

# Olikheterna är för stora

Omrättning av nationella prov i grundskolan och gymnasieskolan, 2013



2013-09-02

## Förord

Statens skolinspektion fick den 19 augusti 2009 i uppdrag av regeringen att genomföra viss central rättning av nationella prov för grundskolan och gymnasieskolan.

Skolinspektionen redovisar här till regeringen resultatet av fjärde årets omrättning. Redovisningen omfattar en presentation och analys av resultatet för det senaste året. Rapporten innefattar även en analys av nåbar bedömaröverensstämmelse vid bedömningen av dels längre uppsatsdelar, dels kortsvarsdelar i de nationella proven och hur väl skolorna som ingått i omrättningen når upp till dessa nivåer.

Vid Skolinspektionen har följande personer arbetat med att genomföra uppdraget: Klara Klingspor (projektledare), Jonna Andersson (gruppledare), utredarna Emmie Chau, Jonathan Hedlund, Jessica Hörnsten och Mattias Svantesson.

Ann-Marie Begler

Generaldirektör

Gabriella Bremberg

Avdelningschef

## Innehållsförteckning

<b>INNEHÅLLSFÖRTECKNING</b> .....	<b>3</b>
<b>1. SAMMANFATTNING – STÖRST AVVIKELSER FÖR UPPSATSDELAR</b> .....	<b>5</b>
<b>2. INLEDNING</b> .....	<b>9</b>
UPPDRAGET .....	11
TERMINOLOGIN .....	11
<b>3. RESULTAT – STÖRST AVVIKELSER FÖR UPPSATSDELAR</b> .....	<b>13</b>
UPPSATSDELAR I GYMNASIESKOLAN AVVIKER ALLRA MEST.....	14
DE NEGATIVA AVVIKELSERNA DOMINERAR .....	15
DE HÖGRE BETYGSSTEGEN STÅR FÖR DE STÖRSTA AVVIKELSERNA.....	16
FÄRRE AVVIKELSER I LÄGRE ÅRSKURSER .....	17
NYA BEDÖMNINGAR VANLIGAST VID UPPSATSPROV .....	18
AVVIKELSER VARIERAR KRAFTIGT MELLAN SKOLOR .....	19
SAMBEDÖMNING PÅVERKAR AVVIKELSEN .....	22
OLIKA SKRIVUPPGIFTER PÅVERKAR INTE AVVIKELSEN .....	24
<b>4. NÄR ÄR OLIKHETERNA FÖR STORA?</b> .....	<b>25</b>
ELEVPERSEKTIV – ELEVEN RÄTT TILL EN LIKVÄRDIG BEDÖMNING .....	25
UPPSATSPROV .....	26
KORTSVARSPROV.....	27
OBEFOGADE AVVIKELSER - DET FINNS EN FÖRBÄTTRINGSPOTENTIAL HOS SKOLORNA.....	29
<b>5. REKOMMENDATIONER KRING UPPSATSDELARNA</b> .....	<b>31</b>
UTREDNING FÖRESLÅS FÖR ATT GENOMLYSA BEHOVET AV UPPSATSDELAR.....	32
EN KOMBINATION AV ÅTGÄRDER FÖR ATT ÖKA LIKVÄRDIGHETEN .....	32
SKOLVERKET BÖR FORTSÄTTA FÖRTYDLIGA BEDÖMNINGSANVISNINGARNA .....	33
LÄRARE MÅSTE UTBILDA SIG I BEDÖMNING .....	34
SAMBEDÖMNING SÅ LÅNGT DET ÄR MÖJLIGT .....	34
BEDÖMA PROV DÄR ELEVEN IDENTITET INTE ÄR KÄND FÖR LÄRAREN .....	35
<b>6. REKOMMENDATIONER INFÖR KOMMANDE OMRÄTTNINGAR</b> .....	<b>36</b>
<b>7. URVAL INFÖR KOMMANDE OMRÄTTNINGAR</b> .....	<b>37</b>
<b>8. GENOMFÖRANDE AV UPPDRAGET</b> .....	<b>37</b>
TILLVÄGAGÅNGSSÄTT .....	38
BORTFALL .....	41
<b>9. KÄLLOR</b> .....	<b>45</b>

**10. DIAGRAM- OCH TABELLFÖRTECKNING..... 48**

## 1. Sammanfattning – störst avvikelser för uppsatsdelar

År 2013 har Skolinspektionen omrättat 31 800 elevlösningar av nationella prov på delprovsnivå från 461 skolor. Den här rapporten beskriver resultatet från denna omrättning, vilket är den fjärde genomförda. Totalt har under fyra år 128 000 prov omrättats.

I omrättningen undersöker Skolinspektionen om det delprovsbetyg som en elev fick på det nationella provet skulle bli detsamma om en annan lärare bedömde provet.

Skolinspektionen konstaterar att avvikelserna mellan ursprungsbedömning och omrättning fortsatt är mycket stora i vissa fall. Dessutom är variationerna mellan skolorna stora. Det bör vara möjligt att med olika typer av insatser i skolorna öka säkerheten i bedömningsförfarandet.

Nytt för denna omrättningsomgång är en analys av vad som är att betrakta som en normal respektive obefogad avvikelse.

### När är olikheterna för stora?

Det ideala för eleverna är en total samstämmighet när olika lärare bedömer samma provsvar, det vill säga 100 procents överensstämmelse mellan bedömningarna. Men eftersom vissa prov är förhållandevis komplexa, det finns inte ett rätt och ett fel, så är det realistiskt att tro att en exakt överensstämmelse kan uppnås.

Skolinspektionen bedömer, utifrån forskning och egna resultat, att skolorna bör utveckla sitt arbete med bedömning av uppsatsdelarna i de nationella proven, så att man i efterhand ska kunna konstatera en överensstämmelse på 70 procent. För kortsvarsdelarna i proven bör man kunna nå över 80 procents bedömaröverensstämmelse vid de mer svårbedömda delarna och över 90 procent bedömaröverensstämmelse vid de mer lättbedömda delarna.

Resultatet från årets omrättning visar att avvikelserna gällande uppsatsdelar och kortsvarsdelar i nationella proven i svenska och engelska i många fall kan ses som obefogad.

Avvikelsen bör, enligt Skolinspektionen, ses som obefogad om den i efterhand kan konstateras var högre än 30 procent vid uppsatsdelar och 20 eller 10 procent vid kortsvarsdelar.

### **Stor förbättringspotential hos skolor i arbetet med bedömning**

Mot bakgrund av ovanstående menar Skolinspektionen att skolorna bör bedriva ett aktivt arbete för att bättra sin förmåga att bedöma proven t.ex. genom fortbildning och sambedömning, så att man i efterhand kan konstatera att man nått en högre överensstämmelse. Det finns helt klart ett utrymme för förbättring.

Till exempel har 95 procent av grund- och gymnasieskolorna, enligt analyser Skolinspektionen har genomfört av omrätningsresultatet, obefogade avvikelser vid uppsatsdelar. Vid de mer svårbedömda kortsvarsdelarna har 91 procent av skolorna en obefogad avvikelse. För de mer lättbedömda kortsvarsdelarna är det 73 procent av skolorna som har en obefogad avvikelse.

Förbättringspotentialen är minst 20 procentenheter vad gäller uppsatsdelar. För kortsvarsdelarna är förbättringspotentialen inte lika stor men varierar mer mellan olika delprov.

En del skolor som ingått i årets omrätningsomgång har kommit relativt långt och har låga andelar avvikelser. Men de skolor som har obefogade avvikelser bör arbeta med att uppnå mer likvärdiga och rättssäkra bedömningar.

Ur ett elevperspektiv är omfattande avvikelser av stor betydelse. Om samma provsvar bedöms så olika kan det på sikt handla om att få ett felaktigt slutbetyg<sup>1</sup>, något som kan påverka möjligheterna för framtida studie- och yrkesval.

### **Större överensstämmelse vid delprov med kortsvar**

Tittar man på andelen avvikelser prov för prov spänner avvikelserna i den senaste omrättningen mellan 11 och 62 procent.<sup>2</sup> Avvikelseerna är avsevärt mindre för de kortsvarsprov där svaren har karaktären av rätt-eller-fel eller innehåller kortare textproduktion. Det gäller bland annat delproven B i engelska för årskurs 9 och gymnasieskolan, som handlar om att läsa kortare texter, lyssna på till exempel radioprogram och nyhetsinslag och välja olika svarsalternativ vid flervalsfrågor.

---

<sup>1</sup> Läraren använder sig av många olika underlag när ett betyg sätts, varav resultaten från nationella proven, är ett. Det som undersöks i denna rapport är avvikelserna i de nationella proven.

<sup>2</sup> Engelska i årskurs 9 har för delprovet B en total avvikelse mellan ursprungs rättaren och omrätaren på 11 procent och svenska1/svenska som andra språk<sup>1</sup>, delprov C har en avvikelse på 62 procent.

### **Oftast ger omrättaren eleven ett lägre betyg**

Omrättaren bedömer oftast att elevens prestation motsvarar ett lägre betyg än vad ursprungs rättaren har bedömt; det vill säga att den nya bedömningen oftast går i negativ riktning. Spridningen är större i gymnasieskolan, men då är också betygsskalan sexgradig (A-F) jämfört med grundskolan, där betygsskalan är fyrgradig (Icke godkänt-MVG). Vi ser även tydligt utifrån denna omgång att det är de högre betygsstegen (VG, MVG, B och A) som lärarna överlag är minst ense om.

### **Stora variationer mellan skolorna**

Avvikelse rna varierar fortfarande kraftigt mellan olika skolor. Den skola där de största skillnaderna finns avviker mer än tjugo gånger så ofta som den skola som har minst antal avvikelser.

### **Forskare har analyserat faktorer kring vad som kan påverka avvikelsen**

Skolinspektionen har låtit Björn Tyrefors Hinnerich och Jonas Vlachos vid Stockholms universitet analysera förekomsten av systematiska skillnader i bedömningarna.<sup>3</sup>

När det gäller elevens bakgrund finns en tendens att ursprungsrättarna jämfört med omrättarna sätter lägre provbetyg på pojkar, elever med utländsk bakgrund, elever vars föräldrar saknar högre utbildning och elever som är födda sent på året. Samtliga dessa grupper är sådana som generellt tenderar att uppvisa relativt svaga studieresultat. Gymnasieelever med höga grundskolebetyg tenderar att få en generösare provbedömning av ursprungsrättare jämfört med omrättares bedömning. Skillnaderna är överlag små till medelstora. Mönstret är inte helt likartat mellan grundskolan och gymnasieskolan.

### **Skolinspektionens rekommendationer**

Skolinspektionens resultat visar att det för uppsatsdelarna finns stora skillnader i bedömning av olika omrättare.

Skolinspektionens bedömning är därför att delprov med uppsatsdelar under rådande förhållanden fungerar dåligt för att stödja likvärdigheten som är ett av två huvudsyften med de nationella proven. Frågan är om uppsatsdelarna ens kan användas för att uppnå likvärdighet. Skolinspektionen menar dock att uppsatserna är viktiga av andra skäl, dock inte för att säkerställa likvärdigheten i betygssättningen.

---

<sup>3</sup> Redovisningen återfinns i sin helhet i resultatbilagan.

Om uppsatsdelarna helt skulle lyftas bort ur nationella prov skulle provet i sin helhet, om det består av kortsvars- och flervalssdelar, bli mer likvärdigt. Samtidigt finns det nackdelar med att inte ha kvar uppsatsdelar. Om de tas bort bör de ersättas med andra typer av prov som kan pröva elevers mer kvalificerade skrivförmåga, som är viktiga delar i kursplanen.

Skolinspektionen föreslår därför att Regeringen överväger att **tillsätta en utredning kring för- och nackdelar med att låta uppsatsdelar ingå i de nationella proven.**

Skolinspektionen anser också att en kombination av åtgärder bör vidtas för att öka samstämmigheten mellan olika bedömningar vad gäller uppsatsdelarna.

Skolverket bör fortsätta med att förtydliga och tillgängliggöra **bedömningsanvisningarna** framförallt i de delar som gäller de högre betygsstegen. Vidare vill Skolinspektionen understryka vikten av att lärare av skolhuvudmän erbjuds kontinuerlig **grund- och fortbildning** i att bedöma. Rektorer bör samtidigt se till att **sambedömning** används så långt det är möjligt.<sup>4</sup> Skolinspektionen föreslår även ett bedömningsförfarande där **elevens identitet är okänd för lärare som bedömer prov.**

Skolinspektionens bedömning är att dessa åtgärder skulle kunna bidra till ökad samstämmighet men att det ändå skulle kvarstå stora differenser mellan olika lärares bedömning framförallt i bedömningen av längre elevtexter.

---

<sup>4</sup> Med sambedömning avses att fler än en lärare bedömer samma prov eller ger sin syn på en aktuell bedömning. Syftet är att öka säkerheten i bedömningen av elevens prestation och minska den enskilda lärarens subjektiva inslag, vilket stärker likvärdigheten i bedömningen.



## 2. Inledning

Enligt Skolverkets bedömning har likvärdigheten mellan Sveriges skolor försämrats under de senaste åren. Slutsatsen bygger på att variationen mellan skilda skolors resultat har ökat och att skolans elevsammansättning spelar en allt större roll för en elevs resultat.<sup>5</sup> Ett av huvudsyftena med de nationella proven är att stödja en likvärdig och rättvis bedömning och betygssättning.

### Är bedömningen likvärdig? – undersökningen fokuserar på avvikelse

Skolinspektionen fick i uppdrag av regeringen att skapa ett "tillförlitligt underlag för att bedöma i vilken utsträckning proven rättas korrekt och likvärdigt". Regeringen konstaterar i uppdragsbeskrivningen till Skolinspektionen att: "Två viktiga syften med de nationella proven är att de ska ge underlag för lärarens betygssättning i ett ämne och att de ska bidra till en likvärdig bedömning över landet. Det finns idag inget tillförlitligt underlag för att bedöma i vilken utsträckning proven rättas korrekt och likvärdigt". I uppdragsbeskrivningen skriver regeringen att:

*Syftet med en central rättning av nationella prov är främst att stödja en likvärdig bedömning och betygssättning av proven över landet. Skolinspektionen ska vidare, utifrån indikationer på betydande avvikelser när det gäller dels rättning av nationella prov, dels betygssättning på nationella prov och i ämnen och kurser, vidta lämpliga åtgärder inom ramen för sin tillsynsverksamhet.*

### Forskare har analyserat faktorer kring vad som kan förklara avvikelsen

Skolinspektionen har låtit forskarna Björn Tyrefors Hinnerich och Jonas Vlachos vid Stockholms universitet analysera förekomsten av systematiska skillnader mellan bedömningar av ursprungsrättare och omrättare av elevernas resultat på nationella prov. Vidare analyseras hur lärartäthet och pedagogisk examen samt huvudmannatyp och konkurrens påverkar avvikelserna. Deras redovisning återfinns i sin helhet i resultatbilagan.

### Omrättare har bra förutsättningar för att göra bedömningar

När det gäller resultaten i denna rapport avgör Skolinspektionen inte om "felbedömningarna" ligger hos de lärare som bedömde proven i första hand eller om de ligger hos de lärare som omrättat proven. Vi kan inte säga att de lärare som av Skolinspektionen fått uppdraget att kontrollrätta elevlösningarna är mer korrekta i sina bedömningar än de lärare som ursprungligen bedömt pro-

---

<sup>5</sup> Skolverket, 2012, s 85.

ven. Vad vi dock vet är att samtliga lärare som deltagit i omrättningen är behöriga i de ämnen de bedömer prov i, de har erfarenhet av att bedöma nationella prov och att de disponerar lång tid för sin uppgift.<sup>6</sup> Därför anser vi att omrättarna har bra förutsättningar för att göra bedömningarna. Flertalet av de lärare som har omrättat prov har medverkat vid samtliga fyra omrättningsomgångar.

---

<sup>6</sup> Ursprungs rättarna har en vecka på sig att bedöma sina prov, omrättarna har drygt tre månader på sig att bedöma upp till 400 prov.

## Uppdraget

Uppdragsbeskrivningen finns i sin helhet i bilaga 6, men kan i korthet sammanfattas till att Skolinspektionen ska:

- ansvara för att ett urval av elevlösningar av nationella prov i grundskolan och gymnasieskolan ska rättas centralt, efter det att de har rättats en första gång på respektive skola
- göra analyser med avseende på bland annat systematiska avvikelser
- använda sig av resultaten från denna analys som underlag för fortsatt tillsyn och kvalitetsgranskning

En kort beskrivning av förändringar vid kommande omrättning, omgång 5, beskrivs i kapitel 6.

## Terminologin

Regeringen ger i sin uppdragsbeskrivning Skolinspektionen i uppdrag att genomföra "viss central rättning" av nationella prov. Skolinspektionen använder i rapporten begreppet *omrättning* i stället för *central rättning*.

Skolinspektionen har genomfört omrättning 2010, 2011, 2012 och 2013. De prov som ingick genomfördes 2009, 2010, 2011 och 2012. För enkelhetens skull kommer vi att benämna första omrättningen *omgång 1*, den andra *omgång 2*, den tredje *omgång 3* och slutligen den fjärde *omgång 4*.

Resultaten av omrättningen redovisas som avvikelser mellan ursprungs rättarens bedömning och omrättarens bedömning. Med *avvikelse* avses en skillnad i bedömning av elevens prestation med minst ett betygssteg. För de delprov där provbetyget inte kan användas för jämförelse utgör avvikelsen skillnad i poängsättning.

*Positiv avvikelse* innebär att omrättaren bedömer att elevens svar motsvarar högre betyg/poäng än vad ursprungs rättaren har bedömt.

*Negativ avvikelse* innebär att kontrollrättaren bedömer att elevens svar motsvarar lägre betyg/poäng än vad ursprungsrättaren har bedömt.

För att underlätta läsningen av rapporten redovisas delproven ämnesvis som *uppsatsdel* och *kortsvarsdel* löpande i texten, diagrammen och tabellerna istället för till exempel enbart delprov B och C. Det är en förenkling och kategorisering utifrån vad proven till största delen innehåller. Kortsvarsdelarna i sig kan vara olika komplexa att bedöma och kan exempelvis innefatta flervalfrågor eller frågor där man enbart ska fylla i ett eller några enstaka ord som svar. Det förekommer även kortare öppna frågor i olika omfattning och längd. I anslutning till det förenklade grupperingsnamnet finns även delprovsbokstaven utskrivna. Vad delproven detaljerat innehåller framgår i resultatbilagan.

### 3. Resultat – störst avvikelser för uppsatsdelar

År 2013 har Skolinspektionen omrättat 31 800 elevlösningar av nationella prov på delprovsnivå från 461 skolor.

#### Resultaten i korthet:

- De största avvikelserna återfinns för de delprov där eleven ska avge sitt svar i längre text – de så kallade uppsatsdelarna. Till exempel för uppsatsdelen i svenska i gymnasieskolan (delprov C) avviker 62 procent av bedömningarna.
- För samtliga omrättade delprov är det vanligare att ursprungsrättaren har satt ett högre betyg eller bedömt att eleven i större utsträckning har nått kravnivån än vad omrättaren gjort. Dessa så kallade negativa avvikelser är störst för uppsatsdelarna i svenska i årskurs 9 och gymnasieskolan (delprov C), med 45 respektive 34 procent avvikelse.
- Det är de högre betygsstegen som står för den största andelen avvikelser. Samstämmigheten är högre vid lägre betyg. Vid till exempel uppsatsdelen i svenska i årskurs 9 (delprov C) är samstämmigheten i bedömning 62 procent för betyget Icke godkänt (IG) och 18 procent för betyget MVG.
- Andelen avvikelser varierar starkt mellan olika skolor – från 0 till 89 procent. Spridningen är större för uppsatsdelarna än för kortsvardsdelarna i grundskolan (årskurs 9). För uppsatsdelarna varierar spridningen på skolorna mellan 31 och 71 procent.

Proven som omrättats avser svenska/svenska som andraspråk i årskurs 6<sup>7</sup> samt ämnena engelska, matematik och svenska/svenska som andraspråk och de naturorienterade ämnena (biologi, fysik och kemi) i årskurs 9.<sup>8</sup> I gymnasieskolan

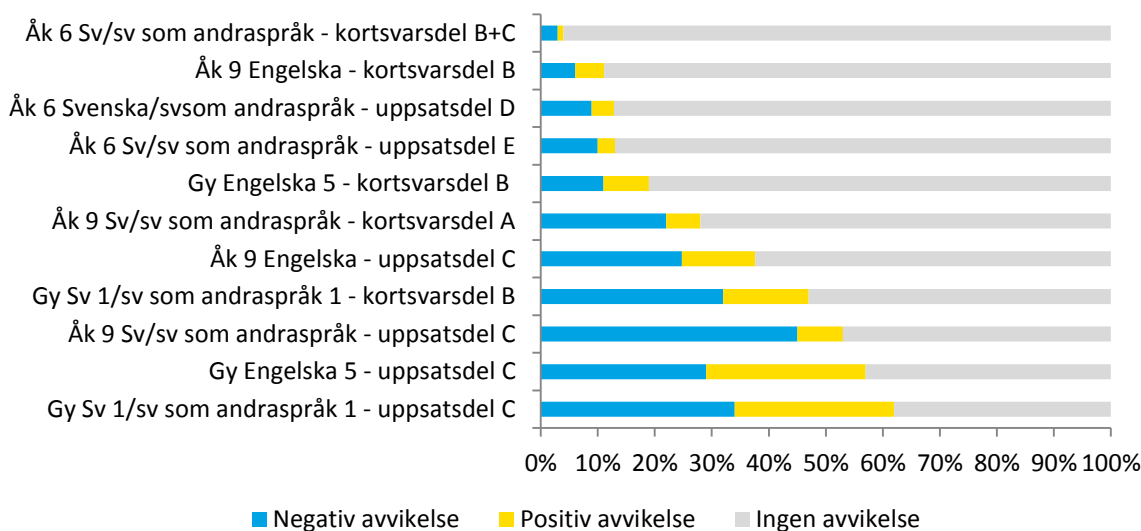
<sup>7</sup> Nationella prov för årskurs 6 genomfördes för första gången 2012 och resultat från en analys av dessa data beskrivs framöver i denna rapport.

<sup>8</sup> I den fjärde omrättningsomgången har inte prov i årskurs tre ingått, vilket de gjort i vissa tidigare omgångar. Anledningen är att för årskurs tre har avvikelserna mellan ursprungsrättningar och omrättningar generellt varit mycket låga.

ingår kursproven engelska 5, matematik 1a, 1b och 1c samt svenska 1/svenska som andraspråk 1. Proven genomfördes under vårterminen 2012.

### Uppsatsdelar i gymnasieskolan avviker allra mest

De flesta avvikelserna finns bland uppsatsdelarna i ämnena svenska och engelska i gymnasieskolan och årskurs 9, som syns i diagram 1 nedan. Avvikelserna är något högre i gymnasieskolan än i grundskolan, vilket kan förklaras av att skrivuppgifterna i uppsatsdelarna blir mer komplicerade och krävande utifrån elevernas skrivvana och stigande ålder.



**Diagram 1: Fördelning av andel avvikelser; positiva, negativa samt ingen avvikelse mellan ursprungsrättarens bedömning och Skolinspektionens bedömning för de olika delproven, omgång 4.**

En jämförelse mellan ämnena visar att uppsatserna i engelska ger något färre avvikelser än de i svenska/svenska som andraspråk. Negativa avvikelser mellan ursprungsrättarens och omrättarens bedömning är störst för uppsatsdelarna i svenska i årskurs 9 och gymnasieskolan (delprov C), med 45 respektive 34 procent avvikelse.

Totalt avviker cirka en tredjedel av proven i ämnena engelska och svenska/svenska som andraspråk. Vid 69 procent av bedömningarna är det inga avvikelser. Generellt ser vi mycket snarlika resultat över de tre första om-

rättningsomgångarna där den totala andelen avvikande prov i stort sett har legat på samma nivå.<sup>9</sup>

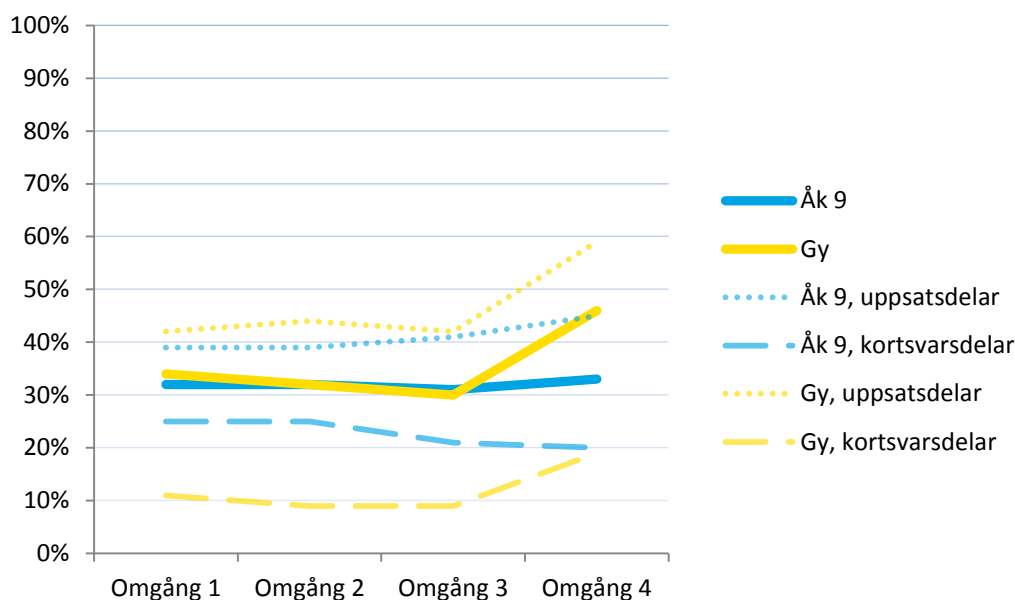


Diagram 2: Total andel avvikelse för de nationella proven i svenska och engelska i årskurs 9 och gymnasieskolan, omgång 1-4.

Mellan omgång tre och fyra ökar andelen avvikelse, främst vid gymnasieskolorna. En anledning till det är att betygsskalan ändrats och nu omfattar sex betygssteg, istället för som tidigare fyra betygssteg. Med en bedömnings skala i fler steg kan fler avvikelser uppkomma.

### De negativa avvikelserna dominerar

Det är vanligare att ursprungs rättaren har satt ett högre betyg eller bedömt att eleven i större utsträckning har nått kravnivån än vad omrättaren gjort. För i stort sett samtliga omrättade delprov är de negativa avvikelserna (upp till 45 procent) långt vanligare än de positiva (upp till 28 procent), se diagram 1 ovan. Vid de två delprov som har högst andel total avvikelse: uppsatsdelarna i svenska och engelska i gymnasieskolan (delprov C), är de negativa och positiva avvikelserna ungefär lika stora.

Forskning har visat att betygssättning i allmänhet är influerad av en rad olika faktorer vid sidan av de kunskaper eleven uppvisat.<sup>10</sup> Lärare tar i sina bedöm-

<sup>9</sup> För tidigare resultat se rapporterna Skolinspektionen 2010a, 2010b, 2011 och 2012.

<sup>10</sup> Klapp Lekholm, 2008, *Grades and grade assignment: effects of student and school characteristics*, Göteborgs universitet.

ningar av prov även hänsyn till hur eleven har presterat under terminen eller kursen i övrigt. Läraren kan också känna till elevens sätt att uttrycka sig och därför tolka ett svar på sätt som en utomstående inte skulle göra. Det kan också vara så att läraren omedvetet väger in irrelevanta fakta om eleven, exempelvis svåra hemförhållanden.

### De högre betygsstegen står för de största avvikelserna

När vi tittar på avvikelse mellan ursprungs rättare och omrättare, finner vi att det är de högre betygsstegen (VG, MVG alternativt B och A) som framförallt står för de största avvikelserna. Lärarna ligger genomgående närmare varandra i sina bedömningar för de lägre betygen.<sup>11</sup> Se diagram 3 nedan.

Forskning har pekat på att det för en bedömare är lättare att vara enig om vad som är dåligt än vad som är bra.<sup>12</sup>

Vid till exempel uppsatsdelen i svenska i årskurs 9 (delprov C) avviker 38 procent av bedömningarna vid betyget Icke godkänt (IG). När det gäller betygen VG och MVG är avvikelserna betydligt högre – 62 respektive 82 procent. Avvikelseerna är lägst vid kortsvarsdelen<sup>13</sup> i engelska (delprov B). Vid motsvarande prov i svenska (delprov A) är spridningen större.

---

<sup>11</sup> Det är enbart för ämnena svenska och engelska som en analys varit möjlig att göra. Betygsskalan är sexgradig (A-F) för gymnasieskolan jämfört med grundskolan, där betygsskalan är fyrgradig (IG-MVG) för de prov som omrättats.

<sup>12</sup> Berge, 2005 och Östlund-Stjärnegårdh, 2002.

<sup>13</sup> För att underlätta läsningen av rapporten redovisas delproven ämnesvis som uppsatsdel och kortsvarsdel löpande i texten, diagrammen och tabellerna istället för till exempel delprov B och C. Det är en förenkling och kategorisering utifrån vad proven till största del innehåller.



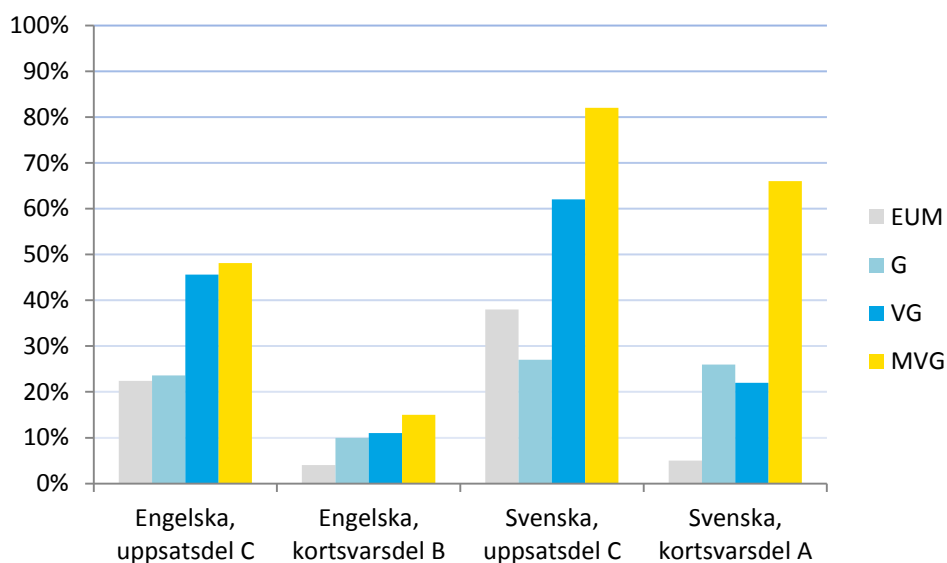


Diagram 3: Andel avvikelse i betygsättning mellan ursprungsrättaren och omrättaren, delprov i svenska och engelska i årskurs 9, omgång 4.

### Färre avvikelser i lägre årskurser

Det är generellt sett färre avvikelser för prov som utförs i årskurs 6 än för de högre årskurserna. Detta gäller särskilt för de delprov som rör svar med fasta alternativ eller tydliga rätt-och-fel-frågor. För kortsvarsdelen i svenska/svenska som andraspråk (delprov B+C) i årskurs 6 ligger den totala andelen avvikelse på 4 procent. Men tendensen är som för andra årskurser liknande, att avvikelserna är större när eleven ska avge sitt svar i form av ett längre svar eller en uppsats. För uppsatsdelarna (delprov E och D) ligger avvikelserna på 13 respektive 11 procent.

En förklaring till att avvikelserna är lägre för de lägre årskurserna är att det till exempel i årskurs 6, vid omrättningarna, enbart finns två steg – att ha uppnått kunskapskraven eller inte.

I gymnasieskolan är avvikelserna störst, vilket sannolikt är en logisk följd av att det där dels är fler betygssteg, dels är mer komplexa bedömningar och mer omfattande svar.<sup>14</sup>

<sup>14</sup> Gunnemyr, 2011, *Likvärdighet till priset av likformighet*. Lunds universitet.

## Nya bedömningar vanligast vid uppsatsprov

Vid de prov där omrättare och ursprungs rättares bedömning avviker mer än ett betygssteg<sup>15</sup> sker en andra omrättning (vid 12 procent<sup>16</sup> av de inskickade proven) och oftast stärks då omrättarens bedömning.<sup>17</sup>

Vid en analys av de prov som omrättades en andra gång skiljer sig bedömaröverensstämmelsen åt beroende på vilket delprov det gäller. Som diagram 4 visar nedan ligger överensstämmelsen på mellan 29 och 61 procent. Ursprungsrättningen stärks i mellan 26 och 40 procent av fallen. I ett av delproven, uppsatsdelen för Engelska i årskurs 9 (delprov C), styrks ursprungsrättarens bedömning i högre grad än omrättarens bedömning.

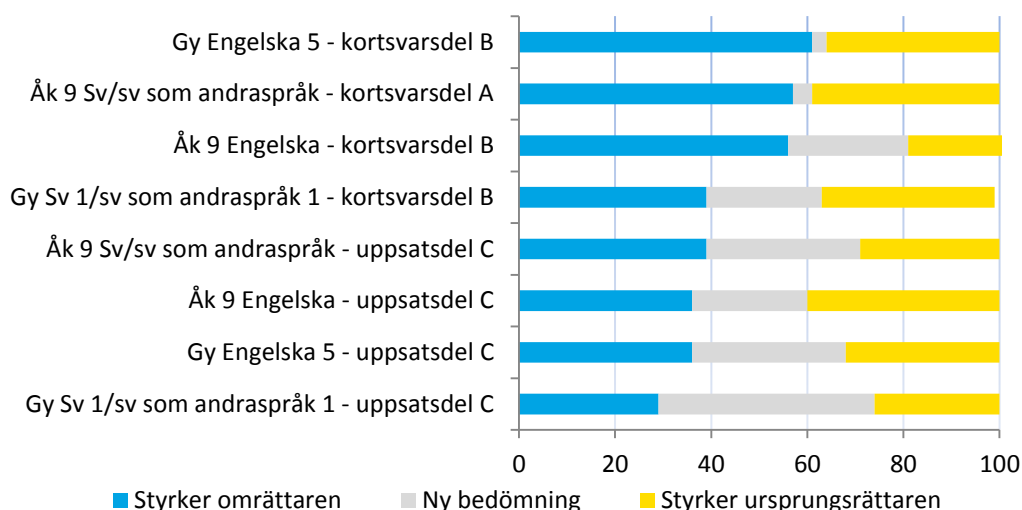


Diagram 4: Prov med avvikelser som omrättats en andra gång, för respektive delprov, omgång 4.

För prov med kortsvarsdelar och mer fasta svarsalternativ, som engelska i gymnasieskolan (delprov B) och svenska/svenska som andraspråk i årskurs 9 (delprov A) görs få nya bedömningar. Det är således endast vid 3-4 procent av

<sup>15</sup> För gymnasieskolan som har fler betygssteg gäller att det ska vara minst två betygsstegs skillnad.

<sup>16</sup> Uppsatsdelen i svenska/svenska som andraspråk i årskurs 9 (delprov C) gick till en andra omrättning i störst utsträckning (23 procent), medan kortsvarsdelen i engelska 5 i gymnasieskolan (delprov) omrättades en andra gång i minst utsträckning (4 procent av inskickade prov).

<sup>17</sup> Det är enbart för ämnena svenska och engelska som en andra omrättning varit möjlig, eftersom det för matematik och NO-ämnena inte finns några delprovsbetyg att jämföra med.

proven som en andra omrättning leder till en ny tredje bedömning. Nya bedömningar är däremot betydligt vanligare vid mer komplexa delprov med uppsatsdelar. Till exempel för uppsatsdelen i svenska 1/svenska som andraspråk 1 (delprov C) i gymnasieskolan sker en ny bedömning i 45 procent av de prov som omrättats en andra gång.

Forskning har visat att det är betydligt svårare att göra likvärdiga bedömningar vid mer komplexa format. Exempelvis ger uppsatsskrivning, mer utrymme för variation i bedömning.<sup>18</sup>

### Avvikelse varierar kraftigt mellan skolor

Avvikelse mellan ursprungs rättarens bedömning och Skolinspektionens bedömning varierar kraftigt mellan olika skolor.<sup>19</sup>

Avvikelse varierar på skolnivå – från ingen avvikelse till 89 procent, vid olika delprov i svenska och engelska. För till exempel kortsvarsdelen i engelska 5 i gymnasieskolan, finner vi en skola med minimivärdet 4 procent – vilket är en mycket låg andel avvikelse totalt sett. Detta kan jämföras med en annan skola där det maximala värdet på avvikelsen ligger på 78 procent.

De skolor som har lägst avvikelse vid olika delprov i engelska har ett minimivärde som varierar mellan 0 och 35 procent. För de skolor som har högst avvikelse varierar maxvärdet mellan 48 och 82 procent.

Motsvarande minimivärde vid svenskans delprov varierar mellan 13 och 36 procent. För de skolor som har högst avvikelse varierar maxvärdet mellan 48 och 89 procent.

### Spridningen mellan olika skolor ser likadan ut för olika delprov i gymnasieskolorna

Sammantaget visar det sig vid en analys av hur gymnasieskolorna ligger i förhållande till varandra att spridningen mellan delproven i engelska och svenska

---

<sup>18</sup> Palmér och Östlund-Stjärnegårdh, 2005, *Bedömning av elevtext. En modell för analys. Svenskämnet i år 8, 9 och gymnasiet*. Natur och Kultur, Stockholm, s.152.

<sup>19</sup> Urvalet är representativt på skolnivå vid de utvalda skolorna i ämnena svenska/svenska som andraspråk, matematik och engelska i årskurs 9 i grundskolan samt i gymnasieskolan. I resultatbilagan återfinns tabeller för samtliga skolor uppdelat på olika delprov.

är likartade. Om man jämför lådornas längd och ”pinnarnas” längd i låddiagram 5 nedan åskådliggörs spridningen.<sup>20</sup>

Vid kortsvarsdelen i engelska (delprov B) är spridningen mellan minimivärdet och maxvärdet störst – 74 procentenheter. Motsvarande spridning är minst för uppsatsdelen i engelska (delprov C) – 47 procentenheter. De allra flesta skolorna, 90 procent, har en total spridning på mellan 24-33 procentenheter.

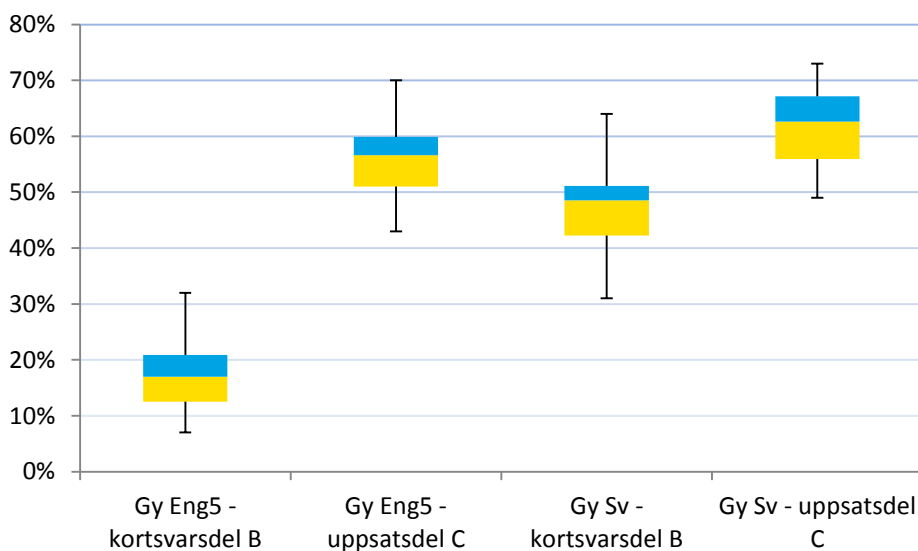


Diagram 5: Kvartilspredning över andel avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning för gymnasieskolan gällande delprov i engelska 5 och svenska 1/svenska som andraspråk 1, för gymnasieskolorna i urvalet, omgång 4.

### Spridningen mellan skolor är större för uppsatsdelarna än för kortsvarsdelarna i grundskolan

Spridningen är större för uppsatsdelarna än för kortsvarsdelarna i grundskolan, enligt diagram 6 nedan. För uppsatsdelen i svenska (delprov C) är spridningen mellan minimivärdet och maxvärdet störst – 61 procentenheter. Mot-

<sup>20</sup> Om samtliga bedömningarna av prov vid skolan avviker, mellan ursprungs rättare och omrätare, får det procenttalet 100, medan procenttalet 0 ges om samtliga provbedömningar vid en skola överensstämmer. Låddiagrammet prickar in lägena för medianen, kvartilerna och därtill 5:e och 95:e percentilerna. Medianen, är det tal som lika många skolor ligger över som under - det vill säga just 50 procent av skolorna ligger under det talet. ”Pinnarna” som går neråt och uppåt på lådan går till 5:e respektive 95:e percentilen. De antyder ett intervall där vi hittar de allra flesta skolorna, i detta fall 90 procent av skolorna i urvalet. Den 75:e percentilen beskriver ett tal som 75 procent av skolorna ligger under, och övriga 25 procent av skolorna ligger över.

svarande spridning är minst för kortsvarsdelarna i engelska och svenska (delprov B och A) – 20 respektive 35 procentenheter. De allra flesta skolorna, 90 procent, har en total spridning på mellan 22-40 procentenheter.

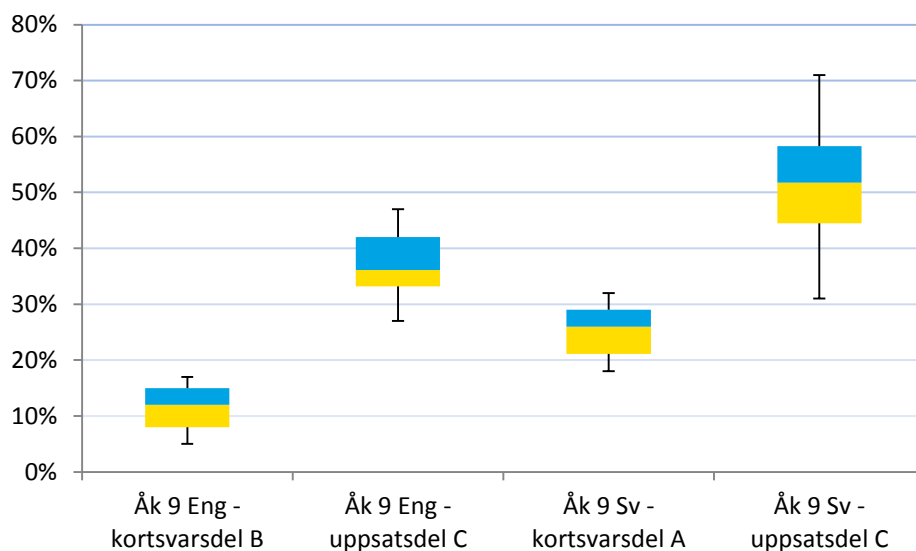


Diagram 6: Kvartilspredning över andel avvikelser mellan ursprungsrättarens bedömning och Skolinspektionens bedömning för grundskolan gällande delprov i engelska och svenska/svenska som andraspråk, för grundskolorna i urvalet, omgång 4.

### Avvikelsen varierar vid olika delprov för de 25 procent skolor som har störst avvikelse

Vi har ovan tittat på minimi- och maxvärden samt vilken spridning avvikelsen på bedömning har för olika skolor och delprov. Det är även intressant att titta på de skolor (25 procent) som har störst avvikelse. De 25 procent gymnasieskolor som har störst avvikelse har en andel avvikelse på över 60 procent för uppsatsdelarna (delprov C) i engelska och svenska (60 respektive 67 procent). Vid kortsvarsdelen i engelska har den fjärdedel skolor som har störst avvikelse mer än 21 procent avvikelse (delprov B).

I årskurs 9 har den fjärdedelen skolor med störst avvikelse mer än 58 procent avvikelse för uppsatsdelen i svenska (delprov C). För kortsvarsdelen är motsvarande avvikelse på 29 procent (delprov A).

### De 25 procent skolor som har lägst avvikelse

Om vi istället tittar på värdena för de 25 procent gymnasieskolor som har lägst avvikelse har de en avvikelse på drygt 50 procent för uppsatsdelarna. Det be-

tyder alltså att för den fjärdedel skolor som har lägst avvikelser bedöms vartannat uppsatsprov olika. Lägst andel avvikelse finns vid kortsvarsprovet i engelska (delprov B), 13 procent.

I årskurs 9 är avvikelsen något lägre för de 25 procent grundskolor som har lägst avvikelser, 44 procent vid provet i svenska och 33 procent vid provet i engelska (delprov C). Vid kortsvarsproven ligger avvikelsen på knappt 10 procent i den fjärdedel av grundskolorna som har lägst avvikelse vid provet i engelska.

### Sambedömning påverkar avvikelsen

Nästan alla lärare som bedömt proven har angett att de diskuterat sina bedömningar med lärare eller annan pedagogisk personal. Det visar en analys av den lärarinformation som samlades in vid varje elevlösning.<sup>21</sup>

Det finns en tendens till att de som sambedömt prov uppvisar något färre avvikelser. Det är dock ett fåtal lärare som *inte* sambedömt. Detta kombinerat med ett internt bortfall gör att tendensen måste tolkas med försiktighet. Analysen visar inte på några skillnader i avvikelser mellan ursprungs rättarnas bedömning och omrättares bedömning utifrån *vilken form* som sambedömningen skett.

Vanligast (i en tredjedel av fallen) är att de nationella proven diskuterats på ett bedömningsmöte eller i annan organiserad form. Nästan lika vanligt är att lärare vid tveksamheter diskuterat sina bedömningar med annan pedagogisk personal (i en knapp tredjedel av fallen). Mindre vanligt (vid ca 20 procent av fallen) är att lärare deltagit i principiella diskussioner inför bedömningen. Att på något annat sätt ha diskuterat sina bedömningar är ovanligt (knappt 10 procent av fallen).

---

<sup>21</sup> Skolinspektionen har i samband med elevlösningarna för varje delprov begärt in uppgifter avseende den lärare som bedömde provet. Vid tolkning av resultatet är det dock viktigt att tänka på att urvalet inte är anpassat för att kunna dra några säkra slutsatser utifrån den information som samlats in om lärarna. Det innebär exempelvis att fördelningen av män och kvinnor i urvalet inte speglar könsfördelningen bland lärare i riket som helhet. Det interna bortfallet – där lärare inte medgett att svaren får användas i analys syfte – är relativt stort (ca 20-30 procent beroende på delprov).

Ungefär en tredjedel av lärarna beskriver att de diskuterat bedömningarna på enbart ett av ovanstående sätt. Cirka en tredjedel har diskuterat på två sätt. Resterande andel, en tredjedel, har diskuterat på tre sätt eller fler.

### **Ordinarie lärare bedömer vanligen proven**

Ursprungsrättarens ålder, kön eller behörighet (att undervisa i aktuell årskurs och ämne) ger i analysen inga tydliga skillnader i avvikelser mellan ursprungsrättning och omrättning.

Forskare som, vid föregående omrättning av nationella prov, tittat på detta för Skolinspektionens räkning har sett att flertalet av sådana bakgrundskaraktäristika inte varit relaterade till hur proven bedömts.<sup>22</sup> Vanligast är att de ursprungsrättade proven bedömts av elevens ordinarie lärare (cirka 75 procent). Näst vanligast är att annan pedagogisk personal på skolan bedömt proven (cirka 25 procent). Det är mycket ovanligt att bedömning skett av annan pedagogisk personal utanför skolan (under 1 procent).

Det finns inget tydligt mönster i skillnad i avvikelser vid uppdelning på vem som ursprungsrättat proven.

### **Forskare har analyserat faktorer kring vad som kan förklara avvikelserna**

Skolinspektionen har också låtit forskarna Björn Tyrefors Hinnerich och Jonas Vlachos vid Stockholms universitet analysera förekomsten av systematiska skillnader i bedömningarna.<sup>23</sup> Deras analys visar att det spelar roll vilken grupp olika elever tillhör för hur stora avvikelserna är mellan bedömningarna av de nationella proven.

Forskarna pekar på en tendens till att ursprungsrättarna jämfört med omrättarna sätter lägre provbetyg på pojkar, elever med utländsk bakgrund, elever vars föräldrar saknar högre utbildning och elever som är födda sent på året. Samtliga dessa grupper är sådana som generellt tenderar att uppvisa relativt svaga studieresultat.

---

<sup>22</sup> Forskarnas rapport *Systematiska skillnader mellan interna och externa bedömningar av nationella prov* finns i föregående års resultatbilaga, *Lika för alla? Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år*. Årets analys av forskarna Björn Tyrefors Hinnerich och Jonas Vlachos, Stockholms universitet, *Systematiska skillnader mellan interna och externa bedömningar av nationella prov – en uppföljningsrapport*, finns i resultatbilagan för denna rapport.

<sup>23</sup> Analysen har gjorts på resultat från omrättningen omgång 2 och 3.

Gymnasieelever med höga grundskolebetyg tenderar att få en generösare provbedömning av ursprungs rättare jämfört med bedömningen av omrättare. Skillnaderna är överlag små till medelstora. Mönstret är inte helt likartat mellan grundskolan och gymnasieskolan.<sup>24</sup>

Generellt sett tyder alltså resultaten på att elever som tillhör grupper som kan förväntas uppvisa goda studieresultat, till exempel flickor och elever vars båda föräldrar är högutbildade, av ursprungsrättaren bedöms generösare än andra elever, jämfört med bedömningen av omrättaren.

Vidare visar analysen av lärartäthet och pedagogisk examen att avvikelserna mellan ursprungsrättaren och omrättarens bedömning är något lägre vid skolor med högre lärartäthet. Det finns även en svag tendens till mindre bedömningskillnader vid skolor med en hög andel lärare med pedagogisk examen. Alltså verkar lärarnas behörighet inte ha någon större betydelse vid bedömningen.

Hur lärare omrättar prov på fristående och kommunala skolor kan skilja sig åt varför forskarna även undersöker om det finns systematiska skillnader mellan ursprungsbedömningarna och omrättarnas bedömningar vid skolor med olika huvudmän.

Det visar sig att fristående skolor i genomsnitt sätter högre provbetyg än de kommunala skolorna, relativt omrättarnas bedömningar. Forskarna finner därutöver inte att bedömningsavvikelsena skiljer sig mellan kommuner där konkurrensen mellan skolorna kan förväntas vara stark respektive svag.

### **Olika skrivuppgifter påverkar inte avvikelserna**

Vid föregående års omrättning av nationella proven visade det sig att graden av avvikelse för uppsatsproven skiljde sig åt beroende på vilken uppsatsuppgift som hade bedömts. Detta gällde såväl för gymnasieskolan som för årskurs 9.

Vid en analys av resultatet från denna omrättning visar det sig att skrivuppgiftens format inte påverkar graden av avvikelse. Den totala avvikelserna ligger på drygt 65 procent respektive cirka 50 procent för skrivuppgifterna oavsett format för gymnasieskolan och årskurs 9. En förklaring till att inga skillnader framträder vid olika skrivuppgifter kan vara att eleverna på gymnasieskolan numera för uppsatsskrivningen i svenska 1/svenska som andraspråk 1 enbart

---

<sup>24</sup> Sambandet som visar att pojkar får lägre provbetyg än flickor av ursprungsrättare gäller gymnasieskolan. I årskurs 9 är effekten liten och inte statistiskt säkerställd. Positiv särbehandling av elever med utländsk bakgrund avser årskurs 9 – i gymnasieskolan är effekten väldigt liten. Vad gäller elever som är födda tidigt på året (första kvartalet) gäller sambandet gymnasieskolan.



kan välja mellan fyra olika skrivuppgifter (utifrån olika genrer, exempelvis debattinlägg och läsarinlägg på nätforum). Föregående år kunde de istället välja mellan åtta olika skrivuppgifter. I ämnet svenska/svenska som andraspråk i årskurs 9 kan eleverna välja mellan fyra uppgifter.

## 4. När är olikheterna för stora?

I regleringsbrevet för Statens skolinspektion 2013 preciserar regeringen att en analys ska genomföras av vad som är att betrakta som en normal respektive obefogad avvikelse:

*[...]Skolinspektionen ska, utifrån en analys av vad som är att anse som normala avvikelser för respektive ämne och delprov, ta ställning till och redovisa avvikelser som framstår som obefogade [...].<sup>25</sup>*

Vi definierar obefogade avvikelser utifrån den samstämmighet som för närvarande kan uppnås för de nationella proven enligt forskningen och våra egna resultat och antar därefter att en obefogad avvikelse är när en skolas bedömning av nationella prov avviker från den nåbara överensstämmelsen.

Nedan gör vi en genomgång av vad forskningen tar upp som nåbar bedömaröverensstämmelse vid bedömningen av uppsatsdelar. Därefter tittar vi på kortvarsdelarnas nåbara överensstämmelse utifrån de 5 procent skolor som har minst avvikelser enligt omrättningen, då det i forskningen inte finns något enhetligt gränsvärde för dessa delar.

En avgränsning är att analysen enbart görs för de delprov där det finns delprovsbetyg – ämnena svenska och engelska. De siffror på nåbar överensstämmelse och obefogad avvikelse som beskrivs avser ett snitt av alla skolor som ingår i omrättningen. På skolnivån kan det variera av olika skäl. Det går därför inte att applicera den exakta gränsen på en enskild skola, men den kan användas som ett jämförelsemått.

### **Elevperspektiv – elevens rätt till en likvärdig bedömning**

Ur ett elevperspektiv skulle lärarna i ett idealläge nå samstämmighet till 100 procent vid bedömningen av nationella proven. Men eftersom vissa prov är mer komplexa, det finns inte ett rätt och ett fel, så är det orealistiskt att tro att en exakt överensstämmelse kan uppnås.

---

<sup>25</sup> Regleringsbrev avseende Statens skolinspektion för verksamhetsåret 2013.

Förenklat uttryckt har uppsatsproven ett mer komplext format, medan prov av kortsvarskaraktär är av sådant format att de ger högst bedömaröverensstämmelse.<sup>26</sup> Om skillnaden är alltför stor mellan olika lärares bedömning kan den till slut bli godtycklig och bero helt på vilken lärare eleven råkar ha. Ur ett elevperspektiv utgör det en stor risk för att få ett delprovsbetyg av en lärare och ett helt annat delprovsbetyg av en annan lärare.

Ur ett elevperspektiv måste man betrakta låg bedömaröverensstämmelse som en riskfaktor som kan ha stor betydelse för eleven framtida studie- och yrkesval. Eleven vill såklart inte att bedömningen av det nationella provet ska bero på vilken lärare han eller hon har.

## Uppsatsprov

Olika typer av prov ger olika nivåer av bedömaröverensstämmelse enligt forskning. Mer komplexa format, som uppsatsskrivning, har mer utrymme för olikheter i bedömningen.<sup>27</sup> Lärares bedömning skiljer sig åt i stor omfattning när det gäller elevers mer fria skrivande, vilket betyder att det är svårare att nå en likvärdig bedömning.<sup>28</sup> Uppsatsdelar är det format som ger lägst bedömaröverensstämmelse. Bedömningsmallarna är också uppbyggda på mer komplexa bedömningsmatriser. Forskningen menar att en nåbar bedömaröverensstämmelse grovt kan beräknas till cirka 70 procent.<sup>29</sup> Störst samstämmighet uppträder enligt forskningen i undersökningar där bedömarna är specialtränade.<sup>30</sup>

I Norge har forskare i ett stort projekt (KAL-projektet<sup>31</sup>) undersökt examensprov och elevers skrivande i den norska grundskolan. Berge beskriver olika bedömningsundersökningar i sin diskussion av de norska proven. Låga nivåer på överensstämmelse och samstämmighet är normalt. I materialet varierade bedömaröverensstämmelsen mellan 63 och 73 procent.

---

<sup>26</sup> Se exempelvis Berge 2005 och Skolverket 2009, s 26.

<sup>27</sup> Palmér och Östlund-Stjärnegårdh 2005, *Bedömning av elevtext. En modell för analys. Svenskämnet i år 8, 9 och gymnasiet*. Natur och Kultur, Stockholm, s.152.

<sup>28</sup> Östlund-Stjärnegårdh 2002. *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter*. Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 57. Uppsala universitet.

<sup>29</sup> Berge 2005, inom ramen för KAL-projektet där ett bedömaröverensstämmelsemått redovisades.

<sup>30</sup> Skolverket 2009, s 25.

<sup>31</sup> Kvalitetssikring av laeringsutbyttet i norsk skriftlig

Det finns ett flertal andra studier där nivåerna på bedömaröverensstämmelse är lägre. Skolverket har i en analys 2009 av 100 elevlösningar för uppsatsdelen (delprov C) i svenska funnit att vid 54 lösningar, gav lärare och bedömare samma betyg, på den fyrgradiga betygsskalan. Vid analysen används en tio-gradig skala vid de enskilda delproven. Bedömaröverensstämmelsen varierade från 36 till 46 procent emellan bedömarna.<sup>32</sup> Sammanfattningsvis blev utfallet ungefär 40 procents bedömaröverensstämmelse vid studien, vilket är lägre än vid forskningen i KAL-projektet, där ett samband på 60 procents bedömningsöverensstämmelse vid uppsatsprov identifierades.

En mindre studie av bedömning av elevtexter i årskurs 9 har gjorts av Ciolek Laerum. Hon undersökte hur 40 elevtexter bedömdes av lärare och övriga granskare från det nationella provet 2006. Vid 35 procent av texterna överensstämde bedömningar av betygen.<sup>33</sup>

Edander har i en undersökning analyserat hur hög överensstämmelse som kan nås då fem av varandra oberoende bedömare betygsätter 30 elevtexter från gymnasieskolans kurs 1, skrivdelen. Lärarna, ursprungsläraryrarna samt fyra externa bedömare, har varit helt överens om bedömningen för 35 procent av elevlösningarna.<sup>34</sup>

**Forskningen pekar på en nåbar bedömaröverensstämmelse på 70 procent**  
Skolinspektionen ser utifrån sin analys av de omrättade nationella proven inga skäl att göra någon annan bedömning än den som hittills gjorts av genomförd forskning när det gäller uppsatsdelar. Skolinspektionen bedömer att det är möjligt att skolor arbetar med sin bedömning av uppsatsdelar, så att man i efterhand ska kunna konstatera en överensstämmelse på 70 procent.

## Kortsvarsprov

Kortsvarsproven är väldigt olika sinsemellan men har det gemensamt att de inte främst innehåller en längre skrivuppgift, utan snarare kortare svarsformat av olika slag. Det kan exempelvis vara flervalfrågor eller frågor där man en-

---

<sup>32</sup> Skolverket 2009, s 22-23.

<sup>33</sup> Ciolek Laerum, 2009. *Elever skriver och lärare bedömer: en studie av elevtexter i åk 9*. Uppsala universitet, Institutionen för nordiska språk, FUMS Rapport. Vid analyser framkommer att det är ovanligt att alla bedömare är överens. Samstämmigheten är störst vid det lägsta betyget IG och det högsta MVG.

<sup>34</sup> Edander, 2013. *Bedömaröverensstämmelse vid Nationella provet i svenska för gymnasiet kurs 1, skrivdelen*. Institutionen för nordiska språk. Uppsala universitet.

bart ska fylla i ett eller några enstaka ord som svar. Det förekommer även kortare öppna frågor i olika omfattning och längd. De olika så kallade kortsvarsproven har alltså en varierande grad av komplexitet vad gäller bedömningen av svaren. Dessa uppgifter kan kombineras i olika delprov och variera från år till år. Att kategorisera proven i detalj skulle kräva en helt annan typ av analys än den vi kunnat göra inom ramen för detta uppdrag.

### **Inga enhetliga nivåer enligt forskningen**

Utifrån forskningen är det svårt att hitta enhetliga nivåer för bedömaröverensstämmelse för olika typer av kortsvarsprov. Det som förenar de studier som genomförts är att överensstämmelsen är högre för kortsvarsdelarna, jämfört med uppsatsdelar.

Skolverket har i en analys 2009 av 100 slumpmässigt utvalda elevlösningar vid kortsvarsdelen (delprov A) i svenska funnit att 72 av 100 prov fick samma delprovsbetyg.<sup>35</sup> Man kan alltså tala om avvikelser i 28 procent av proven. Förutsättningarna var att ombedömarna hade olika typer av bakgrund och förståelse för att få en bredd i bedömningarna.

Vid proven i engelska har bland annat Erickson tittat på bedömarsamstämmighet.<sup>36</sup> Vid ett urval av 100 prov (från 2007) visar en analys av delprov B att överensstämmelsen mellan lärarna och ombedömaren i det närmaste är total, 0,99.<sup>37</sup> Förutsättningarna var att ombedömaren var en medarbetare på institutionen som arbetade med just provkonstruktionen så i detta fall kan man se det som en expertbedömning. Olsson-Wahlsten har gjort en studie som undersökte en uppgift från 1000 elevlösningar som ingick i ämnesprovet för årskurs 9, under 2011.<sup>38</sup> Överensstämmelsen bedömdes som god, 0,81 för lärarbedömningen och 0,80 för ombedömningen.<sup>39</sup>

### **Nåbar bedömaröverensstämmelse enligt de 5 procent skolor som har lägst avvikelse**

Eftersom kortsvarsdelarna ser väldigt olika ut är det nästintill omöjligt att hitta underlag från tidigare studier för hur vi ska bedöma vad som är en nåbar be-

---

<sup>35</sup> Skolverket 2009, s 22-23.

<sup>36</sup> Erickson 2009. *Nationella prov i engelska – en studie av bedömarsamstämmighet*. Göteborgs universitet: Institutionen för pedagogik och didaktik.

<sup>37</sup> Avser sambandet för delprovsbetyget i form av en tiogradig betygsskala.

<sup>38</sup> Olsson-Wahlsten 2002. *Öppna svar – hur funkar det?* Institutionen för pedagogik och didaktik. Göteborg: Göteborgs universitet.

<sup>39</sup> 0,98 var styrkan på sambandet.

dömaröverensstämmelse. För kortsvarsdelarna har vi valt att bedöma nåbar överensstämmelse utifrån de avvikelser vi ser i årets omgång av omrättningen.

När det gäller kortsvarsdelarna har vi tittat på de 5 procent skolor med minst avvikelser. Vissa skolor har nått en hög bedömaröverensstämmelse och det borde även vara möjligt för fler skolor. Vi kan konstatera att en av kortsvarsdelarna sticker ut genom att den har väldigt höga avvikelser. Det är delprovet i svenska 1/svenska som andraspråk 1 i gymnasieskolan (delprov B) där de 5 procent skolor med högst andel avvikelser har 31 procents avvikelse eller mer, vilket får ses som ett extremfall. Det borde kunna ligga betydligt lägre med tanke på att uppsatser enligt forskningen ska kunna nå under 30 procent avvikelse. Delprovet i engelska i gymnasieskolan (delprov B) har en betydligt lägre nivå avvikelser, 5 procent av skolorna har max 7 procent avvikelser. För kortsvarsdelarna i svenska och engelska i årskurs 9 (delprov A och B) har 5 procent av skolorna en avvikelse på max 18 respektive 5 procent.

Vår slutsats är att det för olika typer av kortsvarsprov är möjligt att nå en bedömaröverensstämmelse som är minst 80 procent vid mer svårbedömda kortsvarsdelar, och över 90 procent bedömaröverensstämmelse vid mer lättbedömda kortsvarsdelar.<sup>40</sup>

### **Obefogade avvikelser - det finns en förbättringspotential hos skolorna**

Enligt Skolinspektionen bör en betydligt högre överensstämmelse kunna uppnås än vad fallet är idag om skolor arbetar på ett annat sätt med sin bedömning.

Tittar vi på resultatet från årets omrättning, ser vi att uppsatsdelar och kortsvarsdelar i de nationella proven i svenska och engelska, i genomsnitt avviker betydligt mer än vad vi menar är nåbart. I genomsnitt varierar andelen avvikelser på skolorna mellan 37 och 62 procent för uppsatsproven och mellan 19 och 47 procent för kortsvarsproven.

---

<sup>40</sup> Det finns en mängd variationer vad gäller utformningen av proven som forskarna tittat på och de vi omrättat. För en komplett analys skulle vi behöva följa varje enskild fråga för att se hur olika frågetyper skiljer sig åt i avvikelse. Vi kan dock inte göra en sådan analys då vi i så fall skulle behöva samla in resultaten på uppgiftsnivå från varje skola och varje elev, vilket är ett allt för omfattande arbete.

95 procent av grund- och gymnasieskolorna har en obefogad avvikelse över gränsen på 30 procent vid uppsatsdelar. Vid kortsvarsdelar med kortare öppna frågor i olika omfattning och längd har 91 procent av skolorna en obefogad avvikelse över gränsen på 20 procent. För de kortsvarsdelar med flervalfrågor eller frågor där ett eller några enstaka ord fylls i är det 73 procent av skolorna som har en obefogad avvikelse över gränsen på 10 procent.

Det är alltså mellan 5 och 27 procent av bedömningarna på skolorna som har befogade avvikelser.

Enligt Skolinspektionen bör en rad åtgärder genomföras för att förbättra bedömaröverensstämmelsen och nå ett mer likvärdigt resultat. Förbättringspotentialen är minst 20 procentenheter vad gäller uppsatsdelarna. För kortsvarsdelarna är förbättringspotentialen inte lika stor och varierar det mer. För de delprov som främst innehåller öppna frågor i olika omfattning och längd, finns ett stort utrymme för att förbättra bedömaröverensstämmelsen.

En del skolor som ingått i årets omrätningsomgång har kommit relativt långt och har låga andelar avvikelser. Men de skolor som har obefogade avvikelser bör arbeta med att uppnå mer likvärdiga och rättssäkra bedömningar.

Stor skillnad i bedömning av samma elevlösning, beroende på bedömande lärare, är att betrakta som något allvarligt. Skolinspektionen kan konstatera att vissa skolor uppvisar resultat som föranleder krav på vidare granskning. På ett antal skolor som uppvisat stora avvikelser kommer Skolinspektionen att bedöma hur man arbetar med bedömning och betygssättning.<sup>41</sup>

---

<sup>41</sup> Skolinspektionen 2013d. Den pågående granskningen rör *Bedömning och betygssättning 2013*. Mer information finns på Skolinspektionens hemsida.

<http://www.skolinspektionen.se/sv/Tillsyn--granskning/Kvalitetsgranskning/Skolinspektionen-granskar-kvaliteten/Bedomning-och-betygssattning-2013/>. Granskningen fokuserar på två områden: Om lärare beslutar om betyg med utgångspunkt i de krav som anges i styrdokumentet och om rektor skapar förutsättningar för kunskapsbedömning och betygssättning enligt styrdokumentet. Se även resultat från tidigare granskning utifrån omrättningen: Skolinspektionen 2013a. *Tillsyn av bedömning och betygssättning. Med fokus på skolor med stora avvikelser i förhållande till Skolinspektionens omrättning och i relationen mellan nationella prov och slutbetyg*. Tillsynen är gjord på 30 skolor som visar en stor avvikelse mot Skolinspektionens omrättning och/eller har stor avvikelse i relationen mellan nationellt provbetyg och slut- eller kursbetyg.

## 5. Rekommendationer kring uppsatsdelarna

Skolinspektionen har omrättat nationella prov under fyra år. Resultatet av denna omrättning är entydigt och samstämmigt mellan åren. För uppsatsdelarna finns det stora skillnader i bedömning. Minst två personer har alltså gjort olika bedömningar av ett och samma delprov i väldigt många fall.

Skolinspektionen föreslår därför att Regeringen överväger att **tillsätta en utredning kring för- och nackdelar med att låta uppsatsdelar ingå i de nationella proven.**

### Svårt att uppnå likvärdighet vid bedömning av uppsatsdelar

De nationella proven ska utgå från kursplanerna och har ett flertal syften. Ett första syfte är att proven ska ge underlag för i vilken utsträckning kunskapskraven uppfylls.

De förmågor och kunskaper som prövas i de delprov som innehåller uppsatskrivande och längre självständigt författade redogörelser är centrala och viktiga delar av kursmålen. Uppsatserna är ett underlag som läraren behöver för att bedöma i vilken utsträckning kunskapsmålen uppfylls. Det är viktigt att eleven lär sig bemästra analyser, problemlösningar och eget skrivande, men bedömningen av dessa färdigheter blir ofrånkomligen mer subjektiv än vid till exempel kortsvarsdelar.

Ju mer kreativ och självständig eleven tillåts att vara i sina svar, desto svårare blir det att skapa tydliga och allmänna bedömningsanvisningar och för lärarna att göra likvärdiga bedömningar. Det får konsekvenser för likvärdigheten i bedömningen. Skolinspektionens resultat visar att så många som vartannat omrättat uppsatsdelprov bedöms olika av olika bedömare.

Ett andra syfte med de nationella proven är att de ska "stödja en likvärdig och rättvis bedömning och betygsättning". Skolinspektionens bedömning är att delprov med uppsatsdelar under rådande förhållanden fungerar dåligt för att stödja likvärdigheten, vilket är ett av två huvudsyften med de nationella proven. Frågan är om uppsatsdelarna ens kan användas för att uppnå likvärdighet.

Avvikelserna mellan ursprungsrättarens bedömning och omrättarens bedömning är avsevärt mindre för de delprov som innefattar kortsvarsfrågor eller flervalsdelar. Dessa delprov är således mer lämpade för att stödja en likvärdig bedömning av de kunskaper och kompetenser som prövas i dessa delprov.

## Utredning föreslås för att genomlysna behovet av uppsatsdelar

Om uppsatsdelarna helt lyfts bort i nationella prov skulle provet i sin helhet, om det består av kortsvars- och flervalsdelar, bli mer likvärdigt. Det är en stor fördel. Samtidigt måste man om uppsatsen försvinner som ett underlag för den likvärdiga bedömningen, ha andra prov och former för att pröva elevers mer kvalificerade skrivförmåga, eftersom de är en viktig del i kursplanen. De nationella proven förser läraren med ett verktyg för att pröva dessa viktiga förmågor. Detta är självfallet viktigt även framöver. Uppsatsdelarna behöver dock inte ingå som ett underlag för den likvärdiga betygssättningen.

Det finns också en risk att undervisningen skulle inriktas i större utsträckning mot de innehållsdelar som finns representerade i nationella prov. Om nationella prov bara innefattade kortsvarsdelar skulle det kunna uppstå risker med att elever inte skulle få undervisning utifrån kursplanerna i dessa delar.

Samtidigt kan de nationella proven ge lärarna en falsk trygghet av att proven är ett bra verktyg för att bedöma och pröva elevers skrivförmåga på ett likvärdigt sätt. De stora avvikelser som visar sig vid bedömningen av uppsatsdelar pekar på att uppsatsdelarna inte är ett bra underlag för att göra likvärdiga bedömningar.

Ovan nämnda faktorer skulle behöva genomlysas eftersom det annars finns en risk att tilltron till de nationella proven som ett instrument för likvärdighet på sikt kan urholkas.

## En kombination av åtgärder för att öka likvärdigheten

Så länge uppsatsdelarna kvarstår i de nationella proven finns utrymme för att förbättra bedömarstämmigheten. Det finns en rad faktorer som kan påverka bedömaröverensstämmelsen, som att de lärare som ska bedöma ett visst prov inte har samma definition av den kunskap som ska mätas som konstruktörerna av proven.<sup>42</sup> För att provresultaten ska bli likvärdiga och rättssäkra måste skolor bland annat arbeta systematiskt med att förbättra sin bedömarförmåga som en del av kvalitetsarbetet.

---

<sup>42</sup> Skolverket 2009.



Skolinspektionen rekommenderar nedan en kombination av åtgärder för att förbättra bedömaröverensstämmelsen<sup>43</sup> och öka likvärdigheten i bedömningarna. Rekommendationerna bygger på resultat och erfarenheter från fyra års omrättningar av nationella prov och tidigare forskning. Skolinspektionen vill understryka vikten av att det är kombinationen av rekommendationer som ökar likvärdigheten.

### Skolinspektionen rekommenderar följande åtgärder för att öka samstämmigheten vid bedömningen:

- Skolverket bör fortsätta med att **förtydliga och tillgängliggöra bedömningsanvisningarna** framförallt i de delar som gäller uppsatsdelarna och de högre betygsstegen.
- Skolhuvudmännen bör se till att lärare som ska bedöma elevers prestation på de nationella proven får en god **grund- och fortbildning i att bedöma**.
- Rektorerna bör se till att lärarna använder sig av **sambedömning så långt det är möjligt**, där mer än en lärare i samråd bedömer samma prov.
- Skolverket bör utreda förutsättningarna för om ett bedömningsförfarande där **elevens identitet inte är känd för läraren** som bedömer proven bör införas.

### Skolverket bör fortsätta förtydliga bedömningsanvisningarna

Bedömningsanvisningarnas utformande är en avgörande faktor för att uppnå en ökad samstämmighet och en likvärdig bedömning av elevernas svar.<sup>44</sup> Det är därför av yttersta vikt att anvisningarna är tydliga och använder sig av beskrivningar som lämnar ett minimalt utrymme för enskilda tolkningar. Men med fler exempel och mer omfattande anvisningar följer ökade krav på en tyd-

---

<sup>43</sup> Omrättningsens resultat visar att det är stor skillnad i andel avvikelse vid olika delprov och ämnen för uppsatsproven. Enligt forskningen kan en högre bedömaröverensstämmelse nås även för uppsatsproven.

<sup>44</sup> Se till exempel Andrén och von Horn, 2010. *Att resonera i minsta möjliga mån. Om den lägsta resultatnivån för betyget G på den skriftliga delen av nationella provet i svenska B*. Institutionen för pedagogik, psykologi och idrottsvetenskap. Linnéuniversitetet, s.25-26.

lig struktur.<sup>45</sup> Skolverket bör intensifiera sina insatser för att främja en mer likvärdig bedömning i olika ämnen. Skolinspektionens erfarenheter är att avvikelserna är som störst för uppsatsdelar och de högre betygsstegen vilket tyder på att Skolverket framförallt bör tillgängliggöra och förtydliga anvisningarna i dessa delar.

### Lärare måste utbilda sig i bedömning

Skolinspektionen anser även att skolhuvudmännen bör ta ett större aktivt ansvar för att se till att lärare som ska bedöma elevers prestation på de nationella proven får en god grund- och fortbildning i att bedöma. Bedömning är också en viktig del i grundutbildningen för lärare.

Lärare kan tolka kunskapskrav och kriterier för olika steg på betygsskalan olika.<sup>46</sup> Forskningen har pekat på att bedömaröverensstämmelsen ökar om lärarna tränas i bedömning.<sup>47</sup> Kvaliteten i bedömningarna hos de lärare som bedömer nationella prov behöver förbättras och genom att lärare tränar upp sin bedömningsförmåga kan de både bli medvetandegjorda och självreflekterande.

### Sambedömning så långt det är möjligt

Skolinspektionen anser att rektorerna bör ta ansvar för att lärarna får förutsättningar för och faktiskt använder sig av sambedömning så långt det är möjligt. Skolinspektionen har vid en riktad tillsyn med fokus på skolor med stora avvikelser funnit att skolor ofta brister i arbete med bedömning och betygsättning av elever.<sup>48</sup> För att bedömningen ska bli så likvärdig som möjlig är det viktigt att lärarna regelbundet diskuterar och analyserar hur de bedömer eleverna i förhållande till kunskapskraven.<sup>49</sup> I tillsynen visar det sig att de skolor med mest nöjda elever, vad gäller information kring betygsättning, var de sko-

---

<sup>45</sup> Bedömningsanvisningarna är till exempel mycket omfattande (för svenska 1/svenska som andraspråk 1 överstiger de 100 sidor och för svenska/svenska som andraspråk för årskurs nio överstiger de 70 sidor).

<sup>46</sup> Skolverket 2011. *Kunskapsbedömning i skolan – praxis, begrepp, problem och möjligheter*. Stockholm: Skolverket, s 35.

<sup>47</sup> Gipps 1994 och Alderson 1995.

<sup>48</sup> Skolinspektionen 2013a, s 6-11. I dessa skolor samverkar lärarna oftast inte kring bedömning och betygsättning och i flera fall tar inte heller rektorn sitt fulla ansvar för att leda och styra skolans arbete på området.

<sup>49</sup> Skolinspektionen 2013b, *Skolans kvalitetsarbete ger möjlighet till förändring, Skolinspektionens erfarenheter och resultat från tillsyn och kvalitetsgranskning 2012*. Stockholm: Skolinspektionen. I en riktad tillsyn av bedömning och betygsättning ser vi att rektorer inte vet om betygsättningen på deras skola är likvärd eller inte – det är inget de följer upp. s.41.

lor där lärarna samverkade med varandra kring bedömning och betygssättning.<sup>50</sup>

Samverkan mellan lärare kring bedömning och betygssättning anges ofta som en av grunderna för att skapa förutsättningar för en likvärdig bedömning.<sup>51</sup> Forskning har visat att sambedömning ger högre bedömaröverenstämmelse.<sup>52</sup> Bland annat visar Larson att flera bedömare ger en mer reliabel bedömning samt att det för exempelvis elevtexter är möjligt att uppnå ett gott resultat när tre lärare bedömer samma text.<sup>53</sup> Jönsson menar att samverkan kring bedömning kan leda till att bedömarkompetensen ökar hos lärare.<sup>54</sup>

Det är viktigt att sambedömning problematiseras och inte ses som den enda lösningen för att öka likvärdigheten vid bedömningarna av nationella proven. Om proven sambedöms på elevens skola kan en likvärdig bedömning av eleverna på den aktuella skolan ske, men skillnader kan fortfarande förekomma i förhållande till andra skolor i landet.

Sambedömning kompenserar inte heller provens komplexa syfte eller bedömningsanvisningarnas tolkningsutrymme. Det är också viktigt att sambedömningen definieras utifrån vad forskningen visar är den mest effektiva metoden.

### **Bedöma prov där elevens identitet inte är känd för läraren**

Skolinspektionen anser att förutsättningarna bör utredas för ett bedömningsförfarande där elevens identitet inte är känd för den lärare som bedömer proven.

Forskning av bland annat Tyrefors Hinnerich och Vlachos vid Stockholms universitet har funnit att elever som tillhör grupper som kan förväntas uppvisa goda studieresultat bedöms generösare än andra elever vid bedömningarna av nationella prov.<sup>55</sup>

Selghed konstaterar att relationerna mellan lärare och elev tycks vara centrala i bedömningen. Särskilt viktig upplevs relationsaspekten vara när det gäller ele-

---

<sup>50</sup> Skolinspektionen 2013b, s 41, Skolinspektionen 2013c.

<sup>51</sup> Skolinspektionen 2013a, s 6.

<sup>52</sup> Skolverket 2009, s 25.

<sup>53</sup> Larson 1984, Skrivförmåga. Studier i svenskt elevspråk, Liber, s35.

<sup>54</sup> Jönsson 2011, Jönsson 2012

<sup>55</sup> Se hela analysen i resultatbilagan.

ver som får ett G utan att egentligen ha uppnått målen för det betyget.<sup>56</sup> Vid gränsfall där man har att välja mellan en lägre och en högre poäng på en uppgift kan ursprungs rättaren ha en tendens att utgå från elevens tidigare prestation.<sup>57</sup> Kritiken mot ett förfarande där elevens lärare inte känner till vem som skrivit det prov som bedöms är att elevens lärare behöver få kunskap om elevens kunskapsnivå. Det får man när man själv bedömer de nationella proven.

Skolinspektionen anser att det inte hindrar att läraren vid ett senare tillfälle kan ta del av elevens lösning i full vetskap om vems den är.

## 6. Rekommendationer inför kommande omrättningar

### Delprovsbetyg skulle underlätta omrättning i matematik och NO

Som de nationella proven är utformade i dag finns det stora problem med att analysera resultatet av omrättningen för nationella prov i matematik och de naturvetenskapliga ämnena för årskurs 9 och gymnasieskolan. För de proven är det svårt att bedöma likvärdigheten i bedömningen av omrättade prov, då det i proven ingår delprov som består av muntliga framställningar eller laborationsmoment som inte är möjliga att omrätta. Det utfärdas ett gemensamt provbetyg, där muntliga delar och laborationer ingår. Eftersom delar inte kan omrättas, kan det provbetyget inte användas.<sup>58</sup>

I dag samlar Skolinspektionen istället in resultat på poängnivå för delprov i matematik, fysik, biologi och kemi för årskurs 9. Det medför ett ganska omfattande dubbelarbete för skolorna, eftersom de då måste rapportera in uppgifterna såväl till Skolverket som till Skolinspektionen. Det leder dessutom lätt till ett ökat bortfall och en risk för försämrad kvalitet på de insamlade uppgifterna.

---

<sup>56</sup> Selghed 2004. *Ännu icke godkänt. Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Malmö Studies in Educational Science, No. 15: Lärarutbildningen Malmö högskola.

<sup>57</sup> Skolverket 2008. Om en elev har höga betyg kan bedömningen då bli att den är värd den högre poängen, medan motsatsen blir troligt för en elev med låga betyg. En oberoende bedömare äger inte någon kunskap om eleven. För en lärare som har eleven kan bedömningen påverkas av annan information om den bedömda, även kallat haloeffekten.

<sup>58</sup> I proven i svenska/svenska som andraspråk och engelska finns även muntliga delar. Men eftersom det finns betyg för de olika delproven kan uppgifterna hämtas direkt från SCB/Skolverket.

Resultatredovisningen och analysen har gjorts på poängnivå för de delprov som redovisas skriftligt.

Ett möjligt sätt att handskas med problemet i framtiden är att utfärda delprovsbetyg för prov innehållande delar som inte är praktiskt möjliga att omrätta. Skolinspektionen förespråkar därför att delprovsbetyg utfärdas. Det skulle dock kräva att Skolverket delvis ändrar provens konstruktion.

## 7. Urval inför kommande omrättningar

Skolinspektionen ska enligt uppdragsbeskrivningen i Regleringsbrevet för 2013 koppla urvalet av elevlösningar när det gäller nationella prov till myndighetens planerade regelbundna tillsyn. Det gäller för nationella prov som genomförs från och med läsåret 2012/13 och det innebär att kravet på riksrepresentation utgår. Antalet skolor som omfattas av Skolinspektionens uppdrag om central rättning, och för vilka resultatet av rättningen är representativt på skolnivån, ska utökas. Från och med nästkommande omgång kommer omrättningen av nationella prov att kopplas ihop med de skolor som ingår i den regelbundna tillsynen ett givet år.

Urvalet ska kompletteras utifrån en behovsanalys, så att de prov och delprov där risken för avvikelser är störst ingår i så hög utsträckning som möjligt. Det innebär att det framför allt är uppsatsdelarna och vid behov vissa kortsvardsdelar, som kommer att omrättas. Proven som genomförs i årskurs 3 exkluderas, på grund av få avvikelser i tidigare omrättningsomgångar. Många frågor i proven för de lägre årskurserna är just frågor med fasta svarsalternativ. Även för de frågor där svaret inte lika tydligt har karaktären av rätt och fel är svaren korta och koncentrerade vilket gör att tolkningsutrymmet i bedömningen inte hamnar på samma nivå som för de högre årskurserna. Prov utan delprovsbetyg (exempelvis proven i NO och matematik) är svåra att analysera och kommer därför också att exkluderas.<sup>59</sup>

## 8. Genomförande av uppdraget

Detta kapitel beskriver Skolinspektionens genomförande av uppdraget. Vi har delat in beskrivningen i två delar: *tillvägagångssätt och bortfall*.

---

<sup>59</sup> Detta gäller prov i dessa ämnen som delvis genomförs i form av en laboration eller muntliga framställningar, som inte kan omrättas. Eftersom det inte sätts delprovsbetyg utan endast provbetyg blir någon direkt jämförelse mellan provens ursprungliga betyg och de resultat som erhållits i samband med omrättningen inte möjlig.

## Tillvägagångssätt

Tillvägagångssättet vid omrättningen innehåller följande moment: Urval, insamling av elevlösningar och information om vem som bedömt proven, bedömning av prov, en andra omrättning, överväganden och avgränsningar. De redovisas var för sig nedan.

### Urval

Ett obundet slumpmässigt urval av elevlösningar gjordes bland de obligatoriska proven i grundskolan och gymnasieskolan. Proven omfattar ämnena svenska/svenska som andraspråk för årskurs 6 och ämnena engelska, matematik, svenska/svenska som andraspråk och de naturorienterande ämnena biologi, fysik och kemi i årskurs 9. I gymnasieskolan omfattar urvalet kursproven i engelska 5, matematik 1a, 1b och 1c och svenska 1/svenska som andraspråk 1. Att avstå från omrättningen av de nationella proven i årskursen 3 motiveras med att avvikelserna mellan ursprungs rättarens bedömning och Skolinspektionens bedömning generellt i tidigare omrättningsomgångar visat sig vara mycket små.

Urvalet genomfördes på ett sådant sätt att erhållna resultat skulle vara representativa:

- dels på nationell nivå per ämne och årskurs,
- dels på skolnivå för de utvalda skolorna ämnena engelska, matematik och svenska/svenska som andraspråk i årskurs 9 i grundskolan och gymnasieskolan.

Urvalet drogs från de insamlingar av resultat från nationella prov som SCB genomförde under sommaren och hösten 2012. Totalt sett omfattar urvalet 36 780 prov, fördelat på 3 006 prov från årskurs 6, 19 512 prov från årskurs 9 och 14 262 prov från gymnasieskolan. När ursprungsrättningen av nationella proven genomfördes visste inte skolorna om de var med i urvalet eller inte.

### Insamling av elevlösningar

Det genomförda urvalet innebar att 490 skolor kontaktades och ombads skicka in kopior av bedömda elevlösningar till Skolinspektionen. Skolorna som varit med i urvalet under 2010 eller 2011 belastas i lägre grad än övriga skolor. Omfattningen av arbetsinsatsen varierade för de utvalda skolorna. Målet var att ingen skola skulle skicka fler än runt 100 provkopior.

Med hjälp av SCB fick Skolinspektionen tillgång till det register med provresultat som sammanställs av SCB för Skolverkets räkning. På så sätt kunde skolornas arbetsbörda till viss del minskas. De enda resultat som samlades in direkt från skolorna avsåg poäng och belägg för olika betygsnivåer för prov i matematik, fysik, biologi och kemi för årskurs 9 och gymnasieskolan. Anledningen var att proven i de ämnena inte har några delprovsbetyg utan endast provbetyg. I och med att de proven även innehåller en muntlig eller en laborativ del skulle en jämförelse mellan den ursprungliga bedömningen och Skolinspektionens bedömning vara missvisande.

Kopiorna av elevlösningarna skannades av Rikspolisstyrelsen till digital form för att lättare kunna spridas till de lärare som skulle omrätta proven. Den ursprungliga bedömningen och annan information som kunde leda till identifiering av enskilda elever togs också bort. Det gjorde vi för att lärarna som kontrakteras för uppdraget inte skulle påverkas av den ursprungliga bedömningen.

### **Insamling av information om vem som bedömt proven**

Till varje insamlat prov samlade Skolinspektionen in information kring vilken lärare som bedömt provet och hur bedömningen utförts:

- Lärarens ålder
- Lärarens kön
- Om läraren är behörig att undervisa i det aktuella ämnet och årskursen
- Om läraren diskuterat bedömningar med annan pedagogisk personal
- I vilken konstellation av pedagogisk personal som provet bedömts.

### **Bedömning av proven**

På Skolinspektionens uppdrag kontrakterade företaget Lärarförmedlarna drygt 140 ämnesutbildade och erfarna lärare på visstid för att omrätta proven. Lärarna utförde bedömningen i sina hem. Kommunikationen med lärarna var huvudsakligen digital.

I avtalet mellan Lärarförmedlarna och Skolinspektionen ställdes krav på att lärarna skulle ha relevant lärarexamen och utbildning, enligt skollagens bestämmelser 2 kap. 13–19 §§, för att undervisa i aktuellt ämne och årskurs. Lärarna skulle också ha erfarenhet av att bedöma nationella prov i aktuellt ämne och aktuell årskurs.

Av de lärare som anlätades för innevarande års omrättning deltog 69 procent även i förra årets omrättning. Totalt är ungefär en femtedel av landets kommuner representerade bland de anlätade lärarna.

### En andra omrättning

För att öka kvaliteten på omrättningen har vissa prov omrättats en tredje gång<sup>60</sup> av ytterligare en lärare i de fall det förekom stora avvikelser mellan omrättande lärares bedömning och den ursprungliga bedömningen. Det benämns som en andra omrättning. Med stora eller avgörande avvikelser avses här prov där antingen ursprungs rättarens bedömning eller omrättarens bedömning innebär ett underkännande (IG eller F) eller att eleven ej nått kravnivån. Provet har gått till omrättning för årskurs 9 om det avvikit mer än ett betygssteg åt något håll (till exempel så har antingen omrättaren eller ursprungs rättaren gett betyget G och den andra läraren betyget MVG). För gymnasieskolan har prov omrättats en andra gång om det avvikit mer än två betygssteg åt något håll, alltså om till exempel antingen omrättaren eller ursprungs rättaren gett betyget E och den andra läraren B).

Resultatet av den andra omrättningen har tolkats utifrån principen att det är den bedömning med minst avvikelse från ursprungs rättaren som kommer att gälla. Om kriterierna för den andra omrättningen uppfyllts av ett delprov har samtliga delprov inom ett prov genomgått en andra omrättning. Läraren som omrättar provet en andra gång ska inte veta att det är en andra omrättning som görs för att inte omedvetet påverkas i sin bedömning.

### Överväganden

Skolinspektionen värderar inte om ursprungs rättarens bedömning är felaktig i de fall där Skolinspektionen gör en annan bedömning än ursprungs rättaren. Skolinspektionen kan inte fastslå att de bedömningar som genomförts av Skolinspektionens lärare är mer korrekta än ursprungs rättarens bedömning. Därför talas i rapporten i stället om förekomsten av avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning.

I jämförelser mellan resultatet av ursprungs rättarens bedömning och Skolinspektionens bedömning använde Skolinspektionen den fyrgradiga betygsskalan: Ej uppnått målen/Icke godkänt (EUM/IG), Godkänt (G), Väl godkänt (VG) och Mycket väl godkänt (MVG) samt gymnasieskolans sexgradiga betygsskala för dess ämnen: A-F. Betygssteget har beteckningarna A, B, C, D, E och F där A-E står för godkända resultat och F står för ej godkänt resultat. Det är de skalor som SCB använder när de samlar in provresultat på uppdrag av Skolverket.

---

<sup>60</sup> Eftersom delproven i matematik och de naturorienterade ämnena inte har varit möjliga för omrättaren att betygssätta har det inte gjorts någon andra omrättning av de delproven.



Det finns flera steg i provens bedömningsanvisningar, men dessa kan inte användas eftersom de inte samlas in av Skolverket.

### **Avgränsningar**

Av praktiska skäl samlades bara de delprov in som kunde omrättas. Därför gjordes insamlingen på delprovsnivå. Delprov som genomfördes i grupp, redovisades muntligt och/eller innehöll laborativa moment samlades inte in.

I proven i matematik, biologi, kemi och fysik i årskurs nio och gymnasieskolan ges inte några delprovsbetyg. Ett av de delprov som utgör provet i matematik genomförs i grupp och redovisas muntligt. Ett av de två delprov som utgör provet i biologi, kemi och fysik innehåller delvis ett laborativt moment som Skolinspektionen inte har kunnat omrätta. Därför har det inte heller varit möjligt att bedöma provbetyget. Resultatredovisningen och analysen har därför gjorts på poängnivå för de delprov som redovisas skriftligt.

Svenska och svenska som andraspråk samredovisas i rapporten. Någon särredovisning av svenska som andraspråk har alltså inte gjorts. Urvalet av prov i svenska är inte utformat för att vara representativt i svenska och svenska som andraspråk, utan endast i svenska som ämne.<sup>61</sup> Skolinspektionens anlitade lärare har känt till vilka elevlösningar som avser svenska som andraspråk, där sådan information funnits och det är främst lärare som är behöriga i svenska som andraspråk som bedömt dessa prov. Cirka sju procent av de insamlade proven i årskurs nio och cirka fyra procent av de insamlade proven i gymnasieskolan avsåg svenska som andraspråk.

### **Bortfall**

Av samtliga 36 780 nationella prov som ingick i urvalet skickade skolorna in 31 882 prov, fördelade på 19 740 prov från grundskolan och 12 142 från gymnasieskolan. Av den beskrivning som framkommer i tabell 1 nedan kan man konstatera att det finns ett relativt stort bortfall när det gäller vissa prov.

### **Externt bortfall**

Det externa bortfallet (där provet i sin helhet uteblivit) varierade mellan 8 och 40 procent mellan proven. Den vanligaste anledningen till att skolorna inte skickat in prov är att de inte kunnat hittas eller att de förstörts av misstag. En annan vanlig anledning är att vissa prov har skickats in i original i samband

---

<sup>61</sup> Delprovet för de båda ämnena som prövar läsning innehåller delvis olika uppgifter, vilket kan påverka överensstämmelsen vid bedömningen. Ingen analys av avvikelse har varit möjlig att göra.

med de insamlingar som görs av provkonstruktörer. Det ämne som hade högst externt bortfall var matematik i gymnasieskolan. Vid delprovet 1c var det 40 procent av de begärda proven som inte inkom. För delprovet 1a i matematik var det näst högst externt bortfall, vid 27 procent av de begärda proven. För ämnena svenska/svenska som andraspråk i årskurs 9 och engelska 5 i gymnasieskolan var det externa bortfallet minst, vid 8 procent av de begärda proven.

### Internt bortfall

Eftersom provens utformning ser olika ut för olika ämnen ser även det interna bortfallet olika ut beroende på ämne.<sup>62</sup> En del av de prov som inkom visade sig vara bristfälligt kopierade och därmed inte möjliga att omrätta. Det förekom också att prov som skickats in gällde prov från andra år än det som omrättningen avsåg, det vill säga vårterminen 2012. Det utgör det interna bortfallet.

För ämnena engelska och svenska/svenska som andraspråk beräknas det interna bortfallet per delprov, medan det i ämnena matematik, biologi, fysik och kemi beräknas per typ av poäng (G/VG/MVG-poäng respektive E/C/A-poäng).

Det genomsnittliga interna bortfallet för ämnena engelska och svenska/ svenska som andraspråk för åk 6, åk 9 och gymnasieskolan ligger på 8 procent. Högst andel internt bortfall har svenska/ svenska som andraspråk för gymnasieskolan, 24 procent.

Det interna bortfallet för ämnena matematik, biologi, fysik och kemi för åk 9 och gymnasieskolan ökar i regel med de högre betygsstegen. Det genomsnittliga interna bortfallet för dessa ämnen ligger på 33 procent för VG/MVG poäng respektive C/A-poäng medan genomsnittet för G-poäng respektive E-poäng ligger på 8 procent. Högst internt bortfall i VG/MVG-poängen har fysik åk 9, 70 procent.

Utifrån endast bortfallsandelen är det svårt att avgöra om det går att dra några slutsatser av resultaten i omrättningen eller inte. I vissa fall förefaller bortfallet vara relativt litet. Ett litet bortfall kan vara allvarligt, om bortfallet avviker kraftigt från de svarande med avseende på vad man vill mäta. Av erfarenhet från tidigare omgångar av omrättningar av nationella prov vet vi dock att bortfallet inte är systematiskt och därmed är de resultat som återfinns i rapporten tillförlitliga.<sup>63</sup>

---

<sup>62</sup> Internt bortfall på delprovs- och poängnivå redovisas i detalj för varje tabell i resultatbilagan.

<sup>63</sup> En bortfallsanalys gjordes vid tidigare omgångar av SCB. För de ämnen som hade ett stort internt bortfall, skulle resultatet tolkas med försiktighet.

Tabell 1: Externt bortfall fördelat på årskurs, ämnesprov och skolor

ÅRSKURS	Ämne	Antal be- gärda prov	Urval skolor	Inkomna skolor	Externt bortfall
<b>ÅRSKURS 6</b>		3 006	70	66	
	Svenska	3 006	70	66	10% (310)
<b>ÅRSKURS 9</b>		19 512	300	263	
	Engelska	4 495	70	62	7% (331)
	Matematik	5 255	80	75	11% (581)
	Svenska/svenska som andraspråk	4 510	70	64	8% (341)
	NO (totalt)		80	71	
	Biologi	2 061		26	22% (449)
	Fysik	1 646		24	25% (404)
	Kemi	1 545		21	23% (362)
<b>GYMNASIESKOLAN</b>		14 262	190	151	
	Engelska 5	4 520	60	49	8% (375)
	Matematik (totalt)	5 242	70	75	
	1a				27% (632)
	1b				14% (350)

1c					40% (193)
Svenska 1/svenska som andraspråk 1	4 500	60	53		13% (570)
<b>TOTALT</b>	<b>36 780</b>	<b>560</b>	<b>461</b>		

## 9. Källor

- Alderson mfl. (1995). *Language Test Construction and Evaluation*. Cambridge University Press.
- Andrén och von Horn. (2010). *Att resonera i minsta möjliga mån. Om den lägsta resultatnivån för betyget G på den skriftliga delen av nationella provet i svenska B*. Institutionen för pedagogik, psykologi och idrottsvetenskap. Linnéuniversitetet.
- Berge. (2005). Studie 3 Skriveprøvenes pålitelighet, *Ungdommers skrivekompetanse* Bind I. Red. av Berge m.fl. Oslo.
- Ciolek Laerum. (2009). *Elever skriver och lärare bedömer: en studie av elevtexter i åk 9*. Uppsala universitet, Institutionen för nordiska språk, FUMS Rapport.
- Edander. (2013). *Bedömaröverensstämmelse vid Nationella provet i svenska för gymnasiet kurs 1, skrivdelen*. Institutionen för nordiska språk. Uppsala universitet.
- Erickson. (2009). *Nationella prov i engelska – en studie av bedömarsamstämmighet*. Göteborgs universitet: Institutionen för pedagogik och didaktik.
- Gipps. (1994). *Beyond testing*. London: The Falmer Press.
- Gunnemyr. (2011). *Likvärdighet till priset av likformighet?: en studie av hur och varför svenska och finländska historielärare på gymnasiet uppfattar att de påverkas av externa prov i historia*. Lunds universitet: Historiska institutionen.
- Jönsson. (2011). *Likvärdig bedömning och nationella prov*. Skola och Samhälle. <http://www.skola-ochsamhalle.se/lararutbildning/anders-jonsson-likvardig-bedomning-och-nationella-prov/>
- Jönsson. (2012). *Redskap för rättvis och likvärdig bedömning*. Lärande bedömning. <http://www.larandebedomning.se/presentationer.html>
- Klapp Lekholm. (2008). *Grades and grade assignment: effects of student and school characteristics*. Göteborg studies in educational science, Vol. 269. Göteborg: Acta Universitatis Gothoburgensis.

- Larson. (1984). *Skrivförmåga. Studier i svenskt elevspråk*. Liber.
- Olsson-Wahlsten.(2002). *Öppna svar – hur funkar det?* Institutionen för pedagogik och didaktik. Göteborg: Göteborgs universitet.
- Palmér och Östlund-Stjärnegårdh.(2005). *Bedömning av elevtext. En modell för analys. Svenskämnet i år 8, 9 och gymnasiet*. Natur och Kultur, Stockholm.
- Selghed. (2004). *Ännu icke godkänt. Lärares sätt att erfarar betygsystemet och dess tillämpning i yrkesutövningen*. Malmö Studies in Educational Science, No. 15: Lärarutbildningen Malmö högskola.
- Skolinspektionen. (2010a). *Redovisning av regeringsuppdrag att kartlägga användandet av nationella prov i gymnasial vuxenutbildning samt att undersöka förutsättningarna för central rättning, 2010*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2010b). *Kontrollrättning av nationella prov i grundskolan och gymnasieskolan, 2010*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2011). *Lika eller olika? Omrättning av nationella prov i grundskolan och gymnasieskolan, 2011*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2012). *Lika för alla? Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år, 2012*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2013a). *Tillsyn av bedömning och betygsättning. Med fokus på skolor med stora avvikelser i förhållande till Skolinspektionens omrättning och i relationen mellan nationella prov och slutbetyg. Tillsynsrapport, Rapport T-2013:01*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2013b). *Skolans kvalitetsarbete ger möjlighet till förändring, Skolinspektionens erfarenheter och resultat från tillsyn och kvalitetsgranskning 2012*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2013c). *Navet i skolornas utvecklingsarbete. Ett utdrag ur Skolinspektionens årsrapport. Två exempel på skolors kvalitetsarbete*. Stockholm: Skolinspektionen.
- Skolinspektionen. (2013d). *Bedömning och betygsättning 2013*. Skolinspektionens hemsida: <http://www.skolinspektionen.se/sv/Tillsyn-->

granskning/Kvalitetsgranskning/Skolinspektionen-granskar-kvaliteten/Bedomning-och-betygssattning-2013/.

Skolverket. (2008). *Central rättning av nationella prov*. Stockholm: Skolverket.

Skolverket. (2009). *Bedömaröverensstämmelse vid bedömning av nationella prov*. Dnr 2008:286. Stockholm: Skolverket.

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

Skolverket. (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid*. Stockholm: Skolverket.

Östlund-Stjärnegårdh. (2002). *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter*. Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 57. Uppsala.

Förordning (2011:556) med instruktion för Statens skolinspektion.

Offentlighets- och sekretesslagen (2009:400) kap. 24 Sekretess till skydd för enskild inom forskning och statistik.

## 10. Diagram- och tabellförteckning

Diagram 1: Fördelning av andel avvikelser; positiva, negativa samt ingen avvikelse mellan ursprungs rättarens bedömning och Skolinspektionens bedömning för de olika delproven, omgång 4.....	14
Diagram 2: Total andel avvikelse för de nationella proven i svenska och engelska i årskurs 9 och gymnasieskolan, omgång 1-4.....	15
Diagram 3: Andel avvikelse i betygsättning mellan ursprungs rättaren och omrättaren, delprov i svenska och engelska i årskurs 9, omgång 4.....	17
Diagram 4: Prov med avvikelser som omrättats en andra gång, för respektive delprov, omgång 4.....	18
Diagram 5: Kvartilspridning över andel avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning för gymnasieskolan gällande delprov i engelska 5 och svenska 1/svenska som andraspråk 1, för gymnasieskolorna i urvalet, omgång 4.....	20
Diagram 6: Kvartilspridning över andel avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning för grundskolan gällande delprov i engelska och svenska/svenska som andraspråk, för grundskolorna i urvalet, omgång 4.....	21
Tabell 1: Externt bortfall fördelat på årskurs, ämnesprov och skolor.....	43-44