



Testing av responsinhibitorisk evne og post-error slowing hos elever ved Elvebakken vgs

Forfatter: Sakarias Walhovd, Elvebakken videregående skole

Flankertesten tester responsinhibisjon, der deltageren skal respondere på retningen til en fokuspil omgitt av andre piler. Forskningsspørsmålet er om responstiden til deltagerne blir påvirket av hvorvidt alle pilene peker i samme retning. Forsøket ser også på post-error slowing (PES). Etter at en deltager responderer feil på en oppgave, er det forventet at neste oppgaves responstid (RT) vil være høyere. 27 deltagere gjennomførte 40 Flankeroppgaver. Dette viste en signifikant forskjell i responstiden ut ifra hvorvidt pilens retning samsvarte, og 68 ms tregere respons etter at en feil ble gjort. En t-test gjennomført i Python på de 1075 totale resultatene ga en t-verdi på 5,74 og en p-verdi på $4,77 \cdot 10^{-6}$ for RT. Analysen av PES ga en t-verdi på 3,40 og en p-verdi på $8,81 \cdot 10^{-4}$. Resultatene viser at å undertrykke irrelevant stimuli er viktig for å yte gode resultater. Tydelig PES viser at deltagerne evner å justere responsen sin etter feil.

Innledning

Inhibitorisk evne er en sentral ferdighet å mestre. Flere og flere er bekymret for unges konsentrasjons-evne og mulighet til å følge med (Løkken, 2019). Et sentralt aspekt ved konsentrasjon er evnen til å kunne undertrykke forstyrrende stimuli. Flankertesten er en responsinhibitorisktest (Eriksen & Eriksen, 1974). Den måler responstiden i korrelasjon med irrelevant stimuli. Forskningsspørsmålet er om responstiden til deltagerne blir påvirket av hvorvidt alle pilene peker i samme retning. Forsøket ser også på post-error slowing (PES). Etter at en deltager responderer feil på en oppgave, er det forventet at neste oppgaves responstid (RT) vil være høyere.

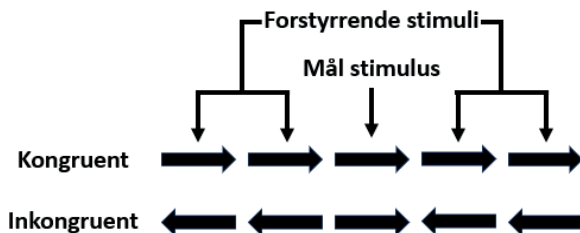
Målet med testen er å måle deltagerens responstid i relasjon til forstyrrende stimuli som blir presentert tett til fokuspielen. Fokuspielen er plassert i midten og er flankert av irrelevant stimuli som enten samsvarer med retningen til fokuset (Kongruentstimuli), den motsatte retningen (inkongruentstimuli), eller ingen av delene (nøytralstimuli) slik som vist i figur 1. Testen har fått navnet sitt etter de amerikanske forskerne Barbara. A. Eriksen & Charles W. Eriksen som var de første til å publisere testen i 1974, og de flankerende stimuliene. Det vanligste er å plassere de flankerende stimuliene til høyre og venstre for fokuset, men det er også mulig å plassere dem i en annen formasjon f.eks. over og under. Flankertesten undersøker deltagerens evne til å undertrykke forstyrrende stimuli. Testen er mye brukt innen psykologi og nevrovitenskap. Flankertesten går ut på at deltageren blir presentert med en serie med stimuli. I den opprinnelige testen som ble publisert i 1974, ble det brukt 7 bokstaver som stimuli (Wikipedia, 2020).

Det er vanlig at det forekommer en rekke feil under gjennomføringen av en Flankertest. Etter en feil forekommer det som regel post-error slowing (PES). Det går ut på at man sakker ned for å sørge for å gjøre det riktig. PES er et viktig fenomen hos mennesker. Hadde vi ikke hatt denne responsen, hadde vi fortsatt på samme måte og utført mange feil. Derfor er lang PES en positiv ting. (Fjell et al., 2012)

Før mennesker gir en respons på stimuli skal en rekke hjerneområder kommunisere på svært kort tid. Visuelle områder i occipitallappen, bakerst i hjernen skal oppfatte stimuluset. Deretter må det tas en beslutning i prefrontal cortex helt forrest i hjernen angående hvilken respons som skal gis. Så må en motorisk respons iverksettes, dette styres i motorisk cortex som ligger mot fremre del, foran midten av hjernen. Flere øvrige hjerneområder er også involvert i prosessen. Blant annet for å styrke oppmerksomheten til oppgaven. Etter dette må det sendes nervesignaler til musklene som skal brukes. Det betyr at beslutningen om å respondere tas i løpet av 200-300 ms. I løpet av den tiden har flere hjerneområder som ligger rundt 15 cm fra hverandre sendt signaler og kommunisert frem og tilbake flere ganger. Derfor er det naturlig å tenke at kvaliteten på nervefibrene er veldig viktig. Det er disse som utgjør hjernens hvite substans (white matter: WM). Vi finner at de som har lang PES har større volum av WM. Å ha større WM-volum er positivt og tyder på en frisk hjerne. Samt større WM-volum har de med lang PES mindre mean diffusion (MD), dette er også positivt. Enkelt sagt tyder mindre MD på høy kvalitet på nervefibrene (Fjell et al., 2012).

Metode

Eriksen Flankertesten ble brukt for å måle responsinhibitorisk evne, og for å danne et grunnlag for å se på PES. I den første versjonen av forsøket ble det benyttet 7 stimuli, men i dette forsøket ble det brukt 5 piler. I dette forsøket benyttes det kun kongruent- og inkongruentstimuli, og ikke nøytral stimuli. Oppsettet var likt som i figur 1 bortsett fra pilenes fargesong. I dette forsøket er de plassert til høyre og venstre for fokuset. Forsøket ble gjennomført på en bærbar PC ved hjelp av en Python-kode skrevet av Erik Marsja (Erik Marsja, u.å.). Koden benyttet seg av biblioteket Expyriment. Koden etablerer først de fire forskjellige stimuliene, hvorav to er kongruente og to er inkongruente. Koden loggførte og plottet deretter deltager-ID, blokken, riktige respons, reelle respons, hvilket stimulus som ble presentert, responstid (RT), hvorvidt testen ble besvart riktig, samt hvorvidt testen var kongruent eller inkongruent.



Figur 1. Figuren viser et enkelt diagram som demonstrerer forskjellen på kongruent og inkongruent stimuli.

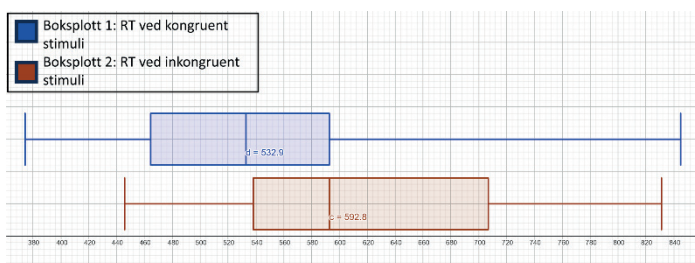
Samtlige deltagere har enten faget Teknologi og forskningslære 1 eller Teknologi og forskningslære 2. Hver deltager fullførte en test bestående av 10 blokker, hver blokk inneholdt 4 tilfeldige kongruente eller inkongruente tester. Etter hver blokk var det en kort pause på 3000 ms før deltageren ble presentert for neste blokk. Totalt hadde forsøket 27 deltagere. Med 40 tester per deltager tilsvarer det 1080 gjennomførte reaksjonstester. Det er flere varianter av stimuli brukt for Flankertesten. I dette forsøket så de fire forskjellige stimuliene slik ut: Stimulus - 0 <<<<<, Stimulus - 1 >>>>>, Stimulus - 2 <<><<, Stimulus - 3 >><>>. Dersom Mål-stimulus peker til venstre «<» skal deltageren trykke på «X» på tastaturet, peker stimulus til høyre «>» skal deltageren trykke på «M». Dette resulterer i at deltageren benytter venstre hånd for å trykke på knappen til venstre når pilen peker til venstre, og at deltageren trykker med høyre hånd på

knappen til høyre når pilen peker til høyre. Deltageren skal kun fokusere på fokuspilen, og skal prøve å ignorere de 4 forstyrrende pilene.

Det gjennomføres en t-test for å analysere dataene. Vi regner resultatene fra studien som signifikant dersom p-verdien er mindre enn 0,05. Det er verdt å merke at dette er en helt vilkårlig grense. Avfeier man alle resultater som kommer over denne grensen kan man miste verdifull forskning som like vel kan ha signifikans (Dahiru, 2008). Nullhypotesen H_0 er at det ikke er noen forskjell i reaksjonstid mellom kongruente og inkongruente stimuli. Hypotesen H_A er at reaksjonstiden er lengere ved inkongruent stimuli. I denne studien ble PES analysert ved hjelp av kode skrevet i Python ved hjelp av bibliotekene; Pandas, Numpy og Scipy. Koden finner alle gangene det ble begått en feil og trekker deretter fra RT-en til den etterfølgende testen fra RT-en der feilen ble begått. Deretter ble gjennomsnittet av forskjellen i RT etter en feil beregnet. Det ble også gjennomført en t-test for å finne p-verdien.

Resultater

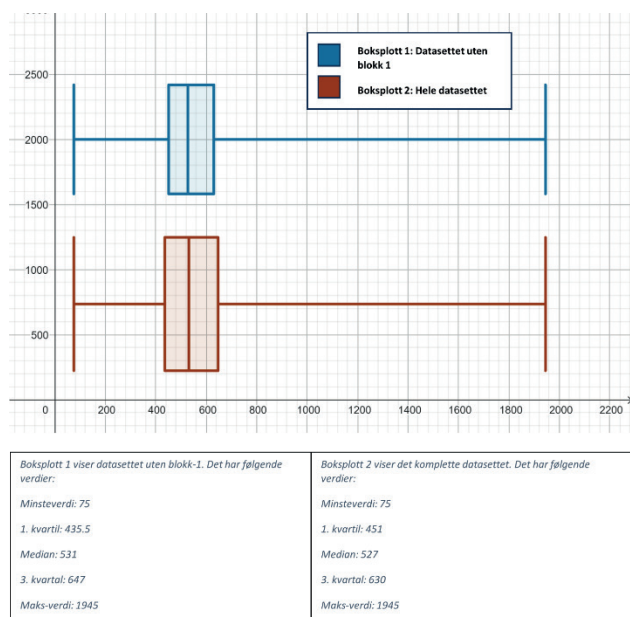
En t-test gjennomført i Python på de 1075 totale resultatene, ga en t-verdi på 5,74 og en p-verdi på $4,77 \cdot 10^{-6}$. Denne verdien er så liten at t-testen opprinnelig ga en p-verdi på 0,00 frem til den ble justert. Ser vi bort ifra første blokk av forsøket hos samtlige deltagere, får vi en t-verdi på 4,45 og en p-verdi på $1,42 \cdot 10^{-4}$. Slik som vist i figur 2 har de kongruente testene en gjennomsnittlig RT på 593ms (1. kvartil: 538 ms, 3. kvartil: 707 ms). De inkongruente testene hadde en gjennomsnittlig RT på 533 ms (1.kvartil: 465 ms, 3.kvartil: 593 ms). Den gjennomsnittlige PES-en er på 68 ms. Analysen av PES ga en t-verdi på 3,40 og en p-verdi på $8,81 \cdot 10^{-4}$. Ser vi bort ifra alle PES som har en absoluttverdi høyere enn snittverdien til alle testene totalt, får vi også en gjennomsnittlig PES på 68 ms, men t-verdien blir 5,68 og p-verdien blir $6,78 \cdot 10^{-8}$.



Figur 2. Figuren viser boksplokk av RT ved kongruent stimuli og boksplokk av RT ved inkongruent stimuli.

Tabell 1. Tabell som samler relevant data på et sted.

	Flankertest	Flankertest uten blokk 1	Post-error slowing	Post-error slowing uten verdier som er større enn absoluttverdien til 527
T-verdi	5,74	4,45	3,40	5,68
P-verdi	$4,77 \cdot 10^{-6}$	$1,42 \cdot 10^{-4}$	$8,81 \cdot 10^{-4}$	$6,78 \cdot 10^{-8}$



Figur 3. Boksplott som viser Flankertesten, sammen med boksplott som viser Flankertesten uten dataene fra blokk1.

Diskusjon

Datasettet skal i utgangspunktet, gitt deltagerantallet og antall tester per deltager, bestå av 1080 tester. I datasettet finner vi bare 1075 resultater. De fem manglende testene er vanskelige å spore opp, men de vil med all sannsynlighet ikke ha noen stor innvirkning på resultatene.

Det ble observert at flere deltagere slet med gjennomføringen av Flankertesten. Det var flere som tilsynelatende ikke skjønnte hva de skulle gjøre. Et fenomen som ble observert flere ganger, selv etter å ha blitt instruert til å bare se på målstimulus, var trykking på x og m fem ganger for å respondere på hvert enkelt stimuli. Dette ble imidlertid rettet på hos samtlige innen første blokk var over. Vi kan derfor si at særlig den første blokken har uforutsigbare data med mulige feilkilder. Likevel vil dette mest sannsynlig ikke ha en for stor påvirkning på resultatene da rekkefølgen på de fire forskjellige testene var tilfeldig. Vi har et stort nok datasett til at dette støyet ikke vil gi noen betydelig endring av resultatene. Med en mindre modifisering av Python-programmet som analyserte dataene fikk vi et nytt resultat. Når vi ser bort ifra blokk 1 hos samtlige deltagere ble P-verdien $1,42 \cdot 10^{-4}$ (tabell 1). Dette gir en noe mindre drastisk konklusjon, dog dataene fortsatt er svært sikre. Det er viktig å nevne at P-verdien naturlig vil øke når datasettet blir mindre. Hadde vi halvert datasettet kunne vi ha regnet med en dobling av P-verdien. Dette avfeier likevel ikke effekten av å se bort ifra blokk-1, da det ga oss omtrentlig 30 ganger høyere P-verdi. Hadde vi kun sett den naturlige effekten og ikke noen reell endring hadde P-verdien kun økt med rundt 10 % da det er omtrentlig det blokk-1 utgjør av det totale datasettet. Dette tyder på at det kan være fordelaktig å ha en prøvetest først hvor deltagerne får erfaring med formatet. Det er fortsatt verdt å merke at å la deltagerne bli kjent med formatet på forhånd kan være vanskelig. Flere av deltagerne syntes de 3 minuttene testen tok var for mye. Det gjør det urealistisk å ha en test av noe særlig større lengde. Det samme gjelder med tanke på å holde deltagerens konsentrasjon ved like. Utover det er det også mulig å på forhånd fastsette at man ikke skal regne med den eller de første blokkene.

Analysen ble gjennomført ved at det ble regnet ut en snittverdi for hver deltager sin RT for de kongruente og de inkongruente oppgavene. Hadde vi analysert dataene med hvert datapunkt som en individuell RT hadde P-verdien vært betydelig mindre meningsfull og slagkraftig da det rene antallet datapunkt da i seg

selv vil redusere p-verdien betydelig. Dermed gir denne metoden resultatene mer kredibilitet fordi man får færre, men sikrere datapunkt og luker ut uteliggere blant RT-ene. Den største RT-en vi fant var på 1945, siden vi regner ut snittverdien vil ikke denne alene kunne forskyve resultatene og gi et feilaktig bilde. I stedet for har den blitt en del av RT-en til den gitte deltageren. Det er også naturlig å tenke at resultatene er svært sikre da p-verdien er så lav og t-verdien er såpass høy som den er. For det betyr at selv hvis dataene skulle ha vært usikre og kunne ha pekt på en større effekt enn det som er reelt, vil de fortsatt være sikre. Det er relativt vanlig å regne en p-verdi på 0,05 som en grense mot nullhypotesen for å si at noe er sikkert. Selv om dette er en helt vilkårlig grense som vist av Tukut Daihru (Dahiru, 2008), kan den ha noe verdi. Benytter vi denne grensen er vi godt innenfor. P-verdien vår er på $4,77 \cdot 10^{-6}$ og viser at det er en forskjell på kongruent- og inkongruentstimuli (tabell 1). Disse resultatene gjør at vi kan konkludere med at det er en signifikant forskjell. Det betyr at forstyrrende irrelevante stimuli påvirker RT ved å gjøre deltagerne tregere. Dette viser og at evnen til å undertrykke irrelevante stimuli er viktig for å yte gode resultater.

Testing av post-error slowing (PES) ga gode og tydelige resultater. P-verdien på $6,78 \cdot 10^{-8}$ (tabell 1) sammen med t-verdien på 5,68 (tabell 1) viser en tydelig trend. Med en gjennomsnittlig responstid på 68 ms mer etter deltagere hadde respondert feil på oppgaven før, er det tydelige tendenser i dataene. Dette kan også bety at deltagerne generelt har god kvalitet på hjernefibrene og en god hjernehelsetilstand (Fjell et al., 2012).

Deltagerne er alle sammen elever ved Elvebakken videregående skole. I tillegg går alle deltagerne studierespesialiserende retning. Dette er interessant da studiespesialiserende ved skolen er en av de linjene med høyest inntakskaraktersnitt i landet (Oslo kommune, u.å.). Dette kan tyde på at deltagerne har god inhibitorisk evne da det er sentralt for å få til skolearbeid. Det er også vist at de som scorer bra på Flankertesten, gjør det bedre i utdanning. Da er det også naturlig å tenke at kvaliteten på hjernefibrene og WM-volumet er stort (Fjell et al., 2012). Derfor kan man tenke at deltagerne ut ifra forholdene kanskje har scoret svært godt.

Videre forskning kan utforske metoder for å forbedre den inhibitoriske evnen hos deltagerne. Samt å indentifisere faktorer som påvirker denne evnen og eventuelle resultater av Flankertesten. Det kan for fremtidige forsøk være interessant å plote resultater oppimot kjønn. Har man et datasett av tilsvarende eller større omfang, kan dette gi betydelige resultater. Viktigheten av å ha en prøvetest først er også signifikant.

Referanseliste

- Dahiru, T. (2008). P - value, a true test of statistical significance? A cautionary note. *Annals of Ibadan postgraduate medicine*, [Internett] 6(1), pp.21–6. Tilgjengelig fra: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4111019/>.
- Erik Marsja (u.å.). Python Programming in Psychology – From Data Collection to Analysis. [Internett] Github Gist. Tilgjengelig fra: <https://gist.github.com/marsja/458274f273bfe999fc49ad9d902a6874> [Hentet 6 Feb. 2024].
- Fjell, A.M., Westlye, L.T., Amlien, I.K. and Walhovd, K.B. (2012). A multi-modal investigation of behavioral adjustment: Post-error slowing is associated with white matter characteristics. *NeuroImage*, 61(1), pp.195–205. <https://doi.org/10.1016/j.neuroimage.2012.03.007>.
- Løkken, N.A. (2019). Professor: Skjermlesing gjør oss utålmodige. [Internett] NRK. Tilgjengelig fra: https://www.nrk.no/innlandet/forskning_-_skjermlesing-gjor-oss-utalmodige-1.14629872 [Hentet 16 Feb. 2024].
- Oslo kommune (u.å.). Nedre poenggrense for inntak til Vg1 (videregående trinn 1) for ungdom med opplæringsrett, skoleåret 2023/2024, 1. inntaksrunde. [Internett] Tilgjengelig fra: <https://www.oslo.kommune.no/getfile.php/13483362-1688541069/Tjenester%20og%20tilbud/Skole%20og%20utdanning/Videreg%C3%A5ende%20skole/S%C3%B8ke%20videreg%C3%A5ende%20skole/Tildeling%20av%20plass%20videreg%C3%A5ende%20skole/2023%20Nedre%20poenggrense%201.%20>

inntak%20Vg1.pdf [Hentet 16 Feb. 2024].

Wikipedia. (2020). Eriksen flanker task. [Internett] Tilgjengelig fra: https://en.wikipedia.org/wiki/Eriksen_flanker_task [Hentet 6 Feb. 2024].

