

Tilsvar i Aftenposten - Nasjonale prøver: Nye analyser er viktige

Utdanningsdirektoratet (Udir) og Frischsenteret jobber begge for å finne de beste analysemetodene for nasjonale prøver. Forskjellene skyldes på ingen måte kun en «simpel regnefeil».

ARTIKKEL | SIST ENDRET: 12.09.2023

Innlegget er skrevet av Udirs direktør, Morten Rosenkvist, og Frischsenterets direktør, Simen Markussen. Dette innlegget ble publisert i Aftenposten 12. september 2023.

Denne uken omtalte Aftenposten hvordan forskere som analyserte tidsutviklingen i nasjonale prøver, kom til en annen konklusjon enn Udirs offisielle tall. Vi vil sammen forklare hva forskjellene består i.

Forskjellen gjelder først og fremst sammenligninger mellom elever som tok prøven i ulike år. Også med dagens offisielle tall får altså elever, skoler og kommuner et godt bilde av hvordan de gjør det sammenlignet med andre dette året.

Det de nye analysene fra Frischsenteret viser, til forskjell fra de offisielle tallene, er at elevenes ferdigheter har vært i endring; De har blitt mye bedre i engelsk og litt svakere i regning og lesing. En slik sammenkobling av prøver på tvers av år, bruker statistiske modeller og et omfattende datamateriale. Når resultatene er ulike, er det derfor misvisende å omtale dette som en simpel "regnefeil". Trolig – vi er nemlig ikke helt sikre ennå – skyldes forskjellene ulike valg av «innstillinger» i statistiske analyseverktøy. Vi jobber for å få klarhet i årsaken til de ulike resultatene.

Underveis i dette arbeidet, har Frischsenteret delt analyser og midlertidige konklusjoner, mens Udir har delt data, resultater og estimeringer fra tidligere år. Dette har vært en konstruktiv, men naturlig nok også krevende dialog. Udir har ingen feil eller mangler å påpeke til analysene som nå er lagt fram av forskerne, men har også behov for å sette seg ytterligere inn i dette i samarbeid med egne kompetansemiljø.

Fra Udirs side, har dette vært en nyttig og verdifull hjelp i en prosess vi påbegynte våren 2022 med å

revidere analysearbeidet som gjøres på nasjonale prøver. Dette betyr også at problemene som nå er avdekket, ikke berører årets elever, siden analysemetodene uansett er endret.

Udir ønsker å berømme Frischsenterets forskere, og alle som fatter interesse for analyser av disse prøvene. I den sammenhengen er det verdt å påpeke at Udirs kronikk med tittelen «Spekulativt om nasjonale prøver» (Aftenposten 7.9) ikke var ment som en kritikk av forskerne - men ble sendt inn med tittelen «Spekulativt fra Aftenposten om nasjonale prøver», med henvisning til Aftenpostens fremstilling av saken.

De siste ti årene har det skjedd en gradvis dreining i retning av mer åpenhet både i forskning og statistikkproduksjon. Det er nok mange årsaker til dette: Ny programvare med åpen kildekode, høyere kompetanse rundt koding, økte krav til gjennomsiktighet og reproduserbarhet både fra forskningstidsskrifter og i en mer kritisk offentlighet. I praksis gjør dette at mens nye analyser, som den forskerne her har gjort av nasjonale prøver, lar seg dele, dokumentere og etterprøve. Slik er det ofte ikke med analyser fra fortiden, som beregningene av skalapoeng i nasjonale prøver fra 2014 til 2021.

I fremtiden vil Udir lage statistikk for skalapoeng på en annen måte, med bruk av et åpent tilgjengelig analyseprogram og kode som fritt kan deles. Det metodiske rammeverket her vil være det samme som forskerne ved Frischsenteret og CEMO har brukt.