

Tartu Ülikool
Sotsiaalteaduste valdkond
Psühholoogia instituut

Laura Asu

Digitaalsete maatriksite MaRs-IB ja Testmybrain pilootuuring

Uurimistöo

Juhendaja: Uku Vainik, PhD

Läbiv pealkiri: Digitaalne neuropsühholoogia

Tartu 2022

Digitaalsete maatriksite MaRs-IB ja Testmybrain pilootuuring

Kokkuvõte

Uurimistöö eesmärk oli piloteerida MaRs-IB ja Testmybrain maatriksite teste. MaRs-IBd läbis 292 ja Testmybraini 341 katseisikut.

Testmybrain maatriksite sisereliaablus oli kõrge ning sarnane varasemate töödega (Cronbach'i $\alpha = .75$, Spearman-Brown = .75). Kinnitavas faktoranalüüsis selgus, et struktuurimudel ennustas hästi käesolevaid andmeid ($\chi^2 = 311.58$, $p = .06$).

MaRs-IB maatriksite sisereliaablus oli kõrge ning eelnevaid tulemusi kordav (Spearman-Brown = .88, Kuder-Richardson 20 = .88). Struktuurimudel töötas käesolevate andmetega ($\chi^2 = 561.48$, $p = .02$, CFI = 0.98, RMSEA = 0.02).

Seadmetest ega haridusaastatest tingitud efekti testide üldskoorile ei kirjeldatud ning puudus sugudevaheline varieeruvus. Testidel ei tulnud esile lae- ega pörandaeefekti.

Autori ettepanek oli järgmises etapis kasutada Testmybrain maatrikseid.

Märksõnad: MaRs-IB, Testmybrain, maatriksid, digitaalne neuropsühholoogia, piloteerimine

Digital matrix tests MaRs-IB and Testmybrain pilot

Abstract

Goal of this research paper was to pilot MaRs-IB and Testmybrain tests. MaRs-IB was taken by 292 and Testmybrain by 341 subjects.

Internal reliability of Testmybrain test was high and similar with previous works (Cronbach $\alpha = .75$, Spearman-Brown = .75). Factor analysis showed that structural model predicts this dataset ($\chi^2 = 311.58$, $p = .06$).

MaRs-IB internal reliability was high and consistent with previous works (Spearman-Brown = .88, Kuder-Richardson 20 = .88). Factor analysis showed that structural model mostly predicts this dataset ($\chi^2 = 561.48$, $p = .02$, CFI = 0.98, RMSEA = 0.02).

Devices, subject's sex nor education level caused variance in the results. Tests don't have floor or ceiling effect.

Author proposes Testmybrain for the next phase.

Keywords: MaRs-IB, Testmybrain, matrices, digital neuropsychology, piloting

Sissejuhatus

Digitaalne neuropsühholoogia

Testid on psühholoogias olulised töövahendid, mida kasutatakse näiteks sisseastumiseksamitel, diagnostikas ja haiguste dünaamika jälgimiseks. Töövahendist on aga ainult siis abi, kui ta annab usaldusväärset infot. Neuropsühholoogia valdkonnas kasutatakse testimisel usaldusväärse tagamiseks ühetaolist protseduuri, traditsioonilist paberit-pliiatsit. Digitaalse neuropsühholoogia testide puhul saab testitav ise valida endale “pliiatsi” ja “paberi”, aga tulemused peavad olema ikka võrreldavad. Uurimistööd on tõestanud, et veebipõhised andmete kogumised saavad olla sama usaldusväärsed kui traditsioonilised meetodid (Germine jt, 2012; Meyerson & Tryon, 2003).

Digitaalne neuropsühholoogia on valdkond, mis tegeleb käitumise ning kognitiivsete võimete hindamisega kasutades digitaalseid seadmeid. Valdkond võimaldab luua uusi innovatiivseid võimalusi hindamiseks neuropsühholoogilist staatust suurel valimil ja pakub lahendusi üksikindiviidi korduvaks hindamiseks tagades ülevaate dünaamikast. Digitaalne neuropsühholoogia pakub testimiseks ning dünaamika jälgimiseks kulutõhusat lahendust. (Germine jt, 2019)

Suurandmete analüüs on tulevik ka psühholoogia valdkonnas, kus sadade või kümnete laboritingimustes olevate katseisikute asemel on digitaalseid vahendeid kasutades võimalik analüüsida tuhandeid inimesi. Selliste valimite saavutamiseks on võimalik teadlastel kasutada mängustamist, mille eesmärk on teha ülesanne isikute jaoks mänguks. Üheks viisiks on pakkuda täitjatele enda kohta kohest tagasisidet. Seda teeb näiteks Testmybrain veebileht. (Woods jt, 2015)

Digitaalse neuropsühholoogia valdkonna oluliseks probleemiks on tagada testide kaasaegsus ja valiidsus samaaegselt tehnoloogia arenguga. Seega kaasaegset seadet kasutades peavad skoorid olema valiidsed. Seadmete eripärad võivad mõjutada testi tajumist ja nõuda erinevaid motoorseid või kognitiivseid võimeid. Seadmete puhul on pidevas muutumises ekraani suurus, resolutsioon, seadmega interakteerumise võimalused, ekraani tüüp ja kaadrisagedus. Seega on tegemist justkui agiilse tarkvaraarendusega, kus kõik on muutumises ning tarkvara pole kunagi valmis. Kuid pidevalt tuleb arvestada testide valiidsuse ning reliablusega. Seega

peab digitaalse neuropsühholoogia valdkond koondama kahte erinevat maailma ning leidma selles sobiva tasakaalu. (Germine jt, 2019)

Digitaalses meediumis on olulised autoriõigused. Kasutuses olevad mitte-verbaalsed testid on sageli kaetud autoriõigustega, mille tõttu ei saa neid vabalt adapteerida ega digitaliseerida. Üks lahendus oleks testide loomisel kasutada vabavavalisuse põhimõtteid koos agiilse arendusprotsessi reeglitega. Testide väljatöötamise põhimõtted ei saa siis vastata majandusliku kasu eesmärgile, vaid on suunatud teaduse arengule. Koostöös on võimalik luua tehnoloogiaga kaasas käivad ning rohkete uuringutega kaetud teste, mille tulemused on seega ka usaldusväärsemad. (Chierchia jt, 2019)

Maatriksid

Maatriksite testid mõõdavad mitte-verbaalset ja visuaalset analüüsioskust, üldist kognitiivset võimekust ning voolavat intelligentsust (Lezak jt, 2012). Maatriksid on ühed parimad testid vaimse võimekuse hindamiseks (Carroll, 1993). Maatriksite skoorid on eluaja jooksul suhteliselt stabiilsed ning vähe mõjutatud ajutistest faktoritest, näiteks on tulemused stabiilsed ka pärast neuroloogilisi kahjustusi ja neid peetakse sageli ka heaks haigusele eelneva intelligentsuse staatuse fikseerimise testiks (Schoenberg jt, 2006; Tranel jt, 2008). Lisaks võrreldes teiste kognitiivsete testidega on antud testi skoor vähe mõjutatud lühiajalistest muutustest tervises (Ryan jt, 2005).

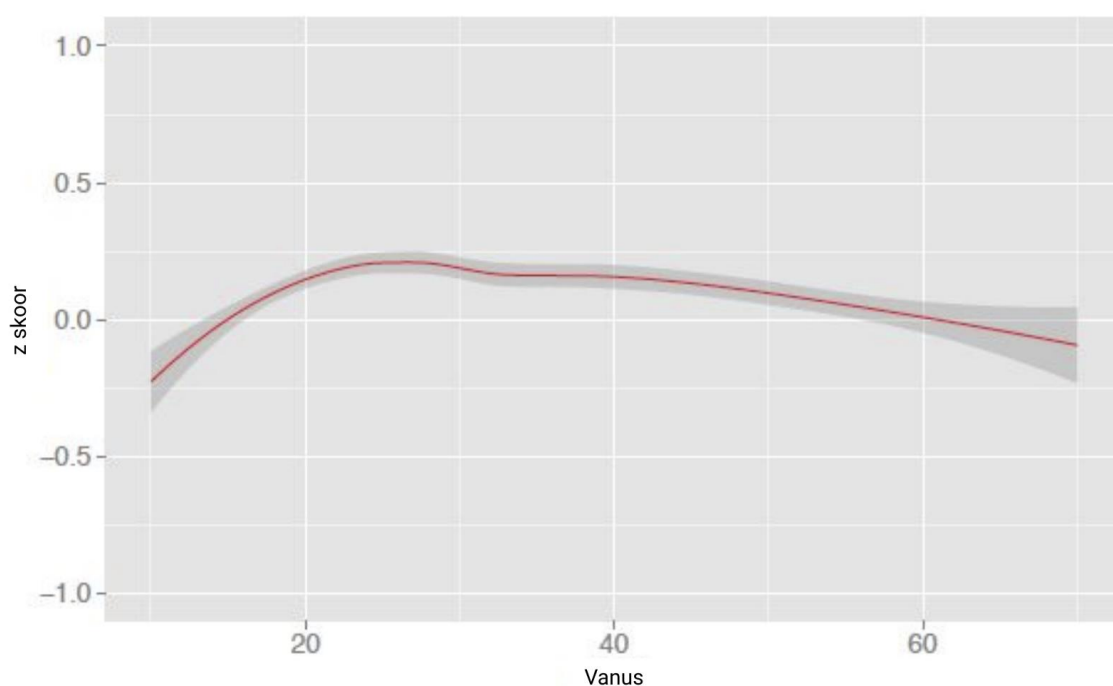
Maatriksite teste on kasutatud neuropsühholoogias aastakümneid ning neid on põhjalikult valideeritud (Passell jt, 2019; Carroll, 1993). Digitaalses meediumis on nende kasutamise eeliseks see, et need on katseisikutele subjektiivselt meeldivad ja hästi talutavad sõltumata nende ajakulukusest ning nende digitaliseerimisel puuduvad olulised tehnilised takistused (Passell jt, 2019). Digitaliseeritud maatriksite test on usaldusväärne meetod hindamaks kognitiivseid võimeid erinevatel digitaalsetel seadmetel näiteks arvutis, nuhvliis või nutitefonis (D'Ardenne jt, 2020).

Käesoleva uurimistöö testipatarei valik põhineb USA terviseinstituudi soovitusel käitumuslike testide lisamisest terviseuuringutesse (Stoeckel jt, 2019). Uurimistöös piloteeritakse maatriksite teste, mis on võetud kahest erinevast allikast - MaRs-IB ja Testmybrain. Eelnevates uuringutes on esinenud Testmybrain maatriksitel laeefekt, seega

käesolevasse uurimistöösse on valitud lisaks MaRs-IB test, millel puudub lae- ning põrandaeft (Passell jt, 2019; Chierchia jt, 2019).

Testmybrain maatriksite test on mudeldatud hea valiidsusega *Wechsler Abbreviated Scale of Intelligence II* (WASI II) testi alusel. Valideerimisel hinnati 20 510 tulemust. Uuringu käigus leiti, et kõrgema haridusega katseisikud said paremaid tulemusi. Inimese arenguperioodil toimub skoori kiire tõus ning alates 30ndast eluaastast toimub aeglane langus (Joonis 1).

Testide sisereliaabluse (Cronbach'i $\alpha = .77$, Spearman-Brown = $.89$) on piisav testi edasiseks kasutamiseks. (Passell jt, 2019)



Joonis 1. Testmybrain maatriksite Z-skoori ja vanuse seose graafik pärineb Passelli 2019. aasta uuringust

Testmybrain digitaliseeritud maatriksitel puuduvad olulised seadmespetsiifilised takistused. Isegi nutitelefonidel ei ole kujundid liiga väikesed ülesande lahendamiseks. Püsiva asukohaga seadmete grupp erineb kaasas kantavate seadmete grupiga, mida võib põhjendada stabiilsema töökeskkonnaga. (Passell jt, 2019)

Testmybrain test on aju tervise indeksi hindamisel kasutusel järgEVates projektides - “NIH Core Neuropsychological Measures of Diabetes and Obesity Project” ja “The Broad Institute Neuropsychological Phenotyping Project” (Passell jt, 2019). Näiteks on ühe projekti eesmärk

luua vaimsete võimete hindamise indeks, mis põhineb maatriksite ning numbri-sümboli vastavuse testil ning kasutada seda diabeedi ning ülekaalu diagnoosiga patsientidel (D'Ardenne jt, 2020). Lisaks on pandeemia perioodil Testmybrain keskkonnal põhinedes loodud neuropsühholoogiline testipatarei vaimsete võime hindamiseks kliinilises praktikas (Singh jt, 2021).

Testmybrain testipatareis on maatriksite test nõrgas positiivses ($r = .38$) seoses numbri-sümboli vastavuse testiga (Passell jt, 2019). Numbri-sümboli vastavuse test mõõdab visuaalset lühimälu ja info töötlemise kiirust ning Testmybrain versioon põhineb WAIS numbri-sümboli vastavuse testil, mida on pikalt kasutatud kliinilises neuropsühholoogias (Joy jt, 2004). WAIS numbri-sümboli vastavuse test on paberil põhjalikult valideeritud ning on laialdaselt kasutuses töötlemiskiirust mõõtva testina, aga digitaalse testi tulemusi mõjutab katseisiku poolt kasutatud seade, sest tulemus on mõjutatud osaliselt reaktsioonikiirusest, mis on omakorda mõjutatud latentsist (Passell jt, 2019).

MaRs-IB sisaldab 80 uut maatriksit, mis on loodud vabatarkvarana. Testi sisereliaablus (Spearman-Brown = .82, Kuder-Richardson 20 = .78) on valideeritud 659 katseisikuga. Kirjeldavate statistikute alusel on tulemused kognitiivsete printsiipidega vastavuses. Näiteks on vanusel tugev lineaarne positiivne efekt ja oluline negatiivne trend vanuse ruuduga. (Chierchia jt, 2019)

Testmybrain ja MaRs-IB uuringutes ei ole meeste ja naiste skooride vahel statistiliselt olulisi erinevusi (Passell jt, 2019; Chierchia jt, 2019). Intelligentsuse skooride puhul on aga leitud, et meeste skoorid on varieeruvamad (Johnson jt, 2008). Kuna Testmybrain ja MaRs-IB uuringud ei kirjeldanud meeste skooride varieeruvuse olemust, siis käesolevas uurimistöös on plaanis seda analüüsida.

Kooli sisseastumiseksamitel kasutatavad testid mõõdavad maatriksitele sarnaselt vaimset võimekust (Silm jt, 2013). Kui maatriksite teste peetakse üheks parimaks testiks vaimse võimekuse hindamiseks (Carroll, 1993), siis tekkis autoril küsimus sisseastumiseksamitel kasutatud vaimsete võimekuse testide seosest maatriksitega. Käesolevas töös kasutatakse Tartu Ülikooli lühendatud akadeemilist testi, mida on võimalik sarnaselt Testmybrainile ja MaRs-IBile täita digitaalselt.

Töö eesmärk ja hüpoteesid

Digitaliseeritud maatriksite testid võimaldaksid kuluefektiivset testimist suurematel valimitel, näiteks TÜ GI Eesti geenivaramuga koostöös. Viimase puhul on võimalik analüüsida seoseid isiksuse, geneetiliste markerite ja terviseiga. Seega saab uurida geenide, elukeskkonna ja eluviisi vahelisi seoseid. Lisaks loob digitaalne testimine võimaluse hinnata isikute staatust dünaamikas, mis annab võimaluse vähendada konkreetsest hetkest tingitud tulemuste ebakorrektsust.

Pilootprojekti käigus teostati kasutajamugavuse testid ning nende analüüs. Uurimistöö autor teostas 12 kasutatavuse testi. Testimised teostati poolstruktureeritud meetodil ehk lisaks skripti küsimustele võis läbiviija küsida juurde täpsustavaid küsimusi. Kasutatavuse testi käigus tegi katseisik läbi kogu testipatarei ning samal ajal kommenteeris enda kogemust, seda siis samal ajal jälgis intervjuueeritav ning küsis jooksvalt täpsustavaid küsimusi. Analüüsi käigus avastatud olulisemad fakte kasutati uurimisküsimuste esitamiseks. Katseisikud esitasid korduvaid märkusi testi pikkuse osas. Esines katseisikute väsimist ning viimaseid ülesandeid lahendati kohati juhuslikkuse alusel. Seega tekkis küsimus, kas ja kuidas oleks võimalik testile kuluvat aega vähendada. Chierchia (2019) uuringus oli MaRs-IB maatriksite puhul kasutuses ajalimiit, aga see täitus ainult 2% testitutest. Lisaks testimise käigus katseisikud kommenteerisid, et MaRs-IB maatriksite puhul on neil probleeme nutitelefoni tajutavate väikeste elementide loetavuses ning arusaadavuses. Katseisikud üritasid suurendada kujundeid, muuta ekraani seadeid ning vaadata kujundeid erinevate ekraaninurkade alt. Testmybrain maatriksite puhul esines neid kommentaare harvem.

Käesoleva uurimistöö eesmärk on pilootuuringu abil valideerida eestikeelset maatriksitesti digitaalseks kasutamiseks. Pilootprojekti käigus adapteeriti kahest erinevast allikast maatrikseid - Testmybrain ja MaRS-IB. Testi saab teha arvutis, nuhvlil ja nutitelefonis. Käesolev pilootuuring valmistab ette testipatarei rakendamist geenivaramu doonoritega.

Loetud kirjandusele ja kasutatavuse testidele toetudes püstitati järgnevad hüpoteesid:

Hüpoteesid 1-5: Võrdlus eelnevate uuringutega

Hüpotees 1. Testmybrain andmebaasi maatriksite sisereliaabluse on korratav (Cronbach'i $\alpha > 0.7$, Spearman-Brown > 0.8) sarnaselt Passelli (2019) uuringuga.

Hüpotees 2. MaRS-IB andmebaasi maatriksite sisereliaablus on korratav (Spearman-Brown > 0.8 , Kuder-Richardson $20 > 0.7$) sarnaselt Chierchia (2019) uuringuga.

Hüpotees 3. Kognitiivsete võimete muutumise seost vanusega kinnitab maatriksite skoori positiivne lineaarne seos vanusega ja negatiivne seos vanuse ruuduga. Vanuse ruut peaks iseloomustama vanaduses tekkivat kognitiivsete võimete langust.

Hüpotees 4. Pikema haridusteedga katseisikul on kõrgem maatriksite skoor.

Hüpotees 5. Maatriksite skooride variatsioon on suurem meestel kui naistel.

Hüpotees 6: Struktuuri hindamine

Hüpotees 6. MaRS-IB ja Testmybrain maatriksite struktuurvõrrandite ühefaktorilised mudelid kirjeldavad tegelikke andmeid piisavalt hästi.

Hüpoteesid 7-9: Testidevahelised seosed

Hüpotees 7. MaRS-IB ja Testmybrain skooridel on tugev positiivne seos ($r > .7$).

Hüpotees 8. MaRS-IB ja Testmybrain skoor on nõrgas positiivses seoses numbri-sümboli vastavuse testi tulemusega ($r > .3$)

Hüpotees 9. MaRS-IB ja Testmybrain skooril on tugev positiivne seos ($r > .7$) Tartu Ülikooli lühendatud akadeemilise testi matemaatika ning ruumilise mõtlemise alamtesti skooriga.

Hüpoteesid 10-11: Skoori mõjutatavus uurimisvahenditest

Hüpotees 10. MaRS-IB ja Testmybrain skooril puudub statistiliselt oluline seos ekraani laiusega, välja arvatud kui suurus on alla 500 piksli.

Hüpotees 11. MaRS-IB ja Testmybrain skoorid on väiksema variatiivsusega lauaarvutil teostatuna võrreldes ülejäänutega (nutitelefonil, nuhvliil, sülearvutiga).

Hüpoteesid 12-14: Piloteeritavad hüpoteesid

Hüpoteeside 12.1-6 eesmärk on võrrelda katseisikute subjektiivset arvamust maatriksite testidest.

Hüpoteeside 13.1-6 ja 14.1-6. eesmärk on hinnata võimalust testide kestvuse lühendamiseks ning hinnata seda maatriksi tüüpidel ning seadmetel eraldi (Tabel 1). Esiteks maatriksite testi

lõplikku skoori ei mõjuta testi katkestamine peale järjestikust seitsmendat või viiendat vale vastust (hüpoteesid 13.1-6). Teiseks maatriksite testi skoor on negatiivses seoses võrreldes testi katkestamisega peale kolmandat vale vastust (hüpoteesid 14.1-6).

Tabel 1

Ettepanekute analüüs ja seadme tüübi efekt skoorile

Maatriksi tüüp	Arvuti	Nuhvel	Nutitelefoni
MaRs-IB	Hüpotees 12.1. Testitavad hindavad maatriksite testi subjektiivselt meeldivamaks võrreldes teiste testide keskmisega ($M > 2.5$).	Hüpotees 12.2.	Hüpotees 12.3. Ebameeldivamaks võrreldes teiste testide keskmisega ($M < 2.0$)
	Hüpotees 13.1. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.	Hüpotees 13.2. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.	Hüpotees 13.3. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.
	Hüpotees 14.1. $r = -0.3$	Hüpotees 14.2. $r = -0.3$	Hüpotees 14.3. $r = -0.3$
Testmybrain	Hüpotees 12.4. Subjektiivselt meeldivamaks võrreldes teiste testide keskmisega ($M > 3.5$).	Hüpotees 12.5.	Hüpotees 13.6. Subjektiivselt meeldivamaks võrreldes teiste testide keskmisega ($M > 2.5$).
	Hüpotees 13.4. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.	Hüpotees 13.5. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.	Hüpotees 13.6. Testi katkestamine viie või seitsme järjestikuse eksimusega ei mõjuta statistiliselt oluliselt üldskoori.
	Hüpotees 14.4. $r = -0.3$	Hüpotees 14.5. $r = -0.3$	Hüpotees 14.6. $r = -0.3$

Meetodi kirjeldus

Valim

Valim oli mugavusvalim, kuhu kuulusid mehed ja naised alates 18. eluaastast. Valimit piiras arvutikasutuse oskus. Katseisikud pidid olema võimelised iseseisvalt osalema uuringus kasutades nutitelefon, nuhvlit või arvutit. Katseisikud pidid osalema testimisel kahel korral ning soovitatavalt erinevate seadmetüüpidega. Kahe testimise vahel olev paus oli vähemalt seitse päeva. Katseisikud ei saanud osalemise eest rahalist ega mitte-rahalist tasu.

Lühendatud akadeemilise testi andmekogumise teostas Kenn Konstabel oma kursuse raames. Kenn Konstabel edastas maatriksite testi läbinud katseisikute akadeemilise testi tulemused uurimistöö autorile.

Käesoleva uurimistööga paralleelselt analüüsis kogutud andmeid uurimistöö kontekstis Adams (2022) teemal „Elusündmuste seosed Testmybrain kognitiivsete testidega“ ja Berezin (2022) teemal „Digitaalse punktide ühendamise testi, lihtreaktsioonaja testi ning valikreaktsioonaja testi valideerimine“.

Andmekogumisviis

Katsekutset levitati Tartu Ülikooli psühholoogia instituudi meililistides, suurfirmade töötajate seas, uurijate suhtlusvõrgustikes ja sotsiaalmeedias perioodil detsember 2020 kuni mai 2021. Katseisikute leidmiseks kasutati ka tasulisi reklaame sotsiaalmeedias.

Testide kohandamise protsess oli teostatud enne autori liitumist projektiga. Uuringu disainiga saab tutvuda veebiaadressil:

<https://cognit.psych.ut.ee/tmb1/run/run.php?study=piloot1a&cc=testkasutaja&sequence=3> .

Lahenduse lähtekoodi asub järgneval aadressil:

https://gitlab.ut.ee/uku.vainik/present_Testmybrain .

Mõõtmisvahend

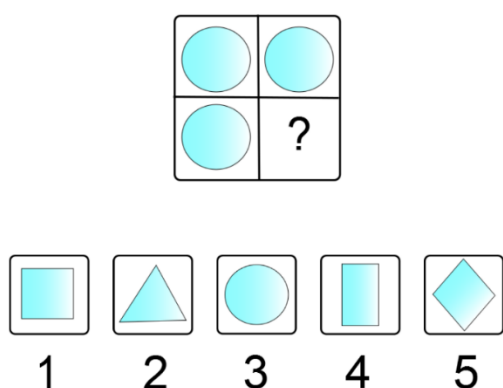
Pilootuuringu käigus adapteeriti vaimse võimekuse testipatarei eesti keelde. Teste sai katseisik teha enda seadmega ning selleks võis olla arvuti, nuhvel ja nutitelefon.

Katsemõõdikud keskendusid vaimsetele võimetele ja kogu testipatarei täitmine võttis keskmiselt 30 minutit. Käesolev uurimistöö keskendub maatriksite testile ning võrdleb neid

ka numbri-sümboli vastavuse testiga. Testimisel oli kasutatud Testmybrain ja MaRs-IB maatriksite kujundeid.

Testmybrain keskkonna maatriksite test mõõdab tähelepanu, töömälu, taju, üldist kognitiivset võimekust, üldist intelligentsust, voolavat intelligentsust ja mitte-sõnalist analüüsi. Katseisik pidi tuvastama, milline pilt või muster lõpetab loogikareeglitele vastava jada (Joonis 2).

Passelli uuringus oli Testmybrain maatriksite testi läbimise aeg keskmiselt 8.5 minutit. Skooriks oli õigete vastuste arv. (Passell jt, 2019)

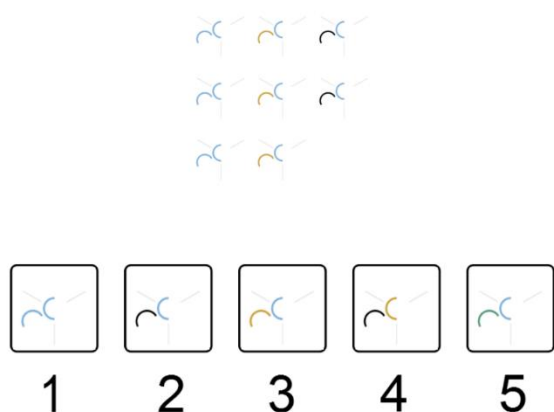


Joonis 2. Testmybrain keskkonna maatriksi näidis. Testmybrain keskkonna vabavalalised maatriksid.

MaRs-IB maatriksite test sisaldas 3 x 3 maatriksit (Joonis 3). Katseisikule kuvati kaheksa abstraktset kujundit ning all paremal oli üks kujund puudu. Katseisiku ülesandeks oli valida viie variandi alusel tühja kohta sobiv maatriks. Kujundite puhul oli kasutatud nelja muutujat - kuju, värvi, suurust ja asukohta. Skooriks oli õigete vastuste arv. (Chierchia jt, 2019)

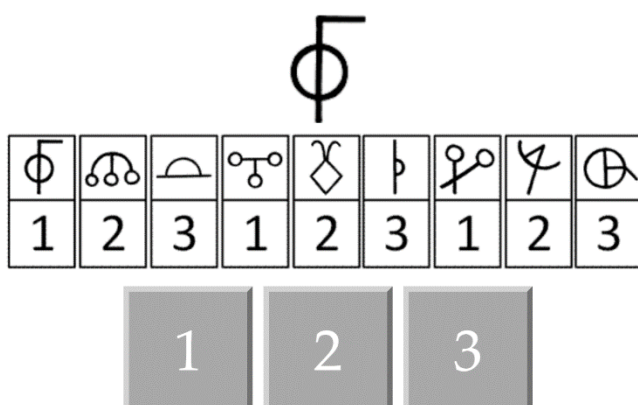
Osalejatele esitati MaRs-IB kujundeid kasutades Testmybrain esitlustarkvara. Suuremast kujundite pangast valiti välja stiimulid, mis oleks nutitefonil võimalikud selgesti näha ja kataks võimalikult erinevaid raskusastmeid. Samuti vähendati vastusevariantide arvu, et see sobiks Testmybrain tarkvaraga ning jälgiti, et vastusevariandid oleks selgelt erinevad.

Kaasatud stiimulite nimekiri on välja toodud (Tabel A1).



Joonis 3. MaRs-IB maatriksi näidis. MaRs-IB andmebaasi vabavaralised maatriksid.

Numbri-sümboli vastavuse test (Joonis 4) mõõdab visuaalset lühimälu ja töötlemise kiirust. Testi kogu kestus oli keskmiselt 2.9 minutit, mis sisaldas juhendiga tutvumist, harjutusülesandeid ja testi. Testi käigus oli vaja vajutada klaviatuuril klahvi 1, 2, 3 või vastavate nimetusega kaste ekraanil, mis vastab otsitavale sümbolile. Testis kasutati üheksat sümbolit. Testi aeg oli piiratud 90 sekundiga. Skooriks oli õigete vastuste arv. (Passell jt, 2019)



Joonis 4. Numbri-sümboli vastavuse testi näidis. Testmybrain keskkonna vabavaraline number-sümboli vastuse test.

Tartu Ülikooli lühendatud akadeemilise test sisaldas kolme alamskaalat – sõnavara, matemaatika ja ruumiline mõtlemine. Matemaatiline ja ruumiline mõtlemine mõõdavad sarnaselt maatriksitele voolavat intelligentsust. Igas alaskaalas oli 15 ülesannet. Test oli loodud kasutades õpilaste tulemusi aastatel 2008– 2012 saavutamaks empiirilisel alatestides võrdse keerulisuse taseme. Tegemist oli digitaliseeritud testiga, mille ajapiiriks oli 60 minutit.

Katseisikutel oli aega kõigile küsimustele vastata aga polnud võimalik teha pikemaid pause küsimuste vahel. Testi viis läbi Kenn Konstabel Tartu Ülikooli tudengitega (Silm jt, 2013).

Protseduur

Sotsiaalmeediast ja meililistidest suunati katseisikud uuringu kodulehele. Veebilehel said katseisikud ülevaate uuringust ning osalemiseks oli vaja sisestada e-posti aadress.

Katseisikutele saabus ühe kuni kümne minuti jooksul e-kiri, mis sisaldas randomiseeritud soovitusi alustada testi süle- ja lauarvuti või nutitelefoni ja nuhvliga. Katse alustamiseks pidi isik vajutama e-kirjas olevale lingile ning ta suunati uuringu detailsema kirjelduse vaatesse. Detailsem kirjeldus sisaldas uuringu ülevaadet, üldist informatsiooni, andmekaitse nõudeid ning osalemiseks nõusoleku andmist.

Enne testipatarei juurde asumist küsiti formr.org keskkonnas katseisikult demograafilisi ja sotsiaalmajanduslikke andmeid: näiteks vanus, sugu, haridus, elukohta ning töötamise staatus.

Järgnevalt osalesid katseisikud individuaalselt vaimset võimekust hindava testipatarei täitmisel cognit.psych.ut.ee keskkonnas. Testipatarei testide järjekord oli randomiseeritud, kuid sisaldas teatud reeglid: joonetest A ja joonetest B ning reaktsiooniaja ja valikreaktsiooniaja test olid alati üksteise järel. Mälutesti meeldejätmise ja meenutamise vahel oli alati numbri-sümboli vastavuse test.

Maatriksite test algas juhendi ning kahe harjutusülesandega. Katseisik pidi läbima mõlemad harjutustestid korrektselt, et alustada ülesande lahendamist. Maatriksi alamtesti puhul oli randomiseeritud esimesena kas Testmybrain või MaRs-IB maatriksid. Kujundite järjekorral oli aluseks võetud Testmybrain protseduur, kus üldises plaanis olid need järjestatud kergemast raskemani, kuid järjekord oli mõnevõrra segamini aetud, et see oleks vähem tajutav. Kõigile katseisikutele oli järjekord sama. Maatriksite testi täitmise aeg ei olnud piiratud. Testi tegemisel puudus dialoogiaken kiirustamiseks katseisikut otsust tegema. Katses olevate maatriksite arv oli 35.

Numbri-sümboli vastavuse testile eelnevalt kuvati katseisikule juhend ning kolm harjutusülesannet.

Testimise lõpus andis katseisik tagasisidet ning hindas testipatarei alamteste formr.org keskkonnas. Lisaks küsiti katseisikute käest kasutatud seadmete, motivatsiooni, strateegiate ning tehnikate kohta.

Kordusuuring oli planeeritud seitse päeva hiljem, kui katseisikule edastati e-posti vahendusel meeldetuletus ning info, et nüüd peaks ta testi teist tüüpi seadmega läbima. Testivatele edastati kahepäevase pausiga maksimaalselt kolm meeldetuletust.

Kordusuuringu protseduur oli esmase läbimisega võrreldes sama. Ainukeseks erandiks oli teise sessioon esimene test, milleks oli alati eelmise korra mälustiimulite meenutamine.

Statistiline analüüs

Statistilise analüüsi käigus kasutati kirjeldavate statistikute leidmiseks, andmete puhastamiseks, struktuuri hindamiseks, sisereliaabluse ning korrelatsioonide arvutamiseks statistikaprogrammi R. Skriptide loomisel kasutati R Studiot ning koodi hallati kasutades Gitlabi.

Statistilise analüüsi esimeseks etapiks oli andmete puhastamine. Esimene puhastamise faas toimus enne uurimistöö autori kaasamist. Esiteks eemaldati katseisikud, kellel oli alla viie õige vastuse maatriksite testis. Teiseks eemaldati katseisikud, kes number-sümbol testis valisid üle 90% sama vastusevariandi või vastasid alla 33% korrektselt. Eemaldati ka katseisikud, kes number-sümbol testis jätsid üle 15% ülesannetest vastamata. Lisaks eemaldati katseisikud, kelle number-sümbol testi reaktsiooniaeg oli alla 600 ms või üle 6000 ms.

Autor eemaldas valimist katseisikute andmed, kui kasutajad teostasid rohkem kui ühe katse (MaRs-IB ja Testmybrain kolm isikut) või kinnitasid välise abi kasutamist (MaRs-IB 9 ja Testmybrain 13 isikut). Andmete analüüsi käigus selgus, et üks kasutaja teostas mõlemad testid samal päeval ning teostas veel korduvkatse, seega tema andmeid sai analüüsist eemaldatud. Eemaldatud said kasutajad, kellel puudus haridusaastate (MaRs-IB ja Testmybrain kuus isikut) või vanuse (MaRs-IB 7 ja Testmybrain 8 isikut) informatsioon.

Haridusaastatega seotud hüpoteeside jaoks esindusliku valimi saavutamiseks jäeti välja alla 25-aastased katseisikud, kuna nende vastamise hetkel saavutatud haridustase ei pruugi olla maksimaalne. Vastasel juhul oli nende üldskooride ning hetke haridusaastate analüüsimisel

võimalik teha valesid järeldusi, kuna ühes grupis oleks nii tudengid kui ka keskharidusega isikud, kes ei läbi järgmist haridusastet. MaRs-IB vastustest eemaldati 80 ning Testmybrain puhul 84 isikut. Muude hüpoteeside puhul olid kaasatud ka alla 25-aastaste isikute skoorid.

Valimi kirjeldamiseks kasutati R funktsioone – *mean()*, *median()*, *summary()*, *min()*, *max()*, *length()*, *sd()*. Protsendiarvutusteks ning tabeli visualiseerimiseks kasutati Microsoft Excelit.

Testi sisereliaablus hindab, mil määral mõõdavad üksikülesanded ehk antud juhul pildid üht ja sama omadust. Eesmärk oli korrata eelnevate uuringute tulemusi ehk analüüside valik põhines Chierchia (2019) ja Passell'i (2019) uuringutel. MaRs-IB maatriksite puhul analüüsiti Kuder-Richardson 20 ja Spearman-Browni koefitsiente ning võrdlemaks Testmybrain testiga ka Cronbach'i α . Testmybrain maatriksite puhul arvutati Spearman-Browni koefitsient ja Cronbach'i α ning võrdlemaks MaRS-IB testiga ka Kuder-Richardson koefitsient. Kõrge sisereliaabsusega on test, kui koefitsiendi väärtus on > 0.7 (Cronbach, 1951). Cronbach'i α arvutati kasutades psych paketist funktsiooni *alpha()* (Revelle, 2021). Kuder-Richardsoni ning Spearman-Browni väärtuse arvutamiseks kasutati Castro (2021) loodud koodi. Sisereliaabluse kontrollimiseks teostati ka McDonald'i Omega kasutades psych paketist olevat funktsioon *omega()*, mis on alternatiivne sisereliaabluse hindamise vahend ühefaktorilise üldskoori puhul (Dunn jt, 2014; Revelle, 2021).

Vanuse, vanuse ruudu (hüpotees 3) ja haridusaastate (hüpotees 4) analüüsiks kasutati Spearmani astakorrelatsioonanalüüsi, mis teostati kasutades *psych* paketist *corr.test()* (Revelle, 2021). Lisaks teostati lineaarse regressiooni mudelid kasutades R tarkvara baasfunktsionaalsuse alla kuuluvat *stats* paketi funktsiooni *lm()* (R Core Team, 2021). Nii korrelatsioone kui ka mittelineaarset seost kontrolliti alla 25-aastasteta kui ka nendega koos. Mittelineaarset seoseid visualiseeriti kasutades *ggplot2* teeki (Wickham, 2016).

Hüpoteesi 5 kontrollimiseks ehk sugudevahelise skoori varieerumise kirjeldamiseks kasutati väärtuste vahemikku (*range*) ja standardhälvet (SD). Varieeruvuse homogeensuse hindamiseks teostati baasfunktsionaalses *stats* paketi Bartlett test ja *car* paketi Levene test (Bartlett, 1937; Brown & Forsythe, 1974; Fox & Weisberg, 2019). Normaoljaotuse hindamiseks loodi histogramm. Rühmade varieeruvuse statistiliseks testiks kasutati F-testi, mis teostati kasutades *stats* paketti.

Struktuurvõrrandite mudeli (hüpotees 6) hindamise eesmärk oli koostada mudel ning hinnata sobivust andmetega ehk kui hästi kirjeldas koostatud mudel kogutud andmeid. Analüüs teostati kasutades lavaani paketti (Rosseel, 2012), mille abil hinnati MaRs-IB ja Testmybrain maatriksite mudelit. Lavaani mudeli koostamisel kasutati cfa (confirmatory factor analysis) analüüsi, millel vaadeldavad muutujad olid vastava testi üksikküsimused ning latentsed muutujad olid vastavalt MaRs-IB või Testmybrain maatriksite testisooritused. Lavaani analüüsis kasutati DWLSi meetodit binaarsete andmete tõttu ehk diagonaalselt kaalutud vähimruudu meetodit (Rosseel, 2012). Kinnitava faktoranalüüsi tulemuste hindamisel on oluline näitaja hii-ruut (χ^2), mis testib täpse sobitumise hüpoteesi, mille kohaselt ei erine mudeli põhjal arvatud kovariatsioonimaatriks andmestiku põhjal arvatud kovariatsioonimaatriksist ning selle kinnitamiseks peaks olema hii-ruut väike ning p väärtus olema üle 0.05 (Hu & Bentler, 1999). Kui p on alla 0.05, siis hinnatakse, kui hästi suudab koostatud mudel taastada tegelike andmete põhjal arvatud kovariatsioonimaatriksi. Mudeli sobivusele viitavad järgnevad väärtused - CFI (soovituslikult üle 0.95), TLI (soovituslikult üle 0.9), RMSEA (alla 0.08) ja SRMR (alla 0.08) (Hu & Bentler, 1999). Mudeli loomisel esines probleem üle 95% õigesti vastatud stiimulitega, kuna vähestel valesti vastamistel puudus piisav variatsioon ülejäänud stiimulitega ning töötava mudeli jaoks need stiimulid eemaldati.

Rizopoulos (2006) alusel teostati *item response theory* (IRT) analüüs vastavalt ühe faktori ning 2PL mudeli abil (R funktsioon (*ltm*)), mille tulemusi uuriti *summary()* ning *coef()* abil. Loodi ka 3PL mudeli analüüs (R funktsioon (*tpm*)) (Rizopoulos, 2006). Välistamiseks mudelile katseisiku poolt juhuslikult vastuse valimise mõju võrreldi 3PL ja 2PL mudelite *anova()* abil. IRT tulemuste visualiseerimiseks kasutati testi mõõtmistäpsuse graafikuid (*test information function*). Viimasena arvutati välja faktorskooride (*lavPredict()*) korrelatsioonikoondskoori suhtes kasutades Spearmani astakkorrelatsioonanalüüsi hindamiseks üldskoori usaldusväärust.

Hüpotees 7 ja 8 ehk testide omavaheliseks võrdlemiseks kasutati Spearmani astakkorrelatsioonanalüüsi ja lineaarse regressiooni mudelit. Visualiseerimiseks kasutati ggplot2 teeki (Wickham, 2016). Hüpoteesi 9 jaoks kasutati Spearmani astakkorrelatsioonanalüüsi. Sõltuvate gruppide omavaheliste korrelatsioonide analüüsiks kasutati Cocori teeki ja analüüsi meetodiks Pearson and Filoni z testi (Diedenhofen & Musch, 2015).

Uurimisvahendi mõju üldskoorile uuriti Spearmani astakkorrelatsiooni abil ja loodi lineaarsed regressiooni mudelid. Mittelineaarsete seoste hindamiseks kasutati graafikuid, mis loodi ggplot2 teeki kasutades ning nende diagrammide koondamiseks *patchwork* teeki (Pedersen, 2020). Skooride variatiivsuse kirjeldamiseks arvutati väärtuste vahemik ja standardhälve. Normaalkaotuse hindamiseks teostati histogramm ning kinnitati Shapiro-Wilki testiga. Kuna andmeid ei vastanud normaalkaotuse tingimustele, siis gruppide varieeruvuse statistiliseks hindamiseks kasutati *stats* paketi Mann-Whitney U mitteparameetrilist testi.

Subjektivse meeldivuse kirjeldamiseks kasutati keskmist ja väärtuste vahemikku. Keskmise arvutati MaRs-IB testile, Testmybrain testile, kõigile üksiktestidele välja arvatud maatriksid ning analüüsi ka vastavalt seadme tüüpidele. Rühmadevahelisi erinevusi hinnati statistilise t-testiga. Loodi ka lihtsad lineaarsed regressiooni mudelid.

Viimaseks uuriti kolme, viie või seitsme järjestikeksimuse üldskoori võrreldes ilma katkemiseta saavutatud üldskooriga kasutades Spearmani astakkorrelatsioonanalüüsi ja lineaarse regressiooni mudeli loomist.

Uuringu eetiline külg

Uuringus osalemine oli vabatahtlik ning lubatud alates 18. eluaastast. Enne testi tutvustati lühidalt uuringu eesmärgi ja andmete kogumise põhimõtteid. Kõik katseisikud andsid oma nõusoleku uuringus osalemiseks enne küsimustiku täitmist. Osalejatele ei pakutud rahalist ega mitterahalist tasu. Katseisikul oli võimalik iga hetk osalemisest loobuda.

Uuring oli kooskõlas mittekahjustamise printsiibiga, kuna testide täitmisega ei kaasne uuritavale teadaolevaid riske. Küsimustikule tagasiside saamine oli vabatahtlik.

Andmeid koguti Euroopa Liidu serveritesse, mis vastasid GDPR nõuetele ja kuhu oli ligipääs vaid uuringu läbiviijatel ning eksperimentaalpsühholoogia labori töötajatel. Andmed edastati töötluseks anonümiseeritult. Analüüsi avaldamisel teadusartiklina avaldatakse anonümiseeritud andmed ka <https://osf.io/uv9sw/> veebilehel, kus need säilivad tähtajatult. Lisaks oli katseisikutel võimalik luua testimiseks anonüümne ning ainult selleks kasutatav e-posti aadressi, mille puhul pole võimalik seost isikuga luua.

Uuringu puhul ei säilitata isikuandmeid ning ei küsitud tundlikke andmeid, seega ei vajanud uuring TÜ eetikakomitee juhi hinnangul Tartu Ülikooli Inimuuringute Eetika Komitee poolt kinnitust.

Autori panus

Hüpoteeside loomiseks analüüsis autor teadustöid ja teostas kasutatavuse testid ning analüüsi. Uurimisküsimustele vastamisel oli autori panus andmete puhastamine, süstematiseerimine, analüüs, tulemuste süntees ning uurimistöö kirjutamine.

Tulemused

Ülevaade andmestikust

Andmete puhastamise järgselt oli Testmybrain (TMB) maatrikseid teostanud 341 ja MaRs-IB testi 292 katseisikut. Haridusaastatega seotud hüpoteeside jaoks esindusliku valimi saavutamiseks jäeti välja alla 25-aastased katseisikud ehk arvestati TMB 257 ning MaRs-IB 212 tulemust.

Mõlemaid maatriksite komplekte läbisid 233 katseisikut. Kolm katseisikut lahendasid ainult MaRs-IB maatrikseid. 52 katseisikut läbisid TMB maatriksid aga jätsid pooleli MaRs-IB testi. MaRs-IB pooleli jätnud katseisikute vanuse ($M = 41$, väärtuste vahemik = 20 – 67), soolise (naised = 43, mehed = 9), haridusaastate ($M = 16$, väärtuste vahemik = 13 – 19), seadmete (arvuti = 29, nutitelefon = 21, nuhvel = 2) proportsioon oli sarnane üldisele valimile. Loobunud katseisikute subjektiivne meeldivus loobunud MaRs-IB maatriksi testile oli 2.37 ja raskuse hinde keskmine 4.04. Kogu valimi keskmine meeldivus MaRs-IB testile oli 2.78 ja raskuse hinnangu keskmine oli 4.13.

Valimis oli rohkem naisi (TMB = 292, MaRs-IB = 254) kui mehi ning üks isik valis enda sooks muu. Katses osalesid isikud vanusevahemikus 20 – 69 ($M = 35.9$, mediaan = 35, $SD = 12.2$), kelle läbitud haridusaastad jäid vahemikku 10 – 19 aastat ($M = 15.5$, mediaan = 15). Valimis olid sagedasemad haridustasemed keskharidus ja kõrgharidus (Tabel 2).

Tabel 2

Haridustasemete jaotus maatriksite testidel.

	Haridusaastate arv	TMB		MaRs-IB	
		N	%	N	%
Põhiharidus	10	5	1.47	2	0.68
Keskharidus	13	118	34.60	103	35.27
Keskeriharidus	15	41	12.02	39	13.36
Rakenduslik kõrgharidus	19	44	12.90	36	12.33
Kõrgharidus	22	133	39.00	112	38.36

Katseisikud teostasid teste kõige enam sülearvutil ja nutitefonil (Tabel 3). Kõige vähem kasutati nuhvlit. Kolm kasutajat olid määranud kasutatud seadmeks muu ning küsimustiku vabateksti väljas täpsustasid, et teostasid teste puutetundliku ekraaniga sülearvutil.

Tabel 3

Katseisikute poolt kasutatud seadmete ülevaade.

	TMB		MaRs-IB	
	N	%	N	%
Sülearvuti	126	43.15	128	37.54
Nutitefon	112	38.36	135	39.59
Nuhvel	14	4.79	22	6.45
Lauaarvuti	40	13.70	53	15.54
Muu	0	0.00	3	0.88

MaRS-IB läbimiseks kulus katseisikutel keskmiselt 17.9 minutit (väärtuste vahemik = 2 – 158, mediaan = 16.2) ja TMB puhul 9.0 minutit (väärtuste vahemik = 2 – 145, mediaan = 7.7). Kahe testimise korra vahe päevades oli keskmiselt 11.6 (väärtuste vahemik = 0.8 – 93.9, mediaan = 9.18).

Lühendatud akadeemilise testi ja maatriksi testi läbinud katseisikute arv oli 88 (MaRS-IB = 70, TMB = 75).

Hüpoteesid 1-5: Testmybrain maatriksid

Käesoleva peatükk võrdles eelnevates teadustöodes esinenud olulisi seaduspärasid käesoleva uurimistöökäigus saadud andmetega.

TMB andmebaasi maatriksite sisereliaablus (hüpotees 1) oli kõrge (Cronbach'i $\alpha = .75$, KR20 = .75, Spearman-Brown = .75, McDonald'i Omega = .75). Spearman-Brown'i väärtus oli sama kümne ja kolme järjestikuse eksimusega.

TMB maatriksite puhul esines nõrk negatiivne seos vanusega ja vanuse ruuduga (Tabel 4). Üldskooril puudus statistiliselt oluline seos haridusaastatega ka koos alla 25-aastaste grupiga (Spearmani $\rho = 0$, $p = .93$).

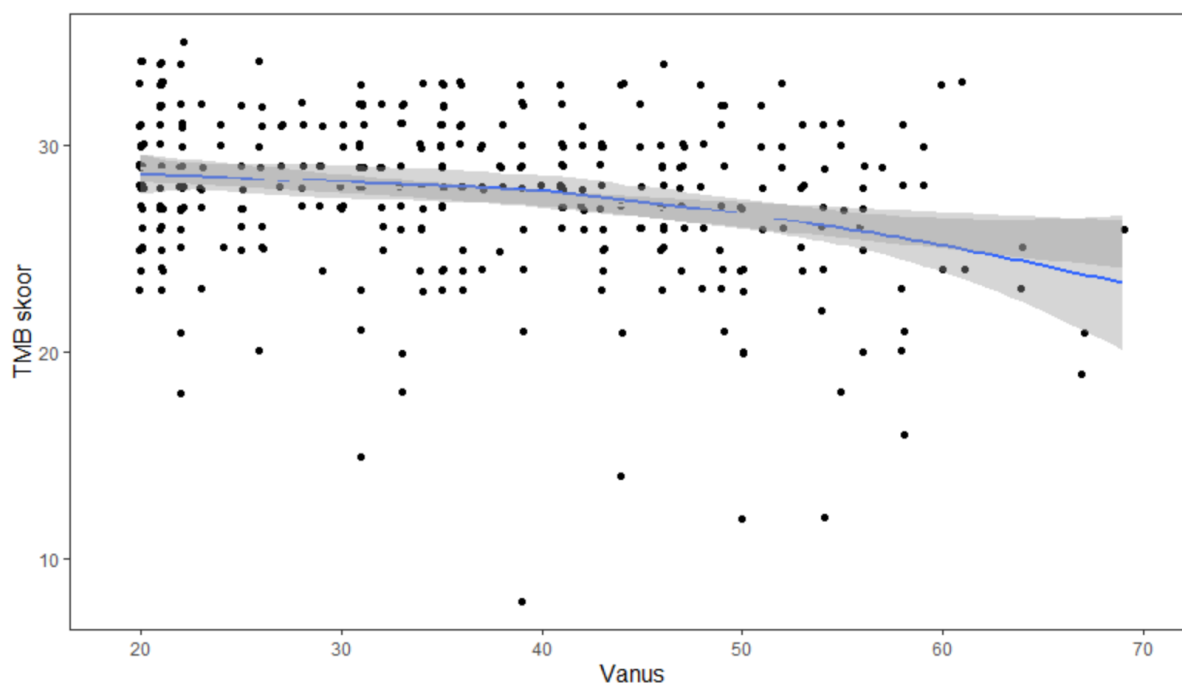
Tabel 4.

MaRs-IB ja TMB üldskooride korrelatsioonimaatriks kasutades Spearmani korrelatsioonikordajat.

	TMB	MaRs-IB	Haridusaastate arv	Ekraani laius	Vanus	Vanuse ruut
TMB	-					
MaRs-IB	.50**	-				
Haridusaastate arv	.06	.12	-			
Ekraani laius	-.09	.03	-.04	-		
Vanus	-.20***	-.03	.10	-.17**	-	
Vanuse ruut	-.20***	-.03	.10	-.17**	1.0	-

Märkus: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Vanuse lineaarse regressiooni mudel ennustas 5.4% skoori varieeruvusest. Mittelineaarse vanuse seose hindamiseks koostati hajuvusdiagramm, kus nähti viiteid mittelineaarsele seosele (joonis 5). Hajuvusdiagrammi andmete detailsemal analüüsil nähti vanuse tõusuga laienenud usalduspiire, kuid antud piirkonnas väikest katseisikute hulka arvestades ei pidanud autor antud seost tähenduslikuks.



Joonis 5. TMB hajuvusdiagramm koos regressioonijoonega.

Tabelis 5 on näha mitmesel regressioonimudelil (kohandatud $R^2 = .08$, $F(7, 242) = 4.17$, $p < .001$) statistilised olulised erinevused sugude vahel. TMB maatriksite regressioonimudeli referentsideks oli seadme grupil arvuti ja soo puhul mees. Regressioonimudelid eemaldati alla 25-aastased vastajad ehk 84 katseisikut.

Tabel 5

TMB skoori (sõltuv tunnus) kirjeldamine järgnevate sõltumatute tunnuste alusel: vanus, sugu, haridusaastate arv, seadmetüüp, ekraani laius.

	Hinnang	Standardviga	t	p
(vabaliige)	28.34	4.82	5.88	0
Sugu: Naine	-2.18	0.72	-3.03	0
Haridusaastate arv	0.1	0.09	1.14	0.26
Vanus	0.14	0.2	0.7	0.48
Vanuse ruut	0	0	-1.21	0.23
Nutitelefon	-1.03	0.8	-1.28	0.20
Nuhvel	-0.88	1.02	-0.86	0.39
Ekraani laius	0	0	-0.82	0.41

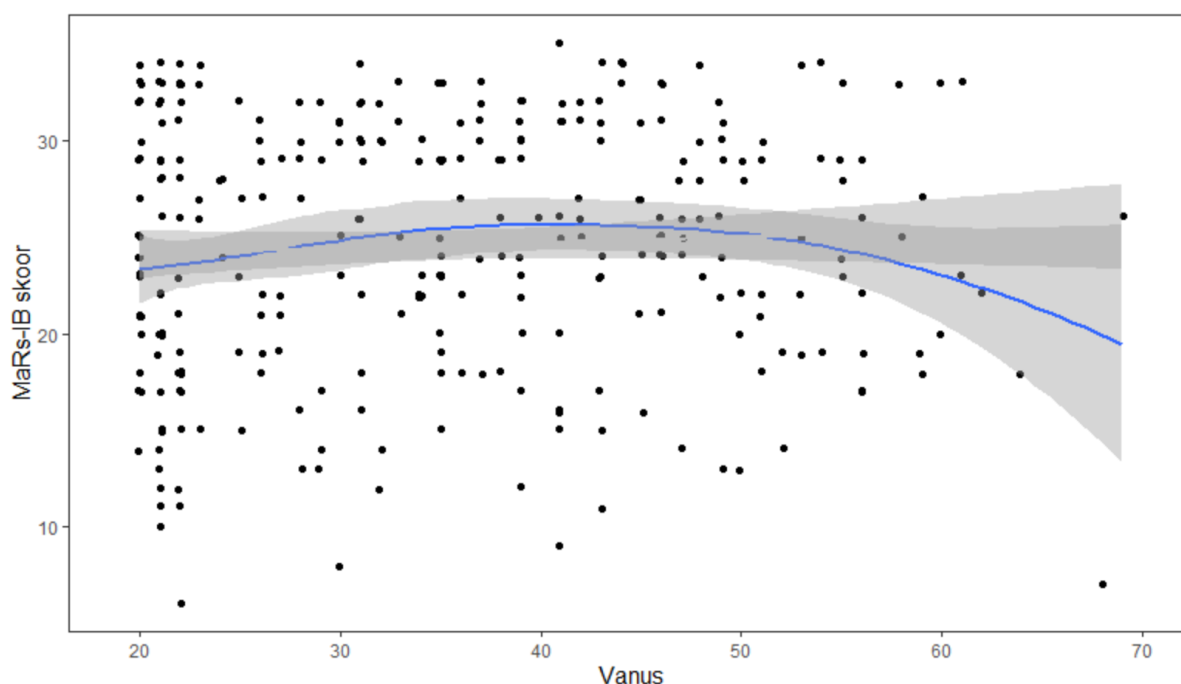
Viies hüpotees kirjeldas TMB skooride varieeruvuse erinevust naiste (väärtuste vahemik = 8 – 35, $M = 27.42$, $SD = 3.89$) ja meeste (väärtuste vahemik = 21 – 34, $M = 29.40$, $SD = 3.31$) vahel. Visuaalselt lähenes skoor normaaljaotusele ning Bartletti testi ($p = .16$) alusel olid rühmad homogeenised (Bartlett, 1937). LeveneTest alusel ei olnud gruppide dispersioonid ($F = .04$, $p = .85$) ehk rühmadispersioonid erinevad (Brown & Forsythe, 1974). Kontrollimaks kirjeldavate statistikute ülevaadet, teostati variatiivsuse statistiline test, milleks kasutati F-testi ($F = .72$, $p = .18$), mis kinnitas statistiliselt olulise sugudevahelise variatiivsuse puudumise.

Hüpoteesid 1-5: MaRS-IB maatriksid

Käesolev peatükk grupeerib MaRS-IB maatriksite hüpoteese, mille eesmärk oli võrrelda eelnevas kirjanduses leitud seaduspärasid käesoleva uurimistöö käigus saadud andmetega.

MaRS-IB andmebaasi maatriksite sisereliaablus (hüpotees 2) oli kõrge (Cronbach'i $\alpha = .88$, KR20 = .88, Spearman-Brown = .88, McDonald'i Omega = .87).

MaRS-IB maatriksite korrelatsioonimaatriksis puudusid statistiliselt olulised seosed üldskoori ning uuritavate tunnuste vahel (Tabel 4). Haridusaastatel oli väga nõrk positiivne seos üldskooriga (Spearmani $\rho = .16$, $p = .05$), kui alla 25-aastaste skoorid olid kaasatud. Vanuse lineaarse regressiooni mudel ei ennustanud skoori varieeruvust ($p = .35$). Vanuse mittelineaarse seose hindamiseks koostati hajuvusdiagramm, kus oli näha viiteid mittelineaarsele seosele (Joonis 6). Hajuvusdiagrammi andmete detailsemal analüüsil oli näha vanuse tõusuga laienenud usalduspiirid, kuid antud piirkonnas väikest katseisikute hulka arvestades ei pea autor antud seost tähenduslikuks.



Joonis 6. MaRs-IB hajuvusdiagramm koos regressioonijoonega.

Tabelis 6 on näha mitmese regressiooni mudelil (kohandatud $R^2 = .05$, $F(7, 199) = 2.61$, $p = .01$) olid statistilised olulised erinevused sugude vahel, haridusaastatel ja vanuse ruudul. MaRs-IB maatriksite mitmese regressiooni mudeli seadmetüübi referents oli arvuti ja soo puhul mees. Regressioonimudelist eemaldati alla 25-aastased vastajad ehk 80 katseisikut.

Tabel 6

MaRs-IB skoori (sõltuv tunnus) kirjeldamine järgnevate sõltumatute tunnuste alusel: vanus, sugu, haridusaastate arv, seadmetüüp, ekraani laius.

	hinnang	standardviga	t	p
(vabaliige)	13.64	8	1.7	0.09
Sugu: Naine	-2.78	1.26	-2.19	0.03
Haridusaastate arv	0.35	0.15	2.29	0.02
Vanus	0.68	0.34	1.98	0.05
Vanuse ruut	-0.01	0	-2.06	0.04
Nutitelefon	-2.45	1.35	-1.82	0.07
Nuhvel	-3.68	2.07	-1.78	0.08
Ekraani laius	0	0	-1.99	0.05

Hüpotees 5 väitis, et MaRs-IB maatriksite skooride variatiivsus oli meestel suurem (väärtuste vahemik = 6 – 35, $M = 27.32$, $SD = 6.57$) kui naistel (väärtuste vahemik = 7 – 34, $M = 24.17$, $SD = 6.45$). Visuaalselt läheneb skoor normaaljaotusele ning Bartletti testi alusel olid andmed homogeensed ($p = .88$) (Bartlett, 1937). LeveneTest alusel ei olnud gruppide dispersioonid ($F = .43$, $p = .51$) ehk rühmadispersioonid erinevad (Brown & Forsythe, 1974). Kontrollimaks kirjeldavate statistikute poolt kirjeldatud varieeruvuse puudumist, teostati F-test ($F = 1.04$, $p = .83$), mis kinnitas statistiliselt olulise sugudevahelise variatiivsuse puudumise.

Hüpotees 6: Struktuuri hindamine

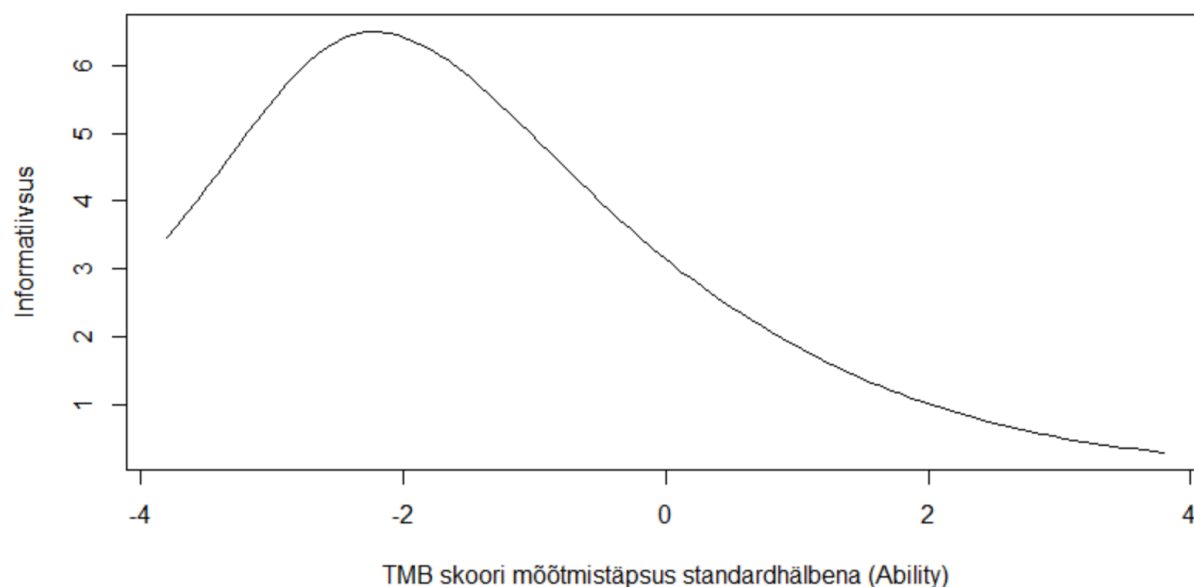
Struktuurvõrrandite mudeli hindamise eesmärk oli koostada mudel ning hinnata sobivust andmetega ehk kui hästi kirjeldas koostatud mudel käesolevaid andmeid.

TMB maatriksite mudeli hii-ruut statistiku 311.58 ($p = .06$) alusel ei erinenud kovariatsioonimaatriks andmestiku põhjal arvutatud kovariatsioonimaatriksist. Seega teised statistikud ei ole olulised, kuid on välja toodud taustainfoks. Mudeli sobivusastet näitasid statistikud CFI oli 0.93 ja TLI 0.92, mis näitasid piiripealset sobitusastet. Absoluutse sobitusastme indeksitest RMSEA oli 0.02 ja SRMR 0.11.

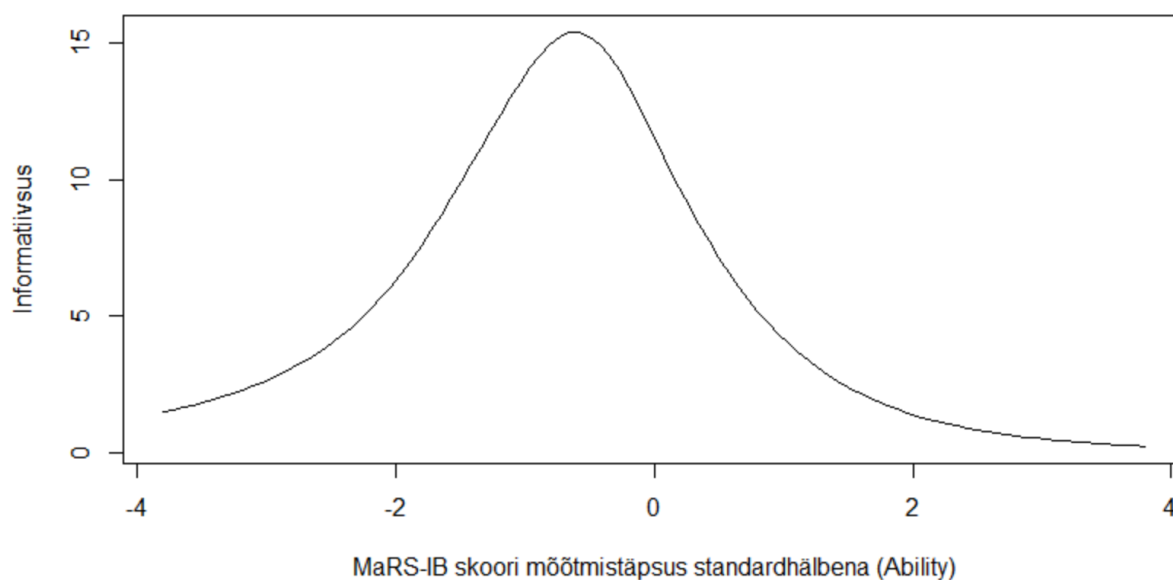
MaRs-IB maatriksite mudeli hii-ruut statistiku 561.48 ($p = .02$) alusel erines kovariatsioonimaatriks andmestiku ($n = 287$) põhjal arvutatud kovariatsioonimaatriksist. Mudeli sobivusastet näitasid statistikud CFI oli 0.98 ja TLI 0.98, mis näitasid head sobitusastet. Absoluutse sobitusastme indeksitest RMSEA oli 0.02 ja SRMR, mille tulemus võiks olla väiksem kui 0.08 oli aga 0.10. Kuna SRMR üksinda ei mõjuta otsust, siis hinnati käesolevad andmed MaRs-IB mudelisse sobivaks.

Testide hindamiseks kasutati ka üksikvastuse teooriat ehk *item response theory* (IRT) dihhotoomsete andmetega. Inimese vastamismuster sõltus nii tema enda võimekusest kui ülesannete raskusastmest. Koostati ühe faktoriga 2PL mudeli, mis kuvas maatriksite raskusaste ning eristatavus, mida on võimalik näha lisas (Tabel B1).

Joonisel 7 on näha, et TMB maatriksi testide puhul oli mõõtmistäpsus kõige suurem -2 standardhälbe juures, seega eristas see hästi madalaid skooore, aga ei erista hästi keskmiste või kõrgete skooridega isikuid. MaRs-IB puhul oli mõõtmistäpsus kõige suurem keskmiste skooride juures, aga test ei eristanud nii hästi madalaid ja kõrgeid skooore (Joonis 8).



Joonis 7. TMB testi mõõtmistäpsuse graafik



Joonis 8. MaRS-IB testi mõõtmistäpsuse graafik

Testi mõõtmistäpsuse graafikute ja 2PL mudeli analüüsi (Tabel B1) alusel selgus, et küsimused olid üldiselt lihtsad, seega teostati ka 3PL mudeli analüüs välistamaks küsimuste lihtsuse põhjuseks katseisikute poolt juhuslikku õigete vastuste ära arvamist. Välistamaks ära arvamise mõju mudelitele, teostati Anova abil 3PL ja 2PL mudelite võrdlemine. Analüüsis

selgus, et mudelit ei mõjutanud katseisiku poolt juhuslik vastuse valimine TMB ($p = .99$) kui ka MaRs-IB ($p = .35$) maatriksite puhul.

Viimasena teostati katseisikute faktorskooride korrelatsiooni analüüs isikute skooriga. TMB faktorskooride ning katseisiku skoori vahel oli tugev positiivne korrelatsioon (Spearmani $\rho = .95$, $p < .001$). MaRs-IB faktorskooride ning katseisiku skoori vahel oli tugev positiivne korrelatsioon (Spearmani $\rho = .99$, $p < .001$).

Hüpoteesid 7-9: Testidevahelised seosed

Käesolev peatükk analüüsis tulemusi, mille eesmärk oli võrrelda MaRs-IB ja TMB maatriksite testide omavahelisi sarnasusi ehk kas need mõõtsid vaimset võimekust sarnaselt.

MaRs-IB ja TMB testide seosed

MaRs-IB andmebaasi ja TMB keskkonna maatriksite skooridel oli keskmise tugevusega positiivne seos ($r = .50$, $p < .001$). TMB skoor selgitas MaRs-IB skoori varieeruvusest 23%.

Numbri-sümboli vastavuse test

TMB maatriksite puhul oli nõrk positiivne seos numbri-sümboli vastavuse testi tulemusega (Spearmani $\rho = .25$, $p < .001$). Passelli 2019 uuringus oli aga keskmise tugevusega positiivne seos numbri-sümboli vastavuse testi tulemusega (Spearmani $\rho = .38$, $n = 1210$, 95% CI [0.33, 0.43]). Käesoleva uuringu ja Passelli 2019 uuringu korrelatsioonide analüüsil olid antud grupid statistiliselt oluliselt erinevad ($z = 2.35$, $p = 0.02$, 95% CI [0.02, 0.24]).

MaRs-IB maatriksite puhul oli väga nõrk positiivne seos numbri-sümboli vastavuse testi tulemusega (Spearmani $\rho = .13$, $p = .04$).

MaRs-IB ja TMB üldskooride ning number-sümbol vastavuse testi korrelatsioonide analüüsi tulemusel ei olnud käesolevad grupid statistiliselt oluliselt erinevad ($z = 1.55$, $p = 0.12$).

Lineaarse regressioonimudeli alusel ennustas number-sümbol vastavuse test MaRs-IB skoori varieeruvust 1,6% ning TMB puhul 7.6%.

Viidi läbi mitmene lineaarne regressioonanalüüs, mille eesmärk oli leida number-sümbol vastavuse testi skoori seos maatriksi skoori ning seadmetüübiga (referents on arvuti), sooga (referents on mees) ja vanusega (Tabel 7). MaRs-IB maatriksite regressioonimudeli statistikud (kohandatud $R^2 = .3$, $F(6, 272) = 28.79$, $p < .001$) ja TMB mudeli omadused (kohandatud R^2

= .3, $F(7, 324) = 21.01$, $p < .001$) kirjeldasid statistiliselt olulisi mudeleid. Mõlema mudeli puhul oli näha, et number-sümbol vastavuse testi skoor oli negatiivselt seotud vanusega ning positiivselt maatriksite skooridega.

Tabel 7

Number-sümbol vastuse skoori (sõltuv tunnus) kirjeldamine järgnevate sõltumatute tunnuste alusel: maatriksi skoor, sugu, vanus, seadmetüüp, ekraani laius.

	hinnang	standardviga	t	p
MaRs-IB mudel				
(vabaliige)	55.05	2.40	22.98	0
MaRs-IB skoor	0.17	0.05	3.33	0
Sugu: Naine	-0.79	1.03	-0.77	0.44
Vanus	-0.34	0.03	-12.19	0
Nutitelefoni	1.74	1.30	1.34	0.18
Nuhvel	-0.22	1.56	0.44	0.89
Ekraani laius	0	0	-0.33	0.75
TMB mudel				
(vabaliige)	50.17	3.58	13.99	0
TMB skoor	0.32	0.09	3.24	0
Sugu: Naine	-0.70	1.02	-0.59	0.56
Vanus	-0.31	0.03	-10.18	0
Nutitelefoni	-0.81	0.97	-0.83	0.40
Nuhvel	1.05	1.41	0.74	0.46
Ekraani laius	0	0	1.10	0.27

Lühendatud akadeemiline test

Lühendatud akadeemilisel testi üldskooril, matemaatika alamskooril, ruumilise mõtlemise alamskooril oli statistiliselt oluline nõrk positiivne seos MaRs-IB testiga ning keskmise tugevusega seos TMB testiga (Tabel 8).

Tabel 8

Vaimse võimekuse üldskoori ja alamskooride seosed maatriksite üldskooriga

	Vaimse võimekuse üldskoor	Matemaatika alamskoor	Ruumilise mõtlemise alamskoor	Sõnavara alamskoor
	Spearmani ρ	Spearmani ρ	Spearmani ρ	Spearmani ρ
MaRs-IB	.34 **	.37 **	.29 *	.12
TMB	.54 ***	.56 ***	.46 ***	.60

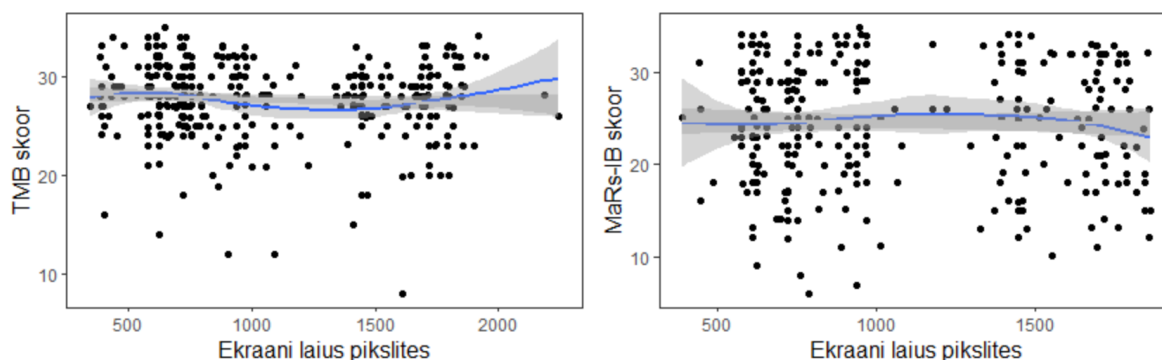
Märkus: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MaRS-IB ja TMB testi ning matemaatika alamskooril korrelatsiooni analüüsil selgus gruppide vaheline statistiliselt oluline erinevus ($z = 2.18$, $p = .03$). Piiripealselt statistiliselt oluline erinevus esines gruppide vahel ka ruumilise mõtlemise ($z = 1.97$, $p = .049$) alamskooril.

Hüpoteesid 10-11: Skoori mõjutatavus uurimisvahenditest

Peatüki eesmärk oli kirjeldada katseisiku poolt kasutatud seadmete - arvuti, nutitelefoniga ja nuhvli efekti skoorile. Hüpoteesides eeldati ekraani laiuse (hüpotees 10) seoseid arvuti, nuhvli ja nutitelefoni. Lisaks eeldas hüpotees 11 lauarvutil teostatud üldskooride väiksemat variatiivsust.

Hüpoteesis 10 kirjeldatud alla 500 piksli laiuste ekraanidega katseisikute arv oli MaRs-IB viis ja TMB puhul 23. Võrreldes väikeste ekraanide üldskooride üle 500 piksli laia ekraaniga ei leitud statistiliselt olulisi erinevusi TMB (Spearmani $\rho = .15$, $p = .49$) ja MaRs-IB (Spearmani $\rho = -.46$, $p = .43$) maatriksitel. Seega teostati analüüs alla 1200 piksli piiriga, aga ka seal puudusid MaRs-IB (Spearmani $\rho = .08$, $p = .26$) ning TMB (Spearmani $\rho = -.11$, $p = .11$) puhul statistiliselt olulised erinevused. Ekraani laiuse lineaarse regressiooni mudel ei ennustanud MaRs-IB ($p = .89$) ning TMB ($p = .19$) skoori varieeruvust. Kontrolliti ka mittelineaarsete seoste olemasolu kasutades hajuvusdiagramme, kus seosed esile ei tulnud (Joonis 9).



Joonis 9. TMB ja MaRS-IB hajuvusdiagrammid koos regressioonijoonega.

Välistamiseks vale parameetri hindamist, sai teostatud ekraani pindala võrdlemine üldskooriga. Ekraani pindala lineaarse regressiooni mudel ei ennustanud MaRS-IB ($p = .70$) ning TMB ($p = .11$) skoori varieeruvust.

Hüpotees 11 väidab, et maatriksite skooride variatiivsus oli lauarvutil suurem kui ülejäänud seadmetel (nutitelefon, nuhvel, sülearvuti). MaRS-IB skooride kirjeldavad statistikud oli lauarvutil (vahemik = 6 – 35, SD = 6.46) ja teistel seadmetel (vahemik = 10 – 34, SD = 6.66). Visuaalselt ei ole skooride puhul tegemist normaaljaotusega ning Shapiro-Wilk testi alusel ei ole lauarvuti ($W = .96$, $p < .001$) ega teiste seadmete ($W = .94$, $p < .001$) andmed normaaljaotuslikud. Kontrollimaks kirjeldavate statistikute poolt kirjeldatud varieeruvuse puudumist, teostati sõltumatute gruppide Mann–Whitney U test ($W = 10867$, $p = .57$), mis kinnitas statistiliselt olulise gruppide vahelise variatiivsuse puudumise.

TMB skoori kirjeldavad statistikud lauarvutil (vahemik = 12 – 35, SD = 3.86) ja muudel seadmetel (vahemik = 8 – 34, SD = 3.87). Visuaalselt lähenesid skoorid normaaljaotusele aga Shapiro-Wilk testi alusel ei ole lauarvuti ($W = .93$, $p < .001$) ega teiste seadmete ($W = .92$, $p < .001$) andmed normaaljaotuslikud. Kontrollimaks kirjeldavate statistikute poolt kirjeldatud varieeruvuse puudumist, teostati sõltumatute gruppide Mann–Whitney U test ($W = 15782$, $p = .15$), mis kinnitas statistiliselt olulise gruppide vahelise variatiivsuse puudumise.

TMB ($t(339) = 1.34$, $p = .18$) ja MaRS-IB ($t(290) = .57$, $p = .57$) testi skoorid ei olnud T-testi alusel statistiliselt oluliselt erinevad lauarvutil võrrelduna teiste gruppidega.

Hüpotees 12-14: Piloteeritavad hüpoteesid

Pilootprojekti üheks eesmärgiks oli teha ettepanekuid testi kasutamiseks järgmiseks faasis. Uurimistöö autor valis parendusettepanekute jaoks subjektiivse meeldivuse analüüsi, mille alusel eelistada ühte maatriksite ülesannet teisest. Lisaks uuriti võimalust testi lühendada järjestike valede vastuste korral.

Subjektiivne meeldivus

Katseisikud said teste hinnata Likerti skaalal meeldivaks (1 – ebameeldiv ja 5 – meeldiv). MaRs-IB keskmine meeldivus oli 2.78 ja TMB puhul 3.61, mis olid võrreldes samal ajal läbitud teiste ülesannete keskmisega (3.93) võrreldes madalamad (hüpotees 12.1-6). T-testi alusel oli tegemist statistiliselt oluliste erinevusega MaRs-IB ($t(582) = 12.7, p < .001$) ja TMB maatriksitel ($t(680) = 4.5, p < .001$).

T-testi alusel oli MaRs-IB ja TMB omavaheliste keskmiste erinevus statistiliselt oluline ($t(631) = 8.2, p < .001$).

Tabelis 9 on toodud subjektiivse meeldivuse keskmised väärtused vastavalt maatriksi ja seadme tüübile. Katseisiku TMB testi subjektiivse meeldivuse hinnang ennustas 11.0% teiste testpatareide keskmisest subjektiivset meeldivusest (kõik teised testid välja arvatud maatriksid) ning MaRS-IB puhul oli see ennustusväärtus 5.2%.

Tabel 9
Testi meeldivuse keskmised

		n	Meeldivuse keskmine (M)
MaRs-IB	Arvuti	166	2.76
	Nuhvel	14	2.86
	Nutitelefon	112	2.79
TMB	Arvuti	181	3.62
	Nuhvel	25	3.63
	Nutitelefon	135	3.36

Järjestikeksimuste arvu mõju üldskoorile

Pilootuuringu käigus salvestati ka katseisiku üldskoor, kui eksitakse järjest kolm, viis või seitse korda (hüpoteesid 13.1-6 ja 14.1-6). See informatsioon võimaldas hinnata võimalust

lühendada järgmises faasis katseisikutele testile kulunud aega. TMB maatriksite puhul kolme, viie või seitsme järjestiku eksimusega skoor oli Spearmani astakorrelatsiooni kasutades tugevas positiivses seoses kümne eksimuse skooriga (Tabel 10). MaRs-IB maatriksite puhul oli tegemist ka tugeva positiivse skooriga, aga korrelatsioonid on madalamad kui TMB testil.

Tabel 10

Testi järjestikuste valede vastuste arvu alusel katkestamise mõju skoorile.

		Kolm järjest eksimust	Viis järjest eksimust	Seitse järjest eksimust
		Spearmani ρ	Spearmani ρ	Spearmani ρ
MaRs-IB	Arvuti	.88 ***	.96 ***	.99 ***
	Nuhvel	.98 ***	.97 ***	.97 ***
	Nutitelefon	.90	.97 ***	.99 ***
TMB	Arvuti	.97 ***	.99 ***	.997 ***
	Nuhvel	.97 ***	.99 ***	.99 ***
	Nutitelefon	.97 ***	.99 ***	.998 ***

Märkus: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TMB testi puhul katseisiku kolme järjestikku vale vastust lõpetatud maatriksite skoor ennustas 93.9% üldskoorist, kui test lõpetatakse kümne järjestikuse vale vastuse korral, viie eksimuse puhul 98.1% skoorist ning seitsme eksimuse korral 99.5% skoorist. MaRs-IB maatriksite testi kolme eksimuse skoor ennustas 79.9% varieeruvusest, viie eksimuse puhul 92.2% ning seitsme eksimuse korral 97.6%.

Arutelu

Käesoleva uurimistö eesmärk oli pilootuuringu abil valideerida eestikeelset maatriksitesti digitaalseks kasutamiseks. Pilootprojekti käigus testiti kahest erinevast allikast maatrikseid - TMB ja MaRs-IB.

Uurimisküsimused olid jaotatud viide rühma – käesolevate andmete analüüsi tulemused võrreldes eelnevate uuringutega, struktuuri hindamine, testide omavaheline võrdlus, seadme mõju üldskoorile ja hüpoteesid muudatusettepanekuteks. Esimese rühma eesmärk oli loetud kirjanduse alusel loodud hüpoteeside korratavuse hindamine ehk kas adapteeritud testil ning antud valimil on võimalik tulemusi korrata - näiteks skooride seos haridusaastate, vanusega ning sisereleiaabluse tulemus. Teise rühma eesmärk oli hinnata, kas MaRS-IB ja TMB

maatriksite struktuurvõrrandite mudelid kirjeldasid tegelikke andmeid. Kolmanda grupi eesmärk oli võrrelda teste omavahel ning uurida, kas mõlemad testid mõõdavad vaimset võimekust sarnaselt. Neljandas grupis hinnati seadme tüübi ning ekraani laiuse mõju maatriksite üldskoorile, mille eesmärk oli hinnata digitaalseid piiranguid. Viimasena oli tuuakse ettepanekud testipatarei rakendamise osas, näiteks testi lühendamine järjestiku eksimuste alusel.

Valim

Käesoleva uuringu valimi puhul oli oluline tähelepanek, et vastajate seas oli mitu korda rohkem naisi (MaRs-IB = 254, TMB = 292) kui mehi (MaRs-IB = 38, TMB = 48) ning nuhvlil teostatud katsete arv oli madal (MaRs-IB = 22, TMB = 14). Seega tegemist ei ole representatiivse valimiga.

Analüüsi käigus selgus, et kolm katseisikut lahendasid ainult MaRs-IB maatrikseid ning 52 katseisikut katkestasid MaRs-IB maatriksite testi. Poolelijätmine ei olnud seotud vanuse, soo ega haridustasemega. Ka katseisikute subjektiivne meeldivus ei põhjendanud poolelijätmist, aga loobunud kasutajate hinnang testi keskmisele raskusele võis selgitada MaRS-IB testi katkestamist, kuna testi hinnati poolelijättnute poolt keerulisemaks.

Hüpoteesid 1-5: Võrdlused eelnevate uuringutega

MaRs-IB maatriksite sisemine reliaablus oli kõrge ja käesolevas uuringus korrati Chierchia (2019) uuringu tulemusi. TMB maatriksite sisemine reliaablus jäi Passelli (2019) uuringust madalamaks, aga oli piisavalt sarnane ning tegemist oli ikkagi kõrge sisereliaablusega. Seega mõlema testi üksikülesannete sisemine kooskõla oli kõrge ja antud testi üksikülesanded mõõdavad sama omadust. Kuna tulemused olid piisavalt sarnased eelnevate uuringutega, siis võib eeldada, et adapteerimise käigus ei kaotanud testid sisemist usaldusväarsust ning neid teste võib edaspidi kasutada.

Teaduskirjanduse alusel tekkis eeldus, et vaimne võimekus on seotud ka haridusaastate arvuga ehk kõrgema vaimse võimekusega isikud saavutavad ka kõrgema haridustaseme. Kuna maatriksid mõõdavad vaimset võimekust, siis saaks eeldada üldskoori ning haridusaastate arvu vahel seost. Haridusaastatega seotud hüpoteeside jaoks esindusliku valimi saavutamiseks jäeti välja alla 25-aastased katseisikud, kuna nende vastamise hetkel

saavutatud haridustase ei pruugi olla maksimaalne. Hüpootees ei leidnud kinnitust kummagi maatriksi puhul. Lisaks ei suutnud käesolev uuring korrata Passelli (2019) uuringute tulemusi. Põhjuseks võib olla valimi mittepiisav representatiivsus, näiteks mitu korda kõrgem naiste osakaal.

Passelli (2019) uuringus tuli esile inimese arenguperioodil skoori kiire tõus ning alates 30-ndast eluaastast aeglane langus. Käesolevas uuringus esines TMB maatriksite puhul nõrk negatiivne seos vanusega ja nõrk negatiivne seos vanuse ruuduga. Seega nooremad saavutasid TMB testil kõrgemaid üldskooore, mis võis olla tingitud suuremast enesekindlusest digitaalsete seadmete kasutamisel. MaRs-IB puhul puudus statistiliselt oluline seos vanuse ja vanuse ruuduga. Seega käesolevate andmetega ei suudetud korrata eelnevate uuringute efekti. Esile ei tulnud ka selget mittelineaarset seost vanuse ning üldskoori vahel. Selline tulemus võib olla põhjustatud mitterepresentatiivsest valimist.

Meeste vaimset võimekust hindavad skoorid varieeruvad rohkem kui naistel (Johnson jt, 2008). Käesolevas uuringus ei olnud MaRs-IB ja TMB üldskoori varieeruvus statistiliselt oluliselt erinev. Selle hüpooteesi kordamise ebaõnnestumine võib olla põhjustatud vähesest meeste osakaalust valimis. TMB ja MaRs-IB uuringutes ei leitud meeste ja naiste skooride vahel statistiliselt olulisi erinevusi (Passell jt, 2019; Chierchia jt, 2019). Käesolevas uuringus oli mõlema testi puhul statistiliselt oluline erinevus sugude vahel, mis erineb eelnevatest uuringutest. Põhjuseks võib olla kallutatud valim ning erineva vaimse võimekusega meeste puudumine valimist.

Hüpootees 6: Struktuuri hindamine

TMB maatriksite mudel ei erinenud kovariatsioonimaatriks andmestiku põhjal arvatud kovariatsioonimaatriksist. Seega TMB struktuurimudel ennustas hästi käesolevaid andmeid ja ühefaktoriline mudel on sobiv.

MaRs-IB maatriksite mudeli erines kovariatsioonimaatriks andmestiku põhjal arvatud kovariatsioonimaatriksist. Mudeli sobivusastmed näitasid piisavalt head sobitusastet. Järelikult võib öelda, et MaRs-IB mudel suutis piisavalt hästi kirjeldada tegelikke andmeid.

TMB maatriksite test eristas hästi madalaid skooore, aga ei eristanud hästi keskmiste või kõrgete skooridega isikuid. MaRs-IB test ei eristanud nii hästi madalaid ja kõrgeid, aga eristas hästi keskmisi skooore. Seega olid mõlemad testid suhteliselt lihtsad ning neil ei

esinenud lae- ega pörandaepekti. Analüüsi käigus kinnitati, et juhuslik vastamine ei mänginud rolli mudeli seletamisel. Seega sobivad mõlemad testid edaspidiseks kasutamiseks.

Katseisikute üldskoorid olid väga tugevalt seotud isiku enda faktorskooriga ning seega üldskoori tulemused on usaldusväärsed ning sobivad edaspidiseks analüüsiks.

Hüpoteesid 7-9: Testidevahelised seosed

MaRs-IB andmebaasi ja TMB keskkonna maatriksite puhul eeldati, et mõlemad mõõdavad kognitiivset võimekust sarnasel viisil ehk testide üldskooride vahel oli tugev positiivne seos. Analüüsi käigus selgus hoopis keskmine positiivne seos ($r = .50$, $p < .01$) ning TMB testi skoor selgitas MaRs-IB skoori varieeruvusest 23%. Nende andmete põhjal tekib kahtlusi, et maatriksite testid hindavad vaimset võimekust erinevalt.

TMB keskkonna testikomplektis oli lisaks maatriksitele ka numbri-sümboli vastavuse testi. Passelli (2019) uuringus oli nende testide vahel nõrk positiivne korrelatsioon ($r = .38$). Käesolevas uuringus korrati üldskooril nõrka positiivset seost TMB maatriksitega ($r = .25$, $p < .001$), aga MaRs-IB puhul oli väga nõrk positiivne seos ($r = .13$, $p = .04$). MaRs-IB ja TMB korrelatsioonide analüüsi alusel ei olnud käesolevad grupid statistiliselt oluliselt erinevad, seega seonduvad kaks testi sarnaselt numbri-sümboli vastavuse testiga.

Käesoleva uuringu TMB üldskoori ning numbri-sümboli vastavuse testi korrelatsiooni analüüsil Passeli 2019 uuringu korrelatsiooniga selgus, et antud grupid olid statistiliselt oluliselt erinevad. Gruppide vahelised erinevused võivad olla tingitud kallutatud valimist või adapteerimise käigus tekkinud vigadest. Adapteerimisel tekkinud vead olid ebatõenäolisemad, kuna protseduur ning tehniline lahendus on TMB keskkonnaga sama ning ainsaks erinevuseks oli keelsus. Kuid antud juhul ei ole võimalik välistada ka adapteerimisega seotud probleeme, siis suurema tõenäosusega selgitab tulemusi kallutatud valim. Valimiga seotud probleemid on täpsemalt käsitletud valimi peatükis.

Lühendatud akadeemilise testi matemaatika ja ruumilise mõtlemise alaskaalal oli keskmise tugevusega seos TMB testiga aga nõrk seos MaRs-IB testiga. Akadeemilise testi võrdluses maatriksite testidega tekkinud erinevatest korrelatsioonide erinevust tõestas ka statistiliselt olulisid gruppide vahelisi erinevusi nii matemaatika kui ka ruumilise mõtlemise alaskaalal. Selle alusel tekib kahtlusi, kas maatriksite testid hindavad vaimset võimekust erinevalt.

Tulemused tekitavad küsimusi, kas ja kui korrektselt mõlemad testid vaimset võimekust mõõtsid ning miks sarnase ülesehitusega testide vahel olid nii madal omavaheline seos. Järgnevates uuringutes soovitab autor võrrelda mõlemat testi näiteks valideeritud WASI II maatriksite testiga. Lisaks antud tulemused suunavad testimise kordamist suuremal ning vähem kallutatud valimil.

Hüpotees 10-11: Seadme mõju üldskoorile

Digitaalse testi puhul sai katseisik ise valida endale “pliiatsi” ja “paberi”, aga tulemused peavad olema ikka võrreldavad. Seega oli üheks oluliseks osaks käesolevas töös hinnata uurimisvahendite mõju üldskoorile.

Uurimistöö hüpoteesides eeldati ekraani laiuse efekti maatriksi skoorile. Ekraani laius mõjutas kõige enam ekraanile kuvatud kujundite suurust. Analüüsi käigus ei leitud statistiliselt olulisi erinevusi alla 500 pikslise ekraani ning suurema ekraaniga gruppide vahel. Seega teostati analüüs alla 1200 piksli, aga ka seal puudusid MaRs-IB ning TMB testidel statistiliselt olulised erinevused. Teostati ka lineaarne regressioonanalüüs ja mittelineaarse seose hindamiseks hajuvusdiagramm, mis kinnitasid samuti mõju puudumist. Käesolevas uuringus ei hinnatud ekraani diagonaali suurust tollides, mis võimaldaks täpsemalt hinnata ekraani suuruse efekti. Hetkel selleks kasutati ekraani laiust pikslites, kuna see piirab kasutajale kuvatavate elementide suurust, aga see ei ole piisavalt täpne väärtus hindamaks kasutaja ekraanil kuvatud kujundite suurust.

Passelli (2019) uuringus leiti erinevus kaasaskantavate ning püüasukohaga seadmete vahel. Põhjuseks toodi see, et kaasaskantavate seadmete puhul ei ole ümbritsev keskkond nii toetav süvenemist vajavate ülesannete lahendamiseks. Käesolevas uuringus analüüsiti gruppidevaheliste skooride variatiivsust, mis võiks olla mõjutatud sellest, et püsival töökohal on vähem segavaid faktoreid ning katseisikul oli võimalik pikemalt keskenduda. Antud uuringu käigus ei leitud statsionaarse seadme ehk lauaarvuti ning teiste seadmete variatiivsuses statistiliselt olulisi erinevusi.

Kokkuvõtlikult ei leidnud käesolev uuring seadmetel statistiliselt olulist mõju katseisiku üldskoorile TMB ja MaRs-IB maatriksite testidel.

Hüpoteesid 12-14: Ettepanekud

Pilootuuringu oluline eesmärk oli teha muudatusettepanekuid jätku-uuringuks. Esiteks eelistati katseisikute subjektiivsel hinnangul TMB maatrikseid sõltumata kasutatavast seadmest. Kasutatavuse testides katseisikud kommenteerisid olulist erinevust testide raskusastmes. MaRs-IB maatriksid tundusid katseisikutele keerulisemad, kuna nad sisaldasid peenemaid, raskemini eristatavad elemente ning neid tajuti subjektiivselt keerulisematena. Nutitelefonidel MaRs-IB maatrikseid läbinud katseisikud tõid esile ka kontrasti ehk ligipääsetavuse probleeme. Autor soovib järgmises faasis rastergraafikas olevad illustratsiooni ümber teha vektorgraafikasse, mis võimaldab katseisikutel paremini eristada ülesandes sisalduvad elemente.

Testide puhul uuriti võimalust testi lõpetamiseks peale kolme, viit või seitset järjest eksimust. Analüüsi käigus selgus, et TMB testi puhul võiks testimise lõpetada peale viite vale vastust ($r = .99, p < .01$) ja MaRs-IB peale seitset vale vastust ($r = .99, p < .01$). Selline muudatus võiks vähendada testide poolelijätmist ja katseisiku ajakulu.

Katseisiku poolt testile kulutatud aeg oli varieeruv. Näiteks MaRS-IB ($M = 17.9$) puhul oli see vahemikus 2 – 158 minutit ja TMB ($M = 9.0$) puhul 2 – 145 minutit. Kuna testile kulunud aeg varieerub katseisikute puhul märgatavalt ning keskmisega võrreldes on maksimaalsed ajakulud MaRs-IB üle 8 korra suuremad ning TMB puhul üle 16 korra, siis võiks kaaluda järgmises etapis maksimaalse ajapiirangu kasutamist sarnaselt MaRS-IB uuringule.

Uurimissuunad

Antud uurimistöö tulemuste usaldusvääruse suurendamiseks tuleks see läbi viia representatiivsemal valimil. Näiteks inglise keelse TMB valideerimisel osales 20 510 inimest (Passell jt, 2019). Uuringu valim oli mõjutatud suurest naiste osakaalust, seega võiks järgnevates uuringutes üritada kaasata rohkem meeskateisikuid. Lisaks võiks proovida testipatarei asemel üksikute testide täitmist ning katseisikutele kohe tagasisidet pakkudes testide kordamist.

Käesolevas uuringus puudus võrdlus digitaalselt ning paberil täidetud maatriksitesti lahendamise vahel. Kinnitamaks digitaalse üldskoori kvaliteeti võiks järgnevalt teostada kontrolliks ka paberil lahendatavate maatriksite võrdlusanalüüsi. Näiteks oleks võimalik seda teha WASI II maatriksi testi abil.

Kokkuvõte

Autor soovib käesoleva uuringu alusel võtta Geenivaramu pilootuuringusse kasutusele TMB maatriksid. TMB maatriksite aluseks on võetud WASI II test ning eelnevad uuringud on olnud suurema valimiga (Passell jt, 2019; Chierchia jt, 2019). Käesolevas uuringus ei tulnud TMB maatriksite puhul esile lae- ega põrandaefekti. TMB maatriksite mudeli puhul ei erine kovariatsioonimaatriks andmestiku põhjal arvutatud kovariatsioonimaatriksist ja seega struktuurimudel ennustab hästi tegelikke andmeid. Sisereliaablus on TMB testil kõrge ja eelneva uuringuga sarnane. TMB testil ning akadeemilise testi matemaatika ning ruumilise mõtlemise alaskaaladel oli keskmisega tugevuse positiivne korrelatsioon. Akadeemilise testi ning maatriksite testide korrelatsiooni analüüsil leitud gruppidevahelised erinevused tekitavad kahtlusi, et MaRS-IB ja TMB mõõdavad vaimset võimekust erinevalt. Seadmetüüpidel ning ekraanilaiusel puudus TMB maatriksite üldskoorile statistiliselt oluline mõju. TMB testi puhul oli katseisikute subjektiivne arvamus positiivsem ning seega meeletatus testi teostada kõrgem. Näiteks loobus MaRS-IB maatriksite teostamisest 52 kasutajat, kes aga läbisid TMB maatriksid, seega on tulevikus ka katseisikute loobumine vähem tõenäoline TMB maatriksite puhul. Lisaks oli TMB testile katseisiku poolt kulutatud keskmine aeg lühem ehk 9.0 minutit võrreldes MaRS-IB 17.9 minutiga. Autoripoolne ettepanek oleks TMB test lõpetada viie järjestiku eksimuse olukorras ilma olulise mõjuta üldskoorile. Lisaks võiks kaaluda testile ajalimiidi lisamist.

Kasutatud kirjanduse loetelu

- Adams, M.-H. (2022). Elusündmuste seosed Testmybrain kognitiivsete testidega. *Uurimistöö*.
- Bartlett, M. S. (1937). Properties of Sufficiency and Statistical Test. *Proceedings of the Royal Society A*, 160, 268-282. <https://doi.org/10.1098/rspa.1937.0109>
- Berezin, A. (2022). Digitaalse punktide ühendamise testi, lihtreaktsioonaja testi ning valikreaktsioonaja testi valideerimine. *Uurimistöö*.
- Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346), 364–367. <https://doi.org/10.1080/01621459.1974.10482955>
- Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies. <https://doi.org/10.1017/CBO9780511571312>
- Castro, S. (2021, oktoober 17). Psychometrics. https://rstudio-pubs-static.s3.amazonaws.com/141954_4774223853064561a65bf2f9dc2256ae.html
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S. J. (2019). The matrix reasoning item bank (MaRs-IB): novel, open-access abstract reasoning items for adolescents and adults. *Royal Society open science*, 6(10), 190232. <https://doi.org/10.1098/rsos.190232>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. <https://doi.org/10.1007/BF02310555>
- D'Ardenne, K., Savage, C. R., Small, D., Vainik, U., & Stoeckel, L. E. (2020). Core Neuropsychological Measures for Obesity and Diabetes Trials: Initial Report. *Frontiers in psychology*, 11, 554127. <https://doi.org/10.3389/fpsyg.2020.554127>
- Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, 10(4): e0121945. [doi:10.1371/journal.pone.0121945](https://doi.org/10.1371/journal.pone.0121945)
- Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *British journal of psychology (London, England : 1953)*, 105(3), 399–412. <https://doi.org/10.1111/bjop.12046>

- Fox, J., Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- 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/1.1080/13854046.2018.1535662>
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Johnson, W., Carothers, A., & Deary, I. J. (2008). Sex Differences in Variability in General Intelligence: A New Look at the Old Question. *Perspectives on psychological science : a journal of the Association for Psychological Science*, *3*(6), 518–531. <https://doi.org/10.1111/j.1745-6924.2008.00096.x>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press..
- Meyerson, P., & Tryon, W. W. (2003). Validating internet research: a test of the psychometric equivalence of internet and in-person samples. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, *35*(4), 614–62. <https://doi.org/1.3758/bf03195541>
- Passell, E., Dillon, D. G., Baker, J. T., Vogel, S. C., Scheuer, L. S., Mirin, N. L., Germine, L. (2019). Digital Cognitive Assessment: Results from the TestMyBrain NIMH Research Domain Criteria (RDoC) Field Test Battery Report. <https://doi.org/10.31234/osf.io/dcszr>
- Pedersen, T. L. (2020). Patchwork: The Composer of Plots. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

- Revelle, W. (2021). psych: Procedures for Psychological, Psychometric, and Personality Research. *R package version 2.1.9*, <https://CRAN.R-project.org/package=psych>.
- Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response Analysis. *Journal of Statistical Software*, *17*(5), 1–25. <https://doi.org/10.18637/jss.v017.i05>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Ryan, J. J., Carruthers, C. A., Miller, L. J., Souheaver, G. T., Gontkovsky, S. T., & Zehr, M. D. (2005). The WASI matrix reasoning subtest: performance in traumatic brain injury, stroke, and dementia. *The International journal of neuroscience*, *115*(1), 129–136. <https://doi.org/10.1080/00207450490512704>
- Schoenberg, M. R., Lange, R. T., Iverson, G. L., Chelune, G. J., Scott, J. G., & Adams, R. L. (2006). Clinical validation of the General Ability Index--Estimate (GAI-E): estimating premorbid GAI. *The Clinical neuropsychologist*, *20*(3), 365–381. <https://doi.org/10.1080/138540491005866>
- Silm, G., Must, O., & Täht, K. (2013). Test-taking effort as a predictor of performance in low-stakes tests. *Trames. Journal of the Humanities and Social Sciences*, *17*(4), 433. <https://doi.org/10.3176/tr.2013.4.08>
- 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*, 1–10. Advance online publication. <https://doi.org/10.1080/13803395.2021.2002269>
- Tranel, D., Manzel, K., & Anderson, S. W. (2008). Is the prefrontal cortex important for fluid intelligence? A neuropsychological study using Matrix Reasoning. *The Clinical neuropsychologist*, *22*(2), 242–261. <https://doi.org/10.1080/13854040701218410>
- Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. <http://ggplot2.org>.
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: a tutorial review. *PeerJ*, *3*, e1058. <https://doi.org/10.7717/peerj.1058>

Lisa A

Tabel A1

MaRS-IB stiimulite nimekiri koos õigete vastuste ning seos TMB stiimulitega

MaRS-IB stiimul	Korrektne vastus	TMB stiimul	TMB pildi järjekord
tf1_4_M_ss3	Vale	1	1
tf1_4_T2_ss3_md	Vale	1-1	1
tf1_4_T1_ss3_md	Õige	1-2	1
tf1_4_T3_ss3_md	Vale	1-3	1
tf1_76_T3_ss3_pd	Vale	1-4	1
tf1_4_T4_ss3_md	Vale	1-5	1
tf1_1_M_ss3	Vale	2	2
tf1_1_T3_ss3_md	Vale	2-1	2
tf1_1_T2_ss3_pd	Vale	2-2	2
tf1_1_T1_ss3_md	Õige	2-3	2
tf1_1_T3_ss3_pd	Vale	2-4	2
tf1_1_T2_ss3_md	Vale	2-5	2
tf1_3_M_ss2	Vale	4	4
tf1_69_T3_ss2_md	Vale	4-1	4
tf1_3_T3_ss2_md	Vale	4-2	4
tf1_3_T2_ss2_md	Vale	4-3	4
tf1_3_T4_ss2_md	Vale	4-4	4
tf1_3_T1_ss2_md	Õige	4-5	4
tf1_13_M_ss3	Vale	15	15
tf1_13_T2_ss3_pd	Vale	15-1	15
tf1_13_T2_ss3_md	Vale	15-2	15
tf1_13_T1_ss3_md	Õige	15-3	15
tf1_13_T3_ss3_md	Vale	15-4	15
tf1_13_T3_ss3_pd	Vale	15-5	15
tf1_6_M_ss2	Vale	6	6
tf1_6_T3_ss2_pd	Vale	6-1	6
tf1_6_T1_ss2_md	Õige	6-2	6

tf1_6_T2_ss2_pd	Vale	6-3	6
tf1_6_T2_ss2_md	Vale	6-4	6
tf1_6_T3_ss2_md	Vale	6-5	6
tf1_2_M_ss1	Vale	3	3
tf1_2_T2_ss1_md	Vale	3-1	3
tf1_2_T4_ss1_md	Vale	3-2	3
tf1_17_T4_ss1_md	Vale	3-3	3
tf1_2_T1_ss1_md	Õige	3-4	3
tf1_2_T3_ss1_md	Vale	3-5	3
tf1_25_M_ss3	Vale	8	8
tf1_25_T1_ss3_md	Õige	8-1	8
tf1_25_T2_ss3_md	Vale	8-2	8
tf1_25_T3_ss3_pd	Vale	8-3	8
tf1_25_T3_ss3_md	Vale	8-4	8
tf1_25_T2_ss3_pd	Vale	8-5	8
tf1_20_M_ss1	Vale	7	7
tf1_20_T3_ss1_md	Vale	7-1	7
tf1_20_T3_ss1_pd	Vale	7-2	7
tf1_20_T2_ss1_md	Vale	7-3	7
tf1_20_T2_ss1_pd	Vale	7-4	7
tf1_20_T1_ss1_md	Õige	7-5	7
tf1_33_M_ss2	Vale	19	19
tf1_33_T3_ss2_pd	Vale	19-1	19
tf1_33_T1_ss2_md	Õige	19-2	19
tf1_33_T3_ss2_md	Vale	19-3	19
tf1_33_T2_ss2_md	Vale	19-4	19
tf1_33_T2_ss2_pd	Vale	19-5	19
tf1_12_M_ss2	Vale	18	18
tf1_12_T3_ss2_pd	Vale	18-1	18
tf1_12_T1_ss2_md	Õige	18-2	18
tf1_12_T2_ss2_pd	Vale	18-3	18
tf1_12_T3_ss2_md	Vale	18-4	18

tf1_12_T2_ss2_md	Vale	18-5	18
tf1_22_M_ss3	Vale	5	5
tf1_22_T3_ss3_pd	Vale	5-1	5
tf1_22_T2_ss3_md	Vale	5-2	5
tf1_22_T2_ss3_pd	Vale	5-3	5
tf1_22_T1_ss3_md	Õige	5-4	5
tf1_22_T3_ss3_md	Vale	5-5	5
tf1_17_M_ss1	Vale	9	9
tf1_17_T1_ss1_md	Õige	9-1	9
tf1_17_T3_ss1_pd	Vale	9-2	9
tf1_17_T2_ss1_pd	Vale	9-3	9
tf1_17_T2_ss1_md	Vale	9-4	9
tf1_17_T3_ss1_md	Vale	9-5	9
tf1_27_M_ss2	Vale	13	13
tf1_27_T3_ss2_pd	Vale	13-1	13
tf1_27_T1_ss2_md	Õige	13-2	13
tf1_27_T3_ss2_md	Vale	13-3	13
tf1_27_T2_ss2_pd	Vale	13-4	13
tf1_27_T2_ss2_md	Vale	13-5	13
tf1_36_M_ss2	Vale	21	21
tf1_36_T3_ss2_pd	Vale	21-1	21
tf1_36_T2_ss2_md	Vale	21-2	21
tf1_36_T3_ss2_md	Vale	21-3	21
tf1_36_T1_ss2_md	Õige	21-4	21
tf1_36_T2_ss2_pd	Vale	21-5	21
tf1_42_M_ss2	Vale	25	25
tf1_42_T2_ss2_md	Vale	25-1	25
tf1_42_T3_ss2_md	Vale	25-2	25
tf1_42_T2_ss2_pd	Vale	25-3	25
tf1_42_T3_ss2_pd	Vale	25-4	25
tf1_42_T1_ss2_md	Õige	25-5	25
tf1_10_M_ss3	Vale	10	10

tf1_10_T3_ss3_md	Vale	10-1	10
tf1_10_T2_ss3_md	Vale	10-2	10
tf1_10_T2_ss3_pd	Vale	10-3	10
tf1_10_T1_ss3_md	Õige	10-4	10
tf1_10_T3_ss3_pd	Vale	10-5	10
tf1_26_M_ss1	Vale	17	17
tf1_26_T1_ss1_md	Õige	17-1	17
tf1_26_T2_ss1_md	Vale	17-2	17
tf1_26_T3_ss1_md	Vale	17-3	17
tf1_26_T2_ss1_pd	Vale	17-4	17
tf1_26_T3_ss1_pd	Vale	17-5	17
tf1_21_M_ss2	Vale	12	12
tf1_21_T3_ss2_md	Vale	12-1	12
tf1_21_T3_ss2_pd	Vale	12-2	12
tf1_21_T2_ss2_md	Vale	12-3	12
tf1_21_T1_ss2_md	Õige	12-4	12
tf1_21_T2_ss2_pd	Vale	12-5	12
tf1_41_M_ss1	Vale	22	22
tf1_41_T3_ss1_md	Vale	22-1	22
tf1_41_T2_ss1_md	Vale	22-2	22
tf1_41_T1_ss1_md	Õige	22-3	22
tf1_41_T3_ss1_pd	Vale	22-4	22
tf1_41_T2_ss1_pd	Vale	22-5	22
tf1_29_M_ss1	Vale	14	14
tf1_29_T2_ss1_pd	Vale	14-1	14
tf1_29_T3_ss1_pd	Vale	14-2	14
tf1_29_T3_ss1_md	Vale	14-3	14
tf1_29_T2_ss1_md	Vale	14-4	14
tf1_29_T1_ss1_md	Õige	14-5	14
tf1_14_M_ss1	Vale	16	16
tf1_14_T1_ss1_md	Õige	16-1	16
tf1_14_T2_ss1_pd	Vale	16-2	16

tf1_14_T3_ss1_md	Vale	16-3	16
tf1_14_T2_ss1_md	Vale	16-4	16
tf1_14_T3_ss1_pd	Vale	16-5	16
tf1_44_M_ss1	Vale	26	26
tf1_44_T1_ss1_md	Õige	26-1	26
tf1_44_T2_ss1_md	Vale	26-2	26
tf1_44_T3_ss1_pd	Vale	26-3	26
tf1_44_T2_ss1_pd	Vale	26-4	26
tf1_44_T3_ss1_md	Vale	26-5	26
tf1_49_M_ss3	Vale	24	24
tf1_49_T3_ss3_md	Vale	24-1	24
tf1_49_T1_ss3_md	Õige	24-2	24
tf1_49_T3_ss3_pd	Vale	24-3	24
tf1_49_T2_ss3_pd	Vale	24-4	24
tf1_49_T2_ss3_md	Vale	24-5	24
tf1_60_M_ss2	Vale	29	29
tf1_60_T1_ss2_md	Õige	29-1	29
tf1_60_T2_ss2_pd	Vale	29-2	29
tf1_60_T2_ss2_md	Vale	29-3	29
tf1_60_T3_ss2_md	Vale	29-4	29
tf1_60_T3_ss2_pd	Vale	29-5	29
tf1_34_M_ss3	Vale	20	20
tf1_34_T1_ss3_md	Õige	20-1	20
tf1_34_T3_ss3_md	Vale	20-2	20
tf1_34_T3_ss3_pd	Vale	20-3	20
tf1_34_T2_ss3_md	Vale	20-4	20
tf1_34_T2_ss3_pd	Vale	20-5	20
tf1_52_M_ss3	Vale	30	30
tf1_52_T3_ss3_md	Vale	30-1	30
tf1_52_T2_ss3_md	Vale	30-2	30
tf1_52_T3_ss3_pd	Vale	30-3	30
tf1_52_T2_ss3_pd	Vale	30-4	30

tf1_52_T1_ss3_md	Õige	30-5	30
tf1_11_M_ss1	Vale	11	11
tf1_11_T3_ss1_md	Vale	11-1	11
tf1_11_T3_ss1_pd	Vale	11-2	11
tf1_11_T2_ss1_md	Vale	11-3	11
tf1_11_T1_ss1_md	Õige	11-4	11
tf1_11_T2_ss1_pd	Vale	11-5	11
tf1_71_M_ss1	Vale	31	31
tf1_71_T3_ss1_pd	Vale	31-1	31
tf1_71_T3_ss1_md	Vale	31-2	31
tf1_71_T2_ss1_pd	Vale	31-3	31
tf1_71_T1_ss1_md	Õige	31-4	31
tf1_71_T2_ss1_md	Vale	31-5	31
tf1_58_M_ss3	Vale	28	28
tf1_58_T2_ss3_md	Vale	28-1	28
tf1_58_T3_ss3_pd	Vale	28-2	28
tf1_58_T1_ss3_md	Õige	28-3	28
tf1_58_T2_ss3_pd	Vale	28-4	28
tf1_58_T3_ss3_md	Vale	28-5	28
tf1_72_M_ss2	Vale	32	32
tf1_72_T1_ss2_md	Õige	32-1	32
tf1_72_T2_ss2_md	Vale	32-2	32
tf1_72_T3_ss2_pd	Vale	32-3	32
tf1_72_T3_ss2_md	Vale	32-4	32
tf1_72_T2_ss2_pd	Vale	32-5	32
tf1_46_M_ss3	Vale	23	23
tf1_46_T2_ss3_md	Vale	23-1	23
tf1_46_T1_ss3_md	Õige	23-2	23
tf1_46_T3_ss3_md	Vale	23-3	23
tf1_46_T2_ss3_pd	Vale	23-4	23
tf1_46_T3_ss3_pd	Vale	23-5	23
tf1_65_M_ss1	Vale	34	34

tf1_65_T3_ss1_md	Vale	34-1	34
tf1_65_T2_ss1_pd	Vale	34-2	34
tf1_65_T2_ss1_md	Vale	34-3	34
tf1_65_T1_ss1_md	Õige	34-4	34
tf1_65_T3_ss1_pd	Vale	34-5	34
tf1_57_M_ss2	Vale	27	27
tf1_57_T1_ss2_md	Õige	27-1	27
tf1_57_T2_ss2_pd	Vale	27-2	27
tf1_57_T2_ss2_md	Vale	27-3	27
tf1_57_T3_ss2_md	Vale	27-4	27
tf1_57_T3_ss2_pd	Vale	27-5	27
tf1_80_M_ss1	Vale	37	37
tf1_80_T3_ss1_pd	Vale	37-1	37
tf1_80_T3_ss1_md	Vale	37-2	37
tf1_80_T1_ss1_md	Õige	37-3	37
tf1_80_T2_ss1_md	Vale	37-4	37
tf1_80_T2_ss1_pd	Vale	37-5	37
tf1_76_M_ss3	Vale	35	35
tf1_76_T2_ss3_md	Vale	35-1	35
tf1_76_T3_ss3_pd	Vale	35-2	35
tf1_76_T1_ss3_md	Õige	35-3	35
tf1_76_T3_ss3_md	Vale	35-4	35
tf1_76_T2_ss3_pd	Vale	35-5	35
tf1_78_M_ss2	Vale	36	36
tf1_78_T2_ss2_md	Vale	36-1	36
tf1_78_T2_ss2_pd	Vale	36-2	36
tf1_78_T3_ss2_md	Vale	36-3	36
tf1_78_T3_ss2_pd	Vale	36-4	36
tf1_78_T1_ss2_md	Õige	36-5	36
tf1_63_M_ss2	Vale	33	33
tf1_63_T1_ss2_md	Õige	33-1	33
tf1_63_T3_ss2_md	Vale	33-2	33

tf1_63_T2_ss2_md	Vale	33-3	33
tf1_63_T3_ss2_pd	Vale	33-4	33
tf1_63_T2_ss2_pd	Vale	33-5	33

Lisa B

Tabel B1

Üksikvastuse teooria analüüsitulemused

TMB stiimul	Keerukus <i>difficulty</i>	Diskrimineeritavus <i>discriminability</i>	MaRs-IB stiimul	Keerukus <i>difficulty</i>	Diskrimineeritavus <i>discriminability</i>
15.png	-6.46	0.61	5.png	-3.78	0.71
22.png	-1.57	1.77	29.png	-0.17	1.56
36.png	1.38	0.53	3.png	0.43	1.07
35.png	0.90	0.99	9.png	-0.56	1.33
21.png	-2.75	1.18	8.png	-2.81	0.63
3.png	-2.77	0.94	35.png	-0.28	1.40
1.png	-6.66	0.53	18.png	0.16	0.99
29.png	-1.41	1.33	12.png	-0.71	1.69
18.png	-2.71	1.03	22.png	-2.91	1.08
32.png	-0.52	0.96	14.png	-0.72	1.54
27.png	0.21	1.22	21.png	-0.47	1.11
23.png	0.20	0.19	13.png	-0.46	1.24
11.png	-1.91	0.84	24.png	-2.32	1.50
2.png	-1.42	0.93	25.png	-0.85	1.40
3.png	31.16	-0.19	2.png	0.34	1.54
26.png	-2.03	1.25	1.png	-1.96	0.65
25.png	-2.53	1.42	33.png	-0.31	2.07
17.png	-2.71	0.95	37.png	-0.78	2.28
28.png	-0.70	1.55	23.png	-0.64	1.79
33.png	1.00	0.81	27.png	-3.12	0.22
12.png	-2.84	1.33	7.png	-2.19	0.12
6.png	-18.45	0.24	11.png	-0.73	0.98
37.png	1.71	0.17	26.png	-0.63	2.61
31.png	-1.35	0.57	17.png	0.77	0.80
8.png	-6.54	0.60	28.png	-1.57	0.92
19.png	-2.79	1.29	19.png	-2.87	0.52

4.png	-1.02	0.34	15.png	0.71	0.50
24.png	-1.85	0.86	16.png	-0.48	1.68
16.png	-2.35	1.07	4.png	-15.05	0.22
5.png	-1.65	0.42	3.png	-6.00	0.88
13.png	-2.74	1.11	32.png	-0.52	3.01
9.png	-2.44	1.98	6.png	-2.25	0.62
34.png	0.43	0.89	34.png	-1.39	2.63
7.png	-4.53	0.82	36.png	-0.36	2.00
14.png	-3.31	0.88	31.png	-1.90	1.46

Käesolevaga kinnitan, et olen korrektselt viidanud kõigile oma töös kasutatud teiste autorite poolt loodud kirjalikele töödele, lausetele, mõtetele, ideedele või andmetele. Olen nõus oma töö avaldamisega Tartu Ülikooli digitaalarhiivis DSpace.

Laura Asu