

Niels Christoffersen  
Rorupvej 9 Rorup  
4320 Lejre

Til undervisningsminister Christine Antorini

Angående svar på min henvendelse af 9. marts 2015 om PISA-undersøgelsen. Reference:  
007.17V.981

Det er med nogen overraskelse, jeg kan konstatere, at ministeren er af den opfattelse, at de danske PISA resultater er repræsentative for de 15-16 årige indskrevet på en uddannelsesinstitution i den danske del af rigsfælleskabet.

Den del af svaret, der er overraskende lyder:

I de samlede analyser tages højde for overrepræsentationen af tosprogede elever ved at vægte stikprøven på en måde, der genskaber repræsentativiteten. Stikprøveudtræk og vægtning foretages af en ekspertgruppe i det internationale konsortium, der står bag PISA-undersøgelsen.

Ekspertgruppen får adgang til de officielle lister med samtlige skoler og antallet af elever på 15-16 år på hver skole og designer herefter stikprøverne i alle deltagerlandene, herunder Danmark. Den internationale ekspertgruppe sikrer også, at vægtene genskaber repræsentativiteten.

Svaret ser dog ud til at være i overensstemmelse med det, der fremgår af de danske rapporter fx tabel 6.5 side 161 i 2012 rapporten, hvor man kan se, at vægtene beregnes som (sum af elever i populationen)/(elever der har gennemført testen).

Min overraskelse bygger på følgende:

**Citat fra svar 1:**

I de samlede analyser tages højde for overrepræsentationen af tosprogede elever ved at vægte stikprøven på en måde, der genskaber repræsentativiteten.

**Bemærkning:**

Det er vist rimeligt normalt, at man i surveyundersøgelser søger at forbedre repræsentativiteten via baggrundsoplysninger fra registre. Baggrundsoplysninger i PISA tilvejebringes imidlertid via deltagerens (forældre, skoleledere og elever) svar på spørgeskemaer og for elevernes vedkommende også svar på spørgsmålene i de faglige opgaver. Det virker derfor noget mærkværdigt, at man kan "genskabe repræsentativiteten" med mindre stikprøven er udtrukket repræsentativt - det er den danske stikprøve som bekendt ikke og derfor er de baggrundsoplysninger, man råder over heller ikke repræsentative. Repræsentativiteten sikres ved at udtrække en repræsentativ stikprøve.

**Citater fra svar 1:**

Stikprøveudtræk og vægtning foretages af en ekspertgruppe i det internationale konsortium, der står bag PISA-undersøgelsen.

**Bemærkning:**

Vægtningen foretages som angivet i kapitel 8 i de internationale rapporter. Der bruges 5 justeringsfaktorer, herunder en justeringsfaktor for skoler, hvor der kun deltager med elever fra klasser beregnet for 15 årige. Da Færøerne netop deltager som en dansk undergruppe med alle elever fra 9. klasse må denne justeringsfaktor være brugt i forhold til de færøske skoler, men denne faktor ændrer ikke ved, at de færøske skoler ikke er en del af det danske skolesystem. Færøerne er

nævnt i den internationale rapport side 221 - skoler i stratum 5.

Det nationale center er for øvrigt ansvarlig for blandt andet at

- forberede den ramme skolerne udtrækkes fra (prepared the school sampling frame and submitted this to the international contractor for the selection of the school sample)
- udvælge eleverne (selected the student sample from a list of eligible students provided by the School Co-ordinators).

Skolelisten bygger på lister fra Danmarks Statistik fra året før undersøgelsen. Udvælgelsen af eleverne foretages øjensynligt af det nationale center.

Listen over det nationale centers opgaver står på side 108 i den internationale rapport.

### **Citat fra svar 1:**

Ekspertgruppen får adgang til de officielle lister med samtlige skoler og antallet af elever på 15-16 år på hver skole og designer herefter stikprøverne i alle deltagerlandene, herunder Danmark. Den internationale ekspertgruppe sikrer også, at vægtene genskaber repræsentativiteten.

### **Bemærkning:**

Stratificeringen er et nationalt anliggende (PISA ser dog ud til at skulle overtales af stærke argumenter for at godkende ændringer i stratificeringen). I DNK har man både i 2009 og 2012 brugt en national option, hvor man oversampler skoler og elever, så man får en større andel af elever med anden etnisk baggrund end dansk i stikprøven end andelen i populationen viser. Da oversamplingen er en national option skal de oversamplede elever/skoler ikke indgå i den internationale sammenligning (se side 453 i den internationale rapport), men det gør de øjensynligt. Standarden for nationale optioner er dermed formentlig ikke blevet implementeret korrekt. Måske har PISA ikke opdaget oversamplingen og fjernet de oversamplede elever/skoler fra udtrækket, fordi der ikke er overensstemmelse mellem PISA's definition af indvandrere og den definition, der bruges i de officielle lister fra Danmarks Statistik.

PISA registrerer en elev som native/indfødt, når eleven angiver, at man er født i testlandet og en af forældrene også er født i testlande. Hvis man ser på tosprogede elever, så vil den stramme udlændinge politik, der ført de sidste 12-15 år, sikkert betyde, at de fleste elever med anden etnisk baggrund end dansk er indfødte set med PISA's definition. Stratificeringen er derfor ikke entydig indenfor den ramme, hvor undersøgelsen foregår, fordi flere elever/skoler kan placeres i flere steder end der, hvor de rent faktisk bliver placeret med den danske definition. PISA har øjensynlig opdaget uoverensstemmelsen, idet der er skiftet navn fra "Minority Enrollment" i 2009 til "Immigrant Levels" i 2012. Men den danske definition/stratificering og udtræksprocedure er ens begge år.

### **Citat fra svar 1:**

Det har dog ingen betydning for (forvirring omkring betegnelser HØJ ... INGEN), hvordan stikprøven er blevet trukket og analyseret, .....

### **Bemærkning:**

Lad mig så se på det i lyset af følgende (mine fremhævelser):

The school sampling frame was defined in the *School Sampling Preparation Manual*<sup>2</sup> as **a frame that would provide complete coverage of the national defined target population without being**

**contaminated by incorrect or duplicate entries or entries referring to elements that were not part of the defined target population.** It was expected that the school sampling frame would include any school that could have 15-year-old students, even those schools which might later be excluded, or deemed ineligible because they had no PISA-eligible students at the time of data collection. **The quality of the sampling frame directly affects the survey results through the schools' probabilities of selection and therefore their weights and the final survey estimates. NPMs were therefore advised to be diligent and thorough in constructing their school sampling frames (side 70 i den internationale rapport).**

### **Fuldstændig dækning**

Det fremgår af de internationale rapporter i 2009 og 2012, at Danmark ikke opnåede fuldstændig dækning. Den danske udtræksramme manglede nogle internationale skoler, som skulle have været inkluderet i rammen: Overall exclusions were greater than 5% (6.18%). Data were fully explained – there was a difficulty in defining the school population – some international schools were not included when they should have been - side 284 i den internationale rapport 2012.

### **Færøerne**

Det er uklart, om deltagelsen af de færøske skoler siden 2006 (jævnfør fx PISA 2006 undersøgelsen – en sammenfatning fremlagt for Uddannelsesudvalget (2. samling) UDU alm. del - Bilag 11 Offentligt) kan betegnes som contamination (**entries referring to elements that were not part of the defined target population**), men det er klart, at de optræder som stratum 5 i den internationale rapport 2012 (side 221) - "and all students in stratum 5 schools (Faroes)" - det er også klart, at det færøske skolesystem ikke er en del af det danske skolesystem. Men deres deltagelse som en del af det danske system må uvægerligt påvirke det danske resultat, da de færøske elever scorer markant lavere end selv de danske tilsvarende elever.

### **Oversampling af elever med anden etnisk baggrund end dansk.**

Det er uklart, om oversamplingen af elever med anden etnisk baggrund end dansk (dansk definition) kan betegnes som en contamination (**incorrect or duplicate entries**) med det fører også til en oversampling af elever med svag socioøkonmisk baggrund jævnfør side 14 i delrapport 1 om kriteriebaseret af de nationale test (elever som både har deltaget i PISA 2012 og de nationale test i læsning/matematik). Socioøkonomisk baggrund har direkte indflydelse på resultatet via non-response mv.

I den tekniske rapport fra 2003 står således følgende - side 211:

In Denmark, students performed surprisingly poorly on booklet 9 when responding to both the science and the reading items. In contrast they performed quite well (relative to other booklets) on problem solving. In addition, it has been noted that the non-responding students (for each domain) have a lower value in the index of economic, social and cultural status (ESCS) than students who did respond to items on each domain. Given the positive correlation between ESCS and achievement, the lower values of ESCS for the students who were not assessed in a domain, and the lower than expected scores on booklet 9, it can be expected that the imputations for the non-assessed students will lead to a reduction in the mean scores in reading and science for Denmark.

Og så kan det i parentes bemærkes, at det også er derfor vægtene, der rapporteres i danske PISA 2012 tabel 6.5 ikke er rigtige (the students who were not assessed in a domain, and the lower than expected scores on booklet 9, it can be expected that the **imputations for the non-assessed students etc.**). Den måde at regne på der vist i tabel 6.5 (sum af elever i populationen)/(divideret med elever som har gennemført testen) forudsætter nok, at alle elever har løst alle opgaver, hvilket

jo lang fra er tilfældet i PISA. Se fx diagrammet her <http://videnskab.dk/kultur-samfund/pisa-ranglisten-er-baseret-pa-computerskabte-data>

### Citat fra svar 1

... og det internationale konsortium fastslår, at de danske PISA-undersøgelser opfylder de faglige krav til repræsentativitet.

### Bemærkning

Det internationale konsortium fastslår, at de danske data er blevet inkluderet i databasen. Repræsentativiteten sikres ved at udtrække en repræsentativ stikprøve, hvilket sikres via udtræksrammen jævnfør side 447 i den internationale rapport og en korrekt implementering af nationale optioner:

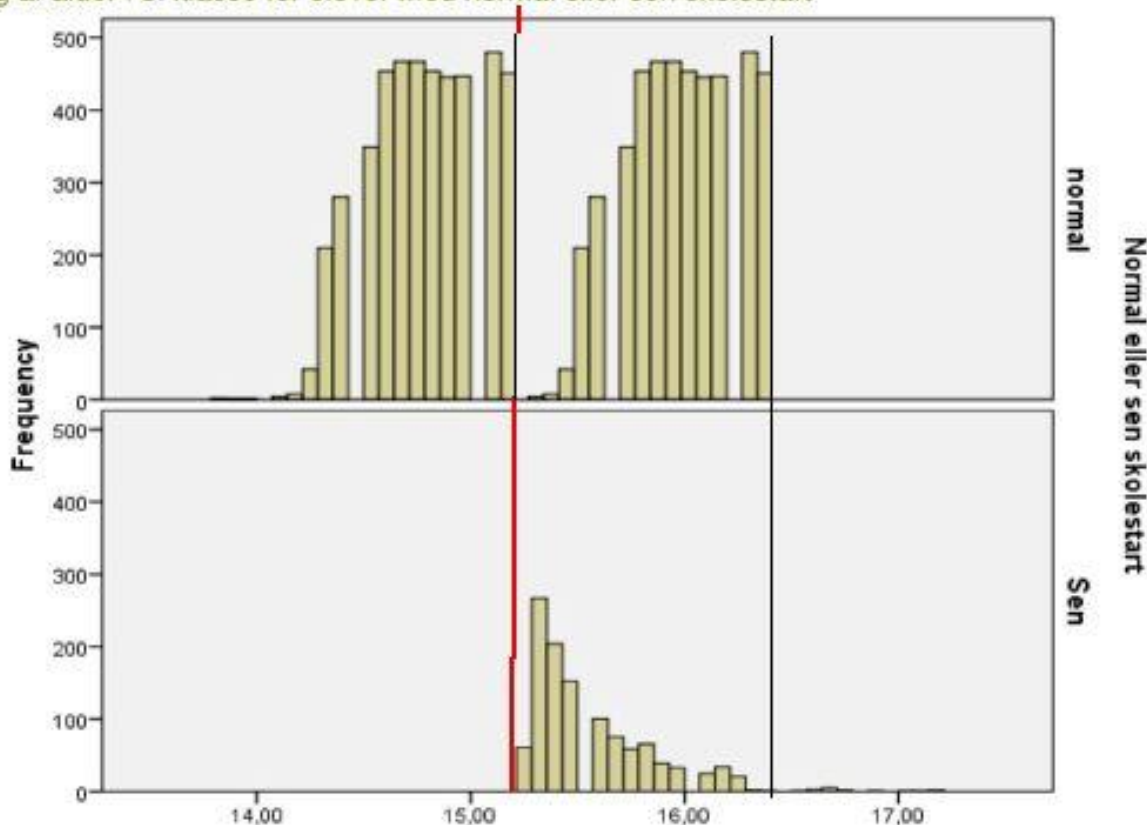
**Generalisability:** Data are collected from specific individuals, in a specific situation, and at a certain point in time. Individuals to be tested, test materials and tasks, etc. should be selected in a way that will ensure that the conclusions reached from a given set of data do not simply reflect the setting in which the data were collected but hold for a variety of settings and are valid in the target population at large. Thus, collecting data from a representative sample of the population, for example, will lead to results that accurately reflect the level of literacy of fifteen-year-old students in a country. (Min fremhævelse)

Det internationale konsortium sikrer ikke repræsentativiteten, de foretager en lang række kontroller, men de forudsætter øjensynlig, at de vedtagne standarder er kendt.

• **Timeliness:** Consistency, precision and generalisability of the data can be obtained in a variety of ways. However, the tight timelines and budgets in PISA, as well as the sheer number of participating countries, preclude the option of developing and monitoring local solutions to be harmonised at a later stage in the project. Therefore, the standards specify one clear-cut path along which data collection and data submission should progress.

Det fremgår af rapporterne om kriteriebaseret af de nationale test figur 4.10, at der har deltaget elever som aldersmæssigt ikke hører med i PISA.

Fordeling af alder i 8. klasse for elever med normal eller sen skolestart



I noten til figuren hedder det:

Alderen i 8. klasse er beregnet ud fra PISA registrering af alderen, idet der er trukket et år fra alderen blandt de elever, der er blevet PISA testet i 9. klasse.

I den øverste del af fig. har jeg "flyttet" (den røde streg) 9. klasse eleverne tilbage, hvor de hører hjemme, idet man har fundet aldersfordelingen i 8 klasse (2011) af de normale skolestartere ved trække et år fra deres alder i 9. klasse (2012). Så stikprøven indeholder elever, som ikke burde deltage. Det er formentlig sket, fordi man udtrukket flere elever end TCS angiver (national option), for eksempel hedder det, at der i stratum 02 blev udtrukket 20 elever (TCS=20) med dansk etnisk baggrund, og alle eleverne med anden etnisk baggrund end dansk blev tilføjet til udtrækket.

TCS er det antal elever (15 årige) man må udtrække.

I den tekniske rapport side 179 angives 104 elever som ineligibile students, men der ser tydeligvis ud til at være flere, som faktisk har deltaget.

Sammenfattende:

1. Den danske udtræksramme manglede nogle internationale skoler som skulle have været inkluderet i rammen: Overall exclusions were greater than 5 % (6.18 %). Data were fully explained – there was a difficulty in defining the school population – some international schools were not included when they should have been - side 284.  
Så alle skoler med potentielle 15 årige deltog ikke i lodtrækningen, jævnfør også standard 1.5: Students are sampled using *agreed upon*, established and professionally recognised principles of scientific sampling and in a way that represents the full population of *PISA-Eligible students*.

2. Det kan vises, at den danske stikprøve indeholder elever, som ikke er PISA elever, de kan være smuttet med, fordi man tilføjer alle tosprogede elever udover TCS til udtrækket og så glemte, at der var tale om en national option. Den internationale rapport angiver 104 elever, som værende ineligibile. Men det ser dog ud til, at der er flere.
3. Den danske stikprøve indeholder skoler (stratum 5), som er beliggende på Færøerne. Færøerne klarer sig langt dårligere end selv tosprogede danske elever. Det påvirker selvfølgelig resultatet i nedadgående retning og det kan man ikke vægte sig fra.
4. Den manglende overensstemmelse mellem definitionerne på indvandrere (PISA/Danmarks Statistik) gør, at man vil få to forskellige stratificeringer og dermed vægte. Det endelige resultat afhænger derfor af hvilken definition man bruger. **The quality of the sampling frame directly affects the survey results through the schools' probabilities of selection and therefore their weights and the final survey estimates.**
5. I realiteten har man bare oversamlet elever med svag socioøkonomisk status (se side 14 delrapport 1 om kriteriebaseret af de nationale test).
6. Det ser faktisk ud til, at de oversamlede skoler/elever (national option) deltager i PISA undersøgelsens internationale del, hvilket åbenbart strider mod standarden for implementering af nationale optioner.

### Spørgsmål

Er ministeren stadig af den overbevisning, at de danske PISA resultater er repræsentative for de 15 årige under uddannelse i den danske del af rigsfælleskabet?

Med venlig hilsen  
Niels Christoffersen  
Rorupvej 9  
4320 Lejre

Kopi til Børne- og Undervisningsudvalget