

NOTAT

FOLKESKOLEREFORMEN DOKUMENTATION AF DATAINDSAMLINGEN

ANNE TOFT HANSEN
KASPER SKOU ARENDT
VIBEKE MYRUP JENSEN
METTE FRIIS-HANSEN

KØBENHAVN 2015

FOLKESKOLEREFORMEN
DOKUMENTATION AF DATAINDSAMLINGEN
Afdelingsleder: Mette Deding
Afdelingen for Skole og Uddannelse

© 2015 SFI – Det Nationale Forskningscenter for Velfærd

SFI – Det Nationale Forskningscenter for Velfærd
Herluf Trolles Gade 11
1052 København K
Tlf. 33 48 08 00
sfi@sfi.dk
www.sfi.dk

SFI's publikationer kan frit citeres med tydelig angivelse af kilden.

SFI-notater skal danne grundlag for en faglig diskussion. SFI-notater er foreløbige resultater, og læseren bør derfor være opmærksom på, at de endelige resultater og fortolkninger fra projektet vil kunne afvige fra notatet.

INDHOLD

1	INDLEDNING	5
2	DATAINDSAMLINGSPROCESSEN	9
	Første dataindsamling (2014)	9
	Anden dataindsamling (2015)	11
3	ANTAL BESVARELSER OG SAMMENKOBLING AF DATA I ÅRET	15
	Deltagelse og besvarelsesprocenter for 2014	16
	Sammenkobling af de enkelte skemaer i 2014	18
	Deltagelse og besvarelsesprocenter for 2015	23
	Sammenkobling af de enkelte skemaer i 2015	25
4	PANELDATA – SAMMENKOBLING AF DATA PÅ TVÆRS AF TID	31
	Sammenkobling af paneldata på skoleniveau	31
	Sammenkobling af data på individniveau	33
5	BORTFALDS- OG REPRÆSENTATIVITETSANALYSE	35

Datagrundlaget	35
Bortfalds- og repræsentativetsanalyse	37
Bortfald	38
Repræsentativitet	42
BILAG	49
LITTERATUR	55

INDLEDNING

I 2014 igangsatte Ministeriet for Børn, Undervisning og Ligestilling (MBUL) et omfattende evaluerings- og følgeforskningsprogram, der skal følge folkeskolereformens implementering. Formålet med programmet er, at politikere samt medarbejdere i kommuner og skoler løbende kan lære af erfaringer og resultater ved folkeskolereformens implementering.

Et vigtigt led i følgeforskningen er en stor dataindsamling ved hjælp af spørgeskemaer. Figur 1.1 illustrerer designet af hele dataindsamlingsprocessen fordelt på respondentgrupper. Det drejer sig om elever, forældre, skoleledere, lærere og andet pædagogisk personale¹ samt skolebestyrelsesformænd (herefter SB-formænd). Alene i 2014 gennemførte lidt mere end 34.000 respondenter undersøgelsen.

Hvert år inviteres elever, forældre, skoleledere, lærere og andet pædagogisk personale og SB-formænd til at deltage i undersøgelsen. Ser vi alene på eleverne, så følges nogle af disse gennem hele perioden for dataindsamlingen, mens andre elever kun følges nogle få eller blot enkelte år.

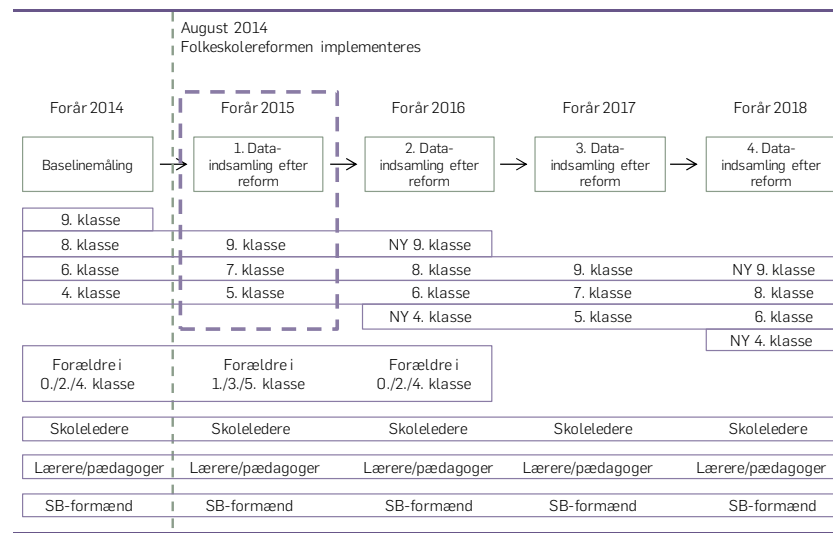
Elever, der fx i skoleåret 2013/2014 gik i 4. klasse, følges frem til de går ud af 8. klasse 5 år senere. For andre årgange, såsom 9. klasse fra skoleåret 2013/14, følger vi kun de enkelte elever ét år, men til gengæld har vi mulighed for at se på udviklingen på de enkelte skoler over tid, idet en ny 9. klasse spørges i 2016 og 2018. For at sige noget om en udvikling er det vigtigt, at et klassetrin følges i mindst 3 år. Derfor inkluderes en ny 4. klasse i 2016, som vi følger, til de går ud af 6. klasse i 2018. Samlet set betyder panelstrukturen, at vi har data fra elever på de fire

1. Betegnelsen "lærere og andet pædagogisk personale" dækker over lærere, pædagoger, pædagogmedhjælpere, børnehaveklasseledere, børnehaveklasseassistenter og undervisningsassistenter. I resten af notatet benytter vi betegnelsen "lærere og andet pædagogisk personale" eller blot "pædagogisk personale" som et samlet udtryk for de forskellige typer personalegrupper.

klassetrin 4., 6. 8. og 9. klassetrin på de enkelte skoler i de 3 år 2014, 2016 og 2018.

FIGUR 1.1

Evaluering af folkeskolereformen, dataindsamling 2014-2018.



Anm.: Figur 1.1 viser deltagerne i dataindsamlingen. Det drejer sig om eleverne på forskellige klassetrin, forældre, skoleledere, lærere/pædagoger og SB-formænd.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

SFI Survey varetog dataindsamlingen i 2014, mens SFI i samarbejde med Danmarks Statistiks survey afdeling (DST Survey) varetager dataindsamlingen i 2015-2016 for MBUL. SFI har efterfølgende foretaget en datarensning inden overlevering af data til Styrelsen for It og Læring (STIL).

STIL, som administrerer dataene på vegne af MBUL, har efterfølgende til opgave at distribuere data til relevante undersøgelser på området, og adgang til data fås ved at ansøge STIL herom. Dette dokument har til formål at beskrive dataindsamlingsprocessen for 2014-2016. Dokumentet opdateres årligt.

Samlet set dækker dataindsamlingen en række centrale emner for analyse af folkeskolereformen. I denne afrapportering af dataindsamlingen går vi ikke nærmere ind i, hvilke emner dataindsamlingen mere konkret vedrører, men henviser til de eksisterende publikationer, der allerede foreligger samt de variabelister, der kan rekvireres hos STIL. Listen opdateres løbende:

- Kjer, M.G., S. Baviskar & S. Winter (2015): *Kortlægning af skoleledelse i forbindelse med implementeringen af folkeskolereformen i 2015*. SFI – Det Nationale Forskningscenter for Velfærd, Rapport, under udgivelse.

- Nielsen, C.P., A.T. Hansen, V.M. Jensen & K.S. Arendt (2015): *Folkeskolereformen, beskrivelse af 2. dataindsamling blandt elever*. SFI – Det Nationale Forskningscenter for Velfærd, Rapport, under udgivelse.
- Keilow, M. & A. Holm (2015): *Udvikling af måleinstrument for elevadfærd og -holdninger, Baseline data fra evaluering af folkeskolereformen*. SFI – Det Nationale Forskningscenter for Velfærd, Rapport 15:09.
- Hansen, A.T., M. Friis-Hansen & V.M. Jensen (2015): *Folkeskolereformen, Beskrivelse af 1. dataindsamling*. SFI – Det Nationale Forskningscenter for Velfærd, Notat. 24.03.2015.

I kapitel 2 beskriver vi processen omkring selve dataindsamlingen, mens vi i kapitel 3 ser på svarprocenterne inden for de enkelte typer af respondenter og de muligheder, der foreligger, når det gælder sammenkoblinger af svar fra de forskellige respondenttyper inden for et enkelt år. Kapitel 4 går skridtet videre og ser på dataenes panelstruktur på tværs af tid. Bilagene er af mere teknisk karakter og beskriver, hvilken type af data, det er muligt at få adgang til, og hvordan de er sammensat.

DATAINDSAMLINGS- PROCESSEN

Primo 2014 inviterede daværende undervisningsminister Christine Antorini 400 skoler til at deltage i en omfattende undersøgelse af implementeringen af folkeskolereformen for perioden 2014-2018. De 400 hovedskoler og dertilhørende afdelinger, svarende til i alt 442 skoler, er tilfældigt udvalgt af STIL. Alle spørgeskemaundersøgelser på nær forældreundersøgelsen er foregået via et webbaseret spørgeskema, mens forældreundersøgelsen er foregået ved hjælp af telefoninterviews. Det er erfarne interviewere ved hhv. SFI Survey i 2014 og DST Survey i 2015, der har gennemført interviews blandt forældre. Skolerne blev tilbudt økonomisk dispensation for at deltage (5.000 kr. årligt).

For alle spørgeskemaundersøgelserne er der medfulgt et informationsbrev, hvor der er blevet orienteret om undersøgelsens baggrund og formål. I brevene er respondenterne ligeledes blevet informeret om, at deres svar vil blive behandlet fortroligt, og at deres anonymitet vil blive sikret i analyseprocessen. En kopi af de konkrete formuleringer i informationsbrevene kan rekvireres hos STIL.

FØRSTE DATAINDSAMLING (2014)

I foråret 2014 blev første dataindsamling gennemført som en baseline-måling inden folkeskolereformens ikrafttræden ved starten af skoleåret 2014/2015. Som tidligere nævnt gennemføres dataindsamlingen blandt en række forskellige respondentgrupper, og for hver gruppe er der anvendt forskellige procedurer i dataindsamlingsprocessen.

ELEVER

Efter at de enkelte skoler accepterede at deltage i undersøgelsen, blev elevlister indhentet hos STIL. Elever i alle klasser på de relevante klassetrin blev inviteret til at deltage (jf. figur 1.1). Dog fik elever i specialtilbud også mulighed for at gennemføre undersøgelsen. Den enkelte skole (i praksis skolelederen) afgjorde selv, om en elev skulle fritages for undersøgelsen. Status og evt. begrundelse er noteret som et led i dokumentationsprocessen. Når en skole godkendte, at eleverne kunne deltage, fik vi følgende oplysninger fra STIL: skolens navn og institutionsnummer, elevens navn, cpr-nummer samt klasse (fx 4B). De elevlister, vi modtog, er de mest opdaterede på det pågældende tidspunkt (primo året).²

Skolerne stod selv for dataindsamlingen blandt eleverne. Det foregik ved, at de enkelte klasser anvendte en undervisningstime på, at eleverne udfyldte online skemaet hver især.

Inden dataindsamlingen modtog skolen informationsbreve og vejledninger til gennemførelse af undersøgelsen. Ligeledes modtog skolen personlige adgangskoder til hver af eleverne, så eleverne kunne logge ind på det elektroniske spørgeskema. Disse koder skulle den spørgeskemaansvarlige lærer videregive til eleverne indledningsvist, samtidig med at han eller hun gav en kort instruks om dataindsamlingens formål.

Alle skoler blev tilbudt teknisk assistance fra en interviewer til gennemførelse af undersøgelsen. Intervieweren har efter aftale forestået dataindsamlingen ude på de enkelte skoler. Det har i den forbindelse ligeledes været muligt at låne et sæt bærbare computere til alle elever i klassen. Interviewerne var ude på ca. 40-50 skoler for at hjælpe med gennemførelse af spørgeskemaundersøgelsen i 2014.

FORÆLDRE

For forældreinterviews udtrak vi en stikprøve blandt forældre til elever i 0., 2. og 4. klasse. For 0. og 2. klasse blev forældrene udtrukket på baggrund af de 238 skoler, der havde accepteret at deltage i undersøgelsen, mens forældrene til 4.-klasses-eleverne blev udtrukket fra de skoler, der havde gennemført elev-delen for 4. klasse. Vi udtrak forældreinterviews til 4.-klasses-eleverne, efter at dataindsamlingen for eleverne havde fundet sted, for at få mest muligt overlap mellem forældrebesvarelser og elevbesvarelser fra samme klassetrin og skole.

Listen over forældreoplysninger fik vi fra STIL. Forældreoplysningerne indeholdt elevens navn, cpr-nummer, skole og klasse, den ene forælders navn (primært moren) og folkeregisteradresse. Hvis eleven ikke boede hos forældrene, leverede STIL institutionsadressen eller plejeforældreadressen. Telefonnummer til forældrene fik vi fra DST's telefon-

2. Det er skolernes ansvar at sørge for, at registre hos STIL er opdaterede. Der forekommer imidlertid undertiden uoverensstemmelser, hvad angår fx klassebetegnelser, som de enkelte skoler anvender til daglig, og de klassebetegnelser, der er registreret hos STIL.

nummerdatabaser. Forældrene modtog et orienteringsbrev og blev efterfølgende ringet op af SFI Surveys interviewerkorps med henblik på at foretage interviewet eller indgå en aftale om et interview.

SKOLELEDERE, SB-FORMÆND, LÆRERE OG ANDET PÆDAGOGISK PERSONALE

I forbindelse med samtykke for deltagelse udleverede skolerne en liste med kontaktoplysninger for dansk- og matematiklærere for de deltagende årgange, andet pædagogisk personale tilknyttet klassen, skoleledere og SB-formænd (navn, e-mail og undervisere på de relevante klassetrin). For de skoler, hvor der eksempelvis blev arbejdet mere bredt i teams fremfor klasser, blev flere lærere inviteret til at deltage. Det var således op til skolen selv at vurdere, hvilke lærere og andet pædagogisk personale der ville være bedst egnet til at besvare spørgsmål om eksempelvis matematikundervisningen på 4. klassetrin.

Skolens lærere og andet pædagogisk personale, skolelederen samt SB-formændene modtog efterfølgende en e-mail med et elektronisk oplysningsbrev med direkte link til undersøgelsen. Skolerne fik to skriftlige påmindelser for at indlevere oplysninger for det relevante pædagogiske personale, skolelederen samt SB-formanden.

TIDSPERSPEKTIV

Dataindsamlingen foregik forskudt for de forskellige typer af respondenter. For eleverne foregik dataindsamlingen fra marts-juni, lærerne og andet pædagogisk personale: april-juni, skoleledere: april-juni, SB-formænd: april-juli, mens dataindsamlingen blandt forældrene foregik juli-september 2014. For alle onlinebesvarelser udsendte vi minimum to skriftlige påmindelser, mens vi kontaktede forældrene op til seks gange, før de blev opgivet.

ANDEN DATAINDSAMLING (2015)

I foråret 2015 gennemførte vi 2. dataindsamling, hvilket vil sige den første dataindsamling efter skolereformens ikrafttræden. For elever, forældre, SB-formænd samt lærere og andet pædagogisk personale var udgangspunktet for 2. dataindsamling bruttolisten for 2014. Det vil sige de 238 skoler, der indleverede respondentlister i 2014.

Primo 2015 blev skolerne indledningsvist mindet om den forestående dataindsamling og det forventede tidspunkt for dataindsamlingen. Processen for dataindsamling for alle respondenter foregik næsten efter samme procedurer som i 2014. Eksempelvis blev skolerne bedt om at indlevere en (opdateret) liste over alle dansk- og matematikundervisere for elever, der nu gik i 1., 3., 5., 7. eller 9. klasse.

I forbindelse med at skolerne indleverede respondentlister, foretog vi to anmodninger pr. mail og efterfølgende op til to telefoniske henvendelser til skolens kontor eller direkte til skolelederen for at bede om respondentlisterne.

SKOLELEDERNE

I 2015 udvidede vi undersøgelsen blandt skoleledere til at indeholde samtlige skoleledere på danske folkeskoler. De er i udgangspunktet ikke en del af panelet for undersøgelse af folkeskolereformen, men inviteret i 2015 og 2016 for at give et bredere indblik på skolelederniveau.

Undersøgelsen blandt de skoleledere, der ikke var med i 2014, foregik sideløbende med den øvrige dataindsamling. Det vil sige, at indholdet i spørgeskemaet var det samme som til de øvrige skoleledere. De modtog ligeledes et direkte link til et online spørgeskema og en personlig kode i deres indbakke.

I sagens natur havde den udvidede skolelederundersøgelse et andet forløb, da de ikke var med i 2014. En anbefaling fra formanden for skolelederforeningen, Claus Hjortdal, understøttede invitationsbrevene til de øvrige skoleledere.

Invitationen blev i første omgang sendt til skolens mail, såfremt vi ikke har skolelederens direkte mailadresse. Ved senere henvendelser fandt og anvendte vi skolelederens egen mail. Vi foretog op til to skriftlige henvendelser og efterfølgende tre telefoniske henvendelser for at bede skolelederne om at besvare det online spørgeskema. I denne forbindelse henvendte vi os både til dem, der endnu ikke havde besvaret skemaet, og dem, der delvist havde besvaret skemaet.

ELEVERNE

Vi inviterede alle elever på de relevante klassetrin og udvalgte skoler til at deltage.

Som i 2014 indhentede vi elevlister direkte hos STIL, men i modsætning til 2014 anvendte eleverne deres personlige UNI-login til at få adgang til undersøgelsen. Dertil programmerede DST en applikation, så det online spørgeskema kunne ”tale sammen” med UNI-login-systemet.

Den primære fordel ved UNI-login-metoden er, at det for skolen og de undersøgelsesansvarlige lærere ikke kræver nær så meget arbejde. Endvidere er UNI-login-metoden i højere grad med til at sikre, at et skema besvares af den person, som det er tiltænkt.

Programmeringen af samkørsel af UNI-login med DST's online spørgeskemasystem var dog ikke problemfrit og medførte, at vi udskød undersøgelsen ca. 3 uger. Vi kan ikke udelukke, at denne udskydelse påvirkede svarprocenten for 2015, til trods for at vi forlængede dataindsamlingsperioden for elev-delen. Dels fordi vi alt andet lige kom tættere på eksamensperioden, dels fordi vi havde meldt en startdato ud til sko-

lerne, som nogle af skolerne anvendte til at planlægge tidspunkt for spørgeskemaundersøgelsen for eleverne.

Anvendelsen af UNI-login har også medført nogle udfordringer. Eksempelvis kunne alle med et UNI-login (elever såvel som lærere) logge sig ind på undersøgelsen, og vi har derfor nogle enkelte besvarelser, der ikke hører til undersøgelsen.

En anden udfordring ved anvendelsen af UNI-login-metoden var, at vi efterfølgende stod med nogle elever, hvor vi har manglende personoplysninger, da der var et mismatch mellem elever i klassen og elever på de modtagne elevlister. Denne udfordring kan ikke løses på forhånd. På grund af eventuelle skoleskift vil de officielle elevlister aldrig være 100 pct. samstemmende med eleverne i den pågældende klasse. Disse personoplysninger kan hentes efterfølgende.

Vi justerer UNI-login-metoden til 2016-dataindsamlingen for de u hensigtsmæssige udfordringer, som vi identificerede i løbet af 2015-dataindsamlingen.

Ligesom for 2014 foretog skolerne primært selv dataindsamlingen i 2015, idet de for hver deltagende klasse anvendte en undervisningstime på at svare på skemaet. Vi tilbød alle skoler assistance fra en interviewer til at gennemføre undersøgelsen og at låne et classesæt af bærbare computere. Cirka 10 skoler tog imod dette tilbud i 2015.

Skolerne har modtaget 2-3 påmindelser vedrørende elevbesvarelser, hvoraf mindst én påmindelse var før sidste frist for besvarelse. For de skoler, der ikke har haft mulighed for at foretage elevundersøgelsen inden for tidsrammen, er der endvidere foretaget særskilte aftaler senere på foråret. Der var således en særdeles fleksibel tidsramme omkring undersøgelsen.

FORÆLDRE

Som i 2014 udvalgte vi forældrene til 1. og 3. klasse på baggrund af bruttolisten af skoler og forældre til elever i 5. klasse efter gennemførelse af elevbesvarelserne. Forældre blev udvalgt, så vi fik minimum 3.000 forældreinterviews. Vi forsøgte endvidere at tage hensyn til, at de samme forældre, der stadig har børn på den pågældende skole, interviewes igen. Dette er for at sikre bedst mulig panelstruktur.

En ulempe ved at udvælge skoler til forældreinterviews ud fra bruttolisten af skoler er, at der altid er risiko for, at vi får nogle skoler, hvor vi kun har forældreinterviews og ikke har besvarelser fra andre typer af respondenter. Det er også sket her.

Vi informerede forældrene om undersøgelsen via brev og har efterfølgende kontaktet dem pr. telefon. Som i 2014 ringede vi primært til moren, og standardproceduren var, at vi forsøgte at kontakte respondenter op til seks gange, før vi opgav.

LÆRERE OG ANDET PÆDAGOGISK PERSONALE SAMT SB-FORMÆND

På baggrund af respondentlisterne med navn og e-mail adresse på relevant personale og SB-formænd sendte vi invitationer ud til den online spørgeskemaundersøgelse, der var relevant for den pågældende respondentgruppe. Invitationen indeholdt et link til undersøgelsen samt en personlig kode. Som for 2014 foretog vi 2-3 skriftlige påmindelser om besvarelse af spørgeskemaet. Disse skriftlige påmindelser var forskellige, i forhold til om respondenterne var påbegyndt sin besvarelse, men ikke færdiggjort den, eller om respondenterne slet ikke havde svaret endnu.

Efter ønske fra skolelederne modtog skolelederen eller skolens kontaktperson to løbende opdateringer mht., hvor langt skolen var kommet med besvarelserne for alle respondentgrupper.

TIDSPERSPEKTIV

Som for 2014 foregik dataindsamlingen lidt forskudt for de forskellige typer af respondenter. For eleverne foregik dataindsamlingen i februar-maj, lærere og andet pædagogisk personale: februar-maj, skoleledere: februar-juni, SB-formænd: marts-maj, mens dataindsamlingen for forældrene foregik marts-maj 2015.

ANTAL BESVARELSER OG SAMMENKOBLING AF DATA I ÅRET

I dette kapitel gennemgår vi svarprocenterne for de forskellige respondentgrupper. Vi beskriver endvidere, hvor stor en andel af besvarelsene der kan kobles sammen på tværs af de forskellige respondenttyper i hhv. 2014 og 2015. Kapitlet præsenterer først data for 2014 og dernæst 2015. Det er vigtigt at bemærke, at vi for 2014 har foretaget lidt flere eksempler på dataoprensning end på 2015-dataene. De rensede datasæt kan rekvireres hos STIL og kan evt. kobles til registerdata fra STIL eller Danmarks Statistik via institutionsnummer (for alle respondenttyper) eller cpr-nummer (elev- og forældrebesvarelser).

Generelt kan det siges om de datasæt, der ligger til grund for nedenstående beskrivelse, at de alle indeholder bruttolister for respondenter (jf. figur 3.1 og figur 3.3). Det vil eksempelvis sige, at datasættet for lærere og andet pædagogisk personale indeholder alle individer, der er inviteret til at deltage i undersøgelsen, og besvarelser for dem, der har valgt at besvare.

Vi har også dannet et grunddatasæt ”skoler442”, der indeholder institutionsnummer og kommune-id på alle de skoler, der blev inviteret til at deltage. Endvidere er der dannet to datasæt, der angiver deltagelse i hhv. 2014 og 2015 for de forskellige typer af respondenter. Tilsammen giver disse tre datasæt et hurtigt overblik over de enkelte skolers deltagelse i undersøgelsen. For en mere udførlig gennemgang af den type af data, som SFI har overleveret til STIL, henvises til bilagene.

DELTAGELSE OG BESVARELSESPROCENTER FOR 2014

Figur 3.1 illustrerer deltagelse i dataindsamlingen for 2014. Ud af de 442 inviterede skoler accepterede 238 invitationen. Det vil sige, at de har indleveret kontaktoplysninger på enten pædagogisk personale (inkl. lærere), skoleledere, SB-formænd og/eller accepteret at lade eleverne deltage.

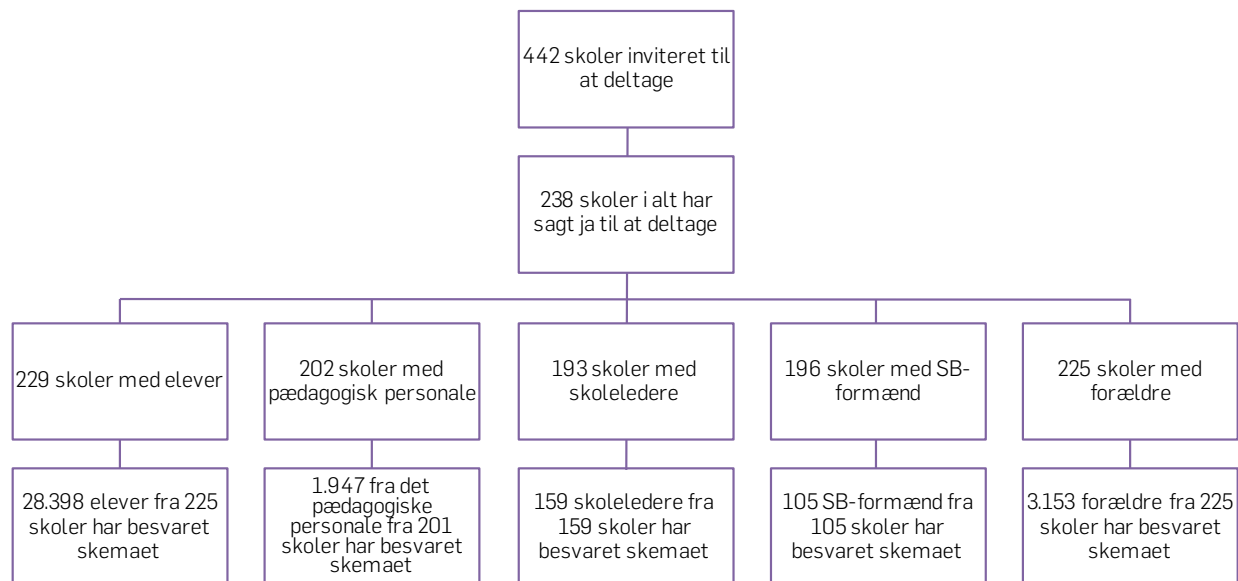
Ud af de 238 deltagende skoler er der modtaget elevlister fra 229 skoler, kontaktoplysninger for det pædagogiske personale fra 202 skoler og kontaktoplysninger fra 193 og 196 skoler for henholdsvis skoleledere og SB-formænd. Der er udvalgt forældre til forældreinterviews på 225 skoler.

Følgende er blevet leveret:

- 225 skoler (98 pct. ud af 229 skoler) har leveret mindst én elevbesvarelse
- 201 skoler (99 pct. ud af 202 skoler) har indgivet mindst én besvarelse fra lærere eller andet pædagogisk personale
- 159 skoler (82 pct. ud af 193 skoler) har indgivet en skolelederbesvarelse
- 105 skoler (54 pct. ud af 196 skoler) har indgivet en besvarelse fra SB-formanden
- For alle 225 skoler er der foretaget mindst ét forældreinterview.

I alt er der modtaget 33.762 besvarelser fra 236 forskellige skoler.

FIGUR 3.1
Deltagende skoler i 2014.



Anm.: Figuren er dannet ud fra "XX_besvar" og INST_NUMMER-variablene i de rensede datasæt og angiver antallet af observationer i bruttolisten samt antallet af respondenter fra skolerne i de forskellige datasæt. I elevdatasættene er der 2 elever ud af de 28.398 besvarelser, hvor vi ikke har oplysninger om institutionsnummer eller cpr-nummer. Disse besvarelser indgår stadig i data som en besvarelse. Endvidere har vi ikke oplysninger om institutionsnummer på 2 lærere eller andet pædagogisk personale og 4 SB-formænd.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

I forlængelse af figur 3.1 viser tabel 3.1 bruttostikprøven for antal skoler, antal respondenter, antallet af besvarelser og opnåelsesprocent for hver type respondent.

TABEL 3.1

Antal skoler og respondenter i bruttolisten samt antal besvarelser (i pct.), 1. dataindsamling, 2014.

Respondenttype	Antal skoler (brutto)	Antal respondenter (brutto)	Antal besvarelser (i pct.)
Elever	229	37.655	28.398 (75 pct.)
Lærere og andet pædagogisk personale	202	4.194	1.947 (46 pct.)
Skoleledere	193	193	159 (82 pct.)
SB-formænd	196	200	105 (53 pct.)
Forældre	225	4.187	3.153 (75 pct.)
I alt	238	46.424	33.762 (73 pct.)

Anm.: For hver respondentgruppe angiver tabellen antallet af skoler og respondenter, der har modtaget spørgeskemaer, og antal respondenter, der har svaret, samt svarprocenten i parentes. Ikke alle respondenter i bruttolisten har et tilgængeligt institutionsnummer. Af de 37.655 elever indgår 2 elever, hvor vi ikke kan identificere, hvilken skole, de går på. Det samme gælder for 2 lærere og 4 SB-formænd. Der er eksempelvis derfor, at vi har 200 SB-besvarelser, men kun 196 skoler. Disse observationer indgår dog stadig i datasættene, hvor deres institutionsnummer er kodet til værdien "999999".

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Vi har for ingen af respondentgrupperne modtaget kontaktoplysninger fra alle deltagende 238 skoler, men for eleverne mangler der kun elevlister fra 13 skoler, hvor der for de andre grupper mangler lidt flere. I alt har hhv. 37.665 elever, 4.194 lærere og andet pædagogisk personale, 193 skoleledere, 200 SB-formænd og 4.187 forældre haft muligheden for at besvare spørgeskemaet i 2014. Heraf har vi i 2014 opnået tilfredsstillende svarprocenter for eleverne, skolelederne og forældrene (hhv. 75 pct., 82 pct. og 75 pct.), mens svarprocenterne er noget lavere for lærere og andet pædagogisk personale (46 pct.) og SB-formændene (53 pct.).

SAMMENKOBLING AF DE ENKELTE SKEMAER I 2014

For at kunne udnytte data er det relevant at vide, hvilke typer af besvarelser der er tilgængelige fra den enkelte skole. Derfor har vi beskrevet en række relevante sammenkoblinger mellem forskellige typer af data (for 2014). Grunddatasættet "besvarelser2014" giver endvidere et overblik over, hvilke skoler der har deltaget i 2014, og med hvilke besvarelser, de har deltaget.

Af tabel 3.2 fremgår antallet af skoler for de forskellige sammenkoblinger af datasæt. Her ses det eksempelvis, at vi modtog minimum én elevbesvarelse og minimum én besvarelse fra lærere og andet pædagogisk personale fra 193 af de udtrukne 442 skoler.

TABEL 3.2

Antal skoler, der går igen på tværs af respondentgruppe. 1. dataindsamling. 2014.

	Elever	Lærere og pædagogisk personale	Skoleledere	SB-formænd	Forældre
Elever		193	155	101	222
Lærere og andet pædagogisk personale	193		157	103	192
Skoleledere	155	157		85	153
SB-formænd	101	103	85		99
Forældre	222	192	153	99	

Anm.: Tabellen angiver antallet af skoler, der har bidraget med minimum én besvarelse i hvert af skemaerne. Eksempelvis har 222 skoler afgivet minimum én besvarelse for forældre og minimum én elevbesvarelse. Antallet af skoler afspejler dog ikke antallet af observationer i de enkelte datasæt. For sammenkobling af besvarelser fra forældre og det pædagogiske personale anvendes alle besvarelser fra det pædagogiske personale og ikke kun besvarelser fra undervisere i 0., 2. og 4. klasse.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Det er dog vigtigt at pointere, at antallet af skoler ikke svarer til antallet af observationer. Når skolelederdatasættet og datasættet for lærere og andet pædagogisk personale sammenkobles, har vi 157 skoler, der indeholder minimum én besvarelse fra hvert skema, mens 1.481 besvarelser fra lærere og andet pædagogisk personale kan kobles til en skolelederbesvarelse.

Når skolelederdatasættet sammenkobles med elevdatasættet, er der i alt 155 skoler, hvor vi har en besvarelse fra både skoleleder og minimum én elev, mens 20.763 elevbesvarelser kan kobles til en skoleleder.

Når datasættet for lærere og andet pædagogisk personale sammenkobles med forældredata, er det muligt at sammenkoble besvarelser fra 192 skoler. Her er der dog ikke taget højde for, hvorvidt besvarelserne fra det pædagogiske personale omhandler de klassetrin (0., 2. og 4.), som forældrebesvarelserne omhandler.³

SAMMENKOBLING AF ELEVBESVARELSER OG BESVARELSER FRA LÆRERE OG ANDET PÆDAGOGISK PERSONALE

Særligt for folkeskolereformpanelet er muligheden for at koble elevbesvarelser med besvarelserne fra deres egne lærere og andet tilknyttet pædagogisk personale. Ønskes denne sammenkobling, må man først tage stilling til, hvor præcist dette match behøver at være. På baggrund af 2014-besvarelserne viser vi i dette afsnit, hvordan denne sammenkobling kan ske (jf. figur 3.2). Det første, der er vigtigt at have for øje, er, at besvarelser fra lærere og andet pædagogisk personale rammer bredere end blot de klassetrin, hvor vi har elevbesvarelser. Eksempelvis er der indsamlet besvarelser fra undervisere på 0. og 2. klassetrin, og disse besva-

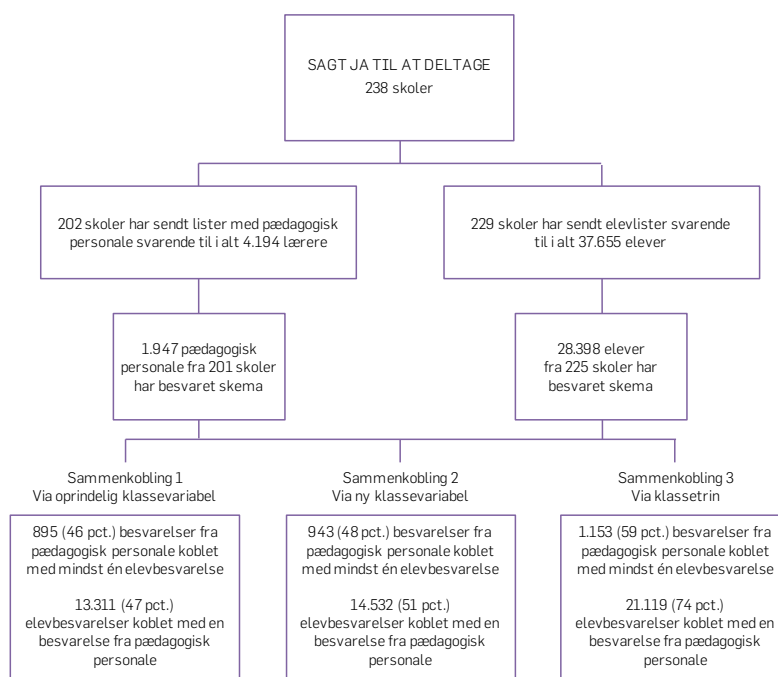
3. Hvis vi forsøger at tage højde for, at et medlem fra det pædagogiske personale skal være tilknyttet enten 0., 2. eller 4. klasse, kan vi sammenkoble 169 skoler.

relser udgår naturligvis i sammenkobling med eleverne.⁴ Manglende besvarelser fra enten eleverne eller det pædagogiske personale har samtidig en langt større påvirkning, når to typer af skemaer skal kobles sammen.

Den største udfordring med hensyn til at sammenkoble de to typer af besvarelser er dog, at der er en vis usikkerhed forbundet med klassevariablen, som gør, at vi ikke er i stand til at lave et præcist match mellem eleverne og deres undervisere. Usikkerheden skyldes ikke manglende oplysninger i klassevariablen, men i højere grad, at der anvendes forskellige betegnelser hos hhv. STIL og respondenterne.

FIGUR 3.2

Sammenkobling af elevdata og datasættet for lærere og andet pædagogisk personale i 2014.



Anm.: Den procentuelle besvarelse er udregnet på baggrund af det samlede antal besvarelser fra hhv. personale- og elevskemaerne. Det vil sige, det er andelen af besvarelser, der kan sammenkobles ud af det samlede antal besvarelser fra hvert datasæt. Eksempelvis for eleverne i sammenkobling 1: 13.311 besvarelser ud af 28.398 besvarelser.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

SAMMENKOBLING 1

I den første sammenkobling anvender vi institutionsnummer og klassevariablene i datasættene. For eleverne har vi en klassevariabel tilgængelig fra STIL, og for det pædagogiske personale anvender vi en let oprenset

4. Vi har elevbesvarelser fra hhv. 4., 6., 8. og 9. klasstrin, mens vi har besvarelser fra lærere og andet pædagogisk personale for 0., 2., 4., 6., 8. og 9. klasstrin.

klassevariabel ”klasse” baseret på den pågældendes angivelse i spørgeskemaet.

Af figur 3.2 fremgår det, at 46 pct. af det pædagogiske personale og 47 pct. af eleverne kan kobles sammen, så der både er én besvarelse fra lærere og andet pædagogisk personale og mindst én elevbesvarelse i en klasse. Denne sammenkobling giver den laveste andel elev- og personalebesvarelser, der kan matches.

SAMMENKOBLING 2

For at forbedre kvaliteten af klassevariablen fra sammenkobling 1 giver vi her et eksempel på, hvordan man kan lave en oprensning af klassevariablene for at kunne sammenkoble flere elever til deres undervisere. Det vil sige, at vi laver en let oprensning af klassevariablen i elevskemaet (dvs. den fra de officielle elevlister fra STIL) og den klassevariabel, som det pædagogiske personale selv angiver i egne spørgeskemabesvarelser.

Helt konkret sammenholder vi de to klassevariabler, og efterfølgende retter vi de steder, hvor der er åbenlyse fejl i klassevariablen for alle elever og det pædagogiske personale, jf. eksempler angivet i næste afsnit. Den form for oprensning giver selvfølgelig en ny usikkerhed, da vi retter efter bedste overbevisning og derfor også kan have matchet en lærer til en klasse, som underviseren ikke underviser. Ikke desto mindre medfører vores oprensning, at vi kan sammenkoble 1.221 flere elever end ved sammenkobling 1.

Mange af rettelserne består i at ensrette de angivne klasser inden for hver skole. Det vil sige, at vi har sammenholdt alle klasser angivet af henholdsvis pædagogisk personale og elever inden for hver skole og rettet der, hvor der synes at være åbenlyse fejl. Det er eksempelvis de steder, hvor der står 04A frem for 4A. Et andet eksempel er, at der for eleverne er angivet tre 6. klasser: 6U, 6V og 6X, mens der for lærerne er angivet: 6., 6U og 6X. I sådanne tilfælde har vi rettet 6. til 6V for lærerne.

Den rensede klassevariabel på baggrund af ovenstående valg indgår ikke i de data, der er tilgængelige.

SAMMENKOBLING 3

Et alternativ til at sammenkoble elevbesvarelser og besvarelserne fra underviseren på den enkelte klasse (en specifik klasse som fx 4A) er at sammenkoble til klassetrin (dvs. 4. klassetrin og ikke enkelte klasser). Det vil sige, at hvis fire undervisere har besvaret, at de underviser på 6. klassetrin, så kobles alle fire til alle elever, der går på 6. klassetrin på en given skole. Denne sammenkobling er derfor mindre præcis end de to ovenstående, men som det fremgår af figur 3.2, bliver langt flere besvarelser også matchet (59 pct. af det pædagogiske personale og 74 pct. af eleverne).

BESVARELSER FRA LÆRERE OG ANDET PÆDAGOGISK PERSONALE
FORDELT PÅ KLASSETRIN

I spørgeskemaet til lærerne indgår der både spørgsmål til undervisningen generelt og specifikke spørgsmål til dansk- eller matematikundervisningen. For de specifikke spørgsmål om dansk og matematik viser tabel 3.3, hvilke klassetrin, lærerne har foretaget deres besvarelser ud fra.

TABEL 3.3

Lærerbesvarelser om dansk og matematik fordelt på klassetrin. 1. dataindsamling 2014.

Klassetrin	Besvarelser ud fra klassetrin	Andel i pct.
2. klasse	320	20
4. klasse	358	22
6. klasse	324	20
8. klasse	269	17
9. klasse	310	20
Andet/irrelevant	21	1
I alt besvarelser	1.602	100

Anm.: Respondenterne er blevet bedt vælge den 2., 4., 6., 8. eller 9. klasse, som han eller hun har den første dansk-/matematik lektion med efter weekenden. For besvarelser fra andet pædagogisk personale indgår der ikke samme specifikke spørgsmål om undervisningen, hvorfor (primært) pædagogernes svar ikke indgår her i. Ej heller for 0. klasse. Kategorien "Andet/irrelevant" dækker fx over indskolingen eller en specialklasse.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Af de 1.602 lærere, der har besvaret spørgeskemaet, er der en nogenlunde lige fordeling af besvarelser på de fem udvalgte klassetrin. Eksempelvis har 17 pct. besvaret skemaet for elever på 8. klassetrin, mens 22 pct. har besvaret skemaet for 4. klasse. Der er derfor ikke tale om en overvægt af besvarelser for fx 2. klasse frem for 8. klasse.

En mindre andel har svaret ud fra et andet klassetrin. Det kan eksempelvis være en specialklasse eller samlet for indskolingen. Tabel 3.3 gælder imidlertid kun for lærerne, da der (primært) for pædagogerne ikke spørges ind til undervisningen for en specifik klasse.

I praksis er både lærere og det andet pædagogiske personale tilknyttet flere klassetrin. Tabel 3.4 viser, hvilke klassetrin, de deltagende lærere og andet pædagogisk personale er tilknyttet. Da en underviser kan være tilknyttet mere end én klasse, så er antallet af observationer større end antallet af personer, der i alt har besvaret skemaet.

TABEL 3.4

Klassetrin, som lærerne eller det pædagogiske personale er tilknyttet. 1. dataindsamling, 2014.

Klassetrin	Har markeret	Andel i pct.
0. klasse	183	4
1. klasse	349	8
2. klasse	440	11
3. klasse	325	8
4. klasse	494	12
5. klasse	393	10
6. klasse	476	12
7. klasse	415	10
8. klasse	481	12
9. klasse	522	13
10. klasse	29	1
Antal observationer	4.107	100

Anm.: Tabellen viser antallet af lærere, børnehaveklasselærere og pædagogisk personale, der har tilkendegivet, at de er tilknyttet et bestemt klassetrin. Tabellen er dannet ud fra spørgsmålene "På hvilke klassetrin underviser du i dansk?", "På hvilke klassetrin underviser du i matematik?" og "Hvilke klassetrin er du tilknyttet?". Respondenterne har haft mulighed for at sætte flere kryds.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Af tabellen fremgår det, at de deltagende lærere og andet pædagogisk personale er tilknyttet de fleste klassetrin, derfor synes deres besvarelser (om undervisningen generelt) overordnet set at være repræsentativ for undervisningen på alle klassetrin.

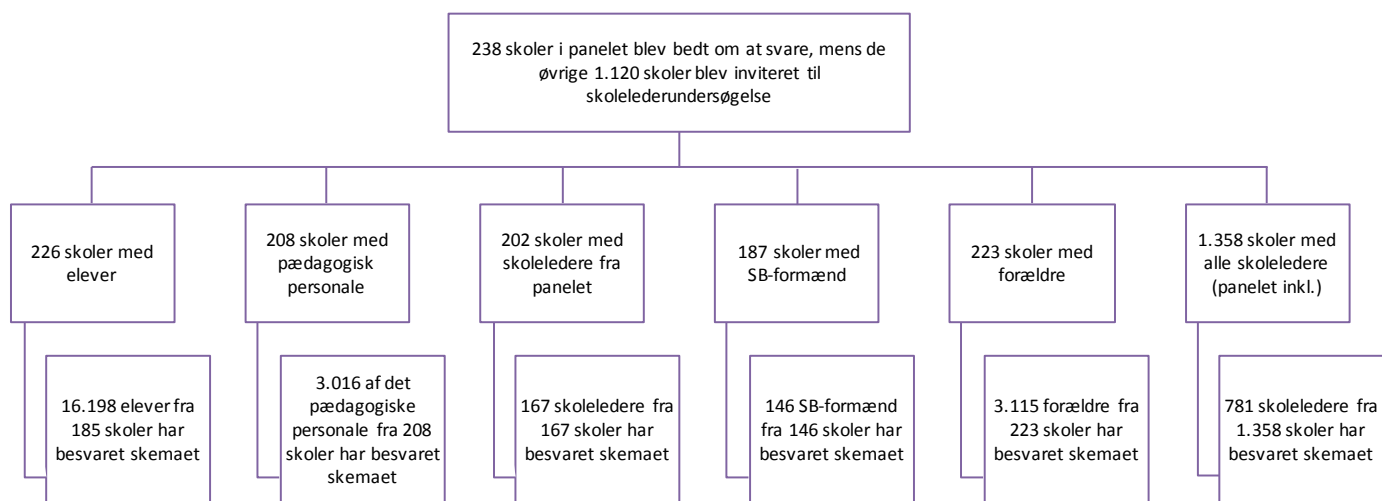
I de tilfælde, hvor underviseren har besvaret skemaet, men ikke er tilknyttet et af de udvalgte klassetrin, ekskluderes denne automatisk, når data sammenkobles med elevdatasættene, men i mange andre henseender, hvor klassevariablen ikke er relevant, kan han/hun indgå på lige fod med de øvrige besvarelser. I alt har 77 pct. af lærerne og andet pædagogisk personale markeret, at de er tilknyttet mindst ét af de udvalgte klassetrin.

DELTAGELSE OG BESVARELSESPROCENTER FOR 2015

I foråret 2015 gennemførte vi 2. dataindsamling, hvilket som nævnt tidligere er den første dataindsamling efter folkeskolereformens ikrafttræden. Figur 3.3 viser, hvordan besvarelserne fordeler sig blandt de forskellige typer respondenter. Udgangspunktet for 2015-dataindsamlingen er de skoler, der accepterede at deltage i undersøgelsen i 2014. Det vil sige, at de 238 skoler, der enten indgav lister i 2014 eller deltog i 2014-undersøgelsen, er inviteret til at deltage i hele undersøgelsen i 2015. Skolelederdelen blev endvidere udvidet til at indbefatte alle folkeskoler i Danmark i 2015, hvorfor bruttostikprøven for skolelederdelen udgør 1.358 skoler.

FIGUR 3.3

Antal deltagende skoler og respondenter i 2. dataindsamling, 2015.



Anm.: Figuren er dannet ud fra "XX_besvar" og INST_NUMMER-variablene i de rensede datasæt og angiver antallet af observationer i bruttolisten samt antallet af respondenter fra skolerne i de forskellige datasæt. Kasserne yderst til højre indeholder informationer om alle skoleledere. Det vil sige, både de skoleledere, der hele tiden har været en del af undersøgelsen, og skolelederne på de skoler, der først blev inviteret i 2015. Der er 72 elever ud af de 16.198, der besvarede spørgeskemaet, der har ukendt institutionsnummer. Det samme gælder for 21 besvarelser fra det pædagogiske personale. I de rensede datasæt er disse observationer kodet med INST_NUMMER=999999. Ydermere har vi i elevbesvarelserne manglende informationer om 71 elevers cpr-numre. Disse indgår stadig i data som besvarelser.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Figur 3.3 viser også, at ud af de 238 inviterede deltog eksempelvis 226 skoler i elevundersøgelsen, mens 208 skoler deltog i undersøgelsen med lærere og andet pædagogisk personale. Alle skoler, der indleverede respondentlister for lærere og andet pædagogisk personale, er repræsenteret med minimum en besvarelse fra lærere eller andet pædagogisk personale.

I alt besvarede 16.198 elever spørgeskemaet i 2. indsamling af data ud af 27.489 mulige. Elevbesvarelserne er fordelt på 185 skoler ud af 226 mulige.

I forlængelse af figur 3.3 angiver tabel 3.5 antallet af respondenter i bruttostikprøven, antal besvarelser og svarprocenter. Det fremgår, at de højeste svarprocenter opnås for skoleledere, der tidligere har deltaget i panelet (83 pct.), SB-formændene (78 pct.), forældrene (75 pct.) og det pædagogiske personale (75 pct.), mens de laveste svarprocenter er opnået for eleverne (59 pct.) og i den landsdækkende skolelederundersøgelse (58 pct.).

TABEL 3.5

Antal skoler og respondenter i bruttolisten samt antal besvarelser (i pct.). 2. dataindsamling, 2015.

Respondenttype	Antal skoler (brutto)	Antal respondenter (brutto)	Antal besvarelser (i pct.)
Elever	226	27.489	16.198 (59 pct.)
Lærere og andet pædagogisk personale	208	4.038	3.016 (75 pct.)
Skoleledere i panelet	202	202	167 (83 pct.)
Alle skoleledere	1.358	1.358	781 (58 pct.)
SB-formænd	187	187	146 (78 pct.)
Forældre	223	4.166	3.115 (75 pct.)
I alt ¹	233	37.238	23.256 (62 pct.)

Anm.: For hver respondentgruppe angiver tabellen antallet af skoler og respondenter, der har modtaget spørgeskemaer, og antal respondenter, der har svaret, samt svarprocenten i parentes. Ikke alle respondenter i bruttolisten har et tilgængeligt institutionsnummer. Data indeholder 72 elever, hvor vi ikke kan identificere, hvilken skole, de går på. Det samme er gældende for 21 besvarelser fra det pædagogiske personale. Disse observationer indgår dog stadig i datasættene, hvor deres institutionsnummer er kodet til "999999".

1. Bemærk, at summen i antal respondenter og antal besvarelser er fratrukket antallet af skoleledere i panelet, da skoleledere i panelet også indgår i kolonnen for alle skoleledere.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

SAMMENKOBLING AF DE ENKELTE SKEMAER I 2015

For at kunne udnytte data fra 2015 og oprette et fyldestgørende panel på tværs af årene er det relevant at vide, i hvor høj grad man har flere forskellige typer af besvarelser fra den enkelte skole. Derfor har vi beskrevet en række relevante sammenkoblinger mellem de forskellige typer af data (for 2015). Grunddatasættet "besvarelser2015" giver endvidere et overblik over, hvilke skoler der har deltaget med hvilke besvarelser.

På samme vis som ved gennemgangen af sammenkoblingerne for 2014-data, illustrerer vi med tabel 3.6 antallet af skoler, der har bidraget med minimum én besvarelse i de to skemaer. Her fremgår det eksempelvis,

at vi har modtaget mindst én elevbesvarelse og mindst én besvarelse for det pædagogiske personale fra 181 af de 238 deltagende skoler.

TABEL 3.6

Antal skoler, der går igen på tværs af respondenttype. 2. dataindsamling, 2015.

	Elever	Pædagogisk personale	Skoleledere	SB-formænd	Forældre
Elever		181	150	134	181
Lærere og andet pædagogisk personale	181		164	143	204
Skoleledere	150	164		119	160
SB-formænd	134	143	119		140
Forældre	181	204	160	140	

Anm.: Tabellerne angiver antallet af skoler, der har bidraget med minimum én besvarelse i hver af de to respondentgrupper, der sammenlignes. Eksempelvis har 180 skoler afgivet minimum én besvarelse for det pædagogiske personale og minimum én besvarelse fra en elev. NB: Skoleledere refererer her til skoleledere i det oprindelige panel og ikke den udvidede undersøgelse af skolelederne.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Med udgangspunkt i tabel 3.6 kommenterer vi i det følgende på interessante koblinger mellem datasættene.

Når skolelederdatasættet og datasættet for lærere og andet pædagogisk personale kobles sammen, har vi 164 skoler, der indeholder minimum én besvarelse fra hver respondentgruppe. Det betyder samtidig, at vi har 2.569 besvarelser fra lærere og andet pædagogisk personale, som kan kobles til skolelederbesvarelserne.

Når skolelederdatasættet sammenkobles med elevdatasættet, er der i alt 150 skoler, hvor vi har en spørgeskemabesvarelse fra både skolelederen og minimum én elev. I alt betyder det også, at vi har 13.853 elevbesvarelser, som kan kobles til skolelederne.

Når datasættet for det pædagogiske personale kobles med besvarelserne fra forældrene, er det muligt at sammenkoble besvarelser fra 204 skoler. Her er der dog ikke taget højde for, hvorvidt besvarelserne fra det pædagogiske personale omhandler de klassetrin (1., 3. og 5.), som forældrebesvarelserne omhandler.⁵

Når vi i det hele taget vælger at sammenkoble de forskellige datasæt på tværs, får vi mulighed for at danne os et overblik over, hvordan panelet kommer til at se ud, og i hvilket omfang, vi får mulighed for at udtale os om folkeskolereformens betydning inden for de enkelte skoler. Jo bedre match vi har mellem de forskellige grupper af respondenter på den enkelte skole, jo bedre bliver vi i stand til at undersøge reformens indvirkning på sammenspillet mellem disse grupper. Som eksempel vælger vi derfor at uddybe, hvordan den bedst mulige sammenkobling mellem elever og lærere og andet pædagogisk personale opnås.

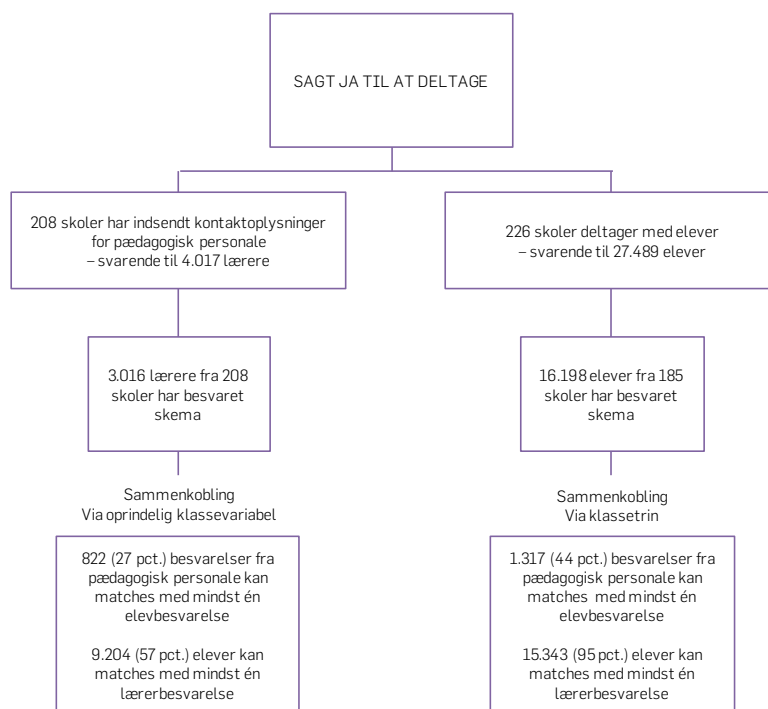
5. Hvis vi tager højde for, at et medlem af det pædagogiske personale skal være tilknyttet et af de klassetrin, som forældreskemaet omhandler, bliver vi i stand til at matche 137 skoler.

SAMMENKOBLING VIA KLASSEVARIABEL

Ligesom i data fra 2014 er der i 2015 også en usikkerhed forbundet med klassevariablen, og vi har derfor foretaget to eksempler på sammenkobling af dataene.⁶ I den første sammenkobling retter vi ikke manuelt i nogen variabler. Det vil sige, at vi sammenkobler på institutionsnummer og klassevariablene, der er angivet i datasættene. For eleverne anvender vi klassevariablen tilgængelig fra STIL, og for det pædagogiske personale danner vi en ny klassevariabel baseret på personalets egne spørgeskema-besvarelser. Hvordan vi danner denne variabel, forklarer vi nærmere i bilaget under variabelforklaringer.

FIGUR 3.4

Sammenkobling af elevdata og datasættet for lærere og andet pædagogisk personale. 2. dataindsamling, 2015.



Anm.: Figuren angiver både besvarelser fra det pædagogiske personale og elevbesvarelser og undersøger, hvordan det bedste match mellem disse grupper opnås. Tallene er baseret på bruttostikprøven og antal besvarelser. Der er 72 elever og 21 lærere eller andet pædagogisk personale, der ikke kan matches, da vi for disse observationer ikke har informationer om deres institutionsnummer.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

6. Til forskel for 2014 eksemplificerer vi ikke den oprensning, hvor vi går igennem hver enkel klassebetegnelse og retter til efter bedste overbevisning. Proceduren er meget tidskrævende og beror på en række valg. Valg, som afhænger af, hvorvidt den enkelte forsker vil gå på kompromis med mulig fejlsammenkoblinger mod et større match mellem besvarelser fra eleverne og lærerne og andet pædagogisk personale.

Figur 3.4 viser, at 27 pct. af det pædagogiske personale og 57 pct. af eleverne kan kobles sammen, så der både er en besvarelse fra det pædagogiske personale og mindst én elevbesvarelse i en klasse. Denne sammenkobling er den dårligste ift. andelen af elever og pædagogisk personale, der kan kobles sammen. Dog må den antages at være den mest præcise sammenkobling ift. at koble pædagogisk personale med elever fra en klasse, som de er knyttet til. Det er muligt manuelt at rette klassevariablene til, ligesom vi viste tidligere for dataene fra 2014. Dette vil højst sandsynligt forbedre matchet en smule, som det også fremgik af 2014-dataene (jf. sammenkobling 2, figur 3.2).

SAMMENKOBLING VIA KLASSETRIN

Den anden sammenkobling sker på skole- og klassetrinsniveau. For de enkelte skoler sammenkobles det pædagogiske personale med alle elever på samme klassetrin. Det vil sige, at hvis fire undervisere har besvaret, at de underviser på 7. klassetrin, så kobles alle fire til alle elever, der går på 7. klassetrin på en given skole. Denne sammenkobling er derfor mindre præcis end den ovenstående, men som det fremgår af figuren, bliver langt flere besvarelser matchet (44 pct. af det pædagogiske personale og 95 pct. af eleverne).

Generelt er det dog vigtigt at have for øje, at sammenkoblingen mellem elevbesvarelser og besvarelser fra lærerne og andet pædagogisk personale ikke skyldes en generel lav svarprocent, men i stedet, at en større del af besvarelserne fra det pædagogiske personale ikke kommer fra de samme klassetrin, som vi har elevbesvarelser fra.

Ud af de 3.016 besvarelser fra lærerne og andet pædagogisk personale underviser 1.487 (49 pct.) af dem på 5., 7. eller 9. klassetrin, og det er således kun deres svar, der i udgangspunktet kan kobles med elevbesvarelser, mens besvarelser fra 1., 3. og 5. klassetrin i princippet kan kobles til forældrebesvarelserne.

BESVARELSER FRA LÆRERNE OG ANDET PÆDAGOGISK PERSONALE FORDELT PÅ KLASSETRIN

I 2015 skulle det pædagogiske personale også svare på konkrete spørgsmål om dansk- og matematikundervisningen. Som i 2014 er besvarelser for det pædagogiske personale jævnt fordelt på de udvalgte klassetrin (jf. tabel 3.7). Så vidt muligt skulle de svare ud fra samme klassetrin (så en underviser eksempelvis svarer for 4. klasse i 2014 og 5. klasse i 2015), men har også haft mulighed for at svare for et nyt klassetrin.

TABEL 3.7

Lærerbesvarelser om dansk og matematik fordelt på klassetrin. 2. dataindsamling, 2015.

Klassetrin	Besvarelser ud fra klassetrin	Andel i procent
1. klasse	442	17,19
3. klasse	522	21,05
5. klasse	590	23,79
7. klasse	463	18,67
9. klasse	463	18,67
I alt besvarelser	2.480	100

Anm.: Respondenterne er blevet bedt om at vælge den klasse, som de besvarede spørgeskemaet for sidste år, men kunne også angive en "ny klasse", hvis den klasse, de tænkte på i 2014, er dimmitteret, eller hvis de har en ny dansk- eller matematikunderviser på et af de udvalgte klassetrin.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Det pædagogiske personale er dog tilknyttet flere klassetrin, og tabel 3.8 viser, hvilket klassetrin, de enkelte undervisere er tilknyttet. Ligesom i 2014 er de fleste klassetrin repræsenteret, og vi forventer derfor, at de generelle besvarelser om undervisningen og forhold på skolen afspejler forhold for de fleste klassetrin og ikke kun fx udskolingen.

TABEL 3.8

Klassetrin, som lærerne og andet pædagogisk personale er tilknyttet. 2. dataindsamling, 2015.

Klassetrin	Har markeret	Andel i pct.
0. klasse	279	4
1. klasse	985	16
2. klasse	610	10
3. klasse	909	14
4. klasse	446	7
5. klasse	808	13
6. klasse	407	6
7. klasse	705	11
8. klasse	474	7
9. klasse	700	11
Andet ¹	16	0
Antal observationer	6.339	100

Anm.: Tabellen viser antallet af lærere, børnehaveklasselærere og pædagogisk personale, der har tilkendegivet, at de er tilknyttet et bestemt klassetrin. Tabellen er dannet ud fra følgende spørgsmål til lærerne: "Hvilke klassetrin er du tilknyttet i dansk/ matematik" samt spørgsmålet "hvilke klassetrin er du tilknyttet" for det resterende pædagogiske personale. Respondenterne har haft mulighed for at sætte flere kryds.

1 Andet angiver fx specialklasser eller 10. klasse.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

I forbindelse med at besvarelserne fra lærerne og andet pædagogisk personale kobles sammen med elevbesvarelserne, ekskluderes de besvarelser, der ikke er tilknyttet de udvalgte klassetrin. I alt har 96,22 pct. af lærerne og andet pædagogisk personale dog markeret, at de er tilknyttet mindst ét af de udvalgte klassetrin, hvilket er en del flere end i 2014.

For 2015 er der derfor både en langt bedre svarprocent blandt lærerne og andet pædagogisk personale samt et bedre match med de udvalgte klassetrin end i 2014.

PANELDATA – SAMMENKOBLING AF DATA PÅ TVÆRS AF TID

I dette kapitel ser vi på sammenkoblinger af data over tid. Det vil sige, at data fra 2014 kobles sammen med data fra 2015. Først viser vi, hvordan data kan sammenkobles på skoleniveau via institutionsnummer, og dernæst gennemgår vi forældredata og elevdata, hvordan data kan kobles på individniveau ved brug af cpr-nummer.

SAMMENKOBLING AF PANELDATA PÅ SKOLENIVEAU

Vi sammenkobler datasættene fra 2014 og 2015 via skolens institutionsnummer for at oprette et samlet datasæt med panelstruktur. I forbindelse med opsætningen af panelet laver vi en opgørelse af antallet af skoler i hhv. 2014, 2015 og de skoler, der både indgår med besvarelser i 2014 og 2015. Det vil sige de skoler, der indgår i panel-datasættet. Af tabel 4.1 fremgår panelets størrelse opgjort på skoleniveau.

TABEL 4.1

Skoler, der har deltaget med besvarelser i 2014 og 2015 samt i både 2014 og 2015.

	Skoler i 2014	Skoler i 2015	Antal skoler i 2014 og 2015
Samlet antal skoler ¹	236	233	232
Skoleledere i panel	159	167	130
Lærere og andet pædagogisk personale	201	208	186
SB-formænd	105	146	77
Forældre	225	223	223
Elever ²	225	185	182

Anm.: Tabellen angiver skoler i de forskellige spørgeskemaer for hhv. 2014 og 2015. Derudover viser kolonnen længst til højre, hvor mange skoler der indgår med mindst en besvarelse i 2014 og 2015. For skolelederne tages kun udgangspunkt i de skoler, der også var inviteret til at deltage i 2014.

- Den oprindelige stikprøve består af alle skoler, der har afgivet en besvarelse for mindst ét skema.
- Eleverne kan indgå i panelet ved enten at blive sammenkoblet gennem skolens institutionsnummer eller gennem elevens eget cpr-nummer. For sammenkobling af elevbesvarelser via deres cpr-nummer, se næste afsnit.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Eksempelvis modtog vi besvarelser fra lærere og andet pædagogisk personale fra 201 skoler i 2014 og fra 208 skoler i 2015. Sammenlagt har 186 skoler bidraget med besvarelser fra lærere og andet pædagogisk personale både i 2014 og 2015.

TABEL 4.2

Antal besvarelser særskilt for skoler i 2014, 2015 og skoler, der både er med i 2014 og 2015.

Respondenttype	Antal besvarelser i 2014	Antal besvarelser i 2015	Antal besvarelser på panelskoler i 2014	Antal besvarelser på panelskoler i 2015
Pæd. personale	1.947	3.016	1.787	2.742
Forældre	3.153	3.115	3.110	3.115
Elever ¹	28.398	16.202	23.381	15.910

Anm.: Tabellen angiver antal besvarelser i de forskellige spørgeskemaer for hhv. 2014 og 2015. Sammenkoblingen af data fra 2014 og 2015 er dannet på skoleniveau. Det vil sige, at det er antallet af individer, der er en del af stikprøven på en skole, der har besvaret med den givne respondenttype i 2014, 2015 eller begge år. Derved angiver de to kolonner til højre ikke antallet af individer, der har svaret i både 2014 og 2015, men det samlede antal af besvarelser i hhv. 2014 og 2015 fra skoler, der i begge år mindst har én besvarelse fra den givne respondenttype. I 2014 og 2015 har hhv. 2 og 72 elever manglende oplysninger om institutionsnummer. Det samme gælder for 21 lærere eller andet pædagogisk personale i 2015.

- Eleverne kan indgå i panelet ved enten at blive sammenkoblet gennem skolens institutionsnummer eller gennem elevens eget cpr-nummer. For sammenkobling af elevbesvarelser via deres cpr-nummer, se næste afsnit.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

For nogle af respondentgrupperne indgår der mere end én besvarelse pr. skole. Tabel 4.2 angiver antal besvarelser for 2014 og 2015 og antallet af besvarelser på skoler, der både deltog med den givne respondenttype i 2014 og 2015. Eksempelvis modtog vi 1.947 besvarelser fra lærere og andet pædagogisk personale i 2014 og 3.016 besvarelser i 2015. Af de 186 skoler, hvor vi har besvarelser fra mindst én lærer eller andet pæda-

gøgisk personale i både 2014 og 2015, indeholder data i 2014 1.787 besvarelser og i 2015 2.742 besvarelser.

SAMMENKOBLING AF DATA PÅ INDIVIDNIVEAU

Den centrale enhed i panelundersøgelsen af folkeskolereformen er den enkelte skole. Ikke desto mindre har man for forældre og elever mulighed for at følge en udvikling i de enkelte personers svar. At analysere udviklingen i de enkelte personers svar er ikke problemfrit, idet det kræver en større omkodning for nogle af respondentgrupperne.

Nemtest er det for eleverne, da de enkelte elevs cpr-numre indgår i de rekvirerede lister fra STIL, mens eksempelvis besvarelser fra forældrene kræver en sammenkobling via barnets cpr-nummer. Der er større udfordringer ved sammenkobling af besvarelser fra lærerne og andet pædagogisk personale, da dette primært skal foregå via personens fødselsdato og skolens institutionsnummer.⁷ I stedet kan man følge besvarelserne for grupper af pædagogisk personale på de enkelte skoler.

Med de rette tilladelser til at have adgang til sådanne følsomme oplysninger og en tilladelse fra MBUL/STIL er det muligt at få adgang til data inklusiv følsomme oplysninger, således at den slags sammenkoblinger kan foretages.

Vi har lavet en sådan sammenkobling på individniveau for forældredata og elevdata. Det samme er ikke foretaget for det pædagogiske personale.

ELEVDATA

Tabel 4.3 angiver antallet af respondenter sammenkoblet med elevernes cpr-numre.⁸ Sammenkobler vi elevdataene for 2014 og 2015 på individniveau, så får vi et samlet panel på 12.674 elever. Det vil sige, at vi for 12.674 elever har mulighed for at følge ændringer i deres besvarelser fra 2014 til 2015, men til gengæld får vi på denne måde et mindre panel ift. sammenkoblingen på skoleniveau (jf. tabel 4.2). Skolepanelet tillader os i stedet at se på udviklingen inden for elevbesætningerne på de 130 skoler, der har deltaget med besvarelser i begge år.

7. For 2015 har vi endvidere navn på personalet, men ikke for 2014. Navnene har vi fået fra skolerens respondentlister, og de er anvendt, når vi har inviteret den enkelte person til at deltage i undersøgelsen.

8. For data på forældrene anvender vi også elevernes cpr-numre til sammenkoblingen.

TABEL 4.3

Antal besvarelser særskilt for respondenter i 2014, 2015 og respondenter, der både er med i 2014 og 2015.

Respondenttype	Antal besvarelser i 2014	Antal besvarelser i 2015	Antal respondenter, der har besvaret både i 2014 og 2015
Forældre	3.153	3.115	2.145
Elever	28.398	16.202	12.674

Anm.: Tabellen angiver respondenter i de forskellige spørgeskemaer for hhv. 2014 og 2015. Sammenkoblingen af data fra 2014 og 2015 er dannet på individniveau. Det vil sige, at det er antallet af individer, der er en del af stikprøven i 2014, 2015 eller begge år. Derved viser kolonnen til højre antallet af individer, der har svaret i både 2014 og 2015. I 2014 og 2015 har hhv. 2 og 71 elever manglende oplysninger om cpr-nummer. Derfor kan disse ikke sammenkobles mellem 2014 og 2015.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Der er altså et trade-off mellem et mere præcist panel (ved en sammenkobling over cpr-nummer) eller et større panel (ved sammenkobling over institutionsnummer). Den mest relevante sammenkobling vil selvfølgelig også afhænge af, om man vil følge eleverne på den pågældende skole, uanset om det er de samme elever, eller man vil følge den enkelte elev, uanset hvilken skole, eleven går på.

BORTFALDS- OG REPRÆSENTATIVITETSANALY SE

I dette kapitel præsenteres en bortfalds- og repræsentativitetsanalyse for dataindsamlingen for folkeskolereformen. Målet er at give et fundament til at vurdere, hvorvidt dataindsamlingen er repræsentativ for alle landets skoler. Endvidere giver analysen et godt overblik over data og er et værktøj til at vurdere forskelle i ekstern validitet på tværs af grupper.

Analyserne foretages for hver af de seks respondentgrupper og for både 2014- og 2015-indsamlingerne. Den grundlæggende analyseenhed i analyserne er skolen og ikke de enkelte respondenter. I analysen af lærerne er det således ikke de enkelte lærere, men de skoler, som lærerne er ansat på, som er analyseenheden. Bortfalds- og repræsentativitetsanalysen går derfor ud på at vurdere, om de skoler, som de deltagende lærere repræsenterer, kan siges at ligne skoler generelt (på landsplan), eller om bortfaldet er skævt, således at man i stedet primært har svar fra lærere på fx store skoler.

DATAGRUNDLAGET

I 2014 blev 442 tilfældigt udvalgte skoler valgt til at deltage i undersøgelsen, hvoraf 238 skoler sagde ja. Udgangspunktet for 2015-indsamlingen er de 238 skoler, der accepterede at deltage i undersøgelsen i 2014, men det var dog ikke alle, der besvarede spørgeskemaerne. Derudover blev der i 2015 udsendt en udvidet spørgeskema til skoleledere på samtlige folkeskoler i Danmark.

Dataindsamlingen er, som tidligere beskrevet, opdelt på 6 respondentgrupper. Det drejer sig om elever, forældre, skoleledere, lærere og andet pædagogisk personale samt skolebestyrelsesformænd (SB-

formænd). Dertil kommer skoleledere, som deltager i den udvidede skolelederundersøgelse i 2015.

Tabel 5.1 viser, hvordan skolerne i hver af de seks respondentgrupper fordeler sig på henholdsvis bruttoliste og besvarelse. Bruttolisten omfatter alle de skoler, som har sagt ja til at deltage i undersøgelsen, mens ”Besvarelse” omfatter de skoler, der faktisk svarede. At en skole har deltaget med besvarelse betyder, at de som minimum har deltaget med en besvarelse fra den pågældende respondentgruppe. Grunden til, at bruttolisten ikke tæller 238 skoler for alle respondentgrupper i 2014, skyldes, at skolerne i nogle tilfælde ikke har fremsendt kontaktoplysninger for alle respondentgrupper.

TABEL 5.1

Antal inviterede og deltagende skoler i 2014 og 2015, dataindsamlingen vedrørende folkeskolereformen.

	Elever	Forældre	Pædagogisk personale	SB-formænd	Skoleledere	Skoleledere (udvidet)
	442 tilfældigt udvalgt til at deltage, 238 har accepteret					
<i>2014</i>						
Brutto	229	225	202	196	193	-
Besvarelse	225	225	201	105	159	-
<i>2015</i>						
Brutto	230	223	208	187	202	1.358
Besvarelse	189	223	208	146	167	781

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

I bortfalds- og repræsentativitetsanalysen baserer beskrivelserne af skolekarakteristika og elevfordelinger sig på registerdata fra Danmarks Statistik. Registerdataene er fra år 2014 og stammer dels fra det Integreerede Elevregister (INTE), dels fra registre vedrørende demografi, indkomst og arbejdsmarkedstilknytning. De 442 skoler kobles til INTE ved hjælp af skolernes institutionskode.

I forbindelsen med koblingen af de 442 skoler til registerdata var der 15 skoler, som ikke kunne genfindes i INTE og derfor ikke kunne kobles med registerdata. Af de 15 skoler var der 9, som havde givet tilsagn om deltagelse i 2014, og 11 havde givet tilsagn i 2015⁹. Skoler der ikke har kunnet genfindes i INTE er udeladt af bortfalds- og repræsentativitetsanalysen. Dette forventes ikke at påvirke de overordnede konklusioner for bortfalds- og repræsentativitetsanalysen.

9. De 15 skoler/institutionsnumre har alle tilknyttede afdelinger, og alle disse afdelinger kan genfindes i INTE. Dertil kommer, at de manglende institutionsnumre har et afdelingsnummer, der er sammenfaldende med hovedinstitutionsnummeret. Det tyder på, at der på nogle skoler er institutionsnumre, som udelukkende har ledelse tilknyttet, mens eleverne er tilknyttet de enkelte afdelinger. Det er en nærliggende forklaring på, at disse institutionsnumre ikke kan genfindes i det integrerede elevregister, når vi matcher på afdelings institutionsnummer.

TABEL 5.2

Antal inviterede og deltagende skoler i 2014 og 2015, som kan genfindes i det integrerede elevregister, dataindsamlingen vedrørende folkeskolereformen.

	Elever	Forældre	Pædagogisk personale	SB-formænd	Skoleledere	Skoleledere (udvidet)
	442 tilfældigt udvalgt til at deltage, 238 har accepteret					
<i>2014</i>						
Brutto	228	225	195	187	188	-
Besvarelse	224	225	195	100	155	-
<i>2015</i>						
Brutto	226	223	206	181	196	1.358
Besvarelse	185	223	206	143	162	745

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

BORTFALDS- OG REPRÆSENTATIVITETSANALYSE

Bortfalds- og repræsentativitetsanalyserne foretages ved at sammenligne deltagende skoler på tværs af de 6 respondentgrupper (jf. tabel 5.2).

BORTFALDSANALYSE

I bortfaldsanalysen sammenlignes de skoler, der har besvaret skemaerne, med de skoler, der havde sagt ja til at deltage (bruttolisten). Fokus er her på, om frafaldet er så skævt, at de skoler, der har besvaret skemaerne, adskiller sig signifikant fra de skoler, der oprindeligt indvilgede i at deltage. Eksempelvis sammenlignes de 196 skoler, der indgår i bruttolisten for SB-formænd, med de 105 deltagende skoler for SB-formænd.

I analysen anvendes statistiske test for at vurdere, om forskelle i skolekarakteristika og elevfordelinger er signifikante eller blot kan tilskrives statistiske tilfældigheder. Vi anvender chi²-test til at teste fordelinger og t-test til at teste forskelle for numeriske variable. For samtlige test angiver vi både, om forskellene er statistisk signifikante, og på hvilket signifikansniveau. Der anvendes et 5-procents-signifikansniveau i testene. Det vil sige, hvis p-værdien er mindre end 0,05, anses sammenhængen som tilstrækkelig sikker til, at det ikke skyldes et tilfælde.

REPRÆSENTATIVITETSANALYSE

I repræsentativitetsanalysen sammenlignes de skoler, der har besvaret skemaerne, med alle folkeskoler i Danmark. Formålet er at undersøge, om de skoler, der har besvarelser for hver af respondentgrupperne, adskiller sig signifikant fra den samlede population af folkeskoler i Danmark – altså om stikprøven af skoler er repræsentativt. Eksempelvis sammenlignes de 105 deltagende skoler med besvarelser fra SB-formænd med alle folkeskoler i Danmark.

Analyserne foretages særskilt for 2014 og 2015. For analysen af den udvidede skolelederundersøgelse i 2015 er bortfaldsanalysen og repræsentativitetsanalysen dog sammenfaldende, da alle folkeskoler blev inviteret til at deltage i undersøgelsen.

BAGGRUNDSKARAKTERISTIKA

For både bortfaldsanalysen og repræsentativitetsanalysen anvendes Danmarks Statistiks registeroplysninger til at karakterisere de deltagende skoler vha. en række skolekarakteristika og elevfordelinger.

Vi anvender følgende skolekarakteristika:

- Karaktergennemsnit fra folkeskolens afgangsprøve
- Skolestørrelse, målt ved antal elever
- Geografisk beliggenhed (regioner).

Vi anvender følgende elevfordelinger (på skoleniveau):

- Køn
- Etnicitet (dansk vs. anden etnisk herkomst)
- Familietype – bor med begge forældre
- Forældres højest fuldførte uddannelse
- Forældres indkomst (placering i indkomstkvarterer)
- Forældres tilknytning til arbejdsmarkedet (i beskæftigelse).

Oplysningerne om forældrene indgår særskilt for mor og far.

BORTFALD

Af tabel 5.3 fremgår det, at der er meget stor forskel på bortfaldet mellem de forskellige respondentgrupper. I 2014 er der stort set intet bortfald blandt skolerne i elev-, forældre- og lærerrespondentgrupperne, mens 47 pct. af skoler i SB-formændsrespondentgruppen og 18 pct. af skolerne i skoleledergruppen faldt fra. I 2015 er mønstret næsten det samme med den forskel, at der nu også er bortfald fra elevrespondentgruppen, hvor 18 pct. af skolerne falder fra. Til gengæld er der i 2015 et markant mindre frafald fra respondentgruppen af SB-formænd, nemlig 21 pct.

Bortfaldet i respondentgrupperne er ikke isoleret set et problem så længe, der ikke er en systematisk sammenhæng mellem frafald og skolekarakteristika. Hvis der er systematik i frafaldet, således at fx store skoler i højere grad falder fra end små skoler, så kan det betyde, at stikprøven af skoler skævvrides så meget, at det går ud over stikprøvens repræsentativitet.

TABEL 5.3

Bortfald fordelt på respondentgruppe og år, antal og procent (skoler).

	Brutto	Bortfald	Pct.
<i>2014</i>			
Elev	228	4	2
Forældre	225	0	0
Pædagogisk personale	195	0	0
SB-formænd	187	87	47
Leder	188	33	18
<i>2015</i>			
Elev	226	41	18
Forældre	223	0	0
Pædagogisk personale	206	0	0
SB-formænd	181	38	21
Ledere	196	34	17
Ledere, udvidet skema	745	0	0

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

For at undersøge, om der er en systematisk sammenhæng mellem skolekarakteristika og frafald, har vi testet, om fordelingen af skolekarakteristika og elevfordelinger er signifikant forskellig for de skoler, der har besvaret spørgeskemaerne sammenlignet med bruttostikprøven. Resultatet for 2014 er vist i tabel 5.4, og resultaterne for 2015 er vist i tabel 5.5. I tabellerne er der kun medtaget de respondentgrupper, hvor der har været bortfald – jf. tabel 5.3.

Tabellerne viser entydigt, at det ikke er muligt at påvise nogen signifikante forskelle i skolekarakteristika og elevfordelinger mellem de skoler, der har givet tilsagn om at deltage, og de skoler, hvor den pågældende respondentgruppe har besvaret spørgeskemaerne. Det gælder for alle respondentgrupper. Tabellerne viser elevfordeling og skolekarakteristika for bruttostikprøven (angivet ”brutto” i tabellerne) samt *forskellen* imellem karakteristika ved bruttostikprøven og karakteristika ved de skoler, hvor der er besvarelser fra de respektive respondentgrupper (angivet med ”differens” i tabellerne) som følge af bortfald. For eksempel ses det i tabel 5.4, at den gennemsnitlige andel elever med dansk oprindelse er 92,2 pct. i bruttostikprøven med SB-formænd, men som følge af bortfald er denne andel 0,6 pct. mindre blandt de skoler, der besvarer SB-formændsspørgeskemaet. Der er testet for, om forskellen er signifikant på et 5-procents-signifikansniveau. Der kan ikke påvises signifikante forskelle i elevfordelingerne og i skolekarakteristika i hverken 2014 eller 2015 (jf. tabel 5.5).

TABEL 5.4

Forskelle i elevfordeling og skolekarakteristika som følge af bortfald opdelt på respondentgrupper, 2014, procent.

	Elever		SB-formænd		Skoleledere	
	Brutto	Differens	Brutto	Differens	Brutto	Differens
Elevkarakteristika						
Dansk oprindelse	91,9	0,1	92,2	-0,6	92,4	0,0
Bor med begge forældre	65,4	0,1	66,2	-0,2	66,4	-0,2
Drenge	51,5	-0,0	51,4	0,0	51,4	0,1
Fars karakteristika						
I arbejde	87,2	0,1	87,9	-0,6	88,1	0,1
<i>Højeste uddannelsesniveau</i>						
Grundskole/uoplyst	19,4	-0,1	18,7	0,2	18,6	0,1
Gymnasial/erhvervsfaglig uddannelse	51,7	0,0	51,5	0,3	51,6	0,2
Videregående uddannelse	28,9	0,1	29,8	-0,6	29,8	-0,3
<i>Indkomst</i>						
Indkomst 1. kvartil	26,1	-0,2	25,0	1,2	24,8	-0,1
Indkomst 2. kvartil	28,4	0,0	28,4	0,2	28,7	0,4
Indkomst 3. kvartil	25,4	0,1	25,7	-0,1	25,8	0,2
Indkomst 4. kvartil	20,1	0,1	20,9	-1,3	20,8	-0,5
Ingen oplysninger om far	6,1	-0,0	5,8	0,2	5,8	0,1
Mors karakteristika						
I arbejde	80,1	0,2	80,9	-0,8	81,0	-0,0
<i>Højeste uddannelsesniveau</i>						
Grundskole/uoplyst	15,7	-0,1	15,0	0,7	14,9	0,0
Gymnasial/erhvervsfaglig uddannelse	44,7	0,1	44,5	-0,5	44,6	0,3
Videregående uddannelse	39,6	0,1	40,5	-0,1	40,5	-0,3
<i>Indkomst</i>						
Indkomst 1. kvartil	26,1	-0,2	25,1	0,9	25,0	0,2
Indkomst 2. kvartil	27,2	0,0	26,9	0,2	27,0	0,1
Indkomst 3. kvartil	25,8	0,1	26,2	0,1	26,3	0,0
Indkomst 4. kvartil	21,0	0,1	21,8	-1,2	21,7	-0,3
Ingen oplysninger om mor	3,2	-0,0	3,0	0,2	3,0	-0,0
Skolekarakteristika						
<i>Karakter</i>						
Lavt karakterniveau	25,0	-0,4	25,1	7,1	23,4	0,2
Mellem karakterniveau	25,4	0,4	26,2	-6,8	26,6	0,1
Højt karakterniveau	19,3	-0,3	20,3	-2,7	20,2	-1,7
Missing	30,3	0,4	28,3	2,3	29,8	1,4
<i>Størrelse</i>						
Lille	33,8	0,3	32,6	3,6	33,0	1,4
Mellem	34,6	-0,2	36,4	-2,6	36,7	2,5
Stor	31,6	-0,1	31,0	-1,0	30,3	-3,9
<i>Region</i>						
Region Hovedstaden	24,6	-0,4	24,1	1,1	22,3	1,7
Region Midtjylland	22,4	0,5	24,6	-1,4	25,5	-0,3
Region Nordjylland	11,4	-0,2	13,4	-1,6	13,3	-0,9
Region Sjælland	19,3	0,1	13,4	0,4	14,4	0,8
Region Syddanmark	22,4	0,0	24,6	1,6	24,5	-1,3

Anm.: Forskelle i gennemsnitlig elevfordeling er testet ved brug af t-test. Forskelle i fordeling af skolekarakteristika er testet ved brug af χ^2 -test.

"Brutto" angiver bruttostikprøven af skoler, der har givet tilsagn om at deltage.

"Differens" angiver forskellen i karakteristika mellem skoler i bruttostikprøven og skoler med besvarelser fra den pågældende respondentgruppe.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

TABEL 5.5

Forskelle i elevfordeling og skolekarakteristika som følge af bortfald opdelt på respondentgrupper, 2015, procent

	Elev		SB-formænd		Skoleledere	
	Brutto	Differens	Brutto	Differens	Brutto	Differens
Elevkarakteristika						
Dansk oprindelse	91,8	-0,2	92,3	0,5	92,2	0,2
Bor med begge forældre	65,3	-0,8	65,9	-0,7	66,1	-0,1
Dreng	51,5	-0,1	51,5	-0,0	51,5	-0,1
Fars karakteristika						
I arbejde	87,1	-0,9	87,9	-0,1	88,1	0,1
Højeste uddannelsesniveau						
Grundskole/uoplyst	19,5	0,7	18,9	0,3	18,8	0,0
Gymnasial/erhvervsfaglig uddannelse	51,7	-0,1	51,9	0,2	51,7	0,3
Videregående uddannelse	28,8	-0,6	29,2	-0,5	29,5	-0,3
Indkomst						
Indkomst 1. kvartil	26,2	1,0	25,3	0,2	25,0	0,1
Indkomst 2. kvartil	28,4	-0,2	28,9	0,5	28,6	0,2
Indkomst 3. kvartil	25,3	-0,3	25,8	0,1	25,9	-0,1
Indkomst 4. kvartil	20,0	-0,5	19,9	-0,8	20,5	-0,2
Ingen oplysninger om far	6,1	0,1	6,0	-0,1	5,9	-0,0
Mors karakteristika						
I arbejde	80,1	-0,7	80,7	0,0	80,9	0,2
Højeste uddannelsesniveau						
Grundskole/uoplyst	15,8	0,6	15,2	0,1	15,1	-0,1
Gymnasial/erhvervsfaglig uddannelse	44,7	0,0	44,8	0,5	44,6	0,3
Videregående uddannelse	39,5	-0,7	40,1	-0,5	40,3	-0,2
Indkomst						
Indkomst 1. kvartil	26,2	0,8	25,4	0,3	25,3	0,2
Indkomst 2. kvartil	27,2	0,1	27,4	0,7	27,2	0,0
Indkomst 3. kvartil	25,7	-0,3	26,1	-0,5	26,1	0,0
Indkomst 4. kvartil	21,0	-0,5	21,1	-0,5	21,4	-0,2
Ingen oplysninger om mor	3,2	0,1	3,1	-0,1	3,1	-0,1
Skolekarakteristika						
Karakter						
Lavt karakterniveau	25,2	-0,2	27,1	1,2	25,5	-1,7
Mellem karakterniveau	25,2	0,4	24,9	3,2	25,0	0,3
Højt karakterniveau	19,5	0,6	18,2	-2,7	20,4	0,0
Missing	30,1	-0,7	29,8	-1,6	29,1	1,3
Størrelse						
Lille	33,6	1,2	33,7	0,1	32,1	1,3
Mellem	35,4	-0,8	37,0	2,8	37,2	-1,6
Stor	31,0	-0,4	29,3	-2,9	30,6	0,4
Region						
Region Hovedstaden	24,8	-1,2	22,7	-0,4	23,5	0,6
Region Midtjylland	22,1	-1,7	26,0	2,2	24,5	-2,1
Region Nordjylland	11,9	-0,5	12,2	-1,1	12,8	-0,8
Region Sjælland	19,0	3,4	15,5	1,5	14,8	0,0
Region Syddanmark	22,1	0,0	23,8	-2,1	24,5	2,3

Anm.: Forskelle i gennemsnitlig elevfordeling er testet ved brug af t-test. Forskelle i fordeling af skolekarakteristika er testet ved brug af χ^2 -test.

"Brutto" angiver bruttostikprøven af skoler, der har givet tilsagn om at deltage.

"Differens" angiver forskellen i karakteristika mellem skoler i bruttostikprøven og skoler med besvarelser fra den pågældende respondentgruppe.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

REPRÆSENTATIVITET

Bortfaldsanalysen viser, at der ikke er tegn på, at bortfaldet har medført en skævvridning af spørgeskemapopulationerne for de 6 respondentgrupper. I dette afsnit undersøger vi, om spørgeskemapopulationerne er repræsentative for alle folkeskoler i Danmark. Dette gøres ved at teste, om der er forskel på elevfordelinger og skolekarakteristika for de skoler, der har besvaret spørgeskemaerne, sammenlignet med alle skoler i Danmark.

Tabel 5.6 og tabel 5.7 viser afvigelsen i elevfordeling og skolekarakteristika for de skoler, der har deltaget dataindsamlingen i 2014, sammenlignet med alle skoler i Danmark. I første kolonne kan man se elevfordeling og elevfordelingerne for alle skoler i Danmark, mens man i de øvrige kolonner kan se, hvor meget tallene for de enkelte respondentgrupper adskiller sig fra disse. For eksempel fremgår det, at den gennemsnitlige andel elever med dansk oprindelse er 90,5 pct. for alle skoler i Danmark, mens den gennemsnitlige andel for de skoler, der har besvaret elevspørgeskemaerne, er 1,3 procentpoint lavere.

Det er for alle respondentgrupper testet, om hver enkelt af elevfordelingerne og skolekarakteristikaene afviger signifikant fra alle skoler i Danmark. Forskelle, der er signifikante på et 5-procents-signifikansniveau, er markeret med ”***” i tabellen.

Der er generelt en meget fin overensstemmelse mellem elevfordelinger og skolekarakteristika på de skoler, der har besvaret spørgeskemaerne, sammenlignet med alle skoler i Danmark. Det er kun i SB-formændsgruppen og for lærerne og det pædagogiske personale, at den gennemsnitlige elevfordeling afviger signifikant fra tallene for alle skoler. Blandt de skoler, der er repræsenteret i SB-formændsbesvarelsene, er der gennemsnitlig en lidt større andel fædre i nederste indkomstkvarantil. For alle skoler i Danmark er der gennemsnitlig 26,6 pct. af fædrene, der befinder sig i 3. indkomstkvarantil, mens der blandt skoler med en besvarelse fra det pædagogiske personale, er 29,4 pct. fædre. Andelen af fædre i de øvrige indkomstkvarntiler adskiller sig dog ikke signifikant fra landsgennemsnittet. Derfor vurderer vi, at afvigelsen ikke har nogen væsentlig betydning for repræsentativiteten.

Blandt skoler med en SB-formandsbesvarelse er den gennemsnitlige andel mødre i 3. indkomstkvarantil lidt lavere end for alle skoler. Ydermere er den gennemsnitlige andel mødre i 3. indkomstkvarantil 25,3 pct. på landsplan, mens skoler, der deltager med en besvarelse fra det pædagogiske personale, har en 1,0 procentpoint lavere andel. Igen finder vi dog, at andelen af mødre i de øvrige indkomstkvarntiler ikke adskiller sig signifikant fra gennemsnittet for alle skoler. Der er ligeledes heller ikke andre elevfordelinger eller skolekarakteristika, der afviger signifikant fra tallene for alle skoler. Vi vurderer derfor, at afvigelsen ikke har nævneværdig betydning for repræsentativiteten for respondentgruppen.

Tabel 5.8 og tabel 5.9 viser afvigelsen i elevfordeling og skolekarakteristika for 2015-indsamlingen for de skoler der har deltaget i dataindsamlingen, sammenlignet med alle skoler i Danmark. Som for 2014-indsamlingen er der generelt en høj overensstemmelse mellem elevfordelinger og skolekarakteristika for de skoler, der har besvaret spørgeskemaerne, sammenlignet med tallene på landsplan. Det er kun i SB-formændsgruppen, at vi finder en signifikant afvigelse, hvor andelen af mødre i 3. kvartil af indkomstfordelingen er 1,3 procentpoint lavere end gennemsnittet for alle skoler i Danmark. Vi konkluderer dog, at der ikke er alvorlige repræsentativitetsproblemer for respondentgruppen, da der ikke er andre signifikante forskelle.

Opsummerende kan det på baggrund af denne bortfalds- og repræsentabilitetsanalyse konkluderes, at dataindsamlingen i al sin væsentlighed er repræsentativ for landets folkeskoler.

TABEL 5.6

Elevfordelinger for skoler, der har besvaret spørgeskemaerne sammenlignet med elevfordelingen for hele landet, 2014, procent.

	Alle skoler i DK	Elev	Forældre	Personale	SB-formænd	Skoleledere
Elevkarakteristika						
Dansk oprindelse	90,5	-1,3	-1,4	-1,6	-2,3	-1,8
Bor med begge forældre	65,7	0,4	0,3	-0,4	-0,7	-0,8
Drenge	51,6	0,0	0,0	0,1	0,2	0,3
Fars karakteristika						
Arbejdsmarkedsstatus: I arbejde	86,7	-0,4	-0,4	-1,1	-1,8	-1,2
Højeste uddannelsesniveau						
Grundskole/uoplyst	19,4	-0,1	-0,0	0,7	0,9	1,0
Gymnasial/erhvervsfaglig uddannelse	50,6	-1,1	-0,9	-0,9	-0,6	-0,9
Videregående uddannelse	30,0	1,2	0,9	0,2	-0,4	-0,1
Indkomst						
Indkomst 1. kvartil	26,6	0,3	0,4	1,7	2,8 **	1,7
Indkomst 2. kvartil	28,0	-0,4	-0,4	-0,4	-0,2	-0,2
Indkomst 3. kvartil	25,0	-0,3	-0,3	-0,8	-0,8	-0,6
Indkomst 4. kvartil	20,4	0,4	0,3	-0,4	-1,8	-0,9
Ingen oplysninger om far	6,6	0,4	0,5	0,7	0,9	0,8
Mors karakteristika						
I arbejde	79,8	-0,2	-0,3	-1,2	-2,0	-1,3
Højeste uddannelsesniveau						
Grundskole/uoplyst	16,0	0,2	0,3	0,9	1,6	1,1
Gymnasial/erhvervsfaglig uddannelse	43,6	-1,1	-0,9	-0,9	-1,4	-0,6
Videregående uddannelse	40,4	0,9	0,7	-0,0	-0,2	-0,4
Indkomst						
Indkomst 1. kvartil	26,3	0,1	0,1	1,3	2,1	1,5
Indkomst 2. kvartil	27,1	-0,1	-0,1	0,2	0,4	0,2
Indkomst 3. kvartil	25,3	-0,4	-0,3	-1,0 **	-0,8	-1,0
Indkomst 4. kvartil	21,3	0,4	0,3	-0,5	-1,7	-0,6
Ingen oplysninger om mor	3,7	0,5	0,5	0,6	0,9	0,7

Anm.: ** angiver, at forskellen er signifikant på et 5-procents-signifikansniveau.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

TABEL 5.7

Karakteristika for skoler, der har besvaret spørgeskemaerne sammenlignet med elevfordelingen for hele landet, 2014, procent.

	Alle skoler i DK	Elev	Forældre	Personale	SB-formænd	Skoleledere
<i>Karakter</i>						
Lavt karakterniveau	23,8	-1,6	-1,5	-0,8	5,8	0,6
Mellem karakterniveau	23,8	-1,2	-1,1	-1,8	-9,2	-2,6
Højt karakterniveau	23,7	4,1	4,6	3,2	0,7	1,8
Missing ¹	28,6	-1,3	-2,0	-0,6	2,6	0,3
<i>Størrelse</i>						
Lille	33,3	-0,1	-0,4	1,0	4,3	1,7
Mellem	33,3	-1,5	-1,3	-3,1	-5,7	-0,9
Stor	33,3	1,6	1,8	2,1	1,3	-0,9
<i>Region</i>						
Region Hovedstaden	21,7	-3,3	-3,2	-2,9	-1,3	1,1
Region Midtjylland	26,7	4,8	4,5	2,1	0,7	0,9
Region Nordjylland	13,0	1,4	1,8	0,1	-2,0	-1,2
Region Sjælland	16,1	-3,1	-3,0	1,7	3,1	2,5
Region Syddanmark	22,6	0,2	-0,1	-1,0	-0,4	-3,2

Anm.: ** angiver, at forskellen er signifikant på et 5-procents-signifikansniveau.

1. Årsagen til, at der på nogle skoler ikke er oplyst karakterer fra folkeskolens afgangsprøver, er, at en del skoler kun har elever op til 6 klasse.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

TABEL 5.8

Elevfordelinger for skoler, der har besvaret spørgeskemaerne sammenlignet med elevfordelingen for hele landet, 2015, procent

	Alle skoler i DK	Elev	Forældre	Personale	SB-formænd	Skoleledere	Skoleledere (udvidet)
Elevkarakteristika							
Dansk oprindelse	90,5	-1,6	-1,4	-1,8	-1,3	-1,5	-0,1
Bor med begge forældre	65,7	-0,5	0,3	0,1	-0,9	-0,4	-0,1
Drenge	51,6	-0,0	0,0	0,1	0,1	0,0	0,2
Fars karakteristika							
I arbejde	86,7	-1,3	-0,4	-1,0	-1,2	-1,3	-0,3
<i>Højeste uddannelsesniveau</i>							
Grundskole/uoplyst	19,4	0,6	-0,0	0,1	0,8	0,6	0,4
Gymnasial/erhvervsfaglig uddannelse	50,6	-1,2	-0,9	-1,2	-1,1	-0,8	0,3
Videregående uddannelse	30,0	0,6	0,9	1,1	0,3	0,2	-0,7
<i>Indkomst</i>							
Indkomst 1. kvartil	26,6	1,4	0,3	0,9	1,5	1,7	0,4
Indkomst 2. kvartil	28,0	-0,6	-0,4	-0,7	-0,5	-0,4	0,2
Indkomst 3. kvartil	25,0	-0,7	-0,3	-0,5	-0,7	-1,1	0,0
Indkomst 4. kvartil	20,4	-0,1	0,3	0,3	-0,3	-0,2	-0,6
Ingen oplysninger om far	6,6	0,5	0,4	0,5	0,5	0,6	0,1
Mors karakteristika							
I arbejde	79,8	-1,0	-0,2	-0,8	-1,0	-1,0	-0,1
<i>Højeste uddannelsesniveau</i>							
Grundskole/uoplyst	16,0	0,8	0,2	0,5	0,9	0,8	0,2
Gymnasial/erhvervsfaglig uddannelse	43,6	-1,1	-0,9	-1,2	-0,7	-0,8	0,3
Videregående uddannelse	40,4	0,2	0,7	0,7	-0,2	-0,0	-0,5
<i>Indkomst</i>							
Indkomst 1. kvartil	26,3	0,9	0,1	0,5	1,1	1,2	0,2
Indkomst 2. kvartil	27,1	-0,1	-0,1	-0,2	0,4	-0,1	0,3
Indkomst 3. kvartil	25,3	-0,7	-0,3	-0,5	-1,3 **	-0,9	-0,2
Indkomst 4. kvartil	21,3	-0,2	0,3	0,3	-0,3	-0,2	-0,3
Ingen oplysninger om mor	3,7	0,6	0,5	0,5	0,5	0,6	0,1

Anm.: ** angiver at forskellen er signifikant på et 5-procents-signifikansniveau.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

TABEL 5.9

Karakteristika for skoler, der har besvaret respondentskemaerne sammenlignet med elevfordelingen for hele landet, 2015, procent.

	Alle skoler i DK	Elev	Forældre	Personale	SB-formænd	Skoleleder	Skoleleder (udvidet)
<i>Karakter</i>							
Lavt karakterniveau	23,8	-1,60	-1,75	-2,40	-2,06	-3,35	0,05
Mellem karakterniveau	23,8	-1,06	-0,41	0,02	2,13	-0,88	-1,56
Højt karakterniveau	23,7	4,81	4,45	4,31	2,75	3,36	-0,16
Missing ¹	28,6	-2,16	-2,29	-1,93	-2,82	0,87	1,67
<i>Størrelse</i>							
Lille	33,3	0,90	-0,75	-0,16	-0,23	2,47	1,66
Mellem	33,3	-2,88	-1,64	-3,56	-0,93	-5,56	0,04
Stor	33,3	1,98	2,39	3,72	1,17	3,09	-1,70
<i>Region</i>							
Region Hovedstaden	21,7	-4,24	-3,41	-3,54	-1,38	-1,14	-0,04
Region Midtjylland	26,7	2,91	4,28	3,40	2,92	0,15	-0,55
Region Nordjylland	13,0	0,53	1,75	1,31	-0,33	-0,62	-0,06
Region Sjælland	16,1	0,41	-2,75	-1,39	2,10	1,27	2,12
Region Syddanmark	22,6	0,40	0,14	0,23	-3,31	0,34	-1,47

1. Årsagen til, at der på nogle skoler ikke er oplyst karakterer fra folkeskolens afgangsprøver, er, at en del skoler kun har elever op til 6 klasse.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd og Danmarks Statistik.

BILAG

TILGÆNGELIGE DATASÆT OG HJÆLPEVARIABLER

Fra SFI til STIL overleveres en række datasæt (jf. bilagstabel B.1.1). For det første et datasæt, der indeholder en liste over de 442 skoler, der blev inviteret til at deltage, deres hovedinstitutionsnummer, afdelingsinstitutionsnummer samt kommune-id ("skoler442").

For det andet et datasæt for hvert år, der giver et overblik over, hvilke skoler der deltager med besvarelser fra de forskellige typer af respondenter ("besvarelser2014" og "besvarelser2015"). Datasættene er masterdatasæt i den forstand, at de har til formål at give et overblik over skolens deltagelse.

For det tredje leveres tre typer af datasæt for hvert år og for hver type af respondenter. Et rådatasæt, et ikke-anonymiseret, men rensed datasæt, og et anonymiseret og rensed datasæt. De rensede datasæt indeholder bruttolisten af respondenter svarende til det antal skoler og personer, hvorpå der er indleveret kontaktoplysninger.

For det fjerde leveres de forskellige bruttoliste datasæt (også benævnt basedatasæt), der anvendes til at danne de tre typer datasæt. Da vi ikke har fået leveret data på samme måde ved hver dataleverance, foretages sammenkoblingen af baseoplysninger (såsom cpr-nummer, institutionsnummer, navn mv.), bruttolister og spørgeskemadata på forskellige måder i de enkelte år og for de enkelte typer af respondenter.

Se bilagstabel B1.1, for hvilke data der anvendes til at danne hvert datasæt. Data leveres i Stata-format med tilhørende programmeringsfiler til at kunne danne de rensede og anonymiserede datasæt ud fra basedatasæt og de rå spørgeskemabesvarelser.

For de rensede datasæt følger variabelbeskrivelser og kodebøger i Excel. For det rå datasæt henvises til de oprindelige spørgeskemaer, som også kan rekvireres hos STIL.

`_RAW`

Datasæt med filnavn `XX_raw` er data leveret i den form, som vi har fået fra Danmarks Statistik. Datasættene indeholder spørgeskemabesvarelser, institutionsnumre, respondenter og alle de øvrige oplysninger, der er erhvervet i forbindelse med dataindsamlingen (fx personens e-mail). Datasættet kan endvidere indeholde dubletter, fejl i institutionsnummer mv. For enkelte datasæt indgår bruttolisten af respondenter direkte i rådatafilen, mens andre bruttolister ligger som særskilte basedatasæt. Se endvidere tabel B1.1 for de datasæt, der anvendes til at danne rensede data.

Nogle spørgsmål indeholder åbne svarkategorier, og disse åbne svar findes i datasættet i rå form, hvilket vil sige, at de kan indeholde personfølsomme oplysninger, såsom lærerens navn.¹⁰

Variabelnummer i de rå datasæt er fortløbende. Det vil sige, at indholdet i spørgsmål 5 i elevskema fra 2014 ikke nødvendigvis svarer til spørgsmål 5 i 2015-elevskemaset, da spørgeskemaet fra 2014 ikke er helt identisk med skemaet for 2015.

`_RENSSET`

Datasæt med navnet `XX_rensset` indeholder bruttostikprøven og alle spørgeskemadata. Data er rensset mht. variabelnavne, labels og value labels. Enkelte variable er også omkodet med henblik på strømligning. Labels forklarer, hvad variablen indeholder, hvor value labels forklarer de enkelte svarmuligheder for hvert spørgsmål. Alle spørgsmål er desuden kodet til at kunne bruges som panel for 2014- og 2015-data.

Det vil sige, at de samme spørgsmål hvert år er kodet til at have samme nummer og navn i både 2014 og 2015. Datasættene er rensset, men ikke anonymiseret. Dog er alle navne på eksempelvis elever, lærere og skoler slettet. Data indeholder cpr-nummer (for elever). Spørgsmål stillet med åbne svar er anonymiseret, så der ikke fremgår navne eller andre personfølsomme oplysninger af svarene.

`_ANONYM`

De anonyme datasæt indeholder bruttostikprøven og alle spørgeskemabesvarelser. Dog indeholder data ingen personfølsomme oplysninger eller id – heller ikke institutionsnumre. I stedet indeholder forældre- og elevdata et anonymiseret `elev_id` (baseret på elevens cpr-nummer) og et anonymiseret institutionsnummer. (`skole_id`). Det er således muligt at identificere eksempelvis, hvilke lærere og elever der tilhører den samme

10. Vær opmærksom på, at DST ikke accepterer åbne svarkategorier, hvis data skal kobles til registrene.

skole, men det er ikke muligt at finde ud af, hvem disse lærere eller elever er, ej heller hvilken skole de præcis er på.

Data har bibeholdt deres panelstruktur på alle niveauer (skole-, forældre- eller elevniveau). Det vil sige, at en elev har samme anonyme id i både 2014- og 2015-dataene. Der er dog ikke tale om en nøgle, der kan føres tilbage til den enkelte elevs cpr-nummer. Det anonyme id for både skoler og personer er skabt ved hjælp af en random generator, hvor nøglen slettes umiddelbart efter, at det anonyme id er genereret.

Det betyder også, at det anonyme id ikke vil være det samme for samme person, såfremt man skulle få tilsendt en opdateret version af data. I forældre- og elevdatasættene er denne id-variabel dannet både på elevniveau (elev_id) og på skoleniveau (skole_id). De andre datasæt indeholder kun skole_id. Spørgsmål stillet med åbne svar er anonymiseret, så der ikke fremgår navne eller andre personfølsomme oplysninger af svarene.

BILAGSTABEL B1.1

Basedatasæt og rådata, der anvendes til at danne rensede data (_renset og _anonym) i 2014 og 2015.

	2014	2015
Elever	skoler442	elevskema5
	basalle_elever	elevskema7
	4kl	elevskema9
	6kl	
	89kl	
Lærere og andet pædagogisk personale	skoler442	basalle_laerer2015
	basalle_laerer	laerer2015_raw
	laererpædagogdata2014_raw	
Skoleledere	skoler442	leder_raw_2015
	basalle_skoleledere	
SB-formænd	skoleleder2014_raw	
	skoler442	skbestyr_raw_2015
	bas_bestyr	
	basalle_skoleledere	
Forældre	SBformand2014_raw	
	stik_0_2	forældre_raw_2015
	forældre_til_mette	
	basalle_forxxxxxx	
	Forældre2014_raw	

Anm.: Tabellen angiver, hvilke rådatasæt og basedata der skal anvendes til at danne de endelige rensede datasæt. De tilhørende programmeringsfiler er centrale ift. at kunne genskabe de leverede data, da de retter eventuelle småfejl i form af dupletter eller i de enkelte variable.

Kilde: SFI – Det Nationale Forskningscenter for Velfærd.

Bilagstabel B1.1 viser, hvilke rådata og baser der anvendes til at danne datasættene xx_rensed og xx_anonym for hver respondenttype hvert år.

MEDFØLGENDE PROGRAMMERINGSFILER

Vi leverer for hver respondenttype hvert år en master programmeringsfil, der kan danne datasættene beskrevet i afsnittet ovenfor. Hvis stierne til data ændres i programmeringsfilen til, hvor data placeres efter dataleve-

rance, kan den køres fra start til slut og vil danne hhv. et rensset ikke-anonymiseret datasæt og et rensset anonymiseret datasæt.

Der bliver i flere af datasættene rettet fejl i data, fx dubletter, fejl i institutionsnumre, cpr-nummer og lignende. Derfor er de medfølgende programmeringsfiler centrale, hvis data skal genskabes mest retvisende.

Følgende programmeringsfiler (do-files) medsendes dataleveringen:

- MASTER do file lærer 2015
- MASTER do file elev 2015
- MASTER do file forældre 2015
- MASTER do file skolebestyrelse 2015
- MASTER do file skoleleder 2015
- MASTER do file elev 2014
- MASTER do file forældre 2014
- MASTER do file skoleleder 2014
- MASTER do file lærer 2014
- MASTER do file skolebestyrelse 2014.

HJÆLPEVARIABLER

Alle datasæt indeholder enkelte hjælpevariable til fx at beregne bortfald, identificere antallet af skoler eller klasser eller til at koble de forskellige datasæt sammen. Dokumentation af de enkelte hjælpevariable indgår også i de excel-ark, der beskriver de enkelte variable.

XX_BESVAR

Variablen er i hvert datasæt navngivet efter identifikationsbogstavet for hvert skema. Det vil sige, at variablen hedder E_besvar i elevdatasættet og S_besvar for SB_formænd, L_besvar for skolelederne og LP_besvar for lærerne og det andet pædagogiske personale. Variablen er binær, og tager værdien 1, hvis observationen findes blandt de deltagende skoler og i spørgeskemadatasættet; har vi eksempelvis oplysninger på, at en elev vil deltage i undersøgelsen, og vi fra spørgeskema har én besvarelse fra samme elev, så er E_besvar=1. Derved kan XX_besvar-variablene bruges til at beregne bortfald for hvert skema.

STATUS (FRA 2015)

Status-variablene angiver, om den givne respondent har besvaret spørgeskemaet, har påbegyndt sin besvarelse, men ikke færdiggjort den eller slet ikke har besvaret. Denne variabel minder om vores konstruerede XX_besvar-variabel, men status-variablen kan skelne mellem fulde og delvise besvarelser. Vi har ikke samme oplysninger fra 2014-data.

KLASSETRIN (KUN FOR ELEVER OG PÆDAGOGISK PERSONALE)

Elevdatasættet og datasættet for lærerne og andet pædagogisk personale indeholder en variabel, der angiver, hvilket klassetrin (årgang), som eleven tilhører, og som lærerne eller andet pædagogisk personale underviser på.

KLASSE (KUN FOR ELEVER OG PÆDAGOGISK PERSONALE)

Elevdatasættet og datasættet for lærere og andet pædagogisk personale indeholder en klassevariabel, der angiver den præcise klasse, som hhv. eleven og det pædagogiske personale er tilknyttet. Eksempelvis om eleven går i 4.a. Klassevariabler er mangelfulde, men er det mest præcise bud på identifikation af enkelte klasser.

Datasættet for det pædagogiske personale i 2014 indeholder to klassevariabler. Den ene er den oprindelige klassevariabel fra de oprindelige kontaktoplysninger (også benævnt 'basen') og hedder *klasse_base*. Denne variabel er yderst mangelfuld og indeholder for størstedelen af personalet kun det klassetrin, som han/hun underviser på. Dette på grund af, at skolerne blot er blevet bedt om at angive, hvilket klassetrin, som lærerne eller det pædagogiske personale underviste på. Derfor har vi på baggrund af besvarelsene for det pædagogiske personale dannet en anden mere præcis variabel, som hedder *klasse*. Variablen er kodet tilsvarende i elevdatasættet og hedder også her *klasse*. Variablen er kodet på baggrund af personalets egne besvarelser for variablerne: LPQ50_Ch og LPQ51_Ch således i STATA¹¹:

BILAGSFIGUR B1.1

Stata-kode til oprensning af klassevariabel i data fra det pædagogiske personale.

```
*Rekoder så ud fra spørgeskema
gen h_klasse=LPQ50_Ch
replace h_klasse=LPQ51_Ch if h_klasse=="
*Hvis vores nye klassevariabel er blank, så overskriver vi med den gamle
replace h_klasse=klasse_base if h_klasse=="

gen k1 = regexs(0) if regexm(h_klasse, "^[0-9]+")
gen k2 = regexs(0) if regexm(h_klasse, "[a-zA-Z]+")
gen klasse=k1+k2
replace klasse=strupper(klasse)
```

Kilde: SFI - Det Nationale Center for Velfærd.

KLASSE0-KLASSE9 OG KLASSE_MIS (KUN PÆDAGOGISK PERSONALE I 2015)

Fra basen af lærere og andet pædagogisk personale har vi i 2015 oplysninger om, hvilke klassetrin den enkelte underviser er tilknyttet. Alle undervisere kan være tilknyttet flere klassetrin. Derfor har vi dannet 10 indikatorvariable, der angiver, hvilke klassetrin læreren er tilknyttet på sin

11. Denne oprensning af klassevariablen vil kunne laves for data for det pædagogiske personale i 2015 ved at anvende Stata-koden fra bilagsfigur B1.1.

skole (eller klasse_mis, hvis vi ikke har sådanne oplysninger tilgængelige for den pågældende lærer/pædagog). Disse variable har ikke noget at gøre med, hvilken klasse eller klassetrin, det pædagogiske personale har besvaret undersøgelsen for, men giver blot et overblik over, hvilke klassetrin, den pågældende lærer er tilknyttet på sin skole.

SKOLE_ID (KUN _ANONYME DATASÆT)

De anonyme datasæt indeholder en skole_id-variabel, der er dannet ud fra skolernes institutionsnummer. Vi har genereret en tilfældig nøgle ud fra alle de deltagende skolers institutionsnummer, som vi har koblet med alle datasæt både i 2014 og 2015. Nøglen er unik og bliver slettet umiddelbart efter sammenkoblingen på alle datasættene, hvilket vil sige, at det ved en ny dataleverance ikke vil være muligt at danne den samme nøgle igen. Fordelen ved denne nøgle er, at datasættene kan kobles sammen på skoleniveau for alle respondenttyper og på tværs af år, til trods for at data er anonymiseret.

ELEV_ID (KUN _ANONYME DATASÆT)

På samme måde som variabelen skole_id har vi dannet et elev_id i alle forældre- og elevdatasæt, som er de datasæt, der indeholder elevernes cpr-numre. Nøglen kan ikke genskabes og er slettet umiddelbart efter sammenkoblingen med forældre- og elevdata fra 2014 og 2015. Variablen gør, at det er muligt at koble forældre- og elevdata, til trods for at data er anonymiseret.

LITTERATUR

- Hansen, A.T., M. Friis-Hansen & V.M. Jensen (2015): *Folkeskolereformen, Beskrivelse af 1. dataindsamling*. København: SFI – Det Nationale Forskningscenter for Velfærd, Notat. 24.03.2015
- Keilow, M. & A. Holm (2015): *Udvikling af måleinstrument for elevadfærd og -holdninger, Baselinedata fra evaluering af Folkeskolereformen*. København: SFI – Det Nationale Forskningscenter for Velfærd, 15:09.
- Kjer, M.G., S. Baviskar & S. Winter (2015): *Kortlægning af skoleledelse i forbindelse med implementeringen af folkeskolereformen i 2015*. København: SFI – Det Nationale Forskningscenter for Velfærd, Rapport, under udgivelse.
- Nielsen, C. P., A.T. Hansen, V.M. Jensen & K.S. Arendt (2015): *Folkeskolereformen, beskrivelse af 2. dataindsamling blandt elever*. København: SFI – Det Nationale Forskningscenter for Velfærd, Rapport, under udgivelse.