

Tartu Ülikool
Psühholoogia instituut

Gerli Silm

**MOTIVATSIOONI ROLL VAIMSE VÕIMEKUSE MÕÕTMISEL JA
TESTITULEMUSTE TÕLGENDAMISEL**

Magistritöö

Juhendajad: Olev Must, PhD
Karin Täht, PhD

Läbiv pealkiri: Testi täitmise motivatsioon madala olulisusega testides

Tartu 2015

Kokkuvõte

Arvestades, kui palju kasutatakse inimeste omaduste ja võimete hindamiseks teste, sh saadud andmeid rahvusvahelisteks võrdlusteks, on tekkinud küsimus, kui võrd näitavad saadud tulemused seda, mida sooviti mõõta, ning kui võrd peegeldavad muid tegureid. Kuna taoliste testide täitmisel puuduvad tavaliselt testitaitja jaoks olulised tagajärjed, mis võib vähendada motivatsiooni ja pingutust testi täitmisel, käsitlesin käesolevas töös kahe motivatsiooninäitaja – enesekohase motivatsiooniküsimustiku (*Student Opinion Scale*, Sundre, 1999) ning APV (ajaline pingutus vastamisel) (Wise & Kong, 2005) – seoseid vaimse võimekuse testi skooriga ning kummagi meetodi omadusi motivatsiooni filtreerimisel.

Valim on kogutud neljast erinevast uuringust, milles osalesid üliõpilased ja abiturientid erinevatest koolidest (N = 714). Leidsin, et kaks meetodit eristavad mittemotiveeritud motiveeritud suhteliselt sarnaselt. Motivatsiooniküsimustik on filtreerimismeetodina rangem, st selle alusel loetakse mittemotiveeritud rohkem testitaitjaid kui APV-indeksi alusel. See-eest võib APV-indeks olla objektiivsem näitaja, kuna tegemist on mitte-sekkuva meetodiga, samas kui enesekohaselt teatatud motivatsioon võib olla mõjutatud sotsiaalselt soovitatavast vastamisest. Viimase puhul esines vastamises ka soolisi erinevusi. Mudel, kuhu peale kummagi motivatsiooninäitaja on lisatud ka kooliedukuse näitaja riigieksamite keskmise tulemuse näol ja sugu, seletab ligi 50% testitulemuse varieeruvusest.

Märksõnad: *testitaitmise motivatsioon, madala olulisusega testid, testimise valiidsus*

Abstract**Title:** THE ROLE OF MOTIVATION IN THE MEASUREMENT OF MENTAL ABILITIES AND THE INTERPRETATION OF TEST RESULTS

The wide use of psychometric tests in assessing the test-takers' characteristics and abilities, and also the use of the obtained data for international comparisons, have raised the question to what extent the interpretation of the results reflects reality. Usually such tests are non-consequential for the test-takers. In my thesis I studied the relationship between two motivation indicators – self-reported motivation, measured with Student Opinion Scale (Sundre, 1999), and RTE (response time effort) (Wise & Kong, 2005) – and mental ability test results, and the properties of the two indicators as motivation filtering methods.

The sample is collected from four different studies involving university students and high school senior students from different schools and universities (N = 714). I found that the two methods distinguish the unmotivated from the motivated relatively similarly. Student Opinion Scale as a motivation filtering method is more strict, i.e. more test-takers will be considered unmotivated by the criteria of the motivation scale than on RTE. On the other hand RTE can be considered more objective, as it is an unobtrusive method, whereas self-reported motivation might be affected by socially desirable responding. There were also gender differences in the responses to the motivation questionnaire. A model, with either of the motivation indicators and national examination results, which were considered as a proxy for mental ability, and gender, explained approximately 50% of the variance in test scores.

Keywords: *test-taking motivation, low-stakes testing, testing validity*

Sissejuhatus

Psühhomeetrilisi teste kasutatakse inimeste erinevate psühholoogiliste omaduste ja võimete hindamiseks aina enam. Taoliste testidega proovitakse objektiivselt mõõta näiteks oskusi ja teadmisi, vaimseid võimeid, hoiakuid, isiksuseomadusi ja haridusalaseid saavutusi. Testide laialdasema kasutusega seoses on üha rohkem hakatud tähelepanu pöörama sellele, mil määral näitavad testide tulemused testitajate tegelikke omadusi. Näiteks on varasemalt leitud, et võimekustesti sooritust võivad peale võimekuse mõjutada ka muud tegurid, nagu vastaja hoiakud, väsimus, ärevus, motivatsioon ning testi formaat ja pikkus (nt Ackerman & Kanfer, 2009). Ajal, mil teste kasutatakse laialdaselt ja igapäevaselt, tekib õigustatult küsimus, kui võrd teste täites pingutatakse ning ausalt ja läbimõeldult vastatakse. Valiidsete järelduste tegemiseks on uurijatele oluline teada, kas testidest saadud vastused peegeldavad seda, mida on tahetud mõõta, või hoopis midagi muud ning kas neid on võimalik üksteisest eristada. Näiteks, kui võrd näitab vaimse võimekuse testi tulemus testitaja tegelikke võimeid ja kui võrd tema motivatsiooni (või selle puudumist) testi täitmisel.

Käesolevas magistritöös võtangi vaatluse alla testi täitmise motivatsiooni vaimse võimekuse testides ning motivatsiooni filtreerimise tehnikad, mis võimaldaks eristada motiveeritud testitajaid mittemotiveeritute. Samuti vaatan, kuidas ennustab testi täitmise motivatsioon vaimse võimekuse testi tulemust mudelis, kuhu on ennustajana lisatud ka teine vaimset võimekust peegeldav näitaja.

Vaimse võimekuse mõõtmine hariduses

Vaimsed võimed ja haridus on tugevas vastastikkuses seoses. Ka esimene süstemaatiline vaimse võimekuse testimine toimus koolikontekstis, kui 1905. aastal püüdis Alfred Binet mõõta, millised lapsed saavad koolis hakkama. Vaimsel võimekusel on mõju kooliedule – paljudes uuringutes on näidatud, et lapsed, kes saavad kõrgemaid tulemusi vaimse võimekuse testides, kipuvad ka koolis õpetatust paremini aru saama, võrreldes on madalamate tulemustega kaaslastega. Korrelatsioon IQ skoori ja hinnete vahel on keskel läbi .50, mis tähendab, et ühe hajuvusest on teisega võimalik seletada 25%. Sellest on järeldatud, et heade hinnete saamine sõltub peale vaimse võimekuse ka teistest teguritest nagu püsivus, koolihuvi ja valmisolek õppida (Neisser jt, 1996).

Arvestades, kui palju viiakse tänapäeval läbi erinevaid rahvusvahelisi haridusalaseid teste (näiteks PISA (*Programme for International Student Assessment*), TIMSS (*Trends in International Mathematics and Science Study*), PIRLS (*Progress in International Reading*

Literacy Study)) ja võrreldakse nende abil erinevate riikide haridustulemusi, on äärmiselt oluline osalejate testimotivatsiooni taolistes kontekstides mõista (Barry, 2010). Näiteks TIMSSi puhul on näidatud, et motivatsiooninäitajad seletavad ära suure osa riikidevahelisest erinevusest: vähemalt kaheksanda klassi õpilaste puhul seletas riiklik keskmine motivatsiooninäitaja skoor (*Student Task Persistence*) üle poole testi tulemuste variatiivsusest, jättes muude tegurite, sealhulgas õpilaste teadmised matemaatikast ja loodusainetest, seletada alla poole (Boe, May, & Boruch, 2002).

Testi täitmise motivatsioon ja motivatsiooni filtreerimine

Testi täitmise motivatsioon näitab, mil määral testitaitja „annab endast testi täitmisel parima, eesmärgiga esitada võimalikult täpselt oma teadmisi ja võimeid valdkonnas, mida test mõõdab“ (Wise & DeMars, 2005). On lihtne mõista, et motivatsioon võib oleneda nii indiviidist kui ka konkreetsest testi täitmise situatsioonist ja seega võib tekkida olukord, kus testitaitja ei näita oma tegelikke teadmisi ega võimeid, vaid hoopis seda, „mida ta suudab teha vähima pingutusega“ (O’Neil, Sugrue, & Baker, 1995).

Mõiste „motivatsiooni filtreerimine“ (ingl *motivation filtering*) tähistab protseduuri, mille käigus eemaldatakse kogutud andmete hulgast tõenäoliselt mittemotiveeritud testitaitjate vastused (Steedle, 2014). Motivatsiooni filtreerimine on seega protseduur, mis võib potentsiaalselt parandada testi tulemuste alusel tehtavate järelduste valiidsust (Sundre & Wise, 2003). Motivatsiooni filtreerimine ei ole meetod, mis eemaldaks tulemustest täielikult motivatsiooni mõju, vaid selle abil tehakse kindlaks ja jäetakse lõplikult valimist kõrvale rühm, kelle testi täitmise motivatsiooni võib lugeda äärmuslikult madalaks. Konkreetsetest motivatsiooni filtreerimise tehnikatest saab lugeda käesolevast tööst alates lk 9.

Erinevad testide täitmise situatsioonid

Olukorras, kus psühhomeetrilised testid on kasutusel väga laialdaselt, võivad ette tulla situatsioonid, mil testidel ei ole alati testitaitja enda jaoks tähenduslikke tagajärgi, küll aga teevad testi tulemuste alusel järeldusi uurijad. Kui eeldada, et saadud tulemused näitavad alati konkreetset omadust, mida test peaks mõõtma, ja ei midagi muud, võivad need järeldused olla mittevaliidsed (Mislevy, 1995; Wainer, 1993; Wise & DeMars, 2005).

Lähtudes sellest, millised tagajärjed on testi täitmisel testitaitja jaoks (või millisena testitaitja neid tajub), räägitakse situatsioonidest, kus testidel on testitaitja jaoks kõrge (ingl *high stakes*) või madal olulisus (ingl *low stakes*) ehk jaotatakse need vastavalt testitaitja jaoks tagajärgedega (ingl *consequential*) või tagajärgedeta (ingl *nonconsequential*) testideks.

Esimesel juhul kaasneb testi tulemusega testitajale oluline tagajärg, milleks võib olla hinne, eksamitulemus, nõutava lävendi ületamine, edu konkursil jm. Teisel puhul taolisi tagajärgi ei ole. Erinevates uuringutes kasutatavad testid, mida inimestel täita palutakse, on enamasti inimese jaoks madala olulisusega. Oluline on tähele panna, et madal või kõrge olulisus ei viita konkreetse testi omadustele, vaid sellele, kuidas konkreetsetes situatsioonides testi tajutakse ja sooritatakse. Üks ja sama test, täpsemalt selle tulemus, võib erinevates olukordades ja erinevatele inimestele olla nii kõrge kui ka madala olulisusega.

Testi täitmise motivatsioon ootuse-väärtuse teooria raamistik

Teoreetiliselt on testi täitmise motivatsiooniga seonduvat käsitletud põhiliselt saavutusmotivatsioonialasest kirjandusest pärit ootuse-väärtuse teooria (ingl *expectancy-value theory*) valguses. Kaasaegsed ootuse-väärtuse teooriad (sh Eccles & Wigfield, 2002; Wigfield & Eccles, 2000) põhinevad Atkinsoni (1964) ootuse-väärtuse mudelil. Ecclesi ja Wigfieldi ootuse-väärtuse teooria kohaselt sõltub pingutus, mida inimesed ülesande täitmiseks rakendavad, nende ootusest edu saavutada ning sellest, kui väärtuslikeks nad testi tulemusi hindavad. Ootuse-väärtuse teooria kohaselt koosnevad motivatsioonilised ootused kahest osast: 1) edu ootusest, mis kätkeb endas inimese tõekspidamisi selle kohta, kui hästi on võimalik ülesannet sooritada, ning 2) uskumustest võimekuse kohta, mida määratletakse kui inimese usku oma pädevusse kindlas valdkonnas. Väärtused koosnevad Ecclesi ja Wigfieldi (2000; 2002) käsitluses omakorda neljast tähtsast osast: 1) saavutusväärtusest (ingl *attainment value*), mis on defineeritud kui ülesande hästi sooritamise olulisus; 2) sisemisest väärtusest (ingl *intrinsic value*), mis tähendab naudingut ja rahulolu, mida inimene ülesande täitmisest saab; 3) kasulikkusest (ingl *utility value*), mis näitab ülesande kooskõla tulevikueesmärkidega; 4) kulust (ingl *cost*), mis tähistab ülesande täitmisega seotud ebaseeldivaid külgi, sh kui palju see teisi tegevusi piirab, kui suurt pingutust see vajab ja millised vastumeelsed emotsionaalsed seisundid sellega kaasnevad (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). Ehk lühidalt: ootuse-väärtuse teooria seletab inimese käitumist selle kaudu, kui tõenäoliseks ta edu saavutamist peab ja millise väärtusega on tema jaoks tegevuse tulemus.

Erinevad testitajate motivatsiooni mõjutavad tegurid

Testi täitmise motivatsiooni uuringutes on leitud, et ülesannete raskusaste või töömaht, mis kulub ülesande lahendamiseks, on samuti seotud testi täitmise motivatsiooniga. Nimelt kaasneb raskemate ülesannetega madala olulisusega testimise situatsioonides tavaliselt madalam

motivatsioon (Bovaird, 2002; Wise, 2006; Wolf & Smith, 1995). Neid tulemusi võib kergesti tõlgendada läbi ootuse-väärtuse teooria: keerukamad ülesanded on seotud suurema kulu ja madalamate ootustega edule, seega võivad kaasa tuua madalama motivatsiooni. Samuti on näidatud, et kui testis on avatud vastustega küsimused, on testitajate motivatsioon madalam kui valikvastustega testi puhul (DeMars, 2000; Sundre & Kitsantas, 2004). Ka neid tulemusi võib läbi ootuse-väärtuse teooria tõlgendada: avatud vastuste andmine nõuab rohkem pingutust ning puudub võimalus juhuslikult õige vastus anda, seega kaasneb avatud vastustega suurem kulu ja väiksem võimalus eduks (Barry & Finney, 2009).

Kokkuvõtvalt võib eelnevast lähtudes arvata, et testitajad on madala olulisusega testides valmis pingutama, kuid on selleks rohkem valmis, kui test ei ole liiga raske. Kui test tundub ülemäära keerukas, siis mingil hetkel pingutus väheneb. Testide läbiviijaid huvitavad tihti aga just keerukamate ülesannete lahendamise võimed (Barry, 2010).

Esineb ka inimesi, kes on testide täitmisel valmis endast parima andma isegi siis, kui tegemist on testiga, mida üldiselt võiks lugeda madala olulisusega testiks. Põhjuseid selleks võib olla erinevaid: tahtmine endast alati parim anda, huvi enda võimekuse vastu, soov teistest parem olla, panustada uurijate töösse, olla „hea kodanik“ (Eklöf, 2006).

Madala testitajate motivatsiooni mõju vähendamine

Erinevates uuringutes on madala testitajate motivatsiooni negatiivset mõju testi tulemustele proovitud vähendada eelkõige kahel viisil: 1) testitajate motiveerimine enne testimist; 2) andmete filtreerimine peale testimist (näiteks mittemotiveeritute andmete eemaldamine andmehulgast). Esimese meetodi puhul on eesmärgiks saavutada see, et testitajad pingutaksid maksimaalselt ka madala olulisusega testide puhul. Selle eesmärgi täitmiseks võidakse näiteks testitajatega enne testi arutada selle olulisust, pakkuda materiaalseid hüvesid testi täitmise eest või lubada tunnustust heade tulemuste eest (Swerdzewski, Harnes, & Finney, 2011).

Kuigi madala olulisusega testide kasutamisel on oht, et osa testitajate madala motivatsiooni tõttu võivad saadud tulemused olla mittevaliidsed (st need ei peegelda ainult mõõdetavat konstrukti, vaid ka motivatsioonilist komponenti), siis arvatavasti nende kasutamisest ei loobuta ning see ei pruugi ka otstarbekas olla. Tihti on madala olulisusega test kõigest hoolimata eelistatavaks ja kättesaadavaimaks andmete kogumise viisiks. Seetõttu on uurijad hakanud otsima viise, kuidas motiveeritute tulemusi mittemotiveeritute omadest eristada ehk motivatsiooni filtreerida.

Motivatsiooni filtreerimine annab võimaluse kasutada andmeanalüüsis ainult neid andmeid, mille puhul on väiksem kahtlus, et madal motivatsioon võis tulemust mõjutada. Tänapäeval kasutatakse motivatsiooni hindamiseks peamiselt kahte meetodit. Esimene neist on lühike enesekohane küsimustik, mida eesti keeles võiks nimetada testimotivatsiooni küsimustikuks (ingl *Student Opinion Scale*) (Sundre, 1999). Teine motivatsiooni mõõtmise meetod põhineb igale üksikule küsimusele vastamiseks kulunud aja hindamisel. Wise ja Kong (2005) pakkusid selleks välja näitaja „ajaline pingutus vastamisel“ ehk APV (ingl *Response Time Effort*).

Soolised erinevused testi täitmise motivatsioonis ja hariduses

Motivatsiooni filtreerimine võib anda huvitavat infot ka testi täitmise motivatsiooni kohta üldisemalt. Näiteks on leitud, et äärmiselt madal motivatsioon on omane pigem poistele kui tüdrukutele (Freund & Rock, 1992). Segal (2006) jaotas testitajad aga kaheks: „kaalutlejad“ (ingl *Economists*) ning „pingutajad“ (ingl *Boy Scouts*). „Kaalutlejad“ on inimesed, kes suudavad olukorras, kus testi täitmise eest pakutakse tasu, parandada märgatavalt oma tulemust, samas kui „pingutajate“ tulemus erinevates tingimustes oluliselt ei muutunud, ehk eeldatavasti püüti mõlema tingimuse puhul, olenemata tasu pakkumisest, endast parim anda. Segali uuringus osutus kaalutlejateks 28,6% naissoost testitajatest ja 46% meessoost testitajatest (pingutajateks vastavalt 71,4% naistest ja 54% meestest). Tulemus räägib sellest, et mehed võivad väiksema tõenäosusega testi täitmisel pingutada, kui sellel ei ole nende jaoks olulist tagajärge.

Hariduskontekstis üldiselt esineb mitmeid soolisi erinevusi. Üks trend viimaste kümnendite jooksul on olnud suur naiste osakaalu kasv hariduse omandamisel – praeguseks hetkeks edestavad naised mehi hariduses pea kõikides arenenud riikides. See väljendub selles, et naised on rohkem keskhariduse lõpetajate ning ka ülikooli sisseastujate ja lõpetajate seas ning mehi on rohkem keskhariduse katkestajate seas (Pekkarinen, 2012). Motivatsioon (sh testitajatel) võib teiste hulgas olla üks tegur, mis aitaks neid erisusi põhjendada. Näiteks Spinath, Freudenthaler ja Neubauer (2010) järeldasid oma uuringust, et peale intelligentsuse mängivad haridusalaste saavutuste juures olulist rolli ka motivatsioon ja isiksuseomadused ning seega pidasid neid võimalikeks teguriteks, mis aitaksid soolisi erinevusi haridusalastes saavutustes seletada.

Testimotivatsiooni küsimustik

Testimotivatsiooni küsimustiku (*Student Opinion Scale*) eelkäijaks oli *Motivation Questionnaire* (Wolf & Smith, 1993, viidatud Sundre & Thelk, 2007 vahendusel), mis koosnes kaheksast küsimusest, millest viis olid seotud testi tajutud olulisusega ning kolm pingutusega. Antud väitestikku loeti ühedimensionaalseks, st kõik kaheksa küsimust koondusid ühe faktori alla. Sundre (1999) täiendas skaalat kahe küsimusega testi täitmise pingutuse kohta ning muutis ka olemasolevate küsimuste sõnastust. Saadud skaala puhul leiti, et eristuvad kaks faktorit: olulisus ja pingutus. Kahefaktorilist struktuuri on erinevate valimite peal ka korduvalt kinnitatud (Sundre & Finney, 2002; Sundre & Thelk, 2007).

Testimotivatsiooni küsimustiku eesmärk on mõõta testitajate testi täitmise motivatsiooni, et anda uurijatele sellekohast informatsiooni konkreetses testi täitmise situatsioonis. Küsimustik antakse täitmiseks peale testi täitmist, järeltestina. See võib järgneda nii üksikule testile kui ka testide partaale. Testimotivatsiooni küsimustik ei olnud algselt mõeldud selleks, et hinnata testitajaid individuaalselt, vaid pigem oli selle funktsiooniks anda ülevaade enesekohaselt teatatud motivatsiooniprotsessidest mingil konkreetsel testimiskorral ehk hinnata, milline on testitajate motivatsioonitase keskmiselt. Sellegipoolest on testimotivatsiooni küsimustikku kasutatud ka selleks, et madala motivatsiooniga testitajate andmeid andmehulgast välja filtreerida (nt Sundre & Wise, 2003), kuid motivatsiooni filtreerimise instrumendina ei ole see veel lõplikku kinnitust saanud ja vajab edasist uurimist. Sundre ja Thelk (2007) leiavad, et parim viis selle vahendi kasutamiseks on testitajate motivatsiooni üldine kirjeldamine madala olulisusega testimise situatsioonis. Võimekustesti tulemuste paremale tõlgendamisele võivad kaasa aidata motivatsiooniküsimustiku allskaalad ning võimekustesti enda teised kirjeldavad statistikud (standardhälve, korrelatsioonid).

Testimotivatsiooni skooride esitamine koos võimekustesti tulemustega aitab hinnata testitulemuste valiidsust: kui tulemused on madalad nii võimekustestis kui ka motivatsiooniküsimustikus, võib kahtlustada, et tulemus ei ole valideeritud. Ning vastupidi: kui testimotivatsiooni skoorid on kõrged, võib oletada, et saadud võimekustesti tulemus peegeldab testitajate võimeid tõelähedaselt.

Enesekohase küsimustiku kasutamisel motivatsiooni mõõtmiseks on ka teatud puudusi. Küsimustiku kasutamine eeldab, et testitajad: a) teavad, milline on nende motivatsiooni tase, b) oskavad kasutada skaalat selle iseloomustamiseks, c) on ausad ja vastavad täpselt (Swerdzewski, Harmes, & Finney, 2011). Enesekohaste küsimustike täitmisel võivad inimesed olla mõjutatud ka sotsiaalsest soovtavusest. Sotsiaalne soovtavus viitab

kalduvusele näidata ennast kehtivate kultuurinormide osas paremas valguses (Krumpal, 2013). On näidatud, et teatud grupid, näiteks naised ja kollektivistlike kultuuride esindajad, vastavad enesekohastes testides keskmiselt rohkem sotsiaalselt soovitatavalt kui teised (Bernardi, 2006; Dalton & Ortegren, 2011).

Ajaline pingutus vastamisel

APV on meetod, mida saab kasutada arvutis täidetavate testide puhul. Selleks määratakse igale küsimusele, lähtudes selle pikkusest ning vastuste sagedusjaotusest, lävi, mis tähistab aega, mis võiks olla minimaalne, et jõuda küsimus läbi lugeda ja sellele vastus anda. Kui vastus antakse määratud lävega võrreldes kiiremini, on tõenäoline, et vastus on antud juhuslikult ning võimalik, et küsimust isegi läbi lugemata.

Vastustele kulunud aegade sagedusjaotusest järeldasid Wise ja Kong (2005), et üldiselt jaotuvad ajad normaaljaotuse kohaselt, kuid võib märgata ka väikest väljatõuget või n-õ jõnksu sagedusjaotuse esimeses pooles ehk hulka vastuseid, mis on antud tunduvalt kiiremini kui ülejäänud. Oma uuringus kasutasid nimetatud autorid lävedena 3 ja 5 sekundit. Lävesid võrreldakse aegadega, mis testitajal kulub konkreetsete testi ülesannete lahendamiseks või küsimustele vastamiseks. Võrdluse alusel luuakse iga küsimuse kohta binaarne tunnus, kus lävest pikem aeg tähistatakse ühega, lühem nulliga. Nende tunnuste alusel jällegi arvutatakse iga vastaja kohta indeks: lävest pikemate vastuste aegadega küsimuste summa jagatakse kogu küsimuste arvuga. Saadud indeks kirjeldab seega, millisele osale küsimustest kulutas vastaja etteantud lävest rohkem aega. Näiteks kui indeks on 0.6, tähendab see seda, et 60% küsimuste puhul oli vastusele kulutatud aeg lävest pikem, 40% puhul on tegemist aga tõenäoliselt juhuslikult antud vastustega. Autorid otsustasid, et motiveerituteks võiks lugeda neid, kes on vähemalt 90% küsimuste puhul kulutanud vastamiseks „piisavalt“ aega. Vastused, millele kulutati vähem aega, loeti liigkiireteks (ingl *rapid response*).

Kirjeldatud meetodi miinuseks on, et seda ei saa kasutada paber-pliiats testide puhul. Eeliseks on aga see, et meetod annab objektiivsemat teavet kui enesekohane küsimustik ning detailsemat teavet kui testi täitmisele kulunud aeg tervikuna.

Kahe motivatsiooni filtreerimise meetodi võrdlus

APV-näitaja väljatöötamisel võrdlesid Wise ja Kong (2005) selle tulemusi ka testimotivatsiooni küsimustiku tulemustega. Nad leidsid, et kaks meetodit eristavad

motiveeritud mittemotiveeritute suhteliselt sarnaselt, ainult et APV võimaldab välja filtreerida vähem testitajaid ning järelejäänud tulemuste keskmine on kõrgem.

Ka Swerzewski, Harnes ja Finney (2011) kasutasid oma uuringus kahte eelnimetatud motivatsiooni filtreerimise tehnikat. Nemadki leidsid, et mõlemad meetodid eristasid mittemotiveeritud motiveeritute suhteliselt sarnaselt. Autorid toovad oma töö ühe puudusena välja selle, et nad ei leidnud eriti uuringuid, mis võrdleks omavahel filtreerimise meetodeid, millest üks põhineb enesekohasel raportil ja teine on mitte-sekkuv. Seetõttu vajab uuring kordamist teistsugusel valimil ja teistes tingimustes, et oleks võimalik teha järeldusi kirjeldatud filtreerimismeetodite praktikas kasutamise võimalikkuse kohta. Oma töös püüan seda käsitlust edasi arendada.

Käesolevas töös on minu eesmärgiks vaimse võimekuse testi tulemuste ennustamine erinevate motivatsiooninäitajatega madala olulisusega testimise situatsioonis ning kahe motivatsiooni filtreerimise tehnika võrdlemine. Tulemuse ennustamise mudeli loomise eesmärgiks on erinevad olulised näitajad tervikuks siduda, et nende vahelist dünaamikat paremini mõista, sealhulgas näidata, kas ja kuidas eristuvad motivatsiooni ning vaimse võimekuse näitajad tulemuse ennustamisel. Teiseks soovin vaadata, kas motivatsiooninäitajate mudelisse lisamine aitab ära seletada suurema osa vaimse võimekuse testi tulemuse hajuvusest kui seda teeks vaimse võimekuse näitaja üksinda. Saadud tulemuste alusel soovin omakorda teha järeldusi nende motivatsiooninäitajate sobilikkuse kohta motivatsiooni filtreerimiseks praktikas.

Töös on kaks põhilist uurimisküsimust:

1. Kuidas erinevad omavahel motiveeritud testitajate mittemotiveeritute eristamisel kaks motivatsiooni filtreerimise meetodit (enesekohane raport ja mitte-sekkuv APV-indeks)?
2. Kuidas ennustab üks vaimse võimekuse näitaja teist mudelis, millesse on kaasatud ka motivatsiooninäitajad?

Meetod

Käesolevas uuringus kasutasin järgnevalt kirjeldatud mõõdikuid.

Akadeemilise testi (AT) lühiversioon (vt Must & Allik, 2002; Silm, Täht, & Must, 2013) on vaimse võimekuse test, mis on kokku pandud Tartu ülikooli akadeemilise testi (2008–2012) valitud ülesannetest. Ülesanded on valitud vastavalt raskusastmele, mis on välja arvatud ülikooli sisseastujate testivastuste põhjal. Testist on kaks sisult ja raskuselt väga sarnast paralleelversiooni. Lühitest koosneb kolmest alltestist: sõnavara, matemaatika ja ruumiline mõtlemine. Igas alltestis on 15 ülesannet, st kokku on ülesandeid 45.

Testi täitmise aeg tähistab aega, mis kulutati kogu testi täitmiseks. Kuna test sooritati arvutis, oli võimalik seda täpselt mõõta. Lisaks eelnevale arvutati välja ka teine ajaga seotud näitaja – **APV-indeks**. Indeksi arvutamisel kasutatud lävede määramisel lähtuti varasematest uuringutest, pilootuuringu tulemustest ning küsimuste iseloomust. Sõnavara alltesti kõikide küsimuste läveks määrati 5 sekundit, matemaatika ning ruumilise mõtlemise alltesti küsimuste läveks 10 sekundit. Motivatsiooni filtreerimisel APV-indeksi alusel kasutasin äralõikepunkti 0.9 sarnaselt meetodi väljatöötajatega (Wise & Kong, 2005).

Enesekohane testimotivatsiooni küsimustik on eestindatud versioon küsimustikust *Student Opinion Scale* (Sundre, 1999). Test koosneb kahest allskaalast: olulisus ja pingutus; kummagi alla kuulub viis küsimust (vt tabel 1). Küsimustik esitati alati vahetult peale vaimse võimekuse testi täitmist. Enesekohase küsimustiku alusel motivatsiooni filtreerimisel kasutasin pingutuse alltesti skoori ning äralõikepunktina skoori 15, nii nagu seda tegid Swerdzewski, Harmes ja Finney (2011), tuues põhjenduseks selle, et kui motivatsiooni saab hinnata skaalal ühest viieni, on loogiline, et motiveerituteks saab lugeda need, kes hindavad oma motivatsioon iga küsimuse puhul keskmiselt vähemalt kolmega ehk üle keskmise. Pingutuse alltesti skoori kasutan, kuna see on sisuliselt sarnasem APV-indeksiga – mõlemad hindavad pingutust ja püsivust.

Tabel 1

Enesekohane testimotivatsiooni küsimustik (Sundre, 1999) eesti keeles (autori tõlge)

Küsimused	Allskaala
1. Minu jaoks oli oluline, et mul läheks testis võimalikult hästi.	Olulisus
2. Ma töötasin pingsalt kogu testi vältel.	Pingutus
3. Mulle ei paku huvi, kuidas mul võrreldes teistega läks.*	Olulisus
4. Ma ei muretse selle pärast, millise tulemuse ma saan.*	Olulisus
5. See test oli minu jaoks oluline.	Olulisus
6. Ma andsin testi täites endast parima.	Pingutus
7. Ma oleksin olnud võimeline testi paremini sooritama, kui oleksin rohkem pingutanud.*	Pingutus
8. Ma tahaksin teada, kui hästi testi sooritasin.	Olulisus
9. Testi täites ei olnud ma sellele täielikult pühendunud.*	Pingutus
10. Ma olin suuteline pingutama kuni testi lõpuni.	Pingutus

*pööratud küsimused

Küsimustiku küsimustele saab vastata 5-palli skaalal, kus 1 tähistab „pole üldse nõus“ ja 5 – „olen täiesti nõus“. Testiskoori esitamisel tuleks ära tuua mõlema allskaala skoorid eraldi. Testi autor ei soovita koguskoori esitamist, kuna on leitud, et testimotivatsioon ei ole ühedimensiooniline konstrukt. Kummagi allskaala skoorid saavad jääda vahemikku 5–25 punkti (Sundre & Thelk, 2007).

Riigeksamite tulemused on olemas ühe alavalimi kohta (vt valimi kirjeldus allpool). Tulemused on testitäitjate endi teatatud. Lähtudes Deary jt (2007) artiklist, kus leiti, et lapsepõlves mõõdetud IQ on tugevas seoses riigeksamite tulemustega, vaatan riigeksamite tulemusi tinglikult kui vaimse võimekuse indikaatoreid. Nimelt leidsid nemad oma rohkem kui 70 000 inglise last hõlmanud uuringus, et 11-aastaselt mõõdetud intelligentsus on tugevas korrelatsioonis ($r = 0.81$) 16-aastaselt sooritatud riiklike eksamite tulemustega. Lisaks võib eeldada, et riigeksamid on enamjaolt nende sooritajatele kõrge olulisusega, mis tähendab, et eksamite tegemisel proovitakse anda endast parim ning näidata oma tegelikku võimekust. Sarnaselt on ka Segal (2006) kasutanud vaimse võimekuse näitajana testiskoori, mis on saadud tingimustes, kus test oli täitja jaoks oluliste tagajärgedega.

Valim

Valim on kogutud neljast erinevast uuringust:

1. „Üliõpilaste enesekohaste hoiakute ja vaimsete võimete uuringu“ viisid läbi Tartu ülikooli psühholoogia instituudi magistrandid Liina Adov, Mari-Liis Mägi ning käesoleva töö autor (Mägi, Adov, Täht, & Must, 2013). Uuringusse olid kaasatud esmakursuslased neljateistkümnest Eesti kõrgkoolist.
2. Uuring „Eesti professionaalse muusiku muusikaline identiteet“ viidi läbi koostöös Eesti Muusika- ja Teatriakadeemiaga ning selles osalesid muusikaüliõpilased kolmest kõrgkoolist.
3. Uuring „IKT karjäärivalikud“ viidi läbi koostöös TÜ haridusteaduste instituudiga ning selles osalesid Tartu ülikooli info- ja kommunikatsioonitehnoloogia üliõpilased.
4. Tartu ülikooli sisseastumise kodulehel said huvilised täita nii akadeemilise testi lühiversiooni kui ka täispikka akadeemilist testi, mis oli mõeldud oma võimete proovilepanemiseks ning akadeemilise testi formaadiga tutvumiseks.

Valimi kirjeldus on esitatud alljärgnevas tabelis iga uuringu kohta eraldi (vt tabel 2). Uuritavad kutsuti uuringusse enamasti e-kirja teel. Ainus erand selles osas oli AT lühiversioon Tartu ülikooli sisseastumise kodulehel, mis oli nähtav kõikidele inimestele, kes kodulehte külastasid. AT lühiversioon oli täidetav Tartu ülikooli psühholoogia instituudi testimiskeskonnas (kaemus.psych.ut.ee). Testi täitmiseks oli aega 60 minutit. Peale testi

täitmist paluti täita 10-väiteline küsimustik testi täitmise motivatsiooni kohta. Testitajad ei teadnud, et mõõdetakse nende testi täitmisele kuluvat aega.

Tabel 2

Valimi kirjeldus

Uuring	Andmete kogumise aeg	Lühitesti versioon	Testitajad	N			Keskmine vanus
				Naised	Mehed	Kokku*	
Üliõpilaste enesekohaste hoiakute ja vaimsete võimete uuring	sept.-okt. 2012; veebr.-märts 2013	1.	esmakursuslased	266 73.7%	93 25.8%	361	21.74 (SD = 2.48)
Eesti professionaalse muusiku muusikaline identiteet	okt. 2013-veebr. 2014	1.	muusikaüliõpilased	32 80%	7 17.5%	40	23.95 (SD = 3.74)
IKT karjäärivalikud	dets. 2013-märts 2014	1.	info- ja kommunikatsiooni-tehnoloogia üliõpilased	39 39%	60 60%	100	23.39 (SD = 4.35)
Akadeemiline test TÜ kodulehel	jaan.-veebr. 2013	2.	potentsiaalsed Tartu ülikooli sisseastujad	158 74.2%	52 24.4%	213	22.53 (SD = 4.55)
KOKKU				495 69.2%	212 29.7%	714	22.32 (SD = 3.61)

Märkus: Kuna kõik vastajad ei ole oma sugu märkinud, siis vastajaid kokku on mõnevõrra rohkem kui on meesvastajate ja naisvastajate summa.

Naisvastajate osakaal on käesoleva töö koguvälis veidi suurem kui Eesti kõrghariduses üldiselt, kus vastuvõetud üliõpilastest ligikaudu 60% on naissoost (Tõnisson, 2011). Käesoleva valimi puhul on naiste osakaal 69.2%.

Admetest eemaldasid testitajad, kus test oli avatud, aga üheleki küsimusele ei olnud vastust antud. Andmeanalüüsiks kasutasin statistikaprogrammi SPSS Statistics 19 ja Mplus 7.2. Andmete analüüsimiseks kasutasin korrelatsioonanalüüsi, keskmiste võrdlemisel mitteparameetrilist Kruskal-Wallis ja Mann-Whitney testi ning mudeli loomisel teeanalüüsi (ingl *path analysis*).

Tulemused

Kuna info riigieksamite tulemuste kohta on olemas vaid uuringust „Üliõpilaste enesekohased hoiakud ja vaimsed võimed“, on tulemuste osa jagatud kaheks eraldi uuringuks, lähtuvalt sellest, milliseid andmeid analüüsid kasutatakse. Uuringus 1 kasutan andmeid kõigist neljast uuringust (N = 714). Uuringus 2 kasutan vaid neid andmeid, kus on olemas ka info riigieksamite tulemuste kohta (N = 361).

Uuring 1

Peamised seosed motivatsiooninäitajate ja testitulemuste vahel

Kõikide uuringute puhul võis märgata seda, et mitte kõik testitajad ei täitnud vaimse võimekuse testi algusest lõpuni, vaid olid selle mingil hetkel katkestanud. Järgnevas tabelis on toodud testitajate protsendid, kes testi lõpuni ei täitnud (vt tabel 3).

Tabel 3

Testitajate katkestamised (%)

Uuring	Katkestajate %
Üliõpilaste enesekohaste hoiakute ja vaimsete võimete uuring	15.8
Eesti professionaalse muusiku muusikaline identiteet	2.5
IKT karjäärivalikud	7.0
Akadeemiline test TÜ kodulehel	18.8
Kokku	14.7

Tabelist on näha, et katkestajate osakaal jääb uuringute lõikes vahemikku 2.5–18.8%. Sellist küllaltki suurt varieeruvust võib seletada sellega, et uuring muusikaüliõpilaste ja IKT üliõpilaste hulgas oli lähenemiselt personaalsem ning uuringu läbiviijatel oli võimalusi üliõpilastele meelde tuletada, et nad uuringust osa võtaksid. Teises kahes uuringus olid uuringus osalejad uurijatele tundmatud ning seetõttu ei tundnud nad väljastpoolt arvatavasti vähimatki survet testide lõpuni täitmiseks.

Alljärgnevas tabelis on toodud testitulemuste keskmised ning motivatsiooninäitajad kõigi nelja uuringu kohta (vt tabel 4).

Tabel 4

Vaimse võimekuse testi keskmised tulemused ning motivatsiooninäitajad

Uuring*	Testi tulemused				TESTI TÄITM. AEG (minutites) (max=60)	APV (max=1)	MOT. OLULISUS (max=25)	MOT. PINGUTUS (max=25)
	SÕNAVARA (max=15)	MATEM. (max=15)	RUUM (max=15)	KOKKU (max=45)	M	M	M	M
1	8.45	8.16	7.14	23.76	33.33	.86	17.39	14.84
	SD = 3.05	SD = 4.75	SD = 4.46	SD = 10.65	SD = 15.28	SD = .25	SD = 3.70	SD = 5.19
	N = 361	N = 361	N = 361	N = 361	N = 361	N = 361	N = 300	N = 300
2	9.13	8.15	7.83	25.10	46.41	.93	16.22	16.96
	SD = 2.88	SD = 4.39	SD = 3.96	SD = 9.74	SD = 13.39	SD = .15	SD = 4.35	SD = 5.24
	N = 40	N = 40	N = 40	N = 40	N = 40	N = 40	N = 37	N = 37
3	9.25	11.99	9.75	30.99	46.28	.93	18.15	17.28
	SD = 2.86	SD = 3.76	SD = 4.12	SD = 9.50	SD = 13.60	SD = .19	SD = 3.51	SD = 4.23
	N = 100	N = 100	N = 100	N = 100	N = 100	N = 100	N = 88	N = 88
4	7.98	7.20	5.69	20.87	36.42	.84	20.16	16.37
	SD = 3.19	SD = 4.55	SD = 4.34	SD = 10.39	SD = 16.74	SD = .26	SD = 3.46	SD = 4.76
	N = 213	N = 213	N = 213	N = 213	N = 213	N = 213	N = 107	N = 107
Kõik uuringud koos	8.46	8.41	7.11	23.98	36.80	.87	17.99	15.70
	SD = 3.08	SD = 4.78	SD = 4.53	SD = 10.81	SD = 16.14	SD = .24	SD = 3.85	SD = 5.05
	N = 714	N = 714	N = 714	N = 714	N = 714	N = 714	N = 532	N = 532

*1 – Üliõpilaste enesekohaste hoiakute ja vaimsete võimete uuring; 2 – Eesti professionaalse muusiku muusikaline identiteet; 3 – IKT karjäärivalikud; 4 – Akadeemiline test TÜ kodulehel

Võrreldes teistes uuringutes osalenutega on kõige kõrgema keskmise koguskoori AT lühiversioonis saavutanud IKT üliõpilased. Ühtlasi on neil kõige kõrgem keskmine skoor motivatsiooniküsimustiku pingutuse allskaalal ning APV-indeksi näol (sama kõrge APV-indeks on ka muusikatudengitel). IKT ja muusikatudengid kulutasid testi täitmiseks ka kõige rohkem aega. Muusikatudengite keskmine tulemus jääb küll alla IKT üliõpilaste omale, kuid on kõrgem võrreldes ülejäänud osalejatega. Üldiselt on kõige kõrgem tulemus saadud sõnavara alltestis ja madalaim ruumilise mõtlemise alltestis. Erandiks on ainult IKT üliõpilased, kes on kõrgeima tulemuse saanud matemaatika alltestis ning madalaima sõnavara alltestis. Kõige olulisemaks on vaimse võimekuse testi pidanud ülikooli kodulehel testi täitnud, samas pingutuse skoor on neil jäänud tunduvalt madalamaks ning ka keskmine testiskoor on ülejäänutega võrreldes kõige madalam.

Nägemaks, kas erinevused on ka statistiliselt olulised, viisin läbi sõltumatute gruppide võrdluse. Kuna andmed ei jaotu normaaljaotuslikult, valisin selleks Kruskal-Wallise mitteparameetrilise astakute võrdlemise testi. Leidsin, et gruppide vahel on olulised erinevused nii testi koguskoori ($H(3) = 54.5$, $p < .001$), motivatsiooni pingutuse allskaala skoori ($H(3) = 21.6$, $p < .001$) kui ka APV-indeksi ($H(3) = 19.9$, $p < .001$) osas. Leidmaks, kus täpsemalt erinevused ilmnevad, viisin läbi paarikaupa Mann-Whitney testid (vt tabel 5). Kuna Mann-Whitney testiga on tehtud kuus eraldi võrdlust, on olulisustõenäosus .05 jagatud kuuega, st tulemused saab lugeda statistiliselt oluliseks, kui $p < .008$. Kui eelnevalt nägime, et IKT üliõpilaste testitulemused on teistest kõrgemad, siis nüüd näeme, et need erinevused on ka statistiliselt olulised. Võrreldes esmakursuslastega on statistiliselt oluliselt kõrgem ka nende skoor testimotivatsiooni küsimustiku pingutuse allskaalal ning APV-indeks. APV-indeksi puhul on oluline ka erinevus potentsiaalsete sisseastujate grupist. Kuigi kõik vaadatud erinevused ei ole statistiliselt olulised, võib oletada, et erinevus skoorides võib olla vähemalt osaliselt tulnud erinevusest motivatsioonis.

Tabel 5

Alavalimite keskmiste AT lühiversiooni tulemuste omavahelised võrdlused paarikaupa (Mann-Whitney test)

		Koguskoor		Mot. Pingutus		APV	
		U	p	U	p	U	p
Esmakurs.	IKT	10732.0	.000	9507.5	.000	13665.0	.000
	Muusika	6801.5	.547	4209.5	.016	6783.5	.502
	Sisseastujad	32101.5	.001	13345.5	.010	37153.5	.473
IKT	Muusika	1290.0	.001	1605.5	.901	1574.0	.019
	Sisseastujad	4852.5	.000	4074.5	.105	7809.5	.000
Sisseastujad	Muusika	3261.5	.019	1806.0	.426	3842.0	.295

Märkus: Halliga on märgitud statistiliselt olulised erinevused ($p < .008$).

Korrelatsioonid kõikide alltestide ning motivatsiooniliste näitajate vahel on esitatud tabelis 6. Tabelist on näha, et motivatsiooninäitajatest korreleeruvad testi koguskooriga kõige tugevamini APV-indeks ning motivatsiooniküsimustiku pingutuse allskaala. Mõnevõrra tugevamas korrelatsioonis koguskooriga on keskmine riigieksamitulemus. Motivatsiooninäitajad on riigieksami tulemusega nõrgas korrelatsioonis, mis viitab sellele, et motivatsiooni seos testiskooriga ei ole vahendatud riigieksami tulemuse kaudu.

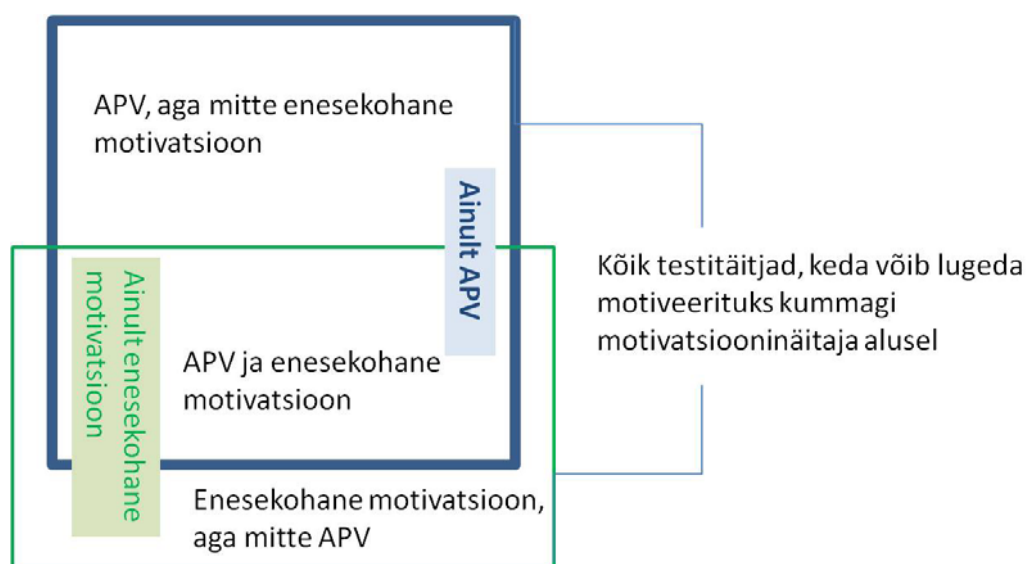
Tabel 6*Korrelatsioonid testitulemuste ja motivatsiooniliste näitajate vahel*

	KOGU- SKOOR	SÕNAVARA	MATEM.	RUUM	KOGU AEG	APV- INDEKS	MOT. OLULISUS	MOT. PINGUTUS	RE KESKMINE
KOGUSKOOR		.768** N = 609	.901** N = 609	.883** N = 609	.487** N = 609	.485** N = 609	.168** N = 519	.326** N = 519	.565** N = 275
SÕNAVARA	.746** N = 714		.546** N = 609	.530** N = 609	.314** N = 609	.301** N = 609	.069 N = 519	.170** N = 519	.520** N = 275
MATEM.	.925** N = 714	.560** N = 714		.698** N = 609	.517** N = 609	.411** N = 609	.206** N = 519	.326** N = 519	.496** N = 275
RUUM	.905** N = 714	.511** N = 714	.772** N = 714		.389** N = 609	.511** N = 609	.132** N = 519	.313** N = 519	.426** N = 275
KOGU AEG	.677** N = 714	.412** N = 714	.685** N = 714	.613** N = 714		.485** N = 609	.284** N = 519	.374** N = 519	.172** N = 275
APV-INDEKS	.703** N = 714	.404** N = 714	.650** N = 714	.719** N = 714	.738** N = 714		.189** N = 519	.144** N = 519	.148* N = 275
MOT. OLUL.	.142** N = 532	.058 N = 532	.179** N = 532	.110* N = 532	.244** N = 532	.117** N = 532		.459** N = 519	.078 N = 275
MOT. PINGUT.	.296** N = 532	.157** N = 532	.299** N = 532	.281** N = 532	.335** N = 532	.095* N = 532	.464** N = 532		.127* N = 275
RIIGIEKSAMI KESKMINE	.510** N = 300	.517** N = 300	.457** N = 300	.385** N = 300	.236** N = 300	.188** N = 300	.076 N = 260	.129* N = 260	

Märkus: Valgel taustal on arvestatud kõikide testitajate tulemusi, hallil taustal ainult testi lõpuni täitnute tulemusi. ** p < .001, * p < .05

Motivatsiooni filtreerimine kahe erineva meetodiga

Järgnevalt vaatan, kuidas eristuvad motiveeritud testitajad mittemotiveeritustest APV-indeksi ja motivatsiooniküsimustiku pingutuse allskaala alusel. APV-indeksi puhul loen sarnaselt teiste autoritega motiveerituteks need, kelle APV-indeks on suurem või võrdne 0.9ga (max 1). Pingutuse allskaala puhul loen motiveerituteks need, kelle skoor on suurem või võrdne 15ga (max 25). Tabelis on toodud lisaks need, keda kehtestatud kriteeriumite alusel saab lugeda motiveerituks APV-indeksi, aga mitte motivatsiooniküsimustiku alusel ja vastupidi; samuti need, kelle puhul APV ja küsimustiku kriteeriumid ühtivad (vt tabel 7). Tabelist näeme, millised oleksid saadud tulemused, kui filtreeriksime andmeid kas ühe, teise või mõlema meetodiga korraga. Lisaks näeme, millised on nende gruppide tulemused, kes võivad olenevalt meetodi valikust analüüsist välja jääda. Tabeli selgituseks on lisatud joonis 1, millel on näha, kuidas tekivad tabelis toodud jaotused.



Joonis 1. Motiveeritute jaotus kahe erineva motivatsiooni filtreerimise meetodi alusel

Tabel 7*Motivatsiooni filtreerimise meetodite võrdlus*

	N	Keskmine koguskoor	SD	Min	Max	Keskmine APV-indeks	SD	Min	Max	Keskmine mot. pingutus skoor	SD	Min	Max
APV \geq 0,9	464	28.61	8.31	9	45	.9901	.0192	.91	1.00	15.99	4.83	5	25
Mot \geq 15	316	29.35	8.91	1	45	.9703	.1071	.07	1.00	19.10	2.82	15	25
APV, aga mitte mot	168	25.68	7.69	9	41	.9865	.0224	.91	1.00	10.61	2.35	5	14
Mot, aga mitte APV	20	15.65	7.84	1	32	.6467	.2613	.07	.89	19.95	3.24	15	25
Mot ja APV mõlemad	296	30.27	8.20	10	45	.9922	.0168	.91	1.00	19.05	2.78	15	25
Kõik tulemused*	519	27.25	9.11	1	45	.9591	.1158	.07	1.00	15.68	5.07	5	25

*Arvestatud on testi lõpuni täitnute tulemusi, mille puhul on täidetud ka motivatsiooniküsimustik.

APV-indeksi alusel saab motiveerituks lugeda 519 testitäitjast 464. Motivatsiooniküsimustiku pingutuse allskaala alusel saab motiveerituks lugeda 316 testitäitjat. Mõlemale kriteeriumile korruga vastab 296 testitäitjat, mis näitab ühest küljest seda, et pingutuse allskaala selline (≥ 15) kriteerium on rangem kui APV-indeksi oma, st selle alusel loetakse rohkem testitäitjaid mittemotiveerituks. Näeme ka seda, et kaks meetodit ei ole omavahel mingis osas kooskõlas – valimis on 168 inimest, keda saab lugeda motiveerituks APV-indeksi, aga mitte pingutuse allskaala põhjal, ning 20 inimest, keda saab lugeda motiveerituks küsimustiku alusel, aga mitte APV-indeksi alusel. Kõige kõrgem keskmine skoor on grupis, kus saab testitäitjaid lugeda motiveerituks mõlema meetodi kriteeriumite alusel. Kõige madalam keskmine skoor on grupis, mille liikmeid saab lugeda motiveerituks küsimustiku, aga mitte APV-indeksi alusel, mis võib tähendada seda, et nende testitäitjate puhul ei pruugi motivatsiooniküsimustikule antud vastused peegeldada tegelikku pingutust testi täitmisel. Ka saadud maksimaalne skoor on selles grupis kõige madalam. Samas võib siin olla ka juhtusid, kus inimene vastas motiveeritult küll osale küsimustest, kuid mingil hetkel lülitus ümber liigkiirele vastamisele. Juhtudel, kus inimest saab lugeda motiveerituks APV, aga mitte küsimustiku alusel, peab arvestama seda, et inimene, kes tegelikult võis testile küll aega pühendada, kuid tundis, et tal ei läinud väga hästi, võis küsimustikus väita, et ta ei pingutanudki. Sellisel juhul ei loetaks teda küsimustiku tulemuse alusel motiveerituks, küll aga APV alusel.

Sugudevahelised erinevused motivatsioonis

Kuna käesolevas uuringus on ühe meetodina kasutatud enesekohast küsimustikku, mille vastused võivad olla mõjutatud sotsiaalselt soovitatavast vastamisest, uurin, kas on soolisi erinevusi motivatsioonis ning motivatsiooniküsimustikule vastamisel. Üks erinevus, mis võib osaliselt ka motiveeritust peegeldada, on see, et mehi osales kolmes uuringus neljast vähem kui naisi (ka siis kui arvestada naiste ja meeste osakaalu kõrghariduses üldiselt).

Võrreldes meeste ja naiste tulemusi vaimse võimekuse testis, on meeste tulemused ligi kuue punkti võrra kõrgemad (efekti suurus Coheni $d = .75$; arvestatud on neid, kelle kohta on olemas ka motivatsiooniküsimustiku skoor). Tulemuste erinevus on statistiliselt oluline ($U = 17100.5$, $p < .001$). Samas on meeste puhul kõrgem ka APV-indeks, kuigi see erinevus ei ole statistiliselt oluline ($U = 26709.5$, $p = .144$), ja motivatsiooniküsimustiku pingutuse alltesti skoor ($U = 22076.5$, $p < .001$).

Motivatsiooniküsimustiku pingutuse allskaala skooride sagedusjaotusest on näha, et kui meeste puhul sarnaneb see normaaljaotusele, siis naiste puhul on tegemist bimodaalse

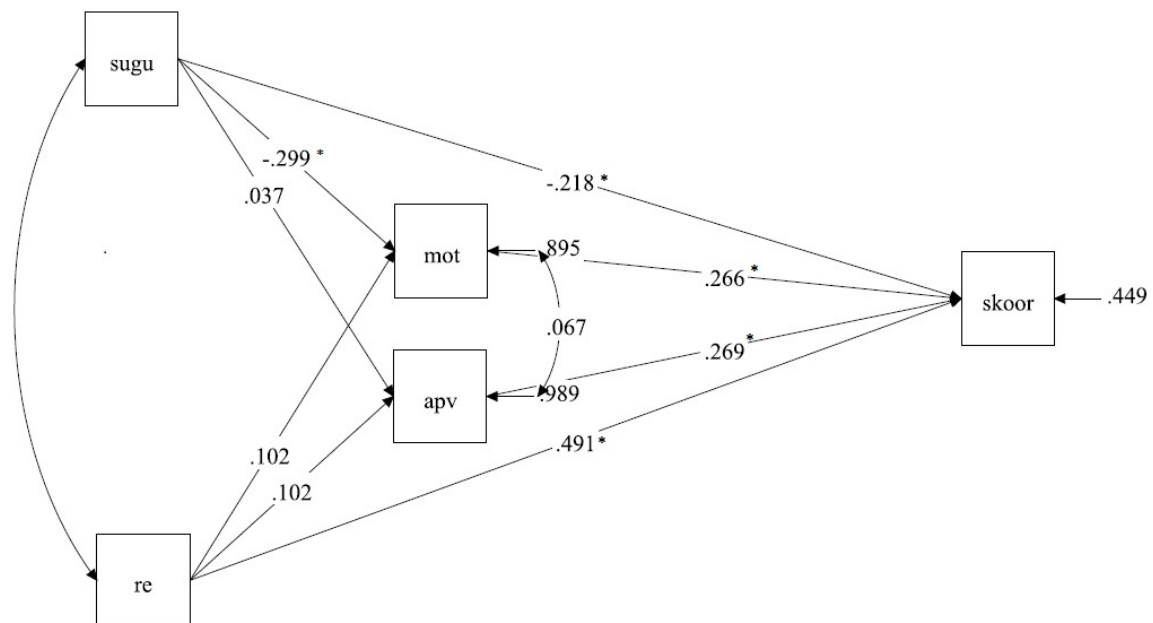
jaotusega. Taoline jaotus võib viidata sellele, et naised vastavad motivatsiooniküsimustikule teistmoodi kui mehed. Samas selget põhjendust, miks taoline jaotus tekib, hetkel ei ole ning see vajaks edasist uurimist ja kontrollimist ka teistel valimitel. APV-indeksi puhul sellist jaotust ei ilmne, mis viitab sellele, et jaotus ei tulene motivatsioonist enesest, vaid pigem küsimustikule vastamise stiilist. Ka korrelatsioon pingutuse allskaala skooride ning testitulemuste vahel on naistel madalam kui meestel (vastavalt $r = .262$, $p < .001$, $r = .354$, $p < .001$)

Uuring 2

Mudel testitulemuste ennustamiseks

Uurimaks, kuidas ennustavad motivatsiooninäitajad vaimse võimekuse näitaja kõrval vaimse võimekuse testi tulemust, vaatasin seda läbi struktuuraalse modelleerimise meetodi, teeanalüüsi. Võrreldes erinevaid mudeleid, kus motivatsiooninäitajad on sees, kas mõlemad koos või ükshaaval, on näha, kas need käituvad sarnaselt või erinevalt ning milline on nende omavaheline dünaamika.

Kuna osa testitajate puhul olid olemas ka andmed nende riigieksamite tulemuste kohta (uuringust „Üliõpilaste enesekohased hoiakud ja vaimsed võimed“), lisisin riigieksamite (RE) keskmise tulemuste ühte mudelisse koos testi täitmise motivatsiooni näitajate ning sooga. Testi täitmise motivatsioon ning riigieksamite keskmine tulemus ei olnud omavahel märkimisväärses korrelatsioonis (RE ja APV-indeks: $r = .149$, $p < .05$; RE ja pingutus: $r = .128$, $p < .05$). Mudel, mida kontrollisin, on toodud alljärgneval joonisel koos regressioonikordajate ja regressioonijääkidega (vt joonis 2). Tärnidega on tähistatud need regressioonikordajad, mis on statistiliselt olulised.



Joonis 2. Mudel testiskoori ennustamiseks soo, riigeksami tulemuse ja motivatsiooninäitajate kaudu (N = 273). Märkused: Tärnidega on tähistatud regressioonikordajad, mis on statistiliselt olulised. Sugu on kodeeritud: mees – 0, naine – 1.

Analüüsi tulemused näitavad, et sugu ennustab teataval määral nii enesekohaselt teatatud motivatsiooni (pingutuse allskaala) kui ka testiskoori – meestel on keskmiselt kõrgem testiskoor ning ka enesekohaselt teatatud motivatsioon. Samas APV-indeksi suurusele sool märkimisväärset mõju ei ole. Keskmise riigeksamitulemus ennustab testiskoori, kuid mitte kumbagi motivatsiooninäitajat. Kaks motivatsiooninäitajat ennustavad testiskoori väga sarnaselt. Kogu mudel ennustab testitulemuse varieeruvusest 55.1% ($R^2 = .551$, $p < .001$). Kui jätta mudelisse sisse vaid statistiliselt olulised seosed (märgitud joonisel tärnidega), ennustaks selline mudel 53.4% testitulemuse varieeruvusest ($p < .001$).

Lisaks testisin sarnaseid mudeleid, kuhu jätsin ainult ühe kahest motivatsiooninäitajast (APV-indeks, enesekohane motivatsioon). Mõlemad mudelid tulid oma seletusvõimelt suhteliselt sarnased. Mudel, kus testiskoori ennustajateks on sugu, testimotivatsiooni skoor ning riigeksami tulemus, ennustab 48% ($p < .001$) testiskoori varieeruvusest. Mudel, kus ennustajateks on APV-indeks, sugu ja keskmine riigeksamitulemus ennustab 48.8% ($p < .001$) testiskoori varieeruvusest.

Arutelu

Käesolevas töös püüdsin vastata kahele küsimusele. Esiteks, kuidas erinevad omavahel kaks motivatsiooni filtreerimise meetodit, millest üks põhineb enesekohasel raportil ning teine testi küsimustele vastamiseks kulunud aegade analüüsil. Teiseks küsisin, kuidas ennustavad need filtreerimiseks kasutatavad motivatsiooninäitajad vaimse võimekuse testi tulemust mudelis, kuhu on lisatud ka sugu ja keskmine riigieksamitulemus, mida käsitlesin tinglikult vaimse võimekuse näitajana.

Antud uuringu puhul oli eelduseks see, et vaimse võimekuse test (akadeemilise testi lühiversioon), mida palusime inimestel täita, oli enamiku inimeste jaoks madala olulisusega, millega omakorda võis kaasned madal motivatsioon testi täita ning oma võimeid adekvaatselt näidata. Madalale motivatsioonile viitab uuringu tulemustes kõigepealt juba testi katkestajate osakaal. Märgata võis ka seda, et testi katkestajate osakaal oli väiksem, kui uuringu läbiviijad saatsid potentsiaalsetele uuringus osalejatele meeldetuletusi uuringus osalemiseks – ehk kontakt uurijate ja uuritavate vahel oli suurem. Motivatsiooninäitajate uurimiseks ja filtreerimismeetodite võrdlemiseks eemaldasin andmehulgast motivatsiooni filtreerimise meetodite abil tõenäoliselt mittemotiveeritute vastused. Sarnaselt varasemate uuringutega (nt Wise & Kong, 2005) leidsin, et nii motivatsiooniküsimustiku skoori kui ka APV-indeksi alusel filtreeritud andmestiku puhul on keskmised skoorid oluliselt kõrgemad kui filtreerimata andmestiku puhul.

Kui võrrelda kahte kirjeldatud motivatsiooninäitajat, siis sarnaselt Swerdzewski, Harnes ja Finney (2011) uuringule leidsin, et nendest rangem on enesekohasel raportil põhinev, kuna selle alusel tuleb valimist välja filtreerida rohkem vastajaid. APV-indeksi puhul loetakse mittemotiveerituks vähem testitajaid ning selle alusel filtreeritud andmestiku keskmine koguskoor ei ole oluliselt madalam võrreldes motivatsiooniküsimustiku alusel filtreerituga. Kõige kõrgem keskmine skoor saadakse, kui arvestada filtreerimisel mõlemat näitajat korraga. Kuid sellisel juhul tuleb andmestikust välja arvata pea poolte testitajate tulemused, mis ei pruugi olla väga praktiline. Samuti ei ole alati võimalik kahte meetodit korraga kasutada (näiteks kui testi ei täideta arvutis).

Lisaks peab filtreerimisel arvestama, et motivatsiooniküsimustikule vastates võisid testitajad olla mõjutatud juba testi täitmise kogemusest, ehk kui testitaja tundis, et test ei läinud hästi, võis ta väita, et ei olnudki motiveeritud seda täitma, isegi kui ta tegelikult pingutas. Seega võib järeldada, et APV alusel filtreerimine võimaldab valimisse rohkem indiviide alles jätta ning ühtlasi ei ole tulemus mõjutatud sotsiaalselt soovitatavast või

testikogemusest kallutatud vastamisest, mis on ohuks enesekohase raporti puhul. Ka Wise ja Kong (2005) leidsid, et need kaks motivatsiooni filtreerimise aluseks olevat meetodit võivad peegeldada mõnevõrra erinevaid konstrukte ning et APV puhul olid konvergente ja diskriminantne valiidsus mõnevõrra kõrgemad. Käesolevas uuringus viitab sellele asjaolu, et enamik testitajaid, keda saab lugeda motiveerituks motivatsiooniküsimustiku alusel, on motiveeritud ka APV-indeksist lähtudes, aga mitte vastupidi. Viimane võibki tähendada seda, et kuigi inimene on objektiivselt kulutanud piisavalt aega testi küsimustele vastamiseks, võib ta vastata, et ei olnud tegelikult motiveeritud seda täitma. Teisest küljest võib see viidata sellele, et APV-indeksi arvutamiseks seatud ajalised läved on liiga madalad, mis tähendab, et kuigi inimene võis kõik ülesanded läbi vaadata ja nendele ka vastata, ei pruugi see näidata tema motivatsiooni seda hästi või põhjalikult teha. Läve määramisel on aluseks võetud aeg, mis eeldatavasti kulub vähemalt ülesande läbilugemiseks, aga mitte otseselt selle lahendamiseks. Suures osas siiski kattuvad need, keda saab motiveerituteks lugeda motivatsiooniskoori alusel, nendega, keda saab lugeda motiveerituteks APV-indeksi alusel, mis omakorda näitab seda, et enamasti oskavad inimesed ise oma motivatsiooni ja pingutust hinnata.

Enesekohase raporti võimalikule kallutatusele tulemustes viitab ka see, et naised ja mehed paistavad sellele erinevalt vastavat. Nimelt, kui vaadata motivatsiooniskooride sagedusjaotust, ilmneb naiste puhul eeldatava normaaljaotuse asemel, erinevalt meestest, bimodaalne skooride jaotus, mille põhjusi üheselt seletada on keeruline. Varasematest uuringutest on teada, et just naised on alimad sotsiaalselt soovitava vastamisele, mis võib olla üks taolise erinevuse põhjustest (vt Bernardi, 2006; Dalton & Ortegren, 2011). Lisaks oli näha, et korrelatsioon motivatsiooniskoori ja testitulemuse vahel oli naiste puhul nõrgem kui meeste puhul. Ka teeanalüüsi mudelist oli näha, et sool on seos enesekohase motivatsiooniskooriga, aga mitte APV-näitajaga.

Sooliste erinevuste poolelt ilmnnes veel, et meeste puhul oli kõrgem nii keskmine testiskoor kui ka motivatsiooninäitajad. Ühtlasi oli meessoost testitajaid pea poole vähem kui naissoost testitajaid. Kuna tegemist oli vabatahtlikult täidetava testiga, võib see viidata sellele, et kui mehed tunnevad, et nad ei ole motiveeritud testi täitma, siis nad ei asugi seda tegema. See-eest naised võivad asuda testi täitma ka siis, kui nad ei ole piisavalt motiveeritud, et kogu testi täitmise ajal pingutada. Tulemus on näilises vastuolus varasemalt leitud tulemusega (Freund & Rock, 1992), mis väitis, et äärmiselt madal motivatsioon on tavalisem meeste kui naiste puhul. Samas kui arvestada, et mehi asus testi täitma tunduvalt vähem kui naisi, võib ka seda lugeda äärmiselt madala motivatsiooni näitajaks ning sel juhul ei ole

vastuolusid ka varasemalt leitud tulemustega. Pigem on see kooskõlas Segali (2006) „kaalutlejate“ ja „pingutajate“ jaotusega, kus näidati, et rohkem mehi kui naisi parandab oma sooritust, kui neile pakutakse selle eest tasu, ehk et meeste motivatsioon on suuremal määral sõltuv tegevusest saadavast kasust – kui testi täitmises ei nähta enda jaoks olulist väärtust, ei hakatagi seda täitma.

Üheks käesoleva töö olulisemaks tulemuseks võib lugeda koostatud teeanalüüsi mudelit, mis aitas ilmestada nii testimotivatsiooni olemust üldiselt kui ka täpsemalt kahe erineva motivatsiooninäitaja näol. Kui võrrelda erinevaid mudeleid, kus testiskoori ennustajateks on sugu, keskmine riigieksamitulemus ja kumbki motivatsiooninäitaja, siis testi tulemuse varieeruvusest suurima osa kirjeldab mudel, kus on sees mõlemad motivatsiooninäitajad korraga, mis võib tähendada seda, et vähemalt mingis osas kirjeldavad need natuke erinevaid asju ja täiendavad teineteist. Kui võtta testiskoori mudelisse ainult üks motivatsiooninäitaja korraga, on mudelite kirjeldusvõime suhteliselt sarnane. Mõnevõrra parem on mudel, milles on prediktorina APV-indeks. Mudel, kus testiskoori ennustajateks on võetud nii motivatsiooninäitajad kui ka riigieksamite keskmine tulemus vaimse võimekuse indikaatorina ning sugu, ennustab üle 50% testiskoori varieeruvusest. Näha oli ka see, et riigieksamite keskmine tulemus ning motivatsiooninäitajad ei ole omavahel märkimisväärse korrelatsioonis, mis näitab seda, et saadud motivatsiooninäitajad peegeldavad vaimsest võimekusest eraldiseisvat konstrukti ning et mõlemal on testiskoori kujunemises oma roll.

Varasemalt on leitud, et kaks vaimset võimekust peegeldavat konstrukti – vaimse võimekuse testi tulemus ja koolihinded – kirjeldavad üksteise hajuvusest 25%. On arvatud, et kirjeldamata osa jääb motivatsiooni, isiksuseomaduste jmt arvele (Neisser jt., 1996). Käesolevas töös saadud mudeli kirjeldusvõime üle 50% on seega üpriski märkimisväärne. Oma töös ennustasin ühe vaimse võimekuse näitaja kaudu teist, lisades sinna ka motivatsiooninäitajad. Esimese vaimse võimekuse näitaja puhul, milleks oli keskmine riigieksamitulemus, eeldasin, et see on olnud testitajate jaoks kõrge olulisusega ning seetõttu peegeldab ka suuremal määral nende tegelikku vaimset võimekust. Teiseks oli aga vabatahtlikult täidetav vaimse võimekuse test, mille täitmise eest testitajatele ühtegi tasu ega olulist tagajärge ei kaasnud. Nägime, et kui lisada mudelisse peale vaimse võimekuse näitaja ka motivatsiooninäitajad, saame kaks korda suurema kirjeldusvõimega mudeli kui on leitud vaimse võimekuse testi tulemuse ja koolihinnete vahelise seose puhul.

Käesolevas töös on vajalik välja tuua ka mõningad piirangud, millega on tarvis edaspidi arvestada ja neile lahendusi otsida. Kasutasin kahte erinevat motivatsiooni filtreerimise meetodit, millel kummalgi on oma puudused. APV-indeksi peamine piirang on

see, et seda saab kasutada ainult arvutis täidetavate testide puhul. Motivatsiooniküsimustiku puuduseks on aga see, et selle tulemus võib olla mõjutatud nii sotsiaalselt soovitatavast vastamisest kui ka vahetult eelnenud testi täitmise kogemusest. Kummagi meetodi puhul ei ole täielikult selge ka filtreerimise aluseks olevate äralõikepunktide kehtestamine, kuna mingil määral jääb see alati subjektiivseks. Kuigi APV-indeksi puhul on võimalik lähtuda konkreetse küsimuse pikkusest, iseloomust ja vastuste sagedusjaotusest, on väga raske hinnata inimeste kiirust ülesannete lahendamisel. Selleks, et ei tekiks olukorda, kus väga kiirete vastajate vastused loetakse liigkiireteks, jääb äralõikepunkt pigem alahinnanguks, millega suudame veenduda vaid selles, et küsimuse jaoks on võetud umbkaudu piisavalt aega, et küsimus läbi lugeda, kuid mitte selles, kuivõrd inimene vastamisse ikkagi süvenes. APV-indeksi puhul ei ole ohtu, et näitaja võiks olla mitte-objektiivne, kuna see põhineb vastamiseks kasutatud aegadel ja reeglina testitajad ei tea, et testi täitmise aeg võetakse vaatluse alla, ning sellest tulenevalt ei oska nad sellega manipuleerida, et ennast paremas valguses näidata. APV-indeksi kasutamist saaks tõhustada, kui leitaks paremad kriteeriumid, mille alusel määrata lävesid liigkiirete vastuste eristamiseks läbimõeldud vastustest.

Motivatsiooniküsimustiku puhul on äralõikepunkti veelgi keerulisem kehtestada – kuna tegemist on subjektiivse vahendiga, on ka äralõikepunkt subjektiivselt määratud. Nimelt on varasemad autorid (nt Swerdzewski, Harmes, & Finney, 2011) lähtunud sellest, et motiveerituks saab lugeda need, kes viiepallisel skaalal on hinnanud oma motivatsiooni üle keskmise, st vähemalt kolme palliga viiest.

Lisaks tasub meelde tuletada ka seda, et kuigi küsimustikust saadud motivatsiooniskoori on hakatud kasutama ka motivatsiooni filtreerimiseks, ei olnud see algselt selleks ette nähtud. See ei olnud mõelnud kasutamiseks individuaalsel tasandil, vaid keskmise skoorina valimi iseloomustamiseks (Sundre & Thelk, 2007).

Käesoleva uuringu üheks miinuseks on suhteliselt vähese informatsiooni olemasolu testitavate vaimse võimekuse kohta – vaid ühest uuringust olid olemas testitajate riigieksamite keskmised tulemused, mida võib tinglikult selle indikaatoriteks lugeda. Seetõttu oleks hea, kui ka edaspidi saaks vaadata testimotivatsiooni näitajaid seoses vaimse võimekuse indikaatoritega, samuti väärivad edasist uurimist soolised erinevused testi täitmise motivatsioonis ning nende võimalikud põhjused.

Olulisim järeldus käesolevat tööst on, et testi täitmise motivatsioon on vaimsest võimekusest eraldiseisev konstrukt, mida on oluline arvesse võtta, kui viiakse läbi testimist testitajate jaoks madala olulisusega situatsioonis. Motivatsiooniga arvestamine on oluline uuringutulemuste tõlgendamise valiidsuse huvides. Üks võimalus motivatsiooniga

arvestamiseks on püüda seda tõsta testimisele eelnevalt. Ka käesoleva töö valimi pealt on näha, et olenevalt testimise situatsioonist on tunduvalt erinev näiteks testitaitmise katkestamine – mida personaalsem oli lähenemine testitaitjatele, seda vähem oli katkestamist.

Teine võimalus testimotivatsiooniga arvestamiseks on käesolevas töös käsitletud motivatsiooninäitajate kasutamine. Mõlemal meetodil on küll omad piirangud, kuid sellegipoolest võib nende kasutamine anda vajalikku informatsiooni testitulemuste tõlgendamiseks. Käesoleva töö tulemuste põhjal võib öelda, et kahe filtreerimismeetodi võime eristada motiveeritud mittemotiveeritud on suhteliselt sarnane. Sellegipoolest tuleks ettevaatlik olla enesekohase raporti alusel filtreerimisega, kuna selle tulemus võib olla mõjutatud sotsiaalsest soovitatavast vastamisest. See-eest saab selle tulemusi kajastada mitte individuaalsel, vaid üldisel tasandil. Objektiivsema vahendina on vähemalt arvutis täidetavate testide puhul võimalik kasutada APV-indeksit.

Tänuõnad

Soovin tänada oma juhendajaid Olev Mustat ja Karin Tähte asjaliku juhendamise ning toetuse eest. Minu suur tänu meeldiva koostöö eest kuulub ka meie uurimisrühma ülejäänud liikmetele: Liina Adovile, Mari-Liis Mägile ning Dmitri Rozgonjukile. Veel tahan väga tänada Margot Lanemani, kes andis head nõu töö keelelistes ja vormilistes küsimustes.

Viidatud allikad

- Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181.
- Atkinson, J. W. (1964). *An Introduction to Motivation*. Princeton, NJ: Van Nostrand.
- Barry, C. L. (2010). Do examinees have similar test-taking effort? A high-stakes question for low-stakes testing. *International Journal of Testing*, *10*, 342–363.
- Barry, C. L., & Finney, S. J. (2009). Exploring change in test-taking motivation. Paper presented at 2009 annual meeting of the Northeastern Educational Research Association. Veebilehelt kättesaadav seisuga 23. mai 2015 [http://www.psyc.jmu.edu/assessment/research/pdfs/barry_nera09.pdf].
- Bernardi, R. A. (2006). Associations between Hofstede's cultural constructs and social desirability response bias. *Journal of Business Ethics*, *65*, 43–53.
- Boe, E. E., May, H., & Boruch, R. F. (2002). Student task persistence in the Third International Mathematics and Science Study: A major source of achievement differences at the national, classroom, and student levels. (Research Rep. No. 2002-TIMSS). Philadelphia: University of Pennsylvania, Graduate School of Education, Center for Research and Evaluation in Social Policy.
- Bovaird, J. A. (2002). New applications in testing: Using response time to increase the construct validity of a latent trait estimate. (Doctoral dissertation, University of Kansas, 2002). *Dissertation Abstracts International*, *64*, 998.
- Dalton, D., & Ortegren, M. (2011). Gender differences in ethics research: The importance of controlling for the social desirability response bias. *Journal of Business Ethics*, *103*, 73–93.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21.
- DeMars, C. E. (2000). Test stakes and item format interactions. *Applied Measurement in Education*, *13*(1), 55–77.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.
- Eklöf, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educational and Psychological Measurement*, *66*, 643–656.

- Freund, D. S., & Rock, D. A. (1992). A preliminary investigation of pattern-marking in 1990 NAEP data. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. ED 347 189)
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant*, 47, 2025–2047.
- Mislevy, R. J. (1995). What can we learn from international assessments? *Educational Evaluation and Policy Analysis*, 17, 419–437.
- Must, O., & Allik, J. (2002). *Tunne oma võimeid. Abivahend eneseanalüüsiks*. Tartu: Tartu Ülikooli Kirjastus.
- Mägi, M-L., Adov, L., Täht, K., & Must, O. (2013). Who is willing to take low-stake assignments? *Trames: Journal of the Humanities and Social Sciences*, 17(4), 417–432.
- Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... & Urbina, S. (1996). Intelligence: knowns and unknowns. *American psychologist*, 51(2), 77.
- O'Neil, H. F., Sugrue, B., & Baker, E. L. (1995). Effects of motivational interventions on the national assessment of educational progress mathematics performance. *Educational Assessment*, 3(2), 135.
- Pekkarinen, T. (2012). Gender differences in education. *Nordic Economic Policy*, 165.
- Silm, G., Täht, K., & Must, O. (2013). Testi täitmise motivatsiooni mõju testi tulemustele. *Psühholoogia ja Kõrgkool*, 78–94.
- Spinath, B., Freudenthaler, H. H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences*, 48(4), 481–486.
- Steedle, J. T. (2014). Motivation filtering on a multi-institution assessment of general college outcomes. *Applied Measurement in Education*, 27, 58–76.
- Sundre, D. L. (1999). Does examinee motivation moderate the relationship between test consequences and test performance? (Report No. TM029964). Harrisonburg, Virginia: James Madison University. (ERIC Documentation Reproduction Service No. ED432588)
- Sundre, D. L., & Finney S. J. (2002). Enhancing the validity and value of learning assessment: Furthering the development of a motivation scale. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

- Veebilehelt kättesaadav seisuga 23. mai 2015
[http://www.jmu.edu/assessment/wm_library/motcfa.pdf].
- Sundre, D. L., & Kitsantas, A. (2004). An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology*, 29, 6–26.
- Sundre, D. L., & Thelk, A. D. (2007). The Student Opinion Scale (SOS): A Measure of Examinee Motivation: Test Manual.
- Sundre, D. L., & Wise, S. L. (2003). 'Motivation Filtering:' An exploration of the impact of low examinee motivation on the psychometric quality of tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.
- Veebilehelt kättesaadav seisuga 23. mai 2015
[http://www.jmu.edu/assessment/wm_library/Filter.pdf].
- Swerdzewski, P. J., Harnes, J. C., & Finney, S. J. (2011). Two approaches for identifying low-motivated students in a low-stakes assessment context. *Applied Measurement in Education*, 24, 162–188.
- Tõnisson, E. (2011). Kõrghariduse valdkonna statistiline ülevaade. Veebilehelt kättesaadav seisuga 23. mai 2015
[https://dspace.utlib.ee/dspace/bitstream/handle/10062/40767/KHaridus_Stat2011.pdf].
- Wainer, H. W. (1993). Measurement problems. *Journal of Educational Measurement*, 30(1) 1–21.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.
- Wise, S. L. (2006). An investigation of the differential effort received by items on a low-stakes computer-based test. *Applied Measurement in Education*, 19(2), 95–114.
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1–17.
- Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18, 163–183.
- Wolf, L. F., & Smith, J. K. (1993). The effects of motivation and anxiety on test performance. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.
- Wolf, L. F., & Smith, J. K. (1995). The consequence of consequence: Motivation, anxiety, and test performance. *Applied Measurement in Education*, 8(3), 227–242.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Gerli Silm,

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

„*Motivatsiooni roll vaimse võimekuse mõõtmisel ja testitulemuste tõlgendamisel*“, mille juhendajad on Olev Must ja Karin Täht,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **25.05.2015**