

Flynn-effekten i Norge og andre land: Praktiske implikasjoner og teoretiske spørsmål

psykologisk.no/sp/2018/08/e6

Gjennomsnittlig IQ i befolkningen er i endring, derfor må vi kontinuerlig sikre oss at våre intelligenstagere har oppdaterte normer, skriver Olav Storsve og kolleger.

TEKST: Olav Storsve, Jon Martin Sundet, Tore M. Torjussen og Ole Christian Lang-Ree

Flynn-effekten dreier seg om intelligenstagere og IQ-skårer. IQ er en samleskåre som beregnes med utgangspunkt i antall riktige svar på en intelligenstagere (råskårer). Tradisjonelt sett gir gjennomsnittet av råskårene en IQ på 100. Råskårer som er mindre enn gjennomsnittet, får IQ-skårer på mindre enn 100, mens råskårer som er høyere enn gjennomsnittet, gir IQ-skårer større enn 100.

Psykologer som har drevet med tester, enten som praktikere eller forskere, vet at reglene for beregning av IQ må forandres fra tid til annen. Grunnen til dette er at prestasjonene på intelligenstagere forandrer seg fra generasjon til generasjon. Det ble publisert lite om dette fenomenet før statsviteren James Flynn skrev to artikler om temaet ([1984](#), [1987](#)), og interessen tok seg da kraftig opp. Forandringene i gjennomsnittlig IQ over generasjoner blir nå ofte omtalt som Flynn-effekten.

I løpet av de siste 30 årene har det kommet hundrevis av publikasjoner om Flynn-effekten. Data fra Norge har vakt adskillig interesse internasjonalt. En av hovedårsakene er at Norge har data fra omtrent samtlige unge menn (ca. 25 000–30 000 pr. år) som var på sesjon fra slutten av 1950-årene, og fra 1969 til 2008. Hensikten med denne artikkelen er å gi en fylldig gjennomgang av hva man vet om Flynn-effekten, og mulige årsaksforklaringer, og ikke minst vil vi se nærmere på hva som har skjedd i Norge. Mye av den norske forskningen er publisert i internasjonale tidsskrifter. I denne artikkelen vil vi også presentere hittil upubliserte data. De nye dataene kan gi en antydning om hva man kan vente seg på IQ-fronten i Norge i årene fremover.

Flynn-effekten har både praktiske, etiske og metodologiske implikasjoner. Vi vil diskutere bruk av intelligenstagere i saker hvor slike inngår i grunnlaget for å ta beslutninger som kan ha store konsekvenser for de som blir berørte. Manglende kjennskap til Flynn-effekten har ført til feilaktige konklusjoner i spørsmålet omkring forskjeller i intelligens relatert til alder. Problemstillinger som berører forskjeller mellom grupper, kan belyses av forskning på

Flynn-effekten.

Utbredelse og forløp av Flynn-effekten

Økning i gjennomsnittlig IQ

Flynn-effekten er blitt grundig dokumentert i USA (Flynn, 1984) og i flere europeiske land, inkludert Norge (Flynn, 1987; Sundet, Barlaug & Torjussen, 2004). Økningene i IQ har vist seg i prestasjonene på en lang rekke forskjellige tester, inkludert Stanford-Binet, Wechsler-testene og Ravens matrisetest. Betegnelsen «Flynn-effekt» ble først foreslått av psykologen Richard Herrnstein og statsviteren Charles Murray i den berømte og til dels kontroversielle boka *The Bell Curve: Intelligence and Class Structure in American Life* (Herrnstein & Murray, 1994), og er siden blitt standard. Det var imidlertid ikke Flynn som var først ute med å vise gjennomsnittssøkninger i IQ. Lynn (2013) påpekte at det finnes minst 10 beskrivelser av dette fenomenet før Flynns publikasjoner. Den eldste av disse ble publisert av Rundquist i 1938. Lynn (2013) har foreslått at man følger vanlig skikk og bruk i vitenskap, og foreslo følgelig betegnelsen «Rundquist-effekten». Andre foretrekker betegnelsen «Lynn-Flynn-effekten» (Woodley, 2012).

Det er naturligvis vanskelig å si når Flynn-effekten inntraff.

Tuddeham (1948) rapporterte økning i gjennomsnittlig IQ (målt med Army-Alpha-testene) på 11 IQ-poeng blant amerikanske soldater fra 1917 til 1943. Engelske data på Ravens matrisetest viser økninger allerede hos fødselskull fra slutten av 1870-årene. Flynn (2007) har foreslått, noe spekulativt, at Flynn-effekten kanskje begynte i og med industrialiseringen, altså godt før de moderne IQ-testene gjorde sitt inntog. Kanskje kan opplysningstiden også ha betydd noe. Dette var som kjent en periode hvor betydningen av kunnskap og rasjonell tenkning ble sterkt framhevet, noe som nå stort sett er blitt dominerende, kanskje spesielt i den vestlige verden.

Forløpet av Flynn-effekten varierer mellom land og over tid. Williams (2013) påpekte at de aller fleste har rapportert økning i gjennomsnittene, men det er også noen som ikke har funnet en Flynn-effekt overhodet. IQ-prestasjonene på forskjellige typer tester forandrer seg på forskjellige måter. Cattell (1971) skilte mellom flytende (fluid) intelligens og krystallisert (crystallized) intelligens. Flytende intelligens viser seg i evnen til å løse nye oppgaver, mens krystallisert intelligens betegner evnen til å løse problemer som krever etablert kunnskap. I det store bildet ser det ut som om prestasjonene på tester som presumptivt måler flytende intelligens, øker mest. Ravens matrisetest er et godt eksempel på en slik type test. Internasjonale studier har vist at det også har vært en ganske stor økning i skårene på deltesten Prestasjonene på deltesten Likheter i Wechsler-testene også har vist en ganske stor økning

(Flynn, 2007). Andre deltester i WAIS som har vist mer beskjedne økninger, er Informasjon, Ordforråd og Forståelse. Regneferdighet har økt lite eller ingenting (Nisbett et al., 2012). Hvor hurtig den gjennomsnittlige IQ-en øker varierer fra land til land. De gjennomsnittlige IQ-skårene på en Raven-liknende test har økt med rundt 20 IQ-poeng blant nederlandske rekrutter i løpet av en 30-årsperiode etter 2. verdenskrig (Flynn, 1987). Gjennomsnittlig økning ligger på rundt 0,3 IQ-poeng pr. år for flere land. Med dette gjennomsnittet som utgangspunkt vil økningen i løpet av en tiårsperiode være i nærheten av 3 IQ-poeng.

Hittil er det funnet Flynn-effekter i over 30 land (Flynn, 2012). Ved siden Nord-Amerika og en rekke land i Europa, inkludert Norge, Sverige, Danmark og Finland, har det etter hvert blitt dokumentert Flynn-effekter i Israel (Flynn, 1998), Saudi Arabia (Batterjee et al., 2013), Sør-Korea (te Nijenhuis et al., 2012), Japan (Lynn & Hampson, 1986), Brasil (Colom et al., 2007), Karibia (Meisenberg et al., 2005), Kenya (Daley et al., 2003), Sudan (Khaleefa, Sulman & Lynn, 2009) og Sør-Afrika (te Nijenhuis et al., 2011). Effekten er observert både hos menn og kvinner og barn i forskjellige aldre (Ang, Rodgers & Wänström, 2010; Mackintosh, 2011). Det er gode grunner til å anta at Flynn-effekten er et fenomen som er bortimot universelt.

Økninger i gjennomsnittlig IQ kan skje på forskjellige måter, blant annet fordi gjennomsnittet øker uten at det skjer vesentlige forandringer i andre statistiske mål, slik som standardavvik (spredning) og skjevhet i fordelingen av IQ-skårer. Slike forandringer fra en generasjon til den neste får vi dersom IQ forandrer seg like mye over hele IQ-skalaen, slik at hele fordelingen flytter seg oppover. Forandringer i gjennomsnitt kan også skje ved at andre egenskaper ved fordelingen endrer seg. For eksempel vil gjennomsnittet bli høyere dersom antallet lavskårere minker fra generasjon til generasjon uten at det skjer en tilsvarende økning blant de som skårer høyt. En slik type forandring fører blant annet til at standardavviket avtar. I flere land, for eksempel USA og Estland, har forandringene stort sett skjedd over hele IQ-skalaen (Flynn, 2012, Trahan et al., 2014). Men det er mange unntak. I Norge (Sundet et al., 2004), Danmark (Teasdale & Owen, 1989), Spania (Colom, Lluís-Font & Andres-Puey, 2005) og Storbritannia (Lynn & Hampson, 1986) skyldes Flynn-effekten at lavskårere forsvinner uten at høyskårere gjør det særlig bedre. Som vi skal se senere er disse funnene en viktig innfallsport for å forstå noen av årsakene som kan ligge til grunn for Flynn-effekten.

Utflating og nedgang

Økningen i gjennomsnittlig IQ kan jo ikke være evig – den må stoppe opp før eller siden. I mange land er det ennå ikke noen tegn på at noe sånt er i ferd med å skje. Økningen i gjennomsnittlig IQ går ufortrødent videre tilnærmet lineært (Flynn, 2012; Raven, 2000). I de siste årene er det imidlertid kommet data som tyder på at vekstraten er i forandring. I blant annet Norge og Danmark (Sundet, et al., 2004; Teasdale & Owen, 2005, 2008) har vekstraten avtatt mer eller mindre kontinuerlig siden siste halvdel av 1970-årene (altså fødselskullene fra rundt slutten av 1950-årene). Blant de som ble testet rundt midten av 1990-årene og framover, har veksten stoppet opp; det har til og med vist seg tegn på nedgang i gjennomsnittene i Norge (Sundet et al., 2004) og Danmark (Teasdale & Owen, 2005, 2008). Samme fenomen er observert i flere andre land (Rindermann, Becker & Coyle, 2017). Den første rapporten om utflating og reversering av Flynn-effekten kom fra Norge (Sundet et al., 2004).

Norske data kan tjene som en illustrasjon på både økningene, utflatingen og nedgangene i gjennomsnittlig IQ. Datasettet stammer fra evnetestingen som inntil nylig ble foretatt av praktisk talt alle unge menn (18–20 år) som var på sesjon i perioden fra midten av 1950-årene til bortimot slutten av det første tiåret etter årtusenskiiftet (fødselskullene fra siste halvdel av 1930-årene til omtrent 1990/1991). Evnetestingen består av tre tidsavgrensede delprøver: Regneferdighet, Figurer og Synonymer. Regneferdighetsprøven består av oppgaver presentert som prosa, Figurprøven er en forkortet variant av Ravens matrisetest og Synonymprøven består i å finne synonymet til et gitt nøkkelord. Figurprøven og Synonymprøven har vært uforandret over hele den perioden datasettet dekker. Regneprøven ble modernisert i midten av 1960-årene, og forandret fra åpne svar til et flervalgsformat i begynnelsen av 1990-årene. Alle dataene er kodet etter normer fra 1954. Test-retest-reliabilitetene som ble beregnet i midten av 1950-årene med utgangspunkt i et utvalg på ca. 800 ungdommer, var $\pm 0,80$.

Med utgangspunkt i prestasjonene på delprøvene lages en samleskåre som gjerne betegnes som Alminnelig Evnenivå (heretter kalt IQ). Korrelasjonen mellom IQ på sesjonstestene og WAIS-IQ i et ganske lite utvalg (N = 48) var 0,75 (Tambs et al., 1988).

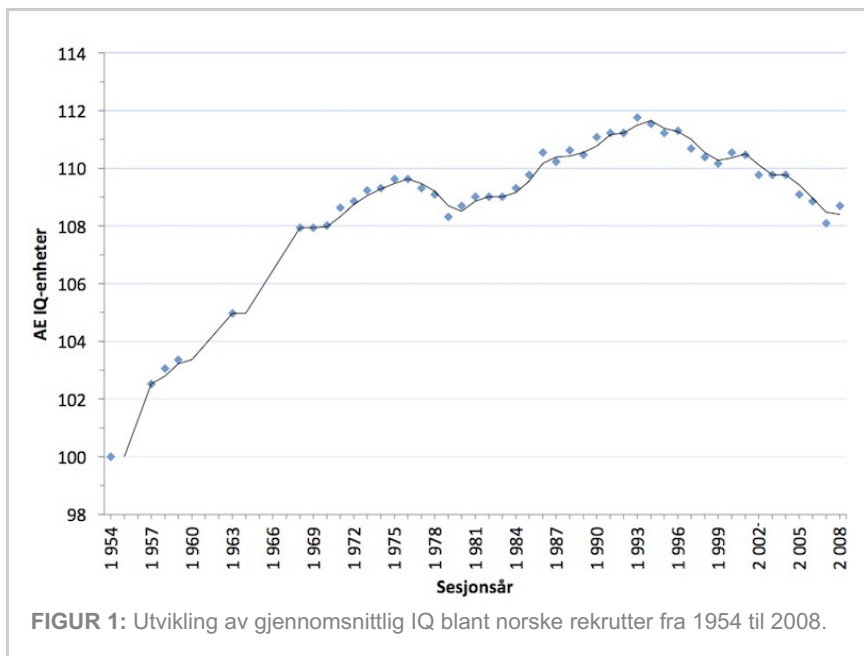
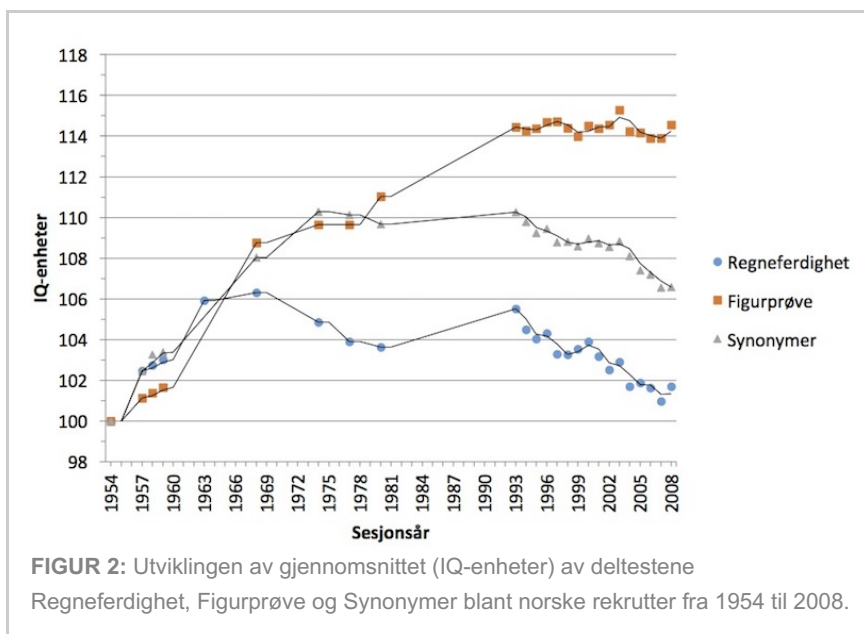


Fig. 1. viser utviklingen i IQ i de nær 50 årene som datasettet dekker. Figuren viser klart at gjennomsnittlig IQ stiger kraftig i sesjonskullene fram til midten av 1970-årene. Stigningshastigheten avtok deretter, og stoppet helt opp i midten av 1990-årene. Med unntak av Figurprøven gjelder de samme trendene også for enkelttestene (figur 2).



Det mest oppsiktsvekkende er likevel det faktum at gjennomsnittlig IQ begynte å gå nedover i Norge i sesjonsårene fra midten av 1990-årene, og nyere data viser at nedgangen i IQ fortsatte fram til og med 2008 (fig. 1). Nedgangen gjelder også for de enkelte testene, om enn i noe mindre grad for Figurprøven. Hittil upubliserte data viser hvordan gjennomsnittsskårene på deltestene har utviklet seg i årene fra 2002 til og med 2008 (fig. 2).

Det er verd å merke seg at det ikke er observert noen utflåting eller nedgang i utviklingsland (Daley et al., 2003; Meisenberg et al., 2005). Datagrunnlaget fra utviklingsland er ennå noe tynt, men dersom disse trendene forsetter, vil forskjellene i IQ mellom industrialiserte land og utviklingsland etter hvert bli mindre og antakelig forsvinne, noe som i så fall ikke er veldig overraskende. På godt og vondt vil antakelig kulturelle påvirkninger fra industrialiserte land føre til at menneskene i utviklingsland tilegner seg de ferdighetene som kreves for å klare oppgavene i «våre» IQ-tester.

Praktiske og etiske konsekvenser av Flynn-effekten

Flynn-effekten og endring av normer for hvordan IQ beregnes

Normer er i denne sammenhengen et sett med regneregler for hvordan en skal beregne IQ ut fra prestasjonene på en intelligens-test. Normering er en omfattende og tidkrevende prosess. Hovedpoenget er at en IQ-skåre plasserer en gitt persons prestasjoner relativt til andre personer. Som tidligere nevnt settes gjennomsnittet til 100. I tillegg settes standardavviket til 15. IQ angir antall standardavvik som for en gitt prestasjon ligger under eller over gjennomsnittet på 100. Flynn-effekten kan føre til komplikasjoner for denne framgangsmåten. Følgende eksempel illustrerer hvordan slike komplikasjoner kan oppstå: Du har laget en ny test og funnet at gjennomsnittlig antall riktige svar er 50. Dette tallet vil da tilsvare en IQ-skåre på 100. Et par tiår etter har gjennomsnittlig antall riktige svar økt til 60. Med de opprinnelige normene vil dette tilsvare en IQ på godt over 100. Nye normer vil føre til at en IQ på 100 nå vil tilsvare 60 riktige svar. De opprinnelige normene er følgelig «snillere» en de nye normene.

I mange land har man innført en minimumsskåre på en IQ-test for når en person oppfattes som strafferettslig tilregnelig. I Norge er personer med en IQ som er høyere enn rundt 55 (tre standardavvik under gjennomsnittet) regnet som strafferettslig tilregnelige. La oss si at retten er i tvil om hvorvidt en tiltalt er strafferettslig tilregnelig eller ikke. Bakgrunnen kan være at det er tvil om vedkommendes kognitive ferdigheter er over eller under grensen for strafferettslig tilregnelighet. Da innhenter retten gjerne en psykolog for å få svar på dette spørsmålet. Psykologen – som ikke er klar over Flynn-effekten – beregner vedkommendes IQ via en IQ-test basert på normer som er et par–tre tiår gamle. Konklusjonen er at den tiltalte har en IQ like over den nedre grensen for strafferettslig tilregnelighet, med alle de konsekvenser dette kan ha ved en eventuelt fellende dom. Ved bruk av nye normer viser det seg at personen har en IQ *under* grensen for strafferettslig tilregnelighet og derfor ikke kan straffes. Grunnen til dette spriket er at økende gjennomsnittlig IQ fra generasjon til generasjon har ført til at identiske prestasjoner (målt som råskårer) gir en høyere IQ ut fra de gamle normene enn ut fra de nye. Dersom

gjennomsnittlig IQ har avtatt, er de gamle normene «for strenge» i forhold til de nye. IQ-skårer som er basert på forskjellige normer, er altså ikke direkte sammenlignbare. I vårt hypotetiske tilfelle får det store konsekvenser for personens videre liv. Ut fra gamle normer vil vedkommende kunne havne i fengsel, mens de nye normene indikerer strafferettslig utilregnelighet.

Det kunne ha vært enda verre. I USA har man en regel som sier at lavere IQ enn 70 indikerer psykisk utviklingshemning (significant reduction in intellectual capacity). Personer med IQ lavere enn 70 kan dermed ikke dømmes til døden (Trahan et al., 2014). Flere dødsdommer er blitt avsagt basert på foreldede normer for IQ-tester. Mange har fått dødsdommer omgjort til livsvarig fengsel etter anker (hvor bl.a. Høyesterett har vært involvert) hvor Flynn-effekten er blitt trukket inn i bildet. Siden 2008 har det i USA vært en regel at Flynn-effekten og feilmålinger av IQ skal være med i vurderingen av strafferettslige reaksjoner. Bruk av korrekte normer på intelligens tester kan altså være et spørsmål om liv og død.

Det generelle poenget er at de normene som brukes, kan være avgjørende for beslutninger om blant annet strafferettslige tiltak i forbindelse med kriminalitet og spørsmål om uføretrygd og omsorgsevne overfor barn. Sivilrettslige spørsmål om intellektuell status i forbindelse med arv og annen overdragelse av formue kan også bli berørt. Feilaktige normer kan også ha konsekvenser i skolen, blant annet for beslutninger om behov for spesialundervisning.

Hva skal vi så gjøre? Første bud er å bruke tester med så nye normer som mulig. I land med en relativt markant forandring i gjennomsnittlig IQ er det en fordel med relativt hyppige renormeringer. Dersom bare relativt gamle normer er tilgjengelige, kan det kanskje være mulig å anslå hvor mye Flynn-effekten har spilt inn i løpet av de årene som er gått siden normene fastsatt. I Norge trenger vi antakelig ikke å bekymre oss for inadequate normer i de nærmeste årene. Wechsler-testene er kommet i nye versjoner og med nye normer i løpet av det siste tiåret. Med tanke på forandringshastigheten for gjennomsnittene fram til 2008 (se fig. 1) kan man antakelig bruke de eksisterende normene. Rundt 2030 kan det muligens bli aktuelt å justere IQ-skåre et par poeng opp eller ned, avhengig av i hvilken retning gjennomsnittene beveger seg.

Flynn-effekten og aldersforandringer i IQ

Forandringer i psykologiske forhold som skyldes aldring er av stor interesse både blant forskere og klinikere. Den vanligste måten å undersøke aldersforandringer på er tverrsnittsstudier, hvor man undersøker mennesker med ulik alder samtidig og tolker differensen mellom skårene for unge og eldre som et uttrykk for

aldersforandringer. Ut fra tverrsnittsstudier har man konkludert med at høyere alder betyr en kraftig nedgang i prestasjonene på intelligens tester. Sammenlikninger av mennesker i 25-års-alderen og opp til 80 + viser at eldre har en nedgang som utgjør opptil 30–40 IQ-poeng i forhold til unge ([Mackintosh, 2011](#)). Dette innebærer at personer som hadde en IQ-skåre på 100 da de var rundt 25 år, kan forvente at IQ har sunket til 60–70 når de er rundt 80 år. Disse resultatene gir et nedslående, men heldigvis misvisende bilde av den faktiske nedgangen. Grunnen er at man har brukt metoder som ikke uten videre er egnet til å studere aldersforandringer.

Tverrsnittsstudier blander sammen to ganske forskjellige forandringstyper, nemlig forandringer som henger sammen med når man er født (kohorteffekter), og forandringer som inntreffer når man blir eldre (alderseffekter). Økningen i gjennomsnittlig høyde (ca. 10 cm) siden begynnelsen av forrige århundre er en utpreget kohorteffekt. Flynn-effekten er også en ren kohorteffekt fordi forandringene skjer fra generasjon til generasjon. En effektiv måte å fastslå hvorvidt forandringene man observerer, faktisk er alderseffekter på, er at man følger opp de samme personene over tid, altså ved longitudinelle studier. Resultatene fra slike studier viser at mesteparten av de forskjellene man får når man sammenligner eldre og yngre som er testet samtidig, faktisk skyldes Flynn-effekten. I en serie longitudinelle studier har man funnet at den aldersrelaterte svekkelsen av prestasjoner på intelligens tester er nærmere 6–8 IQ-poeng ([Schaie, 1995](#)). Sagt litt flåsete: Den gode nyheten er at man ikke blir så mye dummere når man eldes. Den dårlige nyheten er at de som er litt opp i årene, har vært dummere hele tiden.

Det kan også hende at aldersforandringene er forskjellige i forskjellige kohorter. For å finne ut av dette må man studere aldersforandringer i flere kohorter. Slike studier og rene longitudinelle studier er kostbare og tar tid, og frafallet underveis er som regel betydelig.

Mulige årsaker til Flynn-effekten

Gjetting og tilvenning

Et naturlig spørsmål er om Flynn-effekten gjenspeiler forandring av ferdigheter som bare gir utslag på intelligens tester. Det kan være at folk blir mer og mer vant med alskens tester fra generasjon til generasjon. Intelligenstesting var nesten ukjent for vel 100 år siden, og det kan tenkes at tilvenning til testsituasjonen og testformatet kan ha ført til en bedring av prestasjonene i tidlige faser av Flynn-effekten. Testing har blitt mer vanlig, og det virker usannsynlig at øvelse og tilvenning har hatt noen særlig effekt de siste tiårene. Forbedringer fra test til retest er omtrent 5 IQ-poeng, mens økningen fra generasjon til generasjon i mange land utgjør omtrent 30 IQ-

poeng i løpet av 30–40 år. I tillegg er det meget usannsynlig at øvelse og tilvenning kan forklare nedgangen i gjennomsnittlig IQ i Norge og andre land de siste 10–20 årene.

Mange intelligenstagere inneholder oppgaver hvor det riktige svaret er et av flere mulige alternativer, altså muligheten for å gjette blant et gitt antall alternativer. Dersom man gjetter riktig på en oppgave i en intelligenstagere får man belønning i form av plusspoeng. Gale svar som skyldes gjetting eller andre årsaker, blir derimot ikke straffet med minuspoeng. Kanskje er det slik at tilbøyeligheten til å gjette når man ikke vet hva svaret er, har økt fra generasjon til generasjon, og dette har ført til en økning i prestasjonene på intelligenstagere (Brand-hypotesen). Denne hypotesen kan ikke forklare de norske funnene. For det første ble formatet til en av testene som blir brukt på sesjon (Regneferdighet), forandret fra åpne svar til valg blant flere mulige svar på begynnelsen av 1990-tallet, uten at dette medførte merkbare forandringer i skårene på denne testen. For det andre kan heller ikke Brand-hypotesen forklare nedgangen i gjennomsnittlig IQ i løpet av de siste årene. Tilvenning og gjetting kan neppe forklare at de gjennomsnittlige prestasjonene på noen typer tester øker vesentlig mer enn på andre typer tester (fig. 2).

Var våre formødre og forfedre stakk dumme?

Ekstrapolering av Flynn-effekten bakover i tid er tankevekkende (og risikabelt). Går vi 100 år tilbake i tid, vil vi måtte konkludere med at en prestasjon som i 1918 tilsvarte 100 IQ-poeng, i 2018 vil tilsvare en prestasjon som bare vil gi 70 IQ-poeng, altså hele to standardavvik. En normalprestasjon i 1918 tilsvarer altså en prestasjon i nærheten av psykisk utviklingshemning i 2018! Går vi enda 50 år tilbake i tid, vil en normalprestasjon på 100 i 1860-årene tilsvare rundt 55 IQ-poeng på moderne IQ-tester. De fleste av oss vil mene at det er absurd å påstå at mange av våre nære forfedre var psykisk utviklingshemmede.

For selvfølgelig var det ikke slik! Våre forfedre hadde heller måter å tenke på som nok ville ha gitt dårlig uttelling på intelligenstagere. Som tidligere nevnt gjør Flynn-effekten seg aller mest gjeldende på tester som måler flytende intelligens samt kategorisering av organismer og objekter (som måles med deltesten Likheter i Wechsler-testene). Ravens matrisetest er en av de mest brukte som mål på flytende intelligens, og det har vært en vanlig antakelse at denne testen er kulturfri eller i det minste kultureddusert. Et underliggende premiss for denne tankegangen er at abstrakt logisk tenkning er uavhengig av kultur. Flynn (2007, 2012) og Mackintosh (2011) trekker forøvrig fram noen fascinerende funn gjort av den russiske nevropsykologen Alexander Luria i Sentral-Asia for omtrent 80 år siden, som tyder på at dette premisset kan være feilaktig.

Luria prøvde å finne fram til tester på resonneringsevne som passet for bønder som var analfabeter. Følgende er et eksempel på syllogismer som ble prøvd (fritt oversatt fra [Mackintosh, 2011](#), se også [Flynn, 2007, 2012](#)):

Der hvor det alltid er snø, er alle bjørner hvite. På Novaja Semlja er det alltid snø. Hvilken farge har bjørnene på Novaja Semlja?
Svar: Hvordan kan jeg vite det? Jeg har aldri vært så langt mot nord.

For mange vil det være overraskende at respondenten ikke ser den logiske implikasjonen som ligger i spørsmålet, men snarere viser til at hun/han ikke har noen egne erfaringer med bjørner i nord. Respondentene var tydeligvis ikke vant til å forholde seg til hypotetiske situasjoner som denne. Kanskje er det slik at økningen i gjennomsnittene på Raven-testen og Raven-lignende tester nettopp avspeiler et mentalitetsskifte fra konkret, stedbundet tenkning til økende utbredelse av abstrakt tenkning ([Flynn, 2007, 2012, 2016](#)).

Prestasjoner på deltesten Likheter (dessverre har vi ikke norske data på Flynn-effekten for denne deltesten) viser en massiv økning, faktisk helt på linje med økningen på Ravens matrisetest og andre tester som måler flytende intelligens ([Flynn, 2007, Mackintosh, 2011](#)).

Kategoriseringer har tydeligvis endret seg kraftig over tid, noe Lurias funn bekrefter. Følgende er et «odd one out»-eksempel som Luria benyttet ([Mackintosh, 2011](#)):

hammer, sag, øks, tømmerstokk

I henhold til den moderne måten å klassifisere på er det åpenbart at det er «tømmerstokk» som skal ut. Slik var det ikke for bøndene i Sentral-Asia:

Alle hører sammen. Saga og øksa trenges for å hugge ved.

Hammeren trengs for å hamre på tømmerstokken.

Når de fikk opplyst at noen hadde sagt at tømmerstokken ikke hørte med, var svaret gjerne:

Han har sannsynligvis masse ved. Men det har ikke vi.

Når de ble spurt om likheten mellom f.eks. en hare og en hund, var det typiske forslaget at hunden jager haren ([Flynn, 2007](#)).

Tilsvarende måter å kategorisere på finner man også i andre folkegrupper, f.eks. blant Kpelle-folket i Liberia. ([Mackintosh, 2011](#)).

Vi trenger antakelig heller ikke å gå så mange generasjoner tilbake før vi finner liknende måter å tenke på i vår del av verden ([Flynn, 2007](#)).

Kategoriseringssystemer er naturligvis ikke hugget i stein.

Eksemplene viser et kategorisystem som er mer handlingsorientert enn det taksonomiske klassifiseringssystemet vi har en tendens til å ta for gitt. Det taksonomiske system for kategoriseringer har oppstått

som følge av vitenskapelige framskritt i blant annet biologi og botanikk. I «gamle» dager kategoriserte folk planter ut fra hvorvidt de kunne spises eller ikke, og om de kunne brukes som medisin. Flynn (2016) kaller bruk av logikk og moderne kategorisystemer for «vitenskapelig» tenkning, og hevder at overgangen til vitenskapelig tenkning kan ha vært av vesentlig betydning for tidligere økning av gjennomsnittlig IQ. De markante økningene i IQ-gjennomsnittene i Kenya (Daley et al., 2003) og andre utviklingsland, som Den dominikanske republikk (Meisenberg et al., 2005), kan antakelig i noen grad skyldes en tilnærming til tenkemåtene som nå er dominerende i den industrialiserte delen av verden.

Kanskje kan overgangen til vitenskapelig tenkning ha vært en medvirkende årsak til de markante økningene i skårene på Figurtesten blant norske rekrutter fra generasjonene før den siste verdenskrigen til generasjonene etter krigen. De store forandringene i den nedre delen av IQ-fordelingen kan faktisk tyde på at noe sånt er tilfellet.

Demografi og familie

Gjennomsnittlig IQ i en befolkning kan forandre seg av rent demografiske årsaker. Dersom det finnes flere subgrupper innenfor en befolkning med forskjellig gjennomsnittlig IQ, vil gjennomsnittet i hele befolkningen være en funksjon av gjennomsnittet i de forskjellige subgruppene samt andelen av befolkningen som hver subgruppe utgjør. La oss anta at en befolkning består av to subgrupper. I den ene subgruppen er gjennomsnittlig IQ 100, og den utgjør 60 % av befolkningen, mens den andre subgruppen utgjør de resterende 40 %, og har en gjennomsnittlig IQ på 90. Hele befolkningen har da en gjennomsnittlig IQ på 96 ($0.60 \cdot 100 + 0.40 \cdot 90$). Forandrer de relative andelene seg, vil gjennomsnittlig IQ i hele befolkningen forandre seg.

Det har lenge vært kjent at gjennomsnittlig IQ i store søskenflokker er lavere enn i små søskenflokker (Anastasi, 1956). Den norske sesjonsdatabasen inneholder unike data vedrørende forandring i størrelsen på søskenflokker og forandringer i gjennomsnittlig IQ over generasjoner. Dette gjør det mulig å undersøke hvordan gjennomsnittlig IQ forandrer seg fra generasjon til generasjon som funksjon av at andelen av barn i de forskjellige søskenflokker forandrer seg. I Norge har størrelsen på søskenflokker stort sett vært avtagende i løpet av det forrige århundret. Gjennomsnittlig IQ forandret seg over generasjoner innenfor alle søskenflokkene, og mest i de store søskenflokkene. Omtrent 20 % av økningene i gjennomsnittlig IQ fra sesjonsårskullene fra 1950-årene til midten av 1990-årene kom utelukkende av at antall barn i små søskenflokker

økte på bekostning av antall barn i store søskenflokker. Disse forandringene var klart størst for fødselskullene fra før 2. verdenskrig til fødselskullene etter krigen. ([Sundet, Borren & Tambs, 2008](#)).

Korrelasjonen mellom søskenflokkstørrelse og IQ førte til at Francis Galton for godt over 100 år siden bekymret seg for såkalte dysgeniske trender. Slike bekymringer har forekommet med jevne mellomrom helt fram til i dag ([Lynn & Harvey, 2008](#)). Argumentet er omtrent slik: Forskjeller i IQ er delvis genetisk betinget. Dersom det fødes flest barn i deler av befolkningen som har lav IQ, vil dette på sikt føre til uheldige forandringer i arvemassen og at gjennomsnittlig IQ i befolkningen dermed synker. Flynn-effekten (økende gjennomsnitt) viser den motsatte trenden. Enten eksisterer det ingen dysgenisk trend, eller så er størrelsen av Flynn-effekten underestimert. De norske dataene viser at forskjellene i gjennomsnittlig IQ mellom familier av forskjellige størrelser er i ferd med å forsvinne ([Sundet et al., 2008](#)). Korrelasjonen mellom fedres IQ og antall barn er praktisk talt null ([Sundet, 2014](#)). Om det noen gang har vært grunn til bekymring for dysgeniske trender, så er det ingen grunn til denne bekymringen lenger, i hvert fall ikke i Norge.

Det er flere fordeler med å vokse opp med få søsken; mer fritid for folk og gunstige permisjonsordninger for småbarnsforeldre har gitt mer tid til barna. Forandring i oppdragelsesstiler og større bevissthet på viktigheten av å kommunisere med og stimulere barna har økt ([Bronfenbrenner & Ceci, 1994](#)). Stadig bedre barnehagetilbud kan ha økt tilgangen på gunstig stimulering ([Hartmann, 1991](#)).

Utdanning, ernæring og helse

Økende antall skoleår og bedre skoler er uten tvil faktorer som har bidratt til økende gjennomsnittlig IQ. Sannsynligvis har mer og bedre utdanning vært med på å skape forbedringer av resultatene på tester som måler krystallisert intelligens, slik som ordforråd og informasjon. Kanskje er også overgangen til vitenskapelig tenkning delvis en effekt av bedre utdanning. Lynn ([2009](#)) peker på at Flynn-effekten også gjelder førskolebarn, og at dette fenomenet svekker hypotesen om at utdanning bidrar til Flynn-effekten. Dette er riktig dersom man begrenser begrepet utdanning til formell skolegang. Utdanning i videre forstand omfatter også faktorer som foreldre, barnehage og tilgang på kunnskap gjennom TV, radio og bøker. Studier i Skottland viste en tydelig Flynn-effekt i en periode hvor det var små eller ingen forandringer i skolesystemet ([Mackintosh, 2011](#)).

Ernæring, stimulering og sykdomsbelastning har betydning for utviklingen av nervesystemet. Det råder uenighet om hvorvidt bedre ernæring og bedre helse generelt kan være en forklaring på Flynn-effekten. En vanlig måte å undersøke betydningen av ernæring og helse på er å bruke utviklingen i gjennomsnittlig høyde som en

indikator. Vi vet at gjennomsnittlig høyde har økt betydelig i mange industrialiserte land i løpet av det forrige århundret. I Norge dreier det seg om en økning på ca. 10 cm (Bore, 2007). Noen hevder at bedre ernæring og bedring i helse har vært avgjørende årsaker til Flynn-effekten (Lynn, 2009), mens andre mener at denne forklaringen ikke kan være spesielt viktig (Flynn, 2012). Det som er sikkert, er at ernæring ikke har hatt noen vesentlig betydning for forandringene i IQ i Norge i etterkrigs-generasjonene. Flynn-effekten fortsatte lenge etter at økningene i gjennomsnittlig høyde stoppet opp. Ettersom Flynn-effekten er mest markant i den nedre delen av IQ-spekteret, kunne man vente større økning i den nedre delen av fordelingen av høyde enn i den øvre delen, men det er ingen tegn på at dette har skjedd, i hvert fall ikke i fødselskullene etter siste krig (Sundet et al., 2004). Kanskje er det ikke ernæring og helse i seg selv som spiller inn. Fattigdom og stor sykdomsbelastning kan ha ført til at «kroppens visdom» prioriterer andre ferdigheter enn de kognitive.

Mulige genetiske faktorer i økningene i gjennomsnittlig IQ

Den genetiske sammensetningen i en befolkning kan ikke ha endret seg i det tidsrommet hvor Flynn-effekten har blitt observert, men *fordelingen* av gener i befolkningen kan ha forandret seg. Dette kan skje når giftemålspraksisen i en befolkning forandrer seg. Inngifte over flere generasjoner kan føre til at gjennomsnittlig IQ i en befolkning går ned (Bashi, 1975). Dersom frekvensen av utgifte øker på bekostning av inngifte, vil man følgelig forvente en økende IQ i befolkningen. Forbedring av egenskaper gjennom utgifte kalles gjerne heterosis. Potensielt kan altså økningene i gjennomsnittlig IQ skyldes forandring av giftemålmønsteret i en befolkning. Mingroni (2004, 2007) hevdet at en vesentlig andel av Flynn-effekten skyldes heterosis. Denne påstanden kan ikke være helt korrekt – beregninger viser at selv om alle ekteskap i 1900 var inngått mellom bror og søster, ville det ikke kunne forklare hele Flynn-effekten (Flynn, 2007), men det er fortsatt naturligvis mulig at heterosis kan ha bidratt sammen med mange andre faktorer. Data om søsken kan utelukke faktorer som varierer mellom familier (mellom-familie-faktorer), men ikke innenfor familier. Ettersom utgifte bare angår foreldre, og dermed er en mellom-familie-faktor, vil vi vente at barna deres ikke blir påvirket forskjellig, og yngre søsken vil dermed ikke forventes å ha høyere IQ enn eldre søsken. Dette har vi undersøkt ved å sammenligne gjennomsnittlig IQ hos brødre født på forskjellige tidspunkt med den gjennomsnittlige IQ i befolkningen generelt (Sundet, Eriksen & Tambs, 2010). Trenden i de forskjellige fasene av Flynn-effekten viste nøyaktig de samme trendene innenfor brødre-flokker som Flynn-effekten i den generelle befolkningen. Heterosis har dermed *ikke* bidratt til Flynn-effekten, i hvert fall ikke i Norge siden fødselskullene fra begynnelsen av 1950-årene.

Bratsberg og Røgeberg (2018) påpeker at dette resonnementet kan generaliseres til andre mellom-familie-faktorer, som f.eks. etnisitet og sosial klasse, og at endringene i gjennomsnittlig IQ i Norge i løpet av de siste tre–fire tiårene utelukkende skyldes faktorer som varierer innenfor familier.

Nedturen: Implikasjoner og mulige årsaker

Utflatingen og til dels nedgangen i gjennomsnittlig IQ er blitt grundig dokumentert i mange land. Flere mener at nedgangen skyldes en nedgang i «egentlig» intelligens. Et hovedargument er at de miljøfaktorene som har drevet Flynn-effekten, etter hvert har sluttet å ha innvirkning og da vil den «rene» intelligensen komme fram. Årsaken til nedgangen er ifølge dette resonnementet at Flynn-effekten har kamuflert nedgangen i intelligens som har pågått i mange generasjoner. Et kuriøst innspill er at mennesker som levde i Victoriatiden i Storbritannia, var gjennomgående mer intelligente enn nålevende (Woodley, te Nijenhuis & Murphy, 2013). Påstanden er at dysgeniske trender er hovedårsaken til nedgangen. Denne forklaringen ser ikke ut til å stemme for norske forhold. Som tidligere nevnt er det liten eller ingen sammenheng mellom foreldrenes IQ og antall barn i familien i den perioden nedgangen foregår (se fig. 1 og 2). Følgelig er såkalte dysgeniske trender utelukket som forklaring på nedgangen i Norge.

I kretsen rundt berømte og kontroversielle forskere som Richard Lynn og Arthur Jensen og J. Philippe Rushton, er det mange som hevder at det er betydelige forskjeller i intelligens mellom etniske grupper, og at disse forskjellene i betydelig grad skyldes arv (Rushton & Jensen, 2005). Blant disse råder det en bekymring for at innvandring fra ikke-vestlige land kan senke gjennomsnittlig IQ i industrialiserte land. Dette kan ikke stemme for land med lite innvandring fra ikke-vestlige land. Finland er et eksempel på en slik nasjon (Dutton & Lynn, 2013).

La oss kikke litt nærmere på innvandringsargumentet for Norges del. Upubliserte resultater fra sesjonsdatabanken viser at IQ-gjennomsnittet hos rekrutter fra den tredje verden er omtrent 5 IQ-poeng lavere enn hos rekrutter fra vestlige land. Bore (2007) har beregnet at andelen innvandrere er ca. 10 % av befolkningen. Dersom vi setter gjennomsnittet for etnisk norske og mennesker fra andre vestlige land på 100 og ikke-vestlige på 95, er det nye gjennomsnittet 99.5. Dette er bare en liten del av den totale nedgangen, som er rundt 4 IQ-poeng fra midten av 1990-tallet til og med 2008. En hypotetisk innvandrerandel på 20 % vil senke gjennomsnittet med rundt 1 IQ-poeng, altså er det betydelig mer grunn til å bekymre seg over en nedgang på 4 IQ-poeng i befolkningen generelt. Resultatene fra de ovennevnte studiene av

brødre svekker muligheten for at nedgangen i gjennomsnitt har noe med innvandring å gjøre ytterligere (etnisitet er en mellom-familie-faktor). Forklaringer på reverseringen av Flynn-effekten som bygger på genetikk, er dermed utelukket, i hvert fall i Norge. Altså skyldes nedgangen utelukkende miljøforhold ([Bratsberg & Rogeberg, 2018](#)).

Vi tror at nedgangen i noen grad kan skyldes mer hverdagslige forhold. Nedgangen i regneferdigheter skyldes antakelig, i hvert fall delvis, at tilgangen på elektroniske regnemaskiner og PC har økt voldsomt. De unge i dag trenger i mye mindre grad hoderegning og regning med papir og blyant. Dermed oppleves slike «gammeldagse» ferdigheter som unyttige og irrelevante. Synonymtesten er litt over 60 år gammel, og ord kan ha forandret mening underveis.

Hva er det som forandrer seg?

Flynn-effekten og intelligens

Intelligenstester måler selvsagt ferdigheter, og Flynn-effekten dreier seg om kohortforandringer i de ferdighetene som inngår i intelligenstester. Et spørsmål som naturlig dukker opp er om økning i skårer på IQ-tester også har resultert i økninger i den latente variabelen intelligens. Problemet vi da støter på, er at de aller fleste definisjoner av intelligens er veldig abstrakte og ikke kan brukes direkte for å gjøre empiriske undersøkelser om forholdet mellom IQ-økninger og økninger i intelligens ([Nisbett et al., 2012](#)). I praksis har derfor forskere, praktikere og «folk flest» ofte brukt begrepene «intelligens» og «IQ» som synonymer.

En vanlig strategi for å klarlegge underliggende egenskaper for prestasjonene på intelligenstester har vært å undersøke relasjoner mellom tester som brukes til å måle IQ. Et robust og veldokumentert funn er at tester med høyst forskjellig innhold korrelerer med hverandre. Folk som gjør det bra på f.eks. Ravens matrisetest, gjør det gjerne også bra på WAIS. Deltestene i Wechsler-testene korrelerer også med hverandre. Faktoranalyse av korrelasjonsmønstre mellom tester resulterer ofte i en generell faktor samt en del spesifikke evner. Den generelle faktoren blir ofte benevnt som g-faktoren eller *g*. Jensen ([1998](#)) argumenterte for at g-faktoren som man finner i faktoranalyse, er uttrykk for en underliggende enhetlig psykologisk og biologisk realitet. Flere forskere har påpekt at g-faktoren kan oppstå selv om det finnes flere evner som er uavhengige. En mulighet er at de uavhengige evnene er gjensidig forsterkende. Det kan også være at personer med relativt gode evner på ett område lett kan ende opp i sosiale omgivelser som fører til at de utvikler en rekke andre ferdigheter. En tredje mulighet er at leddene i IQ-tester måler flere forskjellige evner. Dersom dette er tilfellet, vil leddene korrelere positivt med hverandre

og gi opphav til en g-faktor i faktoranalyse (Nisbett et al, 2012). En stor fordel med ideen om evner som er uavhengige av hverandre, er at hver av de enkelte evnene kan forandre seg uavhengig av hva som skjer med andre evner. Et av de mest robuste funnene fra studier av Flynn-effekten er jo nettopp at prestasjonene på noen tester forandrer seg mer enn prestasjonene på andre tester. Dessuten er det funn som tyder på at viktigheten av g-faktoren kan ha avtatt over tid (Kane & Oakland, 2000; Sundet et al., 2004). Kanskje dette kan henge sammen med at spesifikke ferdigheter («spisskompetanse») har blitt stadig mer ettertraktet på bekostning av generell kompetanse.

Flere forskere har undersøkt om Flynn-effekten kan føres tilbake til forandringer i g-faktoren fra generasjon til generasjon. Ved hjelp av faktoranalyse kan man finne ut hvor mye hver enkelt deltest har til felles med g-faktoren. Da kan man studere korrelasjonen mellom hvor mye enkelttester har felles med g-faktoren og hvor mye prestasjonene på disse testene forandrer seg fra generasjon til generasjon. Dersom Flynn-effekten er størst for de testene som har mest felles med g-faktoren, er det et argument for at g-faktoren forandrer seg over tid. Flere studier tyder på at korrelasjonene er lave og til dels negative (te Nijenhuis & van der Flier, 2013). Forklaringen på dette funnet er kanskje at g-faktoren oftest er basert på generelle intelligens tester av typen WAIS. Slike tester inneholder gjerne et flertall av deltester for krystallisert intelligens som viser beskjedne forandringer fra generasjon til generasjon (Flynn, 1999). Dersom flere tester som måler flytende intelligens blir inkludert i faktoranalysene, vil tester for flytende intelligens ha mer felles med en generell faktor, og korrelasjonen mellom forandringer i IQ over generasjoner og hvor mye hver enkelt test har felles med en generell faktor kan bli positiv (Colom, Juan-Espinosa & Garcia, 2001). Men korrelasjonene man har funnet, har vært ganske lave. Flynn-effekten kan sannsynligvis ikke fullt ut forstås ut fra en økning i generell intelligens over generasjoner. Wicherts og medarbeidere (2004) fant at det sannsynligvis er forskjellige underliggende prosesser involvert i Flynn-effekten i forskjellige perioder. Dette funnet svekker g-faktorens betydning for Flynn-effekten ytterligere.

Det foreligger (minst) tre forslag til underliggende prosesser som kan gi opphav til en g-faktor: 1. Hvor hurtig man behandler informasjon, 2. forbruk av energi og 3. arbeidshukommelse (working memory). En indikasjon på hurtighet av informasjonsbehandling er reaksjonstid. Et klassisk oppsett for å finne reaksjonstiden er å måle hvor lang tid det tar fra et lys (eller en lyd) blir slått på til forsøkspersonen trykker på en knapp. Energibruk blir blant annet målt som glykoseomsetning i hjernen. I studier av arbeidshukommelse er oppgaven å håndtere flere oppgaver samtidig.

Resultatene har vist moderate korrelasjoner. Reaksjonstid korrelerer rundt $-0,30$ (større hurtighet ved høyere IQ). Resultatene fra studier av energiforbruk er i samme størrelsesorden (men med positivt fortegn). Kanskje er arbeidshukommelse den beste kandidaten – der ligger korrelasjonene i nærheten av $0,50$ (Mackintosh, 2011). En ulempe med de fleste studiene på arbeidshukommelse er at intelligens har blitt målt med Ravens matrisetest eller liknende tester. Samlet er resultatene fra disse studiene på langt nær gode nok til å fullt ut forklare forskjellene i g , men kanskje kan de sammen bidra til å forklare noen av disse forskjellene. En mulig modell er at forandringer i g -faktoren, sammen med forandringer i mer spesifikke evner (som flytende og krystallisert intelligens), er involvert.

Moderne teknikker, særlig funksjonell magnetresonanstomografi (fMRI) og computertomografi (CT), har ført til en rekke studier av strukturer og prosesser i hjernen. Resultatene fra slike studier er ennå inkonsistente og av og til motstridende (Nisbett et al., 2012). Blant de mest etablerte funnene er at noen hjerneområder er mer aktive i problemløsningssituasjoner enn andre hjerneområder, og områder i pannelappen skiller seg ut. Lesjoner i dette området ser ut til å redusere prestasjonene på tester som måler flytende intelligens, mens krystallisert intelligens forblir relativt uberørt. Cattells (1971) skille mellom flytende og krystallisert intelligens er blitt styrket av disse resultatene. Det har også blitt påvist aktivitet i isselappen under problemløsning. Jung og Haier (2007) konkluderte i en metastudie med at det finnes forbindelser mellom frontallappen og isselappen. Hittil er det ikke påvist noen områder i hjernen som kan tenkes å ha noe med g å gjøre, men det betyr jo ikke nødvendigvis at et slikt område ikke finnes.

Hva med andre kognitive ferdigheter?

Intelligenstestene forteller ikke hele historien om kognitive ferdigheter. Sternberg (1985) påpeker at de vanlige intelligenstestene ikke er særlig egnet til å måle praktisk intelligens og kreativitet, mens Gardner (1983) foreslår en radikal utvidelse av intelligensbegrepet – til å omfatte ferdigheter som ikke har så mye med kognitive ferdigheter å gjøre, som f.eks. musikalitet og motorikk. Kognitive psykologer, med nobelprisvinneren i økonomi Daniel Kahneman (og hans avdøde kollega Amos Tversky) i spissen, har påvist at vi har en tendens til å overse tilgjengelig og relevant informasjon når vi skal vurdere sannsynlighet for utfall i usikre situasjoner (Kahneman, 2012). Istedenfor å analysere situasjonene ut fra relevant og tilgjengelig kunnskap tyr vi ofte til tommelfingerregler (heuristikker) som fører til raske, men i noen tilfeller til feilaktige beslutninger. Nye eksperimentelle studier kan tyde på at tilbøyeligheten til å bruke tommelfingerregler ikke

korrelerer særlig høyt med IQ (Stanovich & West, 2014). Disse funnene har ført til et forslag om å skille mellom intelligens og rasjonalitet (Stanovich, 2009).

Konklusjon og veien videre

Konklusjonen må bli at økningen, den etterfølgende utflatingen og nedgangen i gjennomsnittlig IQ utelukkende drives av miljøfaktorer. Den genetiske sammensetningen av en populasjon forandrer seg ikke over såpass korte tidsrom, og effekten av færre tilfelle av inngifte har liten eller ingen betydning, i hvert fall i Norge de siste tiårene. Det er vanskelig å finne ut hvilke faktorer som har vært viktigst. Flynn (2016) har lansert en ide om at de virksomme faktorene er en pakke av faktorer som omfatter blant annet forbedring av ernæring, bedre og mer utdanning, bedre helse, endrede tenkemåter (f.eks. større tilbøyelighet til abstrakt tenkning og moderne kategorier), mindre søskenflokker og forandringer i krav og forventninger til de unge. Sannsynligvis forandrer driverne bak Flynn-effekten seg over tid og etter sted. Avtagende stigningshastighet i Norge og enkelte andre land kan skyldes at flere faktorer har sluttet å være drivere for Flynn-effekten siden de første årene etter siste verdenskrig. For Norges del er det sannsynlig at oppbyggingen av velferdsstaten i årene etter krigen har hatt avgjørende betydning. Bortfall av lavskårener fra generasjon til generasjon tyder på at bedre utdanning og helsevesen og andre forhold først og fremst har kommet de mindre ressurssterke til gode. I fødselskullene etter krigen har antakelig forskjeller i sosiale forhold (mellom-familie-faktorer) hatt mindre betydning for Flynn-effekten. Økningene i fødselskullene fra midten av 1950-årene skyldes i hovedsak faktorer som virker innenfor familier, uten at vi kan peke på hvilke faktorer dette kan være.

Selv om datagrunnlaget foreløpig er noe spinkelt, er det rimelig å tro at forskjellene i gjennomsnittlig IQ mellom industrialiserte land og utviklingsland blir stadig mindre og antakelig er i ferd med å forsvinne.

En viktig konsekvens av Flynn-effekten er at psykologer skal være nøye med å ta denne i betraktning når vi er sakkyndige i saker som har potensielt stor betydning for enkeltmennesker, som f.eks. spørsmål om strafferettslig tilregnet og avgjørelser vedrørende omsorgsevne og uføretrygd – foreldede normer kan medvirke til skjebnesvangre avgjørelser. Det beste sikkerhetsnettet er hyppige renormeringer. Som tidligere påpekt er foreldede normer neppe et problem i Norge ettersom det eksisterer normer som er mindre enn 10 år gamle.

Forskningen på årsaker til Flynn-effekten er i full gang. Årsakene som er blitt foreslått og undersøkt, er ofte av størst interesse for sosiologer, demografer, antropologer og historikere. Det er forholdsvis lite forskning på mulige psykologiske og nevropsykologiske årsaker til Flynn-effekten. Riktignok er det stor aktivitet når det gjelder å finne sammenhenger mellom IQ og elementære psykologiske funksjoner, samt underliggende nevropsykologiske prosesser, men det er tynt med undersøkelser av hvordan slike faktorer *forandrer* seg fra generasjon til generasjon. Slik forskning kan være svært krevende både tidsmessig og ressursmessig. Det ideelle designet er at man tester folk som er omtrent like gamle, men tilhører forskjellige fødselskull. Designet kompliseres ytterligere dersom man i tillegg ønsker opplysninger om forsøkspersoner i forskjellig alder, men innenfor samme kohort.

Referanser

Anastasi, A. (1956). Intelligence and family size. *Psychological Bulletin*, 53, 187–209. doi:[10.1037/h0047353](https://doi.org/10.1037/h0047353)

Ang, S., Rodgers, J. & 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. doi:[10.1016/j.intell.2010.05.004](https://doi.org/10.1016/j.intell.2010.05.004)

Bashi, J. (1977). Effects of inbreeding on cognitive performance. *Nature*, 266, 440–442. doi:[10.1038/266440a0](https://doi.org/10.1038/266440a0)

Batterjee, A. A., Khaleefa, O., Ali, K. & Lynn, R. (2013). An increase of intelligence in Saudi Arabia, 1977–2010. *Intelligence*, 41, 91–93. doi:[10.1016/j.intell.2012.10.011](https://doi.org/10.1016/j.intell.2012.10.011)

Bore, R. R. (2007). Norske rekrutter har skutt i været. Rekruttene gjennomsnittshøyde 1878–2006. *Historisk helsestatistikk* (s. 136–152). ISBN 978-82-537-7294-3. Oslo: Statistisk sentralbyrå.

Bratsberg, B. & Rogeberg, O (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the American of Sciences (PNAS)*, 115(26), 6674–6678. doi:[10.1073/pnas.1718793115](https://doi.org/10.1073/pnas.1718793115)

Bronfenbrenner, U. & Ceci, S. J. (1994). Nature–nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, 101, 568–586. doi:[10.1037/0033-295X.101.4.568](https://doi.org/10.1037/0033-295X.101.4.568)

Cattell, R. B. (1971). *Abilities: their structure, growth, and action*. Boston: Houghton-Mifflin.

Colom, R., Juan-Espinosa, M. & Garcia, L. F. (2001). The secular increase in test scores is a “Jensen effect”. *Personality and Individual Differences*, 30, 553–559. doi:[10.1016/S0191-8869\(00\)00054-4](https://doi.org/10.1016/S0191-8869(00)00054-4)

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. doi:[10.1016/j.intell.2004.07.010](https://doi.org/10.1016/j.intell.2004.07.010)

Colom, R., Flores-Mendoza, C. E., Francisco, J. & Abad, F. J. (2007). Generational changes on the draw-a-man test: A comparison of Brazilian urban and rural children tested in 1930, 2002 AND 2004. *Journal of Biosocial Science*, 39, 79–89. doi:[10.1017/S0021932005001173](https://doi.org/10.1017/S0021932005001173)

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. doi:[10.1111/1467-9280.02434](https://doi.org/10.1111/1467-9280.02434)

Dutton, E. & Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009. *Intelligence* 41, 817–820. doi:[10.1016/j.intell.2013.05.008](https://doi.org/10.1016/j.intell.2013.05.008)

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29–51. doi:[10.1037/0033-2909.95.1.29](https://doi.org/10.1037/0033-2909.95.1.29)

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171–191. doi:[10.1037/0033-2909.101.2.171](https://doi.org/10.1037/0033-2909.101.2.171)

Flynn, J. R. (1998). Israeli military IQ tests: Gender differences small; IQ gains large. *Journal of Biosocial Science*, 30, 541–553. doi:[10.1017/S0021932098005410](https://doi.org/10.1017/S0021932098005410)

Flynn, J. R. (1999). Reply to Rushton: A gang of gs overpowers factor analysis. *Personality and Individual Differences*, 26, 391–393.

Flynn, J. R. (2007). *What is intelligence. Beyond the Flynn effect*. Cambridge: Cambridge University Press. doi:[10.1017/CBO9780511605253](https://doi.org/10.1017/CBO9780511605253)

Flynn, J. R. (2012). *Are We Getting Smarter? Rising IQ in the Twenty-First Century*. Cambridge: Cambridge University Press. doi:[10.1017/CBO9781139235679](https://doi.org/10.1017/CBO9781139235679)

- Flynn, J. R. (2016). *Does your family make you smarter? Nature, nurture and human autonomy*. Cambridge: Cambridge University Press.
- Gardner, H. E. (1983). *Frames of mind. The theory of multiple intelligences*. New York: Basic Books.
- Hartmann, E. (1991). Effects of day care and maternal teaching on child educability. *Scandinavian Journal of Psychology*, 32, 325–335. doi:[10.1111/j.1467-9450.1991.tb00883.x](https://doi.org/10.1111/j.1467-9450.1991.tb00883.x)
- Herrnstein, R. J. & Murray, C. (1994). *The Bell Curve: Intelligence and the Class Structure in American Life*. New York: Free Press.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.
- Jung, R. E. & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, 30, 135–187. doi:[10.1017/S0140525X07001185](https://doi.org/10.1017/S0140525X07001185)
- Kahneman, D. (2012). *Tenke fort og langsomt*. Oslo: Pax forlag.
- Kane, H. & Oakland, T. H. (2000). Secular declines in Spearman's g: Some evidence from the United States. *Journal of Genetic Psychology*, 16, 337–345. doi:[10.1080/00221320009596716](https://doi.org/10.1080/00221320009596716)
- Khaleefa, O., Sulman, A. & Lynn, R. (2009). An increase of intelligence in Sudan, 1987–2007. *Journal of biosocial. Science*, 41, 279–283. doi:[10.1017/S0021932008003180](https://doi.org/10.1017/S0021932008003180)
- Lynn, R. & Hampson, S. L. (1986). The rise of national intelligence: Evidence from Britain, Japan and the USA. *Personality and Individual Differences*, 7, 23–32. doi:[10.1016/0191-8869\(86\)90104-2](https://doi.org/10.1016/0191-8869(86)90104-2)
- Lynn, R. & Harvey, J. (2008). The decline of the world's IQ. *Intelligence*, 36, 112–120. doi:[10.1016/j.intell.2007.03.004](https://doi.org/10.1016/j.intell.2007.03.004)
- Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the Development Quotients of infants. *Intelligence*, 37, 16–24. doi:[10.1016/j.intell.2008.07.008](https://doi.org/10.1016/j.intell.2008.07.008)
- Lynn, R. (2013). Who discovered the Flynn effect? A review of early studies of the secular increase of intelligence. *Intelligence*, 41, 765–769. doi:[10.1016/j.intell.2013.03.008](https://doi.org/10.1016/j.intell.2013.03.008)
- Mackintosh, N. J. (2011). *IQ and human intelligence (2. utg.)*. Oxford: Oxford University Press.

- 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. doi:[10.1016/S0160-2896\(03\)00058-8](https://doi.org/10.1016/S0160-2896(03)00058-8)
- Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review*, 114, 806–829. doi:[10.1037/0033-295X.114.3.806](https://doi.org/10.1037/0033-295X.114.3.806)
- Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J. Halpern, D. F. & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American psychologist*, 67, 130–159. doi:[10.1037/a0026699](https://doi.org/10.1037/a0026699)
- Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, 41, 1–48. doi:[10.1006/cogp.1999.0735](https://doi.org/10.1006/cogp.1999.0735)
- Rindermann, H., Becker, D. & Coyle, T.R. (2017). Survey of expert opinion on intelligence: The Flynn effect and the future of intelligence. *Personality and Individual Differences*, 106, 242–247. doi:[10.1016/j.paid.2016.10.061](https://doi.org/10.1016/j.paid.2016.10.061)
- Rushton, J. P. & Jensen, A. R. (2015). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, 11, 235–294. doi:[10.1037/1076-8971.11.2.235](https://doi.org/10.1037/1076-8971.11.2.235)
- Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist*, 49, 304–313. doi:[10.1037/0003-066X.49.4.304](https://doi.org/10.1037/0003-066X.49.4.304)
- Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Stanovich. K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven: Yale University Press.
- Stanovich. K. E. & West, R. F. (2014). The assessment of rational thinking: IQ ≠ RQ. *Teaching of Psychology*, 41, 265–271. doi:[0098628314537988](https://doi.org/0098628314537988)
- Sundet, J. M., Barlaug, D. G. & 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*, 33, 349–362. doi:[10.1016/S0160-2896\(04\)00052-2](https://doi.org/10.1016/S0160-2896(04)00052-2)

- Sundet, J. M., Borren, I. & Tambs K. (2008). The Flynn effect is partly caused by changing fertility patterns. *Intelligence*, 36, 183–191.
doi:[10.1016/j.intell.2007.04.002](https://doi.org/10.1016/j.intell.2007.04.002)
- Sundet, J. M., Eriksen, W. & Tambs, K. (2010). The Flynn effect in sibships: Investigating the role of age differences between siblings. *Intelligence*, 38, 38–44. doi:[10.1016/j.intell.2009.11.005](https://doi.org/10.1016/j.intell.2009.11.005)
- Sundet, J. M. (2014). The Flynn effect in families: Studies of register data on Norwegian conscripts and their families. *Journal of Intelligence*, 2, 106–118. doi:[10.3390/jintelligence2030106](https://doi.org/10.3390/jintelligence2030106)
- Tambs, K, Sundet, J. M. & Magnus, P. (1988) Genetic and environmental effects on the covariance structure of the Norwegian Army Ability Tests: A study of twins. *Personality and Individual Differences*, 9, 791–799. doi:[10.1016/0191-8869\(88\)90069-4](https://doi.org/10.1016/0191-8869(88)90069-4)
- Teasdale, T. W. & Owen, D. H. (1989). Continuing secular trends in intelligence and a stable prevalence of high intelligence levels. *Intelligence*, 13, 255–262. doi:[10.1016/0160-2896\(89\)90021-4](https://doi.org/10.1016/0160-2896(89)90021-4)
- 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, 837–843.
doi:[10.1016/j.paid.2005.01.029](https://doi.org/10.1016/j.paid.2005.01.029)
- Teasdale, T. W. & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence*, 36, 121–126.
doi:[10.1016/j.intell.2007.01.007](https://doi.org/10.1016/j.intell.2007.01.007)
- te Nijenhuis, J., Cho, S. H., Murphy, R. & Lee, K. H. (2012) The Flynn effect in Korea: Large gains. *Personality and Individual Differences*, 53, 147–151. doi:[10.1016/j.paid.2011.03.022](https://doi.org/10.1016/j.paid.2011.03.022)
- te Nijenhuis, J. & van der Flier, H. (2013). Is the Flynn effect on g? A meta-analysis. *Intelligence*, 41, 801–807.
doi:[10.1016/j.intell.2013.03.001](https://doi.org/10.1016/j.intell.2013.03.001)
- te Nijenhuis, J. T., Murphy, R. & van Eeden, R. (2011). The Flynn effect in South Africa. *Intelligence*, 39, 456–467.
doi:[10.1016/j.intell.2011.08.003](https://doi.org/10.1016/j.intell.2011.08.003)
- Trahan, L. H., Stuebing, K. K., Fletcher, J. M. & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, 140, 1332–1360. doi:[10.1037/a0037173](https://doi.org/10.1037/a0037173)
- Tuddenham, R. D. (1948). Soldier intelligence in World Wars I and II. *American Psychologist*, 3, 54–56. doi:[10.1037/h0054962](https://doi.org/10.1037/h0054962)

Wicherts, J. M., Dolan, C. V., Hessen, D.J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I. & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, 32, 509–537. doi:[10.1016/j.intell.2004.07.002](https://doi.org/10.1016/j.intell.2004.07.002)

Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence*, 41, 753–764. doi:[10.1016/j.intell.2013.04.010](https://doi.org/10.1016/j.intell.2013.04.010)

Woodley, M. A. (2012). A life history of the Flynn effect. *Personality and individual differences*, 53, 152–156. doi:[10.1016/j.paid.2011.03.028](https://doi.org/10.1016/j.paid.2011.03.028)

Woodley, M. A., te Nijenhuis, J. & Murphy, R. (2013). Were the Victorians cleverer than us? The decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence*, 41, 843–850. doi:[10.1016/j.paid.2011.03.028](https://doi.org/10.1016/j.paid.2011.03.028)

Citation

Storsve, O., Sundet, J. M., Torjussen, T. M., & Lang-Ree, O. C. (2018). Flynn-effekten i Norge og andre land: Praktiske implikasjoner og teoretiske spørsmål [The Flynn effect in Norway and other countries: Practical implications and theoretical issues]. *Scandinavian Psychologist*, 5, e6. <https://doi.org/10.15714/scandpsychol.5.e6>

Abstract

The Flynn effect in Norway and other countries: Practical implications and theoretical issues

A number of studies have documented significant gains in mean IQ scores across generations (the Flynn effect). Flynn effects have been found in more than 30 industrialized nations, including Norway, but also in developing nations in Latin America and Africa. In the current paper, we present international and Norwegian results attained from young male military conscripts, encompassing new and unpublished data. In accordance with international data, the mean IQ scores in Norway increased from test years in the late 1950s (birth cohorts \approx 1938–1940), but with decreasing gain rates, to the mid-1990s. After the mid-1990s, increases in the means ceased and exhibited decreasing tendencies. Our new data set shows that these decreases have continued into the new millennium. Practical and methodological consequences of the secular trends in IQ scores are often overlooked. We address issues related to the use of intelligence tests in situations where decisions with potentially serious consequences for participants are involved. In particular, we underscore the need for test norms that take the Flynn effect into consideration. In cross-sectional studies, age and birth cohort are

completely confounded. Effects of birth cohort on intelligence (Flynn effect) have often been wrongly interpreted as age effects. To study age effects, longitudinal studies are needed. There are several theories addressing the possible causes of the Flynn effect. In the present paper we discuss how well existing theories explain known facts, with special emphasis placed on the Norwegian data.

Keywords: family size, nutrition and education, secular trends in IQ, scientific thinking.

Author affiliations: Olav Storsve – Retired psychologist, Vestby, Norway; Jon Martin Sundet – Department of Psychology, University of Oslo, Oslo, Norway; Tore M. Torjussen – Retired psychologist, Oslo, Norway; Ole Christian Lang-Ree – Department of Military Psychology and Leadership Development, The Norwegian Defence University College, Oslo, Norway.

Contact information: Olav Storsve, Vestby, Norway.
Email: olav@storsve.no.

Received: April 26, 2018. **Accepted:** July 23, 2018. **Published:** August 17, 2018.

Language: Norwegian.

Competing interests: The authors report no conflict of interest. The authors alone are responsible for the contents and writing of this paper.

Acknowledgement: The authors would like to thank three anonymous peer reviewers for pointing out mistakes and deficiencies as well as for suggesting changes.

This is a peer-reviewed paper.