

Lika för alla?

Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år



2012-08-31

Förord

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

Skolinspektionen redovisar här till regeringen resultatet av tredje årets rättning. Redovisningen omfattar en presentation och analys av resultatet av de tre årens genomförda rättningar, en fördjupad analys samt en tillbakablick med slutsatser och erfarenheter från alla tre åren.

Vid Skolinspektionen har följande personer arbetat med att genomföra uppdraget: Arletta Plunkett (projektledare), Lovisa Cottrell och Jessica Hörnsten.

Ann-Marie Begler

Generaldirektör

Gabriella Bremberg

Avdelningschef

Innehållsförteckning

INNEHÅLLSFÖRTECKNING	3
1. SAMMANFATTNING	4
2. INLEDNING	8
2.1 UPPDRAGET.....	10
2.2 TERMINOLOGIN.....	11
3. SLUTSATSER OCH RESULTAT EFTER ÅRETS OMGÅNG	12
3.1 DE NEGATIVA AVVIKELSERNA DOMINERAR.....	14
3.2 OLIKA RESULTAT BEROENDE PÅ SKRIVUPPGIFT.....	16
3.3 ELEVPRODUCERADE TEXTER.....	17
3.4 SKILLNADER MELLAN SKOLOR.....	19
4. TRE ÅR MED OMRÄTTNING – EN JÄMFÖRELSE	20
5. SYSTEMATISKA SKILLNADER I LÄRARNAS BEDÖMNINGAR	22
6. REKOMMENDATIONER	23
7. GENOMFÖRANDE AV UPPDRAGET	30
7.1 URVAL.....	30
7.2 BORTFALL.....	35
7.3 BRISTER OCH PROBLEM.....	37
8. KÄLLOR	38
9. DIAGRAMFÖRTECKNING	39

1. Sammanfattning

– lika stora avvikelser under tre år

Under tre år har Skolinspektionen på regeringens uppdrag samlat in och omrättat nationella prov. Sammanlagt har cirka 94 000 prov på delprovnivå samlats in från ungefär 1 800 skolor. Den här rapporten sammanfattar de tre omgångarnas omrättning.

Proven avser i denna omrättningssomgång ämnena engelska, matematik och svenska/svenska som andraspråk och de naturorienterade ämnena (biologi, fysik och kemi) för årskurs nio. I gymnasieskolan ingår kurserna engelska A, matematik A och svenska B/svenska som andraspråk B. Proven har genomförts under 2011. Tidigare år har även delprov i matematik och svenska/svenska som andraspråk för årskurs tre och fem, samt engelska för årskurs fem ingått i omrättningen.

Resultatet av omrättningen redovisas genomgående som avvikelser i bedömningarna mellan den lärare som ursprungligen rättade provet (ursprungsrättare) och den som för Skolinspektionens räkning rättat samma prov (omrättare). Som "avvikelse" räknas de fall där ursprungsrättarens och omrättarens bedömning skilt sig åt med minst ett betygssteg.¹ Andelen avvikelser för ett visst prov anges i procent av det totala antalet omrättade lösningar för samma prov.

Samtliga lärare som omrättat prov för Skolinspektionens räkning är behöriga i de ämnen de rättar prov i och har erfarenhet av att rätta nationella prov. En stor del av dem har medverkat vid samtliga tre omrättningssomgångar. Lärarna som omrättar prov i svenska/svenska som andraspråk har dessutom genomgått en utbildning i Skolinspektionens regi, med provkonstruktörerna som utbildare. Skolinspektionen har därför en stor tilltro till deras bedömningsförmåga. I de fall som Skolinspektionen kommer till en annan bedömning än ursprungsrättaren av hur eleven presterat på provet görs ingen värdering av om ursprungsrättarens bedömning är riktig eller felaktig. Däremot kan man konstatera att minst två lärare i dessa fall har gjort olika bedömningar av samma elevs prestation på provet.

¹ För de delprov där provbetyget inte kan användas för jämförelse, avser avvikelsen skillnad i poängsättning.

Resultat

Resultatet av omrättningarna visar att det, såväl för senaste omrättningsomgången som för de tre omgångarna sammantaget, generellt sett finns omfattande och stora avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning av vissa delprov. Det gäller framför allt de delprov där elevens svar ska vara i form av en längre redogörelse, och särskilt för uppsatsskrivning. Tittar man på andelen avvikelser ämnesprov för ämnesprov spänner avvikelserna i den senaste omrättningen mellan 10 och 45 procent² Det innebär att det finns delprov där det vid nästan vartannat prov finns skillnader mellan ursprungs rättarens bedömning och Skolinspektionens bedömning. Avvikelserna mellan ursprungs rättarens bedömning och Skolinspektionens bedömning är avsevärt mindre och ibland försumbara för de delprov där svaret har karaktären av rätt-eller-fel. Det gäller delprov A reading and listening i engelska på gymnasiet, delprov B i engelska för årskurs nio och delprov A i matematik på gymnasiet samt de naturorienterande ämnena i årskurs nio.

Avvikelserna mellan ursprungs rättarens bedömning och Skolinspektionens bedömning går oftast i negativ riktning, det vill säga att Skolinspektionen bedömer att elevens prestation motsvarar ett lägre betyg än vad ursprungsrättaren har bedömt. Avvikelser i positiv riktning, det vill säga att Skolinspektionen bedömer att elevens prestation motsvarar ett högre betyg än vad ursprungsrättaren gör, förekommer också förhållandevis ofta men inte i samma utsträckning som avvikelser i negativ riktning. Det är dessutom de högre betygsstegen som står för den största andelen avvikelser.

Resultatet från analysen på skolnivå visar att avvikelserna mellan ursprungsrättarens bedömning och Skolinspektionens bedömning varierar kraftigt mellan olika skolor. Den skola som har störst avvikelse avviker mer än tre gånger så ofta som den skola som har minst antal avvikelser.³

² Engelska A på gymnasiet har för delprovet reading och listening en total avvikelse mellan ursprungsrättaren och kontrollrättaren på 10 procent och svenska B/svenska som andra språk delprov B har en avvikelse på 45 procent.

³ 21 procent i den skola som uppvisade minst antal avvikelser, mot 76 procent i den skola som uppvisade flest. I Omrättningsomgång två fanns hos en av skolorna en negativ avvikelse på 95 procent för delprov B i svenska/svenska som andra språk.

I uppdraget till Skolinspektionen preciserar regeringen ett antal analyser som ska genomföras inom ramen för Skolinspektionens omrättning. Skolinspektionen har anlitat forskaren Björn Tyrefors Hinnerich vid Stockholms universitet som tillsammans med Jonas Vlachos har genomfört en analys av vad som kan förklara avvikelserna mellan den ursprungliga rättningen och omrättning av nationella prov utifrån tre huvudområden: elevernas karaktäristika, ursprungsrättarnas egenskaper och olika typer av skolor. Forskarnas analys redovisas i bilaga 4.

Slutsatser

Resultatet efter tre års omrättning är samstämmigt mellan åren: Vid nästan varannan omrättad delprovsuppsats i Svenska i årskurs 9 och på gymnasiet finns skillnader mellan ursprungsrättarens och Skolinspektionens bedömning av proven. Stora skillnader förekommer även vid uppsatser eller längre självständigt författade redogörelser inom andra ämnen.

Forskning (exempelvis Wenche Vagle " Ungdommers skrivekompetanse", 2005) visar att även under mer ideala förhållanden då rättarna intensivt har tränats för att ha en god förmåga att bedöma uppsatser är det svårt att få en så hög samstämmighet att de kan utgöra ett stöd att kunna göra rättsäkra och likvärdiga bedömningar av provresultaten. Skolinspektionens bedömning är därför att uppsatser och längre självständigt författade redogörelser är ett alltför trubbigt verktyg för att kontrollera likvärdigheten i betygssättning och bedömning som är ett av de nationella provens två huvudsyften.

Skolinspektionen lämnar därför rekommendationen att de delprov som är i form av uppsatser eller längre självständigt författade redogörelser lyfts bort ur de nationella proven och genomförs i andra sammanhang.

Skolinspektionen dock vill framhålla att de förmågor och kunskaper som testas genom uppsatsskrivande är viktiga delar av kursmålen. Uppsatser uppfyller därför sin roll för att ge läraren ett underlag för att bedöma i enlighet med de nationella provens andra syfte "i vilken utsträckning kunskapsmålen uppfylls". Skolinspektionens mening är dock att detta inte behöver göras inom ramen för de nationella proven eftersom samma delprov inte stödjer syftet med likvärdig bedömning. Andra modeller och former för att ge lärare stöd att göra dessa bedömningar bör således utvecklas.

Om uppsatsskrivande av andra skäl än att stödja en rättvis och likvärdig bedömning ska finnas kvar i de nationella proven bör en rad åtgärder vidtas för att öka samstämmigheten mellan de olika personerna som rättar proven. Dessa

slutsatser redovisas nedan och överensstämmer i stor utsträckning med de slutsatser som redan drogs utifrån förra årets omrättning.

De *bedömningsanvisningar* som ska vägleda rättningen av proven är alltför omfattande och komplexa och ger bitvis stort utrymme för tolkningar. Skolinspektionen föreslår här att anvisningarna omarbetas med målsättningen att göra dem tydligare och mer tillgängliga framförallt för de högre provbetygen. Exempellösningar bör publiceras elektroniskt. Vidare vill Skolinspektionen understryka vikten av att rättande lärare ska erbjudas kontinuerlig *fortbildning* i att bedöma och att *sambedömning* ska eftersträvas när det är möjligt.⁴

Den stora variationen i (framför allt negativ) avvikelse mellan ursprungs rättning och kontrollrättning vittnar om *systematiska hinder* för att uppnå en likvärdig bedömning. En del av dem kan möjligen bero på svårigheter för den rättande läraren att förhålla sig objektiv vid bedömningen av sina egna elevers prov. Vi kan heller inte utesluta att det föreligger stora kulturella och värde-mässiga skillnader mellan olika skolor. Skolinspektionen föreslår här ett rättningsförfarande där elevens identitet är okänd för rättande lärare.

Vi kan konstatera att omrättningen bidrar till likvärdigheten. De skolor som följts upp på grund av omfattande avvikelser vid första omrättningsomgången, har i dag något färre avvikelser än tidigare. Resultaten pekar på att Skolinspektionen även i fortsättningen bör göra en förstärkt tillsyn av de skolor som varit med i omrättningen i de fall det förekommer stora avvikelser mellan de egna lärarnas och Skolinspektionens bedömning av proven. För att kunna göra detta i samband med den regelbundna tillsynen och samtidigt granska skolornas bedömning och betygssättning i vidare bemärkelse behöver de delar i uppdraget till Skolinspektionen som handlar om analyser på nationell nivå tas bort.

⁴ Med samrättning 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 kraftigt och att skolans elevsammansättning spelar en allt större roll för en elevs resultat (Skolverket, 2012, sid. 85).

Skolinspektionen fick för tre år sedan 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".

Skolinspektionen har sedan 2010 årligen gjort ett urval av elevlösningar och låtit andra lärare än elevens lärare på nytt rätta elevlösningarna. De två olika bedömningarna har sedan jämförts med avseende på avvikelser mellan bedömningarna. Resultaten visar på stora avvikelser för vissa ämnesprov och mindre för andra. Vi kan dock konstatera att nivåerna har varit de samma för de olika proven alla tre åren.

Den här rapporten fokuserar på likvärdigheten i rättningen av nationella proven och sammanfattar tre års omrättningar. Skolinspektionen har också låtit två forskare analysera förekomsten av systematiska skillnader i bedömningarna. Deras redovisning återfinns i resultatbilagan. 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. Den centrala rättningen innebär en kvalitetssäkring av likvärdigheten i provsystemet genom att skapa möjlighet att upptäcka felbedömningar som äventyrar systemets tillförlitlighet.

Stor skillnad i bedömning av samma elevlösning, beroende på rättande lärare, är att betrakta som något allvarligt. Skolinspektionen avgör inte om "felbedömningarna" ligger hos lärarna som rättade proven i första hand eller om de ligger hos lärarna som omrättat proven. Vi kan inte med säkerhet säga att de lärare som av Skolinspektionen fått uppdraget att kontrollrätta elevlösningarna

är mer korrekta i sina bedömningar än de lärare som ursprungligen rättat proven. Vad vi vet är att samtliga lärare som deltagit i omrättningen är behöriga i de ämnen de rättar prov i, de har erfarenhet av att rätta nationella prov och att de disponerar lång tid för sin uppgift.⁵ Lärarna som omrättar proven i svenska har dessutom genomgått en utbildning i Skolinspektionens regi, med provkonstruktörerna som utbildare. Därför har vi all anledning att sätta stor tilltro till deras bedömningsförmåga. Flera av de lärare som har omrättat prov har medverkat vid samtliga tre omrättningsomgångar.

I den tredje omrättningsomgången har inte prov i årskurs tre och fem ingått, vilket de gjort i tidigare omgångar. Anledningen är att för årskurs tre har avvikelserna mellan ursprungs rättningar och kontrollrättningar generellt varit mycket låga. Beslutet att inte rätta prov för årskurs fem beror däremot på att proven läsåret 2011/2012 inte längre var obligatoriska för årskursen, utan i stället kommer att ersatts med nationella prov för årskurs sex. Nationella prov för årskurs 6 genomförs för första gången 2012 och omfattas därför inte av den senaste omrättningsomgången som bygger på prov som genomfördes 2011.

⁵ Ursprungs rättarna har en vecka på sig att rätta sina prov, omrättarna har tre månader på sig att rätta 100 prov.

2.1 Uppdraget

Uppdragsbeskrivningen finns i sin helhet i bilaga 5, men kan i korthet sammanfattas till att Skolinspektionen ska:

- under tre år 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 felbedömningar
- använda sig av resultaten från denna analys som underlag för fortsatt tillsyn och kvalitetsgranskning (Resultaten från tidigare omgångar har använts vid tillsyn under 2012 och redovisas i samband med årsredovisningen 2012)

Vidare står det i Skolinspektionens regleringsbrev 2012 att:

Skolinspektionen ska samråda med Skolverket när det gäller verkets uppdrag att utveckla nationella ämnes-, kurs-, och slutprov [...] Utifrån Skolinspektionens erfarenheter från arbetet med regeringsuppdraget om viss central rättning av nationella prov samt förstärkt tillsyn över huvudmännens och skolornas arbete med nationella prov [...] ska Skolinspektionen föreslå de ändringar i den praktiska utformningen av de nationella proven som myndigheten anser vara nödvändiga för uppdragets effektiva genomförande.

Dessa förslag till praktiska ändringar beskrivs i kapitel 6 – Rekommendationer.

2.2 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*.

Omrättningarna gjordes 2010, 2011 och 2012. De omfattade då prov som genomfördes 2009, 2010 och 2011. För enkelhetens skull kommer vi att benämna första omrättningen *omgång 1*, den andra *omgång 2* och slutligen den tredje *omgång 3*.

Resultaten av omrättningen redovisas som avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens 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 ett högre betyg/poäng än vad ursprungsrättaren har bedömt.

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

3. Slutsatser och resultat efter årets omgång

År 2012 har Skolinspektionen rättat om 29 700 elevlösningar av nationella prov på delprovsnivå från 452 skolor. Proven avser ämnena engelska, matematik och svenska/svenska som andraspråk samt de naturorienterande ämnena (biologi, fysik och kemi) i årskurs nio. Från gymnasieskolan ingår kurserna engelska A, matematik A och svenska B/svenska som andraspråk B. Proven genomfördes under vårterminen 2011.

Resultaten i korthet:

- Största avvikelserna återfinns för de delprov där eleven ska avge sitt svar i längre text.
- Negativa avvikelser är vanligare än positiva.
- Det är de högre betygsstegen som står för den största andelen avvikelser.
- Andelen avvikelser varierar starkt mellan olika skolor.

Liksom tidigare år finns de flesta avvikelserna bland prov i ämnena svenska och engelska. Matematik och de naturvetenskapliga ämnena har fastare ramar för rätt och fel svar. De lämnar ett litet utrymme för eleven att själv författa längre utläggningar och för den rättande läraren att göra egna tolkningar och bedömningar. Det är framför allt uppsatsskrivandet som är behäftat med stora avvikelser i bedömningarna. Det gäller såväl för svenska/svenska som andra språk som för engelska.

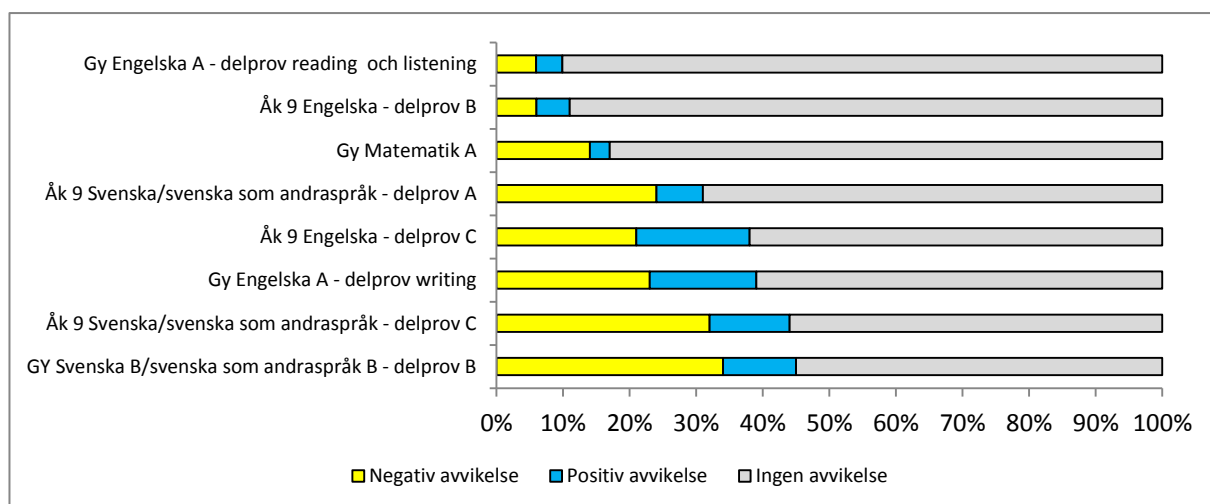
Vi ser även tydligt att det är betygen VG och MVG som de rättande lärarna över lag är minst ense om. De naturvetenskapliga ämnena utgör dock ett undantag även här.

Andelen avvikelser mellan ursprungrättaren och Skolinspektionens bedömning bland samtliga omrättade prov skiljer sig stort mellan olika skolor (från cirka 20 procent till över 70). De skolor som uppvisade störst avvikelser i den första omrätningsomgången, samt några skolor som var föremål för riktad tillsyn på det området, har återigen ingått i den tredje omgångens omräkning. Skolinspektionen kan konstatera att dessa skolor inte längre skiljer sig från genomsnittet. Skolinspektionen ser detta som en tydlig indikation på det fruktbara i

att koppla samman omrätningsresultat på skolnivå med ordinarie tillsynsverk-samhet.

I 45 procent av de rättade proven i svenska B/svenska som andraspråk B i gymnasieskolan och i 44 procent i svenska/svenska som andraspråk i årskurs nio, finner vi avvikelser mellan ursprungs rättarens och kontrollrättarens be-dömning.

Diagram 1: Fördelning av andel avvikelser; positiva, negativa samt ingen avvikelse mellan ursprungsrät-tarens bedömning och Skolinspektionens bedömning för de olika delproven, omgång tre.



Samma tendens, som för de delprov där eleven ska avge sitt svar i form av en uppsats, finns i de delar av kursproven som syftar till problemlösning i mate-matik A i gymnasieskolan. Eleven ska här avge sitt svar i form av en beskri-vande text.⁶ I Skolverkets studie, *Bedömaröverensstämmelse vid bedömning av nat-ionella prov*, beskrivs tre typer av elevlösningar i matematik, som ofta är föremål för skillnader i bedömningarna: a) där eleven har angett ett svar, men där reso-nemanget inte är av den typ som återfinns i bedömningsanvisningarna; b) där elever har räknat rätt men inte kommit fram till någon slutsats; och c) där ele-

⁶ En liknande tendens återfinns även för proven i matematik och naturorienterande ämnen för årskurs nio. Men de ger inte samma möjlighet till jämförelse, då Skolinspektionen inte kan göra en avstämning av ursprungsrättarens bedömning i förhållande till kontrollrättarens bedömning. Detta då det endast är delar av provet som går att omrätta på grund av att vissa svar redovisas muntligt eller genomförs i grupp. Betyg sätts endast för hela provet och inte för enskilda delprov, som är fallet vid de flesta andra prov.

ven har presenterat en bra analys och bra slutsatser, men utifrån felaktiga beräkningar.

Uppsatserna i engelska ger färre avvikelser än de i svenska/svenska som andra språk. Orsaken är oklar men lärdom bör dras från bedömningsanvisningarna i engelska.

3.1 De negativa avvikelserna dominerar

De så kallade "negativa avvikelserna" är långt vanligare än de "positiva". För samtliga rättade delprov gäller att det är vanligare att ursprungsrättaren har satt ett högre betyg eller bedömt att eleven i större utsträckning har nått kravnivån än vad omrättaren gjort, än att omrättaren är den som har satt det högre betyget, för de fall där de inte är överens.

Flera studier visar att lärare i sina bedömningar av prov inte sällan tar 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. I studien *Grades and grade assignment: effects of student and school characteristics* konstaterar Klapp och Lekholm att betygssättning i allmänhet är influerad av en rad olika faktorer vid sidan av de kunskaper eleven uppvisat.⁷

Omvänt kan vi föreställa oss att de lärare Skolinspektionen anlitar gör en striktare tolkning av bedömningsanvisningarna. Flera av dem säger att de upplever att de följer bedömningsanvisningarna striktare när de omrättar på Skolinspektionens uppdrag, än när de bedömer de elever de själva har undervisat. Det kan heller inte uteslutas att det i högre grad är lärare med ett specialintresse för bedömningsfrågor som söker sig till uppgiften att omrätta prov för Skolinspektionens räkning.

De nationella proven är diagnostiska prov, inte examensprov, och elevens slutbetyg eller kursbetyg behöver inte överensstämma med betyget på provet. Vid betygssättningen ska läraren bedöma elevens måluppfyllelse i ämnet utifrån de ställda betygskriterierna. Läraren ska ta hänsyn till elevens prestation på det nationella provet, men läraren förväntas vid betygssättningen även utnyttja all

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

tillgänglig information om elevens kunskaper i förhållande till kraven i kursplanen och göra en allsidig bedömning av kunskaperna (Skolverket 2004). I "Nationella prov i gymnasieskolan – ett stöd för likvärdig betygsättning" (Skolverket 2005) beskrivs:

Hur lärare konkret går till väga när de sätter betyg och vilka underlag som de bygger sin betygsättning på, skiljer sig mellan ämnen i den här studien. Troligen kan skillnaden hänföras så väl till ämnenas olika karaktär, som till att olika kunskapssyn och förhållningssätt till bedömning präglar de olika ämnena. Skillnader finns emellertid även mellan skolor (sid. 123-4).

3.2 Olika resultat beroende på skrivuppgift

Graden av avvikelse mellan ursprungs rättarens och kontroll rättarens bedömning av uppsatsproven skiljer sig åt beroende på vilken uppsatsuppgift som har bedömts. Detta gäller såväl för gymnasieskolan som för årskurs nio, men variationerna är störst för gymnasiet.

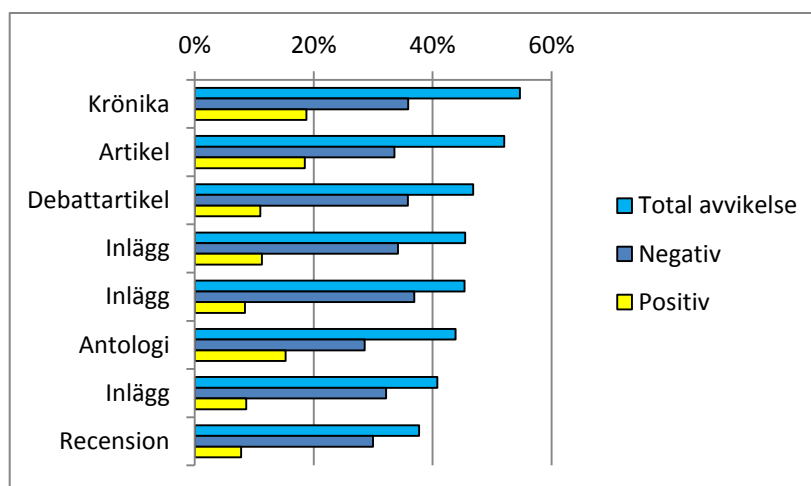
För uppsatsskrivning i svenska B/svenska som andraspråk B kan eleverna på gymnasiet välja mellan åtta olika skrivuppgifter, medan de för svenska/svenska som andraspråk i årskurs nio kan välja mellan fyra. För uppsatsskrivning i engelska har eleverna såväl i årskurs nio och som på gymnasiet att välja mellan två uppsatsämnen. De åtta skrivuppgifterna i svenska för gymnasiet fördelar sig över olika genrer, exempelvis inlägg i en debatt, bidrag till en antologi, en recension, en krönika etc.

Graden av avvikelse mellan ursprungsrättarnas och kontrollrättarnas bedömning varierar kraftigt, beroende på vilken skrivuppgiften är. För vissa skrivuppgifter avviker omrättarens bedömning från ursprungsrättarens med minst ett betygssteg i mer än 60 procent av de omrättade provsvaren, medan andra skrivuppgifter endast har en motsvarande avvikelse om cirka 30 procent. Vi kan även se att den absolut vanligaste avvikelsen är negativ, det vill säga att omrättaren gör en lägre bedömning än ursprungsrättaren.

Vad som däremot är svårt att säga är i vad mån detta är avhängigt av skrivuppgiftens typ av genre, eller om det är andra faktorer som främst spelar in. En jämförelse av skilda skrivuppgifter tillhörande samma genre, men från olika rättningsomgångar, visar på en mycket stor spridning för graden av avvikelse. En faktor som ofrånkomligen bör tas med i beräkningen är att det på samma delprov i vissa fall finns både uppgifter för vilka endast betygen IG–VG ges, och andra med betygsumfånget IG–MVG. Ett MVG är alltså inte möjligt att uppnå i det förra fallet.⁸ Det kan självfallet spela en roll, eftersom vi har sett att det är i de högre betygsskalorna vi generellt finner de största avvikelserna.

⁸ Inom ramen för de prov som har omrättats, gäller det för radiokrönikan från rättningsomgång två och antologiuppgiften från omgång tre. (I omgång ett insamlades inte någon information om vilka skrivuppgifter eleverna hade valt). Vi kan även konstatera att ämnena tillhörde de mindre populära valen.

Diagram 2: Gymnasiet, svenska skriftlig del, olika genre man kan välja mellan att skriva, uppdelat på skrivuppgift (tre skrivuppgifter var i genren inlägg). Andel positiva och negativa avvikelser samt avvikelser totalt per skrivuppgift, omgång 3.



Det aktuella materialet tillåter inte någon djupare analys av skillnaderna i avvikelser mellan olika uppsatsämnen. Men de stora variationer vi funnit antyder att det finns anledning att titta närmare på fenomenet i kommande omrättningsomgångar. Tidigare studier antyder att "enskilda lärare betonar genreaspekten olika och att krönikan tycks vara den mest problematiska genren om man ser till bedömaröverensstämmelsen" (Skolverket, 2009, sid. 25). Men då ska sägas att genren "krönika" var ny när studien genomfördes 2009.

3.3 Elevproducerade texter

Utifrån syftet att uppnå en likvärdig betygssättning av proven, visar resultaten på en problematik som är mest påtaglig för de delprov som avser uppsatser och essäfrågor. En stor del av undervisningen i språk, svenska och engelska i det här fallet, ska resultera i ett för eleven väl utvecklat skrivspråk varför dessa är speciellt problematiska. Borgström och Yassin påtalar i sin forskning kring bedömning av läs- och skrivförmåga, att de nationella provens ambitiösa syftesbeskrivning gör att varje provkonstruktion måste förhålla sig till svårförenliga och ibland motstridiga ingångsvärden, vilket gör att provens syften blir svåra att uppnå.⁹ Provens syften är i dag långt mer avgränsade och tydligt framskrivna, men det hindrar inte att de ibland ändå kan vara svåra att uppnå i ett och samma prov.

⁹ Borgström och Yassin (2010), "Tioåringar skriver nationellt prov", i *Att bygga broar – kulturella, språkliga och mediala möten/sjunde Nationella konferensen i svenska med didaktisk inriktning*. Se även Borgström (2010), "Att skriva prov – Om normer och textstruktur i gymnasieskolans skriftliga nationella prov", i *Språk & Stil* NF 20 2010.

Det är självklart viktigt att eleven lär sig bemästra analyser, problemlösningar och eget författande, men bedömningen av dessa färdigheter blir ofrånkomligen mer subjektiv. Ju mer kreativ och självständig eleven tillåts, eller rent av uppmuntras, att vara i sina svar, desto svårare blir det att skapa tydliga och allmänna bedömningsanvisningar och för lärarna att göra likvärdiga bedömningar. Samtidigt som vi varken vill eller kan hindra att eleven från att lösa problem och uppgifter på sätt som provkonstruktörerna inte kunnat förutspå, får detta konsekvenser för likvärdigheten i betygsättningen.

Även bedömningsanvisningarnas karaktär ger upphov till problem kring likvärdig bedömning. Bedömningsanvisningarna är mycket omfattande (för svenska B/svenska som andraspråk B överstiger de 65 sidor och för svenska/svenska som andraspråk för årskurs nio överstiger de 80 sidor). Det gör dem svåra att ta till sig. Samtidigt vet vi av erfarenhet att många lärare uppskattar exempellösningar och gärna skulle se fler. En möjlig lösning på dessa motstridiga önskemål skulle kunna vara att ge anvisningarna en tydligare struktur, där det är lättare att navigera mellan anvisningar och exempel för olika betyg. Om det på ett säkerhetsmässigt tillfredställande sätt gick att lägga upp anvisningarna på exempelvis på Skolverkets webb-plats, skulle det underlätta möjligheten att snabbt finna det som är relevant för det aktuella sammanhanget.

Anvisningarna för värderingen av elevens prestation mellan de olika betygsstegen utgår till stora delar från olika värdeord som inte alltid är ömsesidigt uteslutande. Därför finns det ett stort utrymme för olika tolkningar. Bedömningen kan därmed påverkas av bedömarens erfarenheter och utgångspunkter. Exempel på detta är följande formuleringar från bedömningsanvisningarna för respektive betygssteg under rubriken "Språk och stil": "Språket är huvudsakligen tydligt, om än inte helt korrekt." (Godkänt). "Språket är klart, varierat och i stort sett korrekt. Stilen är anpassad till situationen." (Väl godkänt). "Språket är stilistiskt säkert och väl anpassat till situationen." (Mycket väl godkänt).¹⁰ I samma bedömningsanvisningar förekommer även beskrivningar av att diskussionen i elevens text ska vara "fokuserad", "läsvärd", "väl genomtänkt" och att "[...] diskussionen i viss mån är koncentrerad [...]"

¹⁰ Bedömningsanvisningar för nationella provet i svenska B/svenska som andraspråk B i gymnasieskolan, vårterminen 2011.

3.4 Skillnader mellan skolor

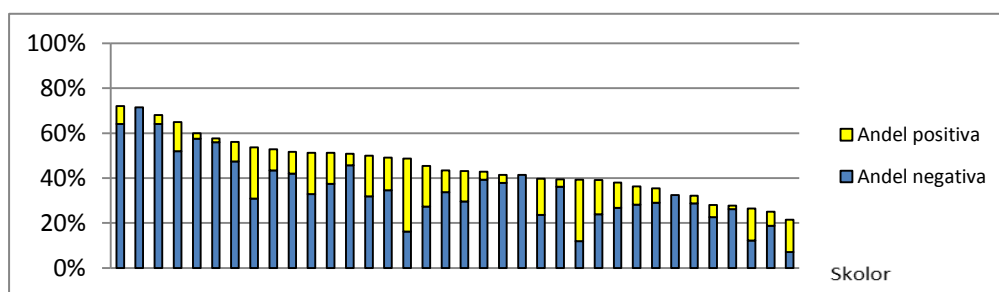
De nationella prov som Skolinspektionen har omrättat har samlats in från ett urval av skolor. Urvalet har i enlighet med uppdragsbeskrivningen utformats så att det är representativt på nationell nivå per ämne (svenska/svenska som andraspråk, matematik¹¹, engelska och naturorienterande ämnen) och årskurs. Det kan också ses som representativt på skolnivå vid de utvalda skolorna i ämnena svenska/svenska som andraspråk, matematik och engelska i årskurs 9 i grundskolan samt i gymnasieskolan.

För vissa skolor finns kraftiga avvikelser mellan ursprungs rättarens bedömning och Skolinspektionens bedömning. Skolinspektionen kommer att kontakta de skolor som i årets omrättningsomgång har uppvisat mest omfattande avvikelser, för att närmare granska hur skolan arbetar med bedömning och betygssättning. På motsvarande sätt följdes i den riktade tillsynen de skolor upp som uppvisade särskilt stora avvikelser i omrättningsomgång två. Resultaten för några av dessa skolor återfinns i resultatbilagan. I korthet kan sägas att de nu inte längre skiljer sig från övriga skolor när det gäller avvikelser.

I diagram 3 visas andel avvikelser mellan ursprungsrättningen och omrättningen för delprov C svenska/svenska som andraspråk av dem som omrättats på respektive skola. Resultatet redovisas som negativ och positiv avvikelse, där varje stapel representerar en skola och den blå delen negativa avvikelser och den gula delen positiva avvikelser.

De ämnesprov som har störst totalt antal avvikelser, är även de där vi återfinner det största spannet av avvikelser mellan olika skolor. I resultatbilagan återfinns diagram för övriga ämnesprov uppdelat på olika skolor.

Diagram 3: Avvikelser mellan ursprungsrättarens bedömning och Skolinspektionens bedömning för gymnasiet gällande svenska, skriftlig del för respektive skola i urvalet, omgång tre.

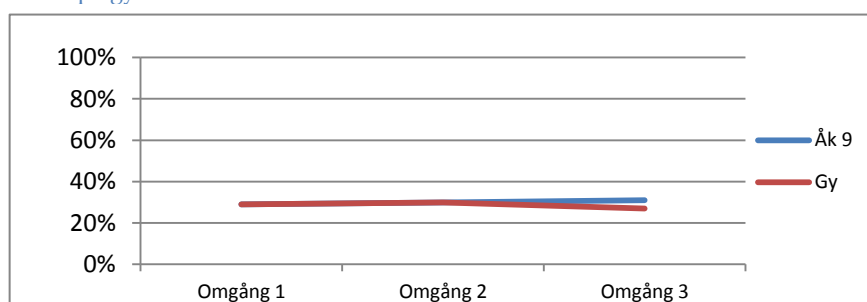


¹¹ För matematikprovet för årskurs nio redovisas inte resultatet av omrättningen på skolnivå. Skolinspektionen kan här inte jämföra ursprungsrättarens bedömning i förhållande till Skolinspektionens, eftersom det endast är delar av provet som går att omrätta på grund av att vissa svar redovisas muntligt och genomförs i grupp. Betyg sätts endast för hela provet och inte för enskilda delprov som är fallet för de flesta andra prov.

4. Tre år med omrättning – en jämförelse

Generellt ser vi mycket snarlika resultat över de tre omrättningsomgångarna. Som framgår av diagram 11, har den totala andelen avvikande prov legat på samma nivå vid alla tre omrättningstillfällen.

Diagram 4: Total andel avvikelse för de nationella proven i svenska och engelska för årskurs nio och för de nationella proven i svenska, engelska och matematik på gymnasiet.



Även Skolverket kan i sin rapport *Provbetyg-Slutbetyg-Likvärdighet* omfattande den betydligt längre tidsperioden 1998–2006, konstatera en mycket liten variation för andelen avvikelser över tid.

Det finns alltså inget som tyder på att ursprungsrättarnas och omrättarnas bedömningar av elevers prestationer, varken för nationella prov eller generellt, skulle ha blivit mer samstämmiga. I ovan nämnda rapport förklarar man detta med det relativt stora utrymme som finns för lärarna och skolorna att göra sina egna tolkningar och bedömningar inom ramen för läro- och kursplaner ((Skolverket, 2007, sid44).

När vi jämför alla de tre omgångarna av omrättning, ser vi att den negativa avvikelsen (där ursprungsrättaren sätter ett högre betyg alternativt i högre grad anser målen uppnådda, än vad omrättaren gör) överlag är större än den positiva (där omrättaren bedömer betyget vara högre än vad ursprungsrättaren gör). För årskurs nio i svenska ser vi att den positiva avvikelsen dock ökat något, medan den negativa minskat något. Samma tendens ser vi även för det nationella provet i ämnet svenska B/svenska som andraspråk B på gymnasiet.

För de nationella proven i engelska, såväl för årskurs nio som för gymnasiet, är fördelningen mellan positiva och negativa avvikelser jämnare. Inte bara som konstaterats tidigare för den tredje rättningsomgången, utan även för de bägge

tidigare omgångarna. Också för det skriftliga delprovet i engelska är andelen positiva och negativa avvikelser förhållandevis jämt fördelade. I graferna nedan, till vänster kan man utläsa avvikelsernas riktning; de negativa avvikelserna visas med ett rött streck och de positiva avvikelserna med ett grönt.

När vi tittar på **samstämmigheten** mellan ursprungs rättare och omrättare över tid, finner vi att de genomgående ligger närmare varandra i sina bedömningar för de lägre betygen och längre ifrån varandra för betygen VG och MVG. Den låga samstämmigheten för högre betyg är ett fenomen som är genomgående för samtliga ämnen och årskurser. I diagrammet 5 exemplifieras detta i diagrammen till höger. Där Ej Uppnått Mål (EUM) visar med ett blått streck, och Godkänt (G) med ett rött streck en samstämmighet mellan ursprungsrättaren och omrättaren som ligger på runt 80% alla tre omrätningsomgångarna.

Diagram 6: Årskurs 9, Engelska skriftlig del. Andel positiva och negativa avvikelser samt avvikelser totalt.

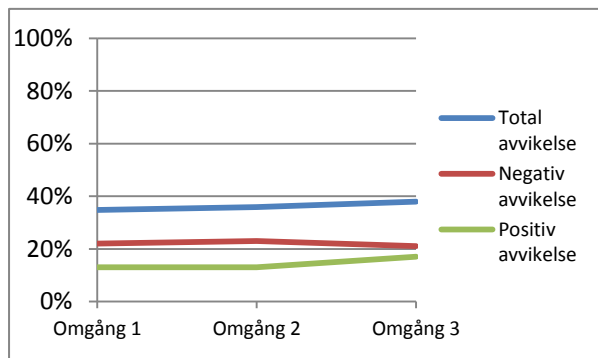
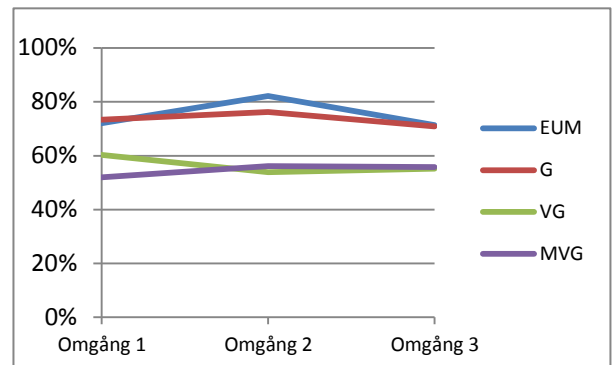


Diagram 5: Årskurs 9, Engelska skriftlig del. Andel samstämmighet mellan ursprungsrättare och omrättare över tid fördelat på betyg.



5. Systematiska skillnader i lärarnas bedömningar

I uppdragsbeskrivningen och i regleringsbrevet avseende Statens skolinspektion 2011 preciserar regeringen ett antal analyser som ska genomföras inom ramen för Skolinspektionens omrättning:

*[...] Skolinspektionen ska så långt möjligt analysera avvikelserna mellan ursprunglig rättning och ny rättning utifrån typ av frågekonstruktion, elevernas kön samt typ av huvudman och skola m.m. Avvikelser ska även analyseras för olika betygsnivåer utifrån ursprungligt provbetyg [...]*¹².

Sekretesslagens bestämmelser om tillgång till statistikuppgifter på individnivå tillåter inte att Skolinspektionen får tillgång till de personuppgifter som behövs för att genomföra analyser av systematiska skillnader i bedömningen av olika grupper. Därför har Skolinspektionen uppdragit åt forskaren Björn Tyrefors Hinnerich vid Stockholms universitet att utföra dessa.

Om analysen

Björn Tyrefors Hinnerich har tillsammans med Jonas Vlachos genomfört en analys av vad som kan förklara avvikelserna mellan den ursprungliga rättningen och omrättning av nationella prov utifrån tre huvudområden: elevernas karaktäristika, ursprungs rättarnas egenskaper och om det fanns någon skillnad mellan olika typer av skolor. Syftet har varit att analysera om någon grupp elever i jämförelse med någon annan grupp diskriminerades i den meningen att de systematiskt bedömdes annorlunda av ursprungs rättaren än av kontrollrättaren. Syftet har också varit att se om någon grupp lärare bland de som gjorde den ursprungliga rättningen systematiskt gör annorlunda bedömningar än omrättarna. Rapporten finns i sin helhet som bilaga 4 i resultatbilagan.

¹² Regleringsbrev avseende Statens skolinspektion för verksamhetsåret 2011.

6. Rekommendationer

Skolinspektionen lämnar följande huvudrekommendation för att de nationella proven i större utsträckning ska kunna stödja en rättvis och likvärdig betygssättning.

Skolinspektionen rekommenderar att de delprov som är i form av uppsatser eller längre självständigt författade redogörelser lyfts bort ur de nationella proven och genomförs i andra sammanhang.

Skolinspektionen har nu omrättat cirka 94 000 nationella prov på delprovsnivå under tre år. Resultatet av denna omrättning är entydigt och samstämmigt mellan åren. Vid nästan varannan omrättad delprovsuppsats i Svenska i årskurs 9 och på gymnasiet finns skillnader mellan ursprungs rättarens bedömning och Skolinspektionens bedömning. Minst två personer har alltså gjort olika bedömningar av ett och samma delprov i väldigt många fall.

Skolinspektionens bedömning är därför att de delprov som är utformade som längre självständigt författade redogörelser men framförallt uppsatsskrivande under rådande förhållanden inte fungerar för att "stödja en likvärdig och rättvis bedömning och betygssättning" som är ett av de två huvudsyftena för de nationella proven. Uppsatser och längre självständigt författade redogörelser är ett alltför trubbigt verktyg när syftet är att kontrollera likvärdigheten i betygssättning och bedömning. När så många som vartannat omrättat delprov bedöms olika av olika rättare finns det risk att tilltron till de nationella proven som ett instrument för likvärdighet på sikt kommer att urholkas.

Skolinspektionen är dock medveten om att de förmågor och kunskaper som testas genom delprov uppsatsskrivande 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. Detta behöver dock inte göras genom de nationella proven eftersom syftet med likvärdig bedömning inte ser ut att kunna uppnås. Andra modeller för att ge lärare stöd att göra dessa bedömningar behöver således utvecklas.

Om uppsatserna ska finnas kvar som delar av de nationella proven finns åtgärder som behöver vidtas för att kunna förbättra samstämmigheten i bedömning-

en av uppsatser så att dessa delprov också skulle kunna stödja en likvärdig och rättvis bedömning och betygssättning. Det finns en rad olika åtgärder, till exempel ökad satsning på utbildning och förbättrade anvisningar för rättning av proven som kan bidra till att höja likvärdigheten i bedömningen av de delprov som består av uppsatser. Dessa åtgärder beskrivs närmare nedan.

Rådande forskning (exempelvis Wenche Vagle " Ungdommers skrivekompetanse", 2005) visar dock att även under mer ideala förhållanden då rättarna intensivt har tränats för att ha en god förmåga att bedöma uppsatser är det svårt att få en så hög samstämmighet så att denna typ av delprov kan stödja en likvärdig och rättvis bedömning och betygssättning.

Om huvudsyftet med de nationella proven även fortsättningsvis primärt ska vara att stödja en likvärdig och rättvis bedömning och betygssättning vill Skolinspektionen rekommendera att de delprov som är i form av uppsatser eller längre självständigt författade redogörelser lyfts bort ur de nationella proven och genomförs i andra sammanhang. Skolinspektionen vill understryka att de förmågor och kunskaper som testas i uppsatsen är centrala och även fortsättningsvis bör testas men dessa tester skulle då genomföras i andra sammanhang än genom de nationella proven.

Avvikelserna mellan ursprungs rättarens bedömning och Skolinspektionens bedömning är avsevärt mindre och ibland försumbara för de delprov där det finns ett facit. Skolinspektionens bedömning är därför att dessa delprov är mer lämpade för att stödja en likvärdig och rättvis bedömning och betygssättning av de kunskaper och kompetenser som testas i dessa delprov. De avvikelser som trots allt finns vid bedömning av dessa delprov skulle kunna minimeras ytterligare genom ett elektroniskt rättningsförfarande införs.

Trots starka argument för att lyfta bort delar som innehåller uppsatssvar i de nationella proven kan det finnas andra skäl till att de ska finnas kvar. Skolinspektionen rekommenderar i sådant fall att följande åtgärder vidtas för att öka likvärdigheten i bedömningen. Rekommendationerna bygger på resultat och erfarenheter från tre års omrättningar av nationella prov, tidigare forskning och årets mer djupgående analys. Skolinspektionen vill understryka vikten av att det är kombinationen av rekommendationer som ökar likvärdigheten, och att till exempel endast genomföra avidentifiering av elevlösningarna ger i sig endast en marginell ökning av likvärdigheten.

Om uppsatsdelen ska vara kvar som en del av de nationella proven vill Skolinspektionen rekommendera följande åtgärder för att öka samstämmigheten vid bedömningen:

- Skolverkets *bedömningsanvisningar behöver bli tydligare och mer tillgängliga* framförallt i de delar som gäller de högre betygsstegen.
- Lärare som ska bedöma elevers prestation på de nationella proven bör få en god *grund- och fortbildning* i att bedöma.
- Skolinspektionen rekommenderar skolorna att i så stor utsträckning som möjligt använda sig av *sambedömning*, där mer än en lärare i samråd rättar samma prov. Denna bedömningsform lämpar sig särskilt väl för essäfrågor och uppsatsskrivande.
- Skolinspektionen anser att ett rättningsförfarande där *elevens identitet inte är känd* för läraren som rättar proven bör införas. Det hindrar inte att läraren vid senare tillfälle kan ta del av elevens lösning i full vetenskap om vems den är.
- Skolinspektionen ser stora vinster med att proven på sikt skulle kunna genomföras *elektroniskt*. Det skulle underlätta möjligheten att identifiera dem, underlätta kontroll- och omrättning samt möjligheten att använda sig av sambedömning inom skolan och mellan skolor. Det skulle därtill förenkla arkiveringen av dem.

Bedömningsanvisningarnas utformande är en avgörande faktor för att uppnå en ökad samstämmighet och en likvärdig bedömning av elevernas svar. Det är därför av yttersta vikt att anvisningarna är tydliga och använder sig av beskrivningar som lämnar ett minimalt utrymme för enskilda tolkningar. Skolinspektionens erfarenheter är att avvikelserna är som störst för de högre betygsstegen vilket tyder på att anvisningarna förförallt bör förtydligas i dessa delar. Exempel på elevlösningar med tillhörande bedömning är mycket uppskattade

av många lärare.¹³ Men med fler exempel och mer omfattande anvisningar följer ökade krav på en tydlig och lättnavigerad struktur. Om bedömningsanvisningarna skulle finnas även i elektronisk form, skulle dels kunna vara mer omfattande vad gäller exempel med mera, utan att behöva bli otympliga.

Skolinspektionen är väl medveten om den komplexitet och de svårigheter som finns kring utformning och bedömning av nationella prov. Föreslagna rekommendationer ska därför läsas som en del i en helhet, snarare än att de var för sig skulle utgöra någon enkel lösning på de problem som Skolinspektionen pekar på. Vi är även medvetna om att vissa av de föreslagna förändringarna kan vara svåra att genomföra på kortare sikt, som att proven skulle genomföras elektroniskt. En sådan åtgärd skulle dock i många avseenden underlätta och förbättra rättning och hantering av proven. Samtidigt är Skolinspektionen väl medveten om de höga kostnader och administrativa förändringar som följer på såväl utveckling som genomförande. Vi är som sagt övertygade om att ett elektroniskt baserat prov skulle underlätta avidentifiering, delning av prov mellan lärare för sambedömning och kontrollrättning. Vi tror även att det skulle främja lika förutsättningar ur elevens perspektiv (genom att minska svårigheterna för elever med särskilda behov, exempelvis dyslektiker).

Urval inför kommande rättningsomgångar

Skolinspektionen ska enligt uppdragsbeskrivningen även föreslå lämpliga grunder för urval i kommande rättningsomgångar.

I uppdragsbeskrivningen till Skolinspektionen framgår att den centrala rättningen ska genomföras på ett urval av samtliga elevlösningar på de nationella prov som är obligatoriska för samtliga elever i grund- och gymnasieskolan. Här framgår också att urvalet ska göras på ett sådant sätt att erhållna resultat är representativa dels på **nationell nivå** per ämne och årskurs, dels **på skolnivå** vid de utvalda skolorna i ämnena svenska/svenska som andraspråk, matematik och engelska i årskurs nio samt i gymnasieskolan.

Uppdraget att genomföra en central rättning innehåller flera olika delar, som på olika sätt syftar till att stödja en likvärdig bedömning och betygssättning av proven över landet. Ett urval på nationell nivå riktar sig till de delar i uppdraget som ska ge ett statistiskt underlag över hur de nationella proven bedöms och betygsätts inom olika ämnen och årskurser. Urval på nationell nivå är även

¹³ Detta har bland annat framkommit i samtal med de omrättare som Skolinspektionen har anlitat.

en förutsättning för att kunna göra analyser med avseende på systematiska felbedömningar som också är en del av uppdraget. Ett urval på skolnivå riktar sig i stället till de delar i uppdraget som fokuserar på att erhålla ett underlag för fördjupad granskning inom ramen för Skolinspektionens tillsyn och kvalitetsgranskning. Med befintliga budgetramar är det inte möjligt att årligen fullt ut tillgodose de båda delarna i uppdraget på ett optimalt sätt. Därför rekommenderar Skolinspektionen att uppdraget fokuseras på en av de två delarna. Alternativt kan de årliga analyserna fokusera på några få begränsade aspekter enligt ett rullande schema.

Statistisk bild och analyser över systematiska felbedömningar

Den centrala rättningen innebär en kvalitetssäkring av provsystemet genom att den ger möjlighet att upptäcka felbedömningar som äventyrar systemets tillförlitlighet. Enligt uppdraget ska en central rättning ge ett statistiskt underlag för hur de nationella proven bedöms och betygsätts inom olika ämnen och årskurser. Analyser ska göras med avseende på bland annat systematiska felbedömningar. En statistisk bild ger en generell lägesbild och genererar kunskap som kan användas för att förebygga systematiska felbedömningar.

I dag är urvalet utformat för att: a) generera en bild av hur stort problemet med felbedömningar är i Sverige; b) möjliggöra vissa övergripande analyser om dessa felbedömningar och c) skapa ett underlag på skolnivå. Målen är satta inom ramen för befintliga budgetramar. Urvalet har däremot inte kunnat utformas för att möjliggöra årliga analyser om systematiska felbedömningar (exempelvis utifrån egenskaper eller bakgrund hos lärare och elever) på ett sådant vis att det går att dra generella slutsatser på nationell nivå. För att kunna göra sådana analyser måste urvalet anpassas och utökas utifrån de aspekter som ska analyseras.¹⁴

Skolinspektionen har nu genomfört omrättning av nationella prov under tre år. Omfattningen av felbedömningar är i lika stor från år till år. Man kan därför ifrågasätta om det finns ett behov av att årligen genomföra de övergripande analyser som nuvarande utformning av uppdraget medger.

¹⁴ Om exempelvis kön, ålder, utbildning och erfarenhet hos läraren ökar risken för felbedömningar ska analyseras, måste urvalet göras bland samtliga lärare (och inte som i dag från prov från en skola). För att kunna dra tillförlitliga slutsatser måste vi även försäkra oss om att det finns ett tillräckligt stort urval av lärare i de olika undergrupperna.

Ett förslag är att överge den delen av uppdraget till förmån för att koncentrera urvalet på skolnivå och då omrätta prov från ämnen där tidigare rättningssomgångar visat på stor avvikelse vart tredje år, enligt ett rullande schema för olika ämnen, och vart sjätte år för ämnen med små avvikelser. Det skulle ge utrymme för att inom befintliga budgetramar öka möjligheten att analysera systematiska felbedömningar.

Utifrån ett sådant upplägg skulle vi under skilda omrättningsomgångar kunna välja att fokusera på analyser utifrån alternerade elever, lärare och skoltyper. De analyser som hittills gjorts av systematiska felbedömningar indikerar att det kan finnas sådana där elever med visst ursprung systematiskt felbedöms. Därför anser Skolinspektionen att fokus bör läggas på detta, trots att det skulle begränsa möjligheten att jämföra nivån på felbedömningar mellan olika ämnen.

Ytterligare ett sätt att analysera systematiska felbedömningar är att lägga ihop uppgifter om omrättning för flera år. Men det skulle kräva att Skolinspektionen får tillgång till material vi på grund av sekretessbestämmelser för närvarande inte har tillgång till.¹⁵

Underlag för fortsatt tillsyn och kvalitetsgranskning

Enligt uppdraget ska den centrala rättningen av nationella prov utgöra ett led i inspektionens löpande verksamhet. Resultat av den bör i förekommande fall vara underlag för fördjupad granskning inom ramen för Skolinspektionens tillsyn och kvalitetsgranskning. Rent konkret arbetar Skolinspektionen med detta på två sätt: 1) I den regelbundna tillsynen av skolorna 2) I en riktad tillsyn för de skolor som utmärker sig i omrättningarna genom att ha väldigt höga avvikelser.

I dag är urvalet primärt utformat för att kunna göra analyser av en skola i ett ämne. Urvalet av skolor är slumpmässigt för att vi ska få ett statistiskt underlag på riksnivå. Urvalet är därmed inte kopplat till den regelbundna tillsynen. Det innebär att när omrättningen visar på stora avvikelser krävs vanligtvis en särskild granskning av skolan. Arbetet skulle kunna effektiviseras om omrättningen av nationella prov kopplades ihop med de skolor som tillsynas i den regelbundna tillsynen ett givet år. En central rättning görs då som kontroll inför att

¹⁵ I och med att Skolinspektionen inte är en statistikansvarig myndighet eller har forskning inskrivet i instruktionerna (Förordning (2011:556) med instruktion för Statens skolinspektion) har Skolinspektionen inte tillgång till personuppgifter enligt Offentlighets- och sekretesslagen (2009:400) kap. 24 Sekretess till skydd för enskild inom forskning och statistik.

en skola tillsynas. Vid stora avvikelser granskas hur skolan arbetar med bedömning och betygssättning i sin helhet.

Analyser av omrättade prov i matematik och de naturvetenskapliga ämnena

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 nio. För de proven är det svårt att bedöma likvärdigheten i betygssättningen på omrättade prov, då de i proven ingår delprov som består av muntliga framställningar eller laborationsmoment som inte är möjliga att omrätta. Skolinspektionen ser två möjliga sätt att handskas med problemet:

- Att samla in resultaten på uppgiftsnivå.
- Att utfärda delprovsbetyg för prov innehållande delar som inte är praktiskt möjliga att omrätta.¹⁶

I dag samlar Skolinspektionen in resultaten för dessa skolor. 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, eftersom Skolinspektionen på grund av sekretessreglerna inte kan göra insamlingen förrän tidigast ett halvår efter det att proven genomförts. Skolinspektionen förespråkar därför att delprovsbetyg utfärdas. Men det skulle kräva att Skolverket delvis ändrar provens konstruktion.

Slutsatser

Urvalet bör formars med ett långsiktigt perspektiv i åtanke. Beroende på vilket syfte för de nationella proven som får företräde finns olika strategier:

Om det primära målet är att analysera systematiska skillnader måste urvalet vara representativt på nationell nivå utifrån de områden man vill analysera.

Väljer man i stället att sätta fokus på likvärdighet och tillsyn bör man samla in ett urval av elevlösningar för omrättning som återspeglar de skolor vi avser granska i den regelbundna tillsynen.

¹⁶ Det är exempelvis inte praktiskt möjligt att omrätta laborationer och muntliga framställningar.

Önskar man behålla utformningen av uppdraget så som det ser ut i dag, det vill säga en kombination av de ovanstående olika vägarna, kräver det en omfattande utökning av såväl urval som budgetanslag.

Skolinspektionen bedömer att för att göra en analys utifrån ett urval som baserar sig på eleverna behöver Skolinspektionen dubbelt så stor budget som dagens. Anledningen är att Skolinspektionen i sådana fall kommer att behöva samla in och omrätta långt fler elevlösningar. Det behövs för att man ska få ett urval som också är representativt för elevgrupper utifrån aspekter som exempelvis kön, föräldrars utbildningsnivå eller etnisk bakgrund. Vill man dessutom ha ett urval för en analys på skolnivå (som Skolinspektionen behöver i tillsynen av skolorna) krävs den tredubbla budgeten. Ett rullande schema där man väljer olika urval och olika fokus varje år skulle vara ett sätt att begränsa kostnaderna.

7. Genomförande av uppdraget

Detta kapitel beskriver Skolinspektionens genomförande av uppdraget. Vi har delat in beskrivningen i de tre delarna: *urval*, *bortfall* samt *brister och problems*.

7.1 Urval

Ett urval av elevlösningar gjordes bland de obligatoriska proven i grundskolan och de prov som är obligatoriska för alla elever i gymnasieskolan. Proven omfattar ämnena engelska, matematik, svenska/svenska som andraspråk och de naturorienterande ämnena biologi, fysik och kemi i årskurs nio. I gymnasieskolan omfattar urvalet kursproven i engelska A, matematik A och svenska B/svenska som andraspråk B. Att avstå från omrättningen av de nationella proven i årskurserna 3 och 5 motiveras med att avvikelserna mellan ursprungsretarens bedömning och Skolinspektionens bedömning generellt är mycket små och att nationella prov i åk 5 inte är obligatoriska för skolorna under vårterminen 2011.

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 i ämnena engelska, matematik och svenska/svenska som andraspråk i årskurs nio i grundskolan samt kurserna i gymnasieskolan.

Urvalet drogs från de insamlingar av nationella prov som SCB genomförde under sommaren och hösten 2011. Totalt sett omfattar urvalet 31 394 prov, fördelat på 22 355 prov från årskurs nio och 9 039 prov från gymnasieskolan.

Utöver detta urval drogs ytterligare 3 103 prov bland skolor som i samband med omrättning av proven i omgång 1 påvisade stora avvikelser i ämnena engelska och svenska/svenska som andraspråk i årskurs nio.

Tillvägagångssätt

Tillvägagångssättet vid omrättningen innehåller följande moment: Insamling, rättning, analys, överväganden och avgränsningar. De redovisas var för sig nedan.

Insamling av elevlösningar

Det genomförda urvalet innebar att 492 skolor kontaktades och ombads skicka in kopior av rättade elevlösningar till Skolinspektionen. Av dem var 37 med i urvalet på grund av stora avvikelser i samband med den första omrättningsomgången av nationella prov. Omfattningen av arbetsinsatsen varierade för de utvalda skolorna. Målet var att ingen skola skulle skicka fler än 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 poängnivå för prov i matematik, fysik, biologi och kemi för årskurs nio. Anledningen var att proven i de ämnena inte har några delprovsbetyg utan endast provbetyg. I och med att de proven även innehåller en muntlig och en laborativ del skulle en jämförelse mellan den ursprungliga bedömningen och Skolinspektionens bedömning vara missvisande.

Till varje insamlat prov utformades ett försättsblad. Samtliga försättsblad var i förväg ifyllda med uppgifter om skolans identitet (skolkod, skol-id och elev-id) och skolans namn för att öka kvaliteten på de uppgifter som skulle samlas in. För proven i matematik, fysik, biologi och kemi innehöll försättsbladet även strukturerade fält för insamlingen av provresultaten. Detta för att säkra att resultaten samlas in på ett enhetligt sätt.

Skolorna fick detaljerade anvisningar för hur inrapporteringen och kopieringen skulle göras.

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

Rättning av proven

På Skolinspektionens uppdrag kontrakterade företaget Lärarförmedlarna 150 - 200 ämnesutbildade och erfarna lärare på visstid för att rätta om proven. Lärarna utförde rättningen 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. 3–4 §§, för att undervisa i aktuellt ämne och årskurs. Lärarna skulle också ha erfarenhet av att rätta nationella prov i aktuellt ämne och aktuell årskurs.

Av de lärare som anlätades för innevarande års rättning deltog 64 procent även i förra årets rättningssomgång. Totalt är ungefär en femtedel av landets kommuner representerade bland de anlidade lärarna.

Skolinspektionen anordnade också ett heldagsseminarium om bedömning av elevtexter i svenska. Seminariet hölls av Eva Östlund-Stjärnegårdh och Anne Palmér från Institutionen för nordiska språk vid Uppsala universitet. Institutionen utarbetar på uppdrag av Skolverket de nationella proven i svenska och svenska som andraspråk. Målet för seminariet var att ge de anlidade lärarna en praktisk övning i att bedöma elevtexter och genom det stärka deras bedömar-kompetens.

För att öka kvaliteten på omrättningen har vissa prov rättats en tredje gång av ytterligare en lärare i de fall det förekom stora eller avgörande skillnader mellan omrättande lärares bedömning och den ursprungliga bedömningen. Det benämns som en andra omrättning. Med stora eller avgörande skillnader avses här prov där antingen ursprungs rättarens bedömning eller omrättarens bedömning inneburit ett IG. Man kan också ha bedömt att eleven inte nått målen, eller så har antingen omrättaren eller ursprungs rättaren gett betyget G och den andra betyget MVG. För årskurs tre och fem har proven omrättats en andra gång om fler än ett delprov under ett och samma prov avvikit mellan omrätta-

rens bedömning och ursprungs rättarens bedömning. Eftersom delproven i matematik och de naturorienterade ämnena i årskurs nio inte har varit möjliga för omrättaren att betygssätta har det inte gjorts någon andra omrättning de delproven.

Resultatet av den andra omrättningen har tolkats utifrån principen att den andra omrättningen avgör om ursprungsrättningen eller omrättningen ska gälla. Om den andra omrättningen har resulterat i ett tredje värde har omrättningen inte påverkats. 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 och därmed omedvetet påverkas i sin bedömning.

Resultat och analys

Resultaten presenteras som en redovisning av avvikelser mellan ursprungsrättarens bedömning och Skolinspektionens bedömning utifrån betyg eller poäng på delprovsnivå för årskurs nio och gymnasieskolan. Dessutom redovisas betygsfördelningen eller poängfördelningen mellan ursprungsrättarens bedömning och omrättarens bedömning.

Resultaten redovisas som negativa och positiva avvikelser. Negativ avvikelse innebär att omrättaren bedömer att elevens svar motsvarar ett lägre betyg/poäng än vad ursprungsrättaren har bedömt. Positiv avvikelse innebär att omrättaren bedömer att elevens svar motsvarar ett högre betyg/poäng än vad ursprungsrättaren har bedömt.

Överväganden

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

Regeringen preciserar i uppdraget till Skolinspektionen att analysen ska baseras på förekomsten av systematiska avvikelser utifrån typ av frågekonstruktion, typ av huvudman och skola samt hur avvikelser fördelar sig mellan olika betygsnivåer.

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). Det är den skala SCB använder när de samlar in provresultat på uppdrag av Skolverket. Eftersom Skolinspektionen beställer provresultaten från SCB är det endast den fyrgradiga skalan som finns att tillgå.

Skolinspektionen har möjlighet att direkt från skolorna samla in provresultaten enligt den utökade betygsskalan, som tillämpas på skolorna i samband med rättning av nationella prov. Men det skulle innebära en mycket utökad arbetsbörda för skolorna. Samtidigt är det inte givet att skolorna skulle kunna uppfylla myndighetens krav, eftersom provresultaten inte alltid förvaras på skolorna på ett systematiskt sätt. Det skulle alltså ha stor påverkan på själva kvaliteten på insamlade uppgifter och i förlängningen ge ett ökat bortfall.

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 ges inte några delprovsbetyg. De två 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 det 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. Skolinspektionens anlitade lärare har känt till vilka elevlösningar som avser svenska som andraspråk, där sådan information funnits. Cirka fyra procent av de insamlade proven i årskurs nio och cirka en och en halv procent av de insamlade proven i gymnasieskolan avsåg svenska som andraspråk.

7.2 Bortfall

Av samtliga 34 497 nationella prov som ingick i urvalet skickade skolorna in 29 707 prov, fördelade på 21 439 prov från årskurs nio och 8 268 från gymnasieskolan. 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 med de insamlingar som görs av provkonstruktörer.

Många 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 2011. Det utgör det interna bortfallet. I motsats till det externa bortfallet (där provet i sin helhet uteblivit), såg det interna bortfallet olika ut för olika delprov.

Tabell 1: Totalt bortfall fördelat på årskurs och ämnesprov respektive kursprov

Årskurs	Ämne	Antal begärda prov	Bortfall	Antal begärda prov p.g.a. stora avvikelser i den första omgången av omrättningen av nationella prov	Bortfall
Årskurs 9	Engelska	6 040	10%	942	17%
	Matematik	5 121	30–31%	-	-
	Svenska	5 707	11–8%	2 161	16%
	Biologi	2 516	23%	-	-
	Fysik	1 855	38–41%	-	-
	Kemi	1 116	24–46%	-	-
Gymnasiet	Engelska	3 278	9–10%	-	-
	Matematik	3 044	22%	-	-
	Svenska	2 717	14%	-	-

Förklaring: I de fall där bortfallet anges som ett intervall innebär det att bortfallet varit olika för olika delprov som ingår i ett specifikt prov.

Av den beskrivning som framkommer i tabell 1 kan man konstatera att det finns ett relativt stort bortfall när det gäller vissa prov.¹⁷

Utifrån endast bortfallsandelen är det svårt att avgöra om det går att dra några slutsatser 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 är tillförlitliga. Men man måste iakta stor försiktighet när man analyserar omrättningens resultat.

¹⁷ Man bör ha i åtanke att fristående skolor inte tidigare omfattas av arkiveringskravet av de nationella proven. Därför kan de redan ha gjort sig av med elevlösningar som vi begär in kopior av. Ytterligare en anledning till visst bortfall är skolor som inte har angett vilket provbetyg de satt på elevlösningar de har kopierat och skickat in. Då kan vi omrätta elevlösningen men har inget provbetyg att jämföra med.

7.3 Brister och problem

För att minska en del av problemen vid insamlingen, har Skolinspektionen inom den ram som uppdraget medger, vidtagit flera åtgärder.

Skolinspektionen meddelade i god tid innan insamlingen skulle äga rum att skolan var föremål för omrättning och att elevlösningar skulle begäras in senare under hösten. Skolorna fick information såväl om vilket ämne som var aktuellt som om det uppskattade antalet prov som skulle bli föremål för insamling. Skolorna uppmanades att ha proven tillgängliga och att avsätta resurser för kopiering. Svarstiden förlängdes samtidigt med två veckor.

I samarbete med Skolinspektionen har Skolverket i den lärarinformation som följer med proven uppmanat lärarna att göra rättningsnoteringar i marginalen, och inte i själva provet.

Tidigare problem med provinsamlingen har minskat kraftigt som en följd av de vidtagna åtgärderna, men kvarstår ändå i flera delar. Det gäller framför allt att skolorna upplever en mycket stor arbetsbelastning för att inkomma med de efterfrågade elevlösningarna inom utsatt tid. De har också problem med att rapportera resultaten för prov i biologi, fysik, kemi och matematik i årskurs nio, det vill säga de prov där Skolinspektionen samlar in resultaten på egen hand.

8. Källor

- Borgström. (2010). *Att skriva prov – om normer och textkultur i gymnasieskolans skriftliga nationella prov* i Språk och Stil, NF 20 2010
- Borgström och Yassin. (2010). *Tioåringar skriver nationellt prov*. I *Att bygga broar – kulturella, språkliga och mediala möten/ Sjunde nationella konferensen i svenska med didaktisk inriktning*.
- Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics*. Göteborg studies in educational science, Vol. 269. Göteborg: Acta Universitatis Gothoburgensis.
- Skolverket. (2003). *Det nationella provsystemet— vad, varför och varhän? Dnr 01-2003:2038*. Stockholm: Skolverket.
- Skolverket. (2004). *Det nationella provsystemet i den målstyrda skolan*. Stockholm: Skolverket.
- Skolverket. (2005). *Nationella prov i gymnasieskolan— ett stöd för likvärdig betygsättning?* Stockholm: Skolverket.
- Skolverket. (2007). *Provbetyg – Slutbetyg – Likvärdig bedömning? Rapport 300*. Stockholm: Skolverket.
- Skolverket. (2009). *Bedömaröverensstämmelse vid bedömning av nationella prov. Dnr 2008:286*. Stockholm: Skolverket.
- Skolverket. (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid*. Stockholm: Skolverket.
- Wenche Vagle (red.)(2005). *Ungdommers skrivkompetanse*. Oslo: Universitetsforlaget
- Förordning (2011:556) med instruktion för Statens skolinspektion.
- Offentlighets- och sekretesslagen (2009:400) kap. 24 Sekretess till skydd för enskild inom forskning och statistik.

9. Diagramförteckning

DIAGRAM 1: FÖRDELNING AV ANDEL AVVIKELSER; POSITIVA, NEGATIVA SAMT INGEN AVVIKELSE MELLAN URSPRUNGSRÄTTARENS BEDÖMNING OCH SKOLINSPEKTIONENS BEDÖMNING FÖR DE OLIKA DELPROVEN, OMGÅNG TRE.	13
DIAGRAM 2: GYMNASIET, SVENSKA SKRIFTLIG DEL, OLIKA GENRE MAN KAN VÄLJA MELLAN ATT SKRIVA, UPPDELAT PÅ SKRIVUPPGIFT (TRE SKRIVUPPGIFTER VAR I GENREN INLÄGG). ANDEL POSITIVA OCH NEGATIVA AVVIKELSER SAMT AVVIKELSER TOTALT PER SKRIVUPPGIFT, OMGÅNG 3.	17
DIAGRAM 3: AVVIKELSER MELLAN URSPRUNGSRÄTTARENS BEDÖMNING OCH SKOLINSPEKTIONENS BEDÖMNING FÖR GYMNASIET GÄLLANDE SVENSKA, SKRIFTLIG DEL FÖR RESPEKTIVE SKOLA I URVALET, OMGÅNG TRE.	19
DIAGRAM 4: TOTAL ANDEL AVVIKELSE FÖR DE NATIONELLA PROVEN I SVENSKA OCH ENGELSKA FÖR ÅRSKURS NIO OCH FÖR DE NATIONELLA PROVEN I SVENSKA, ENGELSKA OCH MATEMATIK PÅ GYMNASIET.	20
DIAGRAM 5: ÅRSKURS 9, ENGELSKA SKRIFTLIG DEL. ANDEL SAMSTÄMMIGHET MELLAN URSPRUNGSRÄTTARE OCH OMRÄTTARE ÖVER TID FÖRDELAT PÅ BETYG.	21
DIAGRAM 6: ÅRSKURS 9, ENGELSKA SKRIFTLIG DEL. ANDEL POSITIVA OCH NEGATIVA AVVIKELSER SAMT AVVIKELSER TOTALT.	21