



IQ – Skrevet i stein eller øvd inn?

Forfatter: Casper Faarlund, Ullern vgs, Oslo

SAMMENDRAG

I dette forsøket har jeg undersøkt om det er mulig å øve til IQ-tester, da det er allmenn oppfatning at slike tester gir en score som ikke endres over tid, og brukes aktivt i samfunnet. En gruppe på 66 elever tok del i forsøket, og de som har tatt IQ-tester før, fikk et bedre gjennomsnitt (119,74 mot 113,52) enn de som ikke hadde tatt slike tester før, med en P-verdi på 0,0028.

INTRODUKSJON

IQ benyttes i dag som et mål (IQ(norsk), u.d.) på menneskets intelligens. Tester som inneholder geometriske mønstre, er tester som i prinsippet ikke skal påvirkes av språk og bakgrunn. Gjennom tidene har IQ-tester hatt viktige roller. Allerede under første verdenskrig ble det brukt intelligenstester for å se hvilke soldater som kunne være aktuelle for offiserstillinger i USA (IQ(norsk), u.d.). I dag brukes IQ-testing ofte i jobbintervjuer (IQ(norsk), u.d.). Dette forekommer ofte i store internasjonale selskaper (DAHLE, u.d.). Derfor vil resultatet av en IQ-test kunne påvirke en persons fremtid.

En større IQ-test kan ha oppgaver som å definere ord, mens de fleste mindre tester som de vi finner på nettet, kun består av matriser. Dette vil si at den som tar testen, skal finne ut hvilken geometrisk figur som kommer neste i rekken. Testen som brukes i dette forsøket består av 35 slike oppgaver som skal løses innen 25 minutter. Alle oppgavene er matriser der man skal svare hva den siste figuren blir.

Siden IQ-tester brukes til å fastslå hvor egnet et menneske er til noe, er det interessant å vite hvor pålitelig en IQ-test faktisk er, og om trening og/eller tidligere erfaring, bidrar til at noen stiller sterkere enn andre. Dersom man faktisk kan øve til en IQ-test, vil dette gjøre resultatene mindre pålitelige, og man burde få vite om de eventuelle effektene av øving, slik at alle stiller likt. Dette kan ha påvirkning på ulike situasjoner, som jobbintervjuer og andre kvalitetstester i samfunnet, og er derfor verdt å undersøke nærmere.

Hvis vi vet at IQ varierer over tid, viser det at den blir påvirket av eksterne faktorer. Et kjent fenomen vedrørende IQ-testers pålitelighet, er Flynn-effekten. Dette fenomenet går ut på at resultatene på IQ-tester øker over tid (Flyneffekten, u.d.). Fenomenet ble oppkalt etter James R. Flynn, en statsviter fra New Zealand. Den viser at IQ-en i befolkningen øker med gjennomsnittlig 3 poeng per tiår, dersom testene er av lik vanskelighetsgrad. Fordi IQ-tester er designet slik at gjennomsnittet alltid skal bli 100 (IQ(norsk), u.d.), betyr dette at vanskelighetsgraden har økt gjennom tiårene. Denne effekten har vært lineær helt frem til 1990 tallet i utviklede land, men da begynte den å avta, og snu i enkelte land. Dette har for eksempel blitt

vist i en undersøkelse med over 500 000 unge menn fra Danmark. Utviklingen viser en klar nedgang fra slutten av 1990 tallet (W.Teasdale, 2005). Dette styrker hypotesen om at IQ kan trenes opp.

I følge Mensa (FAQ om IQ og intelligens, u.d.), kan man lese seg opp til IQ-tester, men gevinsten er minimal, og derfor mener de at de færreste vil gjøre det. Mensa mener denne effekten er på rundt 4-5 poeng. Det er denne gevinsten jeg skal undersøke i dette forsøket.

For å undersøke hypotesen min, utførte jeg en T-test der nullhypotesen H_0 var

- «Å øve til en IQ-test vil gi samme resultat»

Alternativ hypotese, H_A :

- Å øve til en IQ-test vil gi høyere score

METODE

Hensikten med forsøket er å undersøke hvorvidt man vil oppnå høyere score av å øve til IQ-tester. Til dette trengte jeg en forsøksgruppe, i mitt tilfelle bestående av 66 elever på VG3 på Ullern VGS. Videre trengte jeg også en IQ-test. Jeg benyttet Mensa Norges Hjemmetest (Online testen til mensa norge, u.d.).

Testen ble gitt samtlige elever samtidig. De leverte scoren i et Google Forms dokument, der de til slutt også fikk svare på «Har du tatt IQ-test før». Her var det flervalg, der man kunne velge mellom «aldri», «gjort én test» og «gjort flere tester». Elevene tok testene i klasserom, i de klassene de vanligvis hadde time i på dette tidspunktet.

Jeg slo sammen de som hadde tatt «en test én gang før», og de som har tatt «flere ganger». Deretter satt jeg opp to lister i GeoGebra, en til de som hadde tatt tester før, og en til de som aldri hadde tatt tester før. Jeg fant først begge gjennomsnittene. Deretter fant jeg standardavvik og utførte en T-Test, differanse mellom to gjennomsnitt i GeoGebra. Jeg fikk da en P-verdi.

RESULTAT

IQ score for de som har øvd 100, 101, 131, 117, 133, 113, 128, 119, 123, 115, 121, 133, 131, 121, 130, 119, 106, 121, 123, 110, 113, 128, 133, 119, 123, 125, 110, 125, 124, 110, 123, 121, 117, 108, 117

Scoren for de som aldri hadde tatt en IQ-test før, var som følger: 121, 115, 113, 117, 110, 119, 121, 108, 117, 108, 115, 131, 113, 126, 95, 117, 110, 115, 113, 108, 115, 100, 121, 101, 121, 108, 103, 113, 110, 114, 121

Respondentene som hadde tatt IQ-test en eller flere ganger før, fikk et gjennomsnitt på 119,74, og et standardavvik på 8,67. De som aldri hadde gjennomført en test før, hadde derimot et gjennomsnitt på 113,52, med et standardavvik på 7,56. En T-Test mellom de to gjennomsnittene, løst på Gjort det før ≠ Aldri prøvd, ga en P-verdi på 0,0028.

Tabell 1: Gjennomsnittlige IQ-verdier for gruppene som har gjennomført IQ-tester tidligere, og som aldri har gjort det. P-verdi fra T-Test mellom de to gjennomsnittene

	Gjort det før	Aldri prøvd	Total
Gjennomsnitt	119,74	113,52	116,82
Standardavvik	8,67	7,56	8,12
Antall	35	31	66
P-verdi			0,0028

DISKUSJON

Jeg fant at min alternative hypotese ble styrket, altså at de som har gjennomført IQ-tester før, som fungerer som «øving», fikk bedre score enn de som aldri hadde gjort det. Vi ser at gjennomsnittet er 6,22 poeng høyere for de som har tatt IQ-tester før. Etter å ha kjørt det gjennom en T-test som undersøker sannsynligheten for at de to listene faktisk er ulike, og ikke bare resultat av tilfeldigheter, finner vi at sannsynligheten for at det var tilfeldig, er 0,0014. Det er godt under 5 % kravet, og vi kan dermed konkludere med at å ha gjort IQ-tester før, ga et bedre resultat i testen.

Problemstillingen i forsøket var om det var mulig å øve til IQ-tester. Det forsøket faktisk testet, var om de som hadde tatt IQ-tester tidligere, fikk bedre score. Dette er knyttet sammen, da å ha tatt tester tidligere fungerer som øving – men det måler for eksempel ikke hvorvidt det er forskjell på å øve mye eller lite. Det er viktig å ha i tankene når man vurderer hvorvidt resultatene er troverdige.

Etter min mening er resultatene troverdige, da resultatet stemmer overens med tidligere forskning utført av Mensa (FAQ om IQ og intelligens, u.d.), og vi har en P verdi på 0,14 %. Til tross for at jeg mener denne forskjellen i resultat stemmer, er det også usikkerhet og feilkilder i forsøket, og at det kunne vært utført bedre for å bedre svare på problemstillingen.

For det første er oppgavesettet man finner på nettet annerledes enn det man finner på profesjonelle tester. Under utvikling av profesjonelle tester, fjernes de oppgavene som har lav diskriminerende evne, med andre ord oppgaver som har dårlig evne til å skille personer fra hverandre (Dørum). Dette medfører to ting. Den ene; at treningseffekten vil være annerledes på hjemmetester enn disse profesjonelle testene. Det kan være en mulighet at det kun er disse oppgavene som en ikke finner i de profesjonelle settene, som faktisk gir forbedret resultat. I så fall, vil øving til hjemmetester ikke påvirke resultatet på de profesjonelle testene. Den andre konsekvensen av dette, er at hjemmetestene er mindre pålitelige. De mindre diskriminerende oppgavene er det litt tilfeldig hvem som kan løse, og dette fører til mer variasjon i resultatene. Dette øker sjansen for at forskjellen i snitt var tilfeldig, selv om P-verdien var såpass lav.

En annen feilkilde, er at Mensa Norges Hjemmetest er lik hver gang man tar den. Dette, kombinert med at den er kjent og kommer høyt opp i Google søk, åpner muligheten for at enkelte av individene i forsøket har tatt den nøyaktig samme testen før, eller til og med sett på fasiten før. Dette vil gi dem en fordel de ikke ville hatt i en faktisk IQ-test.

Det er flere faktorer som kan ha ført til at ikke alle forsøkskandidatene presterte sitt beste. For det første, så ble testene foretatt i klasserom. Det ble snakket sammen innimellom underveis, til tross for at jeg forsøkte å dempe dette. Elevene kan også ha sett på sidemannen om de slet selv. Det er videre vanskelig å vite om elevene faktisk var interessert i å ta testen, og om de var motiverte til å gjøre sitt beste. Ellers, så er det også ubevisste faktorer som når på dagen testen ble tatt, søvnen natten før og blodsukkeret (glukosenivået) til kandidatene, som også muligens kan ha hatt påvirkning på konsentrasjonen.

Det er viktig å tenke på at ved å gruppere etter de som har gjennomført IQ-tester før og ikke, kan påvirke utgangspunktet. Det er en sannsynlighet for at mennesker med høy IQ oppsøker slike tester hyppigere på fritiden. Det at de har brukt egen fritid på IQ-tester kan videre bety at de har større interesse for IQ-tester enn de som ikke har oppsøkt slike før, og derfor er mer motiverte i mens de tar testen, og presterer bedre. Dette kan påvirke resultatene og er verdt å ha i tankene.

En mer fordelaktig måte å undersøke problemstillingen på kunne vært å finne to oppgavesett av samme vanskelighetsgrad, som deretter ble gitt i to klasser. Så kunne man kommet tilbake ved en senere anledning, og gått igjennom et par oppgaver og forklare hvordan de løses (eller vise en YouTube video med samme formål) for kun den ene klassen. Deretter kunne de begge løst et nytt sett. Da ville man kunnet sammenligne forskjellen i de to snittene. Dersom hypotesen om at øving hjelper, ville for eksempel den klassen som ikke fikk opplæring få en forskjell på rundt 0, mens den andre ville gått opp et par poeng.

Hypotesen var at øvelse fører til bedre resultater på IQ-tester. Ut i fra resultatene ser vi at de som hadde tatt IQ-tester før, fikk et bedre snitt enn de som ikke hadde det. Ettersom det å ta tester fungerer som øving, kan det derfor se ut til at hypotesen ble styrket. Forsøket kunne derimot vært gjennomført bedre, og har flere feilkilder. Det kan derfor virke som at ytterligere forsøk som gjøres annerledes og på et større utvalg, må til for å konkludere noe sikkert.

KILDER

- Dahle, D. Y. (u.d.). *Slik blir du testet før du får jobb i Norges største selskaper*. Hentet fra TU: <https://www.tu.no/artikler/slik-blir-du-testet-for-du-far-jobb-i-norges-storste-selskaper/275649>
- Dørum, O. H. (u.d.).
FAQ om IQ og intelligens. (u.d.). Hentet fra Mensa Norge: <https://www.mensa.no/iq/faq-om-iq-og-intelligens/>
- Flynneffekten*. (u.d.). Hentet fra Wikipedia: <https://no.wikipedia.org/wiki/Flynneffekten>
- Hjemmetest Svensk*. (u.d.). Hentet fra Mensa Sverige: <https://www.mensa.se/provtestet/test/>
- IQ*. (u.d.). Hentet fra Wikipedia: https://en.wikipedia.org/wiki/Intelligence_quotient
- IQnorsk*. (u.d.). Hentet fra Wikipedia: <https://no.wikipedia.org/wiki/IQ>
- IQ og intelligens*. (u.d.). Hentet fra Mensa Norge: <https://www.mensa.no/iq/>
- Online testen til mensa norge*. (u.d.). Hentet fra Mensa Norge: <http://test.mensa.no>
- W.Teasdale, T. (2005, September). *A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse*. Hentet fra Science Direct: <https://www.sciencedirect.com/science/article/pii/S0191886905001145?via%3Dihub>