

Teknisk appendiks til kronik i JP den 9. oktober 2003

Af

Pia Vedel Ankersen, ph.d.-stipendiat i statskundskab

Jørgen Johannes Poulsen, lektor i statskundskab

Siggi Brandt Kristoffersen, ph.d.-stipendiat i statskundskab

I JP's kronik den 9. oktober 2003 fremsættes særligt to påstande om alvorlige fejl i professor Helmuth Nyborgs (HNs) undersøgelse af mænd og kvinders intelligens. Formålet med dette appendiks er på en mere teknisk måde, end det kan lade sig gøre i en kronik, at klargøre for disse fejl.

Gennemgangen af de fremsatte resultater baserer sig udelukkende på artiklen: Nyborg, Helmuth (2003). 'Sex differences in *g*.' in Helmuth Nyborg (ed.): *The Scientific Study of General Intelligence: Tribute to Arthur R. Jensen*. Oxford: Elsevier Science Ltd., pp. 187- 222. I denne refererer HN primært til resultater, der er præsenteret i et paper fra 2001. Dette betyder, at det er svært at gennemskue, i hvilken grad HN opfylder almindelige forskningskriterier, hvis man kun læser dennes nyeste udgivelse. Da HN på trods af gentagne henvendelser har nægtet at offentliggøre sine data eller det centrale paper fra 2001, kommer flere af de fremførte påstande ulukkende til at basere sig på hans fortolkning af egne resultater.

På trods af dette er der på mindst to meget centrale områder helt tydeligt problemer i de konklusioner, han drager. I det følgende vil vi argumentere for, at den meget omtalte kønsforskel i mænds favør i virkeligheden er behæftet med så stor usikkerhed, at HN kunne have skrevet om den manglende forskel mellem kvinder og mænd eller endda om kvindernes overlegne intelligens. Desuden at han i et desperat forsøg på at underbygge mændenes forspring – bevidst eller ubevidst – begår fundamentale fejl i måden at analysere på.

Kønsforskellen

I den omtalte bog fra 2003 præsenterer HN os for de hårde tal, der skal bevise mændenes højere gennemsnitlige score (p. 214). På den berømte *g*-faktor har mændene gennemsnitligt opnået en score på 0,36, mens kvinderne tilsvarende kun scorer -0,01.¹ Det er her imidlertid vigtigt at tage højde for, at der

¹ Den opmærksomme læser vil her lægge mærke til, at den gennemsnitlige *g*-score for de 52 respondenter ikke kan være 0, som vil være den normale gennemsnitsværdi på en faktor. Dette skyldes, at *g*-faktoren er udledt på baggrund af en større

naturligvis ikke er tale om en undersøgelse, der omfatter hele den danske befolkning, men kun et mindre udsnit. I dette tilfælde faktisk et meget lille udsnit på 52 personer i aldersgruppen 16 år og opefter. Som ved alle andre undersøgelser, der har en ambition om at udtale sig mere generelt om befolkningen som sådan, må HN derfor lægge et usikkerhedsinterval omkring størrelsen af kønsforskellen. På trods af, at han intet sted viser os resultatet af dette, så kan det konstateres, at han vælger fortsat at fastholde gyldigheden af den rapporterede forskel. Som nedenstående analyse viser, er det ikke i overensstemmelse med almindelige regler for statistisk inferens.

På baggrund af de oplysninger, HN giver os, kan usikkerhedsmarginen udregnes efter følgende formel:²

$$\text{Forskel mellem køn} \pm 1,96 * \sqrt{\frac{s_{\text{mænd}}^2}{n_{\text{mænd}}} + \frac{s_{\text{kvinder}}^2}{n_{\text{kvinder}}}} \iff 0,37 \pm 1,96 * \sqrt{\frac{1,06^2}{31} + \frac{0,74^2}{21}} \iff (0,37 \pm 0,49) g$$

Resultatet af denne udregning er altså, at mænds intelligensmæssige forspring udtrykt i *g* ligger i intervallet fra -0,12 til 0,86 *g*. Dette svarer til, at den gennemsnitlige mand i en IQ-test scorer et sted mellem 1,8 point *mindre* og 12,9 point mere end den gennemsnitlige kvinde. Det kan altså ikke afvises, at mænd faktisk kan have en højere *g*-score, men det modsatte kunne også være tilfældet. Det normale – og det eneste statistisk holdbare – ville derfor være at konkludere, at der ikke er fundet en pålidelig kønsforskelle. Laver man direkte en statistisk test for, om vi kan være sikre på, at kønsforskellen ikke blot er lig nul, så opnås en *p*-værdi på 0,139.³ Dette ville ikke normalt føre til en accept af kønsforskellens eksistens, da det ville betyde, at 13,9 % af alle rapporterede videnskabelige resultater var direkte forkerte. Inden for intelligensforskningen, hvor der er stor uenighed om disse sammenhænge, og hvor konsekvensen af at melde forkerte resultater ud kan være så omfattende, kan der modsat tales for, at den almindelige accept af 5 % fejl også er for højt sat. Særligt hvis man taler for så omfattende samfundsmæssige konsekvenser med baggrund i forskningsresultater, som HN flere gange har gjort sig til talsmand for. Derfor kunne man argumentere for et endnu bredere interval end det netop udregnede, hvilket ville invalidere HN's resultat yderligere.

gruppe respondenter (n=376), som samlet har et gennemsnit på 0 *g*. Da en kønsforskelle udelukkende kunne findes blandt de 52 deltagende voksne (mindst 16 år gamle), er det kun denne gruppe, der er anvendt i den videre analyse.

² Der anvendes et 95 % konfidensinterval, som er standard inden for størstedelen af den samfundsvidenskabelige forskning. Helmuth Nyborg oplyser ikke, hvordan kønsfordelingen af de 52 respondenter er, blot at de er jævnt fordelt. Da en fordeling med 31 mænd og 21 kvinder taler mest til hans fordel i udregningen, tager vi dette som vores udgangspunkt. Da det er mere sandsynligt, at Nyborg arbejder med en ligelig kønsfordeling, har vi imidlertid også gennemført analysen med denne antagelse. Konfidensintervallet viser i det tilfælde, at mænd scorer fra 1,9 point mindre til 13,0 point mere end kvinderne i en IQ-test (udtrykt i *g* går intervallet fra -0,13 til 0,87 point).

³ Ved en ligelig kønsfordeling vil *p*-værdien være lig 0,144.

Dermed kan det konkluderes, at Helmuth Nyborg enten ikke fortæller, at han med et almindeligt usikkerhedsniveau får afkræftet sin centrale hypotese, eller at han har sat kravene til et videnskabeligt bevis så usædvanligt langt ned, at han har fået det ønskede resultat. Ingen af delene ville tale til hans fordel.

Redningsforsøget

Måske netop i en stille erkendelse af de skuffende resultater beslutter HN sig i stedet for, at en 'simpel' test som ovenstående ikke er tilstrækkelig til at afsløre en sand kønsforskel. I stedet anvendes en fremgangsmåde, der i kraft af en særligt eksotisk Schmid-Leiman transformation og den manglende redegørelse for principperne bag analysemetoden, hensætter læseren i enten en højere grad af forvirring eller – og det er måske hensigten – i en opgivende accept af, at HN må vide, hvad han taler om. Mens det udelukkende ud fra beskrivelsen i 2003-bogen er svært at gennemskue, præcis hvad der foregår i analysen, så står det imidlertid klart, at udregningen af g -faktoren pludselig ændrer sig. Hvor der tidligere anvendtes en hierarkisk faktoranalyse, der ganske fornuftigt vægtede de forskellige typer af tests – og dermed af former for intelligens – i analysen ligeligt, så introduceres køn nu som en selvstændigt variabel i faktoranalysens udledningen af g . I forlængelse af dette mener HN så, at man kan afgøre spørgsmålet om kønsforskellen ved at se på køns bidrag (*loading*) til den samlede faktor (p. 209). Dette er imidlertid at snyde på vægten; hvis køn indgår som et af de delelementer, der bestemmer den enkeltes g -score, så kan man ikke bagefter uafhængigt teste, om køn kan bidrage med noget til g -faktoren – det har man allerede bestemt, at den skal gøre. Man kan sammenligne fremgangsmåden med en intelligenstest, hvor man giver alle mænd bonuspoint for deres køn for derefter at undersøge, om forsøgspersonernes køn har betydning for deres score.

I betragtning af denne lille håndsrækning til mændene er det opsigtsvækkende, at resultatet kun viser en forskel i deres favør, der ligger på grænsen af, hvad der normalt anses at være statistisk holdbart ($p \approx 0,052$). Derfor beslutter HN på forhånd, at det udelukkende giver teoretisk mening, at en eventuel forskel i intelligens skulle kunne favorisere mændene, hvorved usikkerheden fikst kan halveres ($p = 0,026$). Denne fremgangsmåde passer på mange måder ind i det generelle billede af hans analyser, der som det vigtigste validitetskriterie skal kunne underbygge den konklusion, der synes at være skrevet på forhånd: At mænd bør besætte de vigtigste positioner i samfundet, fordi de beviseligt er bedst egnede til det gennem deres overlegne intelligens.