

Tartu Ülikool  
Sotsiaalteaduste valdkond  
Psühholoogia instituut

Vivian Aun

**Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi valideerimine**

Magistritöö

Juhendaja: Uku Vainik, PhD

Läbiv pealkiri: Digitaalne neuropsühholoogia

**Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi valideerimine****Kokkuvõte**

Magistritöö on osa projektist „Käitumise detailsete mõõdikute eesti keelde adapteerimine”. Töö käigus valideeriti paralleeltesti meetodil sünonüümide sõnavara ja sõnapaaride ajendatud meenutamise testid. Sünonüümide sõnavara mõlemat testiversiooni sooritas 325 katseisikut ning sõnapaaride ajendatud meenutamise mõlemat testiversiooni 317 isikut.

Sarnaselt originaalversioonidele leidis kinnitust testide kõrge sisereliaablus – sünonüümide sõnavara testi Cronbach'i  $\alpha = .85$  (paralleelversiooni  $\alpha = .86$ ) ning sõnapaaride ajendatud meenutamise testi Cronbach'i  $\alpha = .88$  (paralleelversiooni  $\alpha = .87$ )

Töö eesmärk oli tuvastada ülesannete kogust mõõdetava võimega tugevalt ning nõrgalt seotud testiülesanded. Selleks viidi läbi mitmevälja (*polychoric*) faktoranalüüs. Tulemuste baasilt tehti ettepanekud testide täiendamiseks.

Autor soovib võtta täiendatud kujul testid kasutusele järgmisse etappi.

Märksõnad: Sünonüümide sõnavara, Sõnapaaride ajendatud meenutamine, digitaalne neuropsühholoogia, *Test My Brain*.

## Assessing the Validity of Vocabulary and verbal Paired Associates Tests

### Abstract

The master's thesis is a part of the project "*Käitumise detailsete mõõdikute eesti keelde adapteerimine*" In the course of the research, Vocabulary and Verbal Paired Associates tests were validated using the parallel testing method. Both versions of the vocabulary test were taken by 325 participants, and both versions of the word-pair tests were taken by 317 individuals.

In line with the original versions, the tests demonstrated high internal consistency - the Cronbach's  $\alpha$  for the synonym vocabulary test was .85 (parallel version  $\alpha = .86$ ), and the word-pair prompted recall test showed a Cronbach's  $\alpha$  of .88 (parallel version  $\alpha = .87$ ).

The aim of the study was to identify test tasks that were both strongly and weakly correlated with the measured ability. This was achieved through the implementation of a polychoric factor analysis. Based on the results of this analysis, recommendations were proposed for test enhancements.

The author recommends implementing the revised tests into the next stage.

Keywords: Vocabulary test, Verbal Paired Associates, Digital Neuropsychology, Test My Brain.

## Sissejuhatus

### **Digitaalne neuropsühholoogia**

Digitaalne neuropsühholoogia on valdkond, mis tegeleb kognitiivsete võimete ja käitumise uurimisega digitaalsete seadmete abil. Tüüpiliselt on selleks seadmeks kas laua- või sülearvuti, tahvelarvuti või nutitelefon (Bauer et al., 2012; Germine et al., 2019). Kognitiivsed võimed pakuvad uurimisalast huvi, kuna need on seotud edu saavutamise ja paljudes eluvaldkondades. Kognitiivsed võimed on seotud nii haridusliku kui ka ametialase edenemisega, sotsiaalmajanduslike saavutustega, igapäevase toimetulekuga ja tervise ning elueaga (Fuente et al., 2019; Lövdén et al., 2020).

Traditsiooniliselt on kognitiivsete võimete mõõtmine toimunud nii-öelda pliiatsi ja paberi meetodil ning aja jooksul on välja töötatud arvukalt usaldusväärseid mõõtevahendeid. Siiski on traditsioonilisel meetodil mitmeid puudusi, mis muudavad nende kasutamise suurte uuringute ja andmekogumite puhul keeruliseks. Traditsiooniline mõõtmisviis on ressursikulukas, kuna nõuab kvalifitseeritud spetsialistide kaasamist, sobivate ruumide olemasolu ning hilisemat andmete skoorimist, millega võivad kaasuda inimlikud vead. Lisaks võib traditsiooniline mõõtmisviis olla koormav ka katseisikutele, kellest paljud võivad uuringus osalemisest loobuda, kui füüsiline kohale ilmumine on raskendatud, näiteks töögraafiku või elukoha kauguse tõttu (Backx et al., 2020; Chaytor et al., 2021; Ferrar et al., 2021).

Digitaalse tehnoloogia laialdane levik on loonud võimaluse mõõta vaimseid võimeid kaugteel, kasutades selleks isiklikku digitaalset seadet (Passell et al., 2019). Uus meetoodika pakub mitmeid eeliseid, mis aitavad ületada traditsioonilise meetodi piiranguid. Veebipõhiseid teste saavad inimesed sooritada endale sobival ajal, oma kodus ilma kõrval viibiva spetsialistita (Chaytor et al., 2020). Digitaalne testimine pakub võimaluse mõõta inimekäitumise aspekte standardiseeritult ja ülima täpsusega (Germine et al., 2019), automatiseerida andmekogumist ja salvestamist (Silverstein et al., 2007) ning vähendada administreerimise ja skoorimise kulusid. Digitaalne testimine võimaldab läbi viia uuringuid suurematel valimitel (Bauer et al., 2012; Chaytor et al., 2020). Mitmed biopangad maailmas, näiteks Soomes ja Suurbritannias, on liitnud veebi kaudu kogutud kognitiivsete testide tulemused geenidoonorite andmetega (Liu et al., 2020; Fawns-Ritchie & Deary, 2020; Rodosthenous et al., 2022).

Kuigi digitaalne neuropsühholoogia loob mitmeid uusi võimalusi tuleb arvestada sellega kaasnevate piirangutega (Germine et al., 2012). Individuaalselt läbiviidavad veebipõhised testid nõuavad sooritajatelt minimaalset tehnoloogia tundmist, eeldavad neilt digitaalse seadme olemasolu ning võimalust kasutada internetti. Üks suurimaid erinevusi võrreldes traditsioonilise kliinilise testimisega on see, et testi sooritamise keskkonda pole võimalik täielikult kontrollida (Chaytor et al., 2020).

Standardseid neuropsühholoogilisi teste on tänu nende skaleeritavusele ja standardiseeritavusele lihtne digitaalsele platvormile kohandada. See eeldus kehtib juhul kui ülesanded ei nõua joonistamist, kirjutamist või vaba meenutamist (Passell et al., 2019). Kuigi digitaalseks kasutamiseks kohandatud testide puhul võib tekkida kiusatus pidada neid samaväärseks ja võrreldavaks klassikaliste mõõtevahenditega mille järgi need loodud on, tuleb arvestada, et tegu on siiski muudetud ja seega uue testiga (Bauer et al., 2012). Digitaalne neuropsühholoogia areneb kiiresti ning vajadus uute valiidsete ja usaldusväärsete testide järgi on kasvavas trendis.

Neuropsühholoogiliste digitaalsete testide loomisega tegeleb *Test My Brain* (TMB), mis on Harvardi Meditsiinikooli mittetulunduslik programm. TMB raames on välja töötatud lai valik neuropsühholoogilisi mõõdikuid, mis on leidnud laialdast rakendust uurimistöodes. Programmi kodulehel <https://www.testmybrain.org/> on avaldatud erinevaid tasuta sooritavaid kognitiivseid teste, mida 12 aasta jooksul on täitnud üle 2,5 miljoni vabatahtliku (Chaytor et al., 2021; Singh et al., 2021). Koroonaviiruse leviku ajal loodi TMB edasiarendusena digitaalne kogumtest *TestMyBrain Digital Neuropsychological Toolkit*, millesse koondatud ülesanded sobivad vaimsete võimete hindamise abivahenditena ka kliiniliseks kasutamiseks (Singh et al., 2021).

Eestis puuduvad eestikeelsed kognitiivseid võimeid mõõtvad testid, mida oleks võimalik iseseisvalt täita nii arvutis kui ka telefonis. Siinolev töö on osa projektist „Käitumise detailsete mõõdikute eesti keelde adapteerimine”. Projekti raames adapteeriti eesti ja vene keelde üheksa kognitiivset ja isiksuslikku mõõdikut, mida saab hõlpsasti kasutada nii telefonis kui ka arvutis. Kognitiivsed mõõdikud on pärit *Test My Brain*’i keskkonnast ning need pakuvad uudseid võimalusi inimese käitumise ja isiksusejoonte mõõtmiseks. Ülesannete valikul oli eesmärgiks koostada testikogum, mis mõõdaks kognitiivseid võimeid võimalikult laiaulatuslikult, kuid samas lühikese aja jooksul. Lõplik valik tehti tuginedes USA terviseinstituudi teadlaste soovitudele (d’Ardenne et al., 2019).

### Sõnapaaride mäluülesanne

Sõnapaaride mäluülesanne, tuntud kui *Verbal Paired Associates*, on laialt kasutatav episoodilise ja verbaalse mälu hindamisvahend. See kuulub klassikalise kliinilise neuropsühholoogia kogumtesti Wechsleri mäluskaala ülesannete hulka (Clark et al., 2018; Pike et al., 2013). Kuna episoodilise mälu kahjustus on Alzheimeri tõve esimeste sümptomitest, kasutatakse testi sageli Alzheimeri tuvastamiseks selle varajases staadiumis (Lowndes et al., 2008; Pike et al., 2013).

Suurem osa standardiseeritud sõnapaaride mäluülesannetest põhinevad hilise ajendatud meenutamise (*delayed cued recall*) formaadil (Lowndes et al., 2008; Clark et al., 2018). Ülesandes tuleb meelde jätta loetud või kuulud sõnade paarid (nt raamat – lennuk). Mõne aja möödudes, näiteks poole tunni pärast, esitatakse testi sooritajale ajend - esimene sõna varem esitatud paarist ja palutakse meenutada milline sõna sellega kokku kuulus (Clark et al., 2018)

Sõnapaaride ajendatud meenutamine (*TMB Verbal Paired Associates*) on individuaalseks veebipõhiseks läbiviimiseks kohandatud sõnapaaride mäluülesanne. Test iseloomustab verbaalset ja episoodilist mälu sarnaselt Wechsleri mäluskaalale. Testi autorid on välja toonud, et see mõõdab ka deklaratiivset mälu ja töömälu. Ülesande ülesehitus muudab selle digitaalsele platvormile lihtsalt adapteeritavaks ning erinevate seadmete kasutamine ei mõjuta skooore olulisel määral. Katseisikute hinnangul on test suhteliselt meeldiv ning hästi talutav (Passell et al., 2019). Testil on kõrge sisereliaablus,  $r = 0.89$ ,  $N = 8496$  ning see on valitud *TestMyBrain Digital Neuropsychological Toolkit*'i mõõdikute hulka (Singh et al., 2021).

*TMB Verbal Paired Associates* 'i skoorid püsivad eluea raames suhteliselt stabiilsena. Noorukieas esineb sooritusvõimes tõus, mille järel püsivad skoorid üsna muutumatutena ning pärast 60 eluaastat on märgata väikest langust. Esineb kerge sooline erinevus - naiste skoorid kalduvad veidi kõrgemad kui meestel. Nagu kognitiivseid võimeid mõõtvate testi puhul eeldada võib, on tulemustele iseloomulik korrelatsioon haridusega – madalama haridustasemega isikute skoorid on madalamad ning sooritusvõime paraneb iga järgneva haridustasemega (Passell et al., 2019; Uttl et al., 2002). TMB keskkonna sõnapaaride meenutamise testi tulemuste puhul esineb nõrk korrelatsioon lühiajalist mälu mõõtvate ülesannetega ning Sõnavara testi tulemustega ( $r = 0.37$ ,  $n = 521$ ) (Passell et al., 2019).

## Sõnavara

Sõnavara testid mõõdavad üldist kognitiivset võimekust, üldist vaimset võimekust, kristalliseerunud intelligentsust ning verbaalset taibukust (Passell et al., 2019). Erinevalt enamikest kognitiivsetest võimetest, mis saavutavad tipu umbes 20 aasta vanuselt on sõnavara teadmisele iseloomulik sooritusvõime tipnemine alles hilises keskeas. Pärast seda võib võimekus langema hakata, kuid ei pruugi (Bowles et al., 2005; Hartshorne & Germine, 2015). Sõnavara testid on levinud vahendid üldise vaimse võimekuse määramiseks ja need kuuluvad muu hulgas ka *Wechsler Adult Intelligence Scale* (WAIS) alatestide hulka. Arvestades, et neurodegeneratiivsete haiguste kulgu iseloomustab sõnavara teadmise püsimine stabiilsena ning langust on oodata alles haiguse hilises faasis, kasutatakse sõnavara teste premorbiidse vaimse võimekuse määramisel dementsuse, kerge kognitiivse kahjustuse ja Alzheimeri tõve puhul (Oliveira et al., 2014; Passell et al., 2019).

TMB Sõnavara test (*TMB Vocabulary*) on kohandatud digitaalsele platvormile WAIS-IV Sõnavara alatesti baasil. Viimases palutakse katseisikul suuliselt defineerida talle ette loetud sõnad. TMB Sõnavara test aga tugineb valikvastuste formaadil, säilitades samas WAIS-IV sõnavara alatestiga mõõdetava konstrukti (Chaytor et al., 2021). Hartshorne ja Germine (2015) märgivad, et digitaalne sõnavara test ja WAIS-IV Sõnavara alatest on psühhomeetriliselt võrreldavad.

Testi sooritanud inimeste hinnangul on tegu meeldiva ja huvitava ülesandega, mida kinnitab ka üliväike katkestajate hulk. TMB Sõnavara testi kasutajakogemus on erinevates seadmetes stabiilne ning kasutatav seade ei mõjuta tulemusi märkimisväärselt. Testil on kõrge sisereliaablus,  $\alpha = 0.83$ . (Passell et al., 2019; Chaytor et al., 2020). TMB sõnavara testi tulemused peegeldavad varasemaid leide, mille kohaselt saavutavad verbaalsed võimed haripunkti keskeas – testi skoorid tõusevad kooskõlas vanusega ning vanemate vastajate puhul võib tulemustes esineda lae-efekt. Tulemused näitavad kerget soolist erinevust naiste kasuks ning on selgelt näha, et skooride tõusevad koos haridustaseme tõusuga. TMB sõnavara testi tulemused on nõrgas korrelatsioonis TMB maatriksitega ( $\rho = 0.29$ ,  $n = 1686$ ) ja TMB sõnapaaride meelespidamise testiga ( $\rho = 0.33$ ,  $N = 9046$ ).

## Töö eesmärk ja hüpoteesid

Mõõdikud võtab kasutusse Geenivaramu ning edastab need oma 220 000 geenidoonorile. Sellises ulatuses andmekogumeid, mis hõlmavad suurt hulka inimesi ja ühendavad ühtlasi kognitiivseid võimeid, isiksusejooni ning geneetikat, on maailmas väga vähe. Andmekogum saab toetada geenipõhiste tervisekäitumise ennustuste täiendamist, tervisekäitumist puudutavate sekkumiste planeerimist ja personaalmeditsiini arengut. Lisaks annab töö panuse neuropsühholoogiliste mõõdikute arendusse ja nende kasutamise laiendamisse teadustöös.

Siinne töö piloteerib sõnapaaride ajendatud meenutamise testi ja sünonüümide sõnavara testi kasutamist mõõtevahenditena. Hüpoteesid on püstitatud tuginedes varasemale teemakohasele kirjandusele eesmärgiga hinnata mõõtevahendite kooskõllalisust originaalidega – kas sisereliaablus järgib *Test My Brain* autorite seatud aktsepteeritavaid piire ning kas esinevad korrelatsioonid mida on täheldatud originaalide puhul (Passell et al., 2019). Lisaks hüpoteesidele vastamisele on eesmärgiks kindlaks teha, millised ülesanded on liiga kerged või liiga rasked, ja tuvastada sõnastiimulite seotus mõõdetavate konstruktidega. Tulemusi kasutatakse, et teha ettepanekud testide parandamiseks, eesmärgiga optimeerida nende tõhusust ja usaldusväarsust.

Hüpotees 1: Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi sisereliaablus on *Passell'i (2019)* seatud aktsepteeritavuse piirile vastavalt vähemalt  $\alpha = 0.7$ .

Hüpotees 2: Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi kooskõllalisust nende paralleeltestidega kinnitab tugev positiivne korrelatsioon.

Hüpotees 3: Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi omavaheline korrelatsioon kordab *Test My Brain*'i tulemust – Pearsoni  $r = 0.37$ .

Hüpotees 4: Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi skoorid on positiivses seoses TMB maatriksite tulemustega.

Hüpotees 5: Sünonüümide sõnavara testi tulemused paranevad vanuse kasvades.

Hüpotees 6: Hariduse mõju tulemustele kinnitab mõlema testi puhul kõrgema haridustasemega vastajate kõrgem skoor.



## Meetod

### Valim

Tegu on mugavusvalimiga millesse kuulus 345 katseisikut vanuses 19 – 69 aastat. Osalemine eeldas katseisikutelt ligipääsu internetile ning suutlikkust kasutada arvutit või nutiseadet. Katseisikud sooritasid teste kahel korral, kasutades võimalusel erinevaid seadmetüüpe. Soorituskordade vahele jäi vähemalt seitse päeva.

### Andmekogumisviis

Andmete kogumine toimus 2020 detsembrist 2021 maini. Kutseid uuringusse levitati sotsiaalmeedias reklaami teel, uurijate suhtlusringkondades, Tartu Ülikooli psühholoogiatudengite meililistides ning suuremate tööandjate kollektiivides. Töös kasutatud ning projekti puudutavad andmed on kättesaadaval lehtedel <https://osf.io/uv9sw/> ja <https://osf.io/xc65p>.

### Protseduur

Kutse suunas osalejad uuringut tutvustavale veebilehele kuhu tuli osalemiseks sisestada oma e-maili aadress. Kümne minuti jooksul saadeti osalejale e-kiri mis sisaldas linki uuringukeskkonda formr.org ning randomiseeritud juhust alustada testide sooritamist kas nutitelefonist või arvutist. Juhise järgimine oli vabatahtlik, kuid enamik osalejatest nõustus soovitusel. Lingi kaudu avanev uuringukeskkond sisaldas detailsemat uuringu ülevaadet, andmekaitse tutvustust ning informeeritud nõusoleku kinnitamist. Nõusoleku andmise järgselt täitsid katseisikud demograafilist ning sotsiaalmajanduslikku tausta uuriva eelküsimustiku. Kogumtesti täitsid katseisikud lehel [cognit.psych.ut.ee](http://cognit.psych.ut.ee). Loosiga määrati osalejatele üks kuuest võimalikust esitamisjärjestusest mis jäi mõlemal sessioonil samaks. Testide puhul millel oli paralleelversioon (Sünonüümide sõnavara, Sõnapaaride ajendatud meenutamine ning Maatriksid) randomiseeriti a või b versioonist alustamise järjekord. Vahetult enne iga ülesande sooritamist tuli tutvuda juhendiga ning lahendada edukalt paar harjutusülesannet. Testide täitmist oli võimalik ülesannete vahele jääval ajal katkestada ning hiljem jätkata. Teine testimiskord algas eelmise korra sõnapaaride ajendatud meenutamise testiga Kogumtesti läbiviimiseks kulus umbes 40 minutit. Mõlema soorituskorra järgselt paluti katseisikutel formr.org lehel lisaks täita tagasiside küsimustik mis uuris muu hulgas iga testi

meeldivuse kohta. Seitsme päeva pärast saadeti katseisikule e-maili teel meeldetuletus ning palve täita teste seekord teise seadmega. Osalejad said tänutäheks soovi korral e-maili teel tagasisidet millistes testides nad kõige tugevamad olid. Küsimuste tekkimise korral oli võimalik uurijatega e-maili teel ühendust võtta.

Projekti käigus kogutud andmete baasil on kirjutanud uurimistööd: Asu (2022) teemal „Digitaalsete maatriksite MaRs-IB ja Testmybrain pilootuuring”, Adams (2022) „Elusündmuste seosed Testmybrain kognitiivsete testidega“ ja Berezin (2022) „Digitaalse punktide ühendamise testi, lihtreaktsioonaja testi ning valikreaktsioonaja testi valideerimine“.

### **Mõõtmisvahendid**

Projekti käigus adapteeriti eesti keelde TMB keskkonnast pärit vaimseid võimeid mõõtev kogumtest mis koosneb üheksast mõõdikust: Numbri-sümboli vastavuse test, Sõnapaaride ajendatud meenutamine, Tasu edasilükkamine, Maatriksid, Sünonüümide sõnavara, Punktide ühendamise test A ja B, Lihtreaktsiooniaeg ning Valikreaktsiooniaeg. Siinolev töö kirjeldab Sünonüümide sõnavara ning sõnapaaride ajendatud meenutamise testide adapteerimist.

### **Adapteerimisprotsess**

Adapteerimise üheks eesmärgiks oli säilitada originaaltestide raskusaste ja eristusvõime. Stiimulite raskusastme võrdsustamise aluseks võeti Malhotra, Krosnick'i ja Haertel'i (2007) töö eeskujul sõnade esinemissagedused. Sõna sageduse efekti kohaselt sõltub inimeste suutlikkus mõnda väljendit teada, nimetada ja meenutada suurel määral sellest kui sageli on antud väljendit võimalik keeles kohta. Sageli esinevaid sõnu teab suurem hulk inimesi ja nende kognitiivne töötlus toimub kiiremini. Umbes 30% kuni 40% varieeruvusest sõnade tundmise ülesannetes on seletatav sõna sagedusega. See on üks olulisemaid tegureid ka mälusoorituse puhul. Ülesannetes kus tuleb sõnu meelde jätta ja hiljem meenutada või teiste hulgast ära tunda on harva esinevaid sõnu keerulisem reprodutseerida kuid lihtsam ära tunda (Badham et al., 2017; Brysbaert et al., 2018).

Sõnade esinemissagedust keeleruumis on võimalik määrata leksikaalsete andmebaaside abil (Hessler et al., 2016; Brysbaert et al., 2018) Siinses töös toimus sõnade esinemissageduse tuvastamine veebipõhise korpuspäringusüsteemi Sketch Engine abil<sup>1</sup>. TMB sõnavara ingliskeelsete

---

1 Saadaval lehel <https://www.sketchengine.eu/>

stiimulite puhul võeti aluseks English Web Corpus 2015 enTenTen ning eestikeelsete stiimulite puhul Estonian National Corpus 2019. Päringud esitati sõnade algvormide ehk lemmade baasil.

### Sünonüümide sõnavara test

*Test My Brain Vocabulary* mõõdab üldist kognitiivset võimekust, üldist vaimset võimekust, kristalliseerunud intelligentsust ning verbaalset võimekust (Passell et al., 2019). Ekraanile kuvatakse stiimulsõna ja palutakse viie all oleva vastusevariandi hulgast valida sõna mis on sellele tähenduselt kõige lähedasem. Test koosneb kolmekümnest küsimusest ja skoori annab õigete vastuste arv. Ülesande täitmise aeg ei ole piiratud ning keskmiselt kulub selle läbimiseks umbes 4.5 minutit (Passell et al., 2019).

Testi a-versiooni stiimulite loomise esimese sammuna tõlgiti ingliskeelsed ülesanded *Google Translate*'i abil eesti keelde. Korpusest küsiti originaaltesti stiimulite ning selle tõlke läbi saadud sõnade esinemissagedused. Kahe keeleruumi esinemissageduste ühtselt võrreldavaks muutmiseks viidi leitud sagedused Younesi ja Reipsi (2019) töö eeskujul logaritmilisele skaalale. Referentsiks võeti mõlemas keeles esineva kaheksa tüüpsõna (näiteks „*such*” – „*nagu*”; „*not*” – „*ei*”) sageduste keskmine, mida kasutati teiste sõnade sageduse jagamiseks. Seejärel korrutati tulemus protsentuaalse väärtuse saamiseks sajaga.

Kui tõlgitud stiimuli raskusaste lahkes eesmärgipärasest rohkem kui kahe ühiku võrra asendati sõna sobivama sünonüümi või uue, raskusastme ja konteksti asetumise poolest sobiva sõnaga. Sünonüümide sõnavara testi võeti originaalist üle neli stiimulit ja viis sünonüümi. Piloteerimiseks pikendati originaalis kolmekümnest küsimusest koosnevat testi viie küsimuse võrra ning juurde otsiti kokku 26 uut stiimulit. Eesmärgiks seati, et ülesannete hulgas oleks vähemalt kümme rasket ehk harva esinevat sõna. Testide stiimulite tavasõnade suhtes normeeritud ja logaritmitud sageduse skoorid ulatusid vahemikust -8 kuni 2 ingliskeelsete sõnade puhul (mean = -1.98, SD = 2.08) ja -7.89 kuni 3.21 eestikeelse A-versiooni ja b-versiooni stiimulite puhul (mean = -1.64, SD = 2.04). Eestikeelseks tavasõnade normeerimise sageduseks oli 13220504 ja inglise keele jaoks 303741584 (Younes & Reips, 2019).

Uute ülesannete otsimisel jälgiti WAIS-III eesti keelde adapteerimise põhimõtteid, et testis oleks samas proportsioonis nimi-, omadus- ja tegusõnu ning et ülesannete hulgas oleks mitmesuguse raskusega ülesandeid. Sobivuse sisulisel hindamisel jälgiti, et ülesanne ei oleks vanuselisel, sooliselt või valdkonniti erapoolik. Juhul kui küsimus nõudis fakti tundmist pidi sellega olema

võimalik kokku puutuda ka väljaspool kitsast valdkonda (Burk 2007). Distraktorite valikul jälgiti, et need sobiksid sõnastiimuliga, ning et nende hulgas oleks nii harva kui sageli esinevaid sõnu.

Kuna sõnavara testi korduval sooritamisel ilmneb õppimise efekt (Passell et al., 2019) loodi Sünonüümide sõnavara testile paralleelversioon. Paralleelversioon koostati 37 küsimusest mille esinemissagedus järgis TMB Sõnavara eeskuju ning testi koostamisel järgiti samu põhimõtteid mida a-versiooni puhul. Sünonüümide sõnavara testide stiimulid on leitavad lehtedelt:

[https://kodu.ut.ee/~a40950/tmb\\_ee/Vocab\\_2020/testInput\\_2021.txt](https://kodu.ut.ee/~a40950/tmb_ee/Vocab_2020/testInput_2021.txt) ja

[https://kodu.ut.ee/~a40950/tmb\\_ee/Vocab\\_2020/testInput\\_2021\\_v2.txt](https://kodu.ut.ee/~a40950/tmb_ee/Vocab_2020/testInput_2021_v2.txt)

### **Sõnapaaride ajendatud meenutamise test**

Sõnapaaride ajendatud meenutamine (TMB *Verbal Paired Associates*) mõõdab töömälu, verbaalset mälu ja episoodilist mälu (Passell et al., 2019). Test koosneb kahest osast. Esimeses, õppimise faasis, tuleb meelde jätta 25 sõnapaari. Osalejaid informeeritakse eelnevalt, et nende meelepidamist testitakse hiljem. Õppimise faasis kuvatakse sõnapaarid kahesekundiliste intervallidega ükshaaval kolmeks sekundiks ekraanile. Umbes kahe minuti pärast toimub meenutamise faas, milles esitatakse varem kuvatud stiimulipaaridest ajend, ehk esimene sõna, ning palutakse nelja vastusevariandi hulgast ära tunda selle juurde kuulunud paariline (*target*).

Vastamiseks kuluvat aega ei piirata. Testi sooritamiseks kulub kokku umbes viis minutit, mis ei hõlma vahepealset pausi. Testi skoor baseerub õigete vastuste arvul (Germine et al., 2012; Singh et al., 2021a). Uuringus sisustas meeldejätmise ja meenutamise vahele jäävat vaheaega numbrisümboli vastavuse test.

TMB Sõnapaaride testil on kaks erineva raskusastmega versiooni. Raskem versioon koosneb abstraktsetest sõnadest (näiteks nagu „uudishimu” ja „vabadus”), ehk kontseptsioonidest, mida on keeruline tajuda. Kergem versioon sisaldab konkreetseid nimisõnu (näiteks „õun” ja „jalgratas”), mis kirjeldavad käega katsutavaid ja selgelt tajutavaid objekte, millest on võimalik endale meenutamist hõlbustav mentaalne kujutluspilt luua (Passell et al., 2019; Yan et al., 2023). Kuna raskema testi skoorid on madalamate tulemuste poole kaldu valiti kogumtesti kergem versioon.

*TMB Verbal Paired Associates*’i stiimulid tõlgiti Google Translate’i abil eesti keelde. Stiimulite valikul lähtuti põhimõtetest, et tegu peab olema ühetähenduslike ning laialt levinud nimisõnadega, mis ei moodusta omavahel otseseid semantilisi seoseid (nagu näiteks „taevas” ja „pilv”). Tõlkimise järgselt asendati kultuuriliselt vähem sobivad ning eesti keeleruumis harvem esinevad sõnad

sagedamini esinevate alternatiividega. Sõnade „morsk”, „ookean” ja „dollar” asemel kaasati testi levinumate väljendid „hüljes”, „meri” ja „euro”. Sõna „tulp” omas kahte tähendust (lill ja post) ning vahetati sarnase kuid ühetähendusliku sõna „roos” vastu. Lisaks jälgiti, et ajend oleks unikaalne ning ei esineks vastusevariantide hulgas. Kuid iga sihtmärk esineb ühel korral ka distraktorina valikuvariantide seas. Pooled distraktoritest on unikaalsed ning pooled neist esinevad kahel korral (Passell et al., 2019).

Kuna TMB VPA mitmekordse sooritamiseega kaasneb õppimise efekt (Lo et al., 2012; Passell et al., 2019) loodi testile paralleelversioon. Paralleelversiooni koostamisel järgiti samu põhimõtteid mida a-versiooni luues. Lisaks jälgiti, et ajendite, sihtmärkide ja distraktorite raskusastmed jaotuksid mõlema testi puhul ühtlaselt. Testide stiimulitega saab tutvuda lehtedel:

[https://kodu.ut.ee/~a40950/tmb\\_ee/VerbalPA\\_2019/testInput\\_2021.txt](https://kodu.ut.ee/~a40950/tmb_ee/VerbalPA_2019/testInput_2021.txt) ja

[https://kodu.ut.ee/~a40950/tmb\\_ee/VerbalPA\\_2019/testInput\\_2021\\_v2.txt](https://kodu.ut.ee/~a40950/tmb_ee/VerbalPA_2019/testInput_2021_v2.txt)

## **Uuringu eetiline külg**

Uuringus osalesid vabatahtlikult täisealised inimesed, kes olid andnud oma informeeritud nõusoleku. Kuigi osalemist ei tasustatud oli soovi korral võimalik saada tagasisidet oma tulemuste kohta. Uuringus osalemisega ei kaasnenud katseisikute jaoks teadaolevaid riske. Andmeid koguti GDPR nõuetele vastavatesse serveritesse millele omasid ligipääsu vaid uuringu läbiviijad ning eksperimentaalpsühholoogia labori töötajad. Osalejatelt ei küsitud tundlikke isikuandmeid ning andmeid säilitati ja töödeldi anonümiseeritult, mistõttu Tartu Ülikooli eetikakomitee juht ei pidanud vajalikuks inimuuringute eetikakomitee loa taotlemist.

## **Autori panus**

Töö autori ülesanneteks oli Sünonüümide sõnavara testi ja Sõnapaaride ajendatud meenutamise testi paralleelversioonide ülesannete koostamine. Autor tutvus teaduskirjandusega, püstitas hüpoteesid ja uurimisküsimused, teostas andmeanalüüsi, analüüsis tulemusi ning kirjutas töö.

## **Statistiline analüüs**

Andmeanalüüs viidi läbi statistikaprogrammis R, kasutades R Studio klienti. Andmete puhastamine toimus enne siinse töö autori kaasamist. Andmekogumist eemaldati isikud kes täitsid ühte

testiversiooni mitmel korral või katkestasid testi täitmise poole pealt. Sünonüümide sõnavara testi puhul eemaldati isikud kes valisid 90% ühte vastusevarianti või kelle testi sooritamise aeg oli lühem kui 60 sekundit. Mälutesti andmetest eemaldati isikud kelle reaktsiooniajad jäid all 1000 ms, kes jätsid 20% küsimustele vastamata või valisid 90% ühte vastusevarianti (Joormann et al., 2022).

Siinoleva töö ülesandeks on uurida testide usaldusväarsust mõõtevahenditena ning võrrelda leitud näitajate kooskõlalisust *Test My Brain* keskkonna testidega. Testi sisereliaablus annab infot kui hästi küsimustikus sisalduvad ülesanded mõõdavad sama konstrukti. Sisereliaablust mõõdetakse seemise konsistentsuse alusel, mida saab hinnata Cronbach'i  $\alpha$  abil. Testi reliaablust võib pidada heaks kui  $\alpha$  väärtus on  $\geq 0.7$  (Cronbach, 1951). Töö eesmärgiks oli saavutada Sünonüümide sõnavara testi ja Sõnapaaride ajendatud meenutamise testide puhul Passell et al. (2019) määratud aktsepteeritav reliaablus  $\geq 0.7$ . Esimese ja teise hüpoteesiga uuriti kahe testi sisereliaablust R Studio *psych* paketti kuuluva *cronbach.alpha()* funktsiooni abil. Valimi iseloomustamiseks kasutati R-funktsioone *summary()*, *summarise()*, *mean()*, ja *sd()*.

Järgnevate hüpoteesidega (hüpoteesid 2-4) uuriti testide omavahelisi korrelatsioone. Positiivne korrelatsioon võimekust mõõtvate testide vahel annab infot, et testid mõõdavad sarnast oskust või võimet (Salthouse, 2014). R-funktsiooni *cor()* abil leiti korrelatsioonikordajad testide ja nende paralleelversioonide vahel ning sõnavara testi ja sõnapaaride mäluülesande vahel. Sõnavara testi ja sõnapaaride omavaheline korrelatsioonianalüüsi läbiviimiseks kasutati testide a- ja b-versioonide keskmisi skoorid. Samuti võeti keskmised skoorid aluseks Maatriksite a-versiooniga seoste uurimisel. Pearson'i korrelatsioonikordaja läbiviimise eelduseks oleva andmete normaaljaotuslikkuse visuaalseks hindamiseks loodi testide skooridest histogrammid. Lisaks teostati funktsiooni *shapiro.test()* abil Shapiro-Wilk testid. Kuna andmed ei vastanud normaaljaotusele leiti Pearson'i koefitsientidele täiendavalt ka Spearman'i korrelatsioonikordajad.

Viienda hüpoteesi eesmärk oli tuvastada, kas saadud tulemused kordavad varasemaid uuringuid, mille kohaselt erinevalt teistest kognitiivsetest võimetest tõuseb verbaalne võimekuse peaaegu kogu eluea vältel. Selleks loodi *geom\_smooth()* funktsiooni abil lineaarsel regressioonil põhinev regressioonijoon, mis illustreerib vanuse ja hariduse vahelist seost koos usalduspiiridega.

Võrdluseks teostati sama analüüs sõnapaaride mäluülesande andmetega. Sõnavara testi a- ja b-versiooni skooride ning vanuse seoseid uuriti täiendavalt R-paketi *ppcor* funktsiooni *pcor()* abil, mis võimaldab Spearman'i meetodile tuginedes uurida osaliskorrelatsioone, mis võtavad arvesse ka kolmanda muutuja mõju.

Kuuenda hüpoteesiga uuriti hariduse mõju skooridele. Analüüsisist jäeti välja alla 25 aastased isikud, kuna nende hariduslik potentsiaal ei pruugi olla realiseerunud ning seega ei saa kindel olla, et kinnitatud haridustase vastab võimetele. Seejärel arvutati kahe testiskoori keskmine tulemus ning loodi tulpdiaagramm mis näitab keskmist skoori erinevate haridusastmete gruppides koos usalduspiiridega.

Lisaks hüpoteesidele vastuste leidmisele oli eesmärgiks tuvastada ülesannete kogumist küsimused, mis mõõdavad uurimisalast võimekust kõige paremini, või osutuvad liiga keerukaks või triviaalseks. Selleks viidi läbi faktoranalüüs ja koostati küsimuste raskusastet illustreerivad graafikud. Uurimisküsimustele vastuste saamiseks uuriti testiküsimuste faktorlaadungeid. Faktoraadung on koefitsent, mis iseloomustab iga küsimuse seost üldise mõõdetava konstruktiga. Faktorlaadungid varieeruvad vahemikus -1 kuni 1, kus 0 tähendab seose puudumist ja 1 tähendab täielikku positiivset seost. Kõigepealt viidi läbi mitmevälja (*polychoric*) faktoranalüüsid ühe- ning seejärel kahefaktorilise mudeli eeldusega. (Floyd & Widaman, 1995). Faktoranalüüsis seati käsitsi kriteeriumid millega filtreeriti välja stiimulid mille faktorlaadung on väiksem kui 0.30 ning ülesanded millele vastati õigesti rohkem kui 90% kordadest.

*Item Response Theory* (IRT) on kasutusel latentsete omaduste (näiteks vaimne võimekus) mõõtmiseks testi küsimuste abil. Teooria kohaselt on iga testi küsimuse vastus mõjutatud kahest tegurist: inimese võimekusest ning küsimuse raskusest. *Item Characteristic curve* (ICC) kujutab graafiliselt kuidas küsimuse raskus ja eristusvõime mõjutavad vastuseid. ICC-graafiku x-teljel kuvatakse inimese võimekus ning y-teljel tõenäosus vastata antud küsimusele õigesti. Näiteks järsk ICC joon näitab, et ülesanne on hea eristaja erineva võimekusega inimeste vahel. ICC aitab tuvastada testi nõrgad ja tugevad kohad (Hori et al., 2022). ICC graafikute loomiseks koostati Masuri (2022) loodud koodi baasil kolmeparametriline logistiline mudel, mis arvestab iga küsimuse raskusastet, eristusvõimet ja arvamise tõenäosust.

## Tulemused

### Ülevaade andmestikust

Pärast andmete puhastamist oli kogumtesti täitnud 345 katseisikut. Katseisikute hulgas oli 311 inimest kes täitsid mõlemad sõnavara testid ja mõlemad mäluülesanded. Katseisikute vanused jäid vahemikku 19 – 69 ( $M = 37.1$ , mediaan = 36,  $SD = 12.4$ ). Valim koosnes suuremas osas naistest keda oli 268, mehi oli 40. Kolm isikut ei avaldanud oma sugu. Kogumtesti täideti kahel korral, täitmiskordade vahele jäi keskmiselt 11.6 päeva ( $SD = 7$ ). Kõige rohkem oli valimis keskharidusega (143) ja kõrgharidusega (125) inimesi. Rakendusliku kõrgharidusega katseisikuid oli 45, teaduskraadiga 6 ning põhiharidusega 3 osalejat. Katseisikute keskmine testiskoor sõnavara testis oli 28.4 ( $SD = 5.21$ ) võimalikust 36st ning mäluülesandes 21 ( $SD = 3.8$ ) 25st võimalikust.

### Hüpoteesid

Esimene hüpotees uuris testide sisereliaablust ning väitis, et Passell et al. (2019) eeskujul on Sünonüümide sõnavara testis ja Sõnapaaride ajendatud meenutamise testis saavutatav Crombach'i  $\alpha \geq 0.7$ . Analüüsi tulemused näitasid mõlema testi kõrget sisereliaablust. Sünonüümide sõnavara testi Cronbach'i  $\alpha = .85$  ning selle paralleelversiooni Cronbach'i  $\alpha = .86$ . Sõnapaaride ajendatud meenutamise testi  $\alpha = .88$  ning paralleelversiooni  $\alpha = .87$  ( $N = 343$ ).

Järgnevate hüpoteesidega (hüpotees 2 – 4) uuriti testide omavaheliste seoste tugevusi. Enne analüüside teostamist eemaldati andmestikust isikud kes olid sooritanud ainult ühe testiversiooni. Korrelatsioonide leidmisele eelnevalt teostati andmete normaaljaotuse kontroll. Sünonüümide sõnavara testi a- ja b-versiooni, Maatriksite ning Sõnapaaride ajendatud meenutamise testi a- ja b-versiooni skooride histogrammide visuaalsel vaatlusel selgus, et andmed ei vastanud normaaljaotusele. Kõikidele tulemustele oli omane skooride kalle paremas suunas ning tulemused tipnesid jaotuse paremal poolel. Shapiro-Wilk'i testi tulemused näitasid, et testide skoorid ei järginud normaaljaotust ning kõikide testide skooride  $p < 0.001$  jäi alla tähendusliku läve.

Läbiviidud analüüs viitab Sünonüümide sõnavara testi a- ja b-versiooni vahelisele tugevale positiivsele seosele. Pearson'i  $r = .71$  ja Spearman'i  $\rho = .66$  ( $N = 325$ ).

Sõnapaaride ajendatud meenutamise testi ja selle paralleeltesti vahel leiti mõõdukalt tugev positiivne korrelatsioon. Pearson'i korrelatsioonikordaja  $r = .58$  ja Spearman'i  $\rho = 0.53$  ( $N = 317$ ).

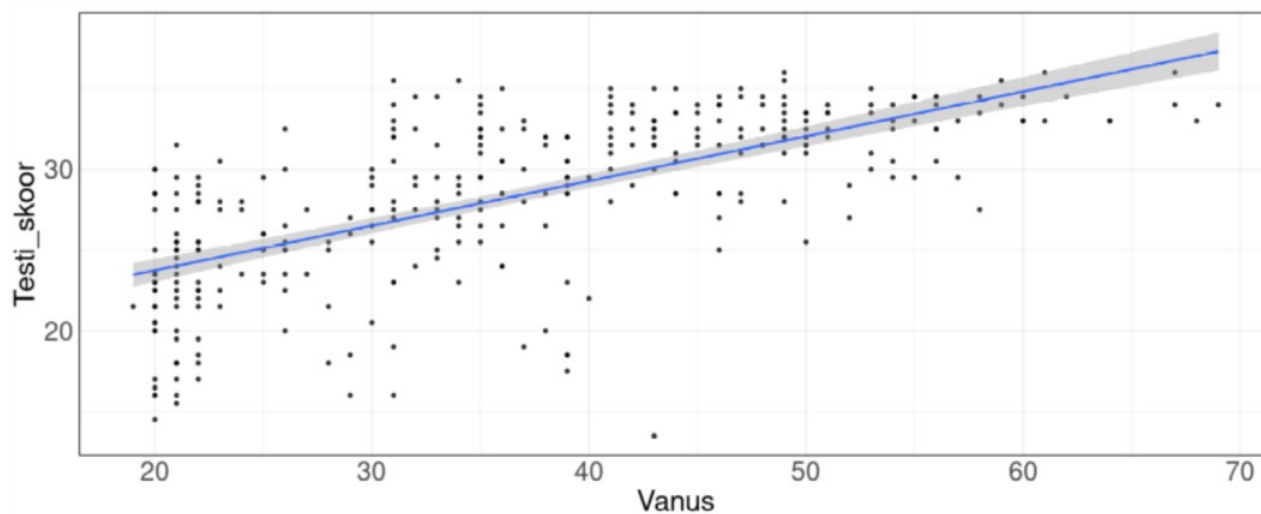


Sünonüümide sõnavara testi ja Sõnapaaride ajendatud meenutamise testi keskmiste tulemuste vahel leiti nõrk positiivne seos. Pearson'i  $r = .10$  ning Spearman'i  $\rho = .13$  ( $N = 345$ ). Tulemus on madalam kui eeskujuks olnud *Test My Brain*'i keskkonna samadel mõõtevahenditel mille  $r = 0.37$ .

Sünonüümide sõnavara keskmise skoori ja Maatriksi skooride vahel leiti Pearson'i  $r = .15$  ning Spearman'i  $\rho = .12$  ( $N=317$ ) mis viitavad nõrgale kuid positiivsele korrelatsioonile kahe testi vahel. Samuti nõrk, kuid veidi tugevam seos esines sõnapaaride testi ja Maatriksite skooride vahel:  $r = .26$  ja  $\rho = .23$  ( $N=317$ ).

Viienda hüpoteesi kohaselt paranevad sõnavara testi tulemused kooskõlas vanuse tõusuga.

Tõusev joon loodud graafikul (joonis 1) näitab positiivset seost vanuse ja kahe sõnavara testi keskmiste skooride vahel. Joonisel on näha vanemate katseisikute kõrgemad skoorid. Usalduspiirid on kitsad, muutudes kergelt laiemaks kõrgemate vanuste kohal, mis on seletatav väikse katseisikute arvuga antud vanusgrupis.

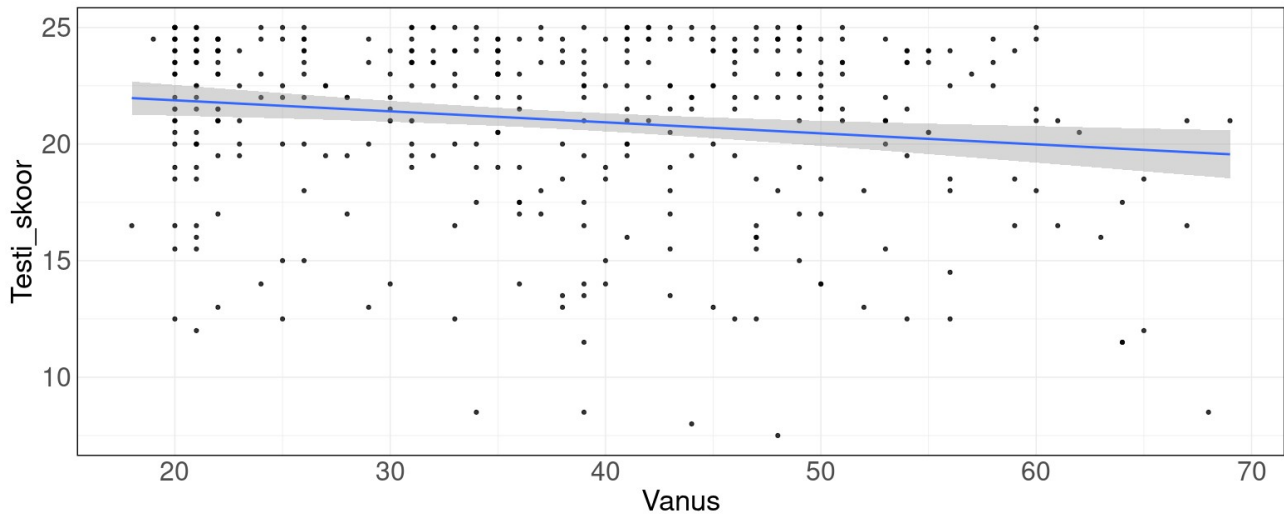


Joonis 1. Sünonüümide sõnavara skooride ja vanuse seos.

Täiendavalt läbiviidud osalise korrelatsiooni analüüs kinnitas statistiliselt olulist korrelatsiooni sõnavara testi ja vanuse ( $\rho = 0.39$ ,  $p < 0.001$ ) ning sõnavara testi paralleelversiooni ja vanuse vahel ( $\rho = 0.37$ ,  $p < 0.001$ ). Siinkohal leidis osalise korrelatsiooni analüüsiga taas kinnitust ka sõnavaratesti mõõdukas positiivne korrelatsioon oma paralleelversiooniga ( $\rho = 0.42$ ).

Võrdluseks uuriti ka Sõnapaaride ajendatud meenutamise testi skooride ja vanuse seost (Joonis 2). Joonisel on kooskõlas vanuse tõusuga laugelt langev joon suhteliselt kitsaste usalduspiiridega.

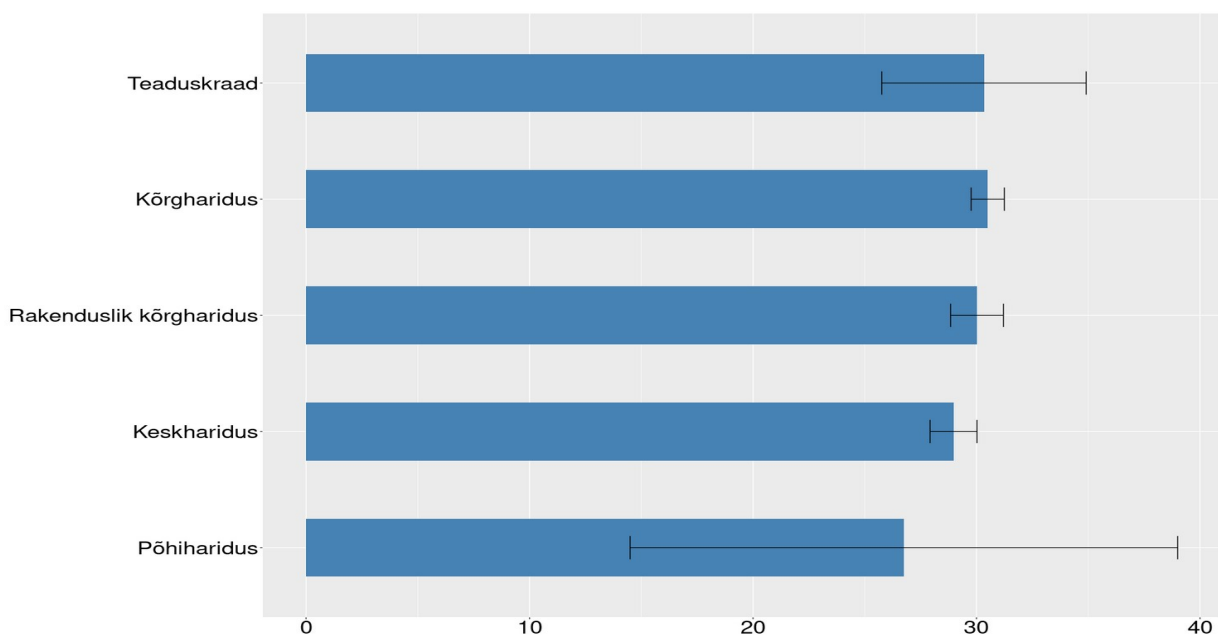
Usalduspiirid on laiemad kõige nooremas ning vanemas vanusgrupis. Vanemate isikute puhul on vähem katseisikuid ning esineb ka hajuvus vastuste osas.



Joonis 2. Sõnapaaride ajendatud meenutamise skooride ja vanuse seos

Hüpootees kuus kohaselt kinnitab hariduse mõju testi tulemustele kõrgema haridustasemega vastajate kõrgem skoor. Hariduse mõju uurimiseks eemaldati andmetest alla 25 aastased isikud.

Sünonüümide sõnavara testi andmetest eemaldati 76 inimese andmed (joonis 3). Joonis 3 illustreerib testide keskmisi tulemusi tulpdiagrammil haridusastmete kaupa. Põhiharidusega

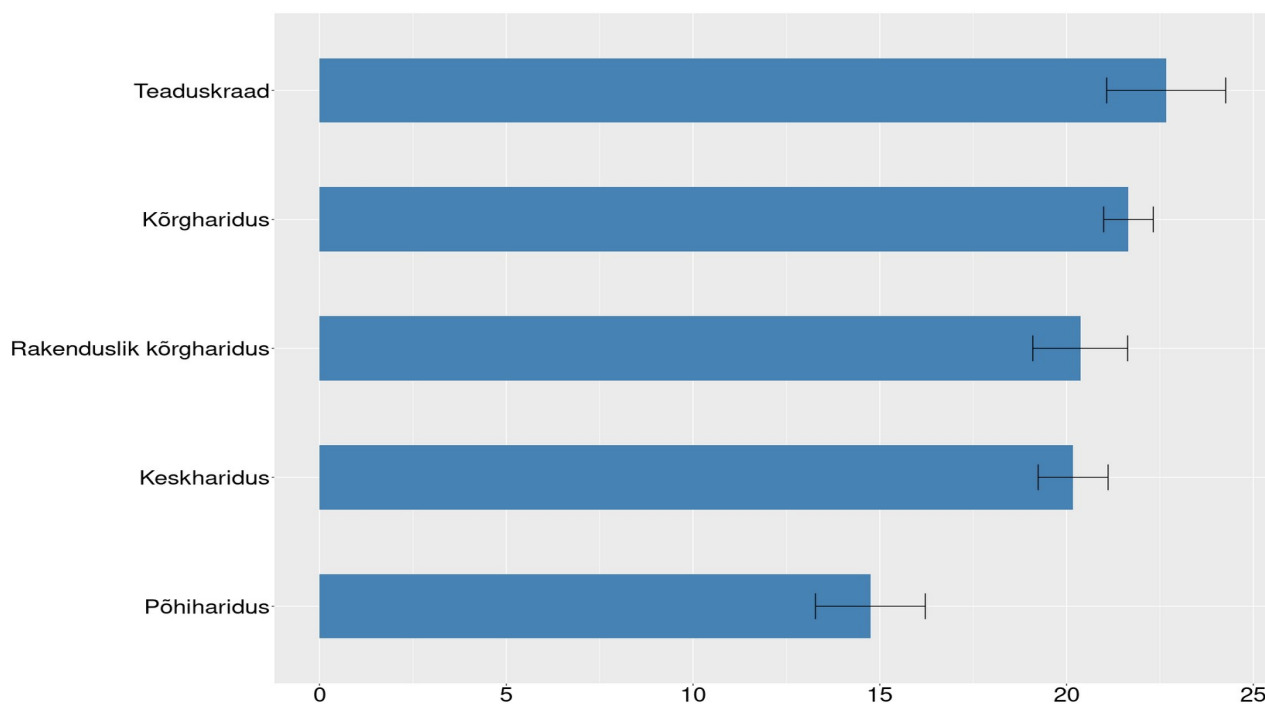


Joonis 3. Sünonüümide sõnavara testi skoorid haridustasemete kaupa

vastajate keskmine skoor oli 26,8 (SD = 8,84, n = 2), keskharidusega vastajatel 29,0 (SD = 4,51, n = 71), rakendusliku kõrgharidusega vastajatel 30,0 (SD = 4,04, n = 45), kõrgharidusega vastajatel 30,5 (SD = 4,19, n = 122) ja teaduskraadiga vastajatel 30,3 (SD = 5,72, n = 6).

Tulpdiagrammilt on näha, et erineva haridusega inimeste keskmised tulemused erinesid vähesel määral. Põhiharidusega ja teaduskraadiga isikute keskmiste skooride vahe oli 3.5 punkti. Keskmised skoorid küll suurenesid koos haridustaseme tõusuga nagu oodatud, kuid kõrgharidusega inimeste tulemused oli teaduskraadiga isikute tulemustest 0.2 punkti võrra kõrgemad. Tuleb arvesse võtta, et teaduskraadiga vastanute arv oli väga väike ning seega ei saa tulemustes väga kindel olla. Seda illustreerivad ka väiksest katseisikute arvust tulenevad laiad usalduspiirid diagrammi ülemises ja alumises osas.

Sõnapaaride ajendatud meenutamise skooride ja hariduse analüüsimiseks eemaldati andmetest 73 alla 25 aastast katseisikut. Joonisel 4 on näha sõnavara testiga võrreldes mäluülesande tulemuste suuremat varieeruvust. Kõige madalama ja kõige kõrgema haridustasemega isikute keskmiste skooride vahe oli 7.9 punkti. Põhiharidusega katseisikute keskmine skoor oli 14.8 (SD = 1.06, N = 2), keskharidusega isikutel 20.2 (SD = 4.03, N = 71), rakendusliku kõrgharidusega isikutel 20.4 (SD = 4.35, N = 45), kõrgharidusega isikutel 21.7 (SD = 3.66, N = 117) ja teaduskraadiga katseisikutel



Joonis 4. Sõnapaaride ajendatud meenutamise testi tulemused haridusastmete kaupa.

22.7 (SD = 1.99, N = 6). Jooniselt on selgelt näha skooride paranemine kooskõlas haridustaseme tõusuga. Usalduspiiride suhteliselt kitsas vahemik väheste vastajatega gruppides on seletatavad saadud skooride sarnasusega.

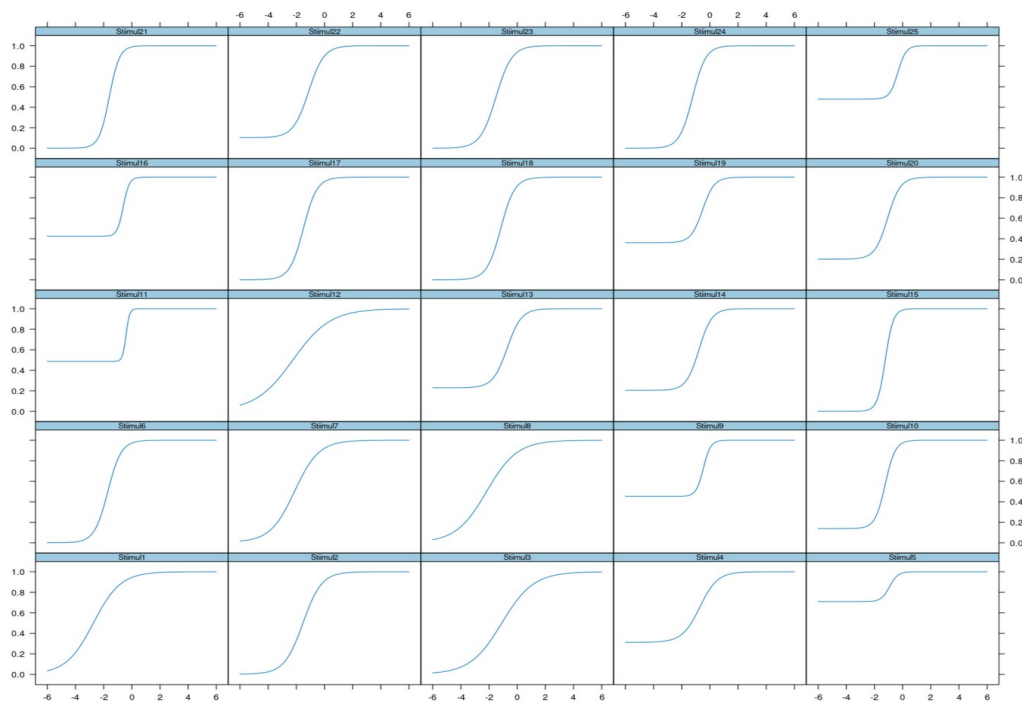
Lisaks hüpoteesidele vastuste leidmisele oli töö eesmärgiks uurida, kui tugevalt on testide stiimulid seotud mõõdetava võimega ja lisaks tuvastada ülesannete hulgast küsimused mis on liiga keerulised või triviaalsed. Mitmevälja (*polychoric*) faktoranalüüsi kasutades selekteeriti välja testiülesanded mille faktorlaadung oli suurem kui 0.30 ning millele õigesti vastamise protsent jäi alla 90. Analüüs viidi läbi nii ühe kui ka kahe faktori eelduse alusel. Võrreldes stiimuleid, mis läbisid seatud kriteeriumitega seleksiooni, oli tuvastatav, et mõlema meetodi puhul laadusid madalad (alla 0.30) faktorlaadungid samadele ülesannetele. Lisaks jäid seatud piirist ülespoole samad stiimulid mõlema meetodi puhul. Tulemused näitasid, et nii Sünonüümide sõnavara testiversioonide kui ka Sõnapaaride ajendatud meenutamise testiversioonide puhul seletab üks faktor testi stiimulite vahelisi seoseid.

Täiendavalt koostati kõikidele testistiimulitele Item Characteristic Curve (ICC) graafikud (joonis x), mis iseloomustavad küsimuse raskust ja eristusvõimet. Graafikud millel on madal alguspunkt ning järsk tõusunurk, näitavad head eristusvõimet. Y-telje kõrgemalt punktilt algav joon viitab, et vastajad on võimekusest sõltumata vastanud küsimusele õigesti. Laugelt kulgev kõver annab infot, et ülesanne on võrdse raskusega kõikide vastajate jaoks sõltumata nende võimekusest.

ICC graafikud kinnitasid faktoranalüüsi tulemusi. Näiteks suurema faktorlaadungiga stiimulite graafikud kaldusid olema järkude tõusunurkadega ning lauged või kõrge algusjoonega graafikutega küsimuste faktorlaadungid jäid alla 0.30. Mõne stiimuli puhul esines lahknevusi faktoranalüüsi ja ICC tulemuste osas. Graafiku järgi hea eristusvõimega küsimused ei läbinud faktoranalüüsi seleksiooni. Sõnavara testis olid nendeks sõnadeks skalpeerima, süüvima, uhmer ja halisema. See on seletatav suure õigesti vastajate osakaaluga (üle 90%) ning testis vähem punkte saanud isikute antud stiimulile valesti vastamisega. Üldiselt töötasid mõlemad meetodid kooskõlas ja andsid sarnaseid tulemusi.

Sõnavara testi a-versioonis eristus faktoranalüüsi seleksiooniga 19 stiimulit 35st, mille faktorlaadungid ületasid soovitud piiri ja jäid vahemikku 0.83 – 0.34. ICC graafikutelt eristusid 8 ülesannet, mille lauge kõrge joon viitas kehvale eristusvõimele. Nimetatud stiimulitele laadunud faktorid jäid alla määratud 0.30 piiri. ICC ja faktoranalüüsi tulemused erinesid viie sõnastiimuli puhul, mille visuaalsel vaatlusel võis eeldada head eristusvõimet kuid mis ei läbinud faktoranalüüsi seleksiooni.

Sõnapaaride ajendatud meenutamise testi a-versiooni 25 stiimulist eristus 21, mis on tugevamalt seotud ühe mõõdetava tunnusega, laadungitega vahemikus 0.87 – 0.36. Võrreldes sõnavara testi graafikutega ei esine mäluülesannete puhul kõrged ja laused jooni. Esmase vaatluse põhjal võiks kaaluda ühe küsimuse kogumist eemaldamist, mida kinnitas ka faktoranalüüs.



Joonis 5. Sõnapaaride ajendatud meenutamise stiimulitele loodud ICC

Sõnapaaride testi b-versioonis faktorlaadungid olid teistega võrreldes kõrgemad, 18 ülesannet läbisid analüüsi laadungitega vahemikus 0.85 – 0.55. ICC graafikud (joonis 5) ei näidanud silmatorkavalt kõrged ja laused jooni. Viis küsimust millel on kõrge alguspunktiga y-graafikul ei läbinud ka faktoranalüüsi.

## Diskussioon

Magistritöö eesmärgiks oli valideerida pilootuuringuga sünonüümide sõnavara ja sõnapaaride ajendatud meenutamise testid. Töö keskendus mõõtevahendite usaldusväarsuse analüüsile ning nende kooskõlalisuse võrdlusele originaalversioonidega. Lisaks uuriti, kui tugevalt on testiküsimused seotud mõõdetava võimega ning millised ülesanded võiksid olla liiga keerulised või triviaalsed, et need tulevikus kogumist välja jätta.

Uuring saavutas *Test My Brain* autorite poolt määratud aktsepteeritava reliaabluse piiri - Cronbach'i  $\alpha \geq 0.7$ . Analüüsi tulemused näitasid, et kõikide testide sisereliaablus on kõrge, ületades eelnevalt määratletud künnist. Sünonüümide sõnavara testide Cronbachi  $\alpha$  väärtused olid vastavalt .85 ja .86 ning Sõnapaaride ajendatud meenutamise testide  $\alpha = .88$  ja .87. Saadud tulemused on veidi kõrgemad *Test My Brain*'i saadud tulemustest, mis olid vastavalt  $\alpha = .83$  sõnavara testi ja  $\alpha = .87$  sõnapaaride mäluülesande puhul. Kõrge sisereliaablus näitab, et testi erinevad küsimused on omavahel hästi seotud ja mõõdavad ühiseid omadusi.

Hüpooteesid, mis uurisid testide omavaheliste seoste tugevusi, leidsid kinnitust, kuid osaliselt. Paralleeltesti meetodi puhul on oluline, et mõlema testiga oleks võimalik saada sarnaseid tulemusi, mis tähendab, et testide vahel peaks olema tugev korrelatsioon. Korrelatsioonianalüüs kinnitas tugevat positiivset seost sünonüümide sõnavara testi ja selle paralleelversiooni vahel ( $\rho = .66$ ). Sõnapaaride ajendatud meenutamise testi korrelatsioon oma paralleelversiooniga osutus mõõdukalt positiivseks ( $\rho = 0.53$ ).

Passell et al. (2019) leidis, et ka sõnavara testi ja sõnapaaride mäluülesande vahel esineb mõõdukas positiivne korrelatsioon ( $r = 0.37$ ). Käesolevas töös leidis see osaliselt kinnitust. Kaks testi on küll omavahel positiivselt seotud ( $\rho = .13$ ), kuid seda seost võib pidada madalaks. Kuigi mõlemad testid hõlmavad ja kasutavad sõnavara, ei saa teha järeldust, et võime mida need mõõdavad kattuks suures osas. Esinev korrelatsioon võib viidata ka mõõdetavate konstruktide all peituvale üldvõimekusele. Lisaks uuriti mõõtevahendite seoste tugevust maatriksitega ning leiti nõrgad positiivsed korrelatsioonid. Sünonüümide sõnavara keskmise skoori ja Maatriksi skooride vaheline  $\rho = .12$ . Samuti nõrk, kuid veidi tugevam seos esines sõnapaaride testi ja Maatriksite skooride vahel:  $\rho = .23$ . Kogumtesti eesmärki arvestades, milleks on kognitiivsete võimete võimalikult laial moel uurimine, on tegu positiivse leiuga, mis annab infot, et testid töötavad eesmärgipäraselt ja

mõõdavad võimeid laialt. Huvitaval moel toovad *Test My Brain* autorid välja seose esinemise sõnavara testi ja maatriksite vahel ( $\rho = 0.29$ ), kuid sõnapaaride mäluülesande puhul korrelatsioone ei mainita. Siinse töö tulemuse kohaselt võiks oodata, et antud seos on ka suurematel valimitel tuvastatav.

Viienda hüpoteesiga uuriti vanusega kaasnevat muutuseid testide skoorides. Analüüsis saadi tulemus mis on vaimsete võimete uurimisel ka varem korduvalt kinnitust leidnud – sõnavara tundmine on võimekus mis tõuseb peaaegu kogu eluea vältel. Vanuse ja hariduse korrelatsioon graafikul näitas kooskõlas vanuse suurenemisega tõusvat joont. See on seletatav tõenäosusega puutuda pikema elu jooksul kokku rohkemate sõnadega. Teistele vaimse võimekuse valdkondadele on iseloomulik vanuse tõusuga kaasnev langus. Töös uuriti võrdluseks ka mäluülesande skooride ja vanuse seoseid. Koostatud graafikul oli näha kerge aga märgatav skooride langus läbi eluea.

Töös püstitati hüpotees, et kõrgema haridustasemega inimeste testide skoorid on kõrgemad. Vastuse saamiseks koostati haridustasemete kaupa testide keskmistest tulemustest tulpdiagrammid koos usalduspiiridega. Hüpotees leidis osaliselt kinnitust – keskmised skoorid tõusid vähesel määral igal järgneval haridustasemel, kuid kõrgharidustega inimeste skoor oli teaduskraadiga isikutest 0.5 punkti võrra kõrgem. Samas oli haridusskaala ülemises ja alumises osas katseisikute arv kindlate järelduste tegemiseks liiga väike. Märkimisväärne on asjaolu, et tulpdiaagramm illustreeris kõrgeid keskmisi skooore kõikides haridustasemetes. Keskmiste tulemuste variatiivsus oli väike ja põhihariduse ning teaduskraadiga katseisikute keskmiste skooride vahe oli ainult 3.5 punkti. Tulemuste tipnemist normaaljaotuse paremal poolel näitasid ka andmete normaaljaotuse kontrollimiseks loodud histogrammid. Seega selgus, et sõnavara testid on pigem lihtsapoolsed ning ei erista selgelt võimekaid vastajaid vähem võimekatest.

Sõnapaaride ajendatud meenutamise testide keskmiste skooride vaatlusel haridustasemete kaupa on näha tulemuste suurem variatiivsus ning erinevus haridustasemetes joonistub selgemalt välja. Samas ei ole valim erinevate haridusega isikute osas representatiivne ning seetõttu ei saa kindlaid järeldusi teha. Sõnavara testile sarnaselt olid ka mäluülesannete skoorid oodatust kõrgemad ning ka see test osutus oodatust lihtsamaks. Samas esines tulemustes rohkem hajuvust kui sõnavara testi puhul.

Selleks, et teha kindlaks, millised küsimused testidest mõõdavad võimeid kõige paremini, teostati mitmevälja faktoranalüüs koos käsitsi seatud kriteeriumitega. Kriteeriumid selekteerisid välja kõrgema faktorlaadungiga ning vastajate jaoks keerulisemateks osutunud küsimused. Lisaks loodi ülesannete visuaalseks hindamiseks ka ICC graafikud. Faktoranalüüs tõi välja testidesse kõige paremini sobivad küsimused, ning ICC graafikute visuaalne vaatlus kinnitas faktoranalüüsiga

saadud tulemusi. Faktoranalüüs selekteeris kõikidest testidest välja 18 – 21 mõõdetava võimega tugevamalt seotud stiimulit. ICC graafikud aitasid visuaalse vaatlusega eristada stiimulid mis tuleks ülesannete kogumist eemaldada. Kõrge alguspunktiga laugelt kulgevad graafikud annavad infot, et ülesanne on vastajate jaoks liiga lihtne ning et sellel ei ole head eristusvõimet.

Kuna testide skoorid demonstreerivad tipnemist normaaljaotuse paremas osas ning skoorid lähenevad maksimumile, tasub mõelda testide täiendamisele. Üks võimalus selleks on eemaldada liiga lihtsad küsimused kuid lihtsalt raskemate küsimuste alles jätmine ei pruugi parandada testi kvaliteeti. Soovituslik oleks kombineerida uued mõõtevahendid faktoranalüüsi abil eristunud kõrgemate faktorlaadungitega ülesannetest.

### **Uurimissuunad**

Käesoleva töö piiranguks oli meeste ning vanemate inimeste väike osakaal valimis ning madala ja kõrgeima haridustasemega inimeste vähene esindatus. Uuringut võiks korrata täiendatud testidega ning esinduslikul valimil.

*Test My Brain* arendustöö tulemusena on loodud *TestMyBrain Digital Neuropsychological Toolkit* – kliiniliseks kasutuseks mõeldud digitaalne hindamise abivahend. On mõeldav, et tulevikus on olemas eestikeelsed digitaalsed mõõtevahendid, mille usaldusväärsus võimaldab ka klinitsistidel neid kasutada. Sellise mõõtevahendi hulka võiksid kuuluda siinse töö mäluülesanne ja sõnavara test kuid eelnevalt tuleks teostada võrdlev uuring mis annaks infot kui palju erinevad digitaalsel meetodil saadud tulemused klassikaliste mõõtmisviisiga saadutest.

### **Kokkuvõte**

Mäluülesanne ja sõnavara test on kogumtesti olulised elemendid, kuna need mõõdavad kognitiivsete võimeid valkonnas mida teised testid ei pruugi katta. Autor soovib sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testid võtta täiendatud kujul geenivaramus kasutusele. Uue testi kombineerimise korral isegi piloteeritud stiimulite korral tuleb siiski arvestada, et tegu on uue testiga mis vajab enne kasutuselevõttu kordusuuringut.



**Kasutatud kirjandus**

- Asu, L. (2022). Digitaalsete maatriksite MaRs-IB ja Testmybrain pilootuuring [Thesis, Tartu Ülikool].
- Backx, R., Skirrow, C., Dente, P., Barnett, J. H., & Cormack, F. K. (2020). Comparing Web-Based and Lab-Based Cognitive Assessment Using the Cambridge Neuropsychological Test Automated Battery: A Within-Subjects Counterbalanced Study. *Journal of Medical Internet Research*, 22(8), e16792. <https://doi.org/10.2196/16792>
- Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized Neuropsychological Assessment Devices: Joint Position Paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology†. *Archives of Clinical Neuropsychology*, 27(3), 362–373. <https://doi.org/10.1093/arclin/acs027>
- Badham, S. P., Whitney, C., Sanghera, S., & Maylor, E. A. (2017). Word frequency influences on the list length effect and associative memory in young and older adults. *Memory (Hove, England)*, 25(6), 816–830. <https://doi.org/10.1080/09658211.2016.1224358>
- Bowles, R. P., Grimm, K. J., & McArdle, J. J. (2005). A Structural Factor Analysis of Vocabulary Knowledge and Relations to Age. *The Journals of Gerontology: Series B*, 60(5), P234–P241. <https://doi.org/10.1093/geronb/60.5.P234>
- Brybaert, M., Mandera, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, 27(1), 45–50. <https://doi.org/10.1177/0963721417727521>
- Burk, K. (2007). *WAIS-III sõnavara, sarnasuste, informeerituse ja taibukuse alatestide kohandamine*, [Thesis, Tartu Ülikool].
- Chaytor, N. S., Barbosa-Leiker, C., Germine, L. T., Fonseca, L. M., McPherson, S. M., & Tuttle, K. R. (2021). Construct validity, ecological validity and acceptance of self-administered online

neuropsychological assessment in adults. *The Clinical Neuropsychologist*, 35(1), 148–164.

<https://doi.org/10.1080/13854046.2020.1811893>

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>

d'Ardenne, K., Savage, C., Small, D., Vainik, U., & Luke E. Stoeckel, P. (2019). *Core Neuropsychological Measures for Obesity and Diabetes Trials Project: Initial Report (Version 2)*. NutriXiv. <https://doi.org/10.31232/osf.io/7jygx>

Estonian National Corpus. 2019 . Available at <https://www.sketchengine.eu/>, accessed May 2, 2021.

Fawns-Ritchie, C., & Deary, I. J. (2020). Reliability and validity of the UK Biobank cognitive tests. *PLoS ONE*, 15(4), e0231627. <https://doi.org/10.1371/journal.pone.0231627>

Ferrar, J., Griffith, G. J., Skirrow, C., Cashdollar, N., Taptiklis, N., Dobson, J., Cree, F., Cormack, F. K., Barnett, J. H., & Munafò, M. R. (2021). Developing Digital Tools for Remote Clinical Research: How to Evaluate the Validity and Practicality of Active Assessments in Field Settings. *Journal of Medical Internet Research*, 23(6), e26004. <https://doi.org/10.2196/26004>

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286–299. <https://doi.org/10.1037/1040-3590.7.3.286>

Fuente, J. de la, Davies, G., Grotzinger, A. D., Tucker-Drob, E. M., & Deary, I. J. (2019). *Genetic “General Intelligence,” Objectively Determined and Measured* (p. 766600). bioRxiv. <https://doi.org/10.1101/766600>

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in

cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, 19(5), 847–857.

<https://doi.org/10.3758/s13423-012-0296-9>

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, 33(2), 271–286. <https://doi.org/10.1080/13854046.2018.1535662>

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26(4), 433–443. <https://doi.org/10.1177/0956797614567339>

Hessler, J., Fischer, A. M., & Jahn, T. (2016). Differential Linguistic Recall Effects in the California Verbal Learning Test in Healthy Aging and Alzheimer's Dementia: Analysis of Routine Diagnostic Data. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*. <https://doi.org/10.1093/ARCLIN/ACW038>

Hori, K., Fukuhara, H., & Yamada, T. (2022). Item response theory and its applications in educational measurement Part I: Item response theory and its implementation in R. *WIREs Computational Statistics*, 14(2). <https://doi.org/10.1002/wics.1531>

Joormann, J., McLean, S. A., Beaudoin, F. L., An, X., Stevens, J. S., Zeng, D., Neylan, T. C., Clifford, G., Linnstaedt, S. D., Germine, L. T., Rauch, S. L., Musey, P. I., Hendry, P. L., Sheikh, S., Jones, C. W., Panches, B. E., Fermann, G., Hudak, L. A., Mohiuddin, K., ... Kessler, R. C. (2022). Socio-demographic and trauma-related predictors of depression within eight weeks of motor vehicle collision in the AURORA study. *Psychological Medicine*, 52(10), 1934–1947. <https://doi.org/10.1017/S0033291720003773>

Liu, M., Rea-Sandin, G., Foerster, J., Fritsche, L., Brieger, K., Clark, C., Li, K., Pandit, A., Zajac, G., Abecasis, G. R., & Vrieze, S. (2020). Validating Online Measures of Cognitive Ability in Genes for Good, a Genetic Study of Health and Behavior. *Assessment*, 27(1), 136–148. <https://doi.org/10.1177/1073191117744048>

- Lo, A. H. Y., Humphreys, M., Byrne, G. J., & Pachana, N. A. (2012). Test-retest reliability and practice effects of the Wechsler Memory Scale-III: Test-retest reliability and practice effect in WMS-III. *Journal of Neuropsychology*, 6(2), 212–231. <https://doi.org/10.1111/j.1748-6653.2011.02023.x>
- Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and Cognitive Functioning Across the Life Span. *Psychological Science in the Public Interest*, 21(1), 6–41. <https://doi.org/10.1177/1529100620920576>
- Malhotra, N., Krosnick, J. A., & Haertel, E. (2007). The psychometric properties of the GSS Wordsum vocabulary test. GSS Methodological Report, 11, 1–63.
- Masur, K. (2022, mai 13). How to run IRT analyses in R. <https://philippmasur.de/2022/05/13/how-to-run-irt-analyses-in-r/>
- Oliveira, M. O. de, Nittrini, R., Yassuda, M. S., & Brucki, S. M. D. (2014). Vocabulary Is an Appropriate Measure of Premorbid Intelligence in a Sample with Heterogeneous Educational Level in Brazil. *Behavioural Neurology*, 2014, e875960. <https://doi.org/10.1155/2014/875960>
- Passell, E., Dillon, D. G., Baker, J. T., Vogel, S. C., Scheuer, L. S., Mirin, N. L., Rutter, L. A., Pizzagalli, D. A., & Germine, L. (2019). *Digital Cognitive Assessment: Results from the TestMyBrain NIMH Research Domain Criteria (RDoC) Field Test Battery Report*. PsyArXiv. <https://doi.org/10.31234/osf.io/dcszr>
- Passell, E., Strong, R. W., Rutter, L. A., Kim, H., Scheuer, L., Martini, P., Grinspoon, L., & Germine, L. (2021). Cognitive test scores vary with choice of personal digital device. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-021-01597-3>
- Pike, K. E., Kinsella, G. J., Ong, B., Mullaly, E., Rand, E., Storey, E., Ames, D., Saling, M., Clare, L., & Parsons, S. (2013). Is the WMS-IV Verbal Paired Associates as Effective as Other Memory Tasks in Discriminating Amnesic Mild Cognitive Impairment from Normal Aging? *The Clinical Neuropsychologist*, 27(6), 908–923. <https://doi.org/10.1080/13854046.2013.809149>

- Rodosthenous, R. S., Niemi, M. E. K., Kallio, L., Perälä, M., Terho, P., Knopp, T., Punkka, E., Makkonen, E. M., Nurmi, P., Mäkelä, J., Wihuri, P., Hautalahti, M., Moffatt, C., Martini, P., Germine, L., Mäkelä, V. A., Karhunen, O. A., Lahti, J., Hiekkalinna, T. S., ... Ganna, A. (2022). *Re-contacting biobank participants: Lessons from a pilot study within FinnGen* (p. 2022.04.07.22273501). medRxiv. <https://doi.org/10.1101/2022.04.07.22273501>
- Salthouse, T. A. (2014). Evaluating the Correspondence of Different Cognitive Batteries. *Assessment*, 21(2), 131–142. <https://doi.org/10.1177/1073191113486690>
- Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology*, 43(8), 786–795. <https://doi.org/10.1080/13803395.2021.2002269>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Uttl, B., Graf, P., & Richter, L. K. (2002). Verbal Paired Associates tests limits on validity and reliability. *Archives of Clinical Neuropsychology*, 17(6), 567–581. [https://doi.org/10.1016/S0887-6177\(01\)00135-4](https://doi.org/10.1016/S0887-6177(01)00135-4)
- Yan, J., Li, W., Zhang, T., Zhang, Q., Zhang, J., Jin, Z., & Li, L. (2023). Precuneus stimulation alters abstract verbal memory encoding. *Journal of Neurolinguistics*, 65, 101107. <https://doi.org/10.1016/j.jneuroling.2022.101107>
- Younes, N., & Reips, U.-D. (2019). Guideline for improving the reliability of Google Ngram studies: Evidence from religious terms. *PLoS ONE*, 14. <https://doi.org/10.1371/journal.pone.0213554>

## **Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks**

Mina, Vivian Aun,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Sünonüümide sõnavara testi ja sõnapaaride ajendatud meenutamise testi valideerimine,

mille juhendaja on Uku Vainik, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

*Vivian Aun*

**15.05.2023**