

Forandringer i gjennomsnittlig IQ i Skandinavia.

Flynn-effektens vekst og fall

Artikkelen tar opp gjennomsnittsforandringer i IQ over generasjoner. I mange andre land øker gjennomsnittene fortsatt, mens de har stoppet opp i Skandinavia. Tester som måler problemløsningsevne øker mest. Årsakene er ennå ikke fullstendig klarlagt.



JON MARTIN SUNDET

Psykologisk institutt,
Universitetet i Oslo

Innledning

I to klassiske artikler dokumenterte statsviteren (political scientist) James R. Flynn (1984, 1987) en forandring av gjennomsnittlige skårer på intelligens tester (heretter kalt IQ) slik at de som var født i yngre fødselskull hadde gjennomsnittlig høyere IQ enn de som var født i tidligere generasjoner. Fenomenet var allerede kjent både av psykologer og pedagoger som brukte tester i sitt arbeid. De ble tvunget til å renormere sine tester fra tid til annen, fordi testene ble for ”lette”. Herrnstein & Murray (1994) døpte fenomenet ”Flynn-effekten” – en betegnelse som blitt standard i vitenskapelige miljøer.

I artikkelen fra 1984 dokumenterte Flynn eksistensen av fenomenet i USA. I den andre klassiske artikkelen (Flynn, 1987) gjennomgikk han data fra 14 europeiske land og fant tilsvarende økninger i gjennomsnittlige intelligens-test-skårer. Norge var et av de landene som inngikk i databasen. Utviklingen i blant annet Storbritannia og Japan ble tidlig studert av Lynn & Hampson (1986). I Danmark utviklet Tom W. Teasdale ved Københavns Universitet (i samarbeid med David R. Owen ved City University of New York) tidlig en egen forskningstradisjon hvor den første artikkelen kom omtrent samtidig med Flynn’s andre klassiske artikkel (Teasdale & Owen, 1987). Også i Sverige var det ganske tidlig forskning på Flynn-effekten (Emanuelsson, Reuterberg & Svensson, 1993). Den norske forskningen på området kom stort sett ikke i gang før dette årtuseten (Sundet, Barlaug & Tørvjussen, 2004). I de senere årene er antall land hvor det er påvist en Flynn-effekt økt jevnt og trutt, blant annet i Asia (Flynn, 1998b) og i land som normalt ikke blir regnet med som en del av den industrialiserte verden. Flynn-effekten er for eksempel observert på lands-

bygda i Kenya (Daley, Whaley, Sigman, Espinosa & Neumann, 2003) og i den Dominikanske republikk i Karibia (Meisenberg, Lawless, Lambert & Newton 2005). Det er altså et meget utbredt fenomen, også sett i et globalt perspektiv. Flynn-effekten gjelder for begge kjønn (Emanuelsson, Reuterberg & Svensson, 1993; Flynn, 1998a), og for barn i forskjellige aldre og etniske subgrupper i USA. (Ang, Rogers & Wänström, 2010).

Styrken på Flynn-effekten har vist seg å være ganske forskjellig fra land til land. En av de største økningene var blant nederlandske soldater. Der fant man en økning på omkring 20 IQ-enheter i en periode på 30 år. Et gjennomsnitt over mange land antyder en økning på ca 3-5 IQ-enheter pr tiår (Flynn, 1999), altså en gjennomsnittsøkning på ca 15-25 IQ-enheter i løpet av en 50-års-periode. Dette er en betydelig økning. Forskjellen mellom en person som har for eksempel 100 i IQ og en person som har rundt 85 er et sprang fra en person som er ganske normalt begavet til en person som kan være noe svakt begavet. Et tanke-eksperiment kan kanskje belyse dette ytterligere: En person som skårer like bra som om lag halvparten av befolkningen på et gitt tidspunkt ville med like mange riktige svar ha skåret omtrent som 34% av befolkningen på samme testen 50 år senere.

Intelligenstester

Studier av Flynn-effekten er basert på intelligens tester. Det er mange typer slike tester, og det vi i dag vet om Flynn-effekten stammer fra studier hvor det er brukt forskjellige intelligens tester. De fleste intelligens tester består av flere deltester med forskjellig innhold. Begrunnelsen for dette er at intelligens viser

seg i ferdigheter av forskjellig art. En av de mest brukte testene er den som ble laget av amerikaneren David Wechsler for ca 80 år siden. Denne testen har flere forskjellige deltester som tester både teoretiske og praktiske ferdigheter. IQ brukes gjerne som betegnelse på prestasjonen på testen som helhet, og grovt sett beregnes den med utgangspunkt i antall riktige svar. IQ på en gitt person beregnes ut fra gjennomsnittsprestasjonen i den aldersgruppen vedkommende hører til. Gjennomsnittsprestasjonen gis en (vilkårlig) IQ-verdi på 100. Prestasjonene over og under gjennomsnittet får IQ henholdsvis over og under 100. Variasjonen rundt gjennomsnittet (standard-avviket) settes vanligvis til 15 (også vilkårlig). Ut fra antakelser om normalfordeling av test-skårer forteller IQ-skåren hvor en gitt person ligger i forhold til andre i sin. En IQ på for eksempel 115 (et standardavvik over gjennomsnittet på 100) forteller at prestasjonen på testen ligger blant de 15-16% beste i den gruppen personen tilhører. Ca. 68% av en befolkning har en IQ-skåre mellom 85 og 115. En annen mye brukt test er Ravens matrisetest. I denne testen er oppgaven å finne ut hvilken figur som passer inn i et sett med figurer som forandrer seg etter bestemte regler. Poenget er å oppdage hvilke(n) regler (regel) som styrer forandringen(e) av figurene. I utgangspunktet skulle en tro at bruk av forskjellige tester vanskeliggjør sammenlikningen av resultater fra forskjellige studier. Det er antakelig ikke et alvorlig problem. Et av de mest robuste funnene innenfor intelligensforskningen er at prestasjoner på tester med forskjellig innhold korrelerer ganske godt med hverandre. I klartekst betyr dette at de som gjør det godt på en test også gjerne gjør det godt på andre tester, de som gjør middels på en har en tendens til å gjøre det middels på andre tester, og så videre. Det er slike funn som ligger til grunn for den vanlige oppfatningen at det finnes en underliggende generell intelligens som slår ut i en rekke ferdigheter (Jensen, 1998).

Flynn-effekten i Skandinavia

Som kjent er det generell verneplikt både i Danmark og Norge. I begge land blir unge menn innkalt til undersøkelser når de er 18-19 år for å avgjøre om de er egnet til å avtjene sin verneplikt. (I Norge kalles dette for sesjon).

Over 90% møter opp til disse undersøkelsene. I Danmark blir de testet med Børge Prien's Prøve (laget i 1953), som består av fire tidsavgrensede enkelttester (cf. Teasdale & Owen, 1989). Den ene (Bokstav-matriser) er modellert etter den ovennevnte Raven's matrise-test og består av 19 oppgaver. Verbale Analogier-testen består av 24 ledd og oppgaven er å fullføre analogier av typen "A forholder seg til B som C forholder seg til?". I Tallsekvens-testen (18 ledd) blir man spurt om å finne neste tall etter 4 foregående tall. Den fjerde deltesten (Geometriske Figurer) blir personen bedt om å konstruere oppgitte figurer ut fra deler som skal settes sammen. Ut fra prestasjonene på deltestene kan en samleskåre beregnes. Børge Prien's prøve har ikke vært forandret siden 1953 (Teasdale & Owen, 2008). Samleskåren på Børge Prien's Prøve korreler høyt (rundt 0.80) med skårene på voksen-versjonen av Wechsler-testen (WAIS, Wechsler Adult Intelligence Scale).

I Norge bruker man 3 tidsavgrensede deltester. Den ene består av 30 regneoppgaver på prosa (Regneprøven), en består av 54 oppgaver som består i å finne synonymet til et gitt ord (Synonymprøven). En test består i å finne det alternativet som passer inn i figurer som forandrer seg etter bestemte regelmessigheter. Figurprøven (36 ledd) har store likheter med Raven's matrisetest. Figurprøven og synonym-prøven har forblitt uforandret over hele perioden som vi har data fra (de ble tatt i bruk i 1954). Regneprøven ble noe forandret i 1963. I begynnelsen av 1990-tallet ble formatet på denne testen forandret fra åpne svar til flervalgs-svar hvor respondenten skulle velge sitt svar blant 6 alternativer. Også de to andre testene (Regneprøven og Figurprøven) har et flervalgs-format. Ut fra antall riktige svar på hver enkelt av disse prøvene beregnes skårer på hver enkelt av delprøvene samt en samleskåre (Alminnelig Evnenivå) for prestasjonene på delprøvene samlet. Korrelasjonen mellom samleskåren Alminnelig Evnenivå og WAIS er funnet å være høy også i Norge (ca. 0.75).

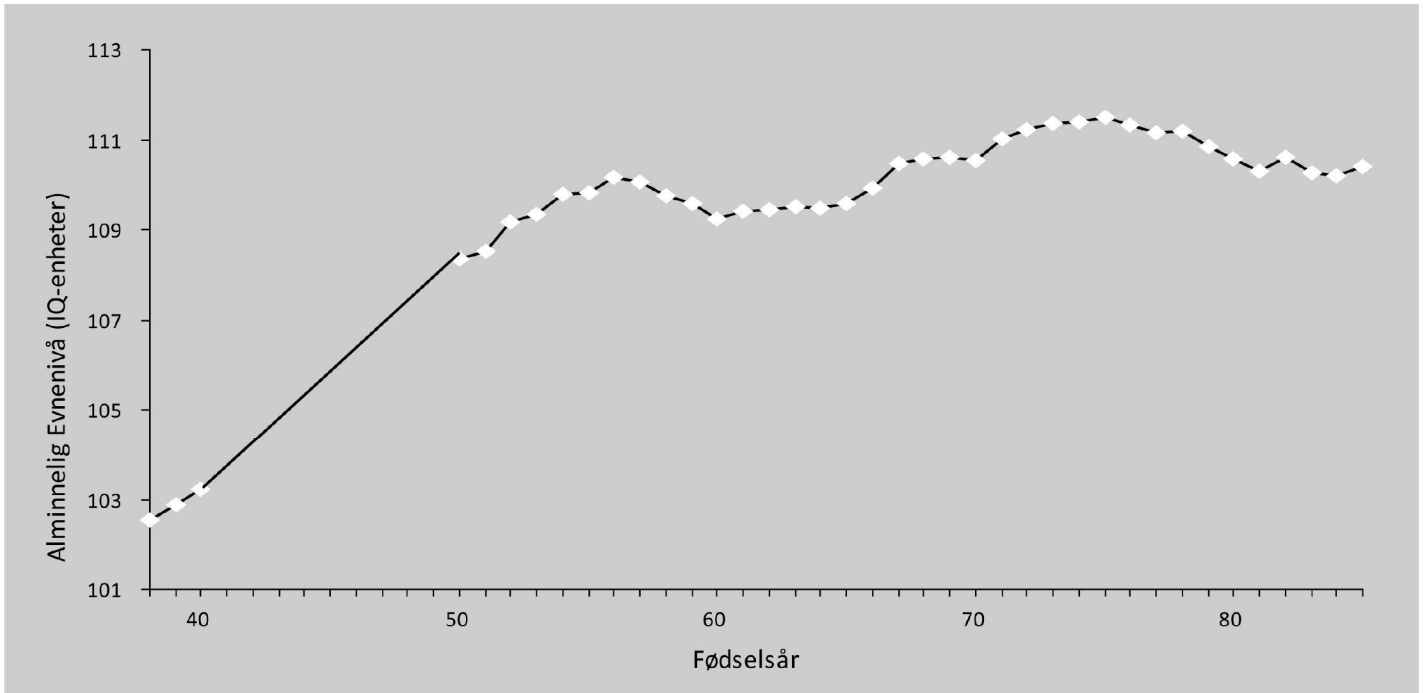
Som man ser, er det klare likheter mellom Børge Prien's prøve og det norske testbatteriet. Dette er ikke tilfeldig. De ble laget omtrent samtidig (Børge Prien's Prøve i 1953 og det norske testbatteriet ble laget i 1954). En av de sentrale aktørene i konstruksjonen av de norske prøvene var faktisk fra Danmark.

De danske og norske dataene regnes som blant de aller beste i forskningsmiljøet rundt Flynn-effekten. Det skyldes flere forhold. For det første bygger dataene på tester som har vært uforandret i bortimot 60 år. Omtrent alle er blitt testet på omtrent samme alder (18-19). Dette er viktig fordi man dermed kan utelukke forandringer som skyldes aldersforskjeller. De danske og norske studiene har begge store utvalg. Hovedtyngden av de norske dataene har 20.000-25.000 i hvert årskull. De seneste studiene fra Danmark har tilsvarende tall. Dette utgjør vesentlige andeler av fødselstallene i begge landene. Store utvalg sikrer at de statistiske usikkerhetsmarginene ekstremt små.

Fig 1 viser forløpet av gjennomsnittet av Alminnelig Evnenivå omregnet til IQ-enheter i kullene som ble født fra og med 1938 til og med 1985, altså en periode på bortimot 50 år (jfr. Sundet, Barlaug & Tørjussen, 2004). Vi har brukt det året som de norske testene ble laget og normert (1954) som referansepunkt.

Fig 1 viser en markant stigning i gjennomsnittet av Alminnelig Evnenivå fram til fødselskullene i midten av 1950-årene. Det vil si at det var en markant økning i gjennomsnittene fra de som var født like før 2. verdenskrig til de som var født i det første tiåret etter krigen. Økningen var ca. 7.6 IQ-enheter, noe som tilsvarer ca. 0.4 IQ-enheter pr år. Fra midten av 1970-årene og utover avtok stigningen merkbart. Derfra og ut til midten av 1970-årene var den "bare" ca. 1.5 IQ-enheter, og tilsvarer en økning på knapt 0.1 IQ-enheter pr år. En fullstendig utflating og nedgang i gjennomsnittene skjedde fra fødselskullene omtrent i midten av 1970-årene. Denne nedgangen tilsvarer omtrent 1 IQ-enhet.

FIGUR 1 Gjennomsnittlig Alminnelig Evnenivå (IQ-enheter) etter fødselsår



Det er bemerkelsesverdige likheter mellom de danske og norske funnene. Teasdale og Owen (1989) rapporterte data for fødselskullene 1939-58 til fødselskullene 1967-1969. Den totale økningen av gjennomsnittlig samleskåre i denne perioden, omregnet til IQ-enheter, var i nærheten av 7.5 IQ-enheter. Dette er ganske nøyaktig den samme stigningen som i omtrent den samme perioden i Norge. Den avtagende stigningshastigheten som er så tydelig i norske data er også tilstede i de danske dataene (Teasdale & Owen, 2000). Nedgangen i gjennomsnittlig IQ fra fødselskullene midten av 1970-årene som er tydelig i Fig 1, er også funnet i Danmark (Teasdale & Owen, 2005, 2008), men kanskje noen få år senere i Danmark enn i Norge. I Fig 1 ser vi også en liten nedgang i gjennomsnittlig Alminnelig Evnenivå i årene 1957-1960. Det er sannsynligvis en lokal norsk effekt som vi ikke vet årsakene til. I tillegg er det et lite fall i gjennomsnittet i fødselskullet fra 1970. Dette kan skyldes en influensa-epidemi (Hong Kong-influensaen) som var på sitt verste i Norge like før og like etter årsskiftet 1969-1970. Dette kan ha ført til nevrologiske skader på fostre i de tilfellene hvor moren var rammet av influensa (Eriksen, Sundet & Tambs, 2009).

Svenske data på gjennomsnittsforandringer i prestasjoner over fødselskull (se nedenfor) ser ut til å peke i noenlunde samme retning som de danske og norske.

Reversingen av Flynn-effekten fra fødselskullene omtrent i midten av 1970-årene kan kanskje være et lokalt skandinavisk fenomen. Det er riktignok rapportert et lignende fenomen i England (Shayer & Ginsburg, 2009). I dette studiet ble det brukt tester som er blitt konstruert med utgangspunkt i den

sveitsiske psykologen Jean Piaget's teorier om barns intellektuelle utvikling. Slike tester regnes som regel ikke som standard intelligens tester, og resultatene kan være vanskelig å sammenligne med andre data innenfor forskningen på Flynn-effekten. På den annen side kan de signalisere at det er en nedgang i gjennomsnittlig IQ også i andre land enn Danmark og Norge. Det er vanskelig å finne noen entydig forklaring på nedgangen som er observert. Teasdale & Owen (2005) antyder at minskende interesse for mer teoretisk orienterte utdanningsveier kan være en mulighet. Nederlandske data viser at rekrutter som er første-generasjons og andre-generasjons innvandrere skårer noe dårligere enn resten av rekruttene (te Nijenhuis, de Jong, Evers & van der Flier 2004). Teasdale & Owen (2008) utelukker ikke at en økende andel av innvandrere i det danske forsvaret kan ha bidratt til nedgangen. Denne faktoren kan utelukkes i Norge. I 2003 hadde bare 0.4% av de norske rekruttene innvandrerbakgrunn. I Norge er det kanskje en mer nærliggende mulighet at mindre vekt på læring av regneferdigheter i skolen kan være en medvirkende faktor.

Vi kan glede oss over at Flynn-effekten fortsetter med uforminskt styrke andre steder. Teasdale & Owen (2005) oppsummerer: "One might therefore speculate that a cessation of the Flynn Effect would now similarly be found in other developed countries. In contrast, there is recent evidence of continuing increases in intelligence in the developing world ... perhaps because, unlike in developed countries, there continues to be widespread and substantial improvement in educational levels, although ... nutritional and health care factors could also play a role. If these two opposing trends continue, then any differentials in cognitive abilities between currently developed and

developing countries ... could be expected to diminish in the future". (s. 842).

Flynn-effekten for de enkelte testene

Flynn-effekten oppfører seg ulikt for de ulike deltestene som inngår i de fleste intelligensstestene. Det typiske funnet i mange land er at de deltestene som grovt sett måler ferdigheter som læres i skolen viser *mindre* (ofte betydelig mindre) utpregede Flynn-effekter enn de man finner på tester som måler ferdigheter som ikke har noen direkte relasjon til skolefag (Flynn, 1999, 2007).

Resultatene på de tre deltestene (jfr. Sundet, Barlaug & Torjus- sen, 2004) som blir brukt på sesjon i Norge er vist i Fig 2.

Gjennomsnitsskårene på de tre deltestene øker omtrent på samme måten fram til fødselskullene tidlig i 1950-årene. Fram til omtrent midten av 1950-årene flater resultatene på Synonymprøven ut, og følges av en tendens til nedgang i gjennomsnittene i fødselskullene fra midten av 1970-årene og utover. Gjennomsnittene på Regneprøven går av ukjente grunner ned i fødselskullene tidlig i 1950-årene og fram til midten av 1950-årene. En liten økning følges av en ganske markant nedgang fra fødselskullene i midten av 1970-årene og utover. Ved slutten av observasjonsperioden er faktisk gjennomsnittet på omtrent samme som fødselskullene før krigen. Resultatene på Figurprøven utvikler seg ganske forskjellig. Der ser vi en mer eller mindre kontinuerlig stigning av gjennomsnittene helt fram til midten av 1970-årene. Deretter flater trenden ut slik

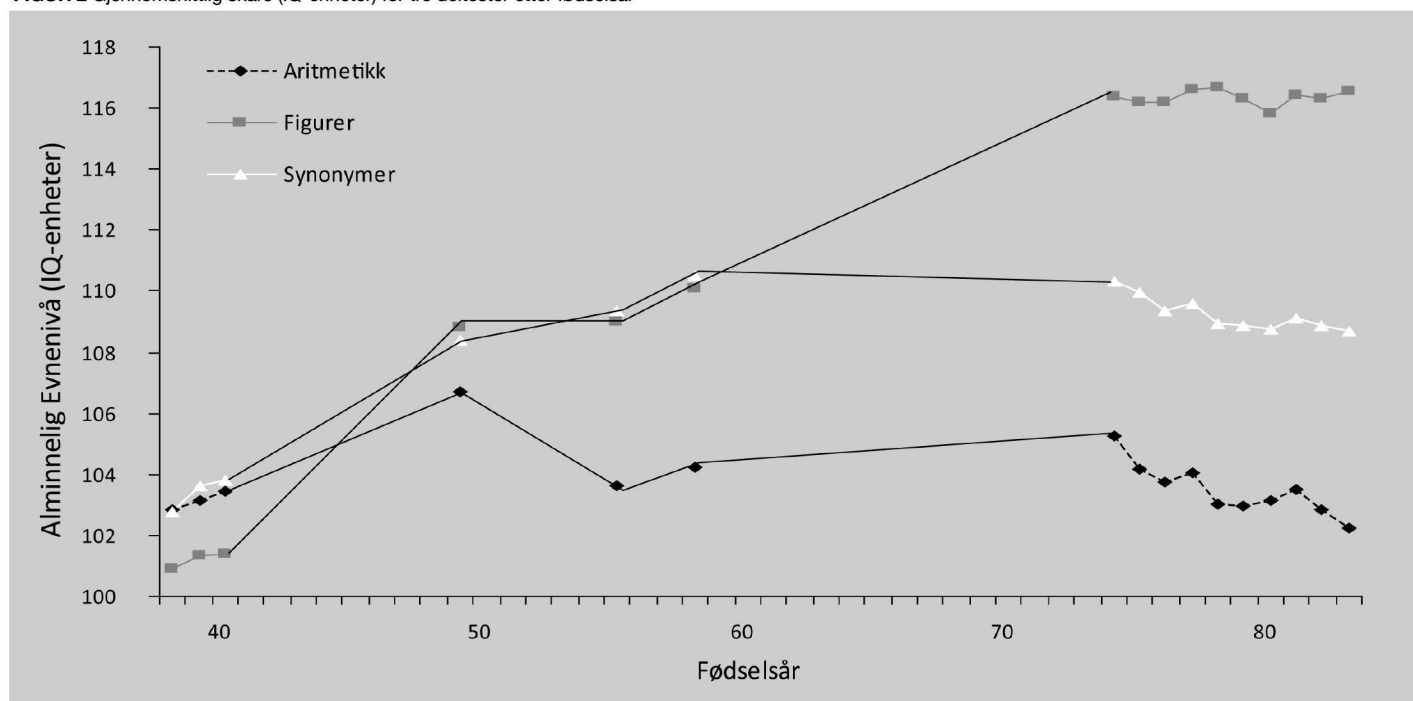
at gjennomsnittene er omtrent like for hele resten av observasjonsperioden.

I Danmark har man funnet noenlunde liknende resultater. Teasdale & Owen (2000) fant at deltesten Geometriske Figurer økte klart mest fra 1988 til 1998 (testår). De andre hadde omtrent samme stigning i gjennomsnittene, men betydelig mindre, kanskje med Bokstav-matriser et lite hestehode foran Verbale Analogi-testen og Tallsekvens-testen. I Sverige studerte de utviklingen blant elever som var 13 år da de ble testet. På verbale evner fant man en oppgang fra de som var testet i 1961 (i hovedsak født i 1948) til de som ble testet i 1966, og en liten nedgang fra 1966-1980 (testår). I perioden 1980-1990 var det en nedgang. Gjennomsnittet på tester som målte spatiale evner og resonneringsevne økte fram til 1980. Deretter var det tendenser til en utflatning og mulig nedgang (Emanuelsson, Reuterberg & Svensson, 1993).

De største forandringene har skjedd i den nedre delen av IQ-fordelingen

Forandringer i gjennomsnitt kan skje på flere forskjellige måter. Den enkleste måten er at forandringene over fødselskull er like store i den nedre enden av fordelingen av IQ-skårer som i den øvre delen. Dersom alle medlemmene i et nytt fødselskull får omtrent samme økning relativt til eldre fødselskull vil vi se en økning i gjennomsnittet uten at noe annet forandrer seg. Men gjennomsnittet kan også øke selv om det er forandringer over fødselskullene i enkelte undergrupper av befolkningen, mens det ikke skjer noen vesentlige endringer i andre under-

FIGUR 2 Gjennomsnittlig skåre (IQ-enheter) for tre deltester etter fødselsår



grupper. I det ekstreme tilfellet kan det tenkes at det blir færre og færre av dem som gjør det dårlig på testen(e), uten at det samtidig blir flere av dem som gjør det veldig bra. Da vil gjennomsnittet i hele befolkningen som helhet øke. En konsekvens av dette vil være at formen på fordelingen av IQ i en befolkning vil forandre seg, og at variasjonen i IQ vil minske. I Norge fant vi at prestasjonene for de som lå under medianen i skårefordelingen økte til dels betydelig mer enn prestasjonene til de som var over medianen (Sundet, Barlaug & Torjussen, 2004). Teasdale & Owen (2000) rapporterte tilsvarende tendenser. Også i England (Lynn & Hampson, 1986) og Spania (Colom, Lluís-Font & Andres-Pueyo, 2005) har man funnet tilsvarende tendenser. I mange andre land ser det derimot ut til at økningene i gjennomsnitt skjer omtrent likt over hele fordelingen av IQ-skårer (Flynn, 1998, 2007). Kanskje noe overraskende er det ikke funnet en like systematisk reduksjon i variasjonen i Sverige (Emanuelsson, Reuterberg & Svensson, 1993).

Avspeiler Flynn-effekten en økning i intelligens er gjelder den bare for de testene vi bruker?

Fortsatt er det ingen alminnelig enighet om en definisjon av intelligens og hvordan intelligens ytrer seg i atferd og andre observerbare prosesser (for eksempel hjernens fysiologi og anatomi). Derfor er det en mulighet for at testene delvis kan måle andre ting enn det vi er ute etter. Flynn-effekten kan kanskje rett og slett skyldes at vi blir flinkere til å ta tester (økende test-sofistikering). Kanskje er de unge nå mer villige til å gjette, og blitt flinkere til å arbeide under tidspress (Brand-hypotesen). Ettersom testene på sesjon er flervalgsoppgaver (dvs at man skal velge et av flere valgalternativer) vil man kunne tjene på å gjette i stedet for å la være å besvare spørsmålet. Disse faktorene kan ha ført til en bedring av testprestasjonene uten at det har skjedd noen forandring av gjennomsnittlig intelligens i befolkningen. Imidlertid svekkes disse hypotesene kraftig av at det er bare noen tester som viser som viser konsistente økninger over tid (Fig 2). Det har lenge vært kjent at prestasjoner på intelligenstester henger sammen med en rekke andre forhold som i og for seg ikke har noe med intelligenstester å gjøre. En av de eldste og mest robuste funnene er at prestasjoner på intelligenstester henger sammen med hvor høy utdanning man får (Colom & Flores-Mendoza, 2007) og hvordan man gjør det i arbeidslivet (Borman, Hanson & Hedge, 1997). Disse sammenhengene synes å bestå omtrent uforandret over fødselskull. Sammenhengene mellom prestasjoner på intelligenstester og suksess senere i livet er langt fra helt perfekt. Flere forhold, slik som motivasjon, interesser, og oppvekstmiljø spiller inn. Ikke med mindre viser disse funnene at intelligenstester avdekker faktorer som har betydning for en lang rekke forhold som er langt videre enn de ganske snevre områdene som intelligenstester vanligvis dekker. Nyere forskning viser at det kan være sammenhenger mellom hjerneprosesser og IQ (Gray, Chabris & Braver, 2003). Howard (1999) har vist at stormestere i sjakk er yngre nå enn de var i tidligere generasjoner. Mye forskning gjenstår for å klarlegge relasjoner mellom prestasjoner på intelligenstester og deres relasjon til prestasjoner i den "virkelige" verden.

Mulige årsaker til Flynn-effekten

Fra pre-vitenskapelig til vitenskapelig tenkemåte

Spørsmålet om årsaker til Flynn-effekten og spørsmålet om hva IQ-tester faktisk måler henger delvis sammen. Dette poenget kommer klart fram i tankene til Flynn (2007) om hvilke faktorer som kan være involvert i gjennomsnittsøkningene. Flynn tar utgangspunkt i at Flynn-effekten kan ha pågått i godt over 100 år (Raven, Raven & Court, 1993, sitert i Flynn, 2007). Dersom man tilbakeskriver gjennomsnittsøkningene 100 år vil en gjennomsnittlig prestasjon på den tiden tilsvare et gjennomsnitt på 50-70 IQ-poeng 100 år senere. Det vil si at de sammenliknet med unge mennesker i dag ville en stor del av dem framstått som mentalt retarderte. Dette er selvsagt en helt absurd konklusjon. Flynn (2007) argumenterer for at folk flest for hundre år eller mer siden tenkte annerledes enn vi gjør nå. Han kaller dette pre-vitenskapelig tenkning. På spørsmålet: "Hva har kaniner og hunder til felles?" ville mange for 100 år eller så siden kanskje svart noe sånt som: Hunder brukes til å jage kaniner. Det har ingenting med dumhet å gjøre, men er sannsynligvis mer rimelig å oppfatte som en kategorisering av verden som var funksjonell for å løse problemer i den verden mange levde i. Kategoriseringer som virker fremmedartede på oss i den vestlige verden eksisterer også i dag (Lakoff, 1987).

Den russiske psykologen Luria (referert i Flynn, 2007, s. 25) intervjuet folk i utkantene av det som var Sovjetunionen og fant påfallende lite bruk av logikk som var uavhengig av erfaring og innhold (min oversettelse):

Forskeren: I områder der det alltid er snø er alle bjørner hvite. På Novaja Semlja er det alltid snø. Hvilken farge har bjørnene der?

Respondenten: Jeg har sett bare svarte bjørner, og uttaler meg ikke om ting jeg ikke har sett.

Disse to svarene ville ikke gitt uttelling på en intelligenstest. Abstrakt tenkning i vår forstand av dette uttrykket var tydeligvis nokså fraværende. Dette er ikke et uttrykk for manglende tenkeevne, men snarere et uttrykk for en tilpasning til en ytterst konkret virkelighet hvor deres tenkemåte og kategorier formodentlig fungerte utmerket. Den vitenskapelige tenkningen er preget av andre kategorier (for eksempel at likheten mellom hunder og kaniner er at de er pattedyr), og tenkning basert på abstrakt logikk. Dette er ikke først og fremst en forskjell i evne til tenkning, men snarere *innholdet* i tenkningen. Flynn (2007) antar at en stor del av Flynn-effekten skyldes at folk blir mer og mer vant til å tenke i vitenskapelige baner. Det er kanskje derfor Flynn-effekten er størst på tester som spør etter begreper (for eksempel ordlikhetstesten i WAIS) og abstrakte resonnementer (Ravens Matriser). De andre som spør om kunnskap og innlærte ferdigheter vil ikke øke så mye. En mulig svakhet med Flynn's interessante spekulasjoner er at en skulle vente at Flynn-effekten var størst for de med minst utdanning. Det stemmer jo for Danmark og Norge samt Storbritannia, men ikke for mange andre land. Flynn's ideer bør følges opp med forskning.

Fluid og crystalized intelligence

Andre forhold kan også ha hatt betydning. Den amerikanske psykologen Raymond Cattell (Cattell, 1987) lanserte et skille mellom fluid og crystalized intelligence. Hovedideen er at vi investerer noen av våre kognitive ressurser til å lære en del ferdigheter som vi får bruk for gjennom livet (cystalized intelligence). Den reserven vi har igjen (fluid intelligence) bruker vi til å håndtere nye problemer. Figurprøven måler antakelig noe av evnen til å løse nye problemer. Det kan altså se ut som om de yngre generasjonene har stagnert eller blir dårligere på lærte ferdigheter, men er blir stadig flinkere til å håndtere nye problemer. Dette er sannsynligvis ganske funksjonelt. I den "gamle" verden var det slik at man kunne lære seg et sett med ferdigheter som stort sett var nok for å klare seg for resten av livet, mens den moderne tid er mer preget av stadige forandringer som krever evne til å tilpasse seg nye situasjoner og forholde seg til nye problemer.

Ernæring og utdanning

Bedre ernæring (som kan ha betydning for utvikling av hjernen) har vært sterkt fremhevet av noen forskere (Lynn, 1990). Det kan ha bidratt til Flynn-effekten i Kenya og andre utviklingsland, og kanskje også i industrialiserte land for ikke så mange tiår siden. Ernæringsforbedringer i Norge fra de som var født før og under siste krig til de som tilhører etterkrigs-generasjonene kan ha hatt betydning. Høyde påvirkes av ernæringsforhold, og økning i gjennomsnittshøyden kan tyde på at ernæring også kan ha påvirket skårene på intelligens-tester. I Norge stoppet økningen i gjennomsnittet i høyde i fødselskullene rundt 1950 (sesjonsårskullene rundt slutten av 1960-årene). Man kan derfor ikke utelukke at ernæringsforbedringer kan ha virket inn på den betydelige økningen av gjennomsnittene fra kullene som var født rett før krigen til kullene rett etter krigen (se Figur 1 og 2). Økningen i testskårer blant fødselskullene etter begynnelsen av 1950-årene skyldes mest sannsynlig i hovedsak andre forhold. Som nevnt tidligere skyldes Flynn-effekten i Norge og Danmark først og fremst at lavskårerne er blitt færre og færre. Det er ingen slik effekt for høyde. Økningen i høyde synes å være betydelig mer jevnt fordelt (Sundet, Barlaug & Torjussen, 2004). Dersom bedre ernæring har spilt en betydelig rolle for Flynn-effekten, skulle en forvente at høyde også skulle ha økt mest i den nedre delen av fordelingen.

Det er mer sannsynlig at forbedringene i den nedre delen kan ha en god del å gjøre med forbedringer av helsevesen, kanskje særlig bedre oppfølging av gravide kvinner og barn. Større bevissthet blant foreldre om hva som er sunt når man er gravid har sannsynligvis også hatt en betydning. Bevisst satsning på å heve nivået blant de savkeste elevene, og større interesse blant foreldrene for barnas utdanning kan også ha vært virksomme faktorer. Derimot ser det ikke ut til at teknologiske nyvinninger, slik som TV og data, har hatt avgjørende innvirkning ettersom de sterkeste økningene i gjennomsnittsskårene på intelligenstester fant sted før disse nyvinningene kom. I mange andre land har Flynn-effekten vært omtrent like kraftig blant høyskårere som på lavskårere. Det er fristende å gi "æren" for

Flynn-effekten i Norge og Danmark (sannsynligvis også i Sverige) til velferdsøkningen og utjevningspolitikken som er blitt ført i disse landene etter krigen. Det kan naturligvis diskuteres om den store satsningen i de skandinaviske sosialdemokratier på de svakere elevene i årene etter krigen har ført til at de flinke elevene er blitt noe lavere prioritert og kanskje ikke fått den utviklingen de ellers kunne ha hatt.

Flynn-effekten og familiestørrelse

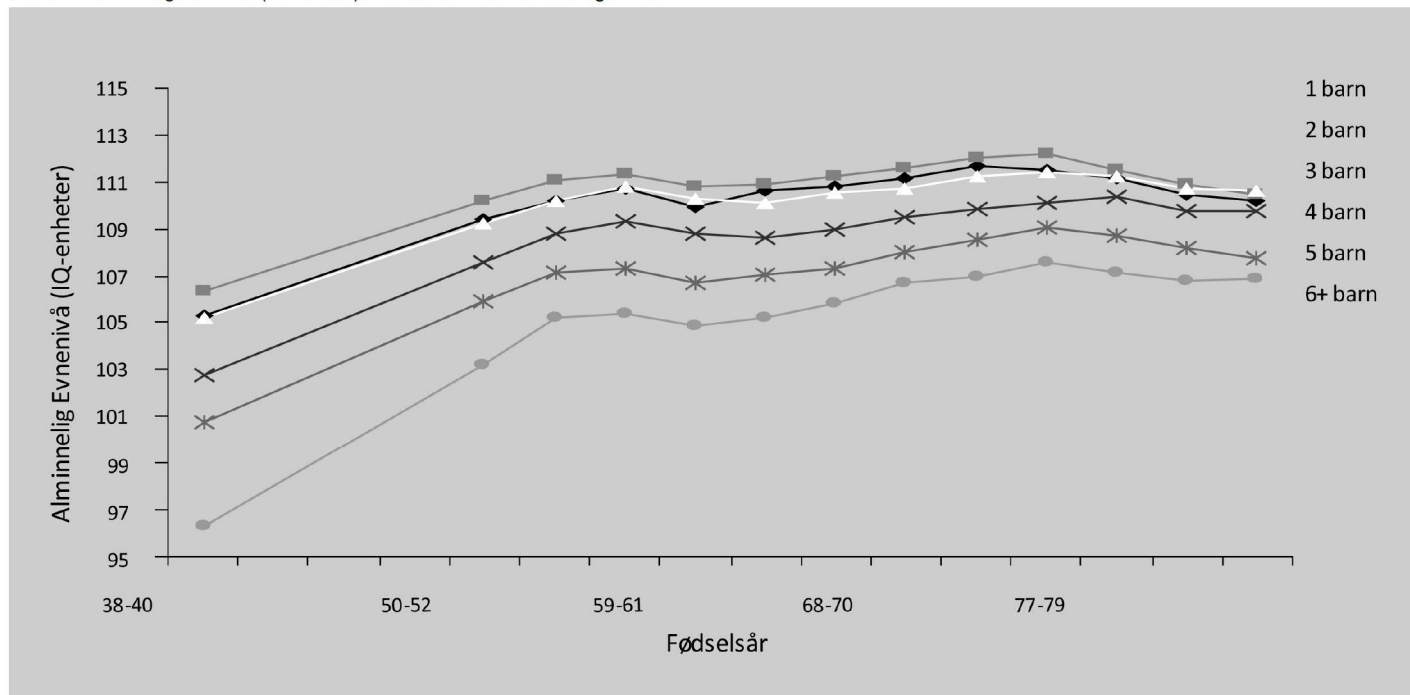
Forandringer av familiekonstellasjonene over tid kan ha hatt en betydning for utviklingen av prestasjonene på evnetester. Det har lenge vært kjent at prestasjonsnivået på tester er avhengig av familiestørrelsen. Gjennomsnittlig evnenivå i en stor søskenflokk er noe lavere enn i små søskenflokker (Belmont & Marolla, 1973). Figur 3 (jfr. Sundet, Borren & Tambs, 2008) viser at det har vært en betydelig positiv Flynn-effekt innenfor alle størrelser av søskenflokkene i Norge. Gjennomsnittlig Alminnelig Evnenivå har imidlertid økt mer i store enn i små søskenflokker. Dette har medført at forskjellene i gjennomsnittlig evnenivå mellom søskenflokker av forskjellig størrelse er blitt stadig mindre i de ferskeste fødselskullene. Det kan faktisk se ut som om at effekten av søskenflokkstørrelse på evnenivå er i ferd med å forsvinne i Norge. Årsakene til dette er sikkert mange, men at foreldre nå til dags har mer tid til barna sine og er mer opptatt av å stimulere barns vitebegjærighet er sannsynlig. Kanskje er det slik at barn i store søskenflokker har tjent mer på dette enn de som hører til i mindre søskenflokker. Sammenhengen mellom foreldrenes utdanningsnivå og antall barn har også avtatt ganske betydelig (Sundet, Borren & Tambs, 2008).

Demografiske forandringer har også hatt en betydning for Flynn-effekten i Norge. Andelen av relativt små søskenflokker har økt på bekostning av andelen av store søskenflokker. Ettersom søskenflokkstørrelsen har påvirket gjennomsnittlig evnenivå vil disse demografiske forandringene alene føre til en økning i gjennomsnittlig alminnelig evnenivå over tid. Denne effekten var størst i perioden fra fødselskullene før krigen til kullene rett etter krigen, og kan forklare rundt 20-25% av Flynn-effekten i denne perioden (Sundet, Borren & Tambs, 2008).

Kan mindre inngifte være en mulig forklaring?

De fleste forskerne har antatt at forandringene i gjennomsnittlig intelligens bare kan skyldes forandringer i miljøet og at dermed genetisk arv kan utelukkes som mulig årsak. Det er riktig at det samlede arvematerialet i en befolkning sannsynligvis ikke forandrer seg vesentlig over de såpass korte periodene hvor det finnes data på Flynn-effekten. En amerikansk forsker (Mingroni, 2004, 2007) påpekte imidlertid at fordelingen av det eksisterende arvematerialet i en befolkning kan ha forandret seg. Det er et velkjent fenomen at innavl hos arter som er arvemessig nær beslektet med mennesket tenderer til å føre til høyere dødelighet blant avkommet og en overhyppighet av sykdommer av forskjellig slag. Data på mennesker kan tyde på at inngifte over flere generasjoner kan føre til at intelligen sen går noe ned (Schull & Neel, 1965). Dersom frekvensen

FIGUR 3 Alminnelig Evnenivå (IQ-enheter) etter søskenflokkstørrelse og fødselsår



av inngifte går ned kan dermed gjennomsnitt-intelligensen øke dermed føre til en positiv Flynn-effekt. Dette fenomenet blir ofte betegnet som heterosis. Ettersom inngifte bare angår foreldrene, vil effekten av inngifte ikke synes innenfor søskenflokker. Det var en betydelig økning i gjennomsnittene i IQ i fødselskullen 1950-1956 (se fig 1), ingen vesentlige endringer i perioden 1960-1965, og en klar nedgang fra omtrent midten av 1970-årene. I følge heterosis-hypotesen skulle aldersforskjellen mellom søsken ikke bety noe for gjennomsnitt-IQ i noen av disse periodene. Dersom hypotesen er gal, vil vi derimot vente at forskjellene i gjennomsnittlig IQ skulle bli mindre med økende aldersforskjell mellom søsken i perioden 1950-1956, mens forskjellen i IQ skulle ventes å øke med økende aldersforskjell i perioden hvor gjennomsnitt-IQ i befolkningen synker. Aldersforskjellen på de som var født i perioden 1960-1965 skulle ikke påvirke IQ-forskjellen mellom søsken. Norske data (fig 4) som omfatter fødselskull fra 1950 viser at forskjellen i IQ mellom søsken som funksjon av aldersforskjell mellom dem tett følger forandringene i gjennomsnitt-IQ i befolkningen som er vist i Fig 1 (Sundet, Eriksen & Tambs, 2010). Forandringene i gjennomsnittene som er observert i Norge har dermed ingenting med inngifte å gjøre. Det er naturligvis fortsatt mulig at minskende tendenser til inngifte kan ha hatt betydning i tidligere tider og i andre land.

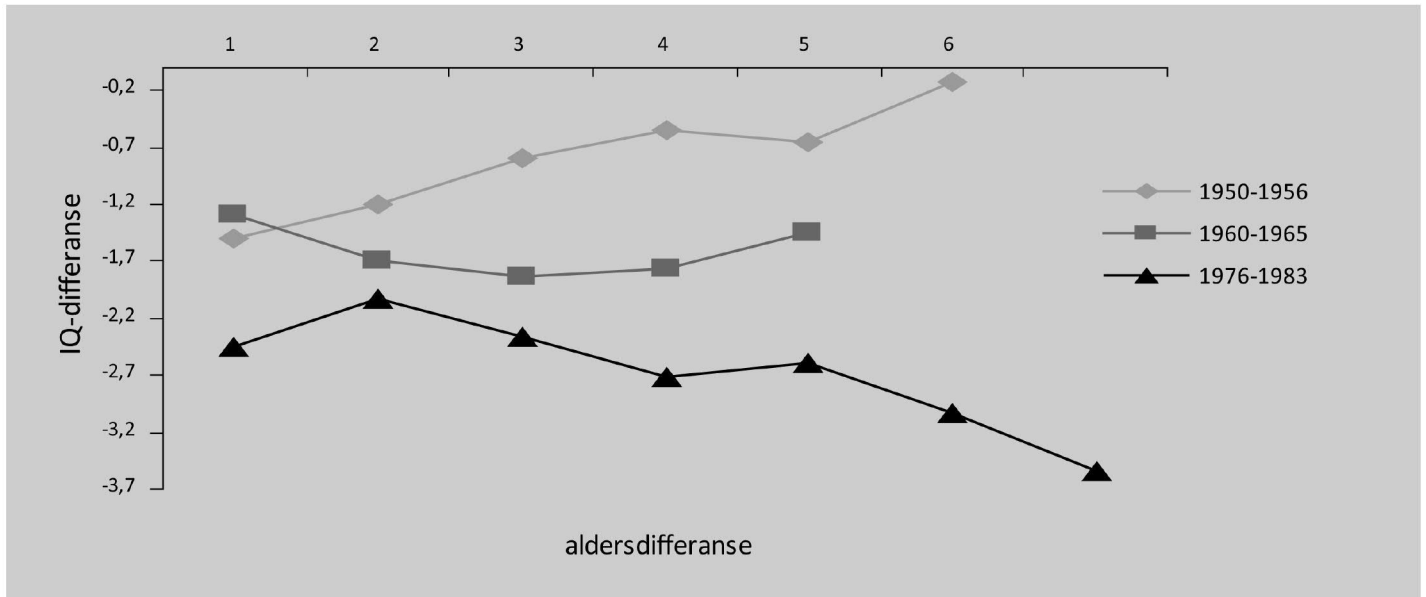
Differansene i Alminnelig Evnenivå som er vist i Fig 4 har alle negativt fortegn. Det kommer av at denne differansen er regnet ut som forskjellen mellom den av to brødre som er født senest og den som er født tidligst. Den av to brødre som er født senest har altså et noe lavere Alminnelig Evnenivå enn de

som er født tidligere. Dette er den mye diskuterte og omtalte søskenrekkefølge-effekten på intelligens (Belmont & Marolla, 1973; Bjerkedal, Kristensen, Skjeret & Brevik, 2007). Flynn-effekten og søskenrekkefølge-effekten har sjelden eller aldri blitt sett i sammenheng. Det er imidlertid en tett sammenheng, i hvert fall i Norge. I perioden 1950-1960 (fødselsår) er det en betraktelig økning i IQ-gjennomsnittene. Fig 4 viser at forskjellen i IQ mellom brødre i denne perioden avtar med økende aldersavstand mellom brødrene. Dette innebærer at de som er født senere i denne perioden tjener på Flynn-effekten, og at differansen nærmer seg null. Dersom Flynn-effekten hadde vært litt mer utpreget ville den senest-fødte broren sannsynligvis ha gjort det bedre enn den som var født tidligere, altså en reversering av søskenrekkefølge-effekten. Fig 4 viser at avstanden mellom senest-fødte og tidligst-fødte øker med økende aldersavstand i perioden hvor gjennomsnittene synker (1976-1983). Det er fortsatt en diskusjon om søskenrekkefølgen eksisterer eller ikke. Noe av forklaringen på denne usikkerheten kan ligge i at forskjellige forskere har hatt data fra tider og land hvor Flynn-effekten har vært i forskjellige faser.

Avsluttende kommentarer

Flynn-effekten synes å være et meget utbredt fenomen i et globalt perspektiv. Den er funnet både i industrialiserte samfunn og i land som er under økonomisk utvikling. Antakelig er Flynn-effekten i forskjellige faser i forskjellige land. Skandinavia, og muligens også Stor-Britannia, skiller seg ut på den måten at gjennomsnittsøkningen har stoppet opp. Muligens er det også en nedgang. I Danmark og Norge samt Stor-Britannia og Spania ser økningene i gjennomsnittene å henge sammen

FIGUR 4



med at lav-skårere blir borte. Et ganske universelt funn er også at tester som ikke er så nær knyttet til lærte skoleferdigheter øker mest.

Årsaksproblemet er ennå ganske åpent, og en rekke muligheter er blitt diskutert (Neisser, 1998). Sannsynligvis er det flere faktorer som spiller inn, og betydningen av disse kan variere over tid og sted. Ernæring kan være av betydning i land under utvikling, og har kanskje også spilt en rolle i den industrialiserte delen av verden for noen tiår siden. Demografiske forandringer har hatt noe betydning i Norge for 50 år siden, men ikke nå lenger. Det kan også være at bruken av moderne kategoriseringer (som er en stor fordel i mange intelligenstester) har økt, og at bruken av abstrakt logisk tenkning er blitt mer og mer

vanlig. Kanskje er det også slik at yngre mennesker lever i en verden som stadig er i forandring hvor det kanskje ikke er så lønnsomt å lære seg ferdigheter, men heller dyktiggjøre seg i å løse nye problemer.

Takk til det norske forsvaret og Norsk samfunnsvitenskapelig datatjeneste som har gitt tilgang til det omfattende datamaterialet på evnetester. En spesiell takk til psykolog Dag Barlaug som nedla et omfattende arbeid for å skaffe deltestdata fra sesjonene 1993-2003. Psykologene Tore Torjussen og Olav Storsve takkes for informasjon om testene og testprosedyrer. Statistisk sentralbyrå ga oss tilgang til informasjon om foreldre å søsken til de som var testet. Psykolog Dag Erik Eilertsen takkes for hjelp til å organisere datafilene.

LITTERATUR

- Ang, S., Rodgers, J.L. & Wänström, L. (2010). The Flynn Effect within subgroups in the U.S.: Gender, race, income, education, and urbanization differences in the NLSY-Children data. *Intelligence*, 38, 367–384.
- Belmont, L., & Marolla, F. (1973). Birth order, family size and intelligence. *Science*, 182, 1096–1101.
- Bjerkedal, T., Kristensen, P., Skjeret, G. A., & Brevik, J. I. (2007). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence*, 35, 503–514.
- Borman, W.C., Hanson, M.A. & Hedge, J.W. (1997). Personnel selection. *Annual Review of Psychology*, 48, 299–337.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam Elsevier.
- Colom, R. & Flores-Mendoza, C.E. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brazil. *Intelligence*, 35, 243–251.
- Colom, R., Lluís-Font, J. M., & Andres-Pueyo, A. (2005). The generational intelligence gains are caused by decreasing variance in the lower half of the distribution: Supporting evidence for the nutrition hypothesis. *Intelligence*, 33, 83–91.
- Daley, T.C., Whaley, S.E., Sigman, M.D., Espinosa, M.P. & Neumann, C. (2003). IQ on the rise. The Flynn Effect in rural Kenyan children. *Psychological Science*, 14, 215 – 219.
- Emanuelsson, I., Reuterberg, S. E., & Svensson, A. (1993). Changing differences in intelligence? Comparisons between groups of 13-year-olds tested from 1960 to 1990. *Scandinavian Journal of Educational Research*, 37, 259–276.
- Eriksen, W. Sundet, J.M. & Tambs, K. (2009). Register data suggest lower intelligence in men born the year after flu pandemic. *Annals of Neurology*, 66, 284–289.
- Flynn, J.R. (1984). The mean IQ of Americans: Massive gains from 1932 to 1978. *Psychological Bulletin*, 95, 29–51.
- Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, s. 171–191.
- Flynn, J. R. (1998a). Israeli military IQ tests: Gender differences small, IQ gains large. *Journal of Biosocial Science*, 30, 541–553.
- Flynn, J.R. (1998b). *IQ gains over time: Toward finding the causes*. In: U. Neisser (Ed.). *The Rising Curve. Long Term Gains in IQ and Related Measures*, 25–66. Washington DC: American Psychological Association.
- Flynn, J. R. (1999). Searching for justice. The discovery of IQ gains over time. *American Psychologist*, 54, 5–20.
- Flynn, J.R. (2007). *What is intelligence. Beyond the Flynn effect*. University Press. Cambridge.
- Gray, J.R., Chabris, C.F. & Braver, T.S. (2003). Neural mechanisms of general fluid intelligence. *Neuroscience*, 6, 316–322.
- Herrnstein, R.J. & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Howard, R.W. (1999). Preliminary real-world evidence that average human intelligence is really rising. *Intelligence*, 27, 235–250.
- Jensen, A.R. (1998). *The g Factor. The Science of Mental Ability*. London: Westport.
- Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind*. University of Chicago Press.
- Lynn, R., & Hampson, S. (1986). The rise of national intelligence: Evidence from Britain, Japan and the U.S.A. *Personality and Individual Differences*, 7, 23–32.
- Lynn, R. (1990). The role of nutrition in secular increases of intelligence. *Personality and Individual Differences*, 11, 273–286.
- Meisenberg, G., Lawless, E., Lambert, E., & Newton, A. (2005). The Flynn effect in the Caribbean: Generational change of cognitive test performance in Dominica. *Mankind Quarterly*, 46, 29–69.
- Mingroni, M.A. (2004). The secular rise in IQ: Giving heterosis a closer look. *Intelligence*, 32, 65–83.
- Mingroni, M.A. (2007). Resolving the IQ Paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review*, 114, s.806–829.
- Neisser, U. (Ed.). (1998). *The Rising Curve. Long Term Gains in IQ and Related Measures*. Washington DC: American Psychological Association.
- Schull, W.J. & Neel, J.V. (1965). *The effects of inbreeding on Japanese children*. New York: Harper & Row.
- Shayer, M. & Ginsburg, D. (2009). Thirty years on – a large anti-Flynn Effect? (II): 13- and 14-year-olds. Piagetian tests of formal operations norms 1976–2006/7? *British Journal of Educational Psychology*, 79, 409–418.
- Sundet, J.M., Barlaug, D. & Torjussen, T.M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, 32, s. 349–362.
- Sundet, J.M., Borren, I., & Tambs K. (2008). The Flynn effect is partly caused by changing-fertility patterns. *Intelligence*, 36, 183–191.
- Sundet, J.M., Eriksen, W., Borren, I. & Tambs, K. (2010). The Flynn effect in sibships: Investigating the role of age differences between siblings. *Intelligence*, 38, s. 38–44.
- Teasdale, T.W., & Owen, D. R. (1987). National secular trends in intelligence and education—A 20-year cross-sectional study. *Nature*, 325, 119–121.
- te Nijenhuis, J., de Jong, M. -J., Evers, A., & van der Flier, H. (2004). Are cognitive differences between immigrant and majority groups diminishing? *European Journal of Personality*, 18, 405–434.
- Teasdale, T.W., & Owen, D. R. (1989). Continuing secular increases in intelligence and a stable prevalence of high intelligence levels. *Intelligence*, 13, 255–262.
- Teasdale, T.W. & Owen, D.R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences*, 39, s. 837–843.
- Teasdale, T.W. & Owen, D.R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence*, 36, 121–126.