi åk 9 delprov A2	1 741	41	23 %	1 956	44	14 %	2 048	46	10 %

\* Andel kompletta ombedömda delprov där svar på frågan saknas.

## Bilaga 3

# Ett harmoniserat mått för att mäta skillnader i bedömning av nationella delprov

Ett nationellt prov består av flera *delprov*. Enligt vad som beskrivs i avsnittet *Urval av prov och skolor* i bilaga 2 omdöms inte alla delprov, utan endast de skriftliga delprov där risken för avvikelser mellan bedömare kan antas vara som störst.

För delproven i svenska och engelska sätts enligt betygsskalan A–F, men för NO-ämnena erhålls i stället ett antal *belägg* (poäng) på E-, C- respektive A-nivå och endast det samlade provbetyget anges på skalan A–F. För att kunna göra en översiktlig sammanställning av omdömningsresultat behöver bedömningarna konverteras till en gemensam skala.

Skalan har konstruerats på så sätt att bedömningarna av samtliga delprov kan uttryckas på skalan 0–20, där 0 representerar lägsta möjliga bedömning och 20 representerar högsta möjliga bedömning. Skillnaden mellan ursprungsbedömning och omdöming kan då uttryckas på skalan -20 till 20. Ett positivt värde indikerar att skolan gjort en generösare bedömning än omdömingaren för en viss elevlösning. Ett negativt värde indikerar motsatsen.

## Beräkning av det harmoniserade resultatmättet

För delprov där bokstavsbedömning (A–F) sätts är konverteringen mellan den faktiska bedömningen och den harmoniserade skalan följande: A=20, B=16, C=12, D=8, E=4, F=0. En genomsnittsskillnad på 2 mellan ursprungsbedömingarens och omdömingarens bedömning för ett delprov betyder att det i genomsnitt skiljer ”ett halvt betygssteg” mellan bedömningarna, där ursprungsbedömingaren sätter ett högre betyg.

För delprov där delprovsbetyg inte sätts omvandlas i stället antalet belägg på E-, C- respektive A-nivå till skalan 0–20 enligt följande modell:

$$((E*10 + C*15 + A*20) / (\max(E)*10 + \max(C)*15 + \max(A)*20)) * 20.$$

E, C respektive A står för antalet belägg som respektive bedömingare angivit för delprovet. Max (E), max (C) respektive max (A) står för det maximala antal belägg som kan uppnås på det aktuella delprovet.

Beräkningen möjliggör att bedömningar som uttrycks i belägg 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 få på ett delprov är lågt väger varje enskilt belägg i framräkningen av det harmoniserade mättet desto tyngre.

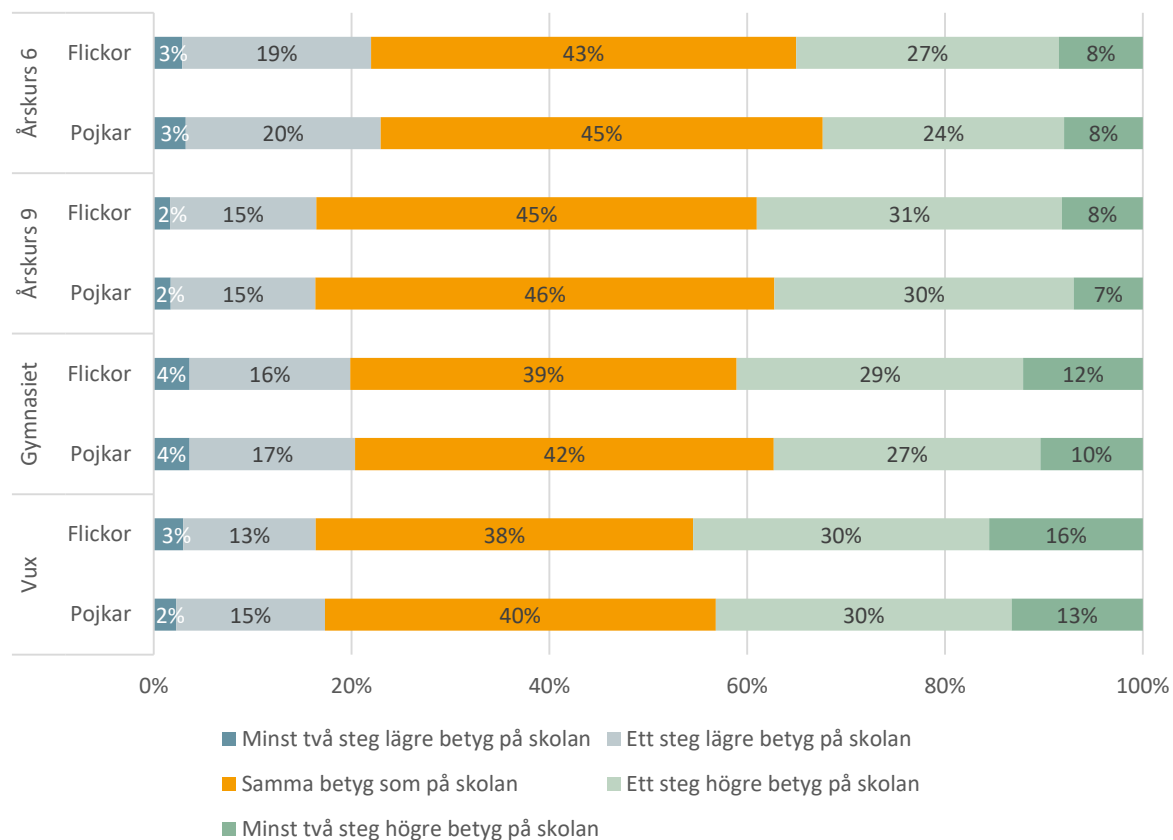
En snabb approximation som underlättar den intuitiva förståelsen är att 2 på den harmoniserade skalan motsvarar tio procent av antalet möjliga belägg för det aktuella delprovet, 10 motsvarar 50 procent av antal möjliga belägg och så vidare. Approximationen tar dock inte hänsyn till att A-, C- och E-belägg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

Det är viktigt att komma ihåg att det harmoniserade resultatmättet som 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 mel-

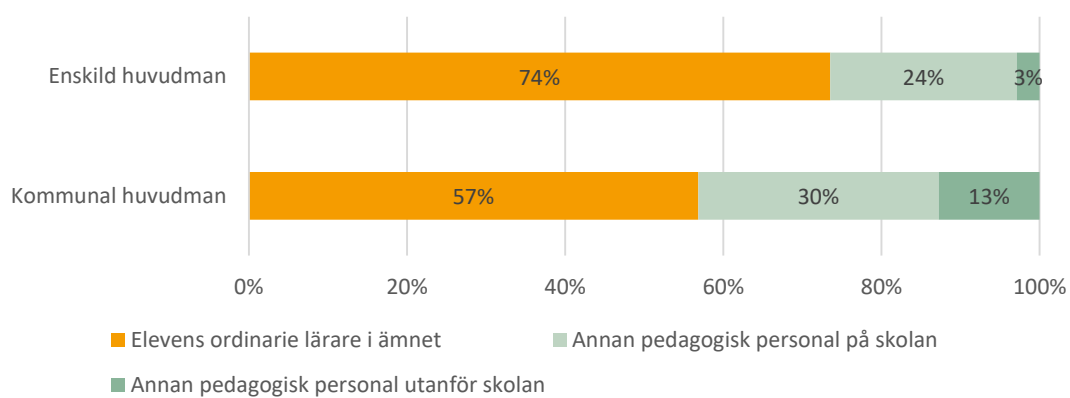
lan detta mått och ett faktiskt betyg på skalan A–F, eftersom reglerna för hur belägg summeras till ett samlat provbetyg ser olika ut för olika prov och dessutom involverar summering av belägg från flera olika delprov.

## Bilaga 4

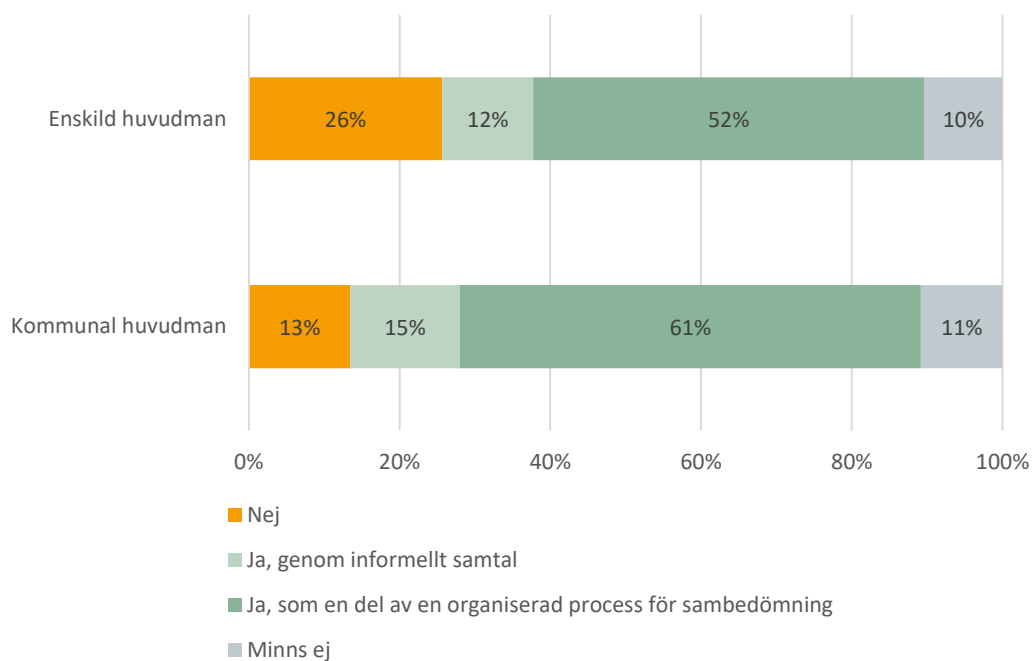
### Övriga figurer



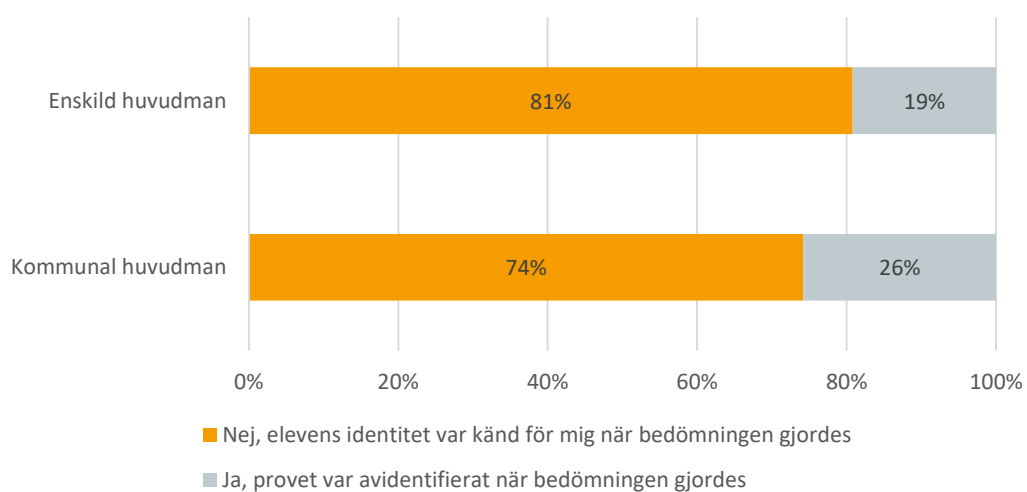
Figur 31: Resultat i omedbömningen, efter årskurs/kurs samt elevens kön. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens omedbörare (belägg i NO-ämnen har räknats om till skalan A–F).



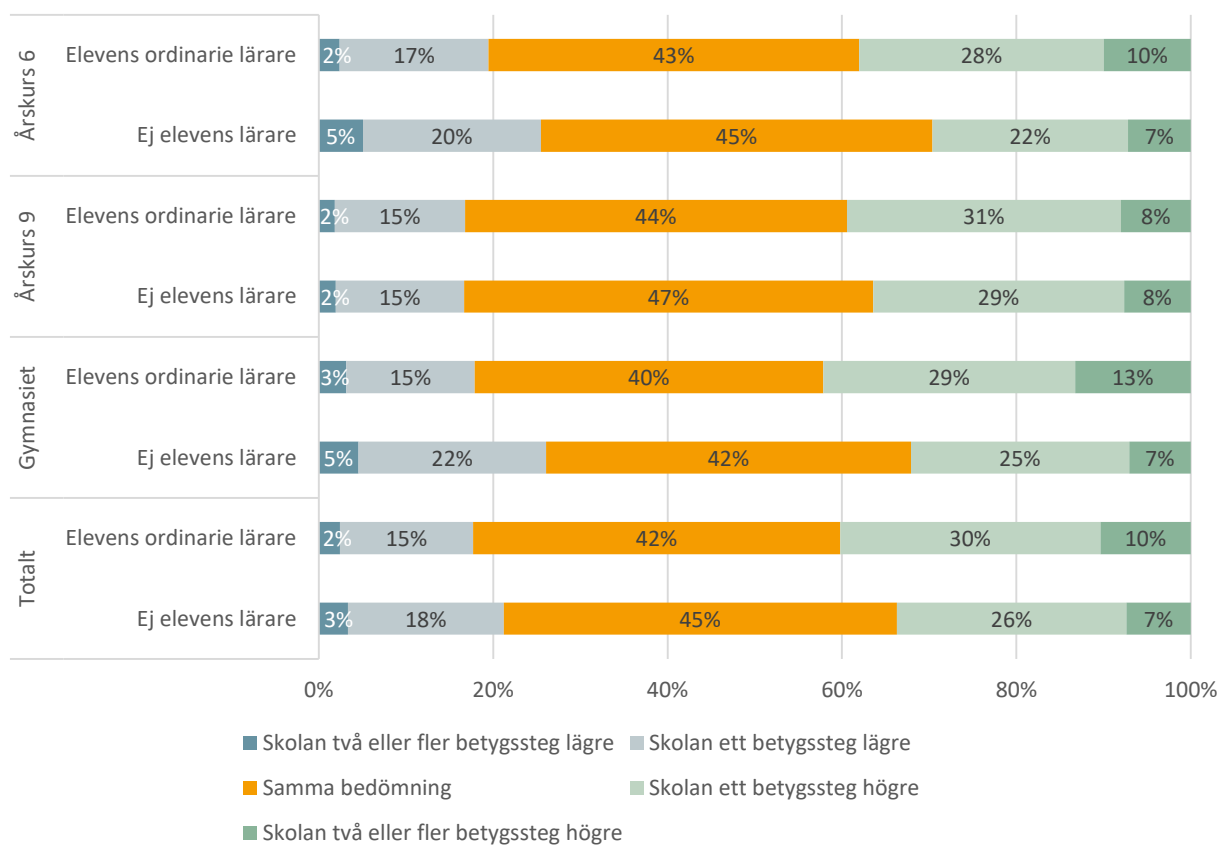
Figur 32: Andel delprov i omedbömningen som ursprungligen bedömts av elevens ordinarie lärare respektive annan pedagogisk personal, efter typ av huvudman.



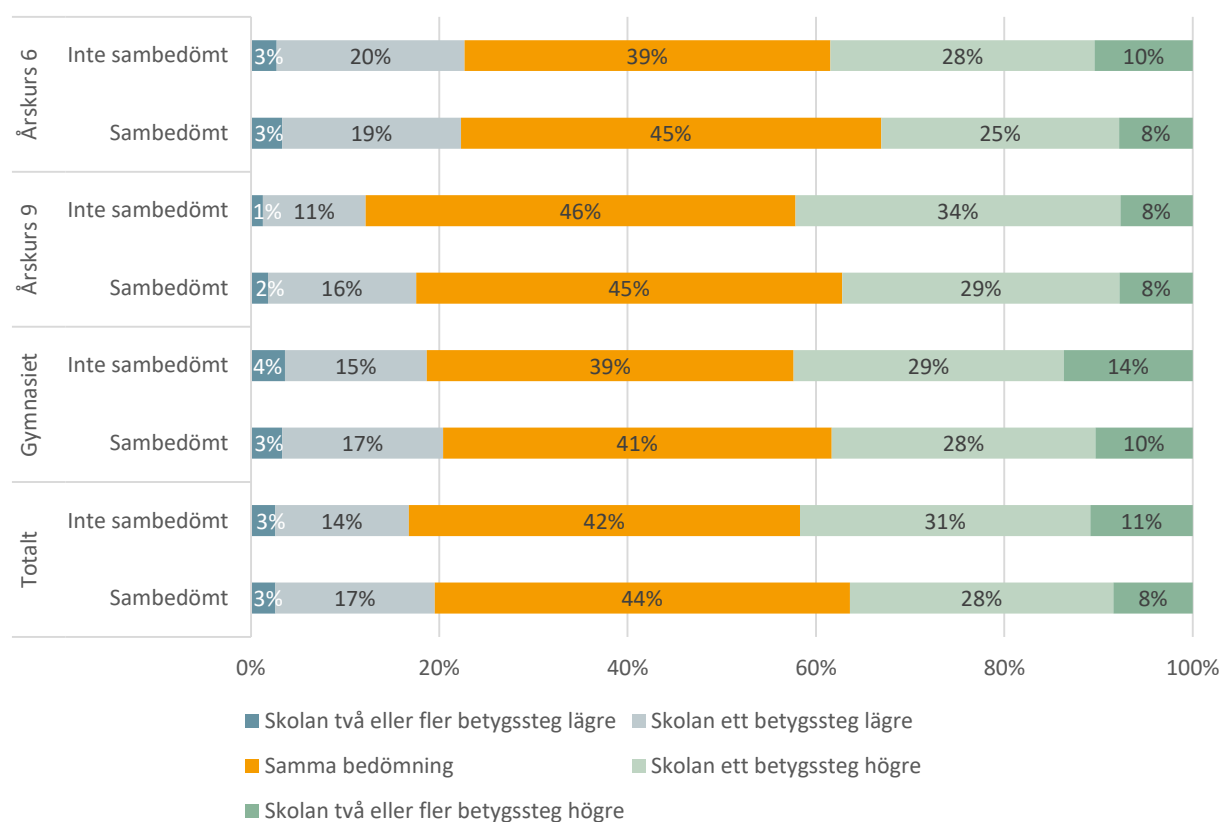
Figur 33: Andel delprov i ombedömningen som sambedömts vid den ursprungliga bedömningen på skolan, efter typ av huvudman.



Figur 34: Andel delprov i ombedömningen som var avidentifierade vid den ursprungliga bedömningen på skolan, efter typ av huvudman.

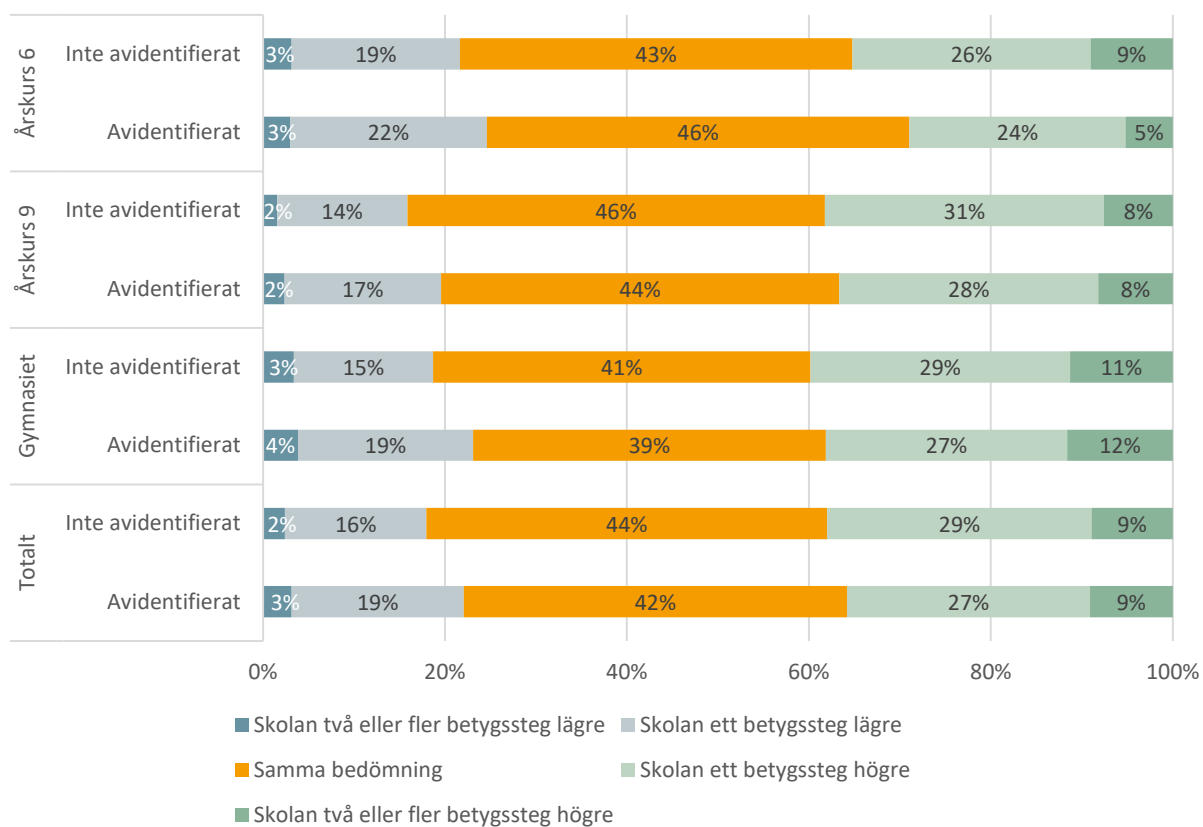


Figur 35. Resultat i ombedömningen för delprov som ursprungligen bedömts av elevers ordinarie lärare respektive av annan pedagogisk personal, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.



Figur 36. Resultat i ombedömningen för delprov som sambedömts respektive bedömts enskilt, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.





Figur 37. Resultat i ombedömningen för delprov som var avidentifierade vid ursprungsbedömningen respektive för delprov där elevens identitet var känd, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare ger lägre, högre eller samma betyg som Skolinspektionens ombedömare.

# Referenslista

Förordning om läroplan för grundskolan, förskoleklassen och fritidshemmet 2011, SKOLFS 2010:37.

Förordning om läroplan för gymnasieskolan 2011, SKOLFS 2011:144.

Regeringsbeslut, *Uppdrag att digitalisera de nationella proven m.m.* Dnr U2017/03739/GV (2017).

SFS 2010:800. *Skollagen*.

SOU 2016:25. Betänkande — *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning*.

Statens skolinspektion (2013), *Olikheterna är för stora - Omrättning av nationella prov i grundskolan och gymnasieskolan*. Dnr 2011:6544

Statens skolinspektion (2014). PM — *Förstärkt tillsyn av skolors arbete med bedömning och betygssättning*. Dnr 400-2013:200.

Statens skolinspektion (2014). *Uppenbar risk för felaktiga betyg*. Dnr 400-2013:200.

Statens skolinspektion (2015). PM — *Ökad kostnadseffektivitet i granskningen av likvärdighet i bedömning och betygssättning*. Dnr 2015:514.

Statens skolinspektion (2016). *Missiv — Statens skolinspektions budgetunderlag 2017–2019 samt regelförteckning*. Dnr 11-2016:1582.

Statens skolinspektion (2018), *Ombedömning av nationella prov 2017 – Fortsatt stora skillnader*. Dnr 2017:342

Statens skolinspektion (2018). *Rättvis och likvärdig betygssättning i grundskolan?* Dnr 2016:11436.

Statens Skolinspektion (2019). *Betygssättning på högskoleförberedande program i kursen svenska 3*. Dnr 400-2017:10213

Statens skolverk, *Sambedömning i skolan – exempel och forskning* (Stockholm: Statens skolverk, 2013).

Utbildningsdepartementet, *Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion*. Dnr U2014/7521/SAM (delvis) U2014/7535/GV (2014).

Utbildningsdepartementet, *Regleringsbrev för budgetåret 2018 avseende Statens Skolinspektion*. Dnr U2017/05023/BS (delvis), U2017/05038/GV (2017).