



# Ombedömning av nationella prov 2017 — Fortsatt stora skillnader

Redovisning av regeringsuppdrag Dnr U2014/7535/GV, U2017/05038/GV

# Innehållsförteckning

<b>Förord</b> .....	<b>4</b>
<b>Sammanfattning</b> .....	<b>5</b>
Bedömningarna skiljer sig åt för majoriteten av elevlösningarna .....	5
Små förändringar i resultatet jämfört med tidigare år, med undantag av delproven i NO .....	5
Störst samstämmighet kring betyget F och lägst kring betyget B .....	5
Stora variationer i hur delproven bedöms .....	6
Bedömningsprocessen har betydelse för likvärdigheten i bedömningen .....	6
Svårt att göra likvärdiga bedömningar av delprov med fritextsvar .....	6
<b>Uppdrag och genomförande</b> .....	<b>7</b>
Bakgrund .....	7
Genomförande .....	7
Ombedömda delprov .....	8
Urvalet följer tillsynscykeln .....	8
Skolinspektionens användning av resultaten .....	8
Denna rapport .....	8
<b>Resultat</b> .....	<b>10</b>
Hur mycket skiljer sig bedömningarna av samma elevlösning? .....	10
Vanligt att skolan gjort en mer generös bedömning än ombedömaren i svenska ..	10
Generösare bedömningar vanligare i högre årskurser och kurser i engelska .....	11
Resultaten för samma årskurs och kurs varierar mellan språkämnen .....	11
Större samstämmighet i NO-ämnena än i språkämnen .....	12
Små skillnader mellan fristående och kommunala skolor .....	13
Överlag små skillnader jämfört med tidigare år .....	13
Stor variation mellan olika skolor .....	16
Störst samstämmighet kring betyget F och lägst kring betyget B .....	18
Bristande likvärdighet i bedömningarna av vissa delprov .....	19
Hur påverkas likvärdigheten av bedömningsprocessen? .....	20
Stora variationer i hur delproven bedöms .....	20
Stora variationer i hur bedömningsmetoderna kombineras .....	23
Bedömningsprocessen har betydelse för likvärdigheten .....	24
Kombinationer av bedömningsmetoder minskar tendensen att skolan gör en mer generös bedömning .....	27
Sambedömning och extern bedömning minskar variationen mellan skolor .....	28

Delprov som ombedömts två gånger .....	32
<b>Slutsatser .....</b>	<b>34</b>
Bristande likvärdighet i bedömningarna av de nationella proven .....	34
Bedömningsprocessen har betydelse för likvärdigheten .....	34
Svårt att göra likvärdiga bedömningar av delprov med fritextsvar .....	34
Avslutningsvis .....	35
<b>Bilagor.....</b>	<b>36</b>
Bilaga 1 Användning av resultaten i tillsyn och kvalitetsgranskning .....	36
Hur avgörs vad som är betydande avvikelser? .....	36
Att använda resultaten i regelbunden tillsyn, regelbunden kvalitetsgranskning och tematisk kvalitetsgranskning .....	37
Skolenhetsrapporter och huvudmannasammanställningar .....	37
Bilaga 2 Genomförandet av ombedömningen .....	38
Urval av prov och skolor.....	38
Ämnen, delprov och årskurser som ingår i ombedömningen .....	38
Hela provet ombedöms inte, endast längre skrivuppgifter.....	38
Urvalets sammansättning .....	39
Fördelning av prov mellan ombedömare .....	39
Omfattning och bortfall.....	40
Vad beror bortfallet på?.....	41
Frågor om bedömningen på skolan .....	41
Bilaga 3 Ett harmoniserat mått för att mäta skillnader i bedömning .....	42
Beräkning av det harmoniserade resultatmåttet .....	42
<b>Referenslista .....</b>	<b>44</b>

# Förord

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

Skolinspektionen redovisar här till regeringen resultatet av nionde årets centrala rättning. Denna rapport avser nationella prov i grund- och gymnasieskolan som genomfördes under 2017 och nationella prov inom den kommunala vuxenutbildningen på gymnasial nivå som genomfördes under 2016.

Vid Skolinspektionen har följande personer arbetat med att genomföra uppdraget. Projektledare: Fanny Hemmingsson. Utredare/statistiker: Jenny Börnin, Mattias Friström, Peter Gustafsson, Marcus Hugosson, Tina Sajjadi. Avdelningschef: Anna Sellin. Enhetschef: Emma Uhman.

Tommy Lagergren

Biträdande generaldirektör

Anna Sellin

Avdelningschef

# Sammanfattning

Skolinspektionen granskar bedömning och betygssättning i skolor på många olika sätt, till exempel i tillsyn och kvalitetsgranskning. Ombedömning av nationella prov är ett uppdrag som Skolinspektionen årligen fått av regeringen sedan 2009. Nationella prov är en av flera delar i betygssättningen på skolorna. Att jämföra skolornas bedömningar med externa lärares bedömningar är ett sätt att mäta likvärdigheten i betygssättningen av nationella prov. Syftet är att stödja en likvärdig bedömning och betygssättning av proven över landet genom att undersöka bedömningarnas likvärdighet samt analysera vilka faktorer som främjar likvärdig bedömning av de nationella proven.

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 årets 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 andelen avvikelser mellan de olika betygsstegen, analyser av skillnader mellan fristående och kommunala skolor samt en fördjupad analys av hur bedömningsprocessen påverkar likvärdigheten. Därtill görs en jämförelse av ursprungsbedömning, första ombedömning och andra ombedömning för delprov som har ombedömts två gånger.

## Bedömningarna skiljer sig åt för majoriteten av elevlösningarna

Resultatet visar att mer än hälften av elevlösningarna får olika bedömning av skolans lärare och Skolinspektionens ombedömare. För delproven i svenska och engelska får 60 procent av elevlösningarna en annan bedömning av ombedömaren. För NO-delproven är motsvarande andel 47 procent. Inom samtliga ämnen, årskurser och kurser är det vanligare att skolans lärare gör en mer generös bedömning än ombedömaren. Andelen delprov där skolan gett ett högre betyg än ombedömaren är något högre bland fristående skolor (40 %) än bland kommunala (35 %).

## Små förändringar i resultatet jämfört med tidigare år, med undantag av delproven i NO

Resultaten från den nionde ombedömningen samstämmer på det stora hela med tidigare jämförbara ombedömningar, det vill säga omgång sex, sju och åtta. Ett undantag är delproven i NO-ämnena, där andelen avvikelser mellan ursprungsbedömning och ombedömning är markant lägre i år jämfört med tidigare år. Minskningen kan troligtvis förklaras av att delproven i NO har ändrat utformning och fått nya bedömningsanvisningar detta år.

## Störst samstämmighet kring betyget F och lägst kring betyget B

En uppdelning per betygssteg visar att skolans lärare och ombedömarna generellt är mest eniga kring elevlösningar som ursprungligen fått delprovsbetyget F, det vill säga de elevlösningar som inte uppfyller kunskapskraven för godkänt betyg. För delprov där skolans

lärare har satt ett F har ombedömaren gjort en annan bedömning i 32 procent av fallen. Skillnaden är störst för de delprov där ursprungsläraren har gett betyget B, 74 procent.

## Stora variationer i hur delproven bedöms

Vid bedömningen av de nationella delproven tillämpade skolorna olika processer. Resultaten visar att tre av fyra delprov sambedömdes. För en tredjedel av delproven tillämpades extern bedömning, det vill säga att delprovet bedömdes av någon annan än elevens ordinarie lärare. En fjärdedel av delproven var avidentifierade vid bedömningen. Olika kombinationer av de tre bedömningsmetoderna förekom också.

## Bedömningsprocessen har betydelse för likvärdigheten i bedömningen

Bedömningsprocessen påverkar bedömningen. Resultaten visar att för delprov som har sambedömts är den genomsnittliga skillnaden mellan ursprungsbedömning och ombedömning lägre än för delprov som inte har sambedömts. Den genomsnittliga skillnaden minskar till viss del även vid extern bedömning och avidentifiering, men effekten är inte lika tydlig. Däremot visar resultaten att alla dessa bedömningsmetoder har en återhållande effekt på bedömningen. Tendensen att skolan gör en generösare bedömning än ombedömarna minskar när någon av dessa bedömningsmetoder har använts. Denna effekt tycks öka när flera bedömningsmetoder kombineras.

Resultaten visar att det finns en stor variation mellan olika skolor i resultatet av ombedömningen. Ungefär tre fjärdedelar av skolorna har satt ett genomsnittligt betyg som är högre än Skolinspektionens, medan resten har satt ett som är lägre. Bedömningsprocessen påverkar även variationen mellan olika skolors avvikelsegrad. Sambedömning och extern bedömning minskar variationen. I viss mån visar även avidentifiering på liknande tendenser.

## Svårt att göra likvärdiga bedömningar av delprov med fritextsvar

För de delprov i svenska där det skilde minst två betygssteg mellan ursprungsbedömning och ombedömning har proven ombedömts ytterligare en gång av en annan ombedömmare. Resultatet visar på stora skillnader mellan de två av varandra oberoende bedömningarna som gjorts av Skolinspektionens ombedömare. Det faktum att även de båda ombedömarna, som inte har en relation till eleven eller skolan, gör olika bedömningar av samma elevlösning indikerar att prov som till stor del består av fritextsvar är svåra att bedöma på ett likvärdigt sätt. Det är rimligt att anta att det är de mest svårbedömda elevlösningarna som får en andra ombedömning, resultatet kan därför inte generaliseras till alla elevlösningar som ingår i ombedömningen.

# Uppdrag och genomförande

Sedan 2009 har Skolinspektionen årligen i sitt regleringsbrev fått i uppdrag av regeringen att omdedöma nationella prov. Syftet med omdedömningen av nationella prov är att stödja en likvärdig bedömning och betygssättning över hela landet.

## Bakgrund

Bakgrunden till att Skolinspektionen fått i uppdrag att omdedöma nationella prov är att det anses finnas en risk för att de bedömningar som görs inte är likvärdiga mellan olika lärare och skolor. Risken för skillnader mellan olika bedömningar är som störst för delprov där eleven ska ange fritextsvar på frågor som inte har ett givet rätt eller fel svar. Resultaten från tidigare års omdedömningar har visat att det för sådana delprov är vanligt att det skiljer ett betygssteg eller mer mellan två olika lärares bedömningar av samma elevlösning.

Med den nya lagstiftning som trädde i kraft 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>1</sup> Att resultaten från de nationella proven ska beaktas särskilt betyder att proven ska ha en större betydelse än andra enskilda underlag vid lärarens allsidiga utvärdering av elevens kunskaper vid betygssättningen.<sup>2</sup> Den nya lagstiftningen innebär därmed att de nationella provens betydelse för elevers ämnes- och kursbetyg stärkts jämfört med tidigare formulering. Samtidigt lämnar dock den nya formuleringen rum för läraren att sätta ett ämnes- eller kursbetyg som avviker från elevens resultat på de nationella proven.<sup>3</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 som anses motivera betyget A på en skola ger C på en annan, få konsekvenser för enskilda elever. Det är därför viktigt och angeläget att de nationella proven bedöms likvärdigt.

## Genomförande

För att genomföra uppdraget samlar Skolinspektionen in kopior på elevlösningar från de nationella proven 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. De ursprungliga bedömningarna av elevsvaren jämförs sedan med omdedömarens bedömningar.

---

<sup>1</sup> SFS 2010:800.

<sup>2</sup> Skolverket. Skolverkets allmänna råd med kommentarer - *Betyg och betygssättning*. s. 37; Skolverket. [<https://www.skolverket.se/undervisning/gymnasieskolan/nationella-prov-i-gymnasieskolan/genomfora-och-bedoma-prov-i-gymnasieskolan#Provresultatetsbetydelseforbetyget>]

<sup>3</sup> Skolverket. Skolverkets allmänna råd med kommentarer - *Betyg och betygssättning*. s. 37 ff.

## Ombedömda delprov

I 2017 års ombedömning ingår delprov i svenska och engelska för grundskolans årskurs 6 och 9, samt prov i biologi, fysik och kemi (NO) för årskurs 9. Vidare innefattar ombedömningen delprov i svenska och engelska i gymnasieskolans grundläggande kurser och motsvarande kurser inom den kommunala vuxenutbildningen på gymnasial nivå. Totalt har 23 458 delprov från 371 skolor ombedömts.

I varje ämne ombedöms endast *ett delprov*.<sup>4</sup> De delprov som ingår är de där risken för skillnader i bedömning mellan olika lärare bedömts som störst. I praktiken betyder 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.

## Urvalet följer tillsynscykeln

Urvalet av skolor för ombedömningen följer från och med 2014 Skolinspektionens tillsynscykel. Urvalet är således 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 10 procent. Bortfallet varierar dock mellan de olika delproven. Högst är bortfallet för vuxenutbildningen. Bortfallet beskrivs mer ingående i bilaga 2.

## Skolinspektionens användning av resultaten

Att Skolinspektionens ombedömare satt ett annat betyg på en elevlösning än skolans lärare betyder inte att ursprungsläraren gjort en felaktig bedömning. En stor variation i bedömningen av likvärdiga elevlösningar mellan *olika skolor* är dock problematisk, eftersom ett av provens huvudsyften är att stödja en likvärdig och rättvis bedömning och betygssättning. Därför är det av stor vikt att skolorna får ett underlag som ger dem möjlighet att jämföra sina resultat från ombedömningen av de nationella proven med andra skolors resultat, och att Skolinspektionen följer upp dessa indikatorer i sin tillsynsverksamhet. En utförlig beskrivning av hur resultaten från ombedömningen används i myndighetens regelbundna tillsyn, regelbundna kvalitetsgranskning och tematiska kvalitetsgranskning finns i bilaga 1.

## Denna rapport

Med denna rapport redovisar Skolinspektionen resultaten från den nionde omgången av ombedömningen av nationella prov. I rapporten presenteras både deskriptiva resultat

---

<sup>4</sup> I svenska årskurs 6 är delprovet uppdelat i två delar (C1 och C2).



och ett antal statistiska analyser med bäring på frågan om likvärdig bedömning av delproven. Liksom i tidigare rapporter mäts skillnaden mellan skolans och omdömnarnas bedömning samt hur bedömningsprocessen påverkar likvärdigheten i bedömningarna. I årets rapport görs dessutom en jämförelse av andelen avvikelser mellan de olika betygsstegen, analyser av skillnader mellan fristående och kommunala skolor samt en fördjupad analys av hur bedömningsprocessen påverkar likvärdigheten. Därtill görs en jämförelse av ursprungsbedömning, första omdömnings och andra omdömnings för delprov som har omdömts två gånger. Den detaljerade redogörelsen för projektets genomförande, urvalet av skolor och delprov samt beräkningstekniska detaljer har sammanställts i bilagor. Den läsare som vill få en djupare förståelse för resultaten uppmanas att läsa bilagorna.

# Resultat

I avsnittet presenteras resultaten från 2017 års omedbedömning. Bland annat redogörs för hur mycket skolornas bedömningar avviker från omedbömarnas, spridningen i andel avvikelser mellan skolor samt hur bedömningsprocessen ser ut på skolorna och hur den påverkar skillnaderna mellan ursprungsbedömning och omedbedömning.

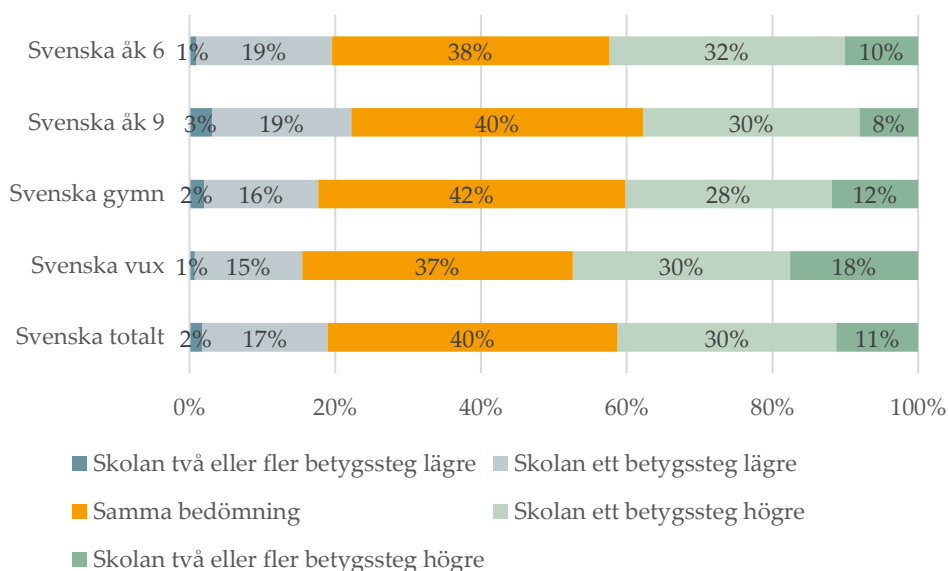
## Hur mycket skiljer sig bedömningarna av samma elevlösning?

### En majoritet av elevlösningarna har bedömts olika

Andelen avvikelser mellan de bedömningar som gjorts av skolans lärare och Skolinspektionens omedbömarna varierar både mellan årskurser och kurser samt ämnen. För både delproven i svenska och engelska är andelen elevlösningar där omedbömarna gjort en annan bedömning än ursprungsbedömarna 60 procent. För delproven i NO-ämnena är andelen lägre, 47 procent. I samtliga ämnen är det vanligare att skolans lärare gjort en mer generös bedömning än omedbömarna.

### Vanligt att skolan gjort en mer generös bedömning än omedbömaren i svenska

Som framgår i figur 1 nedan har omedbömaren gjort en annan bedömning än ursprungsbedömaren för 60 procent av de omedbömnda delproven i svenska. Det är vanligare att skolan har gjort en mer generös bedömning än omedbömaren. För 41 procent av delproven har ursprungsbedömaren satt ett högre betyg än omedbömaren och för 19 procent har ursprungsbedömaren satt ett lägre betyg.

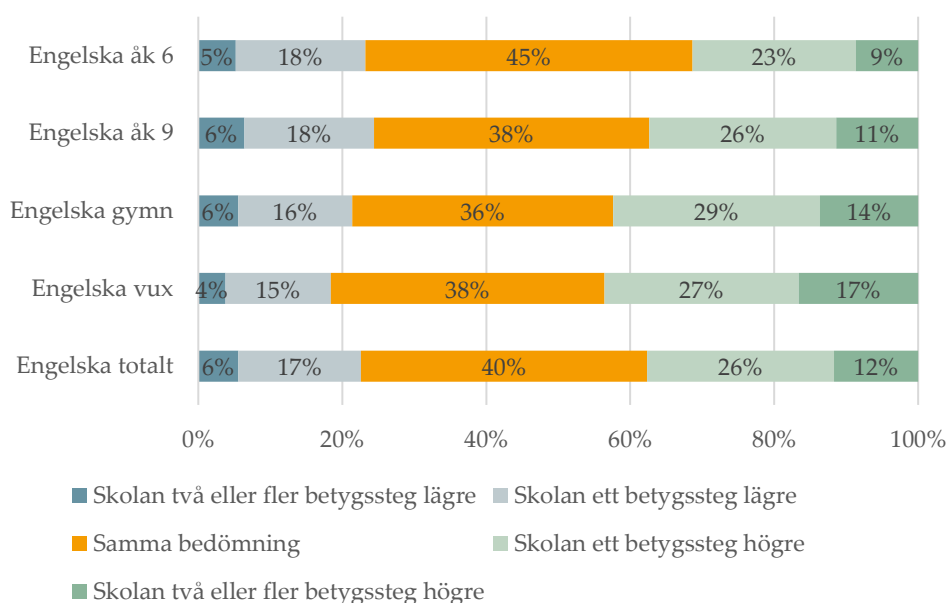


Figur 1: Resultat i omedbedömningen för utvalda delprov i svenska, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens omedbömarna ger samma betyg, samt i vilken utsträckning skolans lärare ger ett lägre respektive högre betyg.

Andelen avvikelser i svenska är ungefär lika omfattande i de olika årskurserna och kurserna. Däremot är andelen delprov där skolan gjort en generösare bedömning något högre inom vuxenutbildningen än i grundskolan och gymnasiet. Bland dessa delprov har 48 procent fått ett högre betyg vid ursprungsbedömning än ombedömning.

## Generösare bedömningar vanligare i högre årskurser och kurser i engelska

Andelen delprov i engelska där skolans lärare och ombedömarna gjort olika bedömningar är totalt 60 procent, vilket framgår i figur 2. Likt resultaten för svenska är det vanligare att skolans lärare gjort en mer generös bedömning än ombedömarna. I 38 procent av fallen har skolan gjort en högre bedömning än ombedömarna och i 23 procent av fallen har skolan gjort en lägre bedömning.



Figur 2: Resultat i ombedömningen för utvalda delprov i engelska, efter årskurs/kurs. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens ombedömarna ger samma betyg, samt i vilken utsträckning skolans lärare ger ett lägre respektive högre betyg.

En jämförelse mellan årskurserna och kurserna i engelska visar att ju högre upp i utbildningssystemet, desto högre blir andelen delprov där skolans lärare gjort en generösare bedömning än ombedömarna. För årskurs 6 har skolan gjort en bedömning som är minst ett betygssteg högre för 32 procent av delproven. För gymnasiet och vuxenutbildningen är motsvarande andel 43 respektive 44 procent.

## Resultaten för samma årskurs och kurs varierar mellan språkämnen

Inom samma årskurs eller kurs ses de största skillnaderna i resultat mellan de två språkämnen i årskurs 6. I svenska årskurs 6 (se figur 1) har skolans lärare i 38 procent av fallen gjort *samma* bedömning som ombedömarna medan motsvarande siffra för eng-

elska är 45 procent (figur 2). För 42 procent av de bedömda delproven i svenska för årskurs 6 har skolans lärare satt ett betyg som är minst ett betygssteg högre än ombedömarena. Motsvarande siffra för engelska är 32 procent.

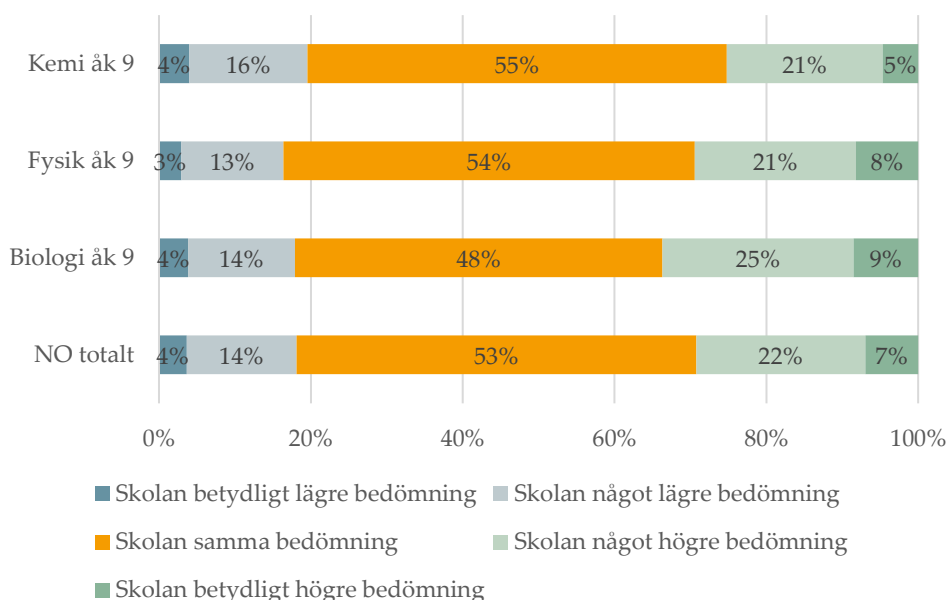
Andelen delprov i svenska för gymnasieskolan där ursprungsläraren och ombedömare gjort samma bedömning är 42 procent. Motsvarande siffra för engelska är 36 procent.

För årskurs 9 skiljer sig inte resultaten i någon större utsträckning mellan svenska och engelska. Detsamma gäller för vuxenutbildningen.

## Större samstämmighet i NO-ämnena än i språkämnena

Som visas i figur 3 nedan är andelen delprov där skolan och ombedömare gjort samma bedömning 53 procent för NO-ämnena, medan motsvarande andel för svenska och engelska är 40 procent (figur 1 och 2). Det innebär att skolans lärare och Skolinspektionens ombedömare i större utsträckning gör samma bedömning i NO-ämnena än i svenska och engelska.

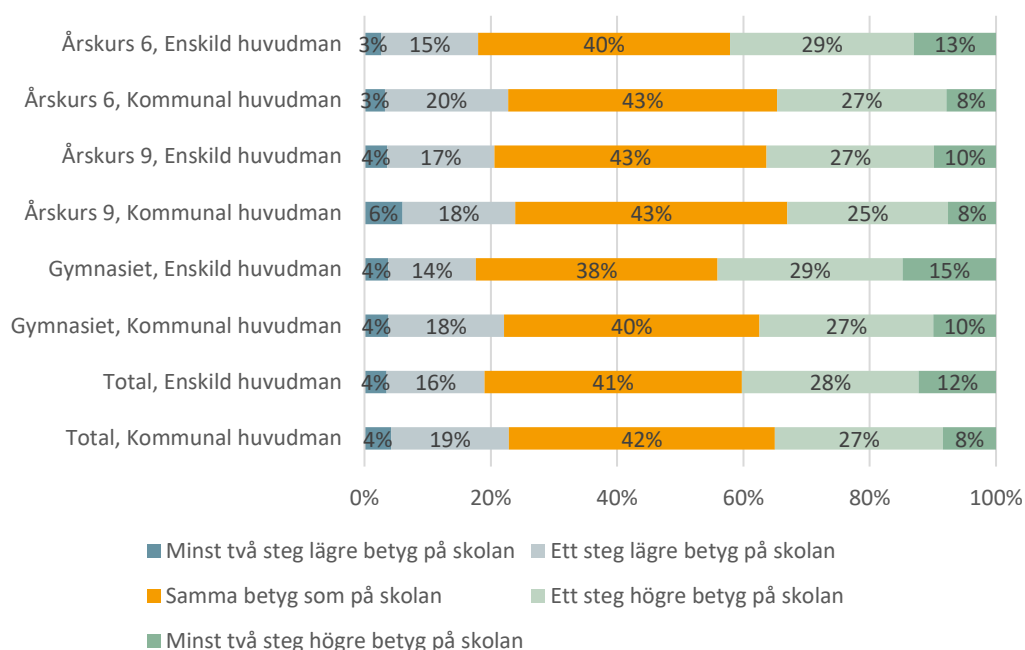
Avvikelsegraden varierar något mellan de olika NO-ämnena. I figuren nedan framgår att andelen delprov där ursprungsbedömarna och ombedömare gjort samma bedömning är något lägre för biologi, 48 procent, jämfört med kemi och fysik, 55 procent respektive 54 procent. Andelen delprov där skolan har gjort en lägre bedömning än ombedömare är liknande för de tre ämnena. Andelen delprov där skolan har gjort en högre bedömning än ombedömare är 26 procent i kemi, 29 procent i fysik och 34 procent i biologi.



Figur 3: Resultat i ombedömningen för utvalda delprov i kemi, fysik och biologi i årskurs 9. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens ombedömare gör samma bedömning, samt i vilken utsträckning skolans lärare gör en lägre respektive högre bedömning.

## Små skillnader mellan fristående och kommunala skolor

Som framgår av de två nedersta staplarna i figur 4 nedan är skillnaderna i resultat mellan skolor med enskild huvudman jämfört med skolor med kommunal huvudman relativt små. Sett till samtliga årskurser och kurser är andelen delprov där skolans lärare gjort en lägre bedömning än Skolinspektionens ombedömare tre procentenheter högre bland kommunala skolor än bland fristående. Andelen delprov där skolans lärare gjort en högre bedömning än Skolinspektionens ombedömare är fem procentenheter högre bland fristående skolor än bland kommunala. Skillnaderna mellan fristående och kommunala skolor är ungefär lika stora i alla årskurser och kurser.



Figur 4: Resultat i ombedömningen, efter årskurs/kurs samt typ av huvudman. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens ombedömare ger samma betyg, samt i vilken utsträckning skolans lärare ger ett lägre respektive högre betyg (belägg i NO-ämnena har räknats om till skalan A–F).

## Överlag små skillnader jämfört med tidigare år

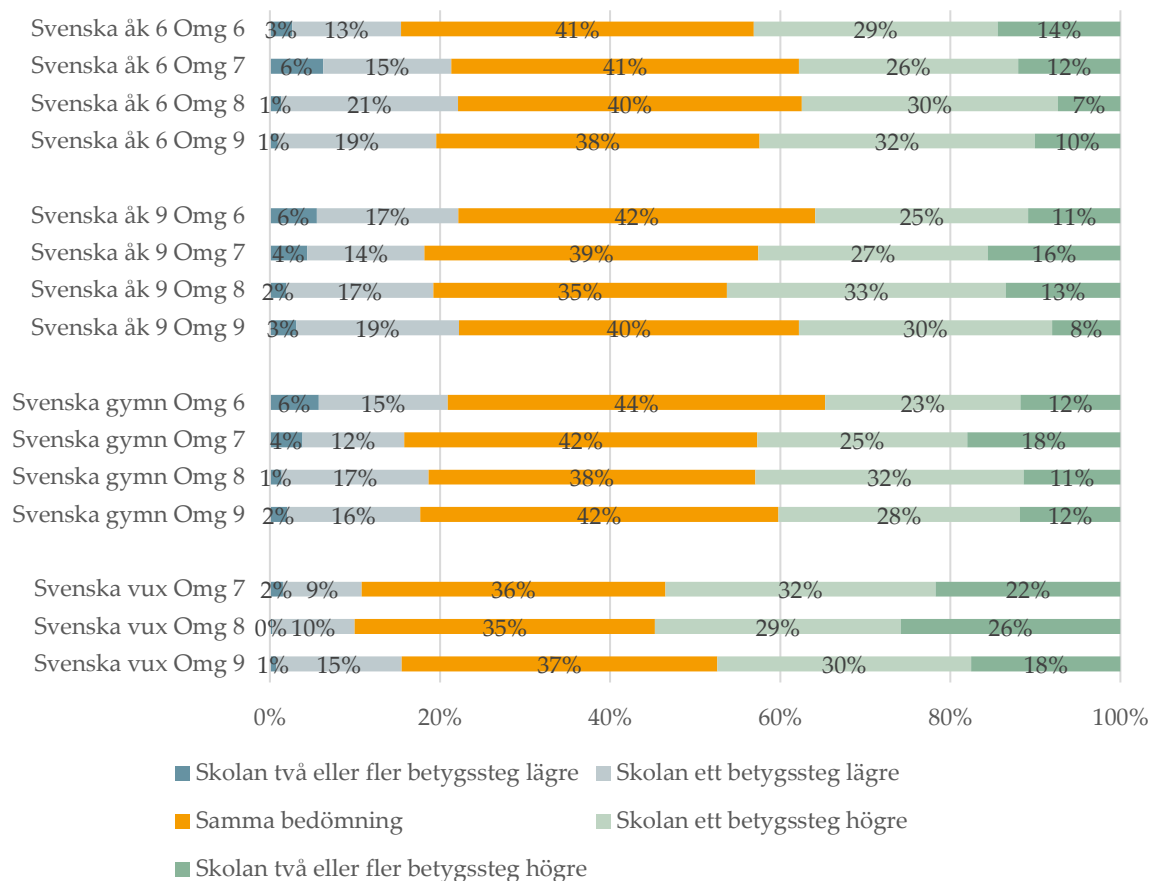
I figur 5–7 nedan ställs resultaten från årets omgång av ombedömningen i relation till resultaten från de tre senaste omgångarna.<sup>5</sup>

Vid läsningen 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 gör att resultaten blir mer användbara i inspektionen, men att nationella jämförelser mellan åren som nedan 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.

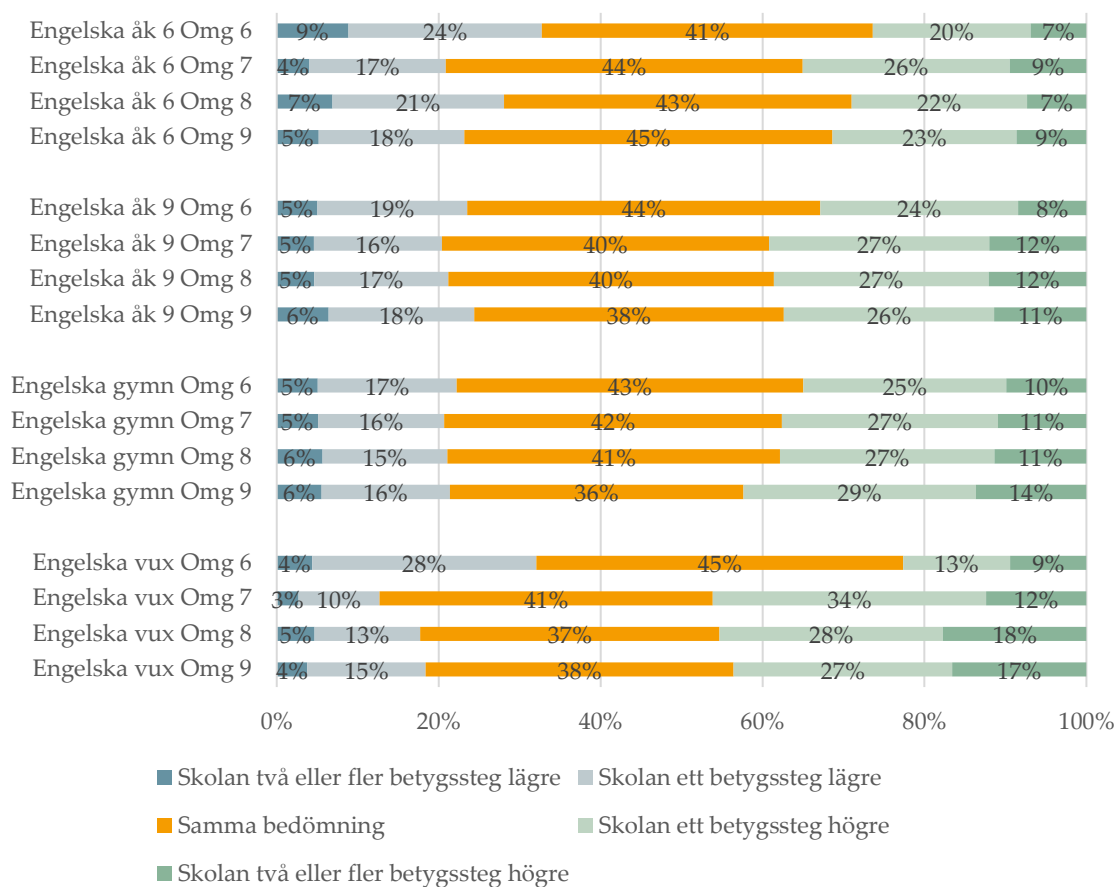
Överlag ses små skillnader i resultatet jämfört med tidigare år. Ett tydligt undantag är resultatet för NO-ämnena (figur 7 nedan). Där har andelen delprov där skolans lärare gjort

<sup>5</sup> Resultatet från ombedömningen i svenska för den kommunala vuxenutbildningen redovisas inte för omgång 6 eftersom ett annat delprov har ombedömts denna omgång jämfört med vid de senare omgångarna.

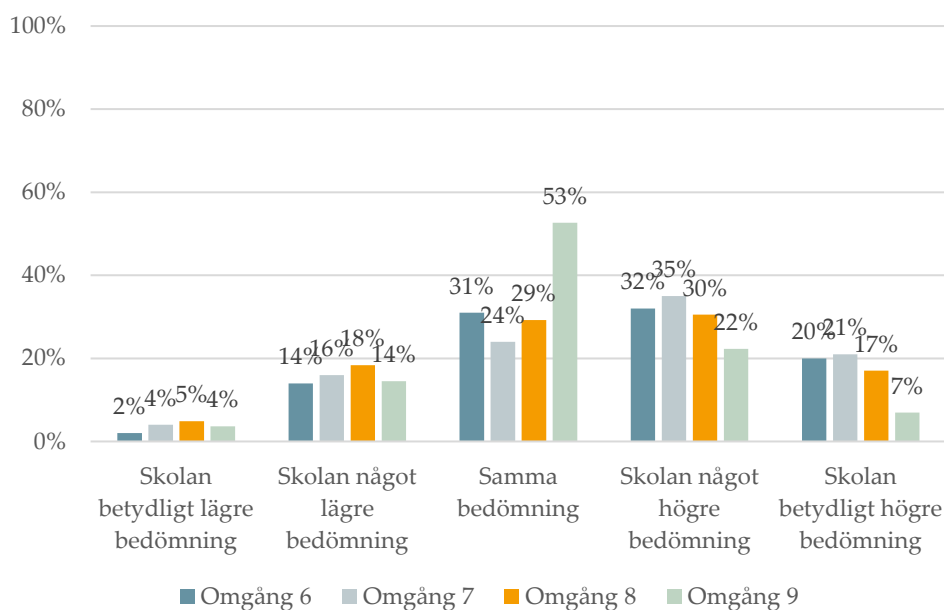
samma bedömning som Skolinspektionens ombedömare ökat från 29 procent förra året till 53 procent i år. I de fall skolans lärare och Skolinspektionens ombedömare gjort olika bedömningar finns tendensen från föregående omgångar dock kvar, att skolans lärare gör en mer generös bedömning än ombedömarna. En trolig förklaring till den mer likvärdiga bedömningen i NO-ämnet är att delproven fått en ny utformning och nya bedömningsanvisningar.



Figur 5: Resultat i ombedömningen för delprov i svenska de senaste fyra omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens ombedömare ger samma betyg, samt i vilken utsträckning skolans lärare ger ett lägre respektive högre betyg.



Figur 6: Resultat i ombedömningen för delprov i engelska de senaste fyra omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens ombedömare ger samma betyg, samt i vilken utsträckning skolans lärare ger ett lägre respektive högre betyg.



Figur 7: Resultat i omedbedömningen för delprov i NO-ämnen i årskurs 9 de senaste fyra omgångarna/åren. Figuren visar i vilken utsträckning skolans lärare och Skolinspektionens omedbedömare gör samma bedömning, samt i vilken utsträckning skolans lärare gör en lägre respektive högre bedömning.

## Stor variation mellan olika skolor

I detta avsnitt redovisas resultat på skolenhetsnivå för att visa den genomsnittliga skillnaden mellan skolans bedömning och omedbedömningen av de utvalda delproven.

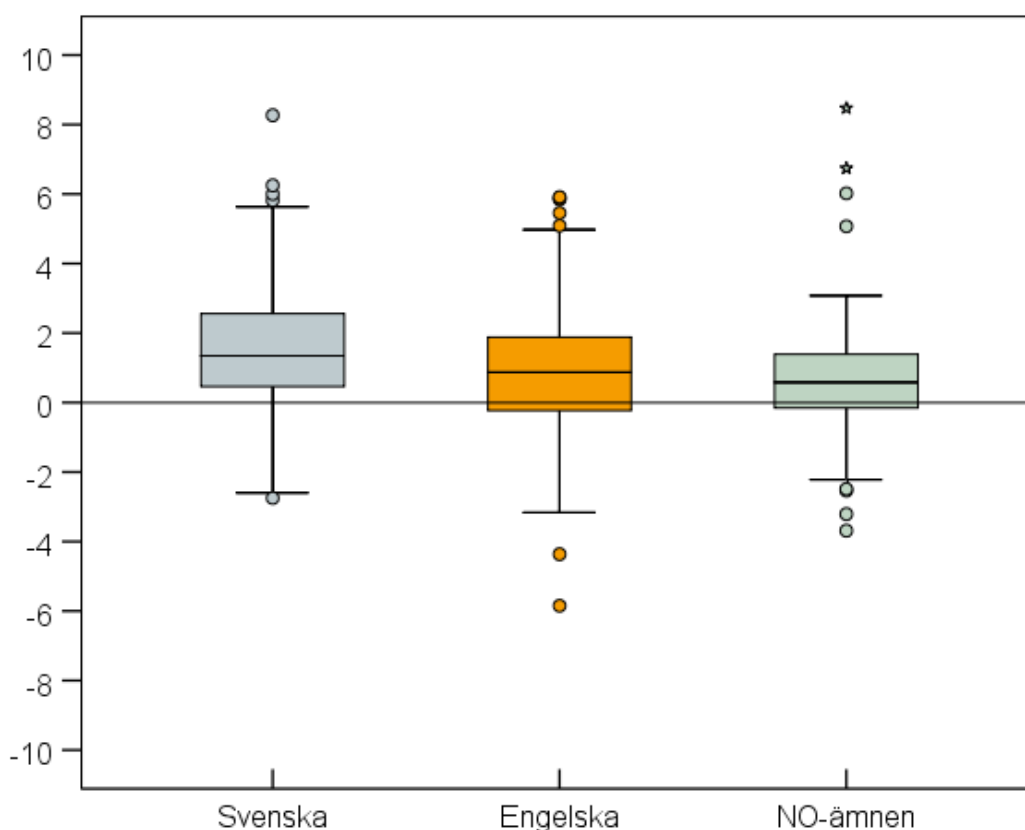
I figur 8 nedan redovisas den genomsnittliga skillnaden mellan olika skolor per ämnesgrupp. Fyra enheter på skalan motsvarar ett steg på betygsskalan A–F.<sup>6</sup> Värdet större än noll innebär att skolan i genomsnitt sätter ett högre betyg än Skolinspektionens omedbedömning.

Resultaten visar att skolorna i genomsnitt sätter högre betyg på de utvalda nationella delproven än Skolinspektionens omedbedömning. Medianen för de tre ämnena ligger mellan 0,58 och 1,35. Det innebär att skolorna i genomsnitt sätter ungefär en fjärdedels betygssteg högre betyg på skalan A–F än vad Skolinspektionens omedbedömning gör. Skillnaden är som störst i svenska och som minst i NO-ämnena. Det är dock stor spridning mellan olika skolor. I alla tre ämnena skiljer det ungefär tre betygssteg på skalan A–F mellan den skola som har störst positiv respektive störst negativ avvikelse i omedbedömningen.

Hälften av skolorna ryms inom boxarna i figuren, som visar medianen samt den 25:e och 75:e percentilen. Variationen inom denna grupp närmast medianen, det vill säga kvartilavståndet, är ungefär en halv betygsenhet på skalan A–F. Figuren visar också att ungefär tre fjärdedelar av skolorna i genomsnitt har satt ett högre betyg än Skolinspektionens omedbedömning på de utvalda delproven, medan en fjärdedel har satt ett lägre betyg. Skolor som ligger extra långt ifrån medianen räknat i antal kvartilavstånd (motsvarar boxens höjd) är markerade med ring eller stjärna.

<sup>6</sup> I NO-ämnena har belägg räknats om till skalan A–F, se bilaga 3.





Figur 8: 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. Totalt ingår därmed 277 skolor i svenska, 265 skolor i engelska och 81 skolor i NO-ämnen. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Resultat större än noll innebär att skolan i genomsnitt har satt högre betyg på de utvalda delproven än Skolinspektionens omdömare. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnen har räknats om till denna skala).

Trots att det finns en stor variation i resultat mellan olika skolor, betyder det inte att en enskild elev med säkerhet skulle få ett högre betyg genom att byta skola. En relativt stor del av spridningen kan nämligen bero på att det är svårt att bedöma de utvalda delproven, som består av frågor med fritextsvar. Därför är det inte otänkbart att samma lärare skulle kunna bedöma samma elevlösning olika, om det gått lång tid emellan och läraren hunnit glömma elevens svar.<sup>7</sup> En annan faktor som bidrar till slumpmässiga variationer är att endast ett begränsat antal delprov vid varje skola har ingått i omdömningen.

En indikation på hur svårt det är att komma fram till samma bedömning av samma elevlösning kan fås genom de analyser Skolinspektionen gjort då två omdömare fått bedöma samma elevlösning var för sig. Det blir då i många fall lika stora skillnader mellan Skolinspektionens båda omdömare som mellan läraren på skolan och respektive omdömare.

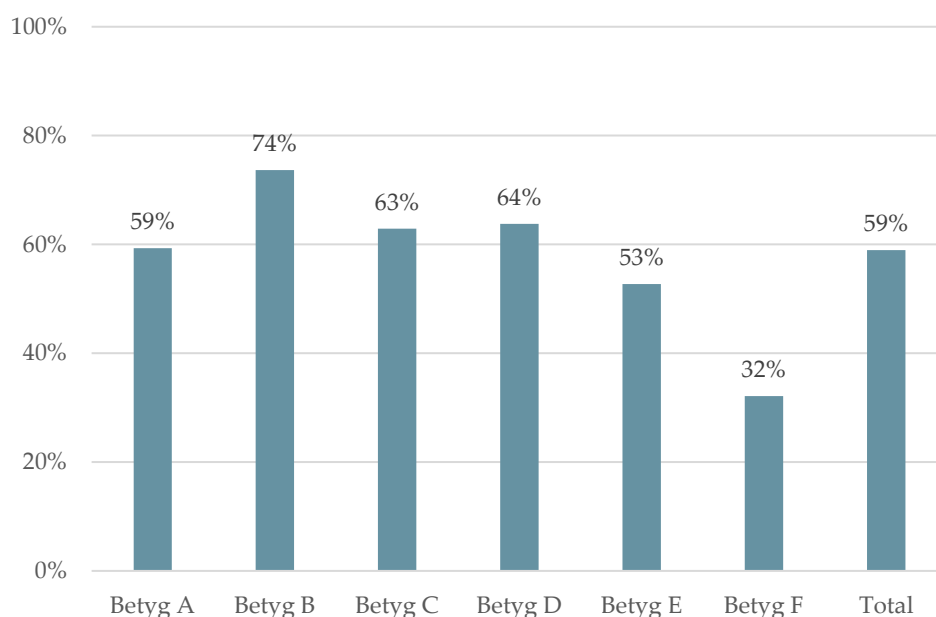
<sup>7</sup> Skolverket. *Sambedömning i skolan – exempel och forskning*. s. 9.

## Störst samstämmighet kring betyget F och lägst kring betyget B

I följande avsnitt jämförs andelen avvikelser mellan ursprungsbedömning och ombedömning mellan de olika betygsstegen (A–F). För de delprov där bedömningen görs i form av belägg (poäng) har dessa räknats om till A–F-skalan. Mer information om belägg och hur omräkningen gjorts finns i bilaga 3.

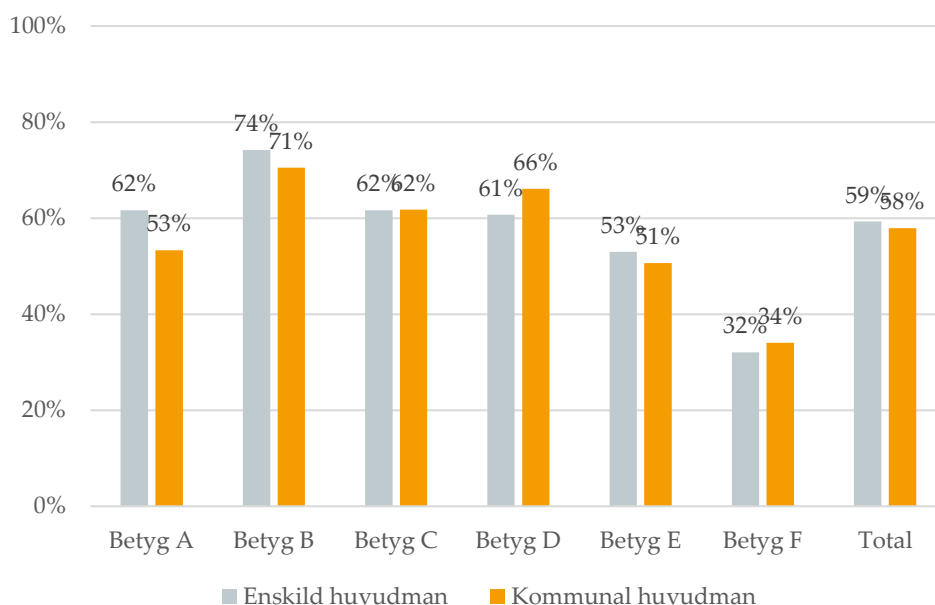
I figur 9 nedan jämförs andelen avvikelser mellan ursprungsbedömning och ombedömning per betygssteg. Jämförelsen utgår från den ursprungliga lärarens bedömning, det vill säga i stapeln "Betyg A" visas andelen elevlösningar där ursprungsläraren har satt ett A och ombedömaren har satt ett annat betyg.

I jämförelsen mellan betygsstegen bör has i åtanke att avvikelser från betygen A och F bara kan gå åt det ena hållet. Det går inte att sätta ett högre betyg än A respektive ett lägre betyg än F. Det är därför intressant att notera att oenigheten mellan ursprungsbedömare och ombedömare är betydligt mindre för betyget F än vad det är för A. Det finns generellt en större enighet bland lärare kring elevlösningar som inte uppfyller kunskapskraven för godkänt betyg än vad det gör för det högsta betyget. Störst oenighet finns bland de delprov som av ursprungsbedömaren tilldelats betyget B. För dessa delprov har ombedömaren satt ett annat betyg i 74 procent av fallen.



Figur 9: Andel delprov där skolans lärare och Skolinspektionens ombedömare ger olika betyg, efter betyg av den ursprungliga läraren (belägg i NO-ämnen har räknats om till skalan A–F)

En uppdelning per typ av huvudman visar att andelen avvikelser för betyget A är högre bland fristående skolor än bland kommunala, vilket visas i figur 10 nedan. Av de delprov som på skolan fått bedömningen A, har 62 procent av delproven från fristående skolor fått ett annat betyg i ombedömningen jämfört med 53 procent av delproven från kommunala skolor.



Figur 10: Andel delprov där skolans lärare och Skolinspektionens ombedömare ger olika betyg, efter betyg av den ursprungliga läraren (belägg i NO-ämnen har räknats om till skalan A–F) samt typ av huvudman. Delprov från komvux ingår inte.

## Bristande likvärdighet i bedömningarna av vissa delprov

Sammanfattningsvis visar resultaten, liksom i tidigare ombedömningar, att de delprov som till stor del består av fritextsvar inte bedöms likvärdigt. En och samma elevlösning har i över hälften av fallen fått olika bedömningar av två olika bedömare. För många delprov skiljer det dessutom två betygssteg eller mer mellan de olika bedömningarna. En ny analys som gjorts i år visar att samstämmigheten mellan ursprungsbedömare och ombedömare är störst för delprov som av ursprungbedömarna fått betyget F och lägst för betyget B.

Eftersom att endast ett av delproven för ett ämnes- eller kursprov ingår i ombedömningen 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 står alltså klart att de avvikelser som uppmäts i ombedömningen på delprovsnivå är av den storleksordningen att de kan komma att leda till att en elevs totala provbetyg påverkas.

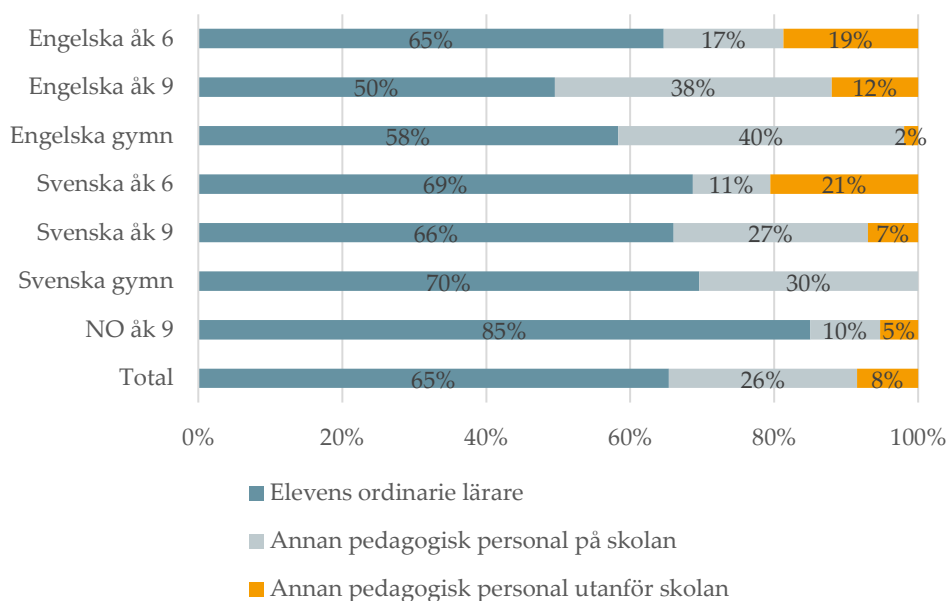
## Hur påverkas likvärdigheten av bedömningsprocessen?

### Stora variationer i hur delproven bedöms

För att ta reda på hur bedömningsprocessen sett ut för de elevlösningar som ombedömts har Skolinspektionen bitt skolorna att för varje insänd elevlösning besvara en enkät.<sup>8</sup> I enkäten ställdes frågor om hur det specifika delprovet bedömts. Enkäten var frivillig att besvara. Därför finns inte uppgifter om hur bedömningen gått till för alla elevlösningar som ombedömts.

### 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 65 procent av fallen, se figur 11 nedan. I samma figur visas att 26 procent av delproven bedömts av annan pedagogisk personal på skolan och 8 procent av delproven av annan pedagogisk personal utanför skolan. Det innebär att det likt tidigare år är vanligast att elevens ordinarie lärare bedömt delprovet. Andelen delprov som bedömts av någon annan än elevens ordinarie lärare (så kallad *extern bedömning*) är dock tio procentenheter högre bland de delprov som ingått i årets ombedömning jämfört med de delprov som ingick i ombedömningen förra året. Det finns betydande variationer mellan ämnen. Att någon annan än elevens ordinarie lärare bedömt provet är vanligast i engelska och minst vanligt inom NO-ämnena.



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

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

<sup>8</sup> Inga uppgifter om bedömningsprocessen har efterfrågats för delprov genomförda inom den kommunala vuxenutbildningen.

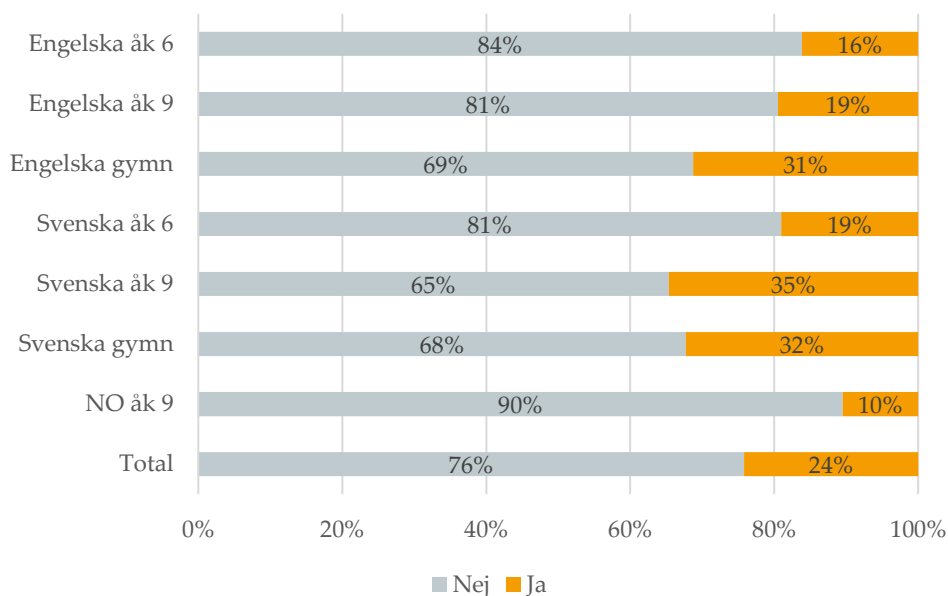
större variationer mellan årskurserna och kurserna. Andelen delprov som bedömts av elevens ordinarie lärare är som störst för årskurs 6, 65 procent, och som minst för årskurs 9, 50 procent. Andelen delprov som bedöms av lärare utanför skolan minskar ju högre årskurs och kurs, både i svenska och i engelska.

Det är vanligare med extern bedömning bland kommunala skolor än bland fristående. Andelen delprov som bedömts av annan lärare på skolan är 37 procent bland kommunala skolor jämfört med 17 procent bland fristående. Andelen delprov som bedömts av lärare utanför skolan är 16 procent bland kommunala jämfört med 2 procent bland fristående.

### En fjärdedel av delproven var avidentifierade

Av enkätsvaren framkommer att sett till alla ämnen så var 24 procent av elevlösningarna avidentifierade vid bedömningen, vilket framgår i figur 12 nedan. Andelen delprov som avidentifierats är högre bland de delprov som ingår i årets rapport jämfört med de som ingick i förra årets rapport, då var 18 procent av delproven avidentifierade. Andel avidentifierade delprov är i år högst för svenska årskurs 9 (35 %) och svenska på gymnasiet (32 %) samt i engelska på gymnasiet (31 %), och lägst i NO-ämnena (10 %).

Inga större skillnader har noterats mellan fristående och kommunala skolor. Bland de fristående skolorna var 26 procent av delproven avidentifierade jämfört med 22 procent bland de kommunala.

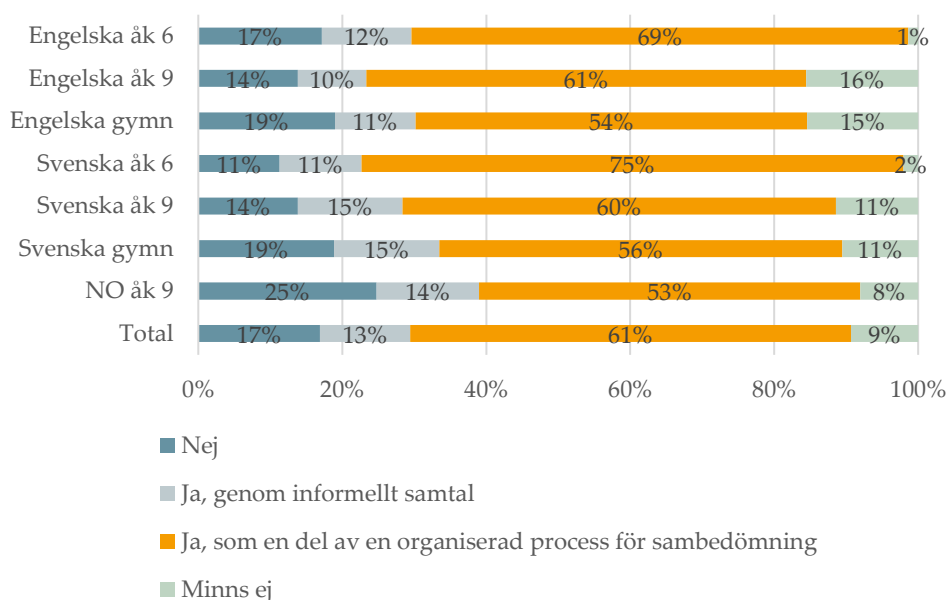


Figur 12: Andel delprov i omedömningen som var avidentifierade (Ja i figuren) vid den ursprungliga bedömningen på skolan, efter ämne samt årskurs/kurs.

### Tre av fyra delprov sambedömdes

Sambedömning innebär att lärare samarbetar om bedömning och betygssättning, till exempel genom att bedöma elevers prestationer tillsammans eller genom att diskutera bedömningen.<sup>9</sup> Med organiserad sambedömning avses här att skolan har strukturer och rutiner för att lärare ska bedöma prov i samarbete eller i diskussion med kollegor. Med informell sambedömning avses här att lärare bedömt prov 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.

För närmare tre av fyra delprov har någon form av sambedömning tillämpats i samband med bedömningen av elevlösningen, det visas i figur 13 nedan. 17 procent av delproven har bedömts utan sambedömning. Andelen delprov som har sambedömts är ungefär densamma som i föregående års ombedömning.



Figur 13: Andel delprov i ombedömningen 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 61 procent av delproven sambedömts på detta sätt. Denna typ av sambedömning var vanligast i årskurs 6, både i engelska (69 %) och svenska (75 %). Det var mindre vanligt med organiserad sambedömning på gymnasiet, där 54 procent av delproven i engelska och 56 procent av delproven i svenska sambedömts som en del av en organiserad process, samt i NO-ämnena där motsvarande andel är 53 procent.

Av enkätsvaren framgår att det varit vanligare att någon form av sambedömning tillämpats på kommunala skolor jämfört med fristående. De kommunala skolorna har uppgett att någon form av sambedömning tillämpats i 79 procent av fallen medan motsvarande

<sup>9</sup> Skolverket. *Sambedömning i skolan – exempel och forskning*. s. 11

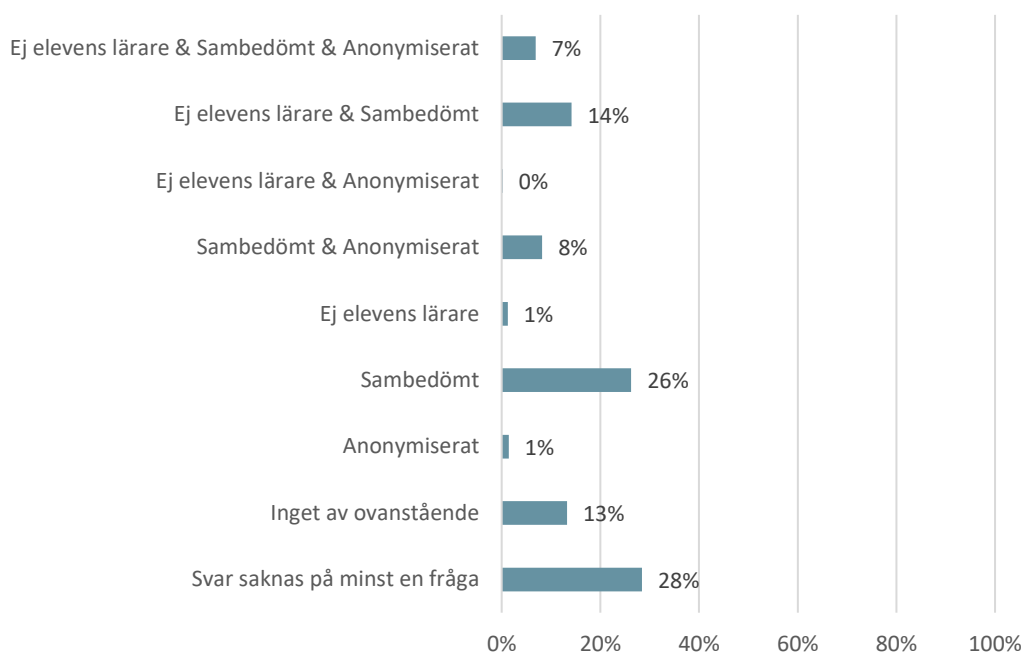
andel för fristående skolor är 68 procent. Tittar vi på formen för sambedömning är organiserad sambedömning vanligare bland kommunala skolor (70 %) än fristående (52 %), medan informell sambedömning är vanligare på fristående skolor (16 %) jämfört med kommunala (9 %).

## Stora variationer i hur bedömningsmetoderna kombineras

I figur 14 nedan redovisas resultat för *kombinationer av svar* på de tre frågorna om hur bedömningen av de nationella proven gått till på skolan.

Det finns stora skillnader när det gäller hur bedömningen av de nationella proven gått till på olika skolor. Som framgår av figuren saknas svar på minst en av de tre frågorna för 28 procent av de ombedömda delproven. Det innebär alltså att vi har kompletta svar på frågorna om bedömningsmetoderna för 72 procent av delproven. Av dessa 72 procent har ungefär en tiondel av delproven (7 procentenheter) 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 (13 procentenheter) 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 28 procentenheter av delproven som haft *en* gynnsam faktor för likvärdig bedömning och den absolut vanligaste faktorn är i dessa fall sambedömning. Bland de 22 procentenheter 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 avidentifierats samt bedömts av någon annan än elevens lärare är däremot mycket ovanlig (noll procentenheter vid avrundning till heltal).



Figur 14: Svar på tre frågor om hur bedömningen av de nationella delproven gått till på skolan. I figuren redovisas i vilken utsträckning faktorer som antas vara *gyynsamma* för likvärdig bedömning av proven förekommer, det vill säga att någon annan än elevens lärare har bedömt delproven, att delproven sambedömts av flera lärare samt att de varit avidentifierade vid bedömningen.

Sammanfattningsvis visar resultatet att inslagen av sambedömning, extern bedömning och avidentifiering varierar kraftigt både 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 för likvärdigheten

I figur 15–17 visas den genomsnittliga skillnaden mellan skolans bedömning och ombedömningen. Fyra enheter på skalan motsvarar ett steg på betygsskalan A–F.<sup>10</sup> Värdet större än noll innebär att skolan i genomsnitt satt ett generösare betyg än Skolinspektionens ombedömare. Värdet under noll innebär att skolan i genomsnitt satt ett striktare betyg.

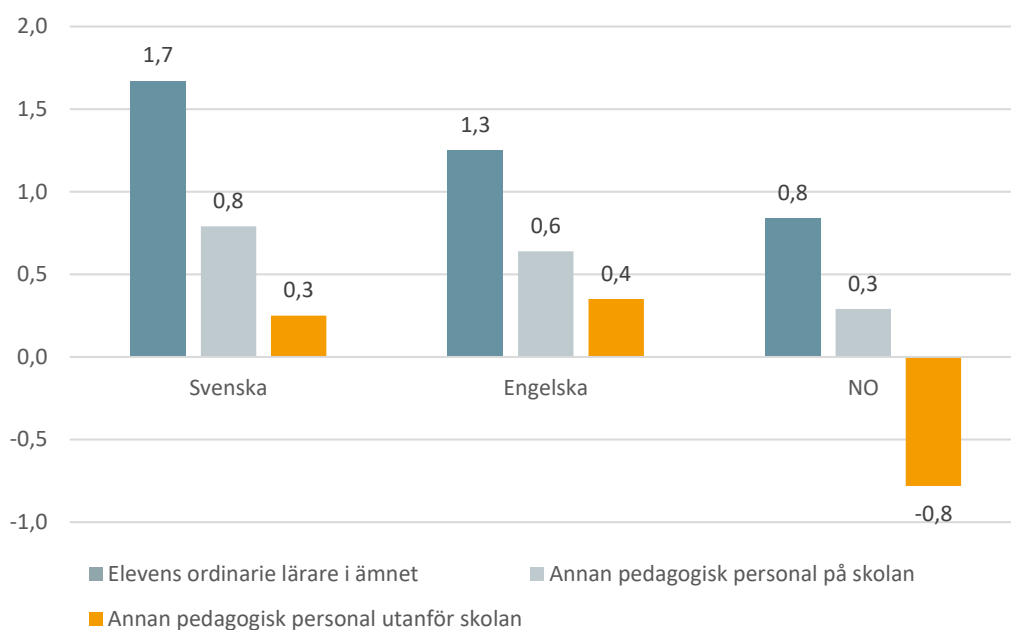
Den genomsnittliga skillnaden mellan delprov som bedöms av elevens ordinarie lärare och av lärare utanför elevens skola är statistiskt signifikant (se figur 15). Detta gäller för samtliga ämnen. I svenska och engelska är skillnaden dessutom statistiskt signifikant mellan delprov som bedöms av elevens ordinarie lärare och av annan lärare på skolan.

I ämnena svenska och engelska innebär det att skolor som låter lärare utanför skolan bedöma delprov i genomsnitt sätter betyg som ligger närmare Skolinspektionens ombedömningar, jämfört med delprov som bedöms av elevens ordinarie lärare. Det gäller även för delprov i svenska och engelska som bedöms av annan lärare på skolan. Inom ämnesgruppen NO går resultatet åt andra hållet. Delprov som bedöms av en annan lärare utanför

<sup>10</sup> I NO-ämnena har belägg räknats om till skalan A–F, se bilaga 3.

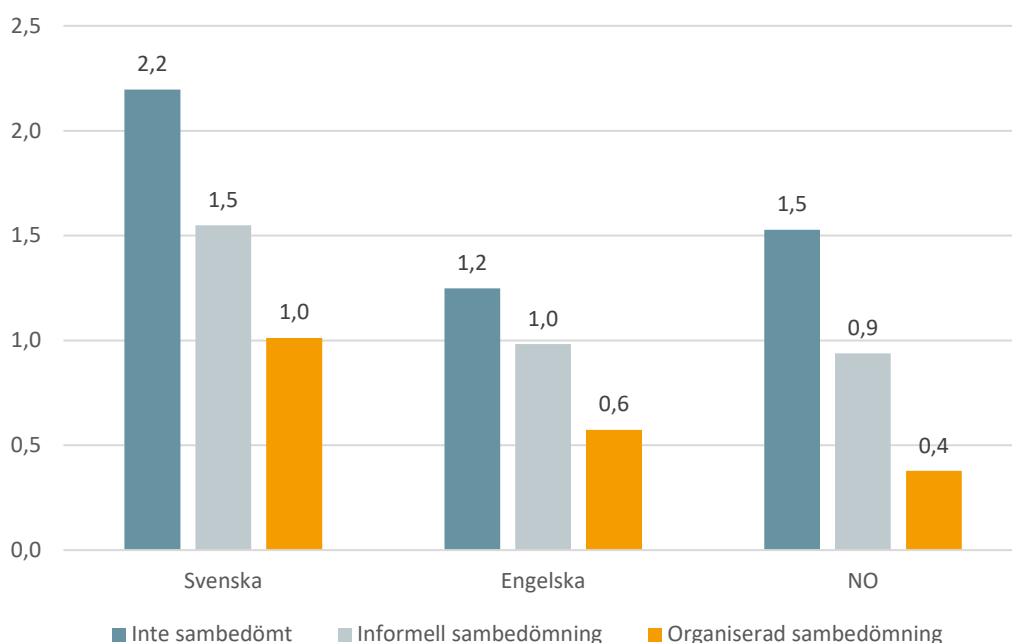


skolan får betyg som i genomsnitt är striktare (minusvärden i figuren) än Skolinspektionens omdömningar.



Figur 15. Genomsnittlig skillnad mellan skolans bedömning och omdömningen för delprov som bedömts av elevens ordinarie lärare eller av annan pedagogisk personal.

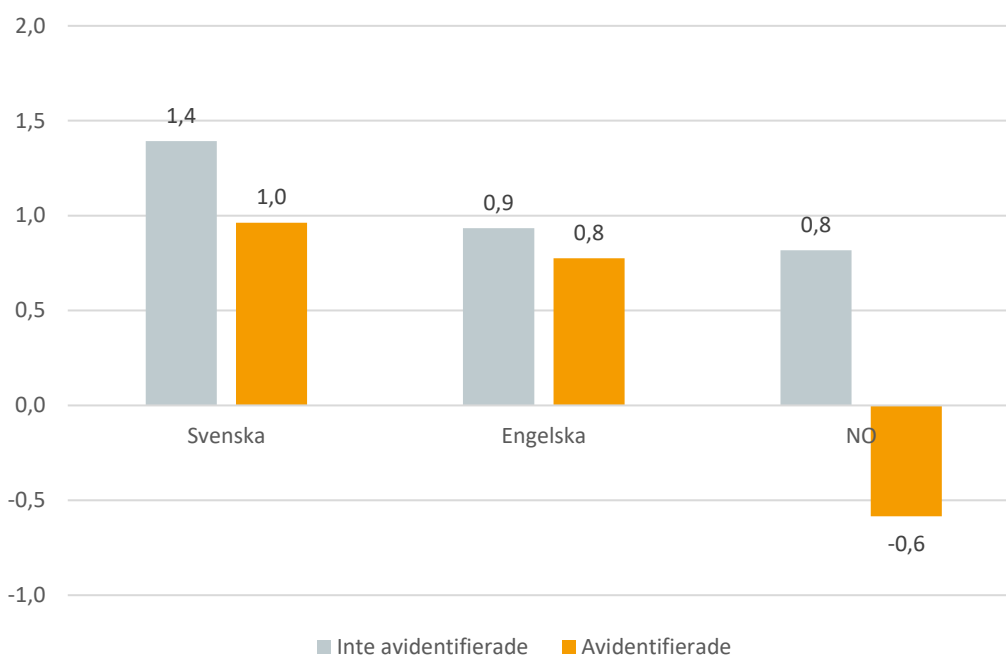
Figur 16 visar att sambedömning också minskar de genomsnittliga skillnaderna i betygsättningen. För samtliga ämnen är skillnaden mellan skolan och omdömningen statistiskt signifikant mindre för delprov som bedömts i en organiserad process för sambedömning än för delprov där ingen sambedömning skett. I ämnena engelska och svenska ser vi också att skillnaden mellan skolans bedömning och omdömningen är statistiskt signifikant mindre om delproven bedömts i en organiserad process för sambedömning jämfört med informell sambedömning. Detta mönster visar sig även för ämnesgruppen NO, men skillnaden är inte statistiskt signifikant.



Figur 16. Genomsnittlig skillnad mellan skolans bedömning och ombedömningen för delprov där bedömaren inte diskuterat bedömningen med någon annan, för delprov där bedömningen diskuterats genom informellt samtal respektive för delprov som bedömts i organiserad sambedömning.

Skillnaden mellan skola och ombedömare för avidentifierade delprov och delprov där elevens identitet är känd för läraren (se figur 17) är statistiskt signifikant för svenska och NO-ämnen, men resultaten skiljer sig ändå åt.

I svenska överensstämmer skolans och ombedömarens bedömningar i större utsträckning när delproven är avidentifierade, genomsnittsskillnaden minskar. Som visas i figur 17 går resultaten för ämnesgruppen NO åt olika håll. Avidentifierade delprov leder till att skolan bedömer proven striktare än ombedömaren medan delprov där elevens identitet är känd för läraren leder till generösare bedömningar på skolan. Det går därför inte säga att genomsnittsskillnaden mellan skola och ombedömare minskar för avidentifierade delprov inom NO-ämnen.

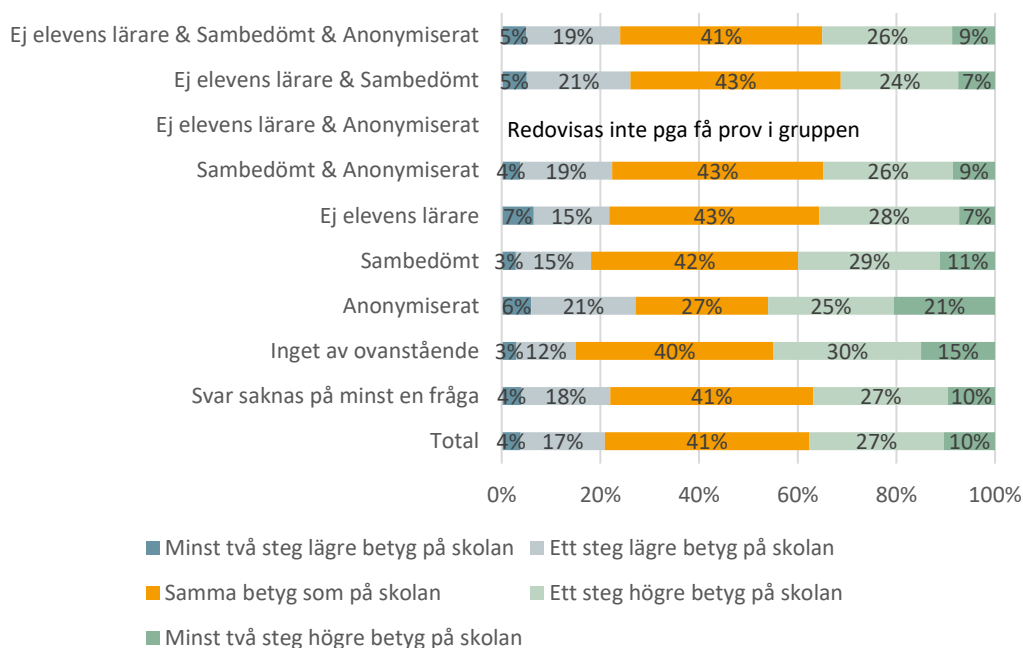


Figur 17. Genomsnittlig skillnad mellan skolans bedömning och ombedömningen för delprov som var avidentifierade vid bedömningen, respektive för delprov där elevens identitet var känd för läraren.

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

I detta avsnitt redovisas hur *kombinationer* av de tre bedömningsmetoderna påverkar resultatet i ombedömningen, till skillnad mot i figur 15–17 där de studerats en och en. 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.

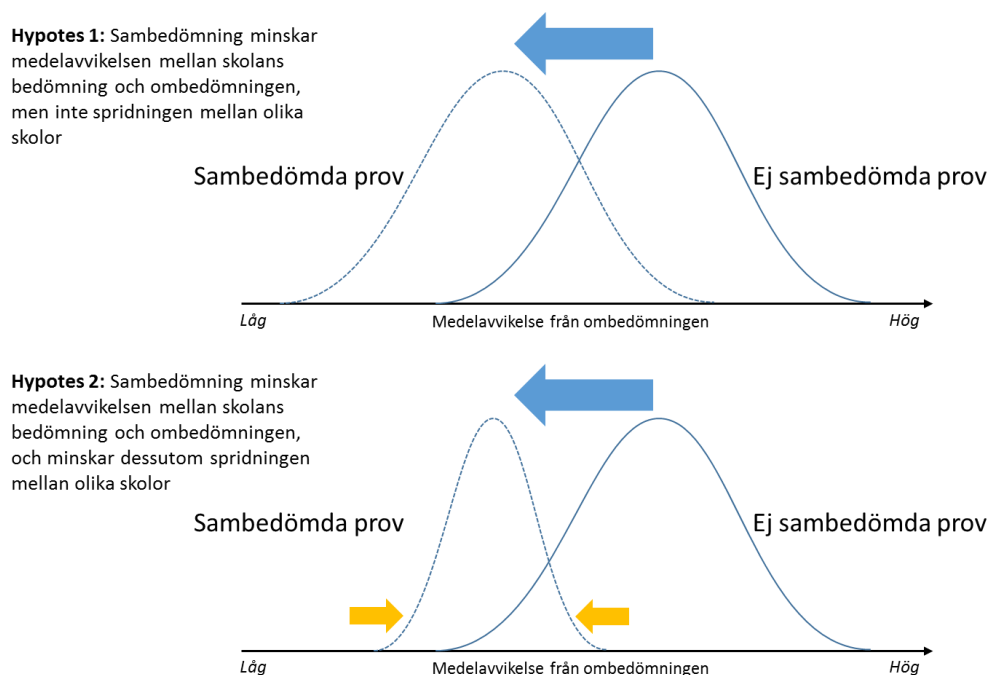
Som framgår av figur 18 får i genomsnitt 41 procent av de utvalda delproven samma betyg vid Skolinspektionens ombedömning som vid den ursprungliga bedömningen på skolan. Det är dock fler som får ett högre betyg vid den ursprungliga bedömningen på skolan (37 %) än ett lägre betyg (21 %). Resultaten i figur 18 visar att tendensen att skolans lärare gör en mer generös bedömning än Skolinspektionens ombedömare minskar om man tillämpar sambedömning och anonymisering samt låter någon annan än elevens ordinarie lärare bedöma delproven. Bland delprov utan dessa faktorer i bedömningsprocessen gör skolan en generösare bedömning i 45 procent av fallen och en striktare bedömning i 15 procent av fallen, vilket kan jämföras med 35 respektive 24 procent då alla de tre gynnsamma faktorerna förekommit när delprovet bedömts. Det tycks också som att den återhållande effekten i betygssättningen ökar ju fler faktorer för likvärdig bedömning som tillämpas, även om det i befintliga data är svårt att fastställa en skillnad mellan om två eller tre faktorer förekommer i bedömningsprocessen.



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). Redovisningsgrupperna anger i vilken utsträckning faktorer som antas vara *gynnsamma* för likvärdig bedömning av delproven förekommer, det vill säga att någon annan än elevens lärare har bedömt delproven, att delproven sambedömts av flera lärare samt att de varit avidentifierade vid bedömningen.

## Sambedömning och extern bedömning 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 betyg i snitt för motsvarande elevlösning än om elevens ordinarie lärare gör bedömningen 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 19: 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 i stor eller liten utsträckning sambedömer delprov på skolan, tillämpar extern bedömning eller avidentifierar delprov. För att visualisera spridningen redovisas resultaten i boxplotsdiagram.

Boxarna visar den 25:e och 75:e percentilen, inom detta spann ryms hälften av skolorna. Se figur 20. Det heldragna strecket genom boxen visar medianen. Medianen för skolor som tillämpar sambedömning är lägre än för skolor som inte sambedömer prov. Motsvarande en fjärdedels betygssteg jämfört med ett halvt betygssteg. Skolor som tillämpar sambedömning har alltså en lägre genomsnittsskillnad i förhållande till Skolinspektionens ombedömningar. Avståndet mellan den 25:e och 75:e percentilen motsvarar längden på boxen och kallas kvartilavstånd. Kvartilavståndet för skolor som sambedömer prov är 1,5 enheter (motsvarande drygt en fjärdedels betygssteg), jämfört med 2,4 enheter för skolor som inte sambedömer prov (motsvarande drygt ett halvt betygssteg). Spridningen mellan skolor är alltså lägre bland skolor som tillämpar sambedömning. För både skolor som tillämpar sambedömning och skolor som inte gör det är skillnaden mellan den skola som är mest generös i bedömningen jämfört med den som är mest strikt cirka ett och ett halvt betygssteg.



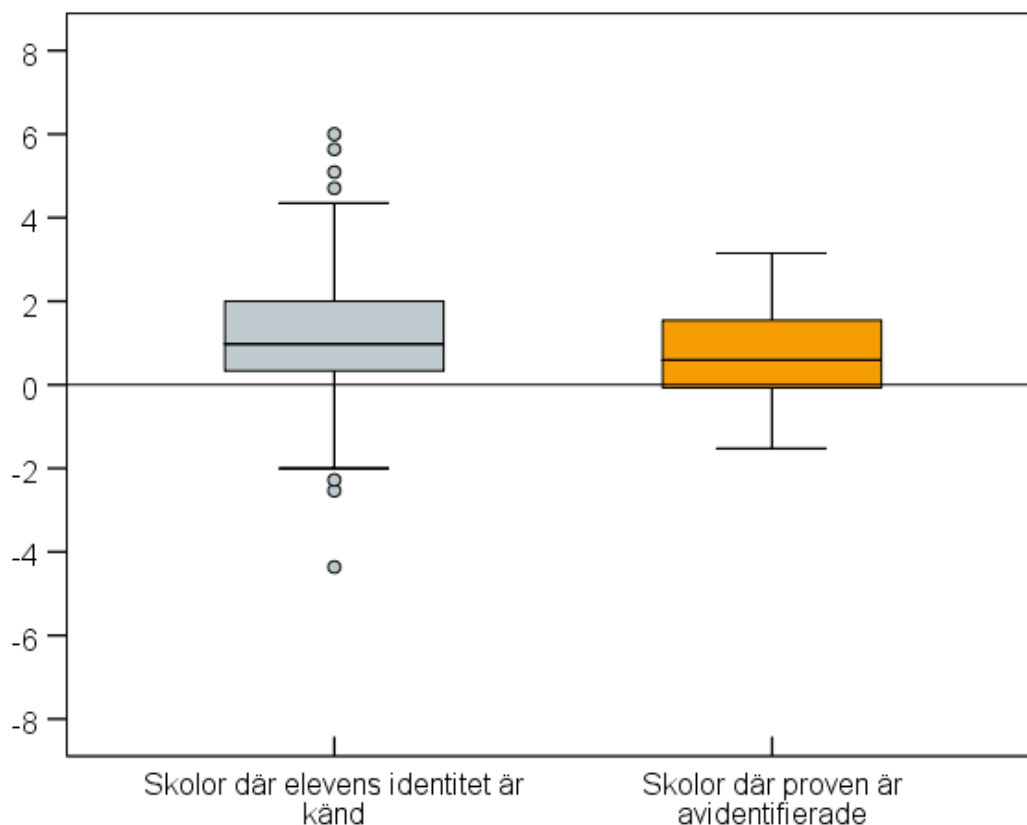
Figur 20: Avvikelse i omdömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många omdömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. I figuren redovisas endast de 35 skolor där under 20 procent av delproven i omdömningen samedöms samt de 202 skolor där minst 80 procent av delproven samedöms. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Resultat större än noll innebär att skolan i genomsnitt har satt högre betyg på de utvalda delproven än Skolinspektionens omdömare. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnen har räknats om till denna skala).

Liksom samedömning minskar extern bedömning genomsnittsskillnaden mellan skola och omdömare i betygssättningen (se figur 21). Medianen för denna grupp ligger på en åttiondels betygssteg jämfört med strax över en fjärdedels betygssteg för prov som bedöms av elevens ordinarie lärare. Extern bedömning minskar även längden på kvartilavståndet. Extern bedömning minskar spridningen mellan skolor men effekten är dock inte lika tydlig som för samedömning. Skillnaden mellan den skola som är mest generös jämfört med mest strikt i bedömningen av elevlösningar är lägre för skolor som tillämpar extern bedömning, ett och ett halvt betygssteg, jämfört med två och ett halvt betygssteg för skolor som inte gör det.



Figur 21: Avvikelse i omdömningen på skolenivå. 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 179 skolor där under 20 procent av delproven i omdömningen bedöms av annan pedagogisk personal än elevens ordinarie lärare samt de 59 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. Resultat större än noll innebär att skolan i genomsnitt har satt högre betyg på de utvalda delproven än Skolinspektionens omdömare. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnena har räknats om till denna skala).

Medianen för skolor som aidentifierar delprov är lägre än för skolor som inte aidentifierar delprov (se figur 22). Aidentifiering minskar den genomsnittliga betygsavvikelsen i förhållande till omdömningen men inte lika tydligt som för sambedömning. Aidentifiering minskar inte längden på kvartilavståndet. Skillnaden i bedömning mellan den mest generösa och mest strikta skolan är ett betygssteg för skolor som aidentifierar delprov jämfört med två och ett halvt betygssteg på skolor där delproven inte aidentifieras.



Figur 22: Avvikelse i omedbedömningen på skolnivå. Figuren är baserad på ett medelvärde per skola för de skolor som haft tillräckligt många omedbedömda delprov i ämnet för att uppfylla kriterierna för att få en skolenhetsrapport i ämnet. I figuren redovisas endast de 230 skolor där under 20 procent av delproven i omedbedömningen är avidentifierade samt de 53 skolor där minst 80 procent av delproven är avidentifierade. Boxarna visar var medianen samt den 25:e och 75:e percentilen ligger. Resultat större än noll innebär att skolan i genomsnitt har satt högre betyg på de utvalda delproven än Skolinspektionens omedbedömare. Fyra enheters skillnad motsvarar ett betygssteg på skalan A–F (belägg i NO-ämnena har räknats om till denna skala).

Sambedömning är den bedömningsmetod som tydligast visar på positiva effekter på minskad spridning mellan olika skolor, likaså extern bedömning men inte i samma utsträckning. I viss mån visar även avidentifiering på liknande tendenser. Bedömningsmetoderna kan dessutom kombineras. Att kombinera olika bedömningsmetoder bidrar till att en lägre andel delprov bedöms generösare av skolorna jämfört med omedbedömningen (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.

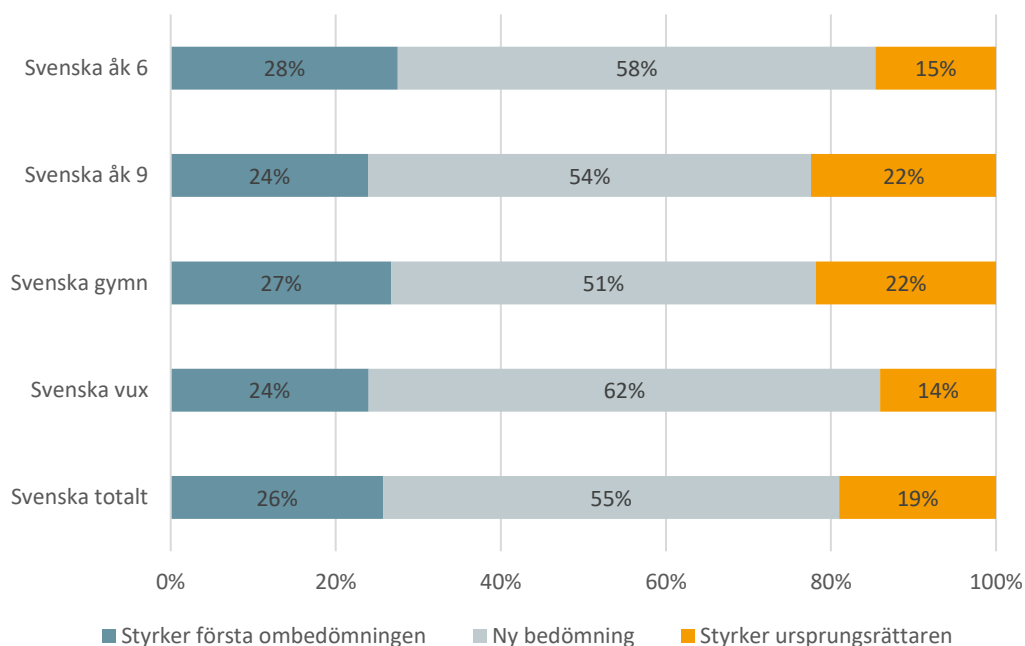
## Delprov som omedbedömts två gånger

Om det för prov i svenska skiljer mer än två betygssteg mellan ursprungsbedömningen och Skolinspektionens bedömning görs ytterligare en oberoende bedömning av en andra omedbedömare. I figur 23 nedan undersöks om andra omedbedömningen styrker Skolinspektionens första bedömning, ursprungsbedömaren eller om det blir ett helt nytt betyg.

Resultaten från jämförelsen visar att över hälften av delproven, i alla årskurser och kurser, får ett nytt betyg i den andra omedbedömningen. Andelen är som störst för komvux (62



%) och lägst för gymnasiet (51 %). Det är alltså vanligare att den andra omdömningen leder till ett nytt betyg än att den styrker någon av de två föregående bedömningarna. I genomsnitt styrker den andra omdömningen Skolinspektionens första bedömning med 26 procent. Mindre vanligt är det att andra omdömningen styrker ursprungsbedömaren, i genomsnitt 19 procent. Det skiljer sig något mellan årskurserna och kurserna. I komvux och årskurs 6 styrker andra omdömningen ursprungsbedömaren i 14 respektive 15 procent av fallen. I årskurs 9 och på gymnasiet är andelen högre, den andra omdömningen styrker i genomsnitt ursprungsbedömaren i 22 procent av fallen.



Figur 23: Resultat från Skolinspektionens andra omdömning av delprov i svenska, efter årskurs/kurs. Den andra omdömningen görs endast för 1) prov i svenska där det 2) var minst två betygsstegs skillnad mellan den ursprungliga bedömningen av skolan och Skolinspektionens första omdömning. Det är alltså endast dessa delprov som ingår i redovisningen (gäller även totalen).

Den låga samstämmigheten kan delvis förklaras av att elevlösningar som går till den andra omdömningen rimligtvis är svårare att bedöma. Det är troligt att överensstämmelsen mellan Skolinspektionens omdömare skulle öka om samtliga elevlösningar ingick i jämförelsen.

Utvalda delprov till omdömningen av nationella prov består till stor del av frågor med fritextsvar, där det inte finns ett tydligt rätt eller fel svar. Det finns med andra ord en stor komplexitet i att bedöma dessa delprov. Jämförelsen mellan de delprov som har fått två av varandra oberoende bedömningar av Skolinspektionens omdömare förtydligar bilden och belyser svårigheten i att bedöma delproven likvärdigt.

## 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, faktiskt blir bedömda utifrån samma kriterier.

### Bristande likvärdighet i bedömningarna av de nationella proven

Resultaten i rapporten visar att mer än hälften av elevlösningarna får olika bedömningar av skolans lärare och omedbarna. För många delprov skiljer det dessutom två betygssteg eller mer. Eftersom endast ett delprov per prov omedboms 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 för ett delprov för att det totala provbetyget ska påverkas. Det står alltså klart att de avvikelser som uppmäts i omedbningen på delprovsnivå är av den storleksordningen att det *kan* leda till att det totala provbetyget påverkas.

### Bedömningsprocessen har betydelse för likvärdigheten

Resultaten från omedbningen visar att sambedömning och extern bedömning ökar samstämmigheten mellan bedömningarna. För de delprov där sambedömning, extern bedömning och avidentifiering har använts är andelen delprov som bedömts generösare av skolan lägre jämfört med övriga delprov. Effekten tycks öka när flera bedömningsmetoder kombineras.

Resultaten från omedbningen visar att inslagen av sambedömning, extern bedömning och avidentifiering varierar kraftigt både 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 bedömningen. De goda exempel på bedömningsmetoder som den här rapporten visar på ger möjlighet att öka likvärdigheten och stärka rättssäkerheten i bedömningarna, i synnerhet om de sker mer enhetligt mellan landets skolor.

### Svårt att göra likvärdiga 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 återstår stora skillnader mellan bedömningarna även för delprov där dessa metoder har tillämpats. Jämförelsen mellan de delprov som har fått två oberoende bedömningar av Skolinspektionens omedbmare belyser komplexiteten i att bedöma proven likvärdigt. Det är rimligt att anta att det är de mest svårbedömda elevlösningarna som får en andra omedbning, resultatet kan därför inte generaliseras till alla elevlösningar som ingår i omedbningen. Det faktum att det för dessa elevlösningar är vanligt att bedömningarna även skiljer sig mellan omedbarna, som inte har en relation till eleven eller skolan, indikerar att delprov som till stor del består av fritextsvar ofta är svåra att bedöma på ett likvärdigt sätt.

## Avslutningsvis

Sedan 2009 har Skolinspektionen årligen fått i uppdrag av regeringen att omdedöma nationella prov. I rapporterna har Skolinspektionen lyft slutsatser och utvecklingsområden för ökad likvärdighet i bedömningarna av proven som varit till nytta i flera olika sammanhang. Resultaten från tidigare omdedömningar har bland annat refererats till i betänkandet *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning*<sup>11</sup> och i Regeringens proposition *Nationella prov – rättvisa, likvärdiga, digitala*<sup>12</sup>.

Skolinspektionen ser det som angeläget att även fortsatt bibehålla en statlig granskning av skolors och huvudmäns arbete med bedömning och betygssättning för att säkerställa elevernas rätt till likvärdighet i detta avseende.

Dock har de övergripande slutsatserna för omdedömningsuppdraget varit återkommande från år till år. I och med att resultaten är likartade vid varje genomförd omdedömning är mervärdet av en fortsatt omdedömning begränsat. Omdedömningens övergripande syfte, att stödja en likvärdig bedömning och betygssättning av elevers prestationer, kan mötas på ett mer verksamt och kostnadseffektivt sätt genom andra former av granskningar. Med utgångspunkt i detta har Skolinspektionen genom tidigare skrivelser, *Ökad kostnadseffektivitet i granskning av likvärdighet i bedömning och betygssättning*<sup>13</sup> och *Budgetunderlag för åren 2017–2019*<sup>14</sup>, föreslagit att uppdraget om central rättning av nationella prov bör ändras så att arbetet med likvärdighet i bedömning och betygssättning förstärks genom att bedrivs som kvalitetsgranskningar.

Skolverket har i uppdrag att digitalisera de nationella proven.<sup>15</sup> Om omdedömningsuppdraget till Skolinspektionen ska fortgå under och efter digitaliseringens införande behöver omdedömningsuppdraget anpassas därefter. Digitaliseringen medför nya förutsättningar samt ett stort omställningsarbete för Skolinspektionen, främst i form av utveckling av IT-lösning. Även frågor om sekretess och personuppgiftsbehandling kan behöva utredas. Därför behöver uppdragsbeskrivningen anpassas i god tid. När extern bedömning införs generellt för de nationella proven behöver även detta tas hänsyn till i uppdragsbeskrivningen. Om proven bedöms av lärare utanför skolan kommer resultatet från omdedömningen inte längre på samma sätt vara relevant för Skolinspektionens tillsynsverksamhet på skolenhetsnivå.

---

<sup>11</sup> SOU 2016:25.

<sup>12</sup> Prop. 2017/18:14.

<sup>13</sup> Skolinspektionen. *PM – Ökad kostnadseffektivitet i granskningen av likvärdighet i bedömning och betygssättning*.

<sup>14</sup> Skolinspektionen. *Missiv – Statens skolinspektions budgetunderlag 2017–2019 samt regelförteckning*.

<sup>15</sup> Regeringen (2017). *Uppdrag att digitalisera de nationella proven m.m.*

# Bilagor

## Bilaga 1

### Användning av resultaten i tillsyn och kvalitetsgranskning

I enlighet med 2017 års regleringsbrev<sup>16</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>17</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 arbetet som leder fram till bedömningen och vidare betygssättningen som granskas i Skolinspektionens inspektion, inte delprovbetygen i sig. Avviker en skolens 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 arbete med bedömning och betygssättning extra noga vid tillsyn och skolan blir särskilt utvald för en 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 ribban 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>16</sup> Utbildningsdepartementet. *Regleringsbrev för budgetåret 2017 avseende Statens skolinspektion.*

<sup>17</sup> Utbildningsdepartementet. *Regleringsbrev för budgetåret 2015 avseende Statens Skolinspektion.*

## Att använda resultaten i regelbunden tillsyn, regelbunden kvalitetsgranskning och tematisk 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 granskning av den aktuella skolenheten.

Inom tillsyn och regelbunden kvalitetsgranskning utgör resultaten från ombedömningen ett av flera underlag som analyseras inför ett skolbesök och som används i utredningen. 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>18</sup> Urvalet av skolor som ingår i de tematiska kvalitetsgranskningarna dras ofta utifrån resultaten från ombedömningen av nationella prov i kombination med resultatet från Skolverkets jämförelser mellan nationella provbetyg och ämnes- eller kursbetyg. För närvarande pågår en kvalitetsgranskning om kvaliteten på betygssättningen med fokus på lärarnas arbete. Kvalitetsgranskningen omfattar gymnasiekursen svenska 3 på skolor med högskoleförberedande gymnasieprogram.

## Skolenhetsrapporter och huvudmannasammanställningar

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 publiceras på Skolverkets hemsida.

Huvudmännen för de skolenheter som medverkat i ombedömningen av nationella prov får en sammanställning med resultaten för samtliga av huvudmannens skolenheter som deltagit i ombedömningen och som fått ett resultat i form av en skolenhetsrapport.

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>18</sup> Se till exempel Skolinspektionen (2018): Rättvis och likvärdig betygssättning i grundskolan? Skolinspektionen (2014): Uppenbar risk för felaktiga betyg samt Skolinspektionen (2014): Förstärkt tillsyn av skolors arbete med bedömning och betygssättning.

## Bilaga 2

# Genomförandet av omdedömningen

### Urval av prov och skolor

2017 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 2016.

Skolor som ingår i omdedömningen väljs ut genom ett slumpvist urval bland skolor vars huvudmän följande år ingår i Skolinspektionens regelbundna tillsyn och granskning. Urvalet av skolor för omdedömning år 2017 gjordes bland skolor vars huvudmän ingår i tillsynsplanen för 2018/2019. 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.

### Ämnen, delprov och årskurser som ingår i omdedömningen

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, om än inte omöjliga att omdedöma. Tidigare erfarenhet har även visat att större avvikelser är mycket sällsynta för denna typ av delprov.

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 skrivfrågor, där eleven fritt ska formulera sig utifrån en öppen frågeställning.

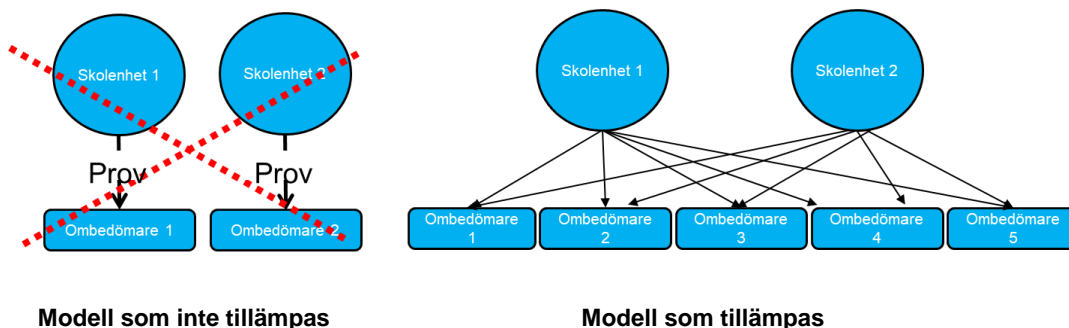
## Urvalets sammansättning

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 52 procent för grundskolor och 59 procent för gymnasieskolor. Vid tiden då urvalet gjordes uppgick andelen grundskolor respektive gymnasieskolor med enskild huvudman till 17 procent respektive 32 procent av det totala antalet grund- respektive gymnasieskolor i landet. Anledningen att fristående skolor är överrepresenterade i urvalet till omdömningen är att urvalet är kopplat till Skolinspektionens tillsyns- och granskningsprocesser.

## Fördelning av prov mellan omdömare

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

Skolinspektionens lärare omdömer prov enskilt. Samedömning 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 5 och 20 olika omdömare.



Figur 24: 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, där det skiljer mer än två betygssteg mellan bedömning och omdömning, går provet 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 proven.

## Omfattning och bortfall

Totalt 23 458 delprov från 371 skolor har omdömts i 2017 års omgång. Ombedömningen genomfördes av totalt 110 lärare som anlätts av Skolinspektionen. Fördelningen över ämne och delprov framgår av tabell 1 nedan.

**Tabell 1:** 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.

Ämne och delprov:	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 omdömare
Engelska åk 6 delprov C	3 298	117	3 184	116	3 %	3 135	116	2 %	5 %	9
Engelska åk 9 delprov C	3 098	65	3 016	65	3 %	2 960	65	2 %	4 %	20
Engelska Gy delprov C	3 948	67	3 638	66	8 %	3 574	66	2 %	9 %	18
Engelska Vux delprov C	852	48	724	45	15 %	663	45	8 %	22 %	7
Svenska åk 6 delprov C1+C2	3 280	117	3 141	116	4 %	3 058	116	3 %	7 %	12
Svenska åk 9 delprov C	3 038	64	2 696	62	11 %	2 570	61	5 %	15 %	18
Svenska Gy delprov C	4 015	68	3 768	68	6 %	3 673	68	3 %	9 %	20
Svenska Vux delprov C	1 860	53	1 596	53	14 %	1 492	53	7 %	20 %	6
Biologi åk 9 delprov A2	913	33	825	31	10 %	787	31	5 %	14 %	6
Fysik åk 9 delprov A2	801	27	654	24	18 %	646	24	1 %	19 %	5
Kemi åk 9 delprov A2	963	29	918	29	5 %	900	29	2 %	7 %	5
<b>Totalt:</b>	<b>26 066</b>	<b>374</b>	<b>24 160</b>	<b>371</b>	<b>7 %</b>	<b>23 458</b>	<b>371</b>	<b>3 %</b>	<b>10 %</b>	<b>110</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 provkopiorna 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?

Det sammanvägda bortfallet varierar från delprov till delprov. Orsaken till varför bortfallet av delprov är stort i den kommunala vuxenutbildningen kan delvis förklaras av 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 2 redovisas antal svar per fråga. Bortfallet utgörs av skolor som inkom med provkopior 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 2:** Bortfall vid frågor om bedömningens genomförande på skolorna.

Frågor om bedömningen:	1. Vem bedömde provet?			2. Var provet avidentifierat vid bedömningen?			3. Diskuterades bedömningen med annan pedagogisk personal?		
	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	2 295	91	27 %	2 876	110	8 %	2 940	112	6 %
Engelska åk 9 delprov C	2 310	54	22 %	2 820	63	5 %	2 731	62	8 %
Engelska Gy delprov C	3 247	60	9 %	3 538	66	1 %	3 459	64	3 %
Svenska åk 6 delprov C1+C2	2 130	89	30 %	2 785	111	9 %	2 899	112	5 %
Svenska åk 9 delprov C	2 070	55	19 %	2 538	60	1 %	2 444	59	5 %
Svenska Gy delprov C	3 133	60	15 %	3 599	68	2 %	3 603	67	2 %
Biologi åk 9 delprov A2	696	29	12 %	774	31	2 %	774	31	2 %
Fysik åk 9 delprov A2	561	22	13 %	598	23	7 %	599	23	7 %
Kemi åk 9 delprov A2	741	25	18 %	829	27	8 %	829	27	8 %

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

## Bilaga 3

# Ett harmoniserat mått för att mäta skillnader i bedömning

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

För delproven i svenska och engelska sätts betyg på skalan A–F, men för delprov inom NO-ämnena är det endast det samlade provbetyget som anges på skalan A–F. För de olika delproven erhålls i stället ett antal *poäng* på E-, C- respektive A-nivå (för NO-proven är *belägg* det ord som används). Hur många poäng/belägg som kan uppnås på respektive nivå varierar mellan olika ämnen och delprov.

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

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

## Beräkning av det harmoniserade resultatmättet

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

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

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

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

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

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

Det är viktigt att komma ihåg att det harmoniserade resultatmåttet 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 mellan detta mått och ett faktiskt betyg på skalan A–F, eftersom reglerna för hur belägg summeras till ett samlat provbetyg ser olika ut för olika prov och dessutom involverar summering av belägg från flera olika delprov.

# Referenslista

Proposition 2017/18:14. *Nationella prov – rättvisa, likvärdiga, digitala.*

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

SFS 2010:800. *Skollagen.*

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

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

Skolinspektionen (2015). *PM – Ökad kostnadseffektivitet i granskningen av likvärdighet i bedömning och betygssättning.* Dnr 2015:514.

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

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

Skolverket (2013). *Sambedömning i skolan – exempel och forskning.*

Skolverket (2018). *Skolverkets allmänna råd med kommentarer – Betyg och betygssättning.*

Skolverket. Hämtat 11 oktober 2018. [<https://www.skolverket.se/undervisning/gymnasieskolan/nationella-prov-i-gymnasieskolan/genomfora-och-bedoma-prov-i-gymnasieskolan#Provresultatetsbetydelseforbetyget>]

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

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

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