

---

# Nationellt prov i grundskolans årskurs 9

**Provrapport Samhällskunskap**  
Årskurs 9  
Vårterminen 2019



**GÖTEBORGS UNIVERSITET**  
**INST FÖR PEDAGOGIK OCH SPECIALPEDAGOGIK**

## Innehållsförteckning:

|                                                               |    |
|---------------------------------------------------------------|----|
| 1. Inledning                                                  | 3  |
| 2. Konstruktionsprocess för provet                            | 3  |
| 2.1 Bakgrund                                                  | 3  |
| 2.2 Provets sammansättning                                    | 4  |
| 2.3 Bedömningsanvisningar                                     | 4  |
| 3. Provresultat med kommentarer                               | 4  |
| 3.1 Resultat fördelat på faktorer kön och undervisning i SvA  | 6  |
| 3.1.1 Skillnader mellan pojkar och flickor                    | 6  |
| 3.1.2. Skillnader mellan de som läser Svenska eller Svenska A | 7  |
| 4. Validitet och reliabilitet                                 | 8  |
| 4.1 Validitet                                                 | 8  |
| 4.2 Reliabilitet                                              | 10 |
| 5. Enkätresultat med kommentarer                              | 12 |
| 6. Avslutning                                                 | 15 |
| Referenser                                                    | 17 |
| Bilaga 1: Utprovningar 2019                                   |    |

## 1. Inledning

Skolverket har i uppdrag av regeringen att ansvara för och genomföra nationella prov i SO-ämnena (dvs. geografi, historia, religionskunskap och samhällskunskap) i årskurs 9. Skolverket har i sin tur gett i uppdrag åt Göteborgs universitet att konstruera de nationella proven i samhällskunskap för årskurs 9. Denna rapport behandlar nationella provet i samhällskunskap i årskurs 9. 2019 genomfördes provet den 7 och 9 maj.

Lärare har digitalt inrapporterat provbetyg och uppgiftsresultat för ett urval av elever. Lärare har också besvarat en digital enkät. Ett urval av hela elevprov har skickats in till provinstitutionen vid Institutionen för Pedagogik och Specialpedagogik, Göteborgs universitet.

Denna rapport bygger på de inrapporterade provresultaten, de hela elevproven, lärarenkäterna och resultaten från SCB. Skillnaden mellan SCBs statistik och vårt urval redovisas också. I år har nästan 200 lärare besvarat enkäten, en förbättring jämfört med förra året.

I materialet nedan nämns flera statistiska metoder och begrepp. Man behöver inte ha kunskaper om dessa för att kunna följa resonemanget. Avsikten med att visa dessa är att de intresserade ska kunna läsa vidare om våra metoder och statistiska överväganden.

## 2. Konstruktionsprocess för provet

### 2.1 Bakgrund

Ämnet samhällskunskap förekommer bara i de nordiska länderna. I övriga världen finns ofta ett ämne som påminner om statsvetenskapsdelen inom samhällskunskap, men inget med vårt ämnes alla olika delar: statsvetenskap, nationalekonomi, sociologi, kulturgeografi och idag också internationella studier, juridik, kriminologi osv. Det finns inga andra obligatoriska nationella prov i ämnet i Norden.

Vid provkonstruktion talar man ofta om god mätsäkerhet. Det kan låta tekniskt men det är samma sak som att eleven ska ges så goda möjligheter som möjligt att visa sina kunskaper och att provet ska spegla elevens faktiska kunskaper. God mätsäkerhet kräver att elevens kunskaper prövas med fler frågor inom varje del av kunskapskraven<sup>1</sup>. Om detta inte sker riskerar man att elevens resultat styrs för mycket av slump. Detta innebär också att vår verksamhet att utprova uppgifter blir stor och kräver stort fokus på verksamhetens kvalitet.

Förutom mindre utprovningar av enstaka uppgifter har fyra större utprovningar genomförts inför proven 2019, under framtagandet av provet har tusentals elever och hundratals lärare medverkat. Se bilaga 1.

Under arbetet med att sälla fram bra uppgifter ur de uppgifter som är med i utprovningar arbetar vi mycket medvetet med olika varianter av IRT, Item Response Theory, s.k. modern mätteori. Med hjälp av olika program analyseras varje utprovning i detalj, hur de olika uppgifterna fungerat i olika elevgrupper.

---

<sup>1</sup> Kunskapskraven i samhällskunskap kan ses som sex olika innehållsmässiga delar. Vi benämner detta ”delar av kunskapskraven.

Förutom ovanstående bearbetningar genomförs också en mer ”mjuk” kvalitetssäkring genom olika utvärderingar på våra uppgifter och prov:

- Elever som deltar i våra utprövningar besvarar en enkät, där får vi information om bland annat upplevd svårighetsgrad av de uppgifter som just denne elev har mött. Det finns också möjlighet att i frisvarsformat skriva övriga synpunkter på de uppgifter som prövats på eleven.
- Lärare som tillsammans med oss prövar ut uppgifter i sina elevgrupper besvarar också en enkät, bland annat angående hur de bedömer att uppgifterna är i relation till kursplanen i samhällskunskap.
- Vi har haft fyra referensgruppsmöten med lärare från hela landet, en del av lärarna har varit helt nya i sammanhanget.
- I vår ”utvidgade” provgrupp har under 2019 fyra verksamma lärare ingått som förstärkning till den personal som arbetar i den fasta gruppen. De kommer till oss en dag/vecka, då är framtagande av nya uppgifter och skrivande av bedömningsanvisningar till utprövade uppgifter i fokus.

Liksom tidigare år har vi arbetat med så kallade stödstrukturer i provet. Vi har noterat att stödstrukturer kan ”stödja” åt fel håll varför vi försökt hitta en lagom nivå. I provet 2019 har vi en s.k. ”hjalpruta” till de öppna uppgifterna. Enligt lärarenkäten har detta fallit väl ut.

## 2.2 Provets sammansättning

Nationella provet i samhällskunskap i årskurs 9 bestod av två delprov. Det fanns ingen principiell skillnad mellan delproven, de innehåller båda uppgifter från de delar av kunskapskraven som prövas samt har en ganska likartad fördelning av öppna och slutna uppgifter. Det nationella provet i samhällskunskap för årskurs 9 har de senaste åren minskat antalet uppgifter, från 30 stycken 2013, till 24 stycken 2019. Vi har arbetat aktivt för att minska andelen helt öppna uppgifter, som ofta kräver ett längre svar. Problemet blir att bibehålla hög validitet och hög reliabilitet med färre antal uppgifter och fler av strukturerad natur. Mer om detta utvecklingsarbete nedan<sup>2</sup>.

| Np SH åk 9 | Uppg./item | Upp till nivå |    |    | Strukturerade |
|------------|------------|---------------|----|----|---------------|
|            |            | E             | C  | A  |               |
| 2013       | 30         | 30            | 29 | 16 | 19            |
| 2014       | 30         | 30            | 30 | 19 | 12            |
| 2015       | 28         | 28            | 24 | 18 | 9             |
| 2016       | 26         | 26            | 26 | 24 | 10            |
| 2017       | 24         | 24            | 24 | 22 | 11            |
| 2018       | 24         | 24            | 24 | 24 | 12            |
| 2019       | 24         | 24            | 24 | 24 | 14            |

Tabell 1. Fördelning av uppgifter, nivåer och format 2013-2019, Nationella provet i samhällskunskap åk 9.

## 2.3 Bedömningsanvisningar

Provbetyg ges inte direkt av elevens svar på uppgifterna, utan svaren måste först bedömas av en lärare med våra bedömningsanvisningar som stöd. Att skapa bra bedömningsanvisningar är minst lika svårt som

<sup>2</sup> I internationell litteratur så kallas en bedömning för ett ”item”. Ordet är numera försvenskat. Ett item kan bestå av flera uppgifter, t.ex. att uppgift 2 och 3 bedöms tillsammans, då blir dessa ett item. Om exempelvis uppgift 3a och 3b bedöms för sig så blir detta två item. I Np Sh 2019 motsvarar en uppgift ett item.

att skapa bra uppgifter. Vi har arbetat mycket med dessa och har använt samma visuella upplägg som tidigare år, dvs tvåspaltigt. Bedömningsanvisningarna valideras vid flera tillfällen av verksamma lärare under framtagningsprocessen och enligt lärarenkäten har detta arbete fallit väl ut.

### 3. Provresultat med kommentarer

Resultaten från provbetygen rapporterades av SCB den 28 november för 2019. De nationella proven i SO har under alla år haft en digital inrapportering där lärare har rapporterat in ett urval av elever (födda 6e, 16e och 26e i varje månad, n=1509). Detta urval har rapporterats in digitalt vilket gjort att vi tidigt har preliminära resultat att analysera, inrapporteringen stängs den 30 juni. Tillgången till detta preliminära resultat medger möjlighet till bearbetning av resultatet ett halvt år före SCB:s resultat publiceras. Den obligatoriska inrapporteringen av ett statistiskt urval är helt avgörande för analys och framtagande av proven, då den till skillnad från SCB:s rapport, också ger oss information på uppgiftsnivå. I texten nedan används beteckningen *urval* för dessa elever. Vi är varje år mycket tacksamma för att lärare gör denna inrapportering, lärare kan också få olika sorters feedback på sina gruppers resultat. Se SO-ämnenas sida för ytterligare information: [www.npsportal.se](http://www.npsportal.se)

Nedanstående diagram visar en jämförelse mellan provbetyget för vårt urval i juni och SCBs redovisning av provbetyg den 30 november 2019. Man ser att vårt statistiska urval står sig mycket bra med avseende på det slutgiltiga resultatet för provbetyget, därav syns också vikten av att dessa insamlingar görs av provinstitutionen så att vi så tidigt som möjligt kan få ett preliminärt resultat att bearbeta.

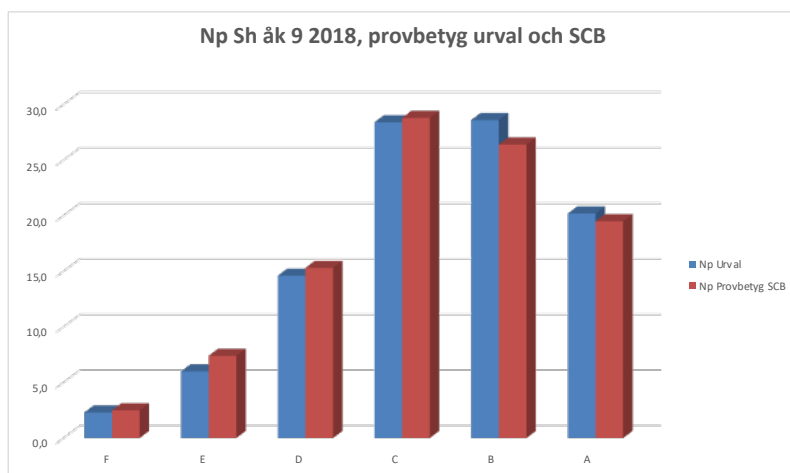


Diagram 1. Jämförelse av provbetyg vårt urval - SCB samhällskunskap i årskurs 9 2019

### 3.1 Resultat fördelat på faktorerna kön och undervisning i SvA

#### 3.1.1 Skillnader mellan pojkar och flickor

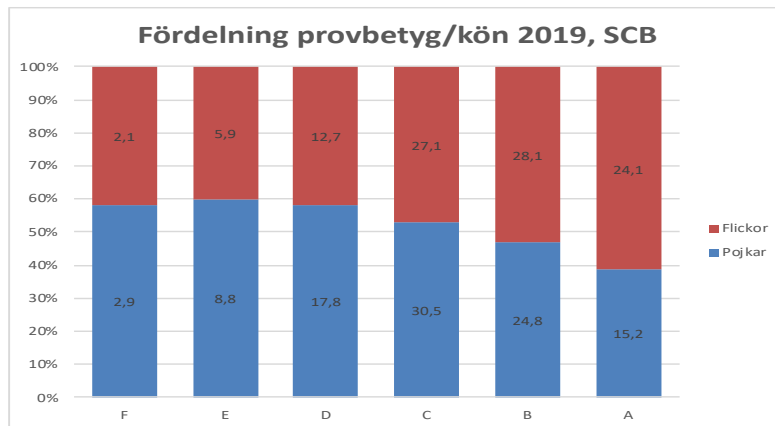


Diagram 2. Fördelning av provbetyg samhällskunskap med avseende på kön/SCB. Årskurs9 2018

Flickor presterar bättre på provet än pojkar. Detta gäller även slutbetyget i årskurs 9 där flickor i många år har haft bättre resultat än pojkar.

Vi har liksom i de tidigare publicerade provrapporterna analyserat om provet missgynnar eller gynnar något av könen. Vi genomför så kallade undergruppsanalyser för att se på hur de två undergrupperna vi har data på i urvalet, kön och om man följer kursplanen för Svenska eller svenska som andraspråk (Sv A). Vi använder Mantel-Haentzels DIF-analys (differential item functioning) för detta med syfte att se om enskilda uppgifter missgynnar/gynnar.

En vanligt förekommande teori är att alla uppgifter ska vara så neutrala som möjligt avseende undergrupper. Detta kan vara svårt att uppnå inom prov av samhällskunskapsämnet eftersom uppgifterna ofta adresserar områden som kan vara olika engagerande för olika grupper. Att åstadkomma helt neutralt ”uppförande” av uppgifterna skulle ibland ge väl ”tunna” uppgifter som inte skulle engagera eleverna i samma utsträckning. Men motsatsen gäller också, uppgifterna får inte heller vara sådana att de berör någon grupp för mycket, det kan skapa orättvisa. Exempelvis kan ett av oss givet val av idrott i en uppgift som handlar om idrott vara olika engagerande för flickor respektive pojkar.

Antalet uppgifter som är ”neutrala” är fler än på nationella provet samhällskunskap 2019 men det är lika många på ”båda sidor” vilket gör att vi är tillfreds med utfallet. Om det finns ungefär lika många uppgifter som gynnar flickor resp. pojkar så kan man tänka att de ”tar ut” varandra. I år är färre uppgifter än förra året neutrala, våra analyser av dessa visar att det beror på innehållet. För att se på skillnader mellan grupper används Mantel-Haentzels DIF-analys och resultatet anger i stegen neutralt, gynnar svagt eller gynnar starkt.

Provet är sekretessbelagt vilket innebär att vi endast kan visa innehåll i det stora, vi bedömer det som att både innehåll och format tilltalar könen på olika sätt.

| NP SH 2019 |                       |              |               |
|------------|-----------------------|--------------|---------------|
| Uppgift    | DIF Kön               | Format       | Innehåll      |
| 1          | Gynnar flickor starkt | Öppen        | Jämställdhet  |
| 2          | Gynnar flickor svagt  | Öppen        | Jämställdhet  |
| 3          | Gynnar flickor svagt  | Öppen        | Källor        |
| 4          | Gynnar flickor svagt  | Öppen        | Samhällsfråga |
| 5          | Gynnar pojkar svagt   | Strukturerad | Styrelseskick |
| 6          | Gynnar pojkar svagt   | Strukturerad | Val           |
| 7          | Gynnar pojkar svagt   | Strukturerad | Källor        |
| 8          | Gynnar pojkar svagt   | Strukturerad | Politik       |
| 9          | Gynnar pojkar svagt   | Strukturerad | Demokrati     |

Tabell 2. DIF-analys avseende kön, Np Sh årskurs 9 2019

Uppgifterna är anonymiserade på så vis att de tilldelats nummer slumpmässigt, fördelningen mellan pojkar och flickor är relativt jämn, dvs det finns uppgifter som tilltalar båda grupperna. Pojkar gynnas något av slutna svarsformat och flickor av öppna, detta är känt sedan tidigare. Denna statistik visar på vikten av att ett prov innehåller både öppna och strukturerade svarsalternativ.

### 3.1.2. Skillnader mellan de som läser Svenska och de som läser Svenska A

Vi har även gjort analyser på uppgifterna avseende på om de gynnar eller inte gynnar de som läser SvA. I denna analys så jämförs de i de olika grupperna som har samma provresultat. De som läser SvA har i snitt sämre resultat på provet. Detta gäller för samtliga ämnen där nationella prov ges, det finns dock skillnader mellan de olika ämnena.

Vi får här ett statistiskt underlag som gör det möjligt att jämföra de med samma förmåga i ämnet, med den särskiljande variabeln att de läser svenska respektive svenska som andraspråk.

| "Uppgift" | DIF                   | Format       | Innehåll           |
|-----------|-----------------------|--------------|--------------------|
| 1         | Missgynnar SvA, svagt | Strukturerad | Val                |
| 2         | Missgynnar SvA, svagt | Strukturerad | Källor             |
| 3         | Missgynnar SvA, svagt | Strukturerad | Samhällsstrukturer |
| 4         | Gynnar SvA, svagt     | Öppen        | Demokrati          |
| 5         | Gynnar SvA, svagt     | Öppen        | Samhällsfråga      |
| 6         | Gynnar SvA, svagt     | Öppen        | Jämställdhet       |

Tabell 3. DIF-analys avseende att följa kursplanen i svenska eller svenska som andraspråk, Np Sh årskurs 9, 2019

Eftersom kunskapskraven i samhällskunskap ofta bygger på att eleven ska kunna resonera om eller redogöra för samband kring olika skeenden i samhället så är det viktigt att ett prov innehåller olika typer av uppgifter som prövar elevernas kunskaper i ämnet. Alla uppgifter kan inte vara öppna, det kräver för mycket skrivande av eleverna och är mer tidskrävande för lärarna vid bedömningen. Alla uppgifter kan inte heller vara strukturerade, de är ofta mer begrepps- och lästunga vilket gör att elever riskerar att missgynnas. Om ett prov ser ut som ovan vid en DIF-analys så kan man vara nöjd, det finns uppgifter som passar olika grupper av elever och dessa uppgifter är också balanserade så att en grupp inte gynnas eller missgynnas i någon större utsträckning.

I år liksom tidigare år kan vi se att de som läser SvA ser ut att klara sig bättre på öppna uppgifter jämfört med slutna.

## 4. Validitet och reliabilitet

### 4.1 Validitet

*Validity refers to the degree to which evidence and theory support the interpretations of test score for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests. The process of validation involves accumulating relevant evidence to provide sound scientific basis for the proposed score interpretations.*<sup>3</sup>

Enligt ovanstående citat innebär validitet att tolkningen av provresultatet stödjer det som provet var avsett för. All information som kan inhämtas om ett prov bör göras för att undersöka om proven är valida. Ibland brukar man på ett enkelt sätt säga att validitet är det samma som att man mäter det som man avser att mäta.

Citatet ovan vidgar dock begreppet något då det även väger in tolkningar som vi gör och det är dessa tolkningar som vi ska se på.

Begreppet valid kommer från latin och har från början betydelsen ”stark”, i svensk betydelse kan man använda ordet ”giltig”.

Ett av problemen med validitet är vilka bevis som används för att mäta korrelationen mellan den teoretiska definitionen och den operationella definitionen. När det gäller provet i samhällskunskap är den teoretiska definitionen kunskapskravens innehåll.

Ett prov konstrueras för att mäta det som internationellt ofta kallas för bakomliggande ”*construct*”. Ett nationellt provs bakomliggande *konstrukt* är kursplanen i ämnet och bedömningen via kunskapskraven.

Validitetsprocessen för ett nationellt prov i samhällskunskap ser ut påföljande sätt:

1. Provggruppen analyserar kursplanen och kunskapskraven noggrant och funderar över operationaliseringar i form av uppgifter och vilka effekter på validiteten som olika tänkbara format skulle kunna ha.
2. Validering gentemot lärarna. Vår arbetsgrupp består av både anställda på Göteborgs universitet och lärare som arbetar med oss en dag i veckan och i skola övriga dagar. Under åren 2018-2019 har vi haft fyra lärare från olika skolor som arbetat med oss en dag/vecka. Vi är måna om att få ny input i den här gruppen, vilket innebär att vi strävar efter att nya personer ska komma in i gruppen med jämna mellanrum och då slutar någon av de andra.
3. Vi har under ett verksamhetsår flera olika tillfällen då våra uppgifter och bedömningsanvisningar valideras:
  - a. Alla lärare och elever som deltar i våra utprövningar ger sina synpunkter till oss i en enkät. Till 2019 års prov så har 4403 elever och 122 lärare deltagit. Lärare som deltar i våra utprövningar anmäler sitt intresse att hjälpa oss via enkäten som medföljer provet när det genomförs i april/maj. Dessa listor förnyas varje år då lärarna återigen får meddela oss om de vill vara kvar på listan.

---

<sup>3</sup> *Standards for Education och Psychological Testing* (2014)

Detta standardverket för amerikansk provverksamhet. Det kan jämföras med ramverk för prov, och litteraturen används också internationellt.



- b. Våra öppna uppgifter som prövas ut bedöms av verksamma lärare i samband med bedömdagar hos oss, dessa äger vanligtvis rum två gånger per år. På dessa lärarsamlingar för provet 2019 har 31 lärare från lika många skolor och kommuner medverkat. Till dessa tillfällen har vi målsättningen att hälften av lärarna ska vara nya och hälften ha varit hos oss tidigare, tyvärr har vi märkt att det blivit svårare att knyta helt nya lärare till oss. Ytterligare en för oss viktig aspekt är att hela landet, från norr till söder, är representerade samt att dessa också kommer från olika typer av skolor.
  - c. Bedömningsanvisningarna till de öppna uppgifterna i provet valideras gentemot verksamma lärare ytterligare en gång i processen. Detta sker ca sju månader före provet genomförs i skolorna. Detta tillfälle benämns bedömningsanvisningsdagar och till detta prov genomförde vi dagarna 3-4 oktober 2018. Gruppen bestod av 16 lärare, varav endast tre var helt nya i vårt sammanhang. Några av deltagarna hade varit med en gång tidigare. Till detta möte bjöd vi in de lärare som deltog i den första utprövningen till provet som ska genomföras 2020, denna utprövning skickades ut till skolorna i september 2019.
  - d. Det sista steget i valideringsprocessen är kravgränssättningsmötet då lärare från olika skolor i landet hjälper oss med att ta fram kravgränserna för probetygen. Mötet ägde rum 15-16 januari 2019, med 16 lärare från 15 olika kommuner. Vår målsättning är att ha fler lärare som är nya eller ”nästan nya” jämfört med de som varit med på våra sammankomster många gånger. I år hade vi 6 nya/nästan nya och 10 som samarbetat med oss minst två gånger..
5. Under senare delen av vår framtagningsprocess så har vår expertgrupp avseende ämnesinnehåll gått igenom alla uppgifter, även detta är ett led i valideringsprocessen. Experterna kommer från Statsvetenskapliga institutionen och Institutionen för nationalekonomi, båda Göteborgs universitet. Dessa personer är ämnesföreträdare men har också vid något tillfälle i sina karriärer undervisat på ämneslärarutbildningarna. Vi har också kontakt med en advokat och en jurist som granskar uppgifter med det specifika ämnesinnehållet.

En stor del av vår kontinuerliga validering utgörs av vårt arbete med våra statistiska metoder. Under arbetet med bedömningsanvisningarna använder vi flera statistiska analysmetoder för att se hur dessa fungerar. IRT (Item Response Theory) är till mycket stor hjälp och vi utvecklar våra analysmetoder kontinuerligt. Till varje prov sammanställs en rapport, där resultatet baseras på urvaleleverna. Denna publiceras inte då den också är uppgiftsspecifik avseende sekretessbelagda uppgifter. Rapporten är omfattande och arbetas fram så snart vi har tillgång till materialet.

Validitet innebär att alla steg i kedjan kunskapskrav – uppgift – prövning – bedömningsanvisning – tolkning av dessa genom lärare ska fungera.

## 4.2 Reliabilitet

Reliabilitet innebär att det valida provet ska mätas på rätt sätt. Ett prov kan inte vara valitt utan god reliabilitet. Provkonstruktörer analyserar ofta sina prov med flera olika korrelationsmått för att kunna belysa provets för- och nackdelar på olika sätt. Provet i samhällskunskap 2019 är analyserat med hjälp av tre vanliga korrelationsmått.

### *Korrelation – Cronbach alpha*

Reliabilitetsmåten avser ibland s.k. inre konsekvens, dvs om uppgifter och elevsvar ”samverkar” så finns det konsekvens i resultatet. Nationella provet Sh 9 2019 har ett reliabilitetsmått på 0,9404, detta är det näst högsta värdet provet har haft.

Tabell 4. Cronbach alpha, Np Sh årskurs 9 2019

| Np Sh |        |
|-------|--------|
| 2013  | 0,911  |
| 2014  | 0,925  |
| 2015  | 0,932  |
| 2016  | 0,941  |
| 2017  | 0,933  |
| 2018  | 0,943  |
| 2019  | 0,9404 |

Reliabiliteten i ett prov är beroende av antalet uppgifter. 2019 hade provet 24 uppgifter jämfört med 30 uppgifter 2013. Reliabiliteten anses som hög, speciellt med tanke på den höga andelen öppna uppgifter.

### *Korrelation – SEM*

Genom att använda IRT vid utprövningar ger det information om de olika uppgifternas svårighetsgrad i förhållande till elevernas olika förmåga/kunskap inom samhällskunskap. Detta ger möjlighet att tillse att elever med olika förmåga alltid har uppgifter att arbeta med i provet. Både de elever som är ”duktiga” i ämnet och de som ”inte så duktiga” i samhällskunskap måste erbjudas uppgifter att arbeta med, annars förmår provet inte att skilja eleverna åt på ett bra sätt. Det måste finnas uppgifter till alla elever för att ett prov ska vara ett bra prov.

För att testa om provet är ett bra prov, kan man mäta SEM, *standard error of measurement*, och TIF, *test information function*. Arbetet med att använda dessa informationskällor för att göra ett bra prov kommer att utvecklas mer kommande år.

Med hjälp av måtten kan vi här, även om vi inte visar detaljer, se en del intressanta resultat:

Provet uppgifter är bra på så vis att de ger elever med både starkare och svagare förmåga i ämnet uppgifter att arbeta med. Detta innebär att provet i samhällskunskap 2019 ger uppgifter till alla elever. Det är viktigt att dessa uppgifter är av god kvalitet och har god mätsäkerhet, så att resultatet för enskild elev kan anses med tillräcklig säkerhet givet.

Ett reliabilitetsmått som används för att mäta detta är SEM. Det anges ibland som mer användbart än t.ex. Cronbach alpha och skrivs fram som ett viktigt mått i Skolverkets systemramverk. Standards for Education and Psychological Testing skriver:

*The SEM is an indicator of a lack of consistency in the scores generated by the testing procedure for some population. A relatively large SEM indicates relatively low reliability/precision.*<sup>4</sup>

Inom modern mätteori, IRT, kan man också mäta SEM över hela den skala man avser att mäta. En vanligt förekommande gräns för mycket god precision är ett CSEM-värde under 0.3 på hela poängskalan där man har sina gränser för olika provbetyg. Detta mått är relativt hårt men ger en tydlig fingervisning om att den elev som fått X poäng också verkligen ska ha det, och inte av slump någon annan gång får fler eller färre poäng.

Np Sh 19 har en lägsta kravgräns på 16 belägg och högsta på 58 belägg (detta är gränserna mellan provbetyg F/E respektive B/A).

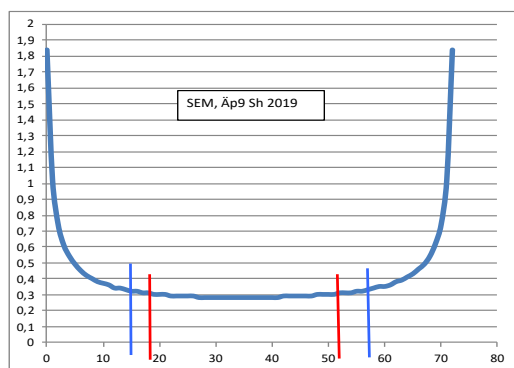


Diagram 4. SEM, NpSh 9, 2019

SEM sjunker under 0.3 vid 19 belägg och stiger över vid 58 belägg, och har 0.34 vid 58 belägg. Lägsta nivå är 0.28 (mellan 27-41 belägg). Detta är ett bra resultat. Provet har en god precision vilket gör att vi kan uttala oss om elevernas antal belägg verkligen stämmer inom de gränser som provet har. Provet 2019 har något bättre mätsäkerhet än provet 2018, avseende de elever som presterar något sämre.

En mycket bidragande del i detta är att bedömningsanvisningarna särskiljer de olika nivåerna inom varje uppgift på ett bra sätt. SEM kan alltså användas för att utvärdera noggrannheten i bedömningsanvisningarna.

Tillförlitligheten (reliabiliteten) i det nationella provet samhällskunskap är god. Detta är resultatet av relativt stora utprövningar som analyserats ur många olika perspektiv och med hjälp av flera olika metoder. Tekniska mått kan ibland te sig tekniska, men om provet inte är tillförlitligt både avseende validitet och mätsäkerhet så blir provet orättvist och mäter inte det som är avsett att mäta. Nationella provet i samhällskunskap åk 9 2019 var ett – på flera olika sätt – tillförlitligt prov.

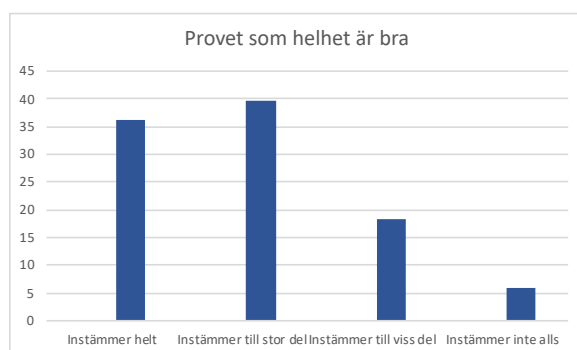
---

<sup>4</sup> Standards for Educational and Psychological Testing, 2014

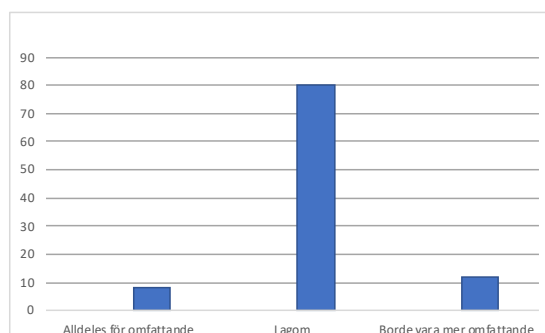
## 5. Enkätresultat med kommentarer

De strukturerade frågorna i enkäten besvarades av 174 lärare och i de frågor det fanns möjlighet att skriva kommentarer utnyttjades det i snitt av 37 lärare. Detta är något högre jämfört med förra året men betydligt lägre än för 4 år sedan då 396 lärare besvarade enkäten.

Det nationella provet i samhällskunskap 2019 mottogs väl. Av de lärare som har svarat på enkäten så har 79,5% av lärarna svarat att de instämmer helt eller instämmer till stor del med påståendet ”provet är som helhet bra”. Detta är något bättre än 2018.



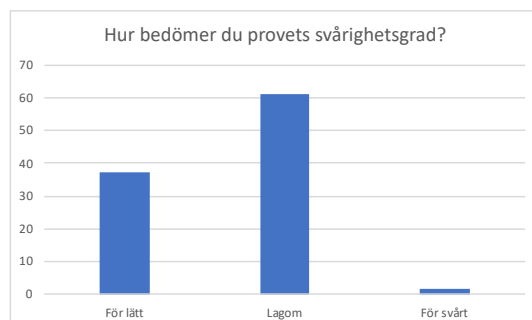
### Vad anser du om omfattningen av ämnesprovet i samhällskunskap?



2019 anser lärarna att provet är mer lagom omfattande än 2018 (då 52%). I frisvarsdelen framkommer att en del lärare anser att provet innehåller för många slutna format i relation till öppna. Detta är något nytt för oss, sedan 2013 har lärarna i enkäten framfört önskemål om fler slutna uppgifter än öppna. Nu kan åsikterna kring detta ha svängt.

Enkäten visar också tydligt att vi prövar det som undervisats om, 98% av lärarna instämmer i detta påståendena ”stämmer i hög grad” och ”stämmer i ganska hög grad”. I kommentarerna kan vi läsa att uppgifterna och innehållet känns välbekant både av lärare och elever, något som provinstitutionen upplever som positivt. Detta är en mycket viktig fråga för oss, vi vill vara säkra på att det vi ställer frågor om också har behandlats i undervisningen. Vi är mycket glada över detta resultat då vi lägger mycket tid och kraft på utformandet av uppgifterna. Om eleverna känner igen innehållet blir de mer positivt inställda till att genomföra provet.

## Hur bedömer du svårighetsgraden i provet?

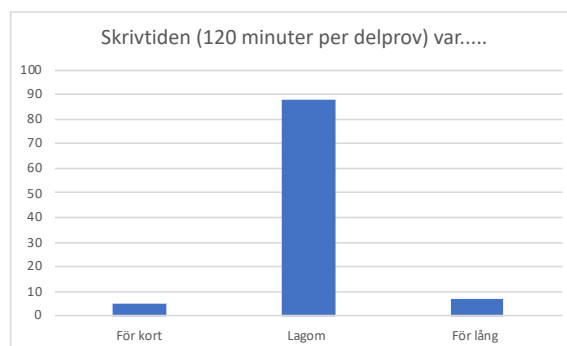


De allra flesta anser att provets svårighetsgrad är lagom, 2019 kan vi dock se att fler tycker det är för lätt än tidigare år. Detta återkommer också i frisvarsdelen, där kommentarerna rör att det var en bra blandning mellan svårare och lättare frågor samt att det upplevdes som att det var relativt lätt att nå E men svårt att nå A. Även här får de slutna formaten viss kritik

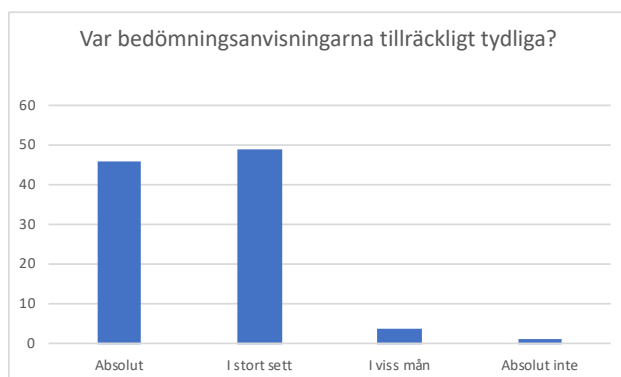
Det är viktigt att provet innehåller uppgifter av varierande format, alltså både uppgifter som kräver skrivinsats och uppgifter som är slutna och inte kräver långa svar. 2019 svarade 80% av lärarna att de var nöjda med balansen mellan slutna uppgifter och öppna, 2018 var den siffran 72%.

### Provets skrivtid

Provets skrivtid är 120 minuter, detta upplevde en större grupp än tidigare som lagom. Något som vi är nöjda med eftersom vi inte vill att många elever ska känna tidsstress under provgenomförandet. En förklaring till detta är troligen den ökade andelen slutna format.



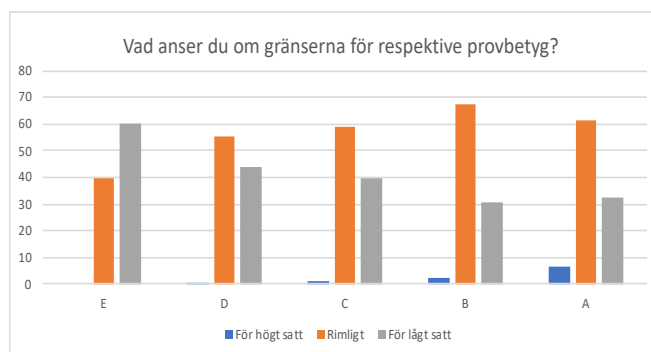
### Var bedömningsanvisningarna tillräckligt tydliga?



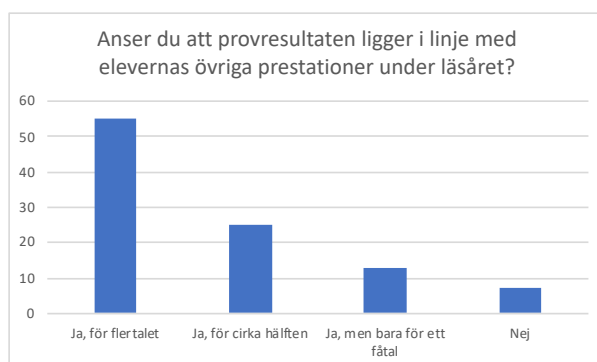
För att ett nationellt prov ska bli reliabelt behövs bedömningsanvisningar som lärarna upplever som stöd. De allra flesta lärarna upplever att stödet gett ett bra eller tillräckligt stöd, 95%. Endast ett fåtal upplever stödet i bedömningsanvisningarna som svagt.

Bland kommentarerna finns att lärarna gärna hade sett exempel på svar som inte når E eller fler elevsvarsexempel. Vi får också beröm för vår tvåspaltiga layout som upplevs som lättförstådd.

### Vad anser du om gränserna för respektive provbetyg?



### Anser du att provresultaten ligger i linje med elevernas övriga prestationer under läsåret?



2019 har gränserna för respektive provbetyg upplevts som för lågt satta. Gränserna är framtagna på

kravgränssättningsmöten där aktiva lärare är med och tar fram gränserna. I år har provet upplevts som relativt enkelt och det syns också i de inrapporterade resultaten. Lärarna upplever ändå att provresultaten till större del (för flertalet alltså för cirka hälften) 80%, ligger i linje med elevernas övriga prestationer under läsåret.

## 6. Avslutning

Utifrån enkäten så har provet tagits emot bra och det återspeglar lärarnas syn på ämnet och deras syn på elevernas kunskapsnivå. Provet har så kallad ”igenkänningsvaliditet”. Utifrån våra analyser så har uppgifterna haft låg tröskel, eleverna har verkligen försökt att lösa uppgifterna, något som är viktigt för oss.

Np Sh 2019 hade fler uppgifter med slutna format än öppna, detta återspeglas i enkäten. Svårighetsgraden på uppgifterna upplevs som lagom och vi har i år ingen uppgift som upplevs som svårförstådd av många elever.

Provet ger eleverna många möjligheter att visa kunskaper för nivå A. I elevhäftena har tankerutan mottagits väl, den fungerar som stöd till eleverna. Bedömningsanvisningarna och vår layout av dessa faller väl ut.

Våra analyser visar på ett prov med relativt höga värden för reliabilitet, bedömningsanvisningarna har kunnat skilja de olika nivåerna i elevsvaren

Områden för fortsatt utveckling:

- Ytterligare analys och utveckling av strukturerade svarsalternativ, så att de på ett reliabelt och valitt sätt mäter det som de är avsedda att mäta.
- Vi har ett stort samarbete med lärare i olika skolor, detta avser vi att behålla och än mer utöka, detta är mycket viktiga inslag i provens valideringsprocess. För samhällskunskapsprovet har denna kontakt varit väldigt viktig inte minst avseende innehåll och utformning på uppgifter.
- Vi avser att öka vår omfattning av mindre utprövningar för att än mer säkra våra uppgifter.
- Våra statistiska bearbetningar kommer att utvecklas och utökas än mer, speciellt med avseende på de slutna formaten.
- För fjärde året i rad har vi haft en frivillig inrapportering. Lärarna kunde där jämföra klassens resultat, på grupp nivå, med rikets. Fler lärare har använt denna möjlighet, i år har nästan 1000 elever inrapporterats. Mycket litet information av detta ges i Lärarinformationen, vi kommer att se över vår hemsida för att försöka sprida denna möjlighet till lärarna.
- Vi behöver se över informationen om betydelsen av lärarenkäten då vi är mycket intresserade av lärarnas synpunkter på vårt prov.

Denna rapport kommer att publiceras på Skolverkets hemsida, SO-ämnenas hemsida och provgruppens hemsida:

[skolverket.se](http://skolverket.se)

[npsportal.se](http://npsportal.se)

[ips.gu.se/forskning/forskningsprojekt/nationella-prov-samhallskunskap/](http://ips.gu.se/forskning/forskningsprojekt/nationella-prov-samhallskunskap/)

## Referenser

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational & Psychological Testing (US). (2014) *Standards for educational an psychological testing*. American Educational Research Association

Löfstedt, A. (2018). *Det första nationella provet i samhällskunskap - en studie i bedömersamstämmighet*. Acta Didactica Norge, 12(4), Art. 13, 16 sidor.



## **Bilaga 1: utprövningar 2019**

### **Utprövning 919**

Antal elever 1104  
Antal skolor 32  
Antal lärare 34

### **Utprövning 929, som var både en digital och analog utprövning**

Antal elever 1253  
Antal skolor 32  
Antal lärare 33

### **Uprövning 939**

Antal elever 1331  
Antal skolor 32  
Antal lärare 34

### **Utprövning 949**

Antal elever 715  
Antal skolor 20  
Antal lärare 21

### **Totalt**

Antal elever 4403  
Antal skolor 116  
Antal lärare 122

