

GÖTEBORGS UNIVERSITET
PSYKOLOGISKA INSTITUTIONEN

**Vidareutveckling av reliabla förändringsindex
för upprepade mätningar med
Rey Auditory Verbal Learning Test**

Hilda Andrén och Rebecka Fernqvist

Examensarbete 30 hp
Psykologprogrammet
PM2519
Vårterminen 2019

Handledare: Per Hellström
Isabelle Rydén

Vidareutveckling av reliabla förändringsindex för upprepade mätningar med Rey Auditory Verbal Learning Test

Hilda Andrén och Rebecka Fernqvist

Sammanfattning. Syftet med föreliggande arbete var att utöka redan befintliga normvärden för förändring på basen av friska kontrollpersoners prestationer på testet Rey Auditory Verbal Learning Test (RAVLT) och att därmed göra de beräkningar som utförs av en så kallad Förändringsbok mer tillförlitliga också för personer under 50 år. Femtioen testpersoner i åldrarna 25-49 testades med tre månaders mellanrum. Resultaten visade mestadels marginella skillnader i förändring mellan olika åldersgrupper samtidigt som de kritiska gränsvärdena för förändring justerades något. Utökningen av normvärdena har gjort Förändringsboken mer representativ. Undersökningen påvisade emellertid större inlärningseffekter bland yngre testpersoner, varför det kan finnas anledning att dela upp Förändringsboken i (minst) två delar – en för äldre och en för yngre.

Under de senaste åren har den kirurgiska behandlingen av låggradiga gliom (LGG), det vill säga långsamtväxande hjärntumörer, förändrats markant. Patienterna behandlas tidigare och mer aggressivt, vilket innebär att man eftersträvar att avlägsna mer drabbad vävnad än vad som var fallet tidigare. Dessa förändringar har visat sig förlänga patienternas överlevnad. Andra effekter än just den förlängda överlevnaden är otillräckligt kartlagda (Jakola, 2018). På Sahlgrenska Universitetssjukhuset har man därför påbörjat ett forskningsprojekt som syftar till att belysa effekterna av den nya behandlingsmetodiken på neuropsykologiska funktioner, psykisk hälsa, mental trötthet/uttröttbarhet, arbetsförmåga och livskvalitet (Jakola, 2018). Ett delprojekt inom forskningsstudien är att utveckla tillförlitliga förändringsindex (reliable change indices; RCI) för elva neuropsykologiska test. Detta görs genom testning av friska individer med intervall om tre månader och ett år, vilket motsvarar intervallen för kliniska uppföljningar av LGG-patienter som opererats. Testningarna används för att fastslå gränsvärden för vad som är att betrakta som normala fluktuationer (Jakola, 2018).

Rey Auditory Verbal Learning Test

Ett av de neuropsykologiska testinstrument som används är Rey Auditory Verbal Learning Test (RAVLT). RAVLT är väl etablerat och används både i kliniska sammanhang och i forskning. Det är ett av de äldsta kognitiva testen som fortfarande används (Lehmann, Marks & Hanstock, 2013; Saury & Emanuelson, 2017). En av anledningarna till dess popularitet är att det mäter flera minnesfunktioner, så som omedelbar och fördröjd retention, inlärning, igenkänning, proaktiv och retroaktiv interferens samt primacy- och recencyeffekter. Att ett flertal minnesfunktioner mäts samtidigt ökar testets känslighet som diagnostiskt verktyg (Lehmann et al. 2013; Saury & Emanuelson, 2017; Vakil & Blachstein, 1997). RAVLT är ett kraftfullt kliniskt instrument trots sin enkelhet. Det har visat sig känsligt för neurologisk försämring (Schoenberg et al. 2006), men relativt okänsligt för affektiva

störningar (Schmidt, 1996). RAVLT är ett tillförlitligt verktyg för att bedöma funktionen i hippocampus samt för tidig upptäckt av kognitiv dysfunktion (Saury & Emanuelson, 2017).

Att bedöma förändring

Upprepade bedömningar är vanligt förekommande i klinisk neuropsykologi och används för att följa den naturliga utvecklingen av ett tillstånd, för att bedöma återhämtning efter en neurologisk sjukdom eller skada eller för att utvärdera effekten av en behandling (Bird, Papadopoulou, Ricciardelli, Rossor & Cipolotti, 2004; Duff, 2012; McCaffrey & Westervelt, 1995). Det är en svår uppgift för kliniska neuropsykologer att avgöra om en patient genomgått en betydande förändring i jämförelse med en tidigare mätning (Duff, 2012; McCaffrey & Westervelt, 1995). En observerad förändring i testresultat vid upprepade mätningar kan vara en indikation på en faktiskt förändrad förmåga, men den kan också vara ett uttryck för mätfel (Duff, 2012; McCaffrey & Westervelt, 1995). Målet är att komma så nära den verkliga förändringen av kognitiv förmåga som möjligt genom att minska eller kontrollera andelen av variansen som kan bero på andra faktorer. De felkällor som är mest relevanta för upprepade bedömningar kan delas in i tre grupper: Faktorer associerade med testet, faktorer associerade med testsituationen och faktorer associerade med testpersonen (tp)(Duff, 2012).

Faktorer associerade med testet

Ett neuropsykologiskt tests värde som diagnostiskt verktyg kan definieras som dess förmåga att korrekt skilja friska personer från personer med neuropsykologisk dysfunktion. Det skall inte klassificera friska individer som neurologiskt dysfunktionella (falska positiva) eller neurologiskt skadade individer som friska (falska negativa) (Schoenberg et al. 2006). Ett tests *sensitivitet* utgörs av dess förmåga att korrekt identifiera neurologisk dysfunktion. Hög sensitivitet är särskilt viktigt då sjukdomen eller skadan som testet avses identifiera är farlig. Testets *specificitet* utgörs av dess förmåga att korrekt identifiera friska personer som friska. Detta är särskilt viktigt ifall den behandling som kan bli aktuell är farlig. *Positivt prediktivt värde* (PPV) utgörs av andelen neurologiskt dysfunktionella inom den grupp som testet pekat ut som avvikande, vilket annorlunda uttryckt innebär att PPV anger hur troligt det är att patienten faktiskt är sjuk givet att han fått ett avvikande resultat. *Negativt prediktivt värde* (NPV) utgörs av andelen friska personer i den grupp som testet identifierat som icke-avvikande (Lagerlund & Zackrisson, 2013). Det är viktigt att ta hänsyn till tp:s ursprungliga resultat (fortsättningsvis benämnt baslinjen) vid bedömning av förändring. Om baslinjen är väldigt låg eller väldigt hög kommer möjligheten för testet att registrera eventuell förändring hos individen att begränsas av *golv- och takeffekter* (Duff, 2012). Om medelvärdet för baslinjen ligger väldigt högt eller lågt för gruppen som helhet kommer testets förmåga att registrera förändring på gruppnivå att begränsas av golv- eller takeffekter. Enkelt uttryckt kommer golveffekter att maskera eventuella försämringar, medan takeffekter kommer att maskera eventuella förbättringar.

I samband med neuropsykologisk bedömning finns det två grundläggande psykometriska egenskaper som avgör ett mätinstruments kvalitet: dess validitet och dess reliabilitet. *Validiteten* är ett mått på i vilken utsträckning testet faktiskt mäter det som testet är avsett att mäta. *Reliabiliteten* avser testets mätprecision eller tillförlitlighet och uttrycks ofta med en reliabilitetskoefficient som (i likhet med korrelationskoefficienter) sträcker sig från +1.0 över 0.0 till -1.0. En hög reliabilitetskoefficient innebär att individer i stor

utsträckning behåller sin relativa position inom fördelningen av poäng från ett testtillfälle till nästa (Duff, 2012; McCaffrey & Westervelt, 1995).

Saury och Emanuelson (2017) har granskat litteraturen om förhållandet mellan RAVLT och underliggande neurala strukturer. De fann stöd för att RAVLT ger tillförlitliga indikatorer på hippocampusfunktion och ansåg att RAVLT är ett bra verktyg för att bedöma och tidigt fånga upp kognitiva förändringar (Saury & Emanuelson, 2017). Poreh, Tolfo, Krivenko och Teaford (2017) undersökte RAVLT:s sensitivitet och specificitet genom att granska arkivdata och kom fram till att RAVLT genererar relativt få falska positiva och en måttlig andel falska negativa svar (Poreh et al., 2017). Magalhaes, Malloy-Diniz och Hamdan (2012) studerade intern konsistens och test-retest-reliabilitet i RAVLT med ett genomsnittligt tidsintervall på 35 dagar mellan testtillfällena. För att bedöma konvergerande (i hur stor utsträckning ett test samvarierar med ett annat test som är avsett att mäta samma sak) och divergerande (i hur stor utsträckning variansen i ett test kan sägas vara oberoende av variansen i ett annat test som är avsett att mäta någonting annat) validitet jämfördes resultaten på RAVLT med Benton Visual Retention Test (BVRT) och Trail Making Test (TMT). Värden på RAVLT var inte signifikant korrelerade med värden på TMT (från -0.19 till -0.03). Däremot uppvisade värden på RAVLT signifikanta men svaga korrelationer med värden på BVRT (från 0.37 till 0.44). Alla test-retest-reliabilitetskoefficienter var signifikanta, mellan 0.36 och 0.68. Summan av försök 1-5, d.v.s. totalvärdet, hade den bästa reliabiliteten ($r = 0.68$). Resten av måtten hade måttligt höga reliabilitetskoefficienter. Cronbachs alpha var emellertid 0.80. (Magalhaes et al., 2012). Resultaten i ovan nämnda studier tyder på att RAVLT har god validitet och reliabilitet och därför kan anses vara ett giltigt och tillförlitligt mått på episodiskt verbalt minne.

Inlärningseffekter. Vid upprepad testning kan förbättrat resultat uppstå på grund av tidigare exponering för testmaterialet (Duff, 2012; Magalhaes et al 2012). Förbättringar på grund av inlärningseffekter är förmodligen relaterade till både deklarativt minne (att komma ihåg de faktiska objekten på testen) och icke-deklarativt minne (vilket kanske bäst låter sig beskrivas som en känsla av förtrogenhet med testsituationen). Inlärningseffekten varierar mellan olika test och är ett av de mest utforskade fenomenen i upprepad testning (Duff, 2012). I många fall är inlärningseffekter viktiga att ta i beaktande. Om ingen korrigering görs, riskerar man att klassificera faktiskt oförändrade individer som förbättrade och faktiskt försämrade individer som oförändrade (Duff 2012; McCaffrey & Westervelt, 1995). Inlärningseffekter kan konceptualiseras endera som en källa till oönskad varians som behöver adresseras eller som en faktor som ska vägas in i den kliniska bedömningen; där avsaknad av inlärningseffekt då skulle utgöra ett observandum (McCaffrey & Westervelt, 1995). Inlärningseffekter är viktiga att ta hänsyn till även då testen har hög test-retest-reliabilitet, eftersom förskjutningar av fördelningar inte påverkar reliabilitetskoefficienten (Duff, 2012; McCaffrey & Westervelt, 1995). RAVLT har flera parallella testversioner vilket minskar inlärningseffekten något (Duff, 2012). De icke-deklarativa aspekterna av inlärningseffekten kvarstår emellertid såvida man inte genomför testningarna vid fler tillfällen än två och sedan bortser ifrån resultaten av den allra första administrationen (vilken då betraktas mer som en ”övningsomgång”), ett förfarande som ibland används i kliniska provningar.

Faktorer associerade med testsituationen

Den omgivande miljön vid testsituationen (t.ex. rumstemperatur, ljus, ventilation, buller etc.) kan påverka resultatet (McCaffrey & Westervelt, 1995). Ett sätt att minimera felkällor på grund av faktorer i testsituationen är att standardisera testsituationen och minimera skillnaderna mellan olika testsituationer. RAVLT innehåller tydliga och specifika

instruktioner, vilket minskar variationer i olika testledares presentation av uppgiften (Magalhaes et al., 2012). Om fokus för den aktuella undersökningen är förändringsvärden så är det viktigt att hålla alla situationella faktorer konstanta över samma individs båda mätillfällen, till exempel genom att se till att det är samma testledare vid individens båda mätningar. Små skillnader i administrering kan få effekt på individens resultat (Hessen, 2011; Magalhaes et al., 2012).

Individer har visat sig prestera bättre på vissa kognitiva uppgifter om de testas på sin föredragna tid på dagen, ett fenomen som kallas *synkroniseringseffekten* (Lehmann et al. 2013). Det finns tecken på en förutsägbar förändring från kvälls- till morgonpreferens under den vuxna livslängden. Synkroniseringseffekten är starkare hos äldre och Lehmann et al. (2013) menar att synkroniseringseffekten delvis kan förklara den väl dokumenterade åldersrelaterade prestationsförsämringen på RAVLT. Yngre vuxnas prestation tenderar att förbättras under dagen medan äldres prestationsnivå tenderar att sjunka. Detta innebär att åldersskillnader blir mindre när äldre och yngre vuxna testas på morgonen men tilltar när testningen genomförs på eftermiddags- eller kvällstid (Lehmann et al. 2013). Tiden på dagen är alltså ytterligare en faktor som kan påverka variationen både mellan individer och mellan två mätningar för samma individ.

I allmänhet leder kortare retestintervall till högre reliabilitetskoefficienter, och längre intervall leder till lägre reliabilitetskoefficienter. Längre intervall minskar, men eliminerar inte nödvändigtvis, inlärningseffekten (Duff, 2012).

Regression mot medelvärdet. Ett statistiskt fenomen som är värt att nämna är regression mot medelvärdet. Detta innebär att vid upprepade försök eller testningar som påverkas av slumpmässiga variabler, så kommer det att finnas individer som vid första testningen erhåller extremvärden, för att vid nästa testning erhålla mindre extrema resultat och således närma sig medelvärdet (Borg & Westerlund, 2012; Duff, 2012; Hinton-Bayre, 2010). Detta innebär i praktiken att man vid upprepade mätningar kan förvänta sig resultat som ligger närmare medelvärdet, utan att någon som helst intervention genomförts (Borg & Westerlund, 2012; Duff, 2012; Hinton-Bayre, 2010). Motsatsen förekommer också, det vill säga att de som vid första mätningen presterat nära medelvärdet tenderar att få något mer avvikande resultat vid andra mätningen (Sjöberg, 2006). Detta innebär att det vid en behandlingssituation kan vara svårt att avgöra vad som faktiskt har förändrats på grund av behandlingen. Detta gör det extra viktigt att ha data från en frisk kontrollgrupp, eftersom behandlingar annars kan visa positiva effekter vare sig behandlingen är effektiv eller ej (Sjöberg, 2006).

Faktorer associerade med den enskilda patienten

Egenskaper hos individen kan påverka testresultatet på flera sätt. Individens allmänna färdigheter och generella förmåga att förstå instruktioner kan vara viktiga faktorer tillsammans med förmågan att lösa de typer av problem som presenteras i testet. Individens känslomässiga reaktion, självförtroende och erfarenheter av testsituationer och liknande uppgifter har också betydelse (Duff, 2012; McCaffrey & Westervelt, 1995). Fluktuationer i uppmärksamhet, koncentration och minne orsakar en del av variansen, liksom trötthet, motivation och emotionellt tillstånd (McCaffrey & Westervelt, 1995). Individernas kliniska tillstånd är också av betydelse för deras prestationer. En individ kan befinna sig i ett sjukdomstillstånd vid den ena mätningen men inte vid den andra, samma eller olika sjukdomstillstånd kan närvara vid båda mätningarna men i olika allvarlighetsgrad. Det är viktigt för neuropsykologer att överväga vikten av sådana förhållanden (Duff, 2012). Skillnader i hur individen tolkar uppgiften och vilken strategi som används kan också få stor

betydelse. Hessen (2011) fann att individer som vid testning med RAVLT använder sig av strategin att upprepa listan med ord mer än en gång tenderar att få bättre resultat. Det är svårt för bedömare att kontrollera sådana faktorer, samtidigt kan de vara avgörande källor till variationen mellan olika undersökningar (McCaffrey & Westervelt, 1995). Vissa av ovan nämnda faktorer kan anses mer stabila och påverkar i större utsträckning variationen mellan individer än variationen mellan två mättillfällen för samma individ. Andra faktorer, så som fluktuationer i uppmärksamhet, koncentration, trötthet, motivation, emotionellt tillstånd och individens kliniska tillstånd kan anses ha större inverkan på variationen inom individen (McCaffrey & Westervelt, 1995).

Demografiska variabler. Tidigare forskning har visat att prestationen på RAVLT påverkas signifikant av demografiska variabler, i synnerhet ålder (Knight, McMahon, Skeaff & Green, 2007; Lehmann et al. 2013; Savage & Gouvier, 1992; Schoenberg et al. 2006; Vakil & Blachstein, 1997; Van der Elst, Van Boxtel, Van Breukelen & Jolles, 2005). Vissa studier indikerar att förändringar i det verbala minnet var måttliga fram till 60 års ålder jämfört med de förändringar som observerats efter 60 års ålder (Vakil & Blachstein, 1997). Snabbare förändringar av minnesförmågan i den äldre åldersgruppen skulle medföra en ökad betoning på vikten av normer baserade på grupper med små åldersintervall, särskilt i höga åldrar (Vakil & Blachstein, 1997). Andra studier har motsatt sig detta och påvisar ett linjärt samband mellan verbalt minne och ålder där en försämring av förmåga syns tidigare (Van der Elst et al. 2005). Hörselnedsättning skulle kunna vara en bidragande orsak till sämre prestation på RAVLT hos äldre (Van der Elst et al. 2005).

Flera studier visar att kvinnor presterar bättre än män på RAVLT i alla åldersgrupper och i alla deluppgifter (Vakil & Blachstein, 1997; Van der Elst et al. 2005), medan andra inte kunnat påvisa några könsrelaterade skillnader (Savage & Gouvier, 1992). När det gäller utbildning är den negativa effekten av en låg utbildningsnivå, jämfört med en genomsnittlig utbildningsnivå, större än den positiva effekten av en hög utbildningsnivå (Van der Elst et al. 2005).

Beräkning av reliabla förändringsindex

Det finns olika metoder för att beräkna reliabla förändringsindex. Den första metoden introducerades 1991 av Jacobson och Truax och rörde dels förslag på hur man bedömer om en person tillhör en frisk eller en avvikande population, dels hur stor förändringen ska vara för att den ska räknas som statistiskt tillförlitlig (Rydén, 2012). Sedan dess har det tillkommit förslag på förbättringar kring hur detta ska beräknas och dessa har främst handlat om hänsyn till inlärningseffekter samt regression mot medelvärdet. Demografiska variabler som till exempel ålder, kön och utbildning anses kunna påverka inlärningseffekten och därför har vissa metoder tagit hänsyn till det i sina beräkningar. Andra menar dock att det finns för få studier som stödjer att dessa variabler har en effekt vid upprepade mätningar (Rydén, 2012).

Maassen (2003) har introducerat en metod som tar hänsyn till regression mot medelvärdet och som således ställer olika krav på förändringens storlek dels beroende på hur värdet vid första mätningen förhåller sig till medelvärdet, dels beroende på hur spridningarna förändras mellan de två mättillfallen. Maassens metod väger inte in demografiska variabler så som t.ex. ålder, kön och utbildning. Metoden använder testresultaten, reliabilitetskoefficienten, spridningen vid de båda mättillfallen och individens ursprungliga prestation i förhållande till kontrollpersonernas medelvärde som faktorer i uträkningen (Rydén, 2012).

Förändringsboken

Ett sätt att hantera utmaningarna med inlärningseffekter och demografiska variabler så som ålder och kön är att upprätta förändringsnormer där graden av inlärningseffekter associerade med olika tidsintervall och demografiska grupper framgår (McCaffrey & Westervelt, 1995). För att avgöra hur stora fluktuationer som kan anses normala behövs normer för vilka skillnader i resultat friska individer uppvisar när de testas vid två tillfällen. Förändringsnormerna kan sedan användas i diagnostiskt och prognostiskt syfte, till exempel för att följa en sjukdomsprogress, bedöma effekterna av ett neurologiskt trauma eller effektiviteten i en behandlingsmetod (Ryden, 2012). För RAVLT finns normer för testning vid ett tillfälle (Schmidt, 1996). Det råder dock brist på normer för upprepade mätningar, vilket gör det svårt att avgöra hur stora resultatförändringar som med säkerhet kan betraktas som faktiska förändringar i den förmåga som avses mätas. Detta är t.ex. aktuellt då man vill studera vilka kognitiva konsekvenser den nya och mer aggressiva behandling många LGG-patienter erhåller, nu får. Inom det tidigare nämnda forskningsprojektet på Sahlgrenska har arbetet med att utforma en sådan databas av förändringsnormer påbörjats. Förändringsboken, som programmet för beräkning av förändring heter, skapades ursprungligen av Isabelle Rydén och Per Hellström inom ramen för ett examensarbete vid psykologiska institutionen (Rydén, 2012).

Förändringsboken är en Excel-fil som används för att kunna utvärdera resultatförändringar mellan två testtillfällen vid testning med RAVLT, Grooved Pegboard och Svenska Stroop. En databas med friska kontrollpersoners testresultat ligger till grund för en konvertering av resultatförändringen till ett z-värde, vilket beräknas enligt två olika metoder för förändring, Maassens och Temkins. När man fyller i en patients testresultat från två testningar i Förändringsboken, anges huruvida resultatet indikerar en signifikant förbättring, en signifikant försämring eller avsaknad av säkerställd skillnad enligt var och en av de båda metoderna på två olika konfidensnivåer; 90 respektive 95%. Med hjälp av Förändringsboken kan neuropsykologer snabbt och överskådligt få en statistisk uträkning och överblick av patientens förändring, vilket är till hjälp vid bedömning av förekomst av förändringar. Förändringsboken är ett komplement till andra metoder så som intervjuer med patienter och anhöriga för att ta del av deras egna förloppsbeskrivningar, andra klinikers omdömen rörande förbättringar/försämringar, jämförelser med avseende på förekomst av tidigare och i nuläget angivna symtom och tecken etc., som samtliga syftar till att avgöra om patienten uppvisar förändringar eller ej.

Syfte

Den friska kontrollgrupp som för närvarande ligger till grund för Förändringsboken består av äldre individer. För RAVLT utgörs kontrollgruppen av 129 individer i åldrarna 50-91 år (Rydén, 2012). Flera studier visar att ålder har en inverkan på prestation i RAVLT (Knight et al., 2007; Lehmann et al., 2013; Savage & Gouvier, 1992; Schoenberg et al., 2006; Vakil & Blachstein, 1997; Van der Elst et al., 2005) och det är tänkbart att det kan föreligga åldersskillnader också med avseende på förändringar mellan mättillfällen. För patienter i yngre åldrar kan därför Förändringsbokens normvärden vara missvisande. Om äldre normalt presterar något lägre på RAVLT, om de har en lägre inlärningseffekt och om deras resultat fluktuerar mer mellan testtillfällen, så riskerar yngre patienter att felbedömas när resultaten jämförs med en äldre normgrupp. Om ålder har en sådan inverkan som tidigare forskning indikerar, riskerar nuvarande version av Förändringsboken att generera falska negativa svar, d.v.s. bedöma en yngre patient med reell neurologisk försämring som oförändrad.

Syftet med det här arbetet var att utöka de redan befintliga normvärdena för friska kontrollpersoner som genomfört testet RAVLT två gånger med tre månaders mellanrum (Rydén, 2012). Genom att tillföra data från yngre kontrollpersoner kan normvärdena göras mer tillförlitliga för en bredare åldersgrupp. Resultatet kan användas för att t.ex. bättre kunna uttala sig om huruvida patienter som erhållit behandling för LGG förändrats mer eller mindre än eller inom ramen för vad som kan anses normalt för en frisk individ: i detta fall vad gäller verbal inlärnings- och minnesfunktion. Detta möjliggör i längden en mer korrekt och enklare utvärdering av den förändrade behandlingsstrategin.

Frageställningar

- Hur förändras normvärdena för test-retest på RAVLT med tre månaders mellanrum när vi adderar testresultat från 51 deltagare i åldrarna 25–49 år till en befintlig kontrollgrupp av 129 friska kontrollpersoner mellan 50–91 år?
- Hur påverkas de kritiska gränsvärdena för när en person anses vara förbättrad, försämrad eller oförändrad vid tillägget av yngre kontrollpersoner?
- Vilka skillnader finns mellan den befintliga kontrollgruppen bestående av i åldrarna 50-91 år och en yngre kontrollgrupp bestående av deltagare i åldrarna 25-49 år med avseende på förändringar mellan två mättillfällen?
- Vilka skillnader finns mellan olika åldersgrupper med avseende på förändringar mellan två mättillfällen?

Metod

Deltagare

Rekrytering och testning av deltagare genomfördes i samarbete med två psykologkandidater, Fredrika Jensen och Louise Edelberg. Sammanlagt rekryterades 52 deltagare, varav vi rekryterade och testade 25. Rekryteringen började i oktober 2018 och pågick under hösten 2018. En person föll bort efter första testningen då det visade sig att den inte uppfyllde hälsokriterierna för deltagande i studien. Hens resultat är därför exkluderade från analysen. Kontrollgruppen bestod således av 51 deltagare.

Deltagarna rekryterades främst genom sociala medier, i form av inlägg på Facebook och personliga meddelanden. I inläggen stod en kort sammanfattning om studiens syfte, att man behövde uppfylla ett antal fysiska och psykiska hälsokriterier för att kunna delta, att vi sökte personer i åldrarna 25–49 år samt att en biobiljett utgick som ersättning vid varje testtillfälle.

Dessa urvalskriterier bestämdes i enlighet med den pågående studien på Sahlgrenska Universitetssjukhus (Jakola, 2018). Åldersspannet 25–49 år valdes dels utifrån vad som bäst skulle komplettera den aktuella Förändringsboken, som då bestod av 129 personer i åldrar över 50 år, dels skulle den bättre matcha gruppen LGG-patienter. Vi försökte få en jämn spridning inom hela åldersspannet, med ca 10 personer inom varje femårs spann, samt en jämn fördelning av män och kvinnor.

Tabell 1

Tabell över ålders- och könsfördelning

Ålder	Fördelning kvinnor/män	Antal
-------	------------------------	-------

25-29	6/7	13
30-34	4/6	10
35-39	6/4	10
40-44	4/6	10
45-49	5/4	9

Hälsokriterierna för att kunna delta som frisk kontrollperson baserades på de kriterier som använts inom det större projektet. Information om detta samlades in genom att deltagarna fyllde i en så kallad "Hälsodeklaration" (se bilaga 1). Detta gjordes för att säkerställa att deltagarna var somatiskt, neurologiskt och psykiskt friska. Hälsodeklarationen innehöll frågor om kön, ålder och antal utbildningsår, samt frågor kring huruvida man varit drabbad av något av följande: Stroke, skallskada, allvarlig hjärnskakning, neurologisk sjukdom som t.ex. epilepsi, demenssjukdom eller ALS, hjärt-kärlsjukdom samt diabetes och högt blodtryck. Dessa fanns med eftersom de är somatiska besvär som anses kunna påverka kognitionen och därmed testresultatet (Rydén, 2012). Alternativen hjärnskakning, hjärt-kärlsjukdomar och diabetes fick förekomma i lindrigare former och inkluderades i studien eftersom tillstånden anses vara tillräckligt medicinskt stabila för att inte påverka kognitionen. Frågor om neuropsykiatriska diagnoser som ADHD, ADD eller autism samt allvarligare psykisk sjukdom lades också till, eftersom detta också riskerar att påverka testresultatet. Lindrigare tillstånd av psykiatriska störningar som exempelvis lindrig depression eller ångestproblematik inkluderades i studien.

De personer som anmälde sitt intresse för att delta fick en mer detaljerad beskrivning av studien. Därefter informerades de om de praktiska detaljerna kring hur testningarna skulle gå till och tid bokades för första testningen.

Tillvägagångssätt

Deltagarna testades två gånger med tre månaders mellanrum. Första testomgången pågick under november och december 2018 och den andra testomgången pågick under februari och mars 2019.

De villkor som skulle uppfyllas rörande testsituationen var att testningen skulle genomföras i ett neutralt rum och i möjligaste mån av samma testledare vid båda tillfällena. Lokalerna skulle vara avskilda, obekanta för deltagaren och så neutrala som möjligt. Samtliga tp testades i avskilda rum på Psykologiska Institutionen, förutom en som testades på sin arbetsplats av praktiska skäl (men även där i ett avskilt och neutralt rum). Deltagarna delades upp så att alla blev testade av för dem obekanta testledare. Alla förutom en testades av samma testledare vid båda tillfällena.

Vid det första testtillfället blev deltagarna ombudda att skriva under ett samtyckesformulär samt fylla i hälsodeklarationen skriftligen. Därefter genomfördes testningen som tog ungefär 90 minuter vid första tillfället, och ungefär 60 minuter vid andra tillfället eftersom ett mindre antal test skulle administreras andra gången. Efter det första testtillfället bokades ett preliminärt datum för andra testningen.

Instrument

Då föreliggande uppsats är en del i ett större forskningsprojekt administrerades ett testbatteri som innehöll sammanlagt elva olika test. Självskattningsformulären MFS (Mental Fatigue Scale), HADS (Hospital Anxiety and Depression Scale) och EORTC QLQ-C30 och

BN20 samt EQ5D (livskvalitetsskattningar) administrerades också vid båda testtillfällena. Tabell 2 presenterar inkluderade test och deras funktion.

Tabell 2

Utförda test och deras funktion

Test	Funktion
BVMT-R	Visuell inläring och minne
RAVLT	Auditiv verbal inläring och minne
WAIS IV, Kodning	Psykomotorisk hastighet, komplex uppmärksamhet
Trail Making Test A och B	Visuell avsökning och kognitiv flexibilitet
Grooved Pegboard	Finmotorisk hastighet
Rey Complex Figure	Exekutiv funktion
WAIS IV, Sifferrepetition	Uppmärksamhet och arbetsminne
D-KEFS Color Word Test	Inhiberingsförmåga och verbal kognitiv flexibilitet
Ordflöde FAS	Verbalt flöde (fonemiskt), exekutiva funktioner
Ordflöde Djur	Verbalt flöde (semantiskt)
Boston Naming Test	Benämningsförmåga, perception

Rey Auditory Verbal Learning Test. Testet bestod av en lista (A) med 15 orelaterade konkreta substantiv som presenterades under fem inlärningsförsök, där listan lästes upp för tp som ombetts lägga orden på minnet och sedan återge så många som möjligt. Detta upprepades fem gånger. Efter det lästes en interferenslista (B) med 15 nya ord upp. Tp ombads återge så många ord som möjligt från lista A direkt efter att ha återgett orden från lista B. Trettio minuter senare ombads tp återge orden från lista A igen och avslutningsvis uppmanades hen att identifiera orden från lista A, från en ordlista med tre olika alternativ (Saury & Emanuelson, 2017; Schmidt, 1996). RAVLT finns i flera parallella versioner med olika uppsättningar av ord. Samtliga deltagare testades med version 1 vid först tillfället och sedan testades varannan deltagare med version 2 respektive 3 vid andra tillfället.

Databearbetning

Resultaten från samtliga tester sammanställdes och fördes in i separata exceldokument. Därefter fördes de över till en lösenordskyddad databas. Därifrån hämtades sedan information om samtliga deltagares ålder samt testresultaten från båda testningarna med RAVLT och överfördes till Förändringsboken.

Analyserna i föreliggande arbete rör förändringar av tre av variablerna i testet RAVLT: Maxprestationen under de fem inlärningsförsöken med lista A (R_{max}), där max är 15, totalpoängen (R_{tot}) som innebär totalsumman av ihågkomna ord från de första fem försöken och där max är 75, samt prestationen vid fördröjd återgivning (R_8) som mäter hur många ord personen kommer ihåg efter 30 minuters fördröjning, där max är 15. Anledningen till urvalet var att dessa variabler visat sig ha de högsta reliabilitetskoefficienterna enligt en tidigare undersökning (Rydén, 2012).

Våra data analyseras med hjälp av Maassens metod för beräkning av reliabla förändringsindex och har sedan också kommit att inkluderas i Förändringsboken så att denna nu är baserad på ett vidare åldersspann än tidigare.

Förändringsboken. Data från 51 kontrollpersoner fördes in i Förändringsboken och databasen utökades därmed med tre RAVLT-variabler. För dessa variabler som tidigare bestod av data från 129 personer mellan 50 och 91 år utökades Förändringsboken till att omfatta 180 personer i åldrarna 25 till 91 år. Därefter beräknades och uppdaterades Förändringsboken med nya medelvärden och standardavvikelser. Resultatet kallar vi Nya förändringsboken och hädanefter har vi därför valt att kalla den tidigare versionen för Gamla förändringsboken. Beräkningar gjordes med hjälp av Microsoft Office Excel version 16.23. Genomsnittliga förändringar i tidigare tillgängliga data (d.v.s. för åldrarna 50–91) jämfördes med de genomsnittliga förändringarna för de nya, yngre kontrollpersonerna. Jämförelserna genomfördes med t-test i Statistical Package for the Social Sciences (IBM, version 25 för Windows, Chicago, USA).

Resultat

Skillnader mellan Gamla och Nya förändringsboken

Skillnader mellan Gamla och Nya förändringsboken illustreras i tabell 3 nedan, där Gamla förändringsbokens värden står först, sedan den Nya förändringsboken och sist differensen mellan de två versionerna. Tabellen visar medelvärde och standardavvikelse för första mättillfället (Mx, sdx) och andra mättillfället (My, sdy) samt medelvärdet av skillnaden mellan första och andra testningen (MDdiff) och slutligen reliabilitetskoefficienten för de två testningarna (rxy).

Tabell 3

Normvärden för Gamla och Nya förändringsboken

Gamla förändringsboken

Kategori	Antal	Mx	sdx	My	sd _y	MDiff	rxy
RMax	n=129	12.88	2.05	13.16	1.89	0.29	0.70
RTot	n=129	49.80	9.25	51.29	9.74	1.50	0.80
R8	n=129	11.12	2.81	11.08	2.86	-0.04	0.72

Nya förändringsboken

Kategori	Antal	Mx	sdx	My	sd _y	MDiff	rxy
RMax	n=180	13.18	1.94	13.48	1.78	0.29	0.72
RTot	n=180	52.17	9.69	54.49	10.45	2.32	0.82
R8	n=180	11.63	2.79	11.71	2.86	0.08	0.75

Differens Gamla och Nya

RMax	0.31	-0.11	0.31	-0.12	0.01	0.02
------	------	-------	------	-------	------	------

RTot	2.37	0.45	3.20	0.70	0.83	0.02
R8	0.52	-0.02	0.63	-0.00	0.12	0.03

Mx = medelvärde för första testningen, sdx = standardavvikelse för första testningen, My = medelvärdet för andra testningen, sdy = standardavvikelse för andra testningen, MDiff = medelvärdet av skillnaden mellan första och andra testningen, rxy = reliabilitetskoefficienten för de två testningarna.

Av tabell 3 framgår att medelvärdena för prestationerna vid båda tillfällena för samtliga tre variabler var något högre i den yngre gruppen, medan standardavvikelserna var något mindre eller oförändrade för RMax och R8, medan standardavvikelsen för RTot var något större. Reliabiliteten var god för samtliga variabler.

Tabell 4 visar fördelningen av signifikant förbättrade (+), försämrade (-) och oförändrade (=) vid tillägg av 51 yngre personer med 90% konfidensintervall. Fördelningen av antal oförändrade, förbättrade och försämrade har betydelse då den är ett sätt att kvalitetssäkra Maassens modell (Maassen, 2003). Om modellen gett en perfekt motsvarighet till data skulle exakt 90% ha klassificerats som oförändrade, 5% som förbättrade och resterande 5% som försämrade om man valt den 90-procentiga konfidensnivån. Tabell 4 visar att värdena inte ligger helt fördelade enligt detta ideala värde, utan avviker något åt båda håll. I synnerhet avviker värdena vid andel förbättrade då det är färre som är förbättrade än vad som är statistiskt förväntat.

Tabell 4

Jämförande tabell över andel signifikant förbättrade (+), försämrade (-) och oförändrade (=) vid tillägg av 51 yngre personer, 90% konfidensintervall

	Gamla			Nya		
	-	=	+	-	=	+
Rmax	5.42	89.92	4.65	5	90	5
RTot	5.42	89.92	4.56	6.11	87.77	6.11
R8	5.42	92.24	2.32	4.44	92.22	3.33

Hur de kritiska gränsvärdena ändras i Nya förändringsboken

Tabell 5

Hur kritiska gränsvärden ändras från Gamla förändringsboken till Nya förändringsboken

	RMax		RTot		R8	
	Gamla	Nya	Gamla	Nya	Gamla	Nya
Antal ord vid första testningen	13	13	52	52	12	12
Antal ord vid andra testningen för att bedömas som förbättrad	Tak	Tak	64	65	Tak	Tak
Antal ord vid andra testningen för att bedömas som försämrad	10	11	43	44	8	8

Tabell 5 illustrerar hur de kritiska värdena, det vill säga med hur många ord en patient behövde förändras för att bedömas som signifikant förbättrad respektive försämrad, i Gamla respektive Nya förändringsboken. Till exempel ser man för variabeln RMax att individen i den Gamla förändringsboken behövde avvika med tre ord från första tillfället för att bedömas som försämrad, medan den i den nya endast behövde avvika med två ord. Likaså för variabeln RTot, medan variabeln R8 förblev oförändrad.

I exemplet nedan har vi för enkelhets skull angett ett resultat vid första mättillfället som motsvarar medelvärdet i den Nya förändringsboken. Det är dock viktigt att ha i åtanke att Maassens modell för förändring genererar olika bedömning beroende på hur nära medelvärdet första testningens resultat befinner sig, bland annat på grund av att modellen kompenserar för regression mot medelvärdet (Rydén, 2012). Detta innebär i praktiken att en individ som hade fått ett mer extremt värde, t.ex. en ovanligt låg poäng, hade behövt en mindre skillnad för att bedömas som försämrad. Tabell 5 visar också att det i vissa fall inte går att bedömas som förbättrad på grund av en takeffekt. Detta eftersom ett resultat som motsvarar medelvärdet ligger så högt att det inte finns utrymme att få tillräckligt avvikande poäng för att uppnå en skillnad, vare sig i Nya förändringsboken eller i Gamla förändringsboken.

Praktiskt exempel. För att illustrera detta ytterligare visar figur 1 vilket utfall dessa resultat gav när de matades in i Förändringsboken. Patientens resultat ansågs inte signifikant förändrade i Gamla förändringsboken och patienten klassificerades därför som "oförändrad" på alla mått utom fördröjd retention (R8) vid 90%. Exakt samma värden klassificerades i Nya förändringsboken som "signifikant försämrad" vid 90% på både Rmax, Rtot och R8. Vid R8 var patienten signifikant försämrad även vid 95%.

Ålder:	40
Kön:	K
Utb:	13

RAVLT max		
X_i	Y_i	D_i
13	11	-2

RAVLT total		
X_i	Y_i	D_i
52	44	-8

RAVLT 8		
X_i	Y_i	D_i
12	8	-4

Gamla förändringsboken			Nya förändringsboken		
Enligt Maassen (AdjRCIsrb)			Enligt Maassen (AdjRCIsrb)		
RCI	90%	95%	RCI	90%	95%
-1.497078	Oförändrad	Oförändrad	-1.6613217	Försämrad	Oförändrad
RCI	90%	95%	RCI	90%	95%
-1.6106641	Oförändrad	Oförändrad	-1.7281969	Försämrad	Oförändrad
RCI	90%	95%	RCI	90%	95%
-1.8916078	Försämrad	Oförändrad	-2.0577699	Försämrad	Försämrad

Figur 1. Exempel på resultat som ligger nära det kritiska gränsvärdet i Gamla respektive Nya förändringsboken.

Skillnader i normdata mellan en äldre och en yngre kontrollgrupp

Tabell 6 visar skillnaderna i normdata mellan den äldre kontrollgruppen med deltagare 50–91 år och den yngre kontrollgruppen 25–49 år.

Tabell 6

Normvärden för två åldersgrupper, 50-91 år samt 25-49 år

50-91 år

Kategori	Antal	Mx	sdx	My	sd _y	MDiff	rx _y
RMax	n=129	12.88	2.05	13.16	1.89	0.29	0.70
RTot	n=129	49.80	9.25	51.29	9.74	1.50	0.80
R8	n=129	11.12	2.81	11.08	2.86	-0.04	0.72

25-49 år

Kategori	Antal	Mx	sdx	My	sd _y	MDiff	rx _y
RMax	n=51	13.96	1.37	14.27	1.11	0.31	0.69
RTot	n=51	58.17	8.13	62.59	7.38	4.41	0.75
R8	n=51	12.94	2.29	13.31	2.16	0.37	0.73

Differens

RMax	1.08	-0.68	1.11	-0.78	0.02	-0.01
RTot	8.37	-1.12	11.3	-2.36	2.91	-0.05
R8	1.82	-0.52	2.23	-0.7	0.33	0.01

Mx = medelvärde för första testningen, sdx = standardavvikelse för första testningen, My = medelvärdet för andra testningen, sd_y = standardavvikelse för andra testningen, MDiff = medelvärdet av skillnaden mellan första och andra testningen, rx_y = reliabilitetskoefficienten.

Resultaten av jämförelser mellan förändringsvärden (MDiff) för yngre (<50 år) och äldre deltagare redovisas i tabell 7. Den yngre gruppen (MDiff = 4.41) hade en signifikant större inlärningseffekt på RTot än den äldre gruppen (MDiff = 1.50), $t(178) = 3.00$, $p < .01$, Cohen's $d = 1.08$.

Tabell 7

Jämförelse av förändringsvärden (MDiff) mellan yngre och äldre

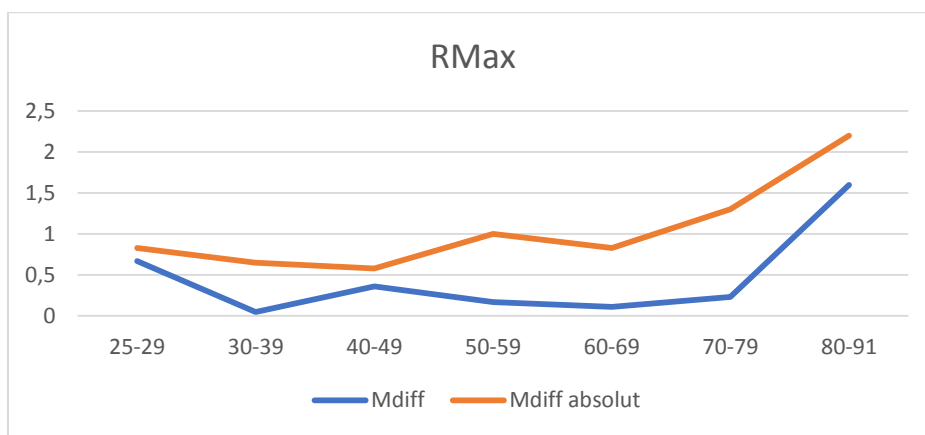
Variabel	Yngre (n=51)	Äldre (n=129)	t-värde	p-värde
RMaxDiff	0.31	0.29	0.14	.891
RTotDiff	4.41	1.50	3.09	.003*
R8Diff	0.37	-0.04	1.40	.165

Not. * $p < .05$

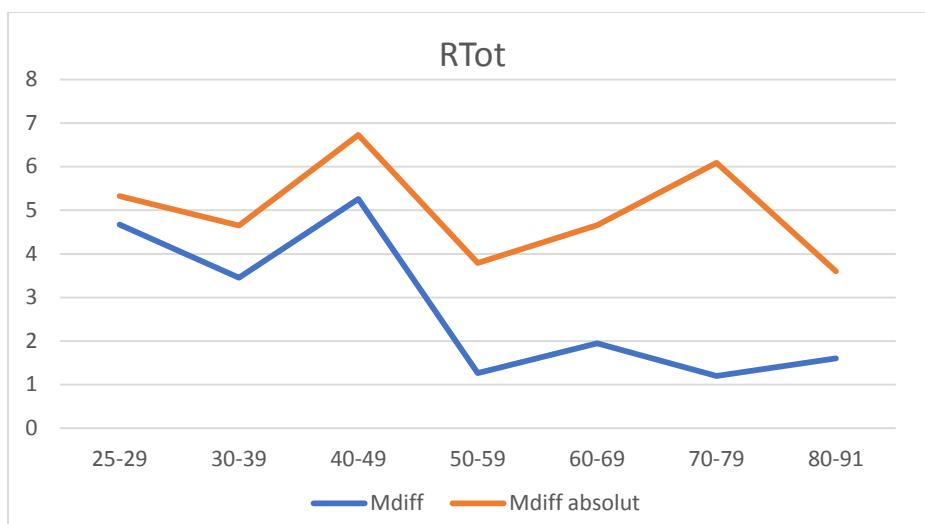
RMaxDiff = medelvärdet av skillnaden mellan första och andra testningen för RMax, RTotDiff = medelvärdet av skillnaden mellan första och andra testningen för RTot, R8Diff = medelvärdet av skillnaden mellan första och andra testningen för R8.

Skillnader i förändring mellan olika åldersgrupper

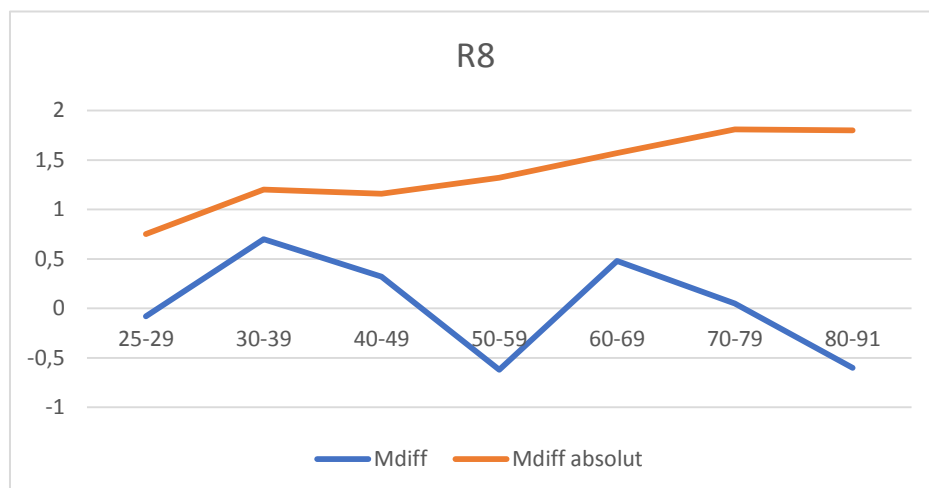
Figur 2, 3 och 4 illustrerar hur stora förändringarna är mellan två mättillfällen i respektive åldersgrupp för variablerna RMax, RTot och R8. MDiff är det totala medelvärdet av förändring mellan två mättillfällen medan MDiff absolut visar genomsnittlig avvikelse, det vill säga medelvärdet av de absoluta avvikelserna från medelvärdet. Detta innebär att plus och minus ignoreras, då vi endast är intresserade av skillnad från medelvärdet oavsett om den är positiv (förbättring) eller negativ (försämring).



Figur 2. Förändringarnas storlek mellan två mättillfällen i respektive åldersgrupp för variabeln RMax.



Figur 3. Förändringarnas storlek mellan två mättillfällen i respektive åldersgrupp för variabeln RTot.



Figur 4. Förändringarnas storlek mellan två mättillfällen i respektive åldersgrupp för variabeln R8.

Diskussion

Vi har, genom tillförandet av 51 friska kontrollpersoner i åldrarna 25–49, utökat den befintliga kontrollgruppen av 129 friska kontrollpersoner mellan 50–91 år som legat till grund för Förändringsbokens normdata. Skillnaderna mellan Gamla och Nya förändringsboken var små. Medelvärdet ökade marginellt för de tre variablerna RMax, RTot och R8. Standardavvikelsen ökade något i RTot men minskade något i RMax och R8.

Även om differensen mellan den Gamla och Nya förändringsboken var liten så flyttas de kritiska gränsvärdena något vid tillägget av de yngre kontrollpersonerna. Vi har illustrerat (figur 1) hur detta kan påverka bedömningen av en patient som ligger nära dessa gränsvärden. För en person vars första testning följer medelvärdet krävs det nu en poäng mindre i försämrad riktning på RMax och RTot för att klassificeras som försämrad medan det på R8 förblir oförändrat. Det indikerar att vi genom tillförandet av yngre kontrollpersoner ökat Förändringsbokens känslighet för försämringar. Det krävs en poäng mer i förbättrad riktning i RTot för att bedömas som förbättrad, medan det i de andra måtten inte går att fånga upp förbättring på grund av takeffektens inverkan när den ursprungliga prestationen motsvarar normgruppens genomsnittliga resultat (vilket dock sällan är fallet i kliniska fall).

Vid jämförandet av normdata från två separata åldersgrupper, en äldre med åldrar 50–91 och en yngre med åldrar 25–49 framgår att medelvärdena i samtliga variabler är högre i den yngre gruppen, framför allt vid andra mätningen. Detta indikerar att i vår data presterar yngre högre än äldre på RAVLT samt att yngre trots högre prestation vid baslinjemätningen uppvisar en större inlärningseffekt. Standardavvikelserna i samtliga variabler är mindre i den yngre gruppen, vilket skulle kunna förklaras av att den yngre kontrollgruppen genom sina höga baslinjeresultat i större utsträckning begränsas av en takeffekt som hindrar dem från att variera uppåt. Slutligen är medelvärdet för skillnaden mellan första och andra mätningen, MDiff, signifikant högre (2.91) hos de yngre i variabeln RTot. Detta indikerar att det, liksom tidigare forskning visat (Knight et al., 2007; Lehmann et al., 2013; Savage & Gouvier, 1992; Schoenberg et al., 2006; Vakil & Blachstein, 1997; Van der Elst et al., 2005), finns en större inlärningseffekt hos yngre och därmed en åldersrelaterad skillnad i vilka förändringsvärden som kan förväntas hos äldre individer jämfört med yngre individer.

Problematisering av resultat

Flera faktorer kan inverka på förändrade poäng vid omtestningar. Inläringseffekter, takeffekter, synkroniseringseffekter och demografiska variabler är några av dessa.

Inläringseffekter. Signifikant skillnad i inläringseffekter ses i kategorin RTot. Det kan argumenteras att denna variabel är den mest lämpade att se till vid jämförelsen mellan de två åldersgrupperna. Det är den variabel som begränsas minst av takeffekten (tabell 5) och den variabel som har högst reliabilitetskoefficient (tabell 6).

Standardadministrationen av RAVLT innebär att testledaren inte uppmanar testpersonen att använda någon särskild minnesstrategi, samtidigt tillåts testpersonen att använda olika strategier spontant. Vilken minnesstrategi testpersonen använder kan få stora konsekvenser på resultatet (Hessen, 2011). Vi noterade under vår administrering av de båda testomgångarna att flera individer bytte minnesstrategi mellan första och andra testtillfället. Vi har inte registrerat eller följt upp detta. Det kan argumenteras att ett byte av minnesstrategi ingår i inläringseffekten. I RAVLT byts ordlistan ut mot en annan version vilket minskar det deklarativa minnets inverkan men det procedurala minnet är fortfarande lika påverkat av inläringseffekter (Duff, 2012). Testpersonen kan utvärdera sin insats på första testtillfället och välja en mer framgångsrik minnesstrategi vid det andra.

Synkroniseringseffekten. Lehmann et al. (2013) menade att synkroniseringseffekten delvis kunde förklara den väl dokumenterade åldersrelaterade prestationsförsämringen på RAVLT, eftersom synkroniseringseffekten var starkare hos äldre. Vi såg en något större standardavvikelse hos den äldre kontrollgruppen jämfört med den yngre, vilket delvis skulle kunna förklaras av synkroniseringseffekten. Det hade varit intressant att samla in data om både när på dygnet testningarna gjordes, samt om detta var den tid som patienten föredrar. Då alla data från äldre individer hämtats från en redan befintlig databas där sådana uppgifter saknas var det dock inte aktuellt för oss.

Takeffekten. Vi såg både i Gamla och Nya förändringsboken att medelvärdet vid första testningen (baslinjen) var väldigt högt. Möjligheten för testet att registrera eventuell förbättring begränsas av takeffekten. Då tidigare forskning indikerat att yngre generellt presterar bättre på RAVLT (Knight et al., 2007; Lehmann et al. 2013; Savage & Gouvier, 1992; Schoenberg et al. 2006; Vakil & Blachstein, 1997; Van der Elst et al., 2005) hade vi förväntat oss att takeffekten skulle ha större påverkan på de yngre och att den eventuella variationen hos de yngre därför inte skulle fångas upp lika väl som hos de äldre. Vi såg en något större standardavvikelse hos den äldre kontrollgruppen jämfört med den yngre vilket delvis skulle kunna förklaras av en takeffekt. Eftersom takeffekten begränsat möjligheten för både Gamla och Nya förändringsboken att fånga upp hela spektrumet av förändring kan det hända att vi hade sett ännu större skillnader i förändring utan dess inverkan.

Det är dock viktigt att poängtera att de patienter med misstänkta verbala inlärnings- och minnessvårigheter som bedöms med hjälp av Förändringsboken kan förväntas ha ett betydligt lägre resultat vid baslinjemätningen och att takeffekter är mindre vanliga i testningar med RAVLT i kliniska sammanhang. Man kan också argumentera för att det är viktigast att fånga upp eventuella försämringar hos patienter t.ex. efter en erhållen kirurgisk och onkologisk behandling. Till det bör sägas att en utebliven förbättring också kan vara ett tecken på en mindre synlig försämring, något som RCI också syftar till att kontrollera för.

Studiens begränsningar

Rekrytering. Att vi främst rekryterat från sociala medier som Facebook kan vara en begränsning för studiens representativitet, då det exkluderar individer som inte har Facebook. Att kontrollpersonerna i huvudsak kommer från vårt kontaktnät kan också innebära en begränsning. Många i kontrollgruppen är högt utbildade, vilket skulle kunna påverka vårt resultat. Vissa studier (Van der Elst et al. 2005) visar dock att när det gäller utbildning är den negativa effekten av en låg utbildningsnivå, jämfört med en genomsnittlig utbildningsnivå, större än den positiva effekten av en hög utbildningsnivå. Dessutom kan man ha i åtanke att högre utbildning är förväntad vid rekrytering av yngre deltagare, då statistik från Statistiska Centralbyrån visar att utbildningsnivån generellt är mycket högre i de yngre åldrarna än bland äldre. (Statistiska Centralbyrån, 2018).

Administrering. Samtliga deltagare testades med version 1 av RAVLT vid första testningen. Vi förstod först efter första testningen att effekterna av RAVLT:s olika versioner skulle få mindre inverkan på resultatet ifall de varierades och alternerade därför mellan version 2 och version 3 vid andra testningen. Det hade varit önskvärt om vi hade varierat testversionerna från början.

Framtida utveckling

Det skulle vara fördelaktigt med fler och större databaser för förändringsvärden mellan test-retest då sådan normdata har en avgörande betydelse i den kliniska bedömningen av en sjukdomsprogress eller effekterna av en behandling. Vårt tillägg av yngre kontrollpersoner gör normdatamängden för förändringar i test-retest på RAVLT större och mer representativ. Nya förändringsboken blir på så sätt mer tillförlitlig. Samtidigt går det att argumentera för att den med vårt tillägg av yngre personer blivit något mindre representativ för äldre. Med tanke på att våra resultat visar på en signifikant skillnad i förändringsvärden mellan personer i åldrarna 25-49 år och personer i åldrarna 50-91 år, samt den forskning som antyder att förändringarna är måttliga före 60 års ålder (Vakil & Blachstein, 1997) för att sedan tillta, finns det anledning att överväga en uppdelning i (minst) två Förändringsböcker, en för yngre och en för äldre. Skiljelinjen mellan de båda Förändringsbäckerna kan dras någonstans runt 50 till 60 års ålder, alternativt att 50-59-åringarna kan inkluderas i båda Förändringsböckerna för att jämna ut glappet för de som är nära gränsen mellan de båda åldersgrupperna.

Referenser

- Bird, C. M., Papadopoulou, K., Ricciardelli, P., Rossor, M. N., & Cipolotti, L. (2004). Monitoring cognitive changes: psychometric properties of six cognitive tests. *British Journal of Clinical Psychology*, 43, 197-210.
- Borg, E., & Westerlund, J. (2012). *Statistik för beteendevetare*. Stockholm: Liber.
- Duff K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. *Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists*, 27, 248–261.
- Hessen, E. (2011). Rehearsal Significantly Improves Immediate and Delayed Recall on the Rey Auditory Verbal Learning Test. *Applied Neuropsychology*, 18, 263-268.
- Hinton-Bayre, A. D. (2010). Deriving reliable change statistics from test-retest normative data: comparison of models and mathematical expressions. *Archives of Clinical Neuropsychology*, 25, 244-256.
- Jakola, A. (2018). *Behandling av långsamtväxande hjärntumörer: dåtid, nutid och framtid*. Projektnummer: 245471. Hämtad 2019-05-09 från

- <https://www.researchweb.org/is/vgr/project/245471>
- Knight, R., McMahon, J., Skeaff, C., & Green, T. (2007). Reliable Change Index scores for persons over the age of 65 tested on alternate forms of the Rey AVLT. *Archives of Clinical Neuropsychology*, 22(4), 513-518.
- Lagerlund, M. & Zackrisson, S. (2013). *Screening: ett tilltalande men problematiskt koncept*. Läkartidningen, 110. Hämtad 2019-05-09 från http://www.lakartidningen.se/OldWebArticlePdf/1/19398/LKT1312s628_630.pdf
- Lehmann, C., Marks, A., & Hanstock, T. (2013). Age and synchrony effects in performance on the Rey Auditory Verbal Learning Test. *International Psychogeriatrics*, 25(4), 657-65.
- Maassen, G. H. (2003). Reliable change indices and practice effects: Principles of defining reliable change. *Nederlandsch Tijdschrift Voor de Psychologie*, 58, 69-79.
- Magalhaes, S. S., Malloy-Diniz, L. F., & Hamdan, A. C. (2012). Validity convergent and reliability test-retest of the Rey Auditory Verbal Learning Test.(Report). *Clinical Neuropsychiatry: Journal of Treatments Evaluation*, 9(3), 129-137.
- McCaffrey, R., & Westervelt, J. (1995). Issues associated with repeated neuropsychological assessments. *Neuropsychology Review*, 5(3), 203-221.
- Poreh, A., Tolfo, S., Krivenko, A., & Teaford, M. (2017). Base-rate data and norms for the Rey Auditory Verbal Learning Embedded Performance Validity Indicator. *Applied Neuropsychology: Adult*, 24(6), 540-547.
- Rydén, I. (2012). *När gör en skillnad en skillnad? Reliabila förändringsindex för upprepade mätningar med tre neuropsykologiska test*. (Masteruppsats). Göteborgs Universitet: Psykologiska Institutionen, Göteborg.
- Saury, J. M. & Emanuelson, I. (2017) Neuropsychological Assessment of Hippocampal Integrity. *Applied Neuropsychology: Adult*, 24(2), 140-151
- Schmidt, M. (1996). *Rey Auditory Verbal Learning Test: A Handbook (RAVLT)*. Los Angeles: Western Psychological Services.
- Schoenberg, M. R., Dawson, K. A., Duff, K., Patton, D., Scott, J. G. & Adams, R. L. (2006). Test performance and classification statistics for the Rey Auditory Verbal Learning Test in selected clinical samples. *Archives of Clinical Neuropsychology*, 21(7), 693-703.
- Sjöberg, R. L., (2006). *Behandlingseffekt eller regression mot medelvärdet?* Läkartidningen, 103. Hämtad 2019-05-09 från http://www.lakartidningen.se/OldWebArticlePdf/4/4391/LKT0626s2057_2058.pdf
- Statistiska Centralbyrån, (2018). *Utbildningsnivån I Sverige*. Hämtad 2019-05-09 från <https://www.scb.se/hitta-statistik/sverige-i-siffror/utbildning-jobb-och-pengar/utbildningsnivan-i-sverige/>.
- Vakil, E., & Blachstein, H. (1997). Rey AVLT: Developmental norms for adults and the sensitivity of different memory measures to age. *The Clinical Neuropsychologist*, 11(4), 356-369.
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of International Neuropsychological Society*, 11, 290-302.

Bilaga 1.

Hälsodeklaration

Namn:

/

Telefonnr och/eller epostadress:

/

Kön: _____

Ålder: _____

Antal skolår: _____

Testad av:

Datum:

Tid tillfälle 2:

Har du varit eller är du drabbad av:

	JA	NEJ
Stroke	<input type="checkbox"/>	<input type="checkbox"/>
Skallskada	<input type="checkbox"/>	<input type="checkbox"/>
Hjärnhinneinflammation	<input type="checkbox"/>	<input type="checkbox"/>
Diabetes	<input type="checkbox"/>	<input type="checkbox"/>
Hjärt-/kärlsjukdom	<input type="checkbox"/>	<input type="checkbox"/>
Högt blodtryck	<input type="checkbox"/>	<input type="checkbox"/>
Neurologisk sjukdom	<input type="checkbox"/>	<input type="checkbox"/>
Psykisk sjukdom	<input type="checkbox"/>	<input type="checkbox"/>

Med neurologisk sjukdom menas t.ex. epilepsi, MS, ALS, demenssjukdom Parkinsons sjukdom.

Med psykisk sjukdom menas t.ex. personlighetsstörning, tvångssyndrom, schizofreni, långvarig depression.

Om JA, får du eller har du tidigare fått någon behandling för detta?

JA **NEJ**

Ange vilken/vilka?

Använder du någon annan medicin regelbundet? **JA** **NEJ**

Ange vilken/vilka?
