



Statistics Sweden

Statistiska centralbyrån

Teknisk Rapport

En beskrivning av genomförande och metoder

Kunskapsmöten unga – lärare - forskare

130111

SCB, Stockholm
08-506 940 00

SCB, Örebro
019-17 60 00

www.scb.se



Inledning

Enheten för statistik om utbildning och arbete vid Statistiska centralbyrån (SCB) genomförde under perioden september 2012 – januari 2013 en undersökning på uppdrag av Vetenskap & Allmänhet.

Syftet med undersökningen var att belysa lärares och skolledares forskningsanknytning och deras attityder till forskare och forskning.

Populationen utgjordes av 215 523 förskollärare, lärare och skolledare (201 595 personer efter korrigerings för skattad övertäckning). Skolledare i förskolan ingår inte då de inte var möjliga att identifiera i urvalsramen vid urvalsdragningen.

Urvalet bestod av 2000 personer och det var totalt 902 personer som besvarade frågeblanketten. I urvalet identifierades 43 personer som övertäckning. Svarsfrekvensen var 46,1 %.

Resultatet i form av datafil, tabeller och diagram levererades vecka 3 enligt överenskommelse.

Undersökningsledare på SCB var Tomas Westling. Johan Löfgren och Fredrik Carlsson var metodstatistiker. VA:s kontaktperson gentemot SCB var inledningsvis Karin Hermansson och Heidi Armbruster-Domeyer och därefter Maria Lindholm.

Omfattning

Population och urval

Populationen, d.v.s. de objekt som man vill kunna dra slutsatser om, utgjordes av individer.

För att kunna dra ett urval från populationen skapades en urvalsram som avgränsade, identifierade och möjliggjorde koppling till objekten i populationen. Urvalsramen i undersökningen skapades med hjälp av data från Registret över pedagogisk personal (Lärrregistret) på SCB och omfattade lärare och skolledare med minst 25 % tjänstgöringsomfattning. Lärrregistret avser 2011/12. En urvalsram för förskollärare i kommunal sektor skapades från undersökningen "Lönestrukturstatistik, primärkommunal sektor" och en urvalsram för förskollärare i privat sektor skapades från undersökningen "Lönestrukturstatistik, privat sektor". "Lönestrukturstatistik, primärkommunal sektor" är en totalundersökning medan "Lönestrukturstatistik, privat sektor" är en urvalsundersökning. Totalt antal förskollärare i privat sektor är känt från Skolverkets statistik över personal i förskolan. Då det inte var möjligt att identifiera totalt antal förskollärare i privat sektor i undersökningen "Lönestrukturstatistik, privat sektor" har den uppgiften istället hämtats från Skolverkets statistik och sedan använts vid estimationen. Skolverkets personalstatistik innehåller inga individuppgifter varför det inte var möjligt att använda den källan som urvalsram. Skolleda-

re i förskolan ingår inte då de inte var möjliga att identifiera i urvalsramen. Antalet personer i urvalsramen var 215 523.

Urvalsramen stratifierades efter befattningskod, skolform och huvudman. Totalt bildades 28 strata. Från urvalsramen drogs ett stratifierat obundet slumpmässigt urval om 2000 personer (se tabell 1 nedan).

Ett stratifierat obundet slumpmässigt urval innebär att alla objekt inom ett stratum har samma sannolikhet att komma med i urvalet.

Urvalsstorleken bestämdes i samråd mellan SCB och Vetenskap & Allmänhet till 2000 personer.

Tabell 1 Population och urval fördelat på stratum

Stratum	Namn	Population	Urval
L01	Lågstadielärare i kommunala skolor	20553	125
L02	Lågstadielärare i fristående skolor	2485	125
L03	Mellanstadielärare i kommunala skolor	11617	125
L04	Mellanstadielärare i fristående skolor	1071	125
L05	Högstadielärare i kommunala skolor, naturv. inr.	5644	50
L06	Högstadielärare i fristående skolor, naturv. inr.	1093	50
L07	Högstadielärare i kommunala skolor, samh. inr.	10009	50
L08	Högstadielärare i fristående skolor, samh. inr.	1963	50
L09	Högstadielärare i kommunala skolor, övr. inr.	14608	50
L10	Högstadielärare i fristående skolor, övr. inr.	2254	50
L11	Gymnasielärare i kommunala skolor, naturv. inr.	4362	40
L12	Gymnasielärare i fristående skolor, naturv. inr.	1438	40
L13	Gymnasielärare i kommunala skolor, samh. inr.	9256	40
L14	Gymnasielärare i fristående skolor, samh. inr.	2841	40
L15	Gymnasielärare i kommunala skolor, yrkesämne	6132	90
L16	Gymnasielärare i fristående skolor, yrkesämne	1527	90
L17	Gymnasielärare i kommunala skolor, övr. inr.	3745	40
L18	Gymnasielärare i fristående skolor, övr. inr.	1171	40
L19	Övriga lärare, kommunala skolor	13601	40
L20	Övriga lärare, fristående skolor	1217	40
L21	Lärare i förskoleklass, kommunala skolor	8276	50
L22	Lärare i förskoleklass, fristående skolor	600	50
L23	Förskollärare i kommunal sektor	64067	100
L24	Förskollärare i privat sektor	18703	100
R01	Rektorer i kommunala grundskolor	4458	100
R02	Rektorer i fristående grundskolor	914	100
R03	Rektorer i kommunala gymnasium	1251	100

R04	Rektorer i fristående gymnasium	667	100
Totalt		215523	2000

Frågor/Variabler

Vetenskap & Allmänhet utformade frågorna i frågeblanketten i samarbete med SCB.

Därefter genomfördes expertgranskning av mätteknisk expertis i syfte att minska risken för mätfel. Grundläggande för bra kvalitet i en undersökning är kvaliteten på de data som samlas in. För att säkerställa att frågorna fungerar så bra som möjligt och enligt intentionerna har därför alla blanketter genomgått ett mättekniskt test.

Blanketten bestod av 18 numrerade frågor, flera av dem hade delfrågor vilket genererade totalt 50 frågor.

Förutom de variabler som samlades in via frågeblanketten hämtades följande variabler från Registret över pedagogisk personal och undersökningarna "Lönestrukturstatistik, primärkommunal sektor" och "Lönestrukturstatistik, privat sektor".

- Skolans kommungruppstillhörighet

Referensperiod för enkätfrågor, se bifogad frågeblankett. Med referensperiod menas vilken tidpunkt svaren avser, t.ex. idag, förra veckan, senaste tre månaderna.

Datainsamling

Frågeblanketterna skickades ut via webb. I ett informationsbrev ombads personen att besvara frågorna och skicka svaret till SCB. Det första utskicket genomfördes den 19 oktober 2012. Sedan skickades två påminnelser till dem som inte besvarat frågeblanketten. En första påminnelse med ny enkät skickades den 1 november och en andra påminnelse med ny enkät den 15 november. Insamlingen avslutades den 10 december 2012.

Tabell 2 Beskrivning av inflödet. Antal och andel

	Antal	xx,x
Efter första utskick	379	18,9
Efter enkätpåminnelse 1	333	16,7
Efter enkätpåminnelse 2	190	9,5
Totalt	902	45,1*

Bortfall	1055	52,7
Övertäckning	43	2,2
Urval	2000	100,0

*) Beräkning av den definitiva svarsandelen görs enligt följande:

$$902/(2000-43) = 46,1 \%$$

I informationsbrevet kunde uppgiftslämnarna läsa om undersökningens bakgrund, syfte och att undersökningen genomfördes i samarbete mellan Vetenskap & Allmänhet och SCB. Informationsbrevet informerade också om att uppgifter hämtades från Registret från pedagogisk personal (Lärarregistret och undersökningarna "Lönestrukturstatistik, primärkommunal sektor" och "Lönestrukturstatistik, privat sektor" och att en aidentifierad datafil levererades till Vetenskap & Allmänhet. Brevet informerade även om personuppgiftslagen samt offentlighets- och sekretesslagen och att det var frivilligt att medverka i undersökningen.

För att SCB ska kunna lämna ut data från en enkätundersökning krävs informerats samtycke av uppgiftslämnarna. Det innebär att de genom att besvara blanketten och skicka in den godkänner att deras svar kompletteras med de bakgrundsvariabler, och behandlas på det sätt, som beskrivs i informationsbrevet.

Datainsamlingen genomfördes av enkätenheten vid SCB.

Datainsamlingen genomfördes via webb samt skanning av de frågeblanketter som kommit in via post.

De enkäter som kommit in via post registrerades med hjälp av skanning.

Kontroller har genomförts bland annat för att säkerställa så att endast valida värden förekommer i materialet.

SCB kan inte garantera att den utvalda personen själv besvarat frågeblanketten. Någon kontroll av att rätt person har besvarat frågeblanketten har inte gjorts.

Svarsdatafilen kompletterades med vikter för uppräkningsnivå.

Bortfall

Bortfallet består dels av objektbortfall, som innebär att frågeblanketten inte är besvarad alls, och dels av partiellt bortfall som innebär att vissa frågor i blanketten inte är besvarade. Om bortfallet skiljer sig från de svarande, med avseende på undersökningsvariablerna, så kan skattningarna som grundar sig på enbart de svarande vara skeva. För att reducera bortfalls-skevheten har vikter beräknats med hjälp av kalibrering (se bilaga).

Objektsbortfall kan bland annat bero på att uppgiftslämnaren inte är villig att delta i undersökningen, att uppgiftslämnaren inte går att nå eller att uppgiftslämnaren är förhindrad att medverka. Objektsbortfallet i denna undersökning redovisas i tabell 3 nedan.

Tabell 3 Beskrivning av objektsbortfall

	Antal
Ej avhörda	1031
Postreturer	8
Avböjd medverkan	15
Ej anträffad	1
Totalt	1055

Med "Ej avhörda" menas att ingen uppgift om varför frågeblanketten/intervjun inte är besvarad har lämnats. Med "Avböjd medverkan" menas att SCB meddelats att uppgiftslämnaren inte vill medverka i undersökningen. "Ej anträffad" innehåller hemlig eller saknad adress i Registret över totalbefolkningen (RTB).

Partiellt bortfall kan bero på att en fråga är svår att förstå, är känslig, att uppgiftslämnaren glömmer att besvara frågan eller att instruktionerna vid hoppfrågor misstolkas. Det partiella bortfallet varierar mellan 1 och 2 procent för de flesta av frågorna. Det högsta partiella bortfallet, 11 procent, finns i fråga 13_01. En förklaring till att det partiella bortfallet är högt på fråga 13 kan vara att den innehåller frågor som inte är relevanta för samtliga svarande.

Viktberäkning och estimation

För varje svarande person (kallas objekt i fortsättningen) har en vikt beräknats. Syftet med detta är att kunna redovisa resultat för hela populationen och inte bara för de svarande. Vikten kallas därför även för uppräkningsstal.

Vikterna har beräknats utifrån urvalsdesignen samt antaganden om objektsbortfall och täckningsfel. Beräkningen gjordes med hjälp av ett av SCB egenutvecklat SAS-makro (CLAN).

Vikterna i denna undersökning kan beskrivas med formeln:

$$w_k = d_k \cdot v_k$$

Där w_k = vikt/uppräkningsstal för objekt k

d_k = designvikt

w_k = kalibreringsvikt baserad på hjälpinformation

Designvikten är den del av vikten som beror på urvalsdesignen. Vid bortfall kan det vara så att vissa grupper av urvalet svarar i större utsträckning än övriga, t.ex. kan kvinnor svara i högre grad än män. Om de grupper som svarat i högre grad har en annan fördelning på undersökningsvariablerna än övriga kan detta ha en snedvridande effekt på resultatet. För att kompensera för detta har kalibreringsvikter använts (se bilaga för mer detaljer). Vikterna har även justerats för observerad övertäckning.

För en mer utförligare beskrivning av kalibreringsestimatorens se Lundström och Särndal (2001): *Estimation in the Presence of Nonresponse and Frame Imperfections*, Statistics Sweden.

Vikterna/uppräkningsstalen multipliceras med objektens variabelvärden för att skapa statistikvärden gällande för populationen. Om vikterna inte används så kan resultaten bli helt missvisande. Vikterna kompenserar för objektsbortfallet men inte för det partiella bortfallet.

För beräkning av skattningen av totaler används följande formel:

$$\hat{Y} = \sum_r w_k y_k$$

där w_k = vikt/uppräkningsstal för objekt k
 y_k = variabelvärde för objekt k

summering sker av de svarande (r)

Och för beräkning av skattningen av medelvärden används följande formel:

$$\hat{\bar{Y}} = \frac{\sum_r w_k y_k}{\sum_r w_k}$$

där w_k = vikt/uppräkningsstal för objekt k
 y_k = variabelvärde för objekt k

summering sker av de svarande (r).

Statistikens tillförlitlighet

Ramtäckning

Täckningsfel, under- och övertäckning, innebär att urvalsram och population inte helt stämmer överens. Undertäckning innebär att vissa objekt som ingår i populationen saknas i urvalsramen. Övertäckning innebär att objekt som inte ingår i populationen ändå finns i urvalsramen. Ett sätt att minska täckningsfelen är att ha bra och uppdaterade register. I denna undersökning var det 43 personer som identifierades som övertäckning vid datainsamlingen. Merparten av dessa (31 st) var sådana som hade ett annat arbete än förskollärare, lärare eller skollärdare. I den återstående övertäckningen var nio personer pensionärer, två personer hade flyttat utomlands och en person var avliden.

Det fanns 31 personer som hade skickat in enkäten men endast besvarat fråga 1 och fråga 2 och där det utifrån svaret på fråga 2 gick att identifiera de som ej tillhörande populationen. De har kodats om till "Övertäckning".

Storleken på undertäckningen är svårare att uppskatta. Undertäckningen utgörs av personer som arbetade som lärare läsåret 2011/12 men inte fanns i registret, t. ex. att personen började att arbeta som lärare efter oktober 2011.

Urval

Denna kvalitetskomponent avser fel som uppkommer på grund av att endast ett urval av populationen undersöks. Urvalsfel är således den avvikelser mellan ett skattat värde och det faktiska värdet som beror på att man inte undersöker alla objekt i populationen. Urvalsfelets storlek minskar med en ökad urvalsstorlek.

Mätning

Ett fel som kan uppstå vid mätning är att lämnade uppgifter skiljer sig från faktiska uppgifter. Felet kallas mätfel och kan uppkomma då uppgiftslämnaren inte minns de faktiska uppgifterna, missförstår frågan eller medvetet svarar felaktigt.

Bearbetning

Vid den manuella och maskinella bearbetningen av datamaterialet kan bearbetningsfel uppstå. Exempel på bearbetningsfel är registreringsfel och kodningsfel. Dessa fel kan förhindras och upptäckas i de kontroller som genomförs vid dataregistreringen. I den här undersökningen bedöms registreringsfelet vara litet eftersom frågeblanketten endast hade fasta svarsalternativ.

Bortfall

Bortfallsfel inträffar om objekten i bortfallet och de svarande skiljer sig åt avseende frågorna/variablerna i undersökningen. En beskrivning av bortfallet finns i den bortfallsanalys som gjorts i samband med valet av hjälpinformation till kalibreringen (se bilaga).

Beskrivning av tabeller/diagram och datafil

Tabeller och diagram

Tabellerna har räknats upp till populationsnivå, vilket innebär att resultatet avser hela populationen och inte endast de svarande. Antalsuppgifterna är således skattningar av antal personer i populationen med den aktuella egenskapen. Att antalet individer är olika för olika frågor beror på hoppinstruktioner då olika delpopulationer avses.

För de följdfrågor som föregås av hoppinstruktioner har selekteringar genomförts vid tabellframställandet. Det betyder att t.ex. för fråga 9 har endast de som svarat "Ja" i fråga 8 ingått i tabellframställningen.

Om en följdfråga är besvarad men dess styrfråga inte är besvarad och svaret på följdfrågan gör att ett visst svarsalternativ på styrfrågan är självklart så har hänsyn till detta tagits vid tabellframställningen.

I andelstabellerna är värdena avrundade till närmaste heltal och värden under 0,5 är därmed avrundade till noll.

Tabellerna har sekretessgranskats så att inga uppgifter för enskilda objekt röjs. Mycket osäkra skattningar, där konfidensintervallet är större än andelsskattningen, har också prickats.

I fråga 1, 2, 3, 9, 10 och 11 fick de svarande ange fler än ett svarsalternativ. Därför summerar inte statistiken till 100 procent.

De tabeller som presenteras i denna rapport har tagits fram enligt de formuler som presenterats i avsnitt "Viktberäkning och estimation" med hjälp av SPSS/SAS och variansprogrammet ETOS.

Datafil

Observera att de bifogade vikterna ska användas vid analys och resultatframställning. I ett stratifierat urval kan datamaterialet inte ses som ett urval av oberoende lika fördelade observationer, som ofta antas i traditionell statistisk teori. Det är därför viktigt att tänka på att de flesta förprogrammerade dataprogram inte klarar av att analysera datamaterialet ifrån en urvalsundersökning på ett korrekt sätt. Var noga med att kontrollera att de dataprogram som används för analys behandlar vikterna på ett riktigt sätt. Att ta fram frekvenser med beaktande av vikter klarar däremot de flesta statistikprogram.

Om det är partiellt bortfall (uppgift saknas) i en fråga har värdet 88 angetts. Vid hoppfrågor har värdet 0 angetts för dem som inte ska besvara följdfrågorna.

Statistiska mått

Materialet lämpar sig bäst att redovisas som totaler eller procentuella andelar för hela målpopulationen eller fördelat på olika redovisningsgrupper.

Redovisningsgrupper

Ofta redovisas statistik inte bara för hela populationen utan också för delgrupper (redovisningsgrupper). Redovisningsgrupper i den här undersökningen kan avgränsas med hjälp av registervariabler (Region) eller bakgrundsfrågor i enkäten (fråga 2, 3 och 4).

Variabeln "Region" har bildats med hjälp av skolans kommunkod och består av tre grupper: Storstad, Stad och Övriga.

I "Storstad" ingår Göteborg, Malmö och Stockholm.

I "Stad" ingår följande kommuner: Borås, Eskilstuna, Falun, Gävle, Halmstad, Helsingborg, Hässleholm, Jönköping, Kalmar, Karlskrona, Karlstad, Kristianstad, Linköping, Luleå, Lund, Norrköping, Nyköping, Skellefteå, Skövde, Sundsvall, Södertälje, Trollhättan, Uddevalla, Umeå, Uppsala, Varberg, Västerås, Växjö, Örebro, Örnsköldsvik och Östersund.

I "Övriga" ingår alla övriga kommuner.

Den regionala indelningen har gjorts i enlighet med indelningar i publikationen "Regionala indelningar i Sverige den 1 januari 2011 (mis 2011.1)".

Variablerna "F2red", "F3red" och "F4red" anger vilken redovisningsgrupp personen tillhör utifrån enkätsvaren på fråga 2, 3 och 4. I de fall personen har svarat flera alternativ på fråga 2 har redovisningsgrupp bestämts utifrån stratumtillhörighet (registervariabler från Registret över pedagogisk personal) vid urvalstillfället. I de fall där enkätsvaren inte överensstämde med registeruppgifterna har det enkätsvar som ligger närmast registeruppgiften valts. Enkätsvaret "Annat" på fråga 2 ingår endast i redovisningen i redovisningsgruppen "Totalt".

På fråga 3 redovisas de personer som har svarat flera alternativ i redovisningsgruppen "Flera olika ämnen". Observera att fråga 3 är en sk hoppfråga som inte ska besvaras av samtliga.

De personer som inte har besvarat fråga 2, 3 och 4 (partiellt bortfall) ingår endast i redovisningen i redovisningsgruppen "Totalt". Det innebär att det skattade antalet personer i de olika huvudsakliga redovisningsgrupperna "Undervisningsnivå", "Sektor", "Undervisningsinriktning" och "Region" skiljer sig åt jämfört med redovisningsgruppen "Totalt". Det partiella bortfallet på frågorna 2, 3 och 4 var:

- *Fråga 2*: 2 personer
- *Fråga 3*: 31 personer
- *Fråga 4*: 24 personer

Jämförbarhet

Jämförbarhet över tid

Detta är första gången som SCB genomför undersökningen.

Sekretess och utlämnande

SCB har gjort en intern sekretessprövning för utlämnandet av data på mikronivå.

Dessutom har ett sekretessförbehåll upprättats mellan Vetenskap & Allmänhet och SCB med avseende på hur det aidentifierade datamaterialet får hanteras.

Övrigt

För att kunna läsa tabellerna måste (oftast) inställningar i Excel ändras.

1. Excel 2003:
För att tabellerna skall fungera måste säkerhetsnivån vara medel.
Detta kontrolleras genom att gå in under "Verktyg - Makro – Säkerhet - Säkerhetsnivåfliken".
När excelfilen öppnas skall även makrot aktiveras.
2. Excel 2007:
Välj Excel-alternativ – Säkerhetscenter – Inställningar för säkerhetscenter – Makroinställningar.
Markera rutorna "Aktivera alla makron" och "Åtkomst till objektmodell för VBA-projekt betrodd".

Bilagor

Frågeblankett och informationsbrev

Kalibreringsrapport