

Tartu Ülikool
Sotsiaalteaduste valdkond
Ühiskonnateaduste instituut
Sotsioloogia õppekava

Liina Eamets

INTEGRATSIOONI MONITOOHINGU ERI AASTATE ANDMETE VÕRRELDAVUS -
UURINGUMETOODIKA, ANKEETKÜSIMUSTE JA ANDMETE ESINDUSLIKKUSE
ANALÜÜS

Magistritöö

Juhendaja: Indrek Soidla, MA
Kaasjuhendaja: Maarja Vollmer, MA

Tartu 2022

Olen koostanud töö iseseisvalt. Kõik töös kasutatud teiste autorite tööd, põhimõttelised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

Liina Eamets

18. august 2022

ABSTRACT

Comparability of the Estonian Integration Monitoring survey data from different years: analysis of survey methodology, questionnaire, and data representativeness.

Since 2000, Estonian Integration Monitoring (EIM) survey has been carried out in every 2 or 3 years, altogether 8 times between 2000-2020. The aim of the EIM survey is to analyse integration processes in Estonian society in order to understand and improve the adaption and integration of immigrants and ethnic minorities. Long-term analysis based on the EIM data has been seen as necessary and very useful to monitor trends and changes in integration processes in Estonia. To carry out these kinds of analysis, the EIM data from different years must be comparable. Therefore, the purpose of this master's thesis is to analyse to what extent the methodology, questionnaire, and data representativeness of the EIM survey have changed over the years.

Author analysed the survey methodology based on documentations provided in the survey reports using content analysis. Changes in questionnaires were analysed by comparing the question wording from different. Data representativeness is a complex concept and was analysed in this thesis from the missing data perspective. Author examined how extensive the non-response in different years is, and whether missing data can be considered missing at random (MAR), missing completely at random (MCAR), or not missing at random (NMAR).

Based on the main results it can be said that the EIM could be documented in much greater detail. Lack of detail in the documentation is the main reason that limits the author to draw definitive conclusions on the comparability of the EIM survey data from different years. Analysis showed that during the years the definition of survey population and the sampling strategy have changed. Moreover, due to lack of detail in the documentation, the author questions the randomness of the samples from different years. Survey questionnaires have also changed remarkably during the years due to many different factors including e.g. changes in research teams or changes in the national integration development programs or changes in the way integration as a concept has developed over the years. Although the question wording can vary over the years the analysis

concludes that in some cases the construct that is measured stays the same and therefore the questions can be used in comparative research regardless of these differences in wording. Missing data analyses showed that item non-response may cause loss in data representativeness and errors in data analysis, as the author concluded that missing data cannot be considered to be missing completely at random (MCAR), but this topic still needs some further in-depth analysis.

Keywords: survey methodology, data representativeness, questionnaire design, data comparability

SISUKORD

| | |
|--|----|
| ABSTRACT..... | 3 |
| SISUKORD..... | 5 |
| SISSEJUHATUS | 7 |
| 1. KIRJANDUSE ÜLEVAADE JA PROBLEEMISEADE | 9 |
| 1.1 Küsitlusuuringute võrreldavus | 9 |
| 1.1.1 Tõenäosuslik ja mittetõenäosuslik valim | 10 |
| 1.1.2 Ankeetküsimustike võrreldavus..... | 12 |
| 1.1.4 Uuringu kvaliteet | 15 |
| 1.1.5 Uuringu dokumenteerimine | 21 |
| 1.2 Ülevaade Eesti ühiskonna integratsiooni monitooringust..... | 22 |
| 1.2.1 Integratsiooni monitooringu meetodika (re)kontseptualiseerimine | 24 |
| 1.3 Probleemiseade ja uurimisküsimused | 25 |
| 2. ANDMED JA METOODIKA | 27 |
| 2.1 Kasutatud andmed..... | 27 |
| 2.1.1 Kitsendused..... | 27 |
| 2.2 Metoodika..... | 28 |
| 2.2.1 Kontentanalüüs | 28 |
| 2.2.2 Ankeetküsimuste analüüsi meetodika..... | 29 |
| 2.2.3 Andmete esinduslikkuse analüüsi meetodika | 30 |
| 3. INTEGRATSIOONI MONITOORINGU ERI AASTATE ANDMETE VÕRRELDAVUSE ANALÜÜS | 32 |

| | |
|---|-----|
| 3.1 Uuringumetoodika analüüs | 32 |
| 3.1.1 Monitooringu läbiviijad ja üldine korraldus | 32 |
| 3.1.2 Küsitlusuuringu üldkogum..... | 34 |
| 3.1.3 Küsitlusuuringute valimid..... | 34 |
| 3.1.4 Küsitlusuuringu vastamismäär ning arvutatud kaalud..... | 38 |
| 3.2 Ankeetküsimuste samasuse analüüs | 39 |
| 3.2.1 Ankeetküsimuste analüüsi üldised tulemused | 40 |
| 3.2.2 Võtmeküsimuste samasuse analüüs | 43 |
| 3.3 Andmete esinduslikkuse analüüs | 59 |
| 3.3.1 Andmelünkade analüüsi üldised tulemused..... | 59 |
| 3.3.2 Andmelünkade juhuslikkuse analüüs..... | 64 |
| 4. ARUTELU | 74 |
| KOKKUVÕTE | 90 |
| KASUTATUD KIRJANDUS..... | 94 |
| LISAD..... | 100 |

SISSEJUHATUS

Statistikaameti andmetel on eestlaste osakaal Eesti rahvastikust viimase 30 aasta jooksul kõikunud ligikaudu 60%-70% vahel (RV0221... i.a., RV021... i.a.). 90ndate algul moodustasid muust rahvusest elanikud pea 40% Eesti rahvastikust, viimased paarkümmend aastat on mitte-eestlaste osakaal püsinud enam-vähem stabiilselt ca 30% juures. Seega on Eesti riik seisnud lõimumisalaste väljakutsete ees juba taasiseseisvumisest alates. Lõimumispoliitika edukus on oluline nii ühiskonna sidususe ja stabiilsuse kui ka riikliku julgeoleku tõttu.

Selleks, et jälgida, kuidas integratsioonipoliitika Eestis on toimunud ning milliseid valikuid edasi teha, on alates 2000. aastast paariaastase intervalliga (kokku kaheksa korda) läbi viidud Eesti ühiskonna integratsiooni monitooringut. Tegemist on laiapõhjalise elanikkonna uuringuga, mis analüüsib Eesti ühiskonna kohanemis- ja lõimumisprotsesse ning aitab mõista nende iseärasusi ja väljakutseid. Monitooringu tulemusi kasutavad nii poliitikakujundajad, teadlased kui ka laiem avalikkus.

Integratsiooni monitooring on üks väheseid Eestis nii pikaajaliselt läbiviidud jätku-uuringuid. Seni ei ole teadaolevalt süstemaatiliselt analüüsinud lõimumismonitooringu eri aastate andmete omavahelist võrreldavust. Üksikute uuringulainete andmeid on monitooringu aruannetes küll omavahel võrreldud, kuid põhjalikku metoodilist võrdlust kõigi aastate lõikes tehtud ei ole. Trendianalüüs 20 aasta andmete pealt annaks võimaluse hinnata muutusi erinevates hoiakutes ja integratsiooni aspektides, kuid selleks peab veenduma, et nende aastate andmed on omavahel võrreldavad. Sellest tulenevalt on käesoleva magistritöö eesmärk välja selgitada, mil määral on lõimumismonitooringu küsitlusmetoodika, ankeetküsimused ning andmete esinduslikkus aastate jooksul muutunud.

Eri aastatel läbi viidud monitooringute võrreldavust saab analüüsida ühelt poolt vaadates kuivõrd on ajas muutunud uuringu sisu, sh ankeetküsimused. Kuna monitooringu teostajad on aastate lõikes olnud erinevad, siis ei piisa ainult ankeetküsimuste muutuste analüüsimisest, vaid lisaks

tuleb vaadata ka uuringu kvaliteeti ehk kuidas on olnud korraldatud andmekogumine ning mil määral esindavad eri aastatel kogutud andmed populatsiooni.

Magistritöö esimeses osas antakse andmete võrreldavuse seisukohtadest lähtudes ülevaade küsitlusuuringu disainimise olulisematest etappidest ning uuringu kvaliteeti määravatest põhialustest. Samuti antakse esimeses peatükis ülevaade integratsiooni monitooringu senisest korraldusest. Töö teises osas kirjeldatakse magistritöös kasutatud andmeid ning meetodikat. Kolmandas osas teostatakse andmete analüüs, eri alapeatükkides analüüsitakse lõimumismonitooringu uurimismetoodikat, ankeetküsimuste samasust ning andmete esinduslikkust (keskenduses andmelünkade analüüsile). Magistritöö analüüsi peatükis on kasutatud integratsiooni monitooringu meetodika analüüsimiseks kontentanalüüsi meetodit, andmete lünklikkuse analüüsis on teostatud peamiselt kirjeldavat statistikat (sagedusjaotused ja risttabelid), hii-ruut-testi ja Fisheri täpset testi. Magistritöö lõpetab analüüsil põhinev arutelu koos autoripoolsete järelduste, rakendustepanekute ning soovitustega edasiseks uurimiseks.

Autor soovib tänada südamest oma juhendajaid Indrek Soidlat ja Maarja Vollmerit väärt nõuannete, kannatlikkuse, mõistva suhtumise ja inspiratsiooni eest.

1. KIRJANDUSE ÜLEVAADE JA PROBLEEMISEADE

Käesoleva magistritöö keskmes on integratsiooni monitooringu küsitlusuuringu andmed ning meetodika, seega antakse järgnevalt lühike ülevaade teemadest, mis on olulised küsitlusuuringu eri aastate andmete omavahelisel võrdlemisel.

1.1 Küsitlusuuringute võrreldavus

Küsitlusuuring on andmete kogumise viis, mida kasutatakse kvantifitseeritud andmete kogumiseks ja mille eesmärk on koguda korraka standardiseeritud kujul infot paljude indiviidide kohta (Beilmann 2020). Selleks, et tagada erinevate küsitlusuuringute andmete võrreldavus, võib olla vajalik andmeid harmoniseerida. Põhiliselt eristatakse kahte tüüpi harmoniseerimist: sisendi harmoniseerimine (*input harmonisation*) ning väljundi harmoniseerimine (*output harmonisation*), viimasel on omakorda kaks alatüüpi - *ex-ante* ja *ex-post* ehk harmoniseerimine enne või pärast andmekogumist (Köner ja Meyer 2005).

Küsitlusuuringute võrdlemise seisukohast tähendab sisendi harmoniseerimine uuringu disaini ja läbiviimisega seotud aspektide (nagu valimi, küsimustiku, andmekogumismeetodi) standardiseerimist juba enne uuringu läbiviimist. Kui võrdlevat uuringut alles hakatakse planeerima, on sisendi harmoniseerimine parim valik (Wolf jt 2016). *Ex ante* väljundi harmoniseerimise lähenemist kasutades võidakse näiteks enne andmekogumist kokku leppida teatud teoreetilised kontseptsioonid, mida mõõdetakse ning empiirilised indikaatorid, mis mõõtmistulemustel saadakse (*ibid*). Näiteks Eurostat annab ette riikidele indikaatorid, mida on vaja esitada, kusjuures ei ole oluline, mil viisil riigid indikaatorite jaoks andmeid koguvad, vaid oluline on, et esitatud indikaatorid vastaks nõuetele ja oleks võrreldavad teiste riikidega. *Ex post* väljundi harmoniseerimist kasutatakse siis, kui andmed on kogutud, ehk enamasti sekundaarandmete analüüsis (*ibid*). Siia alla loetakse näiteks andmete ümber kodeerimist võrreldavate tunnuste saamiseks.

Üheks heaks küsitlusuuringu näiteks võib tuua Eestis läbiviidava Euroopa Sotsiaaluuringu (ESS – European Social Survey). Uuringut on korraldatud Eestis juba alates 2004. aastast ning seda teostatakse iga kahe aasta tagant (ESS tutvustus 2022). ESSi läbiviimiseks on paigas ranged protseduurireeglid, mis tagavad andmete võrreldavuse nii ajas kui ka erinevate riikide vahel. Alati ei ole aga kõik küsitlusuuringud niivõrd hästi planeeritud ja dokumenteeritud. Kui uurija soovib võrrelda erinevate uuringute andmeid, on tavapärase, et enne analüüsi juurde asumist veendub ta, et antud uuringute andmed on omavahel võrreldavad. Jätku-uuringute puhul võidakse aga sageli eeldada, et andmed on võrreldavad. Samas, kui jätku-uuringute disainimisel ja läbiviimisel ei ole jälgitud kindlat reeglistikku, tuleb pealtnäha sarnaste andmete võrdlemisel olla ettevaatlik. Eriti kui uuringu teostajad on erinevad, peaks veenduma uuringumetoodika samasuses ning võrdlema eri uuringute kvaliteeti. Järgnevalt peatutakse lähemalt mõnel küsitlusuuringu disaini olulisemal etapil (valimi moodustamine ja ankeetküsimustiku koostamine) ning vaadatakse, mille alusel hinnata uuringu kvaliteeti.

1.1.1 Tõenäosuslik ja mittetõenäosuslik valim

Kui uuritav sihtrühm on liiga suur, et küsitleda kõiki üldkogumi ehk populatsiooni liikmeid, moodustatakse tavapäraselt valim ehk kaasatakse uuringusse vaid kindlal meetodil eraldatud osa populatsioonist. Valimeid liigitatakse tõenäosuslikeks ning mittetõenäosuslikeks. Valimeid, kus igal üldkogumi liikmel ehk elemendil on nullist suurem tõenäosus valimisse sattuda ning see tõenäosus on teada, nimetatakse tõenäosuslikeks valimiteks (Kalton 1983). Mittetõenäosusliku valimiga on tegemist siis, kui eelnimetatud tõenäosusliku valiku kriteeriume ei järgita. Mittetõenäosuslike valimite suurimaks nõrkuseks peetakse elementide valikusse kaasamise subjektiivsust (*ibid*).

Järgnevalt on toodud käesoleva magistritöö kontekstis olulisemad tõenäosusliku valiku valimitüübid (Lagerspetz 2017):

- Lihtne juhuvalim – tõenäosus valimisse sattuda on iga elemendi jaoks ühesugune;
- Süstemaatiline juhuslik valik – juhusliku valimi edasiarendus, kus valimisse kaasatakse üldkogumi üksus mingi kindla sammu alusel;
- Kihtvalim ehk stratifitseeritud valim – üldkogum jaotatakse osadeks ehk kihtideks mingi olulise tunnuse alusel ning iga kihi sees valitakse elemendid juhuvaliku abil.

- Klastervalim ehk kobarvalim – valimi moodustamise esmaseks valikuühikuks on uurimisobjektide kobarad/klastrid, mitte üksikud objektid (indiviidid). Juhuslikul teel valitakse klastrid, kus küsitletakse kõiki sinna kuuluvaid indiviide.
- Mitmeastmeline valim – kombinatsioon eelpool mainitud tõenäosuslikest valimitest. Näiteks valitakse üldkogumist juhuslikult suuremad üksused (klastrid), mille sees omakorda valitakse juhuslikult uuritavad indiviidid. Mitmeastmeline valim on tõenäosuslik üksnes siis, kui valik on olnud juhuslik kõigis valiku tegemise etappides.

Mittetõenäosuslikest valimitüüpidest on rakendusuringutes sageli kasutuses kvootvalim, mille puhul valitakse indiviidid valimisse etteantud kvootide alusel. Urija määrab valimi struktuuri teatud taustatunnuste alusel, andes valimi jaoks ette kvoodid vastavalt üldkogumi struktuurile (Tiit ja Tooding 2019: 149). Tavapäraselt seatakse kvoodid populatsiooni sotsiaaldemograafiliste tunnuste alusel.

Uuringutes, kus soovitakse saada usaldusväärseid tulemusi, mida saaks üldkogumile üldistada, kasutatakse enamasti tõenäosuslike valimeid (Lohr 2008). Mittetõenäosuslike valimite puhul ei ole võimalik hinnata valimiviga, kuna valimisse sattunud elementide kaasamistõenäosused ei ole teada, ning seetõttu ei saa teha järeldusi valimi esinduslikkuse kohta (Vehovar, Toepoel ja Steinmetz 2016). Tõenäosusliku valimi moodustamiseks on vajalik, et oleks olemas nimekiri populatsiooni elementidest, mida nimetatakse valikuraamiks (Neuman 2006). Valikuraamiks võib olla näiteks maksuameti register, rahvastikuregister, telefoniraamat vms.

Selleks, et saada esinduslik valim, toob Davern (2008) välja kolm kriteeriumi, mis peaksid olema täidetud:

- a) valikuraam peab hõlmama endas kõiki populatsiooni elemente;
- b) kõikide nende elementide kohta peab olema teada tõenäosus valimisse sattuda;
- c) tuleb koguda täielik info kõigilt valimisse sattunud elementidelt.

Reaalsuses on aga ideaalset valimit koostada väga keerukas. Sageli püütakse esinduslikkuse kadu kompenseerida valimi suuruse tõstmisega. Valimi maht on ühtlasi üks enimlevinud küsimusi seoses tõenäosuslike valimite moodustamisega. Ühelt poolt on võimalik valimi maht välja arvutada, kasutades juhuvaliku statistilisi võrrandeid ning tehes sealjuures eeldusi tulemuste täpsuse ning populatsiooni variatiivsuse osas (Neuman 2006). Sageli toetutakse aga valimimahu

määramisel lihtsalt varasemale kogemusele ning lähtutakse uurijate seas üldtunnustatud valimimahtudest (*ibid*).

Eestis on avaliku arvamuse uuringutes tavapraktikana aktsepteeritud valimimaht 1000 vastajat, mille puhul on tulemuste üldistamisel populatsioonile maksimaalne veapiir 95%-lisel usaldusnivool 3,1%. See tähendab, et võttes üldkogumist korduvalt juhuvalikuga 1000-liikmelise valimi, jäävad tulemused 95%-l kordadest $\pm 3,1$ protsendipunkti ulatusse võrreldes tegeliku väärtusega populatsioonis. Maksimaalne veapiir näitab aga ainult valimivea suurust eeldades, et valim on koostatud tõenäosuslikult ning valiku juhuslikkus säilib ka mittevastamist arvesse võttes, mis tähendab omakorda, et mittevastamine oleks absoluutselt kõigi tunnuste suhtes täiesti juhuslik, mida on aga reaalsuses väga raske saavutada. Sageli võetakse lihtsalt võimalikult suur valim, et saada 1000 vastajat kätte ning sealjuures ei arvestata, et mida suurem on mittevastamine, seda suurema tõenäosusega kannatab valimi juhuslikkus ning andmete esinduslikkus.

1.1.2 Ankeetküsimustike võrreldavus

Ankeetküsimustikuga saab koguda andmeid faktide, käitumise ja tegevuse, väärtuste, suhtumiste, teadmiste, uskumuste ning arusaamade kohta (Hirsjärvi jt 2005). Lähtudes püstitatud uurimisülesandest ning eesmärgist, defineerib uurija esmalt konstrukti(d) ehk uuritava(d) nähtuse(d), mida soovib mõõta, ning seejärel koostab küsimuse(d) (Fowler ja Cosenza 2008). Suletud küsimuse korral käsitletakse iga esitatud vastusevarianti võimaliku tulemusena, mis aitavad mõõta algselt defineeritud konstrukti (*ibid*). Seega oleneb nii koostatud küsimuse sõnastusest kui ka etteantud vastusevariantidest, kui täpne on mõõtmistulemus.

Selleks, et võrrelda erinevate küsitlusuuringute andmeid, tuleks uurida, kuidas on erinevates ankeetides küsimused sõnastatud ja millised on etteantud vastusevariandid. Ankeetküsimuste omavahelisel võrdlemisel on mitmeid olulisi aspekte, mida uurijad võiks silmas pidada. Täpsete vastuste (ja sealjuures võrreldavate andmete) saamiseks tuleks lähtuda järgmistest küsimuse sõnastamise põhitõdetest:

1. Küsimus peab olema kõigile vastajatele üheselt arusaadav (Fowler ja Cosenza 2008, Neuman 2006, Kalton ja Schuman 1982, Lagerspetz 2017). Siinkohal on oluline, et vastajad saaksid küsimusest aru nii, nagu uurijad seda planeerisid. Küsimus peaks olema sõnastatud võimalikult lihtsalt ning mitte sisaldama keerulist sõnavara (näiteks tehnilisi termineid või kõnekäande,

mida vastajad ei pruugi mõista). Mitmeti mõistmist võib esineda ka siis, kui küsimuses ei ole kajastatud kindlat ajaperioodi ning määratlemine on jäetud vastajale (Fowler ja Cosenza 2008). Näiteks küsimust „Kui sageli sa tunnend end kurvana?“ võivad vastajad interpreteerida erinevalt – kui kurvameelne inimene on vastaja üldiselt või kui mitu korda ta on kurbust tundnud näiteks viimase nädala jooksul. Erinevate uuringute tulemuste omavahelisel võrdlemisel on seega oluline, et küsimused oleks vastajatele üheselt arusaadavad, kuna vastasel juhul ei pruugi saadud tulemused olla omavahel võrreldavad. Lisaks tuleks tähelepanu pöörata sellele, et küsimuse sõnastamisel ei küsitaks korraga mitut küsimust ning ei tehtaks küsimuse püstitamisel eeldusi vastaja hoiakute või situatsiooni suhtes (*ibid*).

2. Respondent saab vastust anda ainult küsimusele, millele ta oskab vastata (Fowler ja Cosenza 2008, Neuman 2006). Siinkohal peavad uurijad silmas pidama, et vastajatel oleks piisavalt informatsiooni uuritava konstrukti kohta. Näiteks ei pruugi vastaja osata öelda, kui palju tema maja väärt on või kui palju kaloreid ta eelmisel päeval tarbis, kui ta ei ole sellele kunagi varem tähelepanu pööranud. Samuti tuleks ettevaatlikkusega suhtuda küsimustesse, mida küsitakse teiste inimeste kohta (näiteks vastaja elukaaslase, laste jt kohta), sest respondent ei pruugi nendele küsimustele osata adekvaatset vastust anda. Lisaks ei tasu esitada küsimusi kaugemineviku kohta, kui on oht, et vastaja ei pruugi vastust mäletada (*ibid*). Eelnevaga tuleks ka uuringuandmete võrdlemisel arvestada, sest omavahel ei ole võrreldavad näiteks küsimused, mida ühes uuringus küsitakse vastaja enda kohta ning teises tema pereliikmete kohta.
3. Küsimusele peab olema võimalik anda asjakohane vastus (Fowler ja Cosenza 2008, Neuman 2006). Avatud küsimuse puhul tuleb uurijatel sõnastus leida nii, et respondentidel oleks arusaadav, millist vastust neilt oodatakse. Suletud küsimuste puhul peavad vastusevariandid vastama küsimusele ning nende hulgas peab leiduma vastajale sobilik variant. Kui vastaja ei leia enda jaoks sobilikku varianti on ta kas sunnitud vastama variandi, mis ei ole tema kohta tõene või jätma küsimusele vastamata (Fowler ja Cosenza 2008). Seega on küsimustes esitatud vastusevariandid samuti olulised andmete võrreldavuse seisukohast.
4. Küsimuse sõnastus ja vastusevariandid peaks olema koostatud nii, et vastaja on nõus küsimusele ausalt vastama (*ibid*). Sageli on näiteks palga küsimus respondentide jaoks liiga tundlik, et täpset numbrit avaldada, kuid kui vastusevariandid on esitatud vahemikena, võivad nad tõenäolisemalt sobiva variandi valida ja mitte jätta vastamata. Mõnikord ei olene aga vastamise soov ette antud variantidest, vaid vastajad lihtsalt ei soovi isikliku informatsiooni

avaldata, kuna nad ei tea, kuidas seda hiljem tõlgendatakse (*ibid*). Sellistel juhtudel on võimalik näiteks luua vastaja jaoks konteksti eelnevate küsimuste näol või tuua küsimuse juurde sissejuhatus, mis pakub vastajale teatud selgitusi. Samas leiavad mõned uurijad, et küsimust sissejuhatavad laused võivad mõjuda vastajale suunavalt ning seetõttu tuleks nende kasutamisel olla ettevaatlik (Krosnick, Malhotra ja Mittal 2014). Järelikult tuleks ka andmete võrdlemisel jälgida, kas ja mil määral on kasutatud küsimuste sõnastustes täpsustusi või selgitusi, et veenduda tulemuste võrreldavuses.

Inimeste hoiakuid uuritakse kõige sagedamini küsimustega, kus vastusevariantideks on astmeline skaala. Laialt levinud on nii-öelda maatriksküsimused, kus üksteise alla reastatakse hulk väiteid, millele vastaja ette antud skaala alusel märgib, kuivõrd ta mingi väitega nõustub. Schaeffer ja Presser (2003) selgitavad, et sageli jätavad uurijad skaalade koostamisel välja keskmise või nii öelda neutraalse variandi põhjendades seda väitega, et enamikus küsimustes kalduvad vastajad kas nõustumise või mittenõustumise poole. Taoliste küsimuste puhul on uuringud näidanud, et vastajatel esineb selge tendents väidetega pigem nõustuda (Schaeffer ja Presser 2003). See tähendab omakorda, et küsimuse/väite sõnastamisel on oluline, kas see on esitatud negatiivse või positiivseks. Seega on ka erinevate uuringute andmete omavahelisel võrdlemisel oluline, kas algselt esitatud küsimus või väide oli samasuunaline (positiivne/ negatiivne) ning milline oli etteantud vastamise skaala.

Lisaks on Kamoen jt (2013) leidnud, et tulemusi mõjutab ka see, kas küsimus on esitatud ühepoolset või on tasakaaluks teada ka vastandargument. Ehk kas vastajal palutakse märkida, kuivõrd ta esitatud väitega nõustub või palutakse tal end positsioneerida kahe esitatud väite vahele. Näiliselt võivad need küsimused mõõta sama konstrukti, kuid sisuliselt anda väga erinevaid tulemusi. Samuti tuleks jälgida, et skaala küsimuste puhul esitatud skaala oleks tasakaalus, ehk samas ulatuses kui esineb positiivsel poolel astmeid, esineb ka teisele poole ulatuvalt negatiivseid astmeid (Neuman 2006).

Lisaks küsimuste ja vastusevariantide sõnastusele võib tulemusi mõjutada ka vastusevariantide järjekord. Garbarski jt (2016) leidsid, et maatriksküsimuste puhul esineb oluline vahe tulemustes, kui skaala algab negatiivsete variantidega (nt „Üldse ei nõustu“ kuni „Täiesti nõus“) versus kui skaala algab positiivsete variantidega („Täiesti nõus“ kuni „Üldse ei ole nõus“). Kuna respondentidel esineb kalduvus väidetega pigem nõustuda, siis nõustutakse väidetega oluliselt

rohkem juhul kui skaala algab positiivse variandiga („Täiesti nõus“). Kalton ja Schuman (1982) tõid välja, et erinevus siinkohal võib tuleneda sellest, kas respondent vastab küsimusele silmast silma intervjuu käigus või täidab ise ankeeti kirjalikult või veebis. Bowling (2005) kinnitab, et valitud andmekogumismeetodil võib olla oluline mõju uuringutulemustele, sest erinevad küsitlusviisid (silmast silma intervjuu, telefoniintervjuu, veebiküsitlus) nõuavad vastajalt erinevat kognitiivset koormust, sh lugemist, kuulamist, juhiste järgimist, numbrite tuvastamist jne. Seega ei olene erinevate küsitlusuuringute andmete võrdlemine ainult küsimuse ja vastusevariantide sõnastusest, vaid ka näiteks valitud küsitlusviisist.

Kokkuvõttes võib öelda, et vahest piisab ühe sõna muutmisest küsimuse sõnastuses, et muuta täielikult küsimuse tähendust. Ja samas saab vahel ühte küsimust küsida mitmel erineval viisil, mõjutamata sealjuures tulemust (Gendall ja Hoek 1990). Seepärast on oluline kvaliteetse uuringu läbiviimise juures küsimustiku eelnev põhjalik testimine.

Erinevate küsimustike andmete võrdlemisel langeb vastutus analüütikule või uurijale, kes avalikkusele andmeid võrdlevalt esitab. Uuriija roll on mõista, kuidas erinevad küsimused mõjutavad tulemusi ning seda andmete analüüsimisel arvestada ning tulemuste esitamisel ka selgelt avalikkusele kommunikeerida.

1.1.4 Uuringu kvaliteet

Küsitlusuuringute puhul räägitakse sageli erinevat tüüpi vigadest. Salant ja Dillman (1994, viidatud de Leeuw jt 2008:4 kaudu) toovad välja uuringu neli nurgakivi: kaetus, valimi moodustamine, vastamine ja mõõtmise. Teisiti öeldes, võib lähtuvalt nendest neljast nurgakivist tuua välja neli erinevat tüüpi viga, mis uuringu koostamisel võivad esineda (de Leeuw jt 2008):

- **Kaetuse viga** esineb kui populatsioonis leidub indiviide, kelle valimisse sattumise tõenäosus on null.
- **Valimiviga** esineb siis, kui populatsiooni asemel uuritakse vaid osa üldkogumist.
- **Mittevastamise viga** esineb siis, kui mõned valimisse sattunud individid (üldkogumi elemendid) ei vasta küsimustikule ning kui need individid erinevad nendest, kes küsimustikule vastasid.
- **Mõõtmisviga** esineb siis, kui vastaja vastus küsimusele on ebatäpne või vale.

Groves ja Lyberg (2010) toovad uuringu koguvea (*TSE - Total Survey Error*) käsitluses lisaks eelnevalt mainitud neljale veatuübile välja veel kaks:

- **Valiidsus** näitab, kui võrd täpselt mõõdab loodud instrument (küsimustikus küsimus/väide) uuritavat nähtust ehk konstrukti.
- **Järeldootluse viga** tuleneb andmetootluse käigus tekkinud vigadest, sealhulgas andmete sisestamise, kodeerimise ja andmete kaalumise käigus tehtud vigadest.

Erinevad autorid kategoriseerivad uuringus tekkivaid vigu erinevalt, kuid välja toodud vead baseeruvad enamasti samadel alustel (Biemer ja Lyberg 2003, viidatud De Leeuw jt 2008:3 kaudu; Kish 1965, Groves ja Lyberg 2010). Kuigi kvaliteetse uuringu tagamiseks püütakse uuringu viga minimeerida, ei olene uuringu kvaliteet pelgalt eelpool mainitud vigadest, vaid seda peaks käsitlema laiemal kontseptsioonina. Biemeri ja Lybergi (2008) kohaselt tuleb üldist uuringu kvaliteeti hinnata kolmel tasandil: toote tasand, protsessi tasand ning organisatoorne tasand.

Toote tasandil on kõige olulisem, et toode vastaks tellija ootustele meetoodika, läbiviimise ja tulemuste esitamise osas ning et kõik soovitud karakteristikud oleks täpselt kokku lepitud ja hästi kommuniqueeritud (*ibid*). Näiteks kui tellija on huvitatud, et analüüsitulemused oleks esitatud väiksemate alagruppide kaupa, mis sageli tähendab nende gruppide ülesindatust valimis, peab ta mõistma, et sellisel juhul väheneb tulemuste täpsus. Samuti soovivad uuringu tellijad sageli saada andmeid võimalikult kiiresti ning võivad ette anda lühikese ajaraamistiku. Ka siinkohal on oluline tellija ja uuringu läbiviija tihe koostöö, et paika seada realistlik ajaraamistik, mille käigus jõuaks uuringut planeerida, andmeid koguda ja töödelda kvaliteetselt.

Protsessi tasandist oleneb kõige enam lõpptoote kvaliteet ehk kui uuringu protsess on hästi disainitud ja kontrollitud, siis tuleneb sellest ka kvaliteetne uuring (*ibid*). Siia alla kuuluvad nii küsimustiku koostamine kui ka andmekogumise läbiviimine ja andmetootlus. Kontrollimehhanismideks peetakse siin näiteks disainitud küsimustiku testimist ja sellest tulenevate muudatuste sisseviimist. Organisatoorset tasandit seostatakse enamasti uuringu korraldamise ja juhtimisega ning sellel tasandil otsustatakse protsessi kvaliteet (Biemer ja Lybergi 2008). Seega küsitlusuuringu kvaliteeti mõjutavad küll laiemalt erinevad tegurid, kuid kõige olulisemaks ja määravamaks on siiski protsessi kvaliteet.

Järgnevalt on kirjeldatud veidi lähemalt eelnevalt mainitud eri tüüpi vigu, mis võivad küsitlusuuringu protsessi käigus tekkida.

1.1.4.1 Kaetuse- ja valimiviga

Kaetust saab defineerida kui osakaalu populatsioonist, mis on kaasatud valikuraami (Lohr 2008) ehk kaetus näitab, kui hästi vastab valikuraam üldkogumile. Tavapäraselt on kaetuse viga tingitud kehvast valikuraamist, näiteks puudub info kõigi populatsiooni liikmete kohta (de Leeuw jt 2008). Sellisel juhul on tegemist alakaetusega. Ülekaetus esineb juhul kui valikuraamis on rohkem elemente kui üldkogumis või mõni populatsiooni element on valikuraamis esindatud rohkem kui ühel korral. Kaetuse viga ja ennekõike alakaetus on suureks probleemiks just veebiuuringute puhul (Lohr 2008), kuid kaetuse saavutamine võib osutada problemaatiliseks ka elanikkonna uuringute puhul (Tourangeau jt 2013).

Valimiviga esineb, kuna populatsiooni asemel uuritakse vaid osa üldkogumist ehk valimit (de Leeuw jt 2008). Isegi kui tõenäosusliku valiku nõudeid on koostamisel hoolikat järgitud, ei saa ükski valim kunagi 100% vastata üldkogumile (välja arvatud kõikne valim). Seega tuleb uurijatel arvestada lahknevusega populatsiooni ja valimi näitajates. Taolist lahknevust ehk viga, mis tuleneb tõenäosusliku valiku juhuslikkusest, nimetakse valimiveaks (Neuman 2006). Valimivea suurus annab võimaluse hinnata, kui hästi valim üldkogumit esindab ning selle iseloomustamiseks kasutatakse uuringutes sageli veapiiri mõistet. Veapiir näitab, kui palju võivad uurimistulemused erineda tegelikest väärtustest populatsioonis (Toepoel 2016).

1.1.4.2 Mittevastamise viga

Mittevastamisest tingitud kadu võib küsitlusuuringutes jaotada kaheks: objekti kadu (täielik mittevastamine), mis tähendab, et valimisse kuuluvalt elemendilt ei õnnestu vastuseid koguda (*unit non-response*) ja küsimuse või mingi kindla tunnuse väärtuse kadu (osaline mittevastamine) ehk kui respondent jätab küsimustikus vastamata mõnele konkreetsele küsimusele (*item non-response*) (de Leeuw jt 2008). Elemendi või objekti kadu võib esineda näiteks juhtudel, kui respondenti ei õnnestu kätte saada või ta keeldub vastamast. Küsimusele mittevastamine tekitab andmelünkasid, mille hulka loetakse vastamisest nii keeldumist kui arvamuse puudumist („Ei oska öelda“ vastus),

aga ka seda, kui vastaja suunatakse küsimusest mööda või tekib andmelünk mõne tehnilise vea tõttu.

Dong ja Peng (2013) toovad välja, et kuigi uuringutes andmelünkadele sageli väga palju tähelepanu ei pöörata, võivad nad tekitada esinduslikkuse kadu ning nihkeid hinnangutes. Lisaks eeldavad paljud statistilised protseduurid andmeanalüüsis täielikku andmestikku, mistõttu sageli jäetakse puuduvate väärtustega indiviidid lihtsalt analüüsist kõrvale. Kuigi uurijate seas ühtset arusaama ei ole, kui palju puuduvaid väärtusi on „halb“ uuringu kontekstis, siis Bennet (2010) toob välja, et kui küsimuse vastuse seas on vähemalt 10% mittevastamisi, on sellel juba arvestatav mõju hinnangutele ning läbiviidavale analüüsile. Tabachnick ja Fidell (2014: 96) leidsid aga, et andmelünkade mehhanismidel ja muustril on oluliselt suurem mõju uuringutulemustele kui puuduvate väärtuste hulgal.

Little'i ja Rubini (2002) alusel võib andmete analüüsimisel lünklikkust jagada oma olemuselt kolmeks:

- a) juhuslik lünklikkus (MAR – *missing at random*) – lünkade esinemine tunnuses ei sõltu uuritava tunnuse väärtustest.
- b) täiesti juhuslik lünklikkus (MCAR – *missing completely at random*) – andmelüngad ei sõltu ei uuritavast tunnustest ega teistest tunnustest.
- c) mittejuhuslik lünklikkus (NMAR – *not missing at random*) – lünkade esinemine on seotud tunnuse enda väärtustega ning ei ole kirjeldatav andmestikus esinevate teiste tunnustega.

Viimast nendest (NMAR) on aga väga raske uuringutes tuvastada. Kui andmelüngad on juhuslikud või täiesti juhuslikud, on võimalik andmete esinduslikkuse kadu vähendada kaalumisega (Rässler jt 2008). Andmeid kaalutakse reeglina nii, et see tõstaks andmete esinduslikkust objekti kao (*unit non-response*), mitte tunnuse väärtuste kao (*item non-response*) suhtes (Spiess 2016).

Kaalumise protsess koosneb tavaliselt kolmest sammust: disainikaalude arvutamine, mittevastamisest tingitud esinduslikkuse kao kompenseerimine ja kalibreerimine. (Lavallée ja Beaumont 2016). Esimese sammuna arvutatakse disainikaalud, teise nimega valiku kaalud (*sampling weights*), mis on seotud valikuraamiga ning korrigeerivad valimisse sattumise ehk kaasamistõenäosust. Selleks on oluline, et uurijal oleks olemas info valikuraami kohta. Teise sammuna võidakse korrigeerida disainikaale puuduvate väärtuste suhtes, et kompenseerida

nendest tingitud objekti kadu. Viimase sammuna toimub kalibreerimine ehk kaalude kohendamine populatsiooni suhtes. Seda nimetakse ka järelkihistamiseks: valimi objektid jagatakse järelkihtidesse, mille kohta on üldkogumi tasemel teada mahud ning mille alusel saab arvutada igale indiviidile järelkihistuskaalu suuruse (Lepik ja Traat 2013). Kaale arvutatakse, et tõsta andmete esinduslikkust ning hinnangute täpsust.

Arvutatud kaalude suurus sõltub ennekõike sellest, kui esinduslikud on valimi põhjal saadud andmed populatsiooni suhtes, aga ka palju erinevaid tunnuseid kaalude arvutamisse on kaasatud ning kui palju erinevaid alakategooriaid nendel tunnustel esineb. Kui järelkihistuskaalude arvutamise tulemusel saavad mõned valimi elemendid liiga suure väärtuse, võivad uurijad piirata kaalu väärtuste suurusi. Seda protsessi nimetatakse kaalude kärpimiseks (*weight trimming*) (Biemer ja Christ 2008). Näiteks Euroopa Sotsiaaluuringus kärbitakse järelkihistuskaalud 4 peale, mis tähendab, et kõikide suurema järelkihistuskaaluga kui 4 indiviidide kaal asendaks 4-ga (Documentation of... 2014).

1.1.4.3 Mõõtmisviga ja valiidsus

Mõõtmisviga esineb juhul, kui respondendi vastus küsimusele on ebatäpne või vale (de Leeuw jt 2008). Mõõtmisviga seostatakse kõige enam andmekogumise protsessiga, kuid üldiselt võib seda tüüpi viga tuleneda kolmest allikast: küsimustikust, vastajast endast või andmekogumise meetodist. Kui andmekogumise juurde on kaasatud intervjuerijad, siis võib neid pidada neljandaks vea allikaks.

Mõõtmisviga aitab vähendada hästi disainitud ja testitud küsimustik. Kuigi uurija ei saa lõplikult kontrollida, kas või kuidas respondent küsimustikule vastab, on mõõtmisvea vähendamiseks võimalik teha vastaja ülesanne nii lihtsaks ja meeldivaks kui võimalik. Selleks peaks küsimused olema selgelt ja lihtsalt sõnastatud ning üheselt arusaadavad ja vastajal ei tohiks tekkida tõrkeid küsimustele vastamisel (*ibid*).

Mõõtmisviga mõjutab veel andmekogumismeetodi valik. Silmast silma intervjuu puhul kohtub intervjuerija respondentiga ning vastused ankeedi küsimustele paneb enamasti kirja intervjuerija ise (kas siis elektrooniliselt või paber kandjal). Silmast silma intervjuude puhul tuleb arvestada, et intervjuerija võib tahtlikult või tahtmatult mõjutada respondendi vastuseid (Neuman 2006). Tundlike teemade puhul võivad vastajad hoiduda intervjuerijale vastuste andmisest ning

seeläbi suurendada andmelünkade esinemist andmestikus (de Leeuw 2008). Head intervjuerijad peavad olema hinnanguvabad, tekitama vastajas usaldust, säilitama neutraalset positsiooni ning mitte avaldama oma arvamust (Neuman 2006).

Kui intervjuerija sõnalised väljaütlemised või ka kehaline olek peegeldab teatud hoiakuid mingite küsimuste suhtes, võib respondent tunnetada survet teatud viisil vastata. Näiteks on magistritöö autor ise puutunud kokku tunnustatud uuringufirma küsitlejaga, kes on soovitanud teatud vastustes mitte liiga palju variante märkida, sest edaspidi tuleb ka iga variandi kohta eraldi hinnanguid anda. Sama intervjuerija suunas ka autorit küsimustele vastamisel märkama, milliseid vastuseid anda, et erinevate küsimuste vahel tulemuste ebakõlasid vähendada (näiteks küsimustes millist erakonda valiks vastaja järgmistel Riigikogu valimistel ning kes on tema eelistatud peaministri kandidaat).

Seega võib intervjuerija otseselt mõjutada saadud tulemusi ja seeläbi ka andmete võrreldavust nii ühe küsitlusuuringu sees kui ka erinevate uuringute omavahelisel võrdlemisel. Intervjuerija mõju küsitlusandmete kogumisel hinnatakse sageli uuringufirma poolt läbi viidava järelkontrolliga (de Leeuw jt 2008), mille käigus kontakteerutakse vastajaga ning uuritakse intervjuu/küsitluse läbi viimise korrektsust.

Kui mõõtmisviga esines tulenevalt respondendi vastusest esitatud küsimusele, siis valiidsuse all peetakse silmas ennekõike, kuivõrd hästi koostatud instrument mõõdab uuritavat nähtust. Valiidsuse puhul võib eristada näilist valiidsust, ennustavat valiidsust ja sisuvaliidsust, kuid uuringu vea kontekstis räägitakse kõige enam konstruktivaliidsusest, mis näitab kuivõrd instrument mõõdab seda, mida ta mõõtma peaks (Andrews 1984, Babbie 2010). Luues mõõtmisinstrumenti, luuakse ka teoreetiliste ootuste raamistik, kuidas vaadeldav instrument seostub teiste näitajatega (Babbie 2010). Seega on konstruktivaliidsust võimalik hinnata, uurides, kas mõõdetav instrument on seotud teiste tunnustega nii nagu on teoreetiliselt ette nähtud (*ibid*). Näiteks toob Babbie välja, et mõõtes abieluga rahulolu eeldatakse, et abieluga rahulolevad inimesed petavad vähem, kui need, kes oma abieluga rahul ei ole. Kui tulemused seda teoreetilist eeldust ei kinnita, tuleks kontrollida mõõtmisinstrumenti valiidsust. Seega valiidsus peaks kokkuvõttes tähendama, et uurija mõõdab tegelikult ka seda, mida mõõta tahab.

1.1.4.4 Järeltöötuse viga

Järeltöötuse viga ehk töötusviga (*processing error*) esineb siis, kui viga tehakse kogutud andmete sisestamise või töötuse käigus (Biemer 2010). Näiteks võidakse andmete käsitsi sisestamisel ekslikult sisestada valed andmed. Küsitlustes, kus vastaja ankeeti ise täidab, võivad olla ette nähtud protseduurid andmete kontrollimiseks, näiteks valideerib andmetöötleja mingi osa vastustest koos respondendiga üle, kuid seda mitte tehes võivad andmetesse vead tekkida. Töötusviga võib tekkida ka andmetöötlastarkvarast (*ibid*), näiteks kui tehakse arvutusi Excelis kasutades lahtriviidetega valemeid, võib uurija hooletusest panna valemisse sisendina näiteks valed andmed. Järeltöötuse vea alla võib lugeda ka seda, kui andmestiku kaalud on valesti arvatud või kui näiteks arvutamise aluseks võetud tunnused ei võimalda saavutada soovitud esinduslikkust

Mitmed teadlased rõhutavad, et järeltöötusviga on siiani pigem vähe uuritud ja vajaks rohkem tähelepanu (Lyberg 2012, Biemer 2010). Kuigi teadlased arvestavad uuringu koguvea käsitlustes üha enam töötusvea esinemise võimalikkusega, siis praktikas tuleks selle olulisust teadvustada ja rohkem rõhutada just rakendusuringute ja andmetöötusega tegelevate organisatsioonide seas.

1.1.5 Uuringu dokumenteerimine

Erinevate küsitlusuuringute omavahelisel võrdlemisel on väga oluline, kuidas ja mis detailsuses uuringud on dokumenteeritud. Kui uuringul puudub dokumentatsioon, puudub kasutajatel info uuringumetoodika kohta ning ilma selle infota ei saa teha järeldusi erinevate uuringuandmete omavahelise võrreldavuse kohta.

Mohler jt (2008) toovad välja, et uuringu dokumentatsioon peaks koosnema kolmest komponendist: arvulised andmed, metaandmed ning paraandmed. Arvuliste andmete all mõeldakse tavapäraselt küsitlusuuringuga kogutud andmestikku. Metaandmete puhul on tegemist andmetega, mis kirjeldavad kogutud andmeid. Paraandmed viitavad andmetele, mis on andmekogumise protsessi kohta (Couper 1998 viidatud Mohler jt 2008 kaudu). Makrotasandi paraandmeteks võib nimetada näiteks vastamismäära ja kaetuse määra, samas kui mikrotasandi paraandmete alla kuuluvad näiteks intervjuerija märkmed.

Uuringu läbiviijate jaoks on dokumenteerimine oluline, et juhtida uuringu protsessi ning vastata tellija lepingulistele tingimustele. Uuringu andmete kasutajate vaatenurgast on dokumentatsiooni

olemasolu tähtis, kuna tagantjärele on uuringuga mitteseotud inimesel pea võimatu uuringu dokumentatsiooni uuesti kokku panna ja sellest tulenevalt andmete analüüsil adekvaatseid metodoloogilisi valikuid teha. Sellest lähtuvalt võiks uuringu dokumentatsioon kindlasti sisaldada järgnevat infot: uuringu teoreetiline kontseptsioon ja kontekst (sh mis oli püstitatud uurimisülesanne ning millised konstrukte sooviti mõõta), andmed uuringu disaini ja läbiviimise protsessi kohta (sh info valikuraami ja valimimudel koostamise kohta), dokumentatsioon andmekogumisest (sh millist andmekogumismeetodit kasutati, küsitluse korral detailne küsitlustöö aruanne ning info järelkontrolli toimumise ja tulemuste kohta), metaandmed (sh küsimustikud) ning uuringuandmed (sh uuringu andmefail) (Mohler jt 2008).

1.2 Ülevaade Eesti ühiskonna integratsiooni monitooringust

Eesti ühiskonna integratsiooni monitooring on laiapõhjaline elanikkonna uuring, mis esimest korda viidi Eestis läbi aastal 2000. Monitooringu ainulaadsus ning väärtus, võrreldes teiste sarnaste uuringutega, seisneb tema pikaajalises korraldamises ja integratsiooni temaatika süvitsi käsitlemises. Kokku on 20 aasta jooksul tehtud 8 monitooringut (aastatel 2000, 2002, 2005, 2008, 2011, 2015, 2017, 2020). Monitooringu tellijaks on viimasel neljal aastal olnud Kultuuriministeerium, varasemad uuringud tellis Rahvastikuministri büroo Riigikantselei kaudu. Uuringut tellitakse tavapäraselt hankega, mistõttu on uuringu teostajad aastate jooksul muutunud. Eesti ühiskonna integratsiooni monitooringust rääkides ning sellele viidates kasutatakse sageli sõnapaari „integratsiooni monitooring“ või „lõimumismonitooring“, aga ka lühendit „EIM“.

Monitooringu põhieesmärk on analüüsida Eesti ühiskonna kohanemis- ja lõimumisprotsesse, et mõista ja parandada nende edukust Eestis ja anda sisend riikliku kohanemis- ja lõimumispoliitika kavandamise ja rakendamise (Eesti ühiskonna lõimumise monitooring 2020). Uuringu peamiseks sihtrühmaks on Eesti püsielanikkond, kelle seas viiakse läbi elanikkonna küsitlus, viimastel aastatel on lisandunud eraldi gruppina veel uussisserändajad ning tagasipöördunud, kellele on saadetud veebiküsitlus. Monitooringu tulemuste peamiseks tarbijateks on poliitikakujundajad, kuid ka teadlased ning laiemalt kogu Eesti ühiskond, kellele tulemusi meedia kaudu presenteeritakse (Kaldur jt 2019).

Lõimumismonitooring on väga tihedalt seotud olnud valdkondliku lõimumise arengukava ja selle eesmärkidega, pakkudes võimaluse seirata arengukava rakendusplaani ja selle mõõdikuid ning

hinnata eesmärkide saavutamist. Samuti on monitooringu andmed olnud sisendiks uute arengukavade loomisel. Alates 2000. aastast on Eestis olnud nelja erineva arengukava perioodid (sh praegune):

- Integratsioon Eesti ühiskonnas 2000-2007,
- Lõimumiskava 2008-2013,
- Lõimuv Eesti 2020 (arengukava perioodiks 2014–2020),
- Sidusa Eesti arengukava 2021-2030.

Kuigi valdkondliku arengukava laiem eesmärk on läbi aastate olnud sidusa Eesti ühiskonna loomine, võib märgata periooditi alaeesmärkide muutumist ning täpsustumist. Kuna integratsiooni monitooringud on väga tugevalt olnud seotud arengukava eesmärkide ja mõõdikutega, on seetõttu erinevatel arengukava perioodidel olnud ka oluline roll monitooringus toimunud muudatustes.

Lõimumismonitooringu andmete kasutusvõimalused on diskussiooni aluseks olnud ka mitmel toimunud konverentsil ja avalikul üritusel. 2021. aastal toimunud rahvusvahelise konverentsi paneelarutelus tõid nii Kristjan Kaldur kui ka Triin Vihalemm välja, et ankeetküsimused on ajas muutunud (Kaldur jt 2021). Küsimuste muutus on seotud eelkõige sellega, et integratsiooni mõtestamine on 20 aasta jooksul oluliselt muutunud, samas on oluline roll olnud ka teatud võtmesündmustel nagu 2007. aasta Pronksiöö, Ukraina kriis 2014-2015 ning viimaste aastate koroonapandeemia.

Veebiseminaril “20 aastat lõimumismonitooringu andmestikku Eestis – tabamata seosed kaardistamata aladel” tõi Kaldur (2021) välja, et integratsiooni monitooring on muutumas trendiuuringuks ja rõhku tuleks panna tulevikus pikaajaliste trendide analüüsimisele. Monitooringute raames valmib eri teemade lõikes ülevaate andev lõppraport, milles aga andmete pikaajalisele võrdlusele suurt rõhku ei ole pandud. Tavapäraselt esitatakse tulemusi võrdlevalt eelmise või mõne üksiku eelnenud monitooringuga

Rohkem rõhku seatakse eestlaste ja mitte-eestlaste hoiakute ning olukorra võrdlemisele, mistõttu enamasti on tulemused raportites esitatud rahvuse lõikes. Samas võib aruandeid uurides märgata, et olenevalt aastast on teatud küsimuste puhul kasutatud erinevaid lähenemisi ehk rahvuse asemel võrreldud hoopiski vastajaid keele või kodakondsuse alusel. Kokkuvõttes võib öelda, et 20 aasta

jooksul on monitooringut läbi viinud erinevad tegijad ning märgata võib ka aastate lõikes erinevusi uuringu korralduses ja monitooringu teemade kajastamises.

1.2.1 Integratsiooni monitooringu meetodika (re)kontseptualiseerimine

Balti Uuringute Instituut viis aastal 2019 läbi lühiuuringu, mille eesmärk oli läbi vaadata Eesti integratsiooni monitooringute senine korraldus ning meetodika ning seeläbi anda soovitusi tulevaste monitooringute korraldamiseks (Kaldur jt 2019). Uuringu käigus teostati nii sekundaarandmete analüüs kui ka viidi läbi kokku 20 intervjuud valdkondlike ekspertidega. Sealhulgas analüüsiti 2017. aasta ankeetküsimustikku ning esitati muudatussoovitused nii ankeedi kui ka meetodika osas. Järgnevalt on toodud mõned olulisemad tulemused nimetatud raportist (Kaldur jt 2019), arvestades käesoleva magistritöö konteksti:

- Uurijad toovad välja, et lõimumismonitooringu andmed ja küsimustikud võiksid olla vabalt kättesaadavad laiemale avalikkusele edasiste analüüside teostamiseks. Selleks pakuti välja, et andmeid võiks hoida Statistikaamet. Samas märgiti ära, et praeguse korra järgi on Statistikaametilt andmete kättesaadavus piiratud (st enamasti tasuline).
- Monitooring peaks kujunema Eesti integratsioonipoliitikat hindavaks sisukaks tööriistaks, mis võimaldaks analüüsi teostada nn trendiuuringuna.
- Ankeedi pikkus ja küsimuste kattuvus sõltub lisaks teemade valikule ka sihtrühmade valikust. Paljud teemad, nagu meediakasutus või usaldus, on kaetud juba teiste regulaarsete (ja esinduslike) uuringutega. Kui tahetakse uurida kogu elanikkonda, võiks küsitlusuuringus keskenduda vähematele teemadele, mida mujal ei kajastata. Teisisõnu on praegune ankeet ekspertide sõnul liiga pikk.
- Lisaks esineb ankeedi sees mitmeid korduvaid küsimusi (nt keelekasutuse osas), mille osakaalu võiks vähendada. Samas on küsimusi näiteks Venemaa või NATO kohta, mis vajaks põhjalikumat käsitlust, kui praegune paari küsimusega lähenemine võimaldab. Kuna hetkel ankeedis olevate küsimuste pealt ei ole võimalik head ülevaadet anda, võiks need küsimused sama hästi välja jätta.
- Lõimumisvaldkonna jaoks oluliste indikaatorite arvutamise aluseks olevad andmed võiksid tulla eelkõige riiklikest registritest või monitooringust sagedamini toimuvatest uuringutest.

Lõimumismonitooringuga võiks täita tühimiku ehk uurida teemasid, mis on valdkondlikult olulised, kuid ei ole kaetud registri- või teiste uuringute andmetega.

- Ankeedis kasutatud termineid tuleks üle vaadata ja ühtlustada (nt on paralleelselt kasutatud: mitte-estlased, venelased, venekeelne elanikkond, muulased jmt).
- Ankeedis esineb küsimusi, mis on kas liialt äärmuslikud või provokatiivsed, nende puhul tuleks sõnastust muuta neutraalsemaks. Lisaks leidis küsimusi, milles oli ühe küsimuse sees esitatud mitu eraldiseisvat mõtet, mida ei peeta samuti küsimuse esitamisel heaks praktikaks.

Kokkuvõttes tehti hulganisti praktilisi ettepanekuid, millega juba 2020. aasta monitooringu planeerimisel arvestati.

1.3 Probleemiseade ja uurimisküsimused

Monitooringu andmete ning tulemusraportite peamised tarbijad on poliitikakujundajad, teadlased/uurijad ning huvitatud tavakasutajad. Erinevad sihtrühmad on huvitatud monitooringu erinevatest väljunditest: poliitikakujundajad (näiteks ministeeriumi töötajad) kasutavad oma igapäevatöös monitooringu tulemusi ja soovitusi, samas kui uurijad on ennekõike huvitatud monitooringu andmefailidest ja spetsiifilistest teemadest. Aastate jooksul on küsitlusuuringut läbi viinud mitmed erinevad uurimismeeskonnad, samuti on aja jooksul muutunud ka küsimustiku ülesehitus. Samas on olnud tellijapoolseks eelduseks, et üldine monitooringu struktuur ja eri aastate andmete võrreldavus säiliks. Eriti oluline on see just uurimistulemuste tõlgendamisel ja ajaliste trendide jälgimisel.

Monitooringu andmete võrreldavuse küsimus võib tõstatuda erinevates situatsioonides. Küsitlusuuringu eri lained peaksid olema meetoodiliselt sarnaselt disainitud ja hästi dokumenteeritud, et tagada kogutud andmete võrreldavus. Silmas tuleks pidada nii ankeetküsimustiku koostamist, üldkogumi ja valimi defineerimist, andmekogumismeetodi valikut, ning andmete esinduslikkuse tõstmiseks teostatud järeltöötlust. Samuti on oluline monitooringuga kogutud andmete üldine kvaliteet (sh kui ulatuslik on andmete lünklikkus), kuna ka see mõjutab eri aastatel korraldatud küsitluste võrreldavust. Teadaolevalt ei ole Eestis taolist laiapõhjalist uurimust veel läbi viidud, samas üha enam nähakse vajadust ja võimalust monitooringu andmete pealt jälgida pikaajalisi trende lõimumist puudutavates küsimustes.

Sellest tulenevalt on käesoleva magistritöö eesmärk välja selgitada, mil määral on lõimumismonitooringu küsitlusmetoodika, ankeetküsimused ning andmete esinduslikkus aastate jooksul muutunud. Lähtuvalt magistritöö eesmärgist ning teema teoreetilisest käsitlusest on töös püstitatud järgnevad uurimisküsimused:

1. Kuidas on lõimumismonitooringute küsitlusmetoodika muutunud aastatel 2000-2020?

Uuritakse, milliseid erinevusi küsitlusmetoodikas aastate lõikes esineb, millest need erinevused võivad tuleneda ning kuidas mõjutada andmete esinduslikkust ja eri aastate andmete võrreldavust.

2. Kuidas on lõimumismonitooringute ankeetküsimused muutunud aastatel 2000-2020?

Uuritakse, milliseid erinevusi esineb erinevates monitooringutes kasutatud küsimuste sõnastuses ja kuidas võivad erinevused sõnastuses mõjutada antud vastuseid / eri aastate andmete võrreldavust.

3. Missugune on andmete lünklikkus erinevates monitooringu lainetes aastatel 2000-2020?

Andmete lünklikkuse uurimine võimaldab teha järeldusi andmete esinduslikkuse kohta. Uuritakse, kui ulatuslik on lünklikkus eri aastate andmestikes ning kas andmete lünklikkust saab pidada täiesti juhuslikuks (MCAR) või juhuslikuks (MAR)/mittejuhuslikuks (NMAR).

2. ANDMED JA METOODIKA

2.1 Kasutatud andmed

Magistritöö analüüsi läbiviimiseks sai autor lõimumismonitooringutega seotud materjalid ja andmed andmevaldajalt Kultuuriministeeriumilt. Saadud materjalide hulgas olid kõigi aastate põhiküsimustiku ankeedid (nii eesti kui vene keeles, erandina aastad 2005 ja 2008, kui saadeti vaid eestikeelne küsimustik) ning kõigi soovitud aastate andmestikud SPSS failidena (.sav failitüübiga). Nelja esimese monitooringu (2000-2008) puhul on iga aasta kohta üks koondandmestik. Viimase nelja monitooringu (2011-2020) puhul on olemas nii koondandmestik kui ka rahvusepõhiselt eestlaste ning mitte-eestlaste andmestikud, kokku kolm andmestikku iga monitooringu aasta kohta. Kultuuriministeeriumi sõnul põhjalikke kirjeldusi meetodikast eraldi failidena ei ole, kogu info, mis on, peaks kättesaadav olema aruannete raportitest.

Seega on andmetena kasutuses antud magistritöös ka monitooringu lõpp-raportid. Lisaks selgus analüüsi teostamise käigus, et kasulikke infot meetodika kohta võib leida ka hankedokumentidest, seega on analüüsis kasutatud ka Riigihangete registrist avalikult kasutuseks kättesaadavad 2011-2020 aastate integratsiooni monitooringute hankedokumentides esitatud uuringu tehniline kirjeldus ning hanke raames tellijale esitatud küsimused ja vastused.

2.1.1 Kitsendused

Tulenevalt magistritöö mahu piirangutest tuli käesolevas töös teha kitsendusi analüüsitud andmeallikate osas. Tulenevalt ühelt poolt mahust, aga ka autori enda piiratud vene keele oskusest, ei ole ankeetküsimuste analüüsi kaasatud venekeelseid küsimustikke. Kuigi küsimuste adekvaatne tõlge ja venekeelsete küsimuste samasus läbi aja on väga olulised andmete võrreldavuse seisukohast, ei ole neid antud magistritöös kajastatud ning pakub seeläbi hea võimaluse hoopiski teema edasiarenduseks ning täiendavaks uurimiseks mõnele teisele üliõpilasele või uurijale.

Samuti on oluline välja tuua, et magistritöös vaatab autor vaid elanikkonna põhiküsitluse ankeete ning andmestikke. Erinevatel aastatel on monitooringutes läbi viidud ka täiendavaid fookusgrupe,

süvaintervjuusid erinevates sihtrühmades ning viimastel aastatel ka veebiküsitlust uussisserändajate ning tagasipöördujate seas. Neid käesolevas magistritöös ei käsitleta tulenevalt magistritöö eesmärgist, milleks on välja selgitada, mil määral on lõimumismonitooringu küsitlusmetoodika, ankeetküsimused ning andmete esinduslikkus aastate jooksul muutunud.

Viimaks, kuna monitooringute elanikkonna küsimustikud on väga mahukad (nummerdatud küsimusi ankeedis igal aastal ca 100, kuid vaadeldud eraldi küsimustiku objekte, mille hulk erines aastati 200st kuni 400ni), on konkreetsete küsimuste sõnastuse analüüs läbi viidud vaid valitud küsimuste puhul. Täpsem kirjeldus ja põhjendus valiku protsessist on välja toodud ankeetküsimuste analüüsi peatükis (3.2).

2.2 Metoodika

2.2.1 Kontentanalüüs

Monitooringute uuringumetoodika muutust ajas analüüsimisel on kasutatud kontentanalüüsi meetodit. Kontentanalüüsi puhul on tegemist kvalitatiivse analüüsimeetodiga, mille puhul kodeeritakse sisu muutujate abil (sageli arvulisteks) tunnusteks ning mida omakorda on hiljem võimalik kokku lugeda ning statistiliselt analüüsida (Lagerspetz 2017). Teisisõnu kasutatakse kontentanalüüsi selleks, et süstemaatiliselt ja kvantitatiivselt kirjeldada kvalitatiivsete andmete sisu. Kodeerida võib nii analüüsiüksuse nähtavat (ehk manifestset) kui ka peidetud (ehk latentset) sisu (Neuman 2006). Nähtava sisu kodeerimisel paneb uurija kirja selle, mis on tekstist üheselt mõistetav. Peidetud sisu puhul on tegemist niiöelda otseselt välja ütlemata sisuga, kuid mis samas on siiski uurijale tekstist arusaadav (Lagerspets 2017). Kasutada võib ka mõlemat lähenemist korraga ning seeläbi tõsta saadud tulemuse usaldusväarsust.

Analüüsiüksusteks on kontentanalüüsis lõimumismonitooringu raportites toodud metoodika peatükid või nende puudumisel metoodikat käsitlev osa raportist. Valimisse on kaasatud kõigi kaheksa aasta (2000-2020) uuringuraportid ning seetõttu võib öelda, et valim on kõikne. Lisas 1 on toodud käesolevas töös kasutatud kodeerimisjuhise muutujate detailse kirjeldusega.

Kodeerimisjuhise koostamisel lähtus autor magistritöö eesmärgist ja püstitatud uurimisküsimustest. Muutujate määramisel on toetunud ka töö teoreetilises osas tutvustatud kirjandusele. Kodeerimist alustas autor alates hilisematest monitooringu raportitest ning liikudes

ajas tagasi varasemate juurde. Kodeerimise käigus sai jooksvalt täiendatud nii kodeerimisjuhist kui ka muutujate väärtusi ning selgitusi. Analüüsi käigus kodeeris autor nii nähtavat kui peidetud sisu. Nähtavaks sisuks sai pidada näiteks raporti läbiviimise aastat, uuringu teostajate ning raporti autorite nimesid. Peidetud sisu hulka võib lugeda näiteks valimi ja andmekogumise meetodiga seotud muutujate kodeerimise. Andmed kodeeris autor analüüsitarkvaras Microsoft Excel.

Pärast andmete kodeerimist selgus, et varasemate aastate monitooringu raportites ei ole väga palju infot uuringumetoodika kohta. Seetõttu otsustas autor vaadata ka Riigihangete registrist kättesaadavaid hankedokumente ning hanke raames tellijale esitatud küsimusi ja vastuseid. Hankedokumentides keskendus autor tehnilises kirjelduses toodud meetodika ning ankeetküsimustikuga seotud nõuete kirjeldusele. Käesolevas magistritöös on kasutatud noppeid nendest dokumentidest nii meetodika kui ka ankeetküsimuste analüüsi peatükkides, püüdes nende abil pakkuda selgitusi või konteksti monitooringu meetodikas tehtud muudatustele.

2.2.2 Ankeetküsimuste analüüsi meetodika

Ankeetküsimuste samasuse analüüs baseerub autori enda poolt välja töötatud raamistikul, mida kokku pannes võeti arvesse ka töö teoreetilises osas tutvustatud kirjandust. Ankeetküsimuste võrdlemiseks koostas autor esmalt kõigi monitooringu aastate ankeetküsimusi kaardistava maatriksi, kus oli aastate lõikes välja toodud kõik esinenud küsimused koos vastusevariantidega. Ankeetide analüüsimisel ja küsimuste omavahelisel võrdlemisel lähtus autor järgnevast neljast põhilisest punktist:

- Erinevused küsimuste sõnastuses;
- Erinevused küsimuste vastusevariantides/skaalades;
- Küsimuse lihtsus, arusaadavus ja mõistetavus;
- Küsimuse eeldused (suunamised/tingimused).

Autor teadvustab, et hinnang ankeetküsimuste võrreldavusele ja sellele, kui võrd võivad erinevused küsimuste sõnastuses mõjutada uuringutulemuste võrreldavust, jääb teatud määral subjektiivseks. Analüüsi piirab ka tõsiasi, et autoril puudub täpne info küsimuste koostamise teoreetilise tausta ja selle kohta, millist konstrukti on mingi küsimus loodud mõõtma.

2.2.3 Andmete esinduslikkuse analüüsi meetodika

Andmete esinduslikkust mõjutavad mitmed tegurid, sealhulgas valikuraami kvaliteet (kaetuse viga), valimimetoodika, küsitluse korraldus ja välitöö kvaliteet (mittevastamise viga) ja ka valimiviga (kuigi suure valimi korral on selle juhusliku vea mõju reeglina võrdlemisi väike). Küsitlusandmete kvaliteeti ehk andmete põhjal arvutatavate näitajate täpsust (st erinevust nende tegelikest väärtustest) ja seega kaudselt ka esinduslikkust mõjutavad nt instrumentide (ankeediküsimuste) valiidsus, mõõtmise täpsus (sh küsitlustöö metoodilisus ja intervjuerija mõju) ja andmete järeltötluse täpsus (järeltötluse viga, nt avatud vastuste kodeerimine, kaalude tõhusus).

Eeldusel, et valikuraam on täpne, võimaldab andmete esinduslikkust kõige vahetumalt uurida mittevastamise süstemaatilise analüüs. Sealjuures saab eristada täielikust ja osalisest mittevastamisest tulenevat esinduslikkuse kadu. Täielikust mittevastamisest tulenevat esinduslikkuse kadu saaks kõige täpsemini uurida, analüüsides võrdlevalt kõiki valimi liikmeid ja küsitluses osalejaid. See eeldaks oluliste tunnuste väärtuseid kõigi valikuraami indiviidide kohta, mis antud juhul pole võimalik, kuna sellised andmed puuduvad. Osalisest mittevastamisest tulenevat esinduslikkuse kadu saab uurida, analüüsides võrdlevalt neid indiviide, kes on mingile küsimusele vastanud, ja neid, kes on vastamata jätnud. Sellele ka käesolevas töös lünklikkuse analüüs keskendub.

Seejuures tuleb mõnda, et ainult osalisest mittevastamisest tuleneva esinduslikkuse kao analüüs ei pruugi anda täit pilti andmete esinduslikkusest, näiteks juhul kui täielikust mittevastamisest (nt keeldumistest, mittekontaktsete suurest hulgast) tulenev esinduslikkuse kadu on märkimisväärne. Samas on osalisest mittevastamisest tulenev esinduslikkuse kadu oluline aspekt, mis kipub praktikas andmete analüüsil tihtipeale tagaplaanile jääma (nt enamikus analüüsiprogrammides eemaldatakse andmelünkadega indiviidid analüüsist vaikeseadena), kuid millega tuleks kindlasti arvestada.

Käesolevas töös on lünklikkuse analüüsis kasutatud kirjeldavat statistikat, Pearsoni hii-ruut-testi, Fisheri täpset testi ning mitmese testimise puhul on olulisuse nivood korrigeeritud Benjamini ja Hochbergi meetodil.

Hii-ruut-statistik annab võimaluse kahe muutuja vahelise seose iseloomustamiseks, kusjuures uuritav tunnus võib olla diskreetse või klassifitseeritud skaalaga (Tooding 2015:176). Seega on hii-ruut-statistik hea statistiline näitaja, mida kasutada andmelünkade juhuslikkuse uurimiseks teatud teiste kategoriaalsete muutujate suhtes. Magistritöös viiakse läbi kokku 410 hii-ruut-testi monitooringus esinenud küsimuse lünklikkuse indikaatoritunnuse ning taustatunnuse vahel. Hii-ruut-testi läbiviimise eelduseks on, et teoreetilise jaotuse lahtrid, kus $n_{ij} \geq 5$ (i tähistab risttabeli i -ndat rida ja j risttabeli j -ndat veergu), on vähemalt 80% ning kõikides lahtrites $n_{ij} \geq 1$ (Cochran 1954). Käesoleva magistritöö analüüsis ei ole kõikidel juhtudel hii-ruut-testi eeldused täidetud ning nendel juhtudel on kasutatud Fisheri täpset testi. Fisheri täpse testi puhul on kontrollitavate hüpoteeside paar hii-ruut-testile analoogne (võimalik tuvastada tunnuste vahelise seose esinemist). Fisheri täpset testi kasutatakse tavapäraselt väiksemate valimite puhul, kuna testi arvutuskäik on väga mahukas, ja/või siis, kui hii-ruut-testi eeldused ei ole täidetud (Ruxton ja Neuhäuser 2010). Fisheri täpse testi eeliseks hii-ruut-testi ees võib lugeda täpse p -väärtuse välja arvutamist. Lisaks on magistritöös kasutatud Benjamini ja Hochbergi (1995) meetodit, mis pakub mitmese testimise puhul lahendust valeavastuste osakaalu (FDR) kontrollimiseks.

Andmeid on analüüsitud kasutades analüüsitarkvara IBM SPSS Statistics 26 ja Microsoft Excel.

3. INTEGRATSIOONI MONITOOINGU ERI AASTATE ANDMETE VÕRRELDAVUSE ANALÜÜS

Magistritöö analüüs on jagatud kolmeks osaks: uuringumetoodika analüüs, ankeetküsimuste samasuse analüüs ning andmete esinduslikkuse (täpsemalt andmete lünklikkuse) analüüs. Järgnevalt on esitatud analüüsi tulemused vastavalt kolme alapeatüki kaupa.

3.1 Uuringumetoodika analüüs

3.1.1 Monitooringu läbiviijad ja üldine korraldus

Integratsiooni monitooringut on aastatel 2000-2020 läbi viidud kaheksal korral. Esimesel neljal aasta teostasid uuringut Tallinna Ülikooli teadlased (endine TPÜ Rahvusvaheliste ja Sotsiaaluuringute Instituut) ning viimasel neljal korral (2011, 2015, 2017, 2020) on uuringu läbiviimisesse kaasatud mitmeid partnereid ehk võidetud on uuringu teostamine ühishanke vormis. Kõigil viimasel neljal aastal on üheks partneriks olnud SA Poliitikauuringute Keskus Praxis. Balti Uuringute Instituut on kaasatud olnud viimase kolme laine puhul. Aastatel 2011 ja 2020 oli läbiviijate seas ka Tartu Ülikool ning aastatel 2015 ning 2020 Tallina Ülikool.

Elanikkonna küsitlust teostas aastatel 2000-2008 uuringufirma Saar Poll, 2011. aastal viis seda läbi AS Emor ning alates 2015. aastast on läbiviijaks olnud Turu-uuringute AS. Raporti autorite nimekiri ehk analüüsi teostajad on samuti ajas muutunud, silma jääb nende seast üks nimi – Raivo Vetik, kes on monitooringu läbiviimisega seotud olnud 2000. aastast peale ja pole autorite sekka kuulunud vaid 2011. aastal. Ülevaatlik tabel monitooringu läbiviijatest on toodud lisas 2.

Ajalise võrdluse kontekstis on väga oluline, et metoodika kirjeldus ja tehtud metoodilised valikud oleks teada, kuid uuringumetoodika kohta oli aruannetes eraldi peatükk vaid viimasel neljal aastal (2011-2020). Aastatel 2000-2005 oli nipsisõnaliselt mainitud aruande sissejuhatavas peatükis valimi suurust ning rahvuslikku jaotust. Sealjuures tasub märkida, et 2000. aasta raport on veebis kättesaadav ainult peatükkide kaupa ning Marju Lauristini poolt kirjutatud sissejuhatuse leiab vaid paber kandjal aruandest. 2008. aasta raportis puudus info uuringumetoodika kohta. 2011-2020

aasta hankedokumentide läbivaatamisel selgus, et uuringumetoodika peatüki lisamine on olnud nõue alles viimasel neljal korral.

Aruannetest selgus, et aastatel 2000 ja 2015-2020 kasutati kombineeritud uurimisviisi ning 2002-2008 kvantitatiivset lähenemist. Aastal 2000 koosnes monitooring kahest eraldi osast: küsitlusandmetel põhinevast üldmonitooringust ning ajakirjanduse sisuanalüüsil põhinevast meediamonitooringust. Meediamonitooringu osa oli ühekordne ning hilisematel aastatel seda enam ei teostatud. 2002., 2005. ja 2008. aasta monitooringus käsitleti ainult küsitlusuuringu andmeid.

Esimese nelja aasta kohta ei selgu aruannetest, mil viisil küsitlust läbi viidi. Alates 2011. aastast on küsitlusuuring toimunud silmast silma intervjuu vormis. Hankedokumentides on kirjas, et vähemalt viimasel neljal aastal on olnud tellijapoolne nõue teostada järelkontrolli küsitlustöö kvaliteedi kontrollimiseks. Aruannete metoodika peatükkidest selgub, et seda ka teostati (2011-2020). Järelkontrolli käigus võeti ühendust ligikaudu 10%-ga vastajatest. 2011. aastal kinnitasid järelkontrolli tulemused, et küsitlustöö viidi läbi vastavalt vastaja valiku reeglitele ning koolituse nõuetele. 2015. ja 2017. aasta aruannetes toodi välja, et järelkontroll küsitlejate töös eksimusi ei tuvastatud. 2020. aastal kirjeldati aruandes, et reeglite vastu eksinud küsitlejate töö tulemus eemaldati valimist.

Alates 2011. aastast on monitooringu raames korraldatud ka täiendavaid fookusgrupi intervjuusid. Järgnevas tabelis 1 on loetletud aastate lõikes fookusgrupi sihtrühmad.

Tabel 1. Integratsiooni monitooringu raames läbi viidud fookusgrupi sihtrühmad aastate lõikes

| 2011 | 2015 | 2017 | 2020 |
|---|---|--|------------------|
| <ul style="list-style-type: none"> • uusimmigrandid; • kodanikuühendused; • tööandjad; • töötukassa töötajad ja tööturuprojektide läbiviijad; • vene õppekeele koolide õpetajad. | <ul style="list-style-type: none"> • liitgümnaasiumis õppivad noored; • kõrgharidusega venekeelsed noored tööturul; • venekeelse meedia esindajad. | <ul style="list-style-type: none"> • vene- ja eestikeelsed osalejad (tööturu teemadel); • uussisserändajad | Pole täpsustatud |

2020. aasta monitooringu metoodika peatükis ega ka mujal raportis ei olnud mainitud, kellega fookusgrupi intervjuusid läbi viidi. Ainuke viide nende toimumisele oli aruande 9. peatükis järgnevalt kirjeldatud (Eesti ühiskonna lõimumise monitooring 2020: 112):

„Et Narva linna näitel selgitada välja Ida-Virumaa ülejäänud Eestiga parema sidustamise potentsiaal, tegime EIM 2020 käigus hulga fookusrühma- ja süvaintervjuusid ning esitame siin nende olulisimad tulemused.“

Viimasel kahel monitooringu aastal on lisaks elanikkonna küsitlusele ja fookusgrupi intervjuudele korraldatud uussisserändajate veebiküsitlus ning viimasel aastal (2020) ka tagasipöördujate veebiküsitlus, mille andmete analüüsi tulemused on esitatud aruannetes nende jaoks eraldi peatükkides. Fookusgrupi intervjuude ning veebiküsitlusega seonduvad andmeid käesolevas töös ei käsitleta, seega järgnevalt on esitatud andmed vaid elanikkonna küsitlusuuringu kohta.

3.1.2 Küsitlusuuringu üldkogum

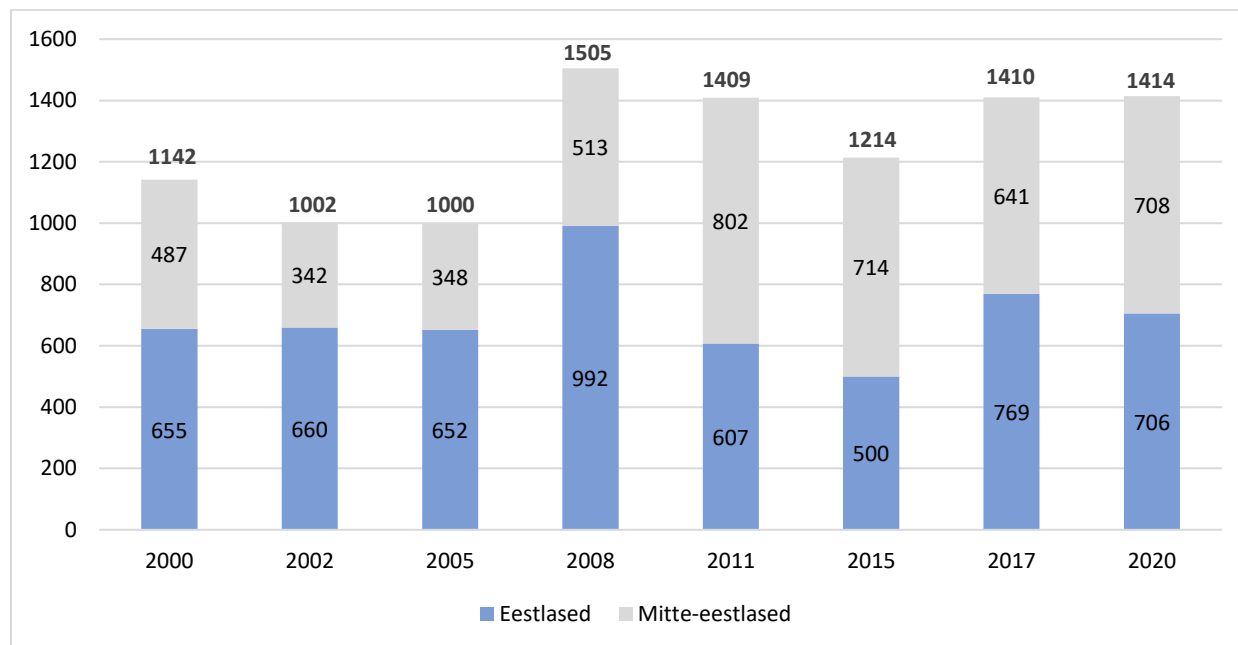
Elanikkonna küsitluse üldkogumit on monitooringu aruannetes defineeritud üldiselt väga napisõnaliselt. Aastatel 2000 ja 2002 on mainitud, et üldkogumi moodustavad Eesti elanikud. Aastal 2005 toodi välja, et küsitleti inimesi vanuses 15-74. Aastal 2008 puudus aruandes info populatsiooni kohta. 2011. aastal oli üldkogum defineeritud kui Eesti alalised elanikud vanuses 15-74 ning tegemist oli ka ainsa aastaga kui üldkogumi arvuline suurus oli aruandes välja toodud. Andmefailidest selgus, et viiel esimesel aastal küsitleti elanikke vanuses 15-74 aastat.

Alates 2015. aastast kaotati vanuseline ülempiir ehk üldkogumi moodustasid Eesti püsielanikud, kes olid vähemalt 15-aastased. Vanuselise ülempiiri kaotamise kohta küsis ka üks pakkujatest 2015. aastal hanke korraldamise ajal tellijapoolset selgitust, kas tegemist oli tahtliku muudatusega. Samuti pööras ta küsimuses tähelepanu trendiküsimuste võrreldavuse probleemile. Pakkujale vastati, et võrreldavuse tagamiseks tuleb endiselt küsitleda 15-74 aastaseid elanikke, kuid tellija jättis ülempiiri avatuks, et oleks võimalik koguda ka täiendavat infot vanematelt inimestelt, kui pakkuja ise seda oluliseks peab. Monitooringu andmetest selgus, et alates 2015. aastast on küsitlusuuringu raames küsitletud 15-aastaseid ja vanemaid Eesti püsielanikke. Seega tuleb eri aastate monitooringu andmete analüüsimisel veenduda, et analüüsitakse samu sihtgrupe.

3.1.3 Küsitlusuuringute valimid

Valimi koostamise ja -tüübi kohta esimesel neljal aastal monitooringutes info puudub. 2000-2005 aasta aruannetes on vaid mainitud, et valimid olid esinduslikud eestlaste ja muu rahvuse suhtes. Selle väite põhjal ei saa aga kindlalt väita, et kasutati tõenäosuslikku valimit, sest ka

mittetõenäosuslikud valimid võivad olla esinduslikud mingi konkreetse tunnuse suhtes. Kui vaadata lisaks monitooringute andmefailide, võib näha, et kõigis andmestikes on valimipunkti tunnus. Sellest lähtudes võiks järeldada, et kuigi valimi koostamise info aruannetes puudub, võis tegemist olla mitmeastmelise juhuvalimiga, kus esimeses faasis valitakse valimipunktid ning teises alles konkreetsed küsitletud inividid. Samas on aga niivõrd puuduliku info pealt järelduste tegemine riskantne ning väita, et tegemist oli esimestel aastatel juhuvalikuga, tegelikult ei saa.



Joonis 1. Valimimahud aastate lõikes

Küll aga oli kõigis aruannetes välja toodud lõplikud valimimahud, mis on ka esitatud joonisel 1. Jooniselt võib näha, et aastate lõikes on valimimahud mõnevõrra erinenud. Esimesel kolmel aastal jäi valimi suurus vahemikku 1000-1142. 2008. aastal läbi viidud küsitlusuuringus oli valimit hüppeliselt suurendatud (1505 vastajat), kuid kuna puudub info metoodika kohta, siis ei ole põhjus siinkohal teada. Aastatel 2011, 2017 ja 2020 oli valimimaht ligikaudu 1400 inimest, mis oli ka hankedokumentides ette antud minimaalne valimimahu nõue. 2015. aastal sellist nõuet kirjas ei olnud ning lõplikuks valimiks kujunes 1214 inimest. Lisaks võib jooniselt märgata, et alates 2011. aastast on muutunud eestlaste ja mitte-eestlaste osakaal valimis. Seda on ka aruannetes välja toodud, et muust rahvusest elanikkonna osakaalu on teadlikult valimis suurendatud (muutes sealjuures valimi üldjaotuse eestlaste ja mitte-eestlaste suhtes ebaproportsionaalseks).

Põhjendatud on seda sooviga tagada muust rahvusest elanikkonna alamgruppides arvatavate näitajate piisav statistiline usaldusväärsus.

Lisaks selgus analüüsist, et 2011. ja 2015. aastal koostati eraldi valimimudelid eestlaste ja muust rahvusest sihtrühmade jaoks. Kahjuks ei selgu, mida selle all täpsemalt mõeldud on. 2017. aastal koostati valim samuti kaheosalisena, kuid mitte enam rahvuse lõikes, vaid esimese osa moodustas üle-eestiline juhuvalim (N= 1000) ja teise osa täiendavalt 400 muu rahvuse esindajat, sealjuures teise osa puhul ei selgu, kas tegemist oli juhuvalikuga. 2020. aasta meetodikapeatükist ei selgu, et valim oleks koostatud mitmeosalisena. Küll aga võib analüüsi tulemusel öelda, et viimase nelja monitooringu laine (2011-2020) valimid kõik moodustati mitmeastmelisena.

2011. aastal oli meetodikaaruandes täpsustatud, et kasutati kihistatud kaheastmelist valikut, kus esmalt jagati üldkogum elukoha järgi territoriaalselt kihtideks ning seejärel teostati igas kihis kaheastmeline valik: esmalt valiti juhuvalikuga valimipunktid ning seejärel teisesed valikuühikud – inimesed. Vastajate leidmiseks kasutati lähteadressi meetodit ning noore mehe-noore naise reeglit. Lähteadressi meetod tähendab, et juhuslikult valitud lähteadressilt liigutakse edasi teatud reeglistiku alusel (näiteks iga teine maja või iga viies korter). Noore mehe-noore naise reegel tähendab, et küsitletakse kodus olevatest meestest kõige nooremast ning kui mehi kodus parasjagu ei ole, eelistatakse nooremast naist. Taoline valikumeetod annab küsitletavale kategooriale, keda on raskem tabada (noored ja mehed), suurema šansi valimisse sattuda ning selle abil soovitakse saavutada paremat soolist ja eri vanusegruppide esindatust valimis.

Sarnast valimi koostamise protseduuri oli ka kirjeldatud 2015., 2017. ja 2020. aasta monitooringus. Mainitakse, et valimipunktid on valitud territoriaalse esinduslikkuse põhjal, kuid ei täpsustata, et Eesti oleks regionaalselt jaotatud kihtideks (nagu seda tehti 2011. aastal) või kuidas täpselt mainitud territoriaalne esinduslikkus saavutati. Seega ei saa eeldada kihistatud valikut. Küll aga toimiti valimipunktides vastaja valikul sarnaselt ehk kasutati lähteadressi meetodit ning noore mehe-noore naise reeglit. Samas tuleb siinkohal mainida, et lähteadressi meetodi reeglistikku aruannetes täpselt välja toodud ei olnud.

Analüüsist selgub, et koostatud valimid on tegelikult aadressi mitte indiviidipõhised. Aadressipõhise lähenemise puhul võetakse valimisse igalt aadressilt vaid 1 indiviid ja seega on indiviididel, kes elavad üksi või väikeses peres, tõenäosus valimisse sattuda suurem. Aruannetes

ei ole aga välja toodud, kas sellega on valimi mudeli koostamisel arvestatud või mitte. Samuti ei ole seda teadaolevalt arvesse võetud kaalude arvutamisel. Noore mehe-noore naise ja lähteadressi meetodi kasutamine seab omakorda kahtluse alla valimi juhuslikkuse. Lähteadressi meetodi puhul on küsitljal tavalisest suurem võimalus süstemaatilist lähenemist valiku tegemisel vältida ja seeläbi valimi juhuslikkust mõjutada.

Metoodikakirjeldustest ei selgu, mil viisil küsitlustöö korraldatud on ehk millistel nädalapäevadel ja kellaaegadel küsitlust läbi viidi või millised olid ette seatud reeglid, kui valikusse sattunud aadressilt kedagi kätte ei saadud, mitu korda seejärel uuesti prooviti ning kas vastajat püüti tabada ka õhtutel ja nädalavahetusel. Seega on raske hinnata millise tõenäosusega tööealised noored (ja ennekõike mehed) võisid näiteks tööl viibimise tõttu valimist välja jääda ning kuivõrd (ja kui süstemaatiliselt) noore mehe-noore naise reeglina valimisse võetud indiviidid nendest erineda võisid. Lisaks oli 2015. ja 2017. aasta metoodikaraportis kirjeldatud, et vanemaealiste üleesindatuse vältimiseks anti küsitlejatele ette kvoodid, kui mitut üle 60aastast võis ühes valimipunktis küsitleda.

Kokkuvõtlikult võib öelda, et monitooringu kõigi aastate (va 2008, mille puhul info puudus) valimid on koostatud mitmeastmelise valikuga ning valiku juhuslikkus on küsitav, kuna varasematel aastatel puudub täpsem info selle hindamiseks ning hilisematel aastatel on kasutatud kvote, noore mehe-noore naise ja lähteadressi reeglit, mis kõik võivad omada mõju juhuslikkusele. Samuti ei saa väita, et need monitooringu aastad, kus valimimaht oli suurem, on populatsiooni suhtes esinduslikumad, kuna esinduslikkus sõltub lisaks valimi mahule ka koostamisel tehtud valikutest ja saadaoleva info pealt on valimi koostamise kvaliteeti raske hinnata. Seega, kui rääkida eri aastate võrreldavusest, on dokumenteeritud info põhjal raske teha põhjapanevaid järeldusi. Võib näha, et eri aastatel (ning erinevate uuringufirmade poolt) on kasutatud mõnevõrra erinevaid lähenemisi valimi koostamisel. Autori tõstatatud küsimused valimi juhuslikkuse ja esinduslikkuse suhtes on osaliselt tingitud siiski dokumentatsiooni puudulikkusest ning ei pruugi tähendada, et monitooringu lained ei ole valimite koostamise põhimõtete poolest üksteisega võrreldavad.

3.1.4 Küsitlusuuringu vastamismäär ning arvutatud kaalud

Analüüsist selgus, et küsitlusuuringu vastamismäär oli välja toodud vaid 2011. aasta raportis. 2015-2020 aasta raportites oli aga olemas küsitlustöö aruanne, mille pealt on võimalik vastamismäärad ise arvutada. Valimiviga oli välja toodud vaid kolmel aastal: 2011, 2015 ja 2017. Aruannetes oli esitatud tulemuste usalduspiiride hindamise tabel, millest sai välja lugeda, et üldvalimi (eestlased + muud rahvused) maksimaalseks veapiiriks oli kõigil 3 aastal ligikaudu 3%.

Samuti oli vaid nendel kolmel aastal (2011, 2015 ja 2017) toodud metoodika peatükis infot kaalude arvutamise kohta, kuigi monitooringu andmefailidest võib näha, et kaalutunnus on kõigi aastate andmestikes olemas. Sealjuures on viimase nelja aasta puhul, kui andmestikud olid lahku löödud rahvuse põhisel, igas andmestikus oma kaal (eestlaste, muu rahvuse ja üldandmestiku kaal). Kuna üheski metoodikapeatükis ei ole mainitud, et arvutatud oleks disainikaalud, siis võib kaalude arvutamise kirjeldusest järeldada, et failides toodud arvutatud kaalude puhul on tegemist järelkihistuskaaludega. Samas ei selgu siiski täpsem lähenemine kaalumisel, kas kasutatud on klassikalist järelkihistuskaalumist või näiteks iteratiivset järelkihistamist (*raking*).

Eelnevalt mainitud kolme aasta aruannetes oli toodud tabel valimi koosseisust kaalumata ja kaalutud kujul järgnevate sotsiaaldemograafiliste tunnuste lõikes: sugu, vanus, rahvus, regioon, asulatüüp. Kuigi aruannetes mainitakse soo, vanuse, rahvuse ja regiooni arvestamist kaalumisel, ei saa päris kindel olla, millised konkreetsed tunnused ja milliste kategooriatega on kaalude arvutamise aluseks olnud, näiteks millised vanusegrupid ja regioonid järelkihistamise aluseks on võetud. 2011. aasta puhul võib arvata, et kaalude arvutamise aluseks on tabelis toodud tunnused (sugu, vanus, rahvus, regioon, asulatüüp), kuna kõrvale on toodud ka vastavate tunnuste jaotused üldkogumis, mis on kõikide tunnuste puhul väga sarnased. 2015. ja 2017. aasta puhul sama järeldust päris kindlalt teha ei saa, kuna näiteks regiooni puhul on tabelis toodud kuuene jaotus (Tallinn, Põhja-Eesti, Ida-Virumaa, Lääne-Eesti, Kesk-Eesti ja Lõuna-Eesti), samas kui metoodikaaruande tekstis on mainitud, et kaalumisel võeti aluseks vaid kolmene jaotus (Tallinn, Ida-Virumaa ja muu Eesti). Samas ei pea kaalumise nüansid olema aastati samasugused, vaid võivad erineda sõltuvalt andmekogumise õnnestumise eripäradest, kuid detailsem info kaalumise kohta annaks kasutajale võimaluse andmeid teadlikumalt analüüsida.

3.2 Ankeetküsimuste samasuse analüüs

Et läbi viia ankeetküsimuste samasuse analüüsi, pidas autor esmalt vajalikuks kaardistada ära kõik aastatel 2000-2020 lõimumismonitooringu elanikkonna küsitluses esinenud küsimused. Selleks sai koostatud küsimuste maatriks, kus veergudes olid vastava monitooringu aasta ning ridadesse sai järjest lisatud esitatud küsimused. Kuna autoril puudus eelnev arusaam, kuivõrd küsimused aastate lõikes muutunud on, osutus see oluliselt ajamahukamaks ülesandeks, kui esialgselt oodatud. Algselt oli planeeritud ära kaardistada 2020. aasta küsimused koos vastusevariantidega ning seejärel kõigi eelnevate monitooringu lainete kohta märkida, kas küsimus esines vastava aasta ankeedis ning kui küsimuse sõnastuses või vastusevariantide hulgas oli erinevusi, siis need kirja panna. Samuti oli jooksvalt plaanis küsimuste nimekirja täiendada, kui eelnevatest aastatest ilmses küsimusi, mida 2020. aastal ei küsitud.

Antud magistritöös on käsitletud eraldi küsimusena igat ankeetküsimustiku objekti, millele vastaja on pidanud vastuse andma. See tähendab, et kui näiteks ankeedis on küsitus usaldust erinevate institutsioonide vastu, siis käesolevas magistritöös on igat väljatoodud institutsiooni, mille kohta vastaja hinnangut soovitakse teada, käsitletud eraldi küsimusena ehk küsimustiku objektina (inglise keeles *survey item*). Seega kaardistas autor näiteks institutsioonide usalduse küsimuse 2020. aasta kohta nii nagu näidatud tabelis 2. Lisaks märkis autor kommentaarina juurde ankeedis välja toodud vastusevariandid/skaala.

Tabel 2. 2020 Institutsioonide usalduse küsimuse kaardistamine maatriksis

| | 2020 |
|---|--------------------------------|
| Kuivõrd Te usaldate järgnevaid institutsioone ja rühmi Eestis? | |
| | <i>Valitsus</i> |
| | <i>Riigikogu</i> |
| | <i>President</i> |
| | <i>Politsei</i> |
| | <i>Kohtusüsteem</i> |
| | <i>Kaitsevägi</i> |
| | <i>Kohalikud omavalitsused</i> |
| | <i>Haridussüsteem</i> |
| | <i>Tervishoiusüsteem</i> |
| | <i>Kodanikuühiskond</i> |

Pärast 2020. aasta küsimuste kaardistamist oli maatriksis 337 rida. Juba 2017. aasta küsimusi kodeerides hakkas autor aru saama, et küsimused on päris palju muutunud ning ainult muudatuste kirja panemine ei anna lõpuks head ülevaadet, sest muutusi küsimustes esines palju. Seega otsustas autor sarnaselt eelnevale aastale kirja panna kõik küsimused koos vastusevariantidega ka siis, kui sõnastus oli jäänud samaks. Iga küsimuse kaardistamisel vaatas autor, kas seda küsimust on sarnaselt eelnevatel aastatel küsitud ning sel juhul märkis esinenud küsimuse kohakuti eelneva aastaga. Kaardistamist tegi autori jaoks raskemaks asjaolu, et küsimused esinesid ankeetides väga erinevas järjekorras ning olles vahepeal kategoriseeritud isegi erinevate teemade alla. Pärast kõikide lainete (8 eestikeelset küsimustikku) kaardistamist kujunes maatriksi lõppsuuruseks 9 veergu ning 1760 rida (varieeruvalt ca. 200-400 küsimustiku objekti ühe aasta kohta).

Seega selgus juba maatriksi koostamisel, et ankeetküsimused on aastate lõikes võrdlemisi palju muutunud. Analüüsist tuli välja, et küsimusi, mida oleks läbivalt monitooringu ankeetides küsitud oli väga vähe. Jättes kõrvale taustaküsimused, saab öelda, et ei leidnud mitte ühtegi küsimust, mida oleks täpselt samas sõnastuses ja samade vastusevariantidega küsitud kõigis kaheksas monitooringus. Seega otsustas autor ankeetküsimuste samasuse analüüsi kaasata just need niinimetatud võtmeküsimused, mida läbivalt on kõige enam aastate jooksul küsitud. Nende küsimuste sõnastuse analüüs on esitatud alapeatükis 3.2.2. Järgnevalt on toodud mõned üldised tulemused, mis maatriksi koostamisel ning küsimuste analüüsimisel esile kerkisid.

3.2.1 Ankeetküsimuste analüüsi üldised tulemused

Monitooringu andmete analüüsi teostatakse enamasti rahvuse lõikes. Sellele viitavad nii eri aastate uuringuraportid kui ka alates 2011. aastast rahvuse põhjal lahku löödud valimid ja andmestikud. Kuigi rahvust oli kõigi aastate ankeetides vastajatelt küsitud, erinesid aastad selle poolest, kas märkida võis end ainult ühe rahvuse esindajaks või said vastajad valida ka mitu vastusevarianti. Aastatel 2000, 2002 ja 2005 sai valida etteantud variantide seast vaid ühe sobiliku. Alates 2008. aastast on võimalik määratleda end mitme rahvuse esindajana. Kui aastatel 2011-2020 esitati küsimus kaheosalisena – millist rahvust pidas vastaja enda jaoks peamiseks ning millise rahvuse esindajaks ta end veel täiendavalt pidas, siis aastal 2008 oli andmestikus ainult 1 küsimus, kus respondentil oli võimalik mitu varianti korraga valida.

Seega tekkis autoril küsimus, mille alusel on 2008. aastal analüütikud vastaja rahvust määratlenud (sest analüüsi tulemused esitati raportis enamasti rahvuse lõikes: eestlased versus muu rahvus). Raporti tulemusi ja andmefailides mõne esitatud tunnuse sagedusjaotust kõrvutades selgus, et 2008. aastal on analüüsid kasutatud vastaja rahvuse määramiseks ankeedi keele tunnust ehk eestikeelsele ankeedile vastajad määratleti eestlastena ning venekeelsele ankeedile vastajad muu rahvusena.

Samuti selgus analüüsi käigus, et monitooringud erinevad aastati selle poolest, mis keeles võisid eestlased ankeedile vastata. 2015. ja 2017. aasta metoodikaraportis tuuakse välja, et kui muust rahvusest respondendid said valida intervjuu keele ise, siis eestlasel oli võimalik vastata vaid eestikeelsele ankeedile, kuna eesti- ja venekeelsed ankeedid erinesid osaliselt küsimuste poolest. Teistel aastatel ei ole sama infot raportites välja toodud, kuid andmefailide analüüsid selgus, et ka 2005. aastal vastasid eestlased vaid eesti keeles. Ülejäänud aastatel paistab, et eestlased said valida, kas vastavad eesti- või venekeelsele ankeedile.

Huvitav on siinkohal ka asjaolu, et 2017. aastal märkis 30 vastajat, kes olid end eestlasena määratlenud, et nende peamiseks suhtluskeeleks on vene keel. Samas oli 2017. aastal eestlastel võimalik vastata vaid eestikeelsele ankeedile. Siin tekib küsimus, kas rahvuse küsimus esitati enne või pärast intervjuu keele määramist. Kui venekeelse emakeelega respondentidel paluti eesti keeles vastata, siis kuivõrd võis nende eesti keele oskuse tõttu olla mõjutatud ankeetküsimuste vastused (kas neil võis tekkida raskusi küsimustest arusaamisel) või kui rahvust küsiti alles hiljem, siis kuidas suhtuti sellesse, kui venekeelne vastaja soovis rahvuselt end määratleda eestlasena.

Seega tasub rahvuse lõikes eri aastate andmete võrdlemisel arvesse võtta eelnevalt väljatoodud asjaolusid. Soovitav on veenduda, et rahvus oleks analüüsis määratletud samadel alustel ning olenevalt analüüsitavatest muutujatest võtta arvesse, kas eestlaste seas oli venekeelseid vastajaid ning hinnata enne analüüsi teostamist, kas ja mil määral see tulemusi võib mõjutada.

Nagu eelnevalt mainitud, selgus küsimuste maatriksi koostamise käigus, et ankeetküsimustikud on aastati üsnagi palju muutunud. Üldiselt on monitooringutes läbivalt käsitletud teemasid nagu tööturg, haridus, kontaktid ja keeleoskus, meedia, ühiskondlik osalus, võrdsed võimalused ning üldised lõimumisalased hoiakud. Monitooringu algusaastad paistavad silma üsnagi eripalgeliste küsimuste poolest. Esines rohkem spetsiifilisi küsimusi, näiteks erinevate seadusmuudatuste kohta

(keeleseadus, valimisseadus jms), samuti teemasid, mida hiljem nii detailselt ei käsitletud nagu majanduslik toimetulek või vastajate vene- või euroopameelsus. Analüüsist selgus, et mida aasta edasi, seda konkreetsemaks muutusid küsimused ja koondusid rohkem põhiteemade ümber.

2011.-2020. aasta hankedokumentidest selgub, et tellija on omalt poolt alati hankesse sisse kirjutanud, et küsimustikud peavad põhinema varasema monitooringu ankeetidel ning sisaldama ajaliste trendide jälgimiseks ka eelnevate monitooringute küsimusi. Lisaks on rõhutatud, et ankeetküsimustiku muudatusi võib sisse viia vaid kokkuleppel hankijaga. Samas ei selgu, millised konkreetset küsimused peaksid ajas muutumatuks jääma. Hankedokumentides liigitatakse ankeetküsimused kolme ploki: 1) põhiplokk, mille küsimusi ei tohi muuta, 2) roteeruv plokk, kuhu alla kuuluvad teemad, mida võiks uurida teatud regulaarsusega, kuid mitte iga monitooring ning 3) lisaplokk ehk hetkel aktuaalsed teemad.

Kahjuks ei ole võimalik küsimustike alusel kindlaks teha, millised küsimused kuhu ploki kuuluvad. Samas võib see pakkuda selgitust, miks on niivõrd vähe küsimusi, mida läbivalt iga aasta küsitud on. Põhiploki kohta on hankedokumentides kommenteeritud, et seal sisalduvad küsimused tulenevad suuresti arengukava eesmärkidest ja mõõdikutest.

Sellest lähtuvalt võiks eeldada, et omavahel on rohkem võrreldavad samal arengukava perioodil läbi viidud monitooringute andmed ehk 2000-2005, 2008 ja 2011 ning 2015-2020. Analüüsist selgub, et teataval määral vastab see eeldus ka tõele. Varasemate aastate (2000-2005) ankeedid on üksteisele rohkem sarnased võrreldes hilisemate ankeetidega. Samas võib see olla tingitud ka asjaolust, et esimese nelja monitooringu teostajad olid samad – Tallinna Ülikooli teadlased.

Monitooringu läbiviija mõjule viitavad ka suured erinevused 2008. ja 2011. aasta ankeediküsimustes. 2011. aastal teostas küsitlust (eelnenud Saar Polli asemel) Emor ning uuringut viisid läbi Tartu Ülikool koostöös Poliitikauuringute Keskusega Praxis (vaata ka magistr töö lisa 2). 2011. aasta aruandes mainitakse esimest korda, et ankeedi piloteerimise tulemusel tuli küsimustikku lühendada. Samuti toovad uurijad välja, et võrdlust varasemate aastatega on teostatud nii palju kui metoodiliselt võimalik ning et ankeetide küsimused on aastate lõikes oluliselt varieerunud (Integratsiooni monitooring 2011: 8).

2015. ja 2017. aasta küsimustikud on võrdlemisi sarnased. 2020. aasta küsimustiku loomisel on juba arvestatud Balti Uuringute Instituudi poolt läbi viidud lõimumismonitooringu metoodika

rekontseptualiseerimise aruandes toodud soovitusetega (Kaldur jt 2019) ja seetõttu on küsimustikus taas märgata olulisi muudatusi võrreldes eelnenud aastatega.

Nagu aga eelnevalt mainitud, siis üldised teemad ja kontseptsioonid, mida monitooringutega on soovitud katta, on aastate lõikes püsinud samad. Muutunud on võib-olla see, kuidas mingit konstrukti mõõdetakse ehk kuidas on mingi konstrukti operatsionaliseeritud küsimuseks. See, et sarnaseid küsimusi aastate lõikes esineb vähe, ei tähenda, et sama konstrukti ei püütud erinevatel aastatel mõõta erinevate küsimustega. Siinkohal tuleb aga andmete võrdlemisel ja järelduste tegemisel olla ettevaatlik ja läheneda igale küsimusele individuaalselt. Mõnikord võib aidata vastusevariantide või eri tunnuste kokku kodeerimine, kuid see vähendab hinnangute täpsust ning sellega peab tulemuste tõlgendamisel arvestama.

Üldisemate tähelepanekute kokkuvõtteks esitab autor mõned olulisemad punktid, millele lisaks küsimuse sõnastusele tähelepanu pöörata, kui hakata eri aastate andmeid võrdlema:

- Kas ning kuivõrd on vastuseskaala muutunud?
- Kellelt küsimust on küsitud – erinevate aastate ankeetides esineb küsimuste suunamist erineval määral. Seega tasub kindlaks teha, et võrreldavad sihtrühmad oleks samad. Näiteks on mõnel aastal küsimust küsitud rahvuse, teisel aastal ankeedi keele alusel. Või näiteks küsiti tööga rahulolu mõnel aastal ainult töötavate respondentide käest, aga mõnel aastal kõigi vastajate käest.
- Kas küsimusele on võimalik vastamata jätta? Ehk kas esineb vastusevariant „Ei oska öelda“, „Raske öelda“ vms. Kui küsimuse sõnastus on aastate lõikes jäänud samaks, aga mõnel aastal on võimalik jätta vastamata ja mõnel mitte, võib skaala muutus kaasa tuua erisusi vastajate jaotuses, millega tuleb tulemuste tõlgendamisel kindlasti arvestada.

3.2.2 Võtmeküsimuste samasuse analüüs

Järgnevalt on toodud kuue niinimetatud võtmeküsimuse samasuse analüüs. Küsimuste valikul lähtus autor kolmest aspektist: 1) küsimust küsiti kas kõigis või vähemalt kuues eri laines 2) küsimused oleksid erinevatest teemadest 3) küsimust küsiti ka 2020. aastal. Viimane kriteerium oli autori jaoks oluline, sest pärast 2019. aastal Balti Uuringute Instituudi poolt läbi viidud metoodikaanalüüsi vähendati küsimustikus olulisel määral küsimuste arvu, ning autor ei pidanud praktikas vajalikuks analüüsida küsimusi, mida tuleviku monitooringutes enam võib-olla ei

kajastata. Nagu ka metoodika peatükis sai kirjeldatud, toetus autor küsimuste analüüsimisel järgnevatele aspektidele:

- Erinevused küsimuste sõnastuses;
- Erinevused küsimuste vastusevariantides/skaalades;
- Küsimuse lihtsus, arusaadavus ja mõistetavus;
- Küsimuse eeldused (suunamised/tingimused).

Valituks osutunud küsimused on esitatud käesolevas peatükis tabelites 3-8. Tabelites on toodud monitooringu aastata lõikes esitatud põhiküsimus, alaküsimused ning vastusevariandid. Samuti on tabelis küsimuse number ankeedis/andmestikus. Tabelites võib ka märgata, et mõne aasta puhul võib olla esitatud rohkem kui üks küsimus. Täpsustused nende küsimuste kohta on toodud juba vastava küsimuse analüüsi juures.

Küsimus 1. Seotus Eestiga tulevikus

Esimene valitud küsimus on toodud tabelis 3 ja sellega on küsitud vastajalt, kuivõrd ta plaanib siduda oma tulevikku Eestiga. Tabelist võib näha, et küsimust on küsitud enam-vähem sarnase sõnastusega alates 2005. aastast: „*Kuivõrd Te näete oma tulevikku seotuna Eestiga? Kas Te lahkuksite Eestist teistesse riikidesse?*“. Problemaatiliseks võib sellise sõnastuse puhul pidada, et vastajalt küsitakse tegelikult korrakahte asja: ühelt poolt võib vastaja sõna *kuivõrd* tõlgendada kui *mil määral* või *mis ulatuses* nähakse oma tulevikku seotuna Eestiga, samas järgnev lause „*Kas lahkuksite Eestist.*“ viitab jällegi jah/ei vastusele. Kuigi sõnastus võib olla läbivalt enam-vähem sarnane, siis võib näha, et aastal 2011 küsiti sama küsimust täpsustavalt mitte ainult vastaja enda, vaid ka tema laste, muude pereliikmete ja sugulaste kohta. Aastal 2002. esitati ankeedis küsimus sõnastuses „*Kus Te tahaksite tulevikus elada ja töötada?*“ ning lisaks eraldi küsimus laste kohta: „*Kus Te tahaksite, et Teie lapsed tulevikus elaksid ja töötaksid?*“.

Nii 2000. aasta kui ka 2020. aasta küsimustes täpsustatakse vastaja jaoks, et mõeldud on, kus vastaja tulevikus elada soovib. Kui valdavalt on küsimuse sõnastus alates 2005. aastast jäänud samaks, tuleb siinkohal tähelepanu pöörata vastusevariantidele, mis annavad tegeliku sisu sellele, mida küsida on tahetud.

Tabel 3. Võtmeküsimus number 1: seotus Eestiga tulevikus

| Aasta | Küs. nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------|--|--|--|
| 2020 | 6.8. | Kuivõrd Te näete oma tulevikku seotuna Eestiga - kas Te sooviksite Eestist lahkuda, püsivalt mujale elama asuda? | | Kindlasti ei soovi Pigem ei soovi (võib-olla kunagi) Sooviksite lahkuda ajutiselt, aga mitte püsivalt Sooviksite lahkuda püsivalt Muu, täpsustage _____ Ei oska / soovi vastata |
| 2017 | Q42 | Kuivõrd Te näete oma tulevikku seotuna Eestiga? Kas Te lahkuksite Eestist teistesse riikidesse ... | | alaliseks, kui selleks vaid avaneks võimalus ajutiselt, seoses töötamise või õppimisega mujal Te ei soovi Eestist lahkuda Ei oska öelda |
| 2015 | Q42 | Kuivõrd Te näete oma tulevikku seotuna Eestiga? Kas Te lahkuksite Eestist teistesse riikidesse? | | Jah, lahkaksin alaliseks, kui selleks vaid avaneks võimalus Jah, lahkaksin ajutiselt, seoses töötamise või õppimisega mujal Ei soovi Eestist lahkuda Ei oska öelda |
| 2011 | q108 | Kas Teie või Teie lapsed/lapselapsed tahaksite Eestist lahkuda, püsivalt mujale elama asuda? | q108_1 Mina ise q108_2 Lapsed või lapselapsed q108_3 Mõni muu pereliige q108_4 Mõni sugulane või tuttav | Kindlasti mitte Ei oska praegu öelda, kunagi võib-olla Mõneks ajaks küll, aga mitte päriselt Jah, tahaksin päriselt ära minna Ei tea, raske öelda |
| 2008 | K75 | Kuivõrd Te näete oma tulevikku seotuna Eestiga? Kas Te lahkuksite Eestist teistesse riikidesse? | | jah, lahkaksin alaliseks, kui selleks vaid avaneks võimalus jah, lahkaksin ajutiselt, seoses töötamise või õppimisega mujal ei soovi Eestist lahkuda RASKE ÖELDA |
| 2005 | K63 | Kuivõrd Te näete oma tulevikku seotuna Eestiga? Kas Te lahkuksite Eestist teistesse riikidesse? | | jah, lahkaksin alaliseks, kui selleks vaid avaneks võimalus jah, lahkaksin ajutiselt, seose töötamise või õppimisega mujal ei soovi Eestist lahkuda RASKE ÖELDA |
| 2002 | K19a | Kus Te tahaksite tulevikus elada ja töötada? | | Ainult Eestis Osalt Eestis, osalt Lääne riikides Osalt Eestis, osalt Venemaal Lahkuks täiesti Lääne riikidesse Lahkuks täiesti Venemaale MUU VASTUS RASKE ÖELDA |
| | K19b | Kus Te tahaksite, et Teie lapsed tulevikus elaksid ja töötaksid? | | |
| 2000 | - | - | | |

Vastusevariante võrreldes selgub, et aastal 2002 küsiti tegelikult, kas vastaja soovib tulevikus elada ainult Eestis, osaliselt Eestis ja Venemaal/Lääne riikides või lahkukuks täielikult Venemaale/Lääne riikidesse. Alates 2005. aastast ei küsita enam, kuhu respondent sooviks elama asuda, vaid kas üleüldiselt on plaanis lahkuda, sealjuures kas või ajutiselt. 2005, 2008, 2015 ja 2017 on seotud „ajutise lahkumise“ variant üksnes töö ja õppimisvõimalusega. Samuti on nendel aastatel kindlat lahkumise soovi avaldav vastusevariant „*Jah, lahkuksin alatiseks, kui selleks vaid avaneks võimalus*“ üsnagi emotsionaalselt laetud, mis tavapäraselt ei ole hea praktika, mida kasutada. Tonaalselt neutraalsemaid vastusevariante on kasutatud 2011. ja 2020. aasta küsimustikus.

Kuna vastusevariandid on aastate lõikes oluliselt erinevad, ei saa korrektselt erinevate aastate andmeid omavahel võrrelda. Vaadata saaks ehk kõikide lainete puhul, kui suur osa vastajatest ei soovi Eestist lahkuda ehk plaanib oma tulevikku siduda ainult Eestiga. Siinkohal tuleks aga tegelikult olla vastuste tõlgendamisel ettevaatlik, kuna alternatiivsed vastusevariandid siiski mõjutavad samuti respondendi otsust, milline vastus anda ning seepärast ei saa öelda, et kõik aastad oleks selles küsimuses hästi võrreldavad. Sarnaselt on esitatud küsimusi (koos vastusevariantidega) aastatel 2005, 2008, 2015 ja 2017, mille puhul võiks võrdlevat analüüsi teostada.

Küsimus 2. Uue töökoha leidmise võimalikkus

Järgmiseks valitud küsimuseks osutus töökoha leidmise võimalikkuse küsimus, mille sõnastused eri aastate lõikes on toodud tabelis 4. Vaadates küsimuste sõnastust võib tegelikult märgata, et küsitud on kahte sisuliselt erinevat küsimust: 1) kui vastaja kaotaks praeguse töökoha, kas tal sellisel juhul õnnestuks leida endale sobiv uus töökoht; 2) kui vastaja otsiks töökohta, kas tal õnnestuks enda hinnangul leida sobiv uus töökoht. Kui varasematel aastatel on küsitud emba-kumba küsimustest, siis 2020. aasta ankeedis võib näha mõlemat küsimust eraldi, mis viitab sellele, et ka ankeedi koostajad näevad neid kui eri küsimusi.

Aastatel 2000 ja 2002 küsiti vastajalt, kas ta leiaks uue töökoha, kui peaks oma töökoha kaotama. Sama küsimus on uuesti veidi täiendatult kujul sisse toodud 2020. aasta ankeedis – küsitakse „*Kui kaotaksite oma praeguse töökoha, kas Teil õnnestuks leida oma ametioskustele ja kogemustele vastav töökoht järgneva 3 kuu jooksul?*“. Vastajale esitatud küsimuses on esitatud selgitus, et mõeldakse oskustele ja kogemusele vastava töökoha leidmist (mitte lihtsalt ükskõik millise töö

leidmist) ning samuti on täpsustatud ajaraam – 3 kuud. Varasemate aastate küsimustes ajaraami esitatud ei ole, seega võivad respondendid küsimust erinevalt mõista – kes võtab seda kui ajamääratluseta põhimõttelist küsimust (kas üldse oleks võimalik töö leida) või kes raamistab enda jaoks mõtteliselt aktsepteeritava ajaraami (näiteks üks kuu või aasta). See on aga väga subjektiivne ning oleneb suuresti vastajast.

2005. ja 2008. aastal on küsitud „*Mida Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht?*“. Küsimus kõlab hüpoteetilise küsimusena ehk vastaja oletab, kas ta võiks töö leida olenemata vajadusest tööd otsida. Alates 2011. aastast täiendatakse küsimust sissejuhatava lausega „*Oletagem, et peaksite praegu uut tööd otsima*“ ning täpsustatakse vajadust tööd otsida „*Mis Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht?*“. „Vastuvõetav töökoht“ nõuab jällegi subjektiivset hinnangut ning vastajad võivad seda erinevalt mõista.

Nagu mainitud, esines 2020. aastal ankeedis kaks erinevat küsimust, millest ühte sai juba kirjeldatud. Teise küsimuse puhul oli sõnastus järgnev: „*Kui otsite või peaksite otsima endale töökohta, kas Teil õnnestuks leida oma ametioskustele ja kogemustele vastav töökoht järgneva 3 kuu jooksul?*“. Ühelt poolt võib lugeda heaks, et täpsustatud on ajaraamistik (3 kuud) ning millist töökohta mõeldakse (oma ametioskustele ja kogemustele vastav). Teisalt aga esitatakse küsimuses kaks erinevat situatsiooni (*kui otsite või peaksite otsima endale töökohta*), millega võidakse viidata eri vastajate võimalikele töösituatsioonidele. Samas oleneb sellise sõnastuse puhul jällegi vastaja tõlgendusest, kummale küsimusele ta tegelikult vastab. Seda küsimust oli 2020. aastal küsitud ainult nendelt respondentidelt, kes tööl käisid.

Oluline erinevus 2000.-2011. ja 2015.-2020. aasta küsimuste vahel seisneb ka selles, et esimeses neljas laines oli esitatud küsimus asukoha täpsustusega. Respondent pidi andma hinnangu oma töö leidmise võimalikkusele elupiirkonnas/Eestis või välismaal. Kusjuures täpsustused olid aastate lõikes erinevad. Välismaa alla võisid kuuluda kõik muud riigid või oli küsitud Soome, Venemaa ja EL riikide kohta eraldi. Skaalad on läbi aastate jäänud sarnaseks, paludes hinnata töökoha leidmise võimalikkust 5-astmelisel skaalal, kus kaks varianti on positiivsed („*kindlasti jah*“, „*tõenäoliselt jah*“), kaks varianti negatiivsed („*kindlasti mitte*“, „*tõenäoliselt mitte*“) ning viienda variandina oli toodud „*Ei oska öelda*“/“*Raske öelda*“/“*Ei soovi vastata*“. Samuti küsiti kõiki küsimusi kõigilt vastajatelt (ainuke erand – aasta 2020 küsimus 2.7).

Tabel 4. Võtmeküsimus number 2: uue töökoha leidmise võimalikkus

| Aasta | Küs. nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------|---|---|---|
| 2020 | 2.7 | Kui kaotaksite oma praeguse töökoha, kas Teil õnnestuks leida oma ametioskustele ja kogemustele vastav töökoht järgneva 3 kuu jooksul? | | Jah, kindlasti Tõenäoliselt jah Tõenäoliselt mitte |
| | 2.8 | Kui otsite või peaksite otsima endale töökohta, kas Teil õnnestuks leida oma ametioskustele ja kogemustele vastav töökoht järgneva 3 kuu jooksul? | | Kindlasti mitte Ei oska / soovi vastata |
| 2017 | Q13 | Oletagem, et peaksite praegu uut tööd otsima. Mis Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht? | | Kindlasti oleks Tõenäoliselt oleks Tõenäoliselt ei oleks Kindlasti ei oleks Ei oska öelda |
| 2015 | Q13 | Oletagem, et peaksite praegu uut tööd otsima. Mis Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht? | | Kindlasti oleks Tõenäoliselt oleks Tõenäoliselt ei oleks Kindlasti ei oleks Ei oska öelda |
| 2011 | q16 | Oletagem, et peaksite praegu uut tööd otsima. Mis Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht ... | q16_1...Eestis? q16_2...Venemaal? q16_3...Mujal välismaal? | Oleks Tõenäoliselt oleks Tõenäoliselt ei oleks Kindlasti ei oleks RÕ |
| 2008 | K18 | Mida Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht? | k18.1...Eestis? k18.2...välismaal? | Kindlasti jah Tõenäoliselt jah Tõenäoliselt mitte Kindlasti mitte RASKE ÖELDA |
| 2005 | K56 | Mida Te arvate, kas Teil oleks vajaduse korral võimalik saada Teile vastuvõetav töökoht ... ? | K56_a...Oma elupiirkonnas? K56_b...Mujal Eestis? K56_c...Soomes? K56_d...Mõnes muus Euroopa Liidu riigis (nt. Saksamaal, Iirimaal)? K56_f...Mõnes muus riigis? K56_e...Venemaal? | Kindlasti jah Tõenäoliselt jah Tõenäoliselt mitte Kindlasti mitte RASKE ÖELDA |
| 2002 | K93 | Kui Te nüüd juhtuksite oma töökoha kaotama, siis kas leiaksite uue, Teile vastuvõetava töökoha | K93_a...Oma elu- või tööpiirkonnast? K93_b...Mujalt Eestis? K93_c...Venemaalt? K93_d...Mujalt välismaalt? | Kindlasti jah Tõenäoliselt jah Tõenäoliselt mitte Kindlasti mitte RÕ |
| 2000 | K112 | Kui Te nüüd juhtuksite oma töökoha kaotama, siis kas leiaksite uue, Teile vastuvõetava töökoha | K112_a...Oma elu- või tööpiirkonnast? K112_b...Mujalt Eestis? K112_c ...Venemaalt? K112_d ...Mujalt välismaalt? | Kindlasti jah Tõenäoliselt jah Tõenäoliselt mitte Kindlasti mitte RÕ |

Kokkuvõttes võib öelda, et küsimused erinevad aastati nii sisu kui täpsustuste poolest (asukoht, ajaraamistik, millist töökohta mõeldakse), mistõttu pikaajalist võrdlust nende andmete pealt teha ei saaks. Küll aga võib öelda, et küsimused alates 2015. aastast (kui kadus ära asukoha täpsustus) on omavahel rohkem võrreldavad kui varasemate aastate küsimused.

Küsimus 3. Võrdsed võimalused erinevates valdkondades

Kolmandaks välja valitud küsimuseks sai tabelis 5 toodud eestlaste ja muu rahvuse võrdsete võimaluste hindamine erinevates valdkondades nagu materiaalne heaolu, haridusvõimalused, poliitilise edu võimalikkus ning võimalus saada juhtivale ametikohale. Eraldi tasub välja tuua, et 2011. aastal antud küsimusi ei küsitud. Lisaks tasub mainida, et tegelikult oli küsimuses esitatud valdkondi eri aastatel tunduvalt rohkem, kuid siinkohal filtreeris autor välja vaid need valdkonnad, mille kohta läbivalt arvamust küsiti.

Aastal 2000 oli küsimus sõnastatud järgnevalt: „*Kas Teie olete Eestis täheldanud, et eestlastel on võrreldes mitte-eestlastega samade eelduste korral kergem saavutada järgnevaid eesmärke?*“ Samuti oli küsimuse järel täpsustatud, mida uurijad peavad silmas samade eelduste all: „*samade eelduste all on mõeldud seda et, näiteks mõlemad on kodanikud, võrdse vanuse, hariduse, keeleoskuse ja kogemusega*“. Vastusevariandid olid järgnevad: „*Jah, alati*“, „*Jah, sageli*“, „*Mõnikord*“, „*Ei ole kergem*“, „*Hoopis mitte-eestlastel on kergem*“, „*Raske öelda*“. Järgneval kahel aastal oli küsimus sõnastatud sarnaselt, kuid olid rahvuse suhtes ümber pööratud. Kui aastal 2000 küsiti, kas eestlastel on võrreldes mitte-eestlastega võrdsed võimalused, siis nüüd küsiti, kas mitte-eestlastel on võrreldes eestlastega võrdsed võimalused.

Sisuliselt on need kaks erinevat väidet. Kui esimesel aastal oli vastusevariantide seas ka võimalik vastata, et eestlaste asemel on paremad võimalused hoopis mitte-eestlastel, siis 2002 ja 2005 sarnast varianti ei olnud. Nendel aastatel oli kasutusel skaala „*Enamasti on*“, „*Sagedamini on*“, „*Sagedamini ei ole*“, „*Enamasti ei ole*“ ja „*Raske öelda*“. Antud küsimuse puhul võib ühe probleemkohana tuua välja, et respondentide jaoks, kes tahaks vastata, et „võimalused on võrdsed“, puudus sobiv variant. Kuna küsimus oli esitatud kallutatuna ehk küsitud, kas ühel rahvusgrupil on kergem eesmärke saavutada kui teisel, siis sisuliselt vastates „*enamasti ei ole*“ või „*sagedamini ei ole*“ ei selgu, kas vastaja peab neid võimalusi võrdseteks või paremateks teisest rahvusgrupist.

Tabel 5. Võtmeküsimus number 3: võrdsed võimalused erinevates valdkondades

| Aasta | Küs. Nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------|--|---|---|
| 2020 | 3.1 | Kuidas hindate eestlaste ja teisest rahvusest inimeste võimalusi, olukorda järgnevates valdkondades? | 3.1_1 Materiaalne heaolu 3.1_3 Haridus- ja enesetäiendamisevõimalused 3.1_4 Poliitilise tegevuse võimalused 3.1_7 Võimalus saada juhtivale ametikohale riigi- ja omavalitsusasutustes 3.1_8 Võimalus saada juhtivale ametikohale erasektoris | Eestlastel palju parem Eestlastel veidi parem Võrdne Teisest rahvusest veidi parem Teisest rahvusest palju parem Ei oska / soovi vastata |
| 2017 | Q76 | Kuidas hindate eestlaste ja venekeelsete inimeste võimalusi, olukorda järgnevates valdkondades? | Q76_1 Sissetulekud, materiaalne heaolu Q76_3 Haridusvõimalused lastele Q76_4 Poliitilise tegevuse võimalused Q76_7 Võimalus saada juhtivale ametikohale riigi- ja omavalitsusasutustes Q76_8 Võimalus saada juhtivale ametikohale erasektoris | Võrdne Eestlastel palju parem Eestlastel veidi parem Vene-keelsetel veidi parem Vene-keelsetel palju parem Ei oska öelda |
| 2015 | Q76 | Kuidas hindate eestlaste ja venekeelsete inimeste võimalusi, olukorda järgnevates valdkondades? | Q76_1 Sissetulekud, materiaalne heaolu Q76_3 Haridusvõimalused lastele Q76_4 Täiendkoolituse võimalused Q76_5 Poliitilise tegevuse võimalused Q76_8 Võimalus saada juhtivale ametikohale riigi- ja omavalitsusasutustes | Võrdne Eestlastel palju parem Eestlastel veidi parem Vene-keelsetel veidi parem Vene-keelsetel palju parem Ei oska öelda |
| 2011 | - | | | |
| 2008 | K44 | Kuidas hindate eestlaste ja mitte-eestlaste võimalusi ja olukorda järgnevates valdkondades? | K44_1 Sissetulekud, materiaalne heaolu K44_3 Haridusvõimalused lastele K44_4 Täiendkoolituse võimalused K44_5 Poliitilise tegevuse võimalused K44_8 Võimalus saada juhtivale ametikohale riigi- ja omavalitsusasutustes | Võrdne Eestlastel palju parem Eestlastel veidi parem Venekeelsetel veidi parem Venekeelsetel palju parem RÕ |
| 2005 | K39 | Kas Teie olete Eestis täheldanud, et mitte-eestlastel on võrreldes eestlastega samade eelduste korral võrdsed võimalused saavutada järgnevaid eesmärke? Samade eelduste all on mõeldud võrdset haridust, oskusi ja kogemusi. | K39_h Saada võrdse töö eest võrdset palka K39_f Saada kõrgemat haridust K39_e Saavutada edu poliitikas K39_c Saada juhtivale ametikohale | Enamasti on Sagedamini on Sagedamini ei ole Enamasti ei ole RASKE ÖELDA |

| | | | | |
|------|-----|---|--|---|
| 2002 | K30 | Kas Teie olete Eestis täheldanud, et mitte-eestlastel on võrreldes eestlastega samade eelduste korral võrdsed võimalused saavutada järgnevaid eesmärke? Samade eelduste all on mõeldud seda et, näiteks mõlemad on kodanikud, võrdse vanuse, hariduse, keeleoskuse ja kogemusega. | K30_g Saada sama töö eest kõrgemat palka K30_f Saada kõrgemat haridust K30_e Saavutada edu poliitikas K30_c Saada juhtivale ametikohale | Enamasti on Sagedamini on Sagedamini ei ole Enamasti ei ole RÖ |
| 2000 | K98 | Kas Teie olete Eestis täheldanud, et eestlastel on võrreldes mitte-eestlastega samade eelduste korral kergem saavutada järgnevaid eesmärke? Samade eelduste all on mõeldud seda et, näiteks mõlemad on kodanikud, võrdse vanuse, hariduse, keeleoskuse ja kogemusega. | K98_g Saada sama töö eest kõrgemat palka K98_h Kindlustada endale ja oma perele materiaalne heaolu K98_f Saada kõrgemat haridust K98_e Saada tunnustatud poliitikuks K98_c Saada juhtivale ametikohale | Jah, alati Jah, sageli Mõnikord Ei ole kergem Hoopis mitte-eestlastel on kergem RÖ |

Alates 2008. aastast muudeti oluliselt küsimuse sisu. Kui eelnevalt paluti hinnata eesmärkide saavutamise võimalikkust võrdsete eelduste olemasolul siis nüüd paluti hinnata millised on eri rahvusest inimeste võimalused ja olukord teatud valdkondades: „*Kuidas hindate eestlaste ja mitte-eestlaste võimalusi ja olukorda järgnevatel valdkondades?*“. Lisaks küsimuse sõnastusele muutusid ka vastusevariandid, andes nüüd võimaluse valida kummal rahvusgrupil on võimalused paremad/natuke paremad või peab vastaja neid võimalusi mõlemas grupis võrdseteks.

Üks aspekt, millele tuleks aga tähelepanu pöörata, on rahvusgruppide defineerimine eri aasta küsimustes. 2000-2008 on palutud hinnata eestlaste ja mitte-eestlaste võimalusi, aastatel 2015 ja 2017 eestlaste ja venekeelsete inimeste võimalusi ning aastal 2020 eestlaste ja muust rahvustest inimeste võimalusi. „Mitte-eestlased“ ning „muust rahvusest inimesed“ on oluliselt laiemad mõisted kui „venekeelsed inimesed“.

Lisaks tuleb välja tuua, et valdkondade määratlused on aastate lõikes samuti oluliselt muutunud. Tabelis 5 on alaküsimuste seas näha valdkonnad, mida autor kaardistas järgmise nelja kategooria alla: materiaalne heaolu ja sissetulekud, haridusvõimalused, poliitilise edu võimalikkus ning võimalus saada juhtivale ametikohale.

Võimalust saada juhtivale ametikohale on küsitud läbivalt kõikidel aastatel, kuid alates 2008. aastast on sisse toodud täpsustus, et mõeldakse juhtivat positsiooni riigi- ja omavalitsusüksuses. Tasakaaluks on alates 2017. aastast juurde lisandunud ka küsimus võimaluse kohta saada juhtivale ametipositsioonile erasektoris. Poliitilise edu võimalikkust on küsitud aastati samuti üsna erinevalt: „*Poliitilise tegevuse võimalused*“ (2008, 2015, 2017, 2020), kusjuures autori hinnangul jääb segaseks, mida poliitilise tegevuse all mõeldakse, „*Saavutada edu poliitikas*“ (2002, 2005) ja „*Saada tunnustatud poliitikuks*“ (2000) puhul on aga tegemist võrdlemisi subjektiivsete hinnangutega, tulenevalt mida vastaja peab eduks ning tunnustamiseks.

Haridusvaldkonna all on 2000, 2002 ja 2005 küsitud võimaluste kohta saada kõrgemat haridust, kusjuures jääb segaseks kas kõrgema hariduse all peetakse silmas näiteks kõrghariduse saamise võimalust või lihtsalt kõrgemat haridust (kellest kõrgemat? millest kõrgemat?). 2008, 2015 ja 2017 küsiti laste haridusvõimaluste kohta ning täiendkoolituste võimaluste kohta ning aastal 2020 oli kombineeritud ühe väite alla: haridus- ja enesetäiendamisvõimalused.

Materiaalse heaolu ja sissetulekute olukorra ja võimaluste kohta on samuti aastati erinevalt küsitud. 2000 ja 2002 paluti hinnata võimalust saavutada eesmärk „*saada sama töö eest kõrgemat palka*“, aastal 2005 „*saada võrdse töö eest võrdset palka*“. Aastal 2000 paluti lisaks hinnata võimalust saavutada eesmärk: „*Kindlustada endale ja oma perekonnale materiaalne heaolu*“. Alates 2008. aastast pidid vastajad hindama võimalusi ja olukorda kategoorias „*Sissetulekud ja materiaalne heaolu*“, kusjuures 2020. aastal jäeti alles vaid hinnang „*materiaalne heaolu*“ võimalustele ja olukorrale.

Seega, kuna antud küsimustes on muudetud sisuliselt nii põhiküsimust kui ka alakategooriad ning vastusevariante, siis kõigi aastate omavahelist võrdlust siin läbi viia ei saa. Küll aga on alates 2008. aastast enamus kategooriaid omavahel küllaltki hästi võrreldavad.

Küsimus 4. Suhtumine teisest rahvusest majanaabritesse

Neljandaks välja valitud võtmeküsimuseks osutus suhtumine muust rahvusest majanaabrisse, mis on toodud tabelis 6. Ka selle küsimuse puhul võib märgata, et küsimuse sõnastus on monitooringu algusaastatel oluliselt erinev hilisematest aastatest. Aastal 2000 ja 2002 küsiti vastajalt sisuliselt valmidust minna elama elamusse, kvartalisse, külasse, kus on palju eestlasi/mitte-eestlasi. Probleemaatiliseks võib siinkohal lugeda, et küsitakse mitu küsimust korraga. Vastaja jaoks võib olla vastumeelne elada samas majas teisest rahvusest inimesega, kuid samas külas elamise vastu ei pruugi tal nii tugevaid tundeid olla. Samuti jäävad natuke segaseks pakutud vastusevariandid („*Esimesel võimalusel*“, „*Rahuldavatel tingimustel*“, „*Ainult soodsatel tingimustel*“, „*Äärmise vajaduse korral*“, „*Mingil juhul ei läheks*“), näiteks ei ole aru saada, mida peetakse silmas „*soodsate tingimuste*“ all.

Alates 2005. aastast küsitakse suhtumist sellesse, kui tuleks ühes majas elada koos eestlaste/mitte-eestlastega ning alates 2015. aastast täpsustatakse küsimust ning küsitakse maja- või korterinaabriks olemise kohta. Samuti täpsustub, et ei peeta silmas seda, kui samas majas elab mõni üksik teisest rahvusest inimene, vaid küsitakse suhtumist sellesse, kui peaks naaber olema paljude teisest rahvusest inimestega. 2005. aastal sai vastaja valida kolme sisulise vastusevariandi vahel („*Mul poleks midagi selle vastu*“, „*Parem kui ei tuleks*“, „*Mul on ükskõik*“) või jätta arvamus ütlemata („*Raske öelda*“).

Tabel 6. Võtmeküsimus number 4: suhtumine teisest rahvusest majanaabritesse

| Aasta | Küs. Nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------------|---|--|--|
| 2020 | 4.7. | Kuidas Te suhtute sellesse, kui Te oleksite korteri- või majanaaber paljude ... ? | venekeelsete inimestega eestikeelsete inimestega | Väga hästi (ei häiriks) Pigem hästi Ükskõiksel Pigem halvasti Väga halvasti (häiriks väga) Ei oska / soovi vastata |
| 2017 | Q21E/ Q21M | Kuidas Teie suhtuksite sellesse, kui ... | Q21E_1 Teil tuleks olla paljude venekeelsete inimestega korteri- või majanaaber Q21M_1 Teil tuleks olla paljude eestlastega korteri- või majanaaber | Täiesti positiivselt, see ei häiriks mind sugugi Pigem positiivselt Mul on ükskõik Pigem negatiivselt Täiesti negatiivselt, see häiriks mind väga Ei oska öelda |
| 2015 | Q21E/ Q21M | Kuidas Teie suhtuksite sellesse, kui ... | Q21_E1 Teil tuleks olla paljude venekeelsete inimestega korteri- või majanaaber Q21M_1 Teil tuleks olla paljude eestlastega korteri- või majanaaber | Täiesti positiivselt, see ei häiriks mind sugugi Pigem positiivselt Mul on ükskõik Pigem negatiivselt Täiesti negatiivselt, see häiriks mind väga Ei oska öelda |
| 2011 | - | | | |
| 2008 | K94 | Kuidas teie suhtuksite sellesse, kui teil tuleks | kp94.1 elada venelastega/eestlastega ühes majas | Täiesti positiivselt, see ei häiriks mind sugugi Pigem positiivselt Mul on ükskõik Pigem negatiivselt Täiesti negatiivselt, see häiriks mind väga RÕ |
| 2005 | KP37 | Kuidas Te suhtuksite sellesse, kui Teil tuleks ... ? | KP37_b Elada venelastega/eestlastega ühes majas | Mul poleks midagi selle vastu Parem kui ei tuleks Mul on ükskõik RASKE ÖELDA |
| 2002 | KP10 | Kas Te ise põhimõtteliselt läheksite ... ? | KP10_b... elama elamusse, kvartalsisse, külasse, kus on naabriteks palju mitte-eestlasi/eestlasi | Esimesel võimalusel Rahuldavatel tingimustel Ainult soodsatel tingimustel Äärmise vajaduse korral Mingil juhul ei läheks RÕ |
| 2000 | K14 | Kas Te ise põhimõtteliselt läheksite ... ? | K14_b Elama elamusse, kvartalsisse, külasse, kus on naabriteks palju mitte-eestlasi/eestlasi | Esimesel võimalusel Rahuldavatel tingimustel Ainult soodsatel tingimustel Äärmise vajaduse korral Mingil juhul ei läheks RÕ |

Alates 2008. aastast laiendati skaala 5-astmeliseks, pakkudes kaks positiivset ning kaks negatiivset vastusevarianti (sidudes need hinnanguga, kuivõrd antud olukord vastajat häiriks/ei häiriks), ning lisaks „*Raske öelda*“ variandile lisandus keskmine hinnang „*ükskõik*“.

Antud küsimuse puhul näeb tabelist 6, et erinevatel aastatel on taaskord kasutatud erinevaid termineid eestlastele/eestikeelsele vastajale küsimuse esitamisel. Aastatel 2000 ja 2002 küsitakse mitte-eestlaste kohta, aastatel 2005 ja 2008 küsitakse venelaste kohta ning 2015-2020 venekeelsete inimeste kohta. Sellel, kuidas muukeelset elanikkonda küsimuses määratletakse, võib olla suur mõju tulemustele, sest kõik venekeelsed inimesed ei ole venelased ning omakorda mitte-eestlaste hulka ei kuulu ainult venelased ning venekeelsed inimesed.

Lisaks tasub tähelepanu pöörata, et mõnel aastal on küsimust eristatud ankeedi keele alusel ja mõnel aastal rahvuse alusel. Kõikide aastate puhul ei selgu küsimustikust, mille alusel erinev küsimus esitati, kas keele või rahvuse. 2020. aasta küsimus on eriline jällegi selle poolest, et mõlemas ankeedis olid olemas mõlemad küsimused, ning eesti ja vene keelest erineva suhtluskeelega (muu) vastajad vastasid mõlemale küsimusele. Seega peaks andmete võrdlemisel kindlasti hoolikalt jälgima, kellelt on küsimust küsitud ning kas küsimuse sihtgrupid on omavahel võrreldavad.

Küsimus 5. Eri rahvusest laste koos õppimine

Tabelis 7 on toodud analüüsi kaasatud viies võtmeküsimus eri rahvustest laste ühes koolis/lasteaias õppimise kohta. Antud küsimust ei olnud 2005. ja 2008. aasta ankeetides. Aastatel 2000 ja 2002 oli küsimus sõnastatud järgnevalt: „*Kuidas Te suhtute sellesse, et üha enam mitte-eesti lapsi pannakse eesti lasteaedadesse, algkoolidesse õppima.*“. Küsimuste vastusevariante vaadates võib näha, et need pole otseselt üksteist välistavad. Kui vastaja pooldab mitte-eesti laste panemist eesti lasteaedadesse ja algkoolidesse, siis võis ta ka iseenesest nõustuda kõigi kolme positiivse variandiga: „*Pooldate igal juhul, sest see aitab mitte-eesti lastel eesti keelt õppida*“, „*Pooldate seda juhul, kui venekeelseid lapsi ei ole rühmas/klassis liiga palju*“, „*Pooldate, kui rühmas/klassis on abiõpetaja*“. Samuti ei ole negatiivsed vastusevariandid üksteist välistavad: „*Ei poolda seda, sest nii eesti kui mitte-eesti laste õppimine ja suhtlemine lasteaias/algkoolis on tulemuslikum oma emakeeles*“, „*Ei poolda, sest see viib rahvuste ühtesulamisele, oma keele ja rahvusliku eripära kaotamisele*“.

Tabel 7. Võtmeküsimus number 5: eri rahvusest laste koos õppimine

| Aasta | Küs. Nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------|---|--------------|---|
| 2020 | 8.4. | Kuidas hindaksite üldiselt seda, kui ühes rühmas/klassis õpiksid koos erinevast rahvusest või eri emakeelt kõnelevad lapsed? | | Väga hea Pigem hea Pigem halb Väga halb Ei oska / soovi vastata |
| 2017 | Q54 | Kuivõrd Te oleksite nõus sellega, et lasteaedades ja koolides õpiksid koos ühes rühmas/klassis erinevast rahvusest või erinevat emakeelt kõnelevad lapsed/õpilased | | Oleksin nõus, kui enamus oleks mitte-eesti emakeelega Oleksin nõus, kui ligi pooled oleksid mitte-eesti emakeelega Oleksin nõus, kui umbes veerand oleksid mitte-eesti emakeelega Oleksin nõus, kui mõned üksikud oleksid mitte-eesti emakeelega Ei ole üldse nõus Ei oska öelda |
| 2015 | Q54 | Kuivõrd Te oleksite nõus sellega, et lasteaedades ja koolides õpiksid koos ühes rühmas/klassis erinevast rahvusest või erinevat emakeelt kõnelevad lapsed/õpilased? | | Oleksin nõus, kui enamus oleks mitte-eesti emakeelega Oleksin nõus, kui ligi pooled oleksid mitte-eesti emakeelega Oleksin nõus, kui mõned üksikud oleksid mitte-eesti emakeelega (alla 20%) Ei ole üldse nõus Ei oska öelda |
| 2011 | q100 | Kuivõrd Te oleksite nõus sellega, et lasteaedades ja koolides õpiksid koos ühes rühmas / klassis erinevast rahvusest või erinevat emakeelt kõnelevad lapsed / õpilased. | | Oleksin nõus, kui rühmas/klassis on enamus mitte-eesti emakeelega lapsed/õpilased Oleksin nõus, kui rühmas/klassis on ligi pooled mitte-eesti emakeelega lapsed/õpilased Oleksin nõus, kui rühmas/klassis on mõned üksikud mitte-eesti emakeelega lapsed/ õpilased Ei poolda üldse Ei oska öelda |
| 2008 | | | | |
| 2005 | | | | |
| 2002 | K70 | Kuidas Te suhtute sellesse, et üha enam mitte-eesti lapsi pannakse eesti lasteaedadesse, algkoolidesse õppima. | | Pooldate igal juhul , sest see aitab mitte-eesti lastel eesti keelt õppida Pooldate seda juhul, kui venekeelseid lapsi ei ole rühmas/klassis liiga palju ja räägitakse eesti keelt Ei poolda seda, sest nii eesti kui mitte-eesti laste õppimine ja suhtlemine lasteaia/ algkoolis on tulemuslikum oma emakeeles Ei poolda, sest see viib rahvuste ühtesulamisele, oma keele ja rahvusliku eripära kaotamisele RASKE ÖELDA |
| 2000 | K90 | Kuidas Te suhtute sellesse, et üha enam mitte-eesti lapsi pannakse eesti lasteaedadesse, algkoolidesse õppima. | | Pooldate igal juhul , sest see aitab mitte-eesti lastel eesti keelt õppida Pooldate seda juhul, kui venekeelseid lapsi ei ole rühmas/klassis liiga palju Pooldate, kui rühmas/ klassis on abiõpetaja Ei poolda seda, sest nii eesti kui mitte-eesti laste õppimine ja suhtlemine lasteaia/ algkoolis on tulemuslikum oma emakeeles Ei poolda, sest see viib rahvuste ühtesulamisele, oma keele ja rahvusliku eripära kaotamisele RÕ |

Alates 2011. aastast muudeti küsimuse sõnastust ning enam ei küsitud suhtumist teatud toimuvasse olukorda, vaid hoiakut/nõustumist tingliku/hüpoteetilise olukorraga: „*Kuivõrd Te oleksite nõus sellega, et lasteaedades ja koolides õpiksid koos ühes rühmas / klassis erinevast rahvusest või erinevat emakeelt kõnelevad lapsed / õpilased.*“. 2020. aasta ankeedis on küsimust mõnevõrra lihtsustatud ja lühendatud, kuid sellegipoolest küsitakse kahe erineva sihtrühma kohta ühes ja samas küsimuses: erinevast rahvusest laste või erinevat emakeelt kõnelevate laste kohta.

Vastusevariandid on aastate lõikes samuti veidi erinenud. Aastatel 2011, 2015 ja 2017 oli skaala tunduvalt nõustumise poole kaldu: esitati kolm varianti (2017 isegi neli erinevat varianti), mille puhul vastaja nõustub väitega kui teatud tingimused on täidetud, ja üks negatiivne variant: „*Ei ole üldse nõus*“. Neutraalne vastusevariant puudus, samas võis märkida, et ta ei oska antud küsimusele vastata. 2020. aasta küsimustikus korrigeeriti ka vastuse skaalat, mis esindas nüüd tasakaalus väärtuseid: „*Väga hea*“, „*Pigem hea*“, „*Pigem halb*“, „*Väga halb*“ ning „*Ei oska öelda/ei soovi vastata*“.

Kuna taaskord on küsimuste sõnastus ning vastusevariandid ajas oluliselt muutunud, võiks omavahel võrreldavad olla esimeste aastate küsimused (kuigi peab arvestama, et 2002. oli võrreldes 2000. aastaga üks vastusevariant vähem „*Pooldate, kui rühmas/ klassis on abiõpetaja*“ ning jaotused võivad olla sellest mõjutatud) ja 2011-2017 aasta küsimused, kui tunnuseid veidi ümber kodeerida ja kõik nõustumise kategooriad üheks siduda.

Küsimus 6. Soovitud kodakondsus

Viimaseks valitud võtmeküsimuseks on analüüsi kaasatud soovitud kodakondsuse küsimus. Antud küsimuse puhul peaks kindlasti alustama sellest, et analüüsimisel vaatama, et võrreldakse samu sihtrühmi. 2020. aasta ankeedis küsiti „*Millist kodakondsust Te endale soovite*“ kõigilt mitte-eesti kodanikelt, samas 2015 ja 2017 küsiti seda küsimust ainult määratlemata kodakondsusega inimestelt. 2011. aastal jällegi kõigilt mitte-eesti kodakondsusega inimestelt. 2008. aastal ja varasemalt küsiti kõigilt, kelle pereliikmete seas oli määratlemata kodakondsusega inimesi. Samuti olid varasemate aastate 2000-2005 ankeetides küsitud mitte ainult vastaja enda, vaid ka pereliikmete (abikaasa, elukaaslane, pojad, tütreid) kohta. 2015. aastal küsiti eraldi küsimuses, millist kodakondsust vastaja oma pereliikmele soovitaks.

Tabel 8. Võtmeküsimus number 6: soovitud kodakondsus

| Aasta | Küs. Nr | Põhiküsimus | Alaküsimused | Vastusevariandid |
|-------|---------|---|---|---|
| 2020 | 9.3 | Millist kodakondsust Te endale sooviksite? | | VÕIB MITU VASTUST Praegune kodakondsus rahuldab Eesti Venemaa Mõne teise riigi, täpsustage Ei soovi kodakondsust Ei oska / soovi vastata |
| 2017 | Q70 | Mis riigi kodakondsust Te endale tahaksite? | | ===== |
| 2015 | Q70 | Millist kodakondsust Te endale tahaksite? | | ===== |
| | Q72 | Kui mõnel Teie pereliikmel ei ole kodakondsust, siis millist kodakondsust Te oma pere liikmele soovitaksite? | | Eesti kodakondsust Venemaa kodakondsust Mõne teise riigi kodakondsust, millise? Ei soovitaks mingit kodakondsust Praegune olukord rahuldab Ei oska öelda |
| 2011 | q72 | Millist kodakondsust Te endale tahaksite? VÕIB OLLA MITU VASTUST | | Eesti Venemaa Mõne teise riigi Ei soovi kodakondsust Praegune kodakondsus rahuldab Raske öelda |
| 2008 | K68 | Kui Teil või mõnel Teie pere liikmel ei ole kodakondsust, siis millist kodakondsust Te endale või oma pere liikmele sooviksite? VÕIMALIK MÄRKIDA MITU VASTUST | | Eesti Venemaa Mõne teise riigi Ei soovi kodakondsust |
| 2005 | K60 | Kui mõnel Teie pere liikmel ei ole veel kodakondsust, siis millist kodakondsust nad endale sooviksid, alaealiste laste puhul aga palun öelge, millist kodakondsust Teie neile sooviksite? | K60_a Teie ise? K60_b Teie abikaasa/ elukaaslane? K60_c Teie poeg/ pojad? K60_d Teie tütar/ tütre? | Eesti kodakondsust Venemaa kodakondsust Mõne teise riigi kodakondsust Ei soovi üldse kodakondsust RASKE ÖELDA |
| 2002 | K22 | Kui mõnel Teie pere liikmel ei ole veel kodakondsust, siis millist kodakondsust nad endale sooviksid, alaealiste laste puhul aga palun öelge, millist kodakondsust Teie neile sooviksite? | K22_a1 Teie ise? K22_b1 Teie abikaasa/ elukaaslane? K22_c1 Teie poeg/ pojad? K22_d1 Teie tütar/ tütre? | Eesti kodakondsust Venemaa kodakondsust Mõne teise riigi kodakondsust Ei soovi üldse kodakondsust RASKE ÖELDA |
| 2000 | K26 | Kui mõnel Teie pere liikmel ei ole veel kodakondsust, siis millist kodakondsust nad endale sooviksid, alaealiste laste puhul aga palun öelge, millist kodakondsust Teie neile sooviksite? | Vastaja ise Vastaja abikaasa/ elukaaslane Vastaja poeg/ tütar Vastaja ema/ isa (ka abikaasa või elukaaslase ema/ isa) Muu | Eesti kodakondsust Venemaa kodakondsust Mõne teise riigi kodakondsust Ei soovi üldse kodakondsust RÕ |

Enamasti oli vastusevariantide seas Eesti ja Vene kodakondsus, mõne muu riigi kodakondsus ning variant, et ei soovita üldse kodakondsust. Alates 2011. aastast sai ka valida varianti „*Praegune kodakondsus rahuldab*“. 2015 ja 2017 oli tegemist avatud küsimusega, kus vastaja sai ise vastuse märkida. Antud küsimuses ei ole vastusevariantide varieerumine aastate lõikes niivõrd olulise tähtsusega, kui see on näiteks erinevaid hinnanguid mõõtvates küsimustes. Sõnastuse osas suurim erinevus tuleb sõnade „*soovin*“ (2000-2005, 2020) või „*tahaksin*“ (2011-2017) kasutusest eri aastatel, mis võib mingil määral vastust mõjutada, kuna „*tahtma*“ on oluliselt tugevam sõna kui „*soovima*“. Üldiselt leiab autor, et soovitud kodakondsuse küsimust on võimalik aastate lõikes analüüsida, arvesse peab võtma lihtsalt, et analüüsi oleks kaasatud ainult määratlemata kodakondsusega inimesed, sest sellisel juhul on eri aastate sihtgrupid võrreldavad.

Kokkuvõttes saab ankeetküsimuste üldisele analüüsile ja konkreetsete küsimuse sõnastuse analüüsile tuginedes öelda, et esineb vaid väga üksikuid küsimusi, mida on läbivalt kõigi 8 monitooringu raames küsitud ja seega on pikaajalise trendianalüüsi teostus mõnevõrra küsitav. Küll aga selgub analüüsist, et teatud lainete võrreldavus on võimalik. Milliseid aastaid omavahel võrrelda saab, sõltub aga juba konkreetsest küsimusest, mida analüüsida soovitakse. Iga küsimuse puhul tuleb uurijal teha kindlaks, et küsimuse sihtgrupp oleks sama ning et küsimuse sõnastuses ning vastusevariantides ei esineks sisulisi erinevusi.

3.3 Andmete esinduslikkuse analüüs

Andmete esinduslikkus on keerukas ning mitmetahuline mõiste. Andmete esinduslikkust mõjutab muuhulgas ka kogutud andmete kvaliteet. Üks aspekt, mille kaudu andmete kvaliteeti hinnata, on analüüsida, kui võrd ulatuslik on lünklikkus uuringuandmetes ning mil määral andmelüngad võivad tekitada esinduslikkuse kadu. Antud magistritöös vaatab autor, missugune on andmete lünklikkus eri aastate küsitlusandmetes ning analüüsib 2020. aasta andmete põhjal täpsemalt, kas andmete lünklikkust saab pidada täiesti juhuslikuks (MCAR) või juhuslikuks (MAR)/mittejuhuslikuks (NMAR).

3.3.1 Andmelünkade analüüsi üldised tulemused

Kuna elanikkonnaküsitlused viidi läbi silmast silma intervjuude käigus (vähemalt viimasel neljal aastal, varasematel kordadel see ei selgu andmetest), võib eeldada, et respondendid ei saanud

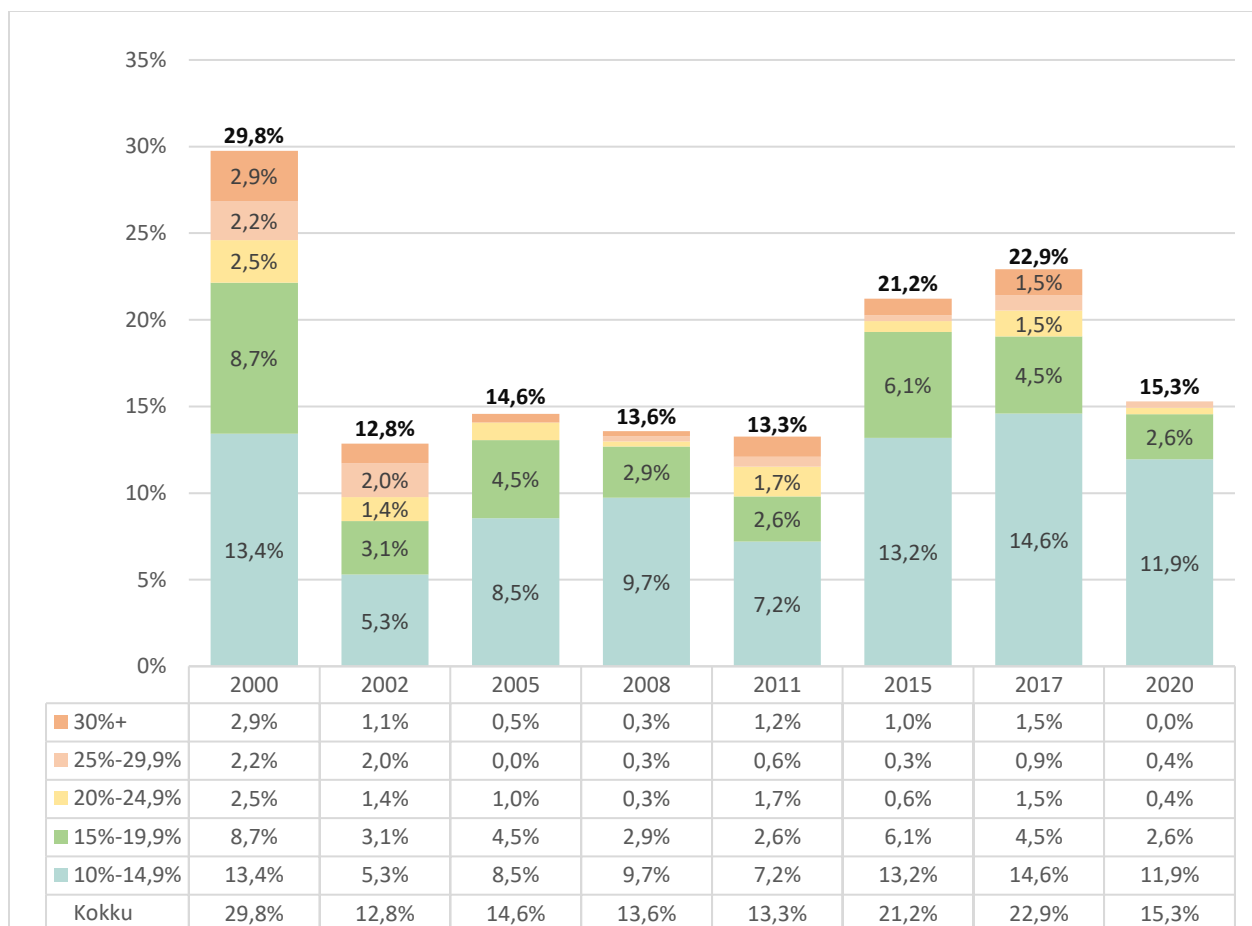
ankeedile vastates küsimusi iseseisvalt vahele jätta/tühjaks jätta. Seega on käesolevas töös puuduvate väärtuste all käsitletud vastusevariante „Ei oska öelda“, „Raske öelda“, „Ei soovi / oska vastata“. Puuduvate väärtustena ei ole käsitletud andmelünkasid, mis tulenesid sellest, et vastaja suunati küsimusest mööda ning konkreetset küsimust tema käest ei küsitud.

Selleks, et saada ülevaade, kui suur võis lünklikkuse probleem olla erinevatel monitooringu aastatel, vaatas autor, kui palju erineva lünklikkuse määraga küsimusi iga monitooringu aasta andmestikus esines. Taolise ülevaate koostamiseks lähtus autor Bennet (2010) seisukohast, et probleemid võivad hinnangute täpsuses tekkida, kui andmelünkasid on küsimusele esitatud vastuste seas rohkem kui 10%.

Lisaks tuleb siinkohal täpsustada, et eraldi küsimuseks on käesolevas analüüsis loetud iga vastust eeldav üksus ehk kui vastajal paluti hinnata sama skaala alusel viit erinevat väidet, siis on neid siinses analüüsis käsitletud kõiki eraldi küsimustena (kuigi ankeedis võidi neid lugeda ühe küsimuse alla). Puuduvate väärtuste analüüsis ei arvestanud autor andmelünkadena küsimuste variante „muu“ (kuigi nad võisid olla andmestikus eraldi tunnuseks).

Kuna monitooringu viimase nelja aasta valimid olid koostatud ebaproportsionaalselt, siis eri aasta üldandmestike võrreldavuse säilitamiseks kaalus autor andmed rahvuse alusel üldkogumiga proportsionaalseks. Parim viis algandmestike võrreldavuse tagamiseks oleks kaaluda andmestikku disainikaaludega, kuid kahjuks ei olnud monitooringu andmetesse disainikaale arvatud ning autoril endal puudus piisav taustainfo nende arvutamiseks. Seetõttu arvutas autor ise rahvuspõhised järelkihistuskaalud (eestlased vs muu rahvus) kasutades selleks Statistikaameti andmebaasist kättesaadavaid rahvastikuandmeid (tabelid RV0221U, RV0221, RV021).

Joonisel 2 on toodud monitooringu aastate lõikes vähemalt 10% andmelünkadega küsimuste osakaal kogu küsimuste arvust. Jooniselt on samuti võimalik näha, milline oli väga suure (enam kui 20% või 30%) andmelünkadega küsimuste osakaal vastava aasta küsimuste seas. Osakaalude arvutamise aluseks oli küsimuste arv vastava aasta küsimustikus, mis varieerus olenevalt aastast vahemikus 199 küsimust - 447 küsimust. Kõige enam küsimusi oli aastal 2000 ning kõige vähem aastal 2005, ülejäänud aastatel jäi küsimuste arv ligikaudu 310-350 vahele, välja arvatud aastal 2020 kui küsimusi oli ligikaudu 268.



Joonis 2. Vähemalt 10% puuduvate väärtustega küsimuste osakaal kogu küsimuste arvust monitooringu aastate lõikes

Küsimuste arv peegeldab maksimaalset võimalikku küsimuste arvu (kui kõik respondendid oleks vastanud kõikidele küsimustele) vastaval aastal. Realsuses erines küsimuste arv vastajatel olenevalt aastast näiteks rahvuse, keele, kodakondsuse või mõne muu eelduse tõttu (kas vastajal oli lapsi vms). Seega võib toodud osakaaludesse suhtuda teatud ettevaatlikkusega. Suurem küsimuste arv võib mõjutada andmelünkade esinemise sagedust andmestikus, kuna mida pikem on küsimustik, seda suurem on ka tõenäosus, et vastaja väsib ja annab pealiskaudsemaid vastuseid (sh näiteks „Ei oska öelda“).

Jooniselt 2 on näha, et andmelünkade probleem võiks kõige enam mõjutada 2000. aasta monitooringu tulemusi ja hinnanguid. Kokku ligi 30%-l küsimustest oli vähemalt 10% vastajatest märkinud, et nad ei oska küsimusele vastata. See omakorda võib mõjutada oluliselt nendes küsimustes antud hinnangute esinduslikkust juhul, kui lünkade esinemine pole täiesti juhuslik.

Võrreldes ülejäänud lainetega esineb andmelünki mõnevõrra enam ka 2015. ja 2017. aasta andmetes, kus veidi enam kui viiendikus küsimustest on puuduvaid väärtuseid enam kui 10%. Ülejäänud monitooringu aastate puhul esineb kõrge lünklikkuse probleemi ligikaudu 13%-15%-l küsimustest.

Nagu ka ankeetküsimuste samasuse analüüsist selgus, siis 2015. ja 2017. aastal olid monitooringu ankeetküsimused võrdlemisi sarnased (võrreldes ülejäänud lainetega). Ka mõlema aasta suurte andmelünkadega (>10%) küsimusi uurides selgub, et need küsimused on kahel aastal suures osas kattuvad. Kui mõningal määral esineb ulatusliku lünklikkusega küsimuste seas keeruliselt sõnastatud küsimusi, siis ei paista sõnastus olevat peamine põhjus. Samas võib ühe keeruliselt esitatud küsimuse näitena tuua 2017. aastal laulupeo traditsiooni puudutanud küsimuse, mis võis vastajate jaoks osutada segaseks, kuna küsimuse sissejuhatus on pikk ja väited kohati liiga ekstreemselt sõnastatud: *„Laulupeol on Eesti kultuuris väga oluline roll. See traditsioon on püsinud eri aegadel ja tingimustes. Samas on tänapäeva Eesti elanikkond mitmekultuuriline ning venekeelsed elanikud moodustavad kolmandiku siinsest elanikkonnast. Palun öelge, kumba väidet laulupeo traditsiooni jätkamisel Te toetate, kas väidet A või B?*

- Väide A: *Laulupeol võiksid kõlada ka venekeelsed laulud, sest venekeelne elanikkond moodustab kolmandiku Eesti elanikkonnast*
- Väide B: *Laulupidu ei ole koht lõimumiseks ja teiste rahvaste kaasamiseks, sest selle ülim väärtus seisneb selles, et ta ühendab omavahel eelkõige eestlasi“*

Kuid nagu eelpool mainitud, ei paista suurte andmelünkadega (>10%) küsimused peamiselt silma keerulise sõnastuse poolest, vaid pigem selguvad analüüsist teatud teemad, millele vastamisel leidis 2015. ja 2017. aastal rohkem andmelünki kui teistes teemades. Järgnevalt on loetletud nimetatud teemad/küsimused ja sulgudes toodud vastaval aastal ulatusliku lünklikkusega (>10%) küsimuste/väidete arv:

- Põhjused, miks suurel hulgal Eestis elavatel mitte-eestlastel ei ole Eesti kodakondsust (2017: 8, 2015: 8);
- Eestikeelsele haridusele üleminekuga seotud väited/küsimused (2017:13, 2015:6);
- Suhtumine sisserändesse (2017: 6, 2015: 4);
- Usaldusväärsus institutsioonide vastu (2017: 7, 2015: 3);

- Eesti ja muust rahvusest laste koos õppimine (2017: 4, 2015: 4);
- Eestlaste ja venekeelsete inimeste võrdsed võimalused erinevates valdkondades (2017:3, 2015: 5);
- Venekeelsete noorte haridusvõimalused võrreldes eestikeelsete noortega (2017:5, 2015:5).

Lisaks eelnevalt mainitud teemadele esines palju andmelünki ka sissetulekuid puudutavates küsimustes, mis on aga küsitlusuuringutes võrdlemisi tavapärane.

Mõningaid sarnasid küsimusi/teemasid leidis ka aastal 2000 suurte andmelünkadega (>10%) küsimuste seas, kuid mitte palju. Kattusid mõned haridusalased või sissetulekuid puudutavad küsimused. Samas tuleb arvestada, et aastal 2000 oli kokku 133 küsimust, kus andmete lünklikkust võib pidada ulatuslikuks (>10%). Seega nende mõne kattuva küsimuse põhjal ei saa öelda, et andmelüngad esinesid võrreldes 2015. ja 2017. aastaga sisuliselt sarnastes küsimustes. Juba ankeetküsimuste samasuse analüüsi peatükis selgus, et monitooringu esimeste aastate küsimustikud on sisult laiemad ja haaravad rohkem teemasid kui hilisemad aastad. Seda oli ka 2000. aasta ning 2015. ja 2017. aasta küsimusi kõrvutades näha, et aastal 2000 esines suurte andmelünkadega (>10%) küsimusi teemades ja küsimustes, mida hilisemates monitooringutes nii detailselt käsitletud ei ole. Siia alla kuuluvad näiteks:

- Valitsemiskorda hindavad küsimused;
- Eestlaste, eestivenelaste ning Venemaa venelaste omavaheline võrdlus erinevate isikuomaduste alusel nagu näiteks puhtus, lohakus, usaldusväärsus, hoolivus jne;
- Erinevad välispoliitikat käsitlevad küsimused (mida on esimeste aastate monitooringutes oluliselt detailsemalt käsitletud kui hiljem), ennekõike Venemaa mõju kohta;
- Kuivõrd kaitsevad mitte-eestlaste huve erinevad institutsioonid nagu valitsus, riigikogu, kohalik omavalitsus ja president.

Samas esines aastal 2000 suurte andmelünkadega (>10%) küsimusi ka hilisematel aastatel käsitletud ning lõimumismonitooringu seisukohast üsna olulistes teemades, nagu kodakondsuspoliitikaga seotud erinevad küsimused ning mitte-eestlaste kaasamine Eesti ühiskonda.

Kokkuvõttes võib lünklikkuse analüüsi põhjal öelda, et andmelünkade esinemine küsimustes paistab rohkem tingitud olevat küsimuste sisust/teemast kui segasest või keerulisest sõnastusest. Arvestada tuleks ka sellega, et küsitlusuuringu valimis on suur osa eestlasi ning paljud monitooringus kajastatud (ja ka käesolevas lünklikkuse analüüsis esile tõusnud) teemad käsitlevad mitte-eestlaste lõimumist või eri rahvuste võrdseid võimalusi, mille osas teatud osal populatsioonist võibki arvamus või teadmised puududa. Väljatoodud teemade seas oli ka mitmeid tundliku sisuga küsimusi/teemasid, mille puhul andmelünkade probleem võibki sagedamini esile tulla. Tundlike küsimuste puhul võib olla võimalik küsimuse sõnastuse abil andmelünki teatud määral vältida, kuid ilma küsimuse aluseks olevat konstrukti ja küsimuse sisu muutmata ei pruugi andmelünkade täielik vältimine realistlik olla. 2000. aastal, kui andmete lünklikkuse probleem oli kõige suurem, jäi silma, et lünklikkust esines palju küsimustes, mis olid erandlikud just monitooringu esimesele aastale. Samas esines suurte andmelünkadega (>10%) küsimusi ka vaadeldud aastate puhul teemades, mida on monitooringus üsnagi läbivalt käsitletud. Just nende küsimuste puhul tekib kahtlus, kas vaadeldud aastad, kui lünklikkuse probleem oli suurem (2000, 2015 ja 2017), erinevad teistest küsitlusmetoodika põhimõtete poolest või võis nendel aastatel tegemist olla madalama andmekogumise kvaliteediga.

3.3.2 Andmelünkade juhuslikkuse analüüs

Järgnevalt uuris autor hii-ruut-testi ja Fisheri täpse testi abil, kas lõimumismonitooringu erinevates tunnustes saab lünklikkust pidada täiesti juhuslikuks või mitte. Kuna kõikide monitooringu aastate analüüs oleks käesoleva magistritöö kontekstis läinud liiga mahukaks, otsustas autor analüüsi kaasata monitooringu viimase aasta kui ühelt poolt kõige relevantsema ning samas mitte kõige suurema lünklikkuse probleemiga aasta. Selleks analüüsiti kõiki 2020. aasta tunnuseid, kus lünklikkus oli >10%, ehk kokku 41 tunnust 268st. Lisas 3 on toodud kõigi 41 küsimuse sõnastused koos vastajate arvuga (N) kaalumata andmestikus. 41-st küsimusest 28-le vastasid kõik respondendid. Nelja küsimuse puhul (K3X5_1- K3X5_4) oli vastajaid 920, nende küsimuste puhul paluti arvamus vaid inimestelt, kes olid eelnevalt märkinud, et ei ole rahul mõne teenuse kättesaadavusega oma linnas/vallas. Kahe küsimuse puhul oli vastajaid 1053 ehk vastasid kõik mitte-eestlased (küsimused K6X3_4, K6X3_5). Seitsme küsimuse puhul (K9X5_1, K9X5_3, K9X5_5, K9X5_6, K9X5_7, K9X5_8, K9X5_9) oli vastajaid vaid 205, kuna küsiti millised põhjused võivad mõjutada soovi mitte-omandada Eesti kodakondsust ning see küsimus esitati vaid

vastajatele, kellel ei olnud Eesti kodakondsust ning kes ka ei soovinud omandada Eesti kodakondsust. Kõik 41 tunnust kodeeris autor enne analüüsi ümber binaarseteks tunnusteks, kus 1 tähistab vastust, mida on käsitletud andmelüngana („Ei oska / soovi vastata“ variant) ja 0 kõiki ülejäänud vastusevariante.

Analüüsis hindas autor juhuslikkust kokku viie tunnuse suhtes, sealhulgas nii sotsiaaldemograafiliste tunnuste (sugu, rahvus, regioon) kui kahe lõimumist hindava tunnuse suhtes. Valitud sai nii sotsiaaldemograafilisi kui hinnanguid mõõtvaid tunnuseid, kuna sotsiaaldemograafilised tunnused on olnud ka järelkihistuskaalude arvutamise aluseks ning võib eeldada, et kaalumiseiga parandatakse nende tunnuste esinduslikkust üldkogumi suhtes; teiste monitooringu tunnuste suhtes, mis ei ole kaalumise aluseks olevate tunnustega seotud, ei pruugi see nii olla.

Tabel 9. Taustatunnuste jaotused

| Tunnus | | Kaalutud rahvusega | | Kaalutud üldkaaluga | |
|---|--|--------------------|------------|---------------------|------------|
| | | N | % | N | % |
| Sugu | Mees | 654 | 46,2 | 654 | 46,2 |
| | Naine | 760 | 53,8 | 760 | 53,8 |
| | Kokku | 1414 | 100 | 1414 | 100 |
| Rahvus | Eestlane | 953 | 67,4 | 957 | 67,7 |
| | Muu rahvus | 461 | 32,6 | 457 | 32,3 |
| | Kokku | 1414 | 100 | 1414 | 100 |
| Regioon | Põhja- Eesti | 617 | 43,6 | 624 | 44,1 |
| | Lääne-Eesti | 150 | 10,6 | 162 | 11,5 |
| | Kesk-Eesti | 130 | 9,2 | 134 | 9,5 |
| | Kirde-Eesti | 154 | 10,9 | 153 | 10,8 |
| | Lõuna-Eesti | 363 | 25,7 | 341 | 24,1 |
| | Kokku | 1414 | 100 | 1414 | 100 |
| K6X8 Kuivõrd Te näete oma tulevikku seotuna Eestiga - kas Te sooviksite Eestist lahkuda, püsivalt mujale elama asuda? | Kindlasti ei soovi | 840 | 59,4 | 830 | 58,7 |
| | Pigem ei soovi (võib-olla kunagi) | 305 | 21,5 | 305 | 21,6 |
| | Sooviksite lahkuda ajutiselt, aga mitte püsivalt | 190 | 13,5 | 196 | 13,9 |
| | Sooviksite lahkuda püsivalt | 42 | 3,0 | 44 | 3,1 |
| | Muu | 3 | 0,2 | 3 | 0,2 |
| | Ei oska / soovi vastata | 34 | 2,4 | 36 | 2,5 |
| | Kokku | 1414 | 100 | 1414 | 100 |
| K6X5_4 Tunnete end osana Eesti ühiskonnast | Täiesti nõus | 683 | 48,3 | 677 | 47,9 |
| | Pigem nõus | 566 | 40,0 | 567 | 40,1 |
| | Pigem ei ole nõus | 108 | 7,6 | 109 | 7,7 |
| | Üldse ei ole nõus | 22 | 1,6 | 23 | 1,6 |
| | Ei oska / soovi vastata | 35 | 2,5 | 38 | 2,7 |
| | Kokku | 1414 | 100 | 1414 | 100 |

Tabelis 9 on toodud valitud taustatunnuste jaotused nii autori poolt arvatud rahvuspõhise järelkiihistuskaalu kui uuringufirma poolt arvatud üldkaaluga (samuti järelkiihistuskaal) läbikaalutud andmestikes. Taustatunnuste jaotusi vaadates võib märgata, et sooline jaotus on ühesugune nii rahvuse kui üldkaaluga kaalutud andmestikes. Tunnuste „Rahvus“ ja K6X5_4 jaotused on samuti võrdlemisi sarnased mõlema kaaluga kaalutud andmestikes. Väikseid erinevusi võib aga märgata ülejäänud kahe tunnuse jaotustes. Kuigi taustatunnuste jaotused rahvuse ning üldkaaluga kaalutud andmestikes oluliselt ei erine (või erinevused on väga väikesed), ei saa seda väita ülejäänud 41 analüüsi valitud tunnuse lünklikkuse jaotuste kohta ning seetõttu viidi edasine analüüs läbi siiski kahe erineva kaaluga.

Puuduvaid väärtusi sotsiaaldemograafilistes tunnustes ei esine, kuid nii tunnuse K6X8 kui ka K6X5_4 puhul on puuduvaid väärtuseid (ehk kategooria „Ei oska / soovi vastata“ vastuseid) 2,4%-2,7% olenevalt kaalumisest ja tunnusest. Kuna analüüsi eesmärk oli testida, kas 41-le tunnusele vastamata jätmine on seotud valitud taustatunnuse valiidsete väärtustega, siis jäeti analüüsis välja tunnuste K6X8 (tuleviku seotus Eestiga) ning K6X5_4 (tunneb end osana Eesti ühiskonnast) puuduvad väärtused ehk kategooria „Ei oska / soovi vastata“ vastused. Lisaks on samadel põhjustel välja jäetud ka küsimuse K6X8 kategooria „Muu“, kuna seda oli valinud kokku vaid 3 vastajat (0,2% kõigist vastajatest).

Hii-ruut-testid tehti nii rahvuse kui üldkaaluga kaalutud andmestikes. 410st testist 56 puhul selgus, et hii-ruut-testi eeldused ei ole täidetud (teoreetilise jaotuse lahtreid, kus $n_{ij} \geq 5$, oli vähem kui 80% või mõnes lahtris oli $n_{ij} < 1$) ja teostati Fisheri täpne test. Lisades 4-8 on toodud hii-ruut-testide ja Fisheri täpsete testide detailsemad tulemused.

Ühe näitena tulemuste tõlgendamisest on järgnevalt esitatud tunnusepaari Sugu ja K6X7 („Kuivõrd edukas on Teie arvates seni üldiselt olnud teisest rahvusest inimeste lõimumine ja kaasamine ehk integratsioon Eesti ühiskonda?“) hii-ruut-testi tulemused (nii rahvusega kui üldkaaluga kaalutud andmestikes). Tabelist 10 on näha, et nii rahvusega kui üldkaaluga kaalutud andmestikes on meeste ja naiste jaotus ühetaoline (vastavalt 46% ja 54%). Kui aga vaadata soolist jaotust integratsiooni hinnangut puudutava küsimuse (K6X7) mittevastanute seas märkame, et naiste osakaal võrreldes meestega on antud küsimusele mittevastanute hulgas oluliselt suurem (rahvusega kaalutud andmestikus naisi 67% ja mehi 33% ning vastavalt 68% ja 32% üldkaaluga kaalutud andmestikus).

Ka hii-ruut-testi tulemus kinnitab, et esineb statistiliselt oluline erinevus (olulisuse tõenäosus alla 0,001) soolises jaotuses nende seas, kes integratsiooni edukusele hinnangu andsid ja kes ei andnud (st esinevad andmelüngad): hii-ruut-statistiku väärtus 15,3 (rahvusega kaalutud) ning 16,8 (üldkaaluga kaalutud) vabadusastmete arvu 1 korral (lisa 4). Võib näha, et andmelünkadega vastajate hulgas on naisi ligikaudu 2/3 ja seega on valiidsete vastuste seas naiste osakaal mõnevõrra madalam (52%) kui kogu vastuste seas (54%). Seetõttu võib öelda, et integratsiooni edukust hindavas tunnuses ei ole andmelüngad täiesti juhuslikud ning järelikult ei saa valiidseid vastuseid automaatselt pidada soos suhtes esinduslikuks. Kuna tulemus näitab, et lünklikkust ei saa pidada täiesti juhuslikuks (MCAR), võib edasisel võimalikul analüüsil, mis põhineb ainult valiidsetel väärtustel, esineda teatav viga.

Tabel 10. Integratsiooni hinnangu valiidsete vastuste ja andmelünkade sooline jaotus

| | | | Kuivõrd edukas on Teie arvates seni üldiselt olnud teisest rahvusest inimeste lõimumine ja kaasamine ehk integratsioon Eesti ühiskonda? | | | | | |
|-------|-------|---|---|-----------------|-------|---------------------|-----------------|-------|
| | | | Kaalutud rahvusega | | | Kaalutud üldkaaluga | | |
| | | | Andmelünk | Valiidne vastus | Kokku | Andmelünk | Valiidne vastus | Kokku |
| Sugu | Mees | N | 60 | 594 | 654 | 62 | 591 | 653 |
| | | % | 33% | 48% | 46% | 32% | 48% | 46% |
| | Naine | N | 123 | 637 | 760 | 129 | 631 | 760 |
| | | % | 67% | 52% | 54% | 68% | 52% | 54% |
| Kokku | | N | 183 | 1231 | 1414 | 191 | 1222 | 1413 |
| | | % | 100% | 100% | 100% | 100% | 100% | 100% |

Kuna analüüsis läbi viidud teste on palju (kokku 410), tuleb arvestada võimalusega, et mõned statistiliselt olulised tulemused võivad esineda juhuse tõttu. Käesolevas töös on võetud otsuste tegemise kriteeriumiks $p < 0,05$, mis tähendab, et ilma mitmese testimise eripärasid arvestamata võivad kuni 5% kõigi testide tulemustest olla statistiliselt olulised juhuse tõttu ning populatsioonis võib nende puhul tegelikult kehtida nullhüpotees (seost vaadeldavate tunnuste vahel ei ole). See tähendab, et hii-ruut-testi tulemusel kinnitatakse monitooringus esinenud tunnuse lünklikkuse seos valitud taustatunnusega, mis tähendaks, et andmelüngad on juhuslikud (MAR) või mittejuhuslikud (NMAR), samas kui seose esinemine võib tuleneda lihtsalt valimi juhuslikkusest ning alust MAR või NMAR lünklikkuse väitmiseks tegelikult ei ole. Kuni 21 juhul (olulisuse nivool 0,05: $410 \cdot 0,05 = 21$) võib seega tegemist olla juhuslikkusest tingitud valepositiivse tulemusega. Seda nimetatakse ka I tüüpi veaks, kui testi tulemusel lükatakse ümber nullhüpotees samas, kui populatsioonis

tegelikult nullhüpotees kehtib (James jt 2021: 559). II liiki veaks loetakse vastupidiselt nullhüpoteesi juurde jäämist juhul, kui tegelikult populatsioonis kehtib alternatiivhüpotees.

Ideaalolukorda, kus nii I kui II tüüpi vea tegemise määr on madal, on väga raske saavutada. Pigem tuleb valida, kumba tüüpi vea tegemise ohtu minimeerida. Praktikas peetakse hüpoteeside testimisel I tüüpi viga enamasti tõsisemaks probleemiks kui II liiki viga, kuna see tähendab seose esinemise väitmist olukorras, kus see ei ole tõene (James jt 2021). Lahendusena, sellise mitmese testimise puhul, kasutatakse sageli olulisuse tõenäosuste või olulisuse nivoo korrigeerimist. Laialt levinud on kasutada näiteks Bonferroni või Holmi paranduse meetodit, keskendudes I liiki vea võimaluse minimeerimisele. Mainitud meetodite puhul kaasneb aga sageli suur II liiki vea tegemise tõenäosus eriti kui testitavaid seoseid on väga palju nagu antud töös. Käesolevas magistris on kasutatud Benjamini ja Hochbergi (1995) parandust, mis garanteerib, et valeavastuste määr (*False Discovery Rate*) ei ületa etteantavat valeavastuste määra q (ehk tagasilükatud nullhüpoteesidest keskmiselt mitte rohkem kui q (näiteks 5%) on valepositiivsed) ja on seetõttu autori arvates parim valik olulisuse nivoo korrigeerimiseks.

Autor arvutas välja korrigeeritud olulisuse nivood järgnevate valeavastuste määrade korral:

- $q \leq 5\%$ $\rightarrow p < 0,01198;$
- $q \leq 1\%$ $\rightarrow p < 0,00183;$
- $q \leq 0,1\%$ $\rightarrow p < 0,00014.$

Tabelis 11 on esitatud tausta- ja kaalutunnuste lõikes testide arvud vastavalt sellele, kas pärast Benjamini-Hochbergi meetodil teostatud olulisuse nivoo parandust hii-ruut-statistiku olulisuse tõenäosus oli $p < 0,05$ või $p \geq 0,05$. Tabel 11 annab seega kokkuvõtliku ülevaate, kui paljude läbiviidud testide puhul esines statistiliselt oluline seos vaadeldud tunnusepaari vahel (korrigeeritud olulisuse nivool 0,05).

Tabel 11. Testide arv ja osakaal taustatunnuste ning kasutatud kaalutunnuse lõikes

| Tunnus | Olulisuse tõenäosus | Kaalutud rahvusega | | Kaalutud üldkaaluga | |
|------------------------------------|---------------------|--------------------|-----|---------------------|-----|
| | | N | % | N | % |
| Sugu | $p_{BH} < 0,05$ | 6 | 15% | 6 | 15% |
| | $p_{BH} \geq 0,05$ | 35 | 85% | 35 | 85% |
| Rahvus | $p_{BH} < 0,05$ | 17 | 41% | 16 | 39% |
| | $p_{BH} \geq 0,05$ | 24 | 59% | 25 | 61% |
| Regioon | $p_{BH} < 0,05$ | 15 | 37% | 18 | 44% |
| | $p_{BH} \geq 0,05$ | 26 | 63% | 23 | 56% |
| Tulevik seotud Eestiga | $p_{BH} < 0,05$ | 7 | 17% | 7 | 17% |
| | $p_{BH} \geq 0,05$ | 34 | 83% | 34 | 83% |
| Tunneb end osana Eesti ühiskonnast | $p_{BH} < 0,05$ | 6 | 15% | 7 | 17% |
| | $p_{BH} \geq 0,05$ | 35 | 85% | 34 | 83% |

$p_{BH} \leq 0,05$ vastab $p < 0,01198$

Esmalt võib tabelist 11 märgata, et läbiviidud testid ei näita märkimisväärseid erinevusi tulenevalt sellest, kas testid on läbiviidud rahvusega või üldkaaluga kaalutud andmestikes. Ainult regiooni puhul võib näha, et statistiliselt olulist seost ($p_{BH} < 0,05$) kinnitavaid teste on mõnevõrra rohkem, kui andmestikku kaaluda üldkaaluga (18 testi 41-st versus 15 testi 41-st ainult rahvusega kaalutud andmestikus). Üldiselt võib siiski öelda, et mittevastamisest tingitud esinduslikkuse kadu on nii rahvuse kui üldkaaluga kaalutud andmestikus sarnane ehk seda ei kompenseeri andmestiku kaalumise uuringufirma poolt arvatud üldkaaludega. Tulemus on mõningal määral ootuspärane, kuna seda võis märgata juba tabelis 9 toodud taustatunnuste jaotuste põhjal, et andmestiku kaalumise üldkaaluga ei mõjuta oluliselt valitud sotsiaaldemograafiliste tunnuste jaotuseid (mis olid aluseks ka kaalutunnuse arvutamisel).

Samas ei ole uuringufirma poolt arvatud kaalud otseselt mõeldudki osalisest mittevastamisest tingitud esinduslikkuse kao vähendamiseks, vaid täielikust mittevastamisest tuleneva esinduslikkuse kao vähendamiseks. Praktikas sageli eeldatakse, et kaalumise on „imerelv“, mis lahendab andmete esinduslikkusega seotud probleemid, kuid ka käesoleva analüüsi tulemusel on näha, et kui tunnuses esineb märkimisväärset hulgal andmelünki, siis kaalumise nendest lünkadest tingitud esinduslikkuse kadu oluliselt ei vähenda.

Tabelist 11 võib näha ka seda, et andmelünkade esinemine on süstemaatilisem valitud 5 taustatunnuse seast rahvuse ja regiooni puhul. Statistiliselt olulist seost Benjamini-Hochbergi meetodil korrigeeritud olulisuse nivool 0,05 kinnitavad rahvuse suhtes testides ligikaudu kaks viiendikku teostatud hii-ruut-testidest: rahvusega kaalutud andmestikus 17 testi 41st ning üldkaaluga kaalutud andmestikus 16 testi 41st. Nagu ka eelnevalt juba mainitud, siis regiooni puhul näitasid statistiliselt olulist seost rahvusega kaalutud andmestikus 15 testi 41st ja üldkaaluga kaalutud andmestikus 18 testi 41st olulisuse nivool 0,05. Soo ning valitud kahe integratsiooni mõõtva tunnuse suhtes kinnitasid statistiliselt olulist seost (korrigeeritud olulisuse tõenäosusega alla 0,05) 6-7 testi 41st iga taustatunnuse puhul. Kuigi regiooni ja rahvuse puhul paistab andmete lünklikkus olevat süstemaatilisem, ei saa lünklikkust ka ülejäänud kolme taustatunnuse suhtes pidada täiesti juhuslikuks.

Järgnevat on tabelis 12 esitatud kõigi 410 testi korrigeerimata olulisuse tõenäosused; tärnidega on näidatud statistiline olulisus korrigeerimata olulisuse tõenäosuste alusel, värviga on näidatud statistiline olulisus pärast Benjamini-Hochbergi parandust. Kokku osutusid pärast Benjamini-Hochbergi parandust statistiliselt oluliseks 105 testi tulemused (olulisuse tõenäosusega alla 0,05), sealhulgas 76 testi tulemused olulisuse nivool 0,01 ning 57 testi tulemused olulisuse nivool 0,001. Tabelist 12 tuleb ka visuaalselt esile, kuidas lünklikkus on süsteemsem rahvuse ja regiooni puhul, kuid samas võib märgata teatavat mustrit ka analüüsidest seoste väljendumist küsimuste kaupa.

Näiteks selgub tabelist 12, et küsimuste K3X5_1-K3X5_4 puhul, kus vastajad hindasid põhjuseid, miks nad ei olnud rahul mõne linna/valla pakutud teenuse kättesaadavusega, on võrreldes teiste tunnustega lünklikkuse süstemaatilisuse probleem suurem, lünklikkus on seotud rohkemate taustatunnustega. Mainitud küsimuste puhul võib märgata statistiliselt olulist seost andmelünkade esinemise ja 3-4 taustatunnuse vahel (rahvus, regioon, kuivõrd vastaja tunneb end osana Eesti ühiskonnast ja K3X5_3 puhul ka sugu). Küsimustes K3X5_1-K3X5_4 toodud põhjused peegeldasid enamasti kas keelebarjääriga seotud probleemi või teenindajate suhtumist vastajasse.

Tabel 12. Hii-ruut-statistiku olulisuse tõenäosused

| Küsimus | Sugu | | Rahvus | | Regioon | | Tulevik seotud Eestiga | | Tunneb end osana Eesti ühiskonnast | |
|----------|--------------------|---------------------|--------------------|---------------------|--------------------|---------------------|------------------------|---------------------|------------------------------------|---------------------|
| | Kaalutud rahvusega | Kaalutud üldkaaluga | Kaalutud rahvusega | Kaalutud üldkaaluga | Kaalutud rahvusega | Kaalutud üldkaaluga | Kaalutud rahvusega | Kaalutud üldkaaluga | Kaalutud rahvusega | Kaalutud üldkaaluga |
| K3X1_4 | 0,024* | 0,029* | 0,004** | 0,003** | 0,000*** | 0,000*** | 0,018* | 0,018* | 0,491 | 0,615 |
| K3X1_5 | 0,475 | 0,400 | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,513 | 0,372 | 0,173 | 0,157 |
| K3X1_7 | 0,046* | 0,093 | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,997 | 0,981 | 0,080 | 0,067 |
| K3X1_9 | 0,002** | 0,001* | 0,103 | 0,102 | 0,015* | 0,01* | 0,847 | 0,806 | 0,750 | 0,794 |
| K3X1_10 | 0,278 | 0,338 | 0,003** | 0,002** | 0,219 | 0,065 | 0,362 | 0,524 | 0,001** | 0,002** |
| K3X1_11 | 0,021* | 0,017* | 0,381 | 0,280 | 0,000*** | 0,000*** | 0,771 | 0,833 | 0,489 | 0,630 |
| K3X3 | 0,103 | 0,124 | 0,401 | 0,648 | 0,026* | 0,006** | 0,479 | 0,427 | 0,964 | 0,960 |
| K3X4_4 | 0,818 | 0,769 | 0,348 | 0,313 | 0,234 | 0,140 | 0,001** | 0,001** | 0,088 | 0,086 |
| K3X4_14 | 0,052 | 0,105 | 0,794 | 0,967 | 0,102 | 0,060 | 0,003** | 0,001** | 0,840 | 0,975 |
| K3X4_15 | 0,279 | 0,322 | 0,01* | 0,012* | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,283 | 0,306 |
| K3X5_1 | 0,037* | 0,016* | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,024* | 0,042* | 0,000*** | 0,000*** |
| K3X5_2 | 0,033* | 0,032* | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,127 | 0,130 | 0,000*** | 0,000*** |
| K3X5_3 | 0,002** | 0,003** | 0,001** | 0,001** | 0,000*** | 0,000*** | 0,264 | 0,199 | 0,000*** | 0,000*** |
| K3X5_4 | 0,036* | 0,054 | 0,003** | 0,002** | 0,000*** | 0,000*** | 0,387 | 0,322 | 0,000*** | 0,000*** |
| K4X7_3 | 0,006** | 0,008** | 0,280 | 0,289 | 0,002** | 0,001** | 0,007** | 0,007** | 0,613 | 0,536 |
| K4X8_3 | 0,003** | 0,002** | 0,894 | 0,798 | 0,005** | 0,005** | 0,000*** | 0,000*** | 0,410 | 0,305 |
| K4X9_3 | 0,166 | 0,137 | 0,059 | 0,061 | 0,333 | 0,320 | 0,044* | 0,071 | 0,480 | 0,204 |
| K5_4 | 0,623 | 0,950 | 0,846 | 0,877 | 0,000*** | 0,000*** | 0,011* | 0,008** | 0,409 | 0,206 |
| K6X1_4 | 0,978 | 0,926 | 0,025* | 0,02* | 0,319 | 0,364 | 0,917 | 0,822 | 0,264 | 0,251 |
| K6X3_4 | 0,785 | 0,620 | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,127 | 0,061 | 0,045* | 0,085 |
| K6X3_5 | 0,875 | 0,776 | 0,001** | 0,001** | 0,044* | 0,067 | 0,672 | 0,550 | 0,875 | 0,872 |
| K6X6_2 | 0,631 | 0,475 | 0,078 | 0,073 | 0,279 | 0,233 | 0,161 | 0,060 | 0,272 | 0,379 |
| K6X7 | 0,000*** | 0,000*** | 0,294 | 0,300 | 0,149 | 0,101 | 0,042* | 0,027* | 0,205 | 0,273 |
| K8X5_1 | 0,859 | 0,840 | 0,082 | 0,110 | 0,940 | 0,766 | 0,376 | 0,552 | 0,281 | 0,223 |
| K8X5_3 | 0,600 | 0,593 | 0,000*** | 0,000*** | 0,03* | 0,021* | 0,099 | 0,077 | 0,016* | 0,008** |
| K8X5_4 | 0,424 | 0,415 | 0,951 | 0,881 | 0,389 | 0,293 | 0,840 | 0,876 | 0,255 | 0,151 |
| K8X5_5 | 0,730 | 0,601 | 0,229 | 0,321 | 0,378 | 0,315 | 0,845 | 0,891 | 0,082 | 0,051 |
| K8X8_7 | 0,833 | 0,614 | 0,002** | 0,003** | 0,107 | 0,103 | 0,733 | 0,668 | 0,609 | 0,593 |
| K8X8_8 | 0,513 | 0,440 | 0,400 | 0,547 | 0,613 | 0,662 | 0,282 | 0,210 | 0,106 | 0,140 |
| K8X8_10 | 0,739 | 0,540 | 0,004** | 0,007** | 0,769 | 0,488 | 0,948 | 0,863 | 0,473 | 0,513 |
| K9X5_1 | 0,347 | 0,332 | 0,035* | 0,035* | 0,192 | 0,193 | 0,672 | 0,513 | 0,959 | 0,812 |
| K9X5_3 | 0,428 | 0,523 | 0,071 | 0,074 | 0,674 | 0,745 | 0,956 | 0,918 | 0,911 | 0,818 |
| K9X5_5 | 0,537 | 0,476 | 0,041* | 0,042* | 0,210 | 0,197 | 0,943 | 0,946 | 0,865 | 0,916 |
| K9X5_6 | 0,901 | 0,831 | 0,029* | 0,029* | 0,080 | 0,082 | 0,737 | 0,707 | 0,744 | 0,735 |
| K9X5_7 | 0,483 | 0,613 | 0,028* | 0,029* | 0,444 | 0,457 | 0,833 | 0,833 | 0,801 | 1,000 |
| K9X5_8 | 0,325 | 0,429 | 0,024* | 0,024* | 0,180 | 0,178 | 0,811 | 0,769 | 1,000 | 1,000 |
| K9X5_9 | 0,543 | 0,597 | 0,112 | 0,116 | 0,205 | 0,474 | 0,742 | 0,730 | 0,783 | 0,765 |
| K9X7_13 | 0,049* | 0,043* | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,299 | 0,151 | 0,072 | 0,062 |
| K9X10_5 | 0,069 | 0,075 | 0,000*** | 0,000*** | 0,085 | 0,066 | 0,032* | 0,055 | 0,901 | 0,996 |
| K9X10_10 | 0,005** | 0,009** | 0,000*** | 0,000*** | 0,000*** | 0,000*** | 0,334 | 0,266 | 0,000*** | 0,000*** |
| T18 | 0,604 | 0,447 | 0,854 | 0,634 | 0,014* | 0,006** | 0,000*** | 0,000*** | 0,275 | 0,298 |

Kursiivis märgitud Fisher'i täpse testi tulemus

*** $p < 0,001$;

** $p < 0,01$;

* $p < 0,05$;

$p_{BH} < 0,001$ (vastab $p \leq 0,00014$)

$p_{BH} < 0,01$ (vastab $p \leq 0,00183$)

$p_{BH} < 0,05$ (vastab $p \leq 0,01198$)

Lisaks selgub tulemustest, et andmete lünklikkus küsimustes, mis mõõdavad suhtumist uussisserändajatega koos ühes majas või korteris elamisse või töötamisse (K4X7_3 ja K4X8_3), on seotud soo, regiooni ja sellega, kuivõrd seotud Eestiga plaanib vastaja tulevikus olla. Nelja taustatunnusega on statistiliselt oluline seos hii-ruut-testide tulemusel (olulisuse nivool 0,05) veel ka tunnusega K9X10_10 ehk lünklikkuse esinemisega küsimuses, kuivõrd vastaja usaldab kodanikuühiskonda Eestis. Nende küsimuste puhul ei saa samuti seetõttu andmete lünklikkust pidada täiesti juhuslikuks.

Tabelis 12 esitatud tulemusi koondades võib tuua välja mõned teemad 2020. aasta monitooringu ulatusliku lünklikkusega küsimuste seast, kus lünklikkust võib pidada süstemaatiliseks:

- rahvuskultuuri ja kultuurilise identiteedi säilimine (K6X3_4, K6X3_5, K8X5_3, K8X8_10);
- eestlaste ja teisest rahvusest inimeste võimalused erinevates valdkondades (K3X1_4, K3X1_5, K3X1_7, K3X1_9, K3X1_10, K3X1_11);
- teenuste kättesaadavusega seotud rahulolematuse põhjused (K3X4_4, K3X4_14, K3X4_15, K3X5_1, K3X5_2, K3X5_3, K3X5_4);
- erakondlikud vaated ja usaldus institutsioonide vastu (K9X7_13, K9X10_5, K9X10_10).

Kokkuvõttes selgus 410 hii-ruut-testi tulemusel, et statistiliselt olulist seost kinnitas ligikaudu neljandik testidest (105), kusjuures seosed esinevad rohkem kahe taustatunnuse (rahvus, regioon) ning teatud tüüpi küsimuste/küsimuste gruppide suhtes (eriti kui vaadata statistiliselt olulisi tulemusi korrigeeritud olulisuse nivool 0,001). 14 küsimuse puhul ei selgunud statistiliselt olulist seost (olulisuse nivool 0,05) ühegi valitud taustatunnuse ja andmelünkade esinemise vahel, mistõttu võime järeldada, et nende küsimuste puhul saab pidada andmete lünklikkust valitud taustatunnuste suhtes täiesti juhuslikuks. 27 küsimuse puhul kinnitas analüüs statistiliselt olulist seost mõne valitud taustatunnusega ($p_{BH} < 0,05$). Kokku oli 2020. aasta monitooringus 268 küsimust, seetõttu võib lünklikkusest tingitud esinduslikkuse kadu olla andmete analüüsimisel problemaatiline vähemalt ligikaudu 10%-l küsimustest. Samas võib lünkade süstemaatilisust (MAR või NMAR tüüpi lünklikkust) esineda ka tunduvalt rohkem kui 10% küsimuste puhul, kuna lünklikkust testiti vaid 5 taustatunnuse suhtes ning ei saa välistada lünklikkuse süstemaatilisust ka tunnustes, mida analüüsi ei kaasatud (kus esines puuduvaid väärtuseid vähem kui 10%-l vastajatest). Mil määral lünkade süstemaatilisus praktikas andmeid analüüsidest tulemusi mõjutab,

oleneb mitmetest teguritest, näiteks sellest, kui võrd tugev on seos konkreetses tunnuses lünkade esinemise ja teiste tunnuste vahel, lisaks milliseid teisi tunnuseid analüüsi kaasatakse ning kas tegu on MAR või MCAR tüüpi lünklikkusega.

Nagu eelnevas peatükis selgus, oli lünklikkus kõige ulatuslikum aastatel 2000, 2015 ja 2017. Kuigi käesolevas peatükis analüüsiti andmelünkade juhuslikkust vaid 2020. aasta andmete pealt, julgub autor teha mõningaid üldistusi ka teiste monitooringu lainete kohta. Eelmises peatükis sai kolme kõige ulatuslikuma lünklikkusega aasta puhul välja toodud ühised teemad (suurte andmelünkadega (>10%) küsimuste seast). Võrreldes neid teemasid/küsimusi 2020. aasta küsimustega, kus lünklikkus oli >10% (41 küsimust välja toodud lisas 3), on näha, et paljud küsimused/teemad kattuvad (rohkem 2015. ja 2017. aastaga, vähem 2000. aastaga), seal hulgas näiteks eri rahvusest inimeste võrdsed võimalused eri valdkondades, usaldus institutsioonide vastu ja eesti kodakondsuse omandamisega seotud küsimused.

Kuna teemad, milles andmete lünklikkuse probleem eri aastatel on suurem, paljuski kattuvad, võiks ka eeldada, et vähemalt aastatel, kus lünklikkuse probleem oli suurem kui 2020. aastal (2000, 2015 ja 2017), korduvad ka sarnased mustrid nagu 2020. aasta andmelünkade juhuslikkuse analüüsist selgus. Täpsemate järelduste tegemiseks tuleks muidugi analüüs läbi viia kõigis aastates eraldi, kuid kokkuvõttes võiks 2020. aasta põhjal järeldada, et andmete lünklikkust ei saa pidada täiesti juhuslikuks (lünklikkus on kas juhuslik (MAR) või mittejuhuslik (NMAR)) ning selgub, et ka andmete kaalumise ei aita lünklikkusest tingitud esinduslikkuse kadu kompenseerida.

Järelikult peaks monitooringu andmete analüüsimisel andmelünkadesse suhtuma ettevaatlikkusega ning arvestama, et andmed ei pruugi ka kaalumise järgselt olla sihtpopulatsiooni suhtes esinduslikud. Samas tuleks ka meeles pidada, et antud järeldused kehtivad eestlaste ja muu rahvuse andmestike koos analüüsimisele (ehk üldandmestiku tulemuste analüüsimisel). Praktikas analüüsitakse monitooringu andmeid sageli rahvuse lõikes eraldi. Sealjuures ilmselt regiooni mõju lünklikkuse analüüsimisel siiski väljendub, kuid teiste tunnuste suhtes tuleks analüüsi eraldi andmestikes korrata. Lisaks on uuringufirmal eraldi arvutatud järelkiihistuskaalud eestlaste ja muu rahvuse andmestike jaoks, mille mõju lünklikkusele võiks samuti mõlema andmestiku puhul eraldi hinnata.

4. ARUTELU

Tuginedes esimeses peatükis esitatud teoreetilisele kirjandusele ning käesoleva magistritöö käigus läbi viidud analüüsile, arutleb autor järgnevalt olulisemate selgunud tulemuste üle. Diskussioonis lähtub autor töö eesmärgist ja arutleb, mil määral on lõimumismonitooringu meetodika, ankeetküsimused ning esinduslikkus ajas muutunud ning kuidas see võiks mõjutada erinevate monitooringu lainete omavahelist võrreldavust. Peatüki lõpus on toodud tööst lähtuvad autoripoolsed üldised soovitused, millele võiksid uurijad tähelepanu pöörata, kui soovivad erinevate küsitlusuuringute andmete alusel võrdlevat analüüsi läbi viia.

Monitooringu teostajad ja uuringu dokumenteerimine

Analüüsist selgus, et aastate jooksul on lõimumismonitooringut läbi viinud mitmed erinevad uurimisrühmad. Samuti on kaasatud erinevaid uuringufirmasid, kelle tööks tavapäraselt on valimi moodustamine ja andmekogumise korraldamine. Kuna uuringu teostajad on aastati vahetunud, siis võivad teatud muudatused monitooringus olla tingitud just uurimisrühmade erinevast lähenemisest. Seda enam on oluline monitooringu meetodika dokumenteerimine.

Kontentanalüüsist ilmnes, et meetodika põhjalikum kirjeldus (eraldi peatükina) on olnud osa uuringuraportist vaid viimasel neljal aastal (2011, 2015, 2017 ja 2020) ja seda ilmselt tellijapoolse nõude tõttu. Aastal 2008 puudus aruandes täielikult info meetodika kohta ning aastatel 2000, 2002 ja 2005 oli kirjeldatud meetodikat väga nipsisõnaliselt vaid raporti sissejuhatuses, mida sisuliselt võib samuti lugeda dokumentatsiooni puudumiseks. Seega on näha, et varasematel monitooringu aastatel ei ole dokumenteerimine olnud prioriteet või on alahinnatud selle tähtsust ja vajadust edaspidisteks aastateks. Ka hilisematel aastatel on meetodika kirjelduse põhjalikkus varieerunud, kuid sellest tuleb juttu lähemalt järgmistes arutelu alapunktides. Kui teha järeldusi monitooringu eri aastate andmete võrreldavuse kohta puhtalt olemasoleva meetodika dokumentatsiooni põhjal, tuleks esimesed neli aastat kõrvale jätta ja edasi vaadata vaid 2011.-2020. aasta monitooringuid, kuna tagantjärele on andmete kasutajal dokumentatsiooni uuesti luua/ taastada pea võimatu.

Samas ei loeta dokumentatsiooni alla vaid metoodika kirjeldust. Mohler jt (2008) selgitavad, et dokumentatsiooni alla kuuluvad muuhulgas ka küsimustikud ning märgenditega andmefailid, mis lõimumismonitooringu puhul kenasti olemas on, ja ka nende põhjal on võimalik teha teatavaid järeldusi eri lainete omavahelise võrreldavuse kohta. Samuti on autor töö kirjutamise ja monitooringuga seotud materjalide kogumise käigus mõistnud, et kuna Eesti uuringumaastik on piisavalt väike, liigub palju infot seoses uuringu metoodika ja korraldusega ilmselt ka suusõnaliselt erinevate läbiviijate ja uurijate seas. Sellegipoolest arvab autor, et lõimumismonitooringud võiks olla paremini dokumenteeritud, kuna see aitaks uuringu andmete kasutajatel teha andmete analüüsil (ja eriti just võrdleva analüüsi puhul) informeeritud ning adekvaatseid metodoloogilisi valikuid.

Üldkogum ja analüüsitavad sihtrühmad

Integratsiooni monitooringu elanikkonna küsitluse üldkogum on aastatel 2000-2011 defineeritud kui Eesti alalised elanikud vanuses 15-74 aastat ja aastatel 2015-2020 kui Eesti alalised elanikud vanuses 15+ aastat. Metoodikaaruannetest ja hankedokumentidest ei selgu täpne vajadus, miks tellija taolise muutuse kasuks otsustas. Samas on muutus olnud vaid vanuselises kaasamises ja seetõttu on võrdlemisi lihtne eri aastate andmete analüüsimisel filtreerida välja üle 74-aastased vastajad selleks, et tagada andmete võrreldavus.

Monitooringu aruandeid ja andmeid uurides selgus, et tavapäraselt viiakse monitooringu analüüsi läbi võrdlevalt rahvuse lõikes (eestlased vs muu rahvus). See jäi autorile silma, kuna sellisel juhul peaks olema võimalik ka võrreldavaid sihtrühmi läbi aastate sarnastel alustel defineerida. Analüüsist selgus aga, et kuigi rahvust on küsitud kõigil aastatel, siis 2008. aastal sai respondent valida end mitme rahvuse esindajaks ning andmetest ei selgu, millist rahvust peeti peamiseks rahvuseks. Seetõttu on integratsiooni monitooringu aruandes rahvus määratud 2008. aastal ankeedi keele alusel. Eestlasi ja muust rahvusest elanikke ei saa aga läbivalt kõigi aastate puhul ankeedi keele alusel eristada, kuna osadel aastatel said eestlased vastata vaid eestikeelsele ankeedile (2005, 2015, 2017) ning osadel jällegi valida, kas soovivad vastata eesti või vene keeles (2000, 2002, 2008, 2011, 2020). Ka sellised muutused võiks tegelikult olla dokumenteeritud, kuna uurijatel, kes pole uuringu korralduse ja läbiviimisega seotud olnud, võivad taolised nüansid kahe silma vahele jääda.

Seega, teostades eri lainete andmeid kaasavat võrdlevat analüüsi, tasub esmalt veenduda, et valitud sihtgrupid ja tunnused oleks erinevatel aastatel sarnaselt defineeritud või et andmete järeltöötlemisega oleks võrreldavus võimalik saavutada.

Valim, valikuraam ja valimi esinduslikkus

Metoodika infot analüüsidest selgus, et ei ole väga lihtne kindlaks teha, kuidas täpselt ja mis meetodil on valimid eri aastatel koostatud. Aastate 2000-2005 kohta oli aruannetes infot minimaalselt, kuid kokkuvõttes võis järeldada, et koostatud on mitmeastmeline juhuvalim. Samas, kuna info oli esimeste aastate kohta niivõrd puudulik, siis valimite juhuslikkust kindlalt väita ei saa. 2008. aasta valimi moodustamise kohta aruandes info puudus. 2011.-2020. aasta monitooringute puhul oli näha, et erinevad uuringufirmad kasutavad valimi koostamisel küll mõnevõrra erinevaid lähenemisi ja ka dokumenteerivad seda erineva detailsusega, kuid sarnaselt 2000.-2005. aasta monitooringutele võib öelda, et valimid on koostatud mitmeastmelise valikuna. Nagu ka Lagerspetz (2017) välja toob, on mitmeastmelise juhuvaliku eelduseks ennekõike see, et valik oleks juhuslik kõikidel valikutasanditel. Ainult sellisel juhul saaks nimetatud valimitüüpi pidada tõenäosuslikuks valimiks.

Analüüsi tulemustest lähtuvalt valiku juhuslikkust täielikult kinnitada aga ei saa. Esimestel aastatel (2000-2008) puudub selle kohta piisav info ning hilisematel aastatel (2011-2020) on kasutatud noore mehe-noore naise ja lähteadressi reeglit ning 2015. ja 2017. aastal lisaks vanemaealiste vastajate puhul kvoote, mis kõik võivad omada mõju valimi juhuslikkusele. Lähteadressi reegel võimaldab põhimõtteliselt teha juhuvalikut ilma indiviidi tasandi valikuraami omamata. See lihtsustab omakorda uuringu korraldust, kuid eeldab samal ajal rangete valikureeglite jälgimist, küsitlejate metoodilist ja selget juhendamist ning täpset ning ausat tööd küsitlejatelt. Neid eelduseid ei ole aga lihtne täita. Kasutatud lähenemised on praktikas võrdlemisi laialt levinud just seetõttu, et nende abil püütakse tagada valimi paremat soolist ning vanuselist esinduslikkust.

Vehovar, Toepoel ja Steinmetz (2016) toovad välja, et mittetõenäosusliku valimi puhul ei ole võimalik hinnata valimiviga, kuna valimisse sattunud elementide kaasamistõenäosused ei ole teada. Valimiveaks peetaksegi tõenäosusliku valiku juhuslikkusest tingitud viga (Neumann 2006). Kuigi magistr töö autor seab kahtluse alla monitooringu valimite juhuslikkuse (ja seega tõenäosuslikkuse), taotlevad uuringu läbiviijad siiski koostatud valimite juhuslikkust. Ainult

aastatel 2011-2017 oli monitooringu aruannetes välja toodud maksimaalne veapiir - ligikaudu 3%, mis aga eeldab, et valim oleks koostatud tõenäosuslikult ning valimi juhuslikkus säiliks ka lisaks mittevastamist arvesse võttes. Seega tuleks ka raportites esitatud veapiiridesse suhtuda teatava ettevaatlikkusega.

Selleks, et esinduslikkust adekvaatselt hinnata peaks dokumentatsioonis olema täpselt määratletud üldkogumi jaotus nende niinimetatud oluliste tunnuste lõikes, mille suhtes esindavust väidetakse, koos samade tunnuste jaotustega vastajate seas. Monitooringu aruannetes olid jaotused vastajate seas välja toodud kolmel aastal (2011, 2015 ja 2017), kuid jaotused üldkogumis ainult 2011. aasta aruandes. Ülejäänud aastatel üldkogumi jaotusi esitatud ei olnud, kuid metoodikakirjelduste põhjal (2015, 2017 ja 2020) võiks eeldada, et esinduslikkust taotleti rahvuse, regiooni, vanuse ning soo suhtes. Seda, kas esindavus nende tunnuste suhtes ka saavutati, pole samas neis metoodikaaruannetes käsitletud.

Lisaks tuleks arvestada, et monitooringu aruannete andmeid analüüsitakse tavapäraselt rahvuse lõikes. Selleks, et oleks võimalik andmeid analüüsida ka muust rahvusest elanikkonna alamgruppides, on alates 2011. aastast muudetud eestlaste ja mitte-eestlaste osakaale valimis, mitte-eestlaste osakaalu teadlikult suurendades, muutes selle populatsiooni üldjaotuse suhtes ebaproportsionaalseks. Valimimahud on aastati kõikunud 1000st indiviidist 1505 indiviidini, mis on kooskõlas ka Eestis levinud tavapraktikaga, et valimimaht võiks olla vähemalt 1000 vastajat. Aastatel 2000-2005 olid valimimahud vahemikus 1000-1142 vastajat. 2011. aastal on valimimahtu oluliselt suurendatud (1505) ning edaspidi püsinud valimimahud umbes 1400 indiviidi juures (va 2015, kui valimis oli 1214 indiviidi). Ilmselt on valimimaht alates 2011. aastast tavapraktikast (1000) märgatavalt suurem seetõttu, kuna sooviti, et ka muust rahvusest vastajate kohta oleks võimalik teha statistiliselt vähemalt sama täpseid järeldusi (st veapiir oleks eestlaste ja muust rahvusest inimeste puhul sarnane) või isegi täpsemaid, sest paljudes küsimustes võib olulisem olla just mitte-eestlaste arvamus.

Kokkuvõttes võib öelda, et valimi koostamise põhimõtted on aastati olnud küllaltki erinevad. Valimite juhuslikkust ja seega ka esinduslikkust võib küll seada kahtluse alla, kuid nagu autor ka analüüsi osas mainis, võib see olla tingitud dokumentatsiooni vähesest detailsusest. Nii palju kui aruannetest välja loeb, taotletakse üldiselt rahvuslikku, vanuselist, soolist ning regionaalset esinduslikkust ning piisavalt suured valimimahud ning asjaolu, et valimimetoodikas on

põhijoontes järgitud valiku juhuslikkuse põhimõtteid, annavad siiski piisavalt alust uskumiseks, et vähemalt monitooringu nelja viimase küsitluse andmed võiks olla omavahel valimi poolest võrreldavad. Esimese nelja küsitluse kohta täpsemate järelduste tegemiseks peaks olema tollase meetoodika kohta rohkem infot ja tuleks vastavate küsitluste andmeid lähemalt uurida.

Andmekogumine ja järelkontroll

Analüüsist selgub, et andmekogumist on alates 2011. aastast korraldatud silmast silma intervjuu vormis. Varasemate aastate kohta aruannetest andmekogumisviisi kohta informatsiooni ei leia. Neuman (2006) rõhutab intervjuueeriya rolli olulisusele mõõtmisvea tekkimisel. De Leeuw (2008) toob omalt poolt välja, et intervjuueeriya mõju küsitlusuuringute puhul püütakse sageli hinnata (ja ka kontrollida) järelkontrolli läbiviimisega. Ka monitooringu meetoodika analüüsis selgus, et (vähemalt) 2011. aastast alates on olnud tellijapoolne nõue, et järelkontrolli tuleb küsitluse järgselt teostada ning aruannetes oli kirjas, et seda ka tehti. 2011., 2015. ja 2017. aasta puhul selgus, et eksimusi küsitlustöös ei tuvastatud. 2020. aasta aruandes oli välja toodud, et reeglite vastu eksinud küsitlajate töö tulemus eemaldati, kuid puudus info selle kohta, mis olid konkreetsed probleemid või kui suurt osa valimist see mõjutas.

Andmete kaalumine

Kaalude arvutamise kohta oli monitooringu aruannetes infot toodud vaid kolmel aastal: 2011, 2015 ja 2017. Samas võis monitooringu andmefailidest näha, et kaalutunnused on arvutatud kõigi aastate puhul, sealjuures alates 2011. aastast, kui andmefailid on rahvuse lõikes lahku löödud, on ka igas andmestikus oma kaalutunnus. Rässleri jt (2008) sõnul arvutatakse kaalud tavapäraselt seetõttu, et vähendada täielikust mittevastamisest tingitud esinduslikkuse kadu andmetes. Analüüsist selgus, et osades meetoodikaaruannetes esitatud jaotused võimaldavad küll hinnata, kui võrd kaalumise sotsiaaldemograafiliste tunnuste jaotust muutis, pole jaotuste vastavust üldkogumile seal enamasti siiski ära toodud. Aruannetes mainitakse soo, vanuse, rahvuse ja regiooni arvestamist kaalumisel, kuid ei selgu täpsed tunnused ja kategooriad, mille alusel kaale arvutatud on. Kirjeldusest võib vaid järeldada, et arvutatud kaalude puhul on tegemist järelkihistuskaaludega.

Ühelt poolt ei saa autor väita, et andmeid kaaludes on monitooringu eri lained omavahel võrreldavad, kuna puudub täpne info (eriti varasemate aastate puhul), milliste tunnuste suhtes andmeid esinduslikuks kaalutakse. Teisalt ei saa autor väita ka vastupidist. Kaalumise puhul tuleb arvestada ka sellega, et isegi kui järelkihistamiskaalude abil saavutatakse loetud tunnuste suhtes esinduslikkus, ei taga see tingimata esinduslikkust muude tunnuste suhtes. Praktikas reeglina küll eeldatakse, et tagab ja kui andmekogumine ja selle meetodika on kvaliteetne, siis on sel eeldusel ilmselt ka tõepõhi all. Tegelikuses sõltub kaalumise tulemusena esinduslikkuse paranemine eelkõige sellest, kas täieliku mittevastamise mehhanism on mõõdetud tunnuste suhtes täiesti juhuslik (MCAR), juhuslik (MAR) või mittejuhuslik (NMAR). Kokkuvõttes on valikute tegemine andmete kaalumisel (millise meetodiga kaale arvutada, milliseid tunnuseid milliste kategooriatega kaalumise aluseks võtta) teatud määral subjektiivne, erinevad valikud võivad (olenevalt ka küsitlusandmete kvaliteedist) viia erinevate tulemusteni. Põhjalikum info valikuraami ning kaalude arvutamise kohta annaks siinkohal kasutajale võimaluse andmeid teadlikumalt analüüsida ning hinnata, ega kaalude arvutamisel ei ole tehtud järeltöötlusviga.

Ankeetküsimustike samasus ajas

Ankeetküsimuste analüüsimisel lähtus autor ennekõike sellest, kui võimalikuks saaks pidada pikaajaliste trendianalüüside teostamist integratsiooni monitooringu andmete põhjal, mida Kaldur korduvalt on välja toonud (2019, 2021), ja mille võimalikkust ja vajadust ka laiemalt nähakse. Seega keskendus autor magistritöös sellele, kuivõrd kõigi 20 aasta jooksul teostatud kaheksa monitooringu andmed omavahel võrreldavad on. Analüüsist selgus, et küsimustikud on aastate jooksul väga palju muutunud ja kui kõrvale jätta nn taustatunnused, siis ei leidu ühtegi küsimust, mida kõigil 8 aastal oleks täpselt samas sõnastuses ning samade vastusevariantidega küsitud. Küll aga selgus küsimustike analüüsis, et monitooringu läbivad teemad on aastate lõikes jäänud enam-vähem samaks: tööturg, haridus, kontaktid ja keeleoskus, meedia, ühiskondlik osalus, võrdsed võimalused ning üldised lõimumisalased hoiakud.

Ka küsimustike pikkused on aastate lõikes palju varieerunud. Meetodikat analüüsinud Balti Uuringute Instituudi uurijad tõid oma aruandes välja, et ankeeti analüüsinud ekspertide sõnul on lõimumismonitooringu ankeet liiga pikk (Kaldur jt 2019). Aastate jooksul on märgata küsimuste hulga vähendamist ankeetides, mis aga võib olla üheks põhjuseks, miks küsimused on niivõrd

palju muutunud. Algas-aastate küsimustikud olid pikemad ja teemad rohkem hajusamad, lisaks leidis palju eripalgelisi küsimusi. Mida aasta edasi, seda konkreetsemaks ja rohkem põhiteemade ümber paistsid küsimustikud koonduvat. Erinevad monitooringu läbiviijad on ka avalike arutelude käigus välja toonud, et monitooringu küsimused on aastate jooksul muutunud, kuna ka integratsiooni mõtestamine on aastate jooksul muutunud. Samuti on teatavad mõju omanud ka toimunud võtmesündmused nagu 2007. aasta Pronksiöö, Ukraina kriis 2014-2015 ja koroonapandeemia.

Kuigi hankedokumentide analüüsist selgub, et vähemalt alates 2011. aastast (mil dokumendid olid autorile kättesaadavad) on tellijapoolne nõue olnud, et teatud küsimused ankeetides jääksid samaks. Hankedokumentides mainitakse, et ankeedi küsimused on jagatud kolme osasse: põhiplokk, roteeruv plokk ning lisaplokk, kuid ankeete analüüsides ei olnud võimalik kindlaks teha, millised küsimused kuhu plokki kuulusid. Põhiploki küsimused peaksid olema aastate lõikes muutumatud ning hõlmama endas vähemalt arengukava mõõdikute aluseks olevaid küsimusi. See aga omakorda tähendab, et küsimused võivad olla muutunud sõltuvalt sellest, kui arengukava periood on vahetunud ning paika pandud uued mõõdikud. Lisaks selgus analüüsist, et ankeetküsimustiku muudatused paistsid olevat seotud ka uurimisrühmade vahetumisega.

Kokkuvõttes võib välja tuua, et küsimuste poolest on omavahel sarnasemad 2000-2005 aasta ankeedid (sama arengukava periood ning ka samad uuringu läbiviijad) ning väga sarnased ka 2015. ja 2017. aasta ankeedid. 2020. aasta ankeedi analüüsides jäi autorile silma, et muudatused võrreldes kahe varasema aastaga on ilmselt suuresti tingitud Balti Uuringute Instituudi 2019. aastal läbiviidud rekontseptualiseerimise aruandes toodud soovitustest. 2008. ja 2011. aasta ankeedid paistsid pigem silma oma eripalgelisuse poolest. Täpselt ei oska autor hinnata, kas see võib olla tingitud varasemate ja hilisemate lainetega võrreldes erinevast arengukava perioodist (Lõimumiskava 2008-2013), erinevatest uuringu teostajatest, mainitud võtmesündmustest või hoopis millestki muust.

Nagu eelnevalt mainitud, lähtus autor analüüsist ennekõike sellest, kui võrd võrreldavad on kõigi monitooringu aastate ankeetküsimused omavahel ehk kui palju on tegelikult 20 aasta jooksul läbivalt esitatud küsimusi. Kuigi tulemustest ilmneb, et ankeete on ajas palju muudetud, leidub ikkagi ka küsimusi, mis on muutumatul kujul küsimustikes olnud juba päris pikalt, kuid lihtsalt mitte 20 aastat, vaid võib-olla viimased 10 aastat. Käsitletavat teemat on siiski jäänud aastate

jooksul enamjaolt samaks, kuid ka nende kohta oli rekontseptualiseerimise aruandes antud soovitusi, mida võiks jätta monitooringust välja ning mida võiks hoopiski põhjalikumalt käsitleda (Kaldur jt 2019). Kuna esitatud muudatussoovituste nimekiri oli pikk (sh metoodika osas), siis on autori jaoks jäänud ebaselgeks, kas 2020. aasta monitooringusse viidi juba kõik muudatused sisse (millega tellija ja uurimismeeskond nõustusid) või on ka tulevates monitooringutes oodata lisanduvaid sisulisi ning metoodilisi muudatusi.

Ankeediküsimuste võrreldavus

Autor analüüsis magistritöös ka lähemalt konkreetsete ankeediküsimuste muutust ajas. Kui üldiselt võib järeldada, et monitooringus käsitletavat teemat on aastate lõikes püsinud samad, siis võib eeldada, et ka (vähemalt osad) uuritavad konstruktid on ajas olnud suhteliselt sarnased. Arutelud selle üle, mida täpsemalt (konstrukti defineerimine) ning kuidas mõõta, on kindlasti toimunud tellija ning vastava aasta uurimisrühma vahel monitooringu protsessi käigus, kuid autori jaoks seda infot kättesaadaval ei olnud. Seega on autoril raske analüüsida ka konstrukti valiidsust ehk seda, kui võrd täpselt mõõtmisinstrumendid mõõdavad uuritavat nähtust. Babbie (2010) toob välja, et luues mõõtmisinstrumenti, luuakse ka teoreetilised eeldused, kuidas mõõdetav tunnus seostub teiste tunnustega. Monitooringu andmete põhjal taoliste seoste ja seeläbi valiidsuse hindamine ei kuulunud aga selle magistritöö skoopi.

Küll aga lähtus autor ankeetküsimuste kaardistamisel ja eri aastate küsimuste kõrvutamisel sellest, mis võis olla mõõdetavaks konstruktik. Nagu ankeetküsimuste sõnastuse analüüsist selgus, võis mõningate küsimuste puhul sõnastused ja vastusevariandid aastate jooksul palju muutuda, kuid on aru saada, et konstrukti, mida mõõta tahetakse, on jäänud samaks. Seega üheks viisiks eri aastate andmete võrdlemiseks on küsimusi ümber kodeerida üldisemale tasemele nii, et mõõdetav konstrukti jääks samaks. Näiteks peatükis 3.2.2 esitatud tabelis 3 võib näha, kuidas läbi aastate on muutunud küsimus, millega autori hinnangul soovitakse teada saada, kui võrd vastaja näeb enda tuleviku seotust Eestiga. Selle küsimuse puhul on nii vastusevariandid kui ka küsimuse sõnastus aastate lõikes muutunud, kuid samas oleks võimalik kõik need küsimused taandada skaalale, kas vastaja pigem plaanib oma tulevikku Eestiga siduda või pigem mitte. Sellisel puhul andmeid ümber kodeerides oleks võimalik erinevate aastate tulemusi siiski võrrelda, kuigi küsimused ise on palju muutunud.

Taolist lähenemist võib nimetada ka *ex post* väljundi harmoneerimiseks, nagu seda kirjeldavad Köner ja Meyer (2005). Wolf jt (2016) arutlevad, et parimaks viisiks andmete võrreldavuse tagamiseks on sisendi harmoneerimine. Hankedokumentide analüüsist selgus, et mingil määral on tellija omalt poolt püüdnud seda ka teha, pannes hankesse kirja erinevaid nõudeid monitooringule, et tagada võrreldavus varasemate aastatega, kuid metoodikat ja küsimustikke analüüsides selgus, et oodatavat mõju sellel siiski olnud ei ole. Näiteks on kõigis vaadeldud hankedokumentides kirjas, et ankeedid peaksid põhinema varasema aasta monitooringul ning sisaldama ka varasemalt küsitud küsimusi, kuid analüüs näitas siiski ankeetide suurt varieeruvust ajas.

Üldiselt oli ankeetküsimuste võrreldavuse analüüsi peamiseks järeltuleks, et igale küsimusele tuleks läheneda individuaalselt ning mingit ühtset võlvitsa, millised küsimused on omavahel võrreldavad ja millised mitte, ei ole. Kuigi tunnuste ümberkodeerimine lõppandmestikus võib olla üks lahendus või lähenemisviis, selgus siiski ankeetküsimuste analüüsist mitmeid olulisi aspekte, millele tuleks enne andmete analüüsimist tähelepanu pöörata.

Üheks kõige olulisemaks tähelepanekuks on autori arvates jälgida, et küsimuste sihtgrupp oleks enne andmete võrdlemise asumist sama. Monitooringu andmete puhul kõige tüüpilisemaks näiteks võib tuua, et ühel aastal on küsitud küsimust kõigilt venekeelsele ankeedile vastajatelt ning teisel aastal hoopiski kõigilt mitte-eestlastelt. Taolist rahvuse ja keele põhjal küsimuste lahknemist tuli aastati päris palju ette. Seega on lõimumismonitooringu eri aastate andmete võrdlemisel oluline jälgida, kas küsimusele vastasid kõik respondendid või oli küsimus suunatud keele, rahvuse, kodakondsuse või hoopis mõne muu eelduse põhjal.

Lisaks ilmnas analüüsist, et eri aastatel on muu rahvuse osas kasutatud mitmeid eri termineid alustades venelastest ja venekeelsest elanikkonnast, kuni muu rahvuse või mitte-eestlasteni välja. Monitooringus kasutatavate terminite ühtlustamise vajadusele osutatakse ka 2019. aastal läbi viidud rekontseptualiseerimise aruandes (Kaldur jt 2019). Kuigi mingil määral on arusaadav, miks kasutatud terminid on aastate jooksul muutunud (tulenevalt lõimumispoliitikas üldiselt kasutatavast sõnavarast ning arengukava eesmärkidest ja indikaatoritest), siis andmete võrdlemisel tasub siiski tähele panna, kas uuritav küsimus on näiteks esitatud ainult venekeelse elanikkonna kohta või kõigi mitte-eestlaste kohta ning vastavalt vajadusele hinnata kuivõrd olulist muutust tulemustes terminite erinevus kaasa võib tuua.

Kalton ja Shuman (1982) toovad välja, et mitmed eri küsimused võivad näiliselt mõõta sama konstrukti, kuid tegelikkuses anda väga erinevaid tulemusi. Näiteks võib tuua siinkohal peatükis 3.2.2 tabelis 5 esitatud võrdsete võimaluste küsimuse, kus 2000. aastal küsiti vastajalt „*Kas Teie olete Eestis täheldanud, et eestlastel on võrreldes mitte-eestlastega samade eelduste korral kergem saavutada järgnevaid eesmärke?*“ ning 2002. aastal sisuliselt sama küsimus, kuid rahvuse suhtes ümber pööratult: „*Kas Teie olete Eestis täheldanud, et mitte-eestlastel on võrreldes eestlastega samade eelduste korral võrdsed võimalused saavutada järgnevaid eesmärke?*“. Pealtnäha tundub, nagu soovitakse mõõta sama konstrukti (võrdsed võimalused), kuid tulenevalt küsimuse sõnastusest, võivad saadud tulemused olla sisult erinevad. Kalton ja Shuman (1982) ning Shuman ja Presser (1996) tõid välja, et vastajatel on kalduvus väidetega üldjuhul pigem nõustuda ning sel juhul on väga oluline, kas väide on esitatud positiivse või negatiivsena. Sama loogikat võib kohandada ka toodud näite puhul: kui vastajatel on tendents pigem nõustuda, siis küsimuse rahvuse suhtes ümber pööramine võib anda eri aastatel väga erinevaid tulemusi.

Lisaks selgus ankeetküsimuste analüüsist, et monitooringutes on esinenud ka keerukaid või mitmeti mõistetavaid küsimusi, mida samuti ei peeta küsimustiku koostamise puhul heaks tavaks (Fowler ja Cosenza 2008). Näiteks ilmnes analüüsis, et aastate jooksul on olnud küsimusi, kus sisuliselt küsitakse mitut asja korraga või näiteks küsitakse küsimust mitte vastaja enda vaid mõne lähedase kohta, kelle kohta respondent ei pruugi osata adekvaatset vastust anda, mida samuti ei saa pidada heaks praktikaks (Neuman 2006). Eri aastate küsimuste puhul võis veel märgata, et vahel oli küsimustes määratletud ajaraamistik ning vahel mitte, mis omakorda võib Fowleri ja Cosenza (2008) hinnangul mõjutada vastaja tõlgendust ja seeläbi vastust küsimusele. Mõnedel juhtudel oli küsimustes esitatud vastajale täiendavaid selgitusi/sissejuhatust, mida jällegi aastate lõikes muudeti või ära kaotati, kuid mis Krosnick, Malhotra ja Mittali (2014) sõnul võib samuti põhjustada erinevusi tulemustes, kuna võivad mõjuda respondentide suunavalt.

Kõik eelnevalt mainitud aspektid võivad seega tekitada mõõtmisviga, mis esineb, kui respondendi vastus küsimusele on ebatäpne või vale. De Leeuw jt (2008) toovad välja, et mõõtmisvea vähendamiseks tuleks küsimustikku eelnevalt hoolikalt testida. Metoodika peatükke analüüsides selgus, et küsimustiku testimist mainiti vaid 2011. aastal, kui vastavalt testimise tulemusele küsimustikus muudatusi tehti. Seega võib küll küsimustiku testimisel saadud tagasiside aidata

kontrollida mõõtmisvea teket, kuid samal ajal mõjutada eri aastate andmete võrreldavust, kui selle tulemusel on küsimusi ankeetides muudetud.

Kokkuvõttes võib öelda, et monitooringu ankeedid varieeruvad küsimuste koostamise kvaliteedi ja hea tava poolest ning sellest tulenevalt jääb autor oma eelnevalt välja toodud järelduse juurde, et eri aastate andmete võrdlemisele tuleks läheneda küsimuse kaupa ning vastavalt vajadusele eelnevalt analüüsida, kas tulenevalt sõnastusest, vastusevariantide või küsimuse eelduste muutumisest on võimalik andmeid võrrelda või mitte. Seetõttu leiab ka autor, et ajaliste trendide või suuremamahuliste võrdlusanalüüside läbiviimine monitooringu andmete pealt oleks väga suur ettevõtmine, arvestades kogu lisatööd, mida andmete võrreldavale kujule viimine kaasa tooks.

Lüklikkus eri aastate andmetes ja sellest tulenev esinduslikkuse kadu

Eelnevalt arutles autor selle üle, kuivõrd esinduslikuks saab monitooringu valimeid pidada selle põhjal, milline info on dokumenteeritud valimite koostamise kohta. Valimi põhjal kogutud andmete esinduslikkust ei mõjuta aga üksnes valimi koostamine, vaid ka kogutud andmete (sh andmekogumise) kvaliteet. Seetõttu uuris autor lisaks ka andmete lüklikkust eri aastate monitooringutes.

Analüüsist selgus, et ulatusliku lüklikkusega (>10% vastustest) küsimusi oli aastati erineval määral, kuid keskmiselt vahemikus 13%-30% kogu küsimuste arvust. Kõige enam oli rohkem kui 10% puuduvate väärtustega küsimusi 2000., 2017. ja 2015. aastal (vastavalt 30%, 27% ja 21%). See tähendab, et lüklikkusest tingitud esinduslikkuse kadu võib olla kõige suuremaks probleemiks just nendel aastatel. Analüüs näitas, et andmelüklade esinemine küsimustes paistis rohkem olevat tingitud küsimuse sisust/teemast, kui keerulisest sõnastusest. See omakorda tähendab, et teatud teemade käsitlemisel peaks andmelüklade probleemi tõsisemalt suhtuma, kuna ulatuslik lüklikkus võib mõjutada küsimustes antud hinnangute esinduslikkust juhul, kui andmelüngad pole täiesti juhuslikud. Seda, et uuringutulemuste täpsus ei sõltu vaid lüklikkuse ulatusest, vaid pigem esinevatest mustritest ning lüklade iseloomust, toovad välja ka Tabachnick ja Fidell (2014).

Little ja Rubin (2002) toovad välja, et mittevastamisi saab oma olemuselt jagada kolmeks: juhuslikud andmelüngad (MAR), täiesti juhuslikud andmelüngad (MCAR) ja mittejuhuslikud andmelüngad (NMAR). Olenevalt andmelüklade iseloomust, tuleb nendega edasises analüüsis

erinevalt arvestada. Kõige vähem probleemsed on andmelüngad, kui neid saab pidada täiesti juhuslikuks ehk andmelüngad ei sõltu uuritavast tunnustest ega teistest tunnustest. Autor analüüsis 2020. aasta andmete põhjal suure lünkade osakaaluga tunnuste (lünki rohkem kui 10% indiviididest) lünklikkust viie taustatunnuse suhtes ja analüüsis selgus, et valitud aastal ei saa lünklikkust pidada täiesti juhuslikuks ühegi valitud taustatunnuse suhtes. Lisaks näitas analüüs, et kahe taustatunnuse suhtes (rahvus ja regioon) on lünklikkuse süstemaatilise probleem ulatuslikum, nad on seotud rohkemate analüüsi valitud (ulatusliku lünklikkusega) tunnustega kui teised valitud taustatunnused (sugu, tuleviku seotus Eestiga ning Eesti ühiskonda kuuluvuse tunne).

Kuna 2020. aasta ei olnud kõige ulatuslikuma lünklikkusega aasta, suuremaks probleemiks võis seda lugeda aastatel 2000, 2015 ja 2017, ning lünklikkus ilmnis suuresti küsimustes, mis esinesid aastati sarnastes teemades, võiks eeldada, et lünklikkuse muster on sarnane ka teistes monitooringu lainetes. Lisaks selgus analüüsist, et mittevastamisest tingitud esinduslikkuse kadu on nii rahvuse alusel kui üldkaaluga kaalutud andmestikus sarnane ehk seda ei kompenseeri andmestiku kaalumise uuringufirma poolt arvatud üldkaaludega. Nagu juba analüüsi osas välja toodud, oli tulemus teatud määral ootuspärane, kuna uuringufirma poolt arvatud kaalud on mõeldud täielikust (mitte osalisest) mittevastamisest tingitud esinduslikkuse kao vähendamiseks.

Kokkuvõtteks selgus tulemustest, et lünklikkust ei saa pidada täiesti juhuslikuks ja seetõttu võib monitooringu andmete analüüsimisel tekkida osalisest mittevastamisest tingitud viga (ning esinduslikkuse kadu). Vea suurus sõltub muuhulgas sellest, kas lünklikkus on juhuslik (MAR) või mittejuhuslik (NMAR), mis omakorda tuleneb sellest, kas tunnuses puuduvate vastuste jaotus erineb valiidsete vastuste jaotusest. Lünklikkust on võimalik korrigeerida näiteks mitmese imputeerimise või edasise osalist mittevastamist arvestava kaalumise, kuid selle täpsus on jällegi tingitud asjaolust, kas lüngad on juhuslikud (MAR) või mittejuhuslikud (NMAR) ja seda on praktikas peaaegu võimatu tagantjärele kindlaks teha.

Samas, kuna autor analüüsis lünkade juhuslikkust vaid ühe monitooringu laine põhjal ja sealgi võttis kriteeriumiks Benneti (2010) poolt soovitatud 10%-lise andmelünkade piiri, mille alusel sai analüüsitud vaid teatud osa tunnustest ning samuti vaid 5 taustatunnuse suhtes, oleks täpsemate järelduste tegemiseks vajalik põhjalikum edasine analüüs. Töös selgus, et ulatuslikku lünklikkust esines nii teemades, mida on läbivalt monitooringus käsitletud, nagu eestlaste ja mitte-eestlaste

võrdsed võimalused eri valdkondades või rahvuskultuuri- ja identiteedi säilitamine, aga ka küsimustes, mis olid omased vaid konkreetsele monitooringu aastale. Kui soovida teada, kas esinduslikkuse kadu võib olla tingitud näiteks madalast andmekogumise kvaliteedist (mis on iseenesest aspekt, mida annaks tulevikus parandada), tuleks analüüsida lünklikkust võrdlevalt aastate lõikes mõnes konkreetses teemas, mida läbivalt kajastatud on ja mis antud töö lünklikkuse analüüsis esile kerkis (näiteks eestlaste ja mitte-eestlaste võrdsed võimalused eri valdkondades). Kuna esindavust tuleks tegelikult käsitleda mitmemõõtmeliseks, võiks ka analüüs olla mitmemõõtmeline, vaadates esinduslikkust mitme tunnuse seisukohast korraga näiteks rahvuse lõikes eraldi valitud taustatunnuste suhtes.

Seega vajaks monitooringu andmete esinduslikkus kindlasti täiendavat analüüsi, et teha veelgi täpsemaid järeldusi monitooringu andmete võrreldavuse kohta. Käesolevas töös läbiviidud analüüsi põhjal võib pigem järeldada, et lünklikkus ei ole monitooringu andmetes täiesti juhuslik. Samas asjaolu, et lünklikkus ei ole lõimumismonitooringu andmetes täiesti juhuslik, ei ole midagi erakordset või midagi, mis automaatselt näitaks, et andmekogumise kvaliteet on madal. Küsitlusuuringute puhul on pigem tavapärane, et kõikides tunnustes ei ole lünklikkus täiesti juhuslik ning seega on oluline andmete analüüsil (sh võrdleval analüüsil) lünklikkuse süstemaatilisust ja selle võimaliku mõju analüüsitulemustele teadvustada ja teada, milliste tunnuste analüüsimisel võib lünklikkus probleemiks osutada.

Soovitused edasiseks uurimiseks

Kuigi käesolevas töös on analüüsitud nii monitooringu metoodikat, ankeetküsimustikke kui ka andmete lünklikkust, on siiski mitmeid erinevaid võimalusi, mida magistritöö edasiarendusena uurida võiks. Üheks kõige olulisemaks teemaks oleks ka venekeelsete ankeetide küsimuste kaardistamine ja võrdlemine ning sealhulgas eri aastate tõlke adekvaatsuse hindamine.

Lisaks oleks heaks edasiarenduseks monitooringu esinduslikkuse analüüs veelgi põhjalikumal kujul. Analüüsida võiks andmelünkade juhuslikkust ka teistes monitooringu lainetes ning võrdlevalt konkreetsetes teemades nagu eelpool väljatoodud. Või siis lünklikkuse asemel hoopiski uurida, mil määral esineb joonvastamist või kes on need respondendid, kes küsimustele vastamata jätavad. Ka taolisest analüüsist võib esile tulla olulisi mustreid või seoseid, mis võivad mõjutada andmete esinduslikkust või anda olulist tagasisidet uuringu disainimise protsessile.

Samuti jäi autorile magistritöö kirjutamise ajal silma, et lõimumismonitooringu aruannetes kasutatakse vahepeal võrdlusbaasina muude sarnaseid teemasid käsitlevate uuringute tulemusi. Seda toodi välja ka Balti Uuringute Instituudi rekontseptualiseerimise aruandes, et kattuvaid teemasid teiste uuringutega on üpris palju. Seega võiks üheks teema edasiarenduseks olla veel ka monitooringu andmete võrreldavuse analüüsimine teiste juba olemasolevate sarnaseid teemasid käsitlevate uuringutega (nt ESS või Statistikaameti isiku-uuringud).

Üldiseid soovitusi erinevate küsitlusuuringute andmete omavaheliseks võrdlemiseks

Kuigi käesolev magistritöö keskendus eri aastate andmete võrreldavuse analüüsimisele konkreetselt lõimumismonitooringu kontekstis, saab analüüsi tulemusel välja tuua mõned olulised tähelepanekud, mida iga uurija võiks kas taoliste jätku-uuringute andmete võrdlemisel või lihtsalt eri uuringute andmetega võrdleva analüüsi teostamisel silmas pidada ning vajadusel eelnevalt analüüsida.

Metoodika võrdlemisel on oluline peatuda järgnevatel aspektidel: üldkogum, valimi koostamine ja esinduslikkus ning andmekogumismeetod. Nendest lähtuvalt võiks enne võrdleva analüüsi teostamist vastata järgnevatele küsimustele:

- Kes moodustavad uuringutes üldkogumi? Kas üldkogumi definitsioon on eri uuringutes sõnastatud ühetaoliselt (ka väikesed erinevused sõnastuses võivad üldkogumit ehk seda, millisele kogumile on uuringutulemused ja järeldused üldistatavad, muuta)? Kui analüüsi all on vaid mingi konkreetne sihtgrupp, siis uurida, kas sihtgrupp on osa uuritavast populatsioonist kõigis vaadeldavates uuringutes.
- Millised andmed on võetud valikuraami aluseks?
- Millised on olnud valimi moodustamise põhimõtted eri uuringutes? Kas valim on moodustatud tõenäosuslikult või mittetõenäosuslikult, milline valimitüüp on valitud ning kas ja milliste populatsiooni karakteristikute suhtes taotletakse esinduslikkust?
- Kuidas andmeid on kogutud? Erinevat tüüpi andmekogumisviisid võivad omada erisugust mõju tulemustele ka juhul, kui valimid, küsimused jm metoodika on identsed. Tasub jälgida, millised reeglid on kehtestatud küsitluse tööle, kuidas on tagatud intervjuerija mõju vähendamine küsitlustöös ja kas andmekogumisele on teostatud järelkontrolli, mida võib lugeda üheks kvaliteetse andmekogumise tagamise indikatsiooniks.

Lisaks tuleks veenduda, et võrdlevasse analüüsi valitud **küsimused** oleks omavahel võrreldavad. Lähtuda võiks küsimuste võrdlemisel järgnevatest küsimustest:

- Kas ning kuivõrd erineb küsimuse sõnastus erinevates uuringutes?
- Kas küsimused on üheselt mõistetavad?
- Kas ning kuivõrd on vastuseskaala/variandid erinevad?
- Kellelt on küsimust küsitud? Kas küsimusele eelneb suunamisi/on eeldusi?
- Kas küsimusele on olnud võimalik vastamata jätta?

Kui ka selgub, et küsimused on võrreldavates uuringutes erinevalt esitatud või näiteks ühest uuringust puudub täpselt sarnaselt küsitud küsimus, tasub mõelda sellele, kuidas on defineeritud analüüsi keskmes olev uuritav nähtus ning kas seda on võimalik mõõta mõne teise uuringus esitatud küsimuse kaudu. Sageli on võimalik küsimusi andmestikus ümber kodeerida ning sellest lähtuvalt teatud tasemel võrdlevat analüüsi siiski teostada.

Kogutud andmete järeltöötlemises võib olla ka oluline, kas uuringute raames on andmekogumise järgselt arvatud andmestikku kaalud, mis peaksid kompenseerima andmekogumisest tingitud esinduslikkuse kadu. Andmete kaalumisel oleks oluline teada, millist meetodikat on kaalude arvutamisel kasutatud (nt nii disainikaalude, mittevastamise kui järelkihistamiskaalude eesmärk on laias laastus alati andmete esinduslikkuse parandamine, kuid täpsem eesmärk ja eeldused kaalude arvutamisele on neil erinevad) ja milliseid tunnuseid on kaalude arvutamisel kasutatud ehk milliste tunnuste suhtes eri uuringu andmed esinduslikkust taotlevad. Sellest lähtuvalt saab uurija ise otsustada, kas andmete kaalumise aitab parandada andmete võrreldavust või peab uurija ise selleks täiendavaid samme ette võtma.

Kui juba jõutakse andmete analüüsimiseni, tasub vaadata, kuivõrd esineb uuritavates tunnustes lünklikkust. Kui lünklikkus on ulatuslik (näiteks üle 10%-i kõigist vastajatest), oleks mõistlik kindlaks teha, kas seda võib pidada täiesti juhuslikuks või mitte. Paljud statistilise andmeanalüüsi meetodid eeldavad täielikku andmestikku ning seetõttu jäetakse andmelüngad sageli analüüsis kõrvale. Kui lünklikkus on süstemaatiline mõne teise tunnuse suhtes, vähendab andmelünkade analüüsist välja jätmise valimi esinduslikkust ja tulemuste täpsust. Seega on oluline lünklikkuse iseloomu (MCAR või MAR/NMAR) kindlaks tegemine, et tulemustest lähtudes saaks otsustada, kas lünklikkust on mõtet korrigeerida näiteks mitmese imputeerimise või mõne muu meetodi abil.

Üheks eelduseks, et eelnevalt väljatoodud soovitusi jälgida, on muidugi võrreldavate uuringute puhul põhjaliku ja detailse dokumentatsiooni olemasolu.

KOKKUVÕTE

Muust rahvusest inimeste integratsioon Eesti ühiskonda on pikaajaline protsess, mis on oluline, et tagada ühiskonna stabiilsus ja sidusus. Selleks, et jälgida, millised arengud lõimumisprotsessides on toimunud ning hinnata integratsioonipoliitika edukust, viiakse alates 2000. aastast läbi Eesti ühiskonna integratsiooni monitooringut. Uuringu näol on tegemist jätku-uuringuga, mida on teostatud juba kaheksal korral ning mille tulemusi kasutavad nii poliitikakujundajad, teadlased kui laiem avalikkus. Kuigi monitooringu läbiviimisel koostatakse igal lainel põhjalik raport, on vähe tehtud pikaajalisi põhjalikumaid analüüse. Aastate jooksul on aga muutunud integratsiooni mõtestamine, samuti on lõimumismonitooringut läbi viinud erinevad uurimisrühmad. Monitooringu andmete ajaliste võrdluste tegemiseks peaks seega eelnevalt veenduma, et uuringud on eri aastatel meetodiliselt sarnaselt disainitud ja läbiviidud ning tagatud on andmete võrreldavus varasemate aastatega.

Sellest lähtuvalt analüüsiti käesolevas magistritöös, mil määral on lõimumismonitooringu küsitlusmetoodika, ankeetküsimused ning andmete esinduslikkus aastate jooksul muutunud. Uuringumetoodikat analüüsis autor raportites esinenud metoodikakirjelduste ning Riigihangete registrist kättesaadavate hankedokumentide ning hanke raames tellijale esitatud küsimuste ja vastuste põhjal kasutades kontentanalüüsi meetodit. Ankeetküsimuste sõnastuse analüüsis toetus autor varasemale erialasele kirjandusele. Andmete esinduslikkuse analüüsis uuris autor lähemalt eri aastate monitooringu andmetes esinevat lünklikkust. Põhjalikumalt analüüsiti lünklikkust 2020. aasta uuringu andmete põhjal, kasutades selleks kirjeldavat statistikat ning seoste hindamiseks hii-ruut-testi (sealjuures korrigeeriti olulisuse nivood Benjamini-Hochbergi meetodil).

Analüüsi tulemustest selgus, et metoodika osa on monitooringu aruannetes väga erineva detailsusega dokumenteeritud. Esimesel neljal aastal mainiti metoodikat enamasti vaid paari lausega raporti sissejuhatuses, alates 2011. aastast on metoodikakirjeldus lisatud eraldi peatükina. Edasine analüüs selgitas, et hoolimata lisatud metoodikapeatükkidest jäi siiski enamiku aastate puhul seal esitatud info ebapiisavaks, et teha täpseid järeldusi eri aasta monitooringute metoodika võrreldavuse kohta.

Tulemused näitasid, et aastate jooksul on muutunud monitooringu elanikkonnaküsitluse üldkogum, kuid muutus on toimunud vaid vanuselises kaasamises: aastatel 2000-2011 küsitleti kõiki Eesti püsielanikke vanuses 15-74 aastat ning alates 2015. aastast kaotati vanuseline ülempiir ehk küsitleti kõiki vähemalt 15-aastaseid elanikke. Taolist muutust populatsioonis saab aga andmete järeltöötlemises arvesse võtta (üle 74-aastased välja jätta) ning aastate omavahelises võrdluses tuleb eelnevalt veenduda, et võrreldavad sihtrühmad oleks samast populatsioonist.

Lisaks selgus analüüsist, et valimi koostamise põhimõtted on aastate jooksul mõnevõrra muutunud. Üks olulisemaid muudatusi toimus 2011. aastal, kui suurendati mitte-eestlaste osakaalu valimis. Sarnaselt on ka kõigil järgnevatel aastatel mitte-eestlaste osakaal valimis olnud üldkogumiga võrreldes ebaproportsionaalselt suurem. Valimi koostamise kohta sai autor olemasoleva dokumentatsiooni põhjal (ainult aastate 2011-2020 kohta) järeldada, et valimid on koostatud mitmeastmelise valikuna, kuid kahtluse alla võib seada valimite juhuslikkuse, kuna puudub piisav dokumentatsioon (täpsete valikureeglite ning näiteks valikuraami kohta). Mahu poolest on valimid aastati kõikunud 1000-1500 indiviidi vahel, mida üldiselt peetakse piisavaks valimimahuks.

Ankeetide analüüsimisest selgus, et ankeetid on aastate lõikes oluliselt muutunud. Põhjuseid võib siinkohal välja tuua mitmeid: uurimisrühmade ja sealjuures küsimustiku koostajate vahetumine; olulised toimunud võtmesündmused (näiteks Pronksiöö 2007. aastal, Ukraina kriis 2014-2015, koroonapandeemia); aastate jooksul on olnud erinevaid arengukava perioode, mille lõikes on muutunud integratsiooni mõtestamine ja hinnatavad mõõdikud; 2020. aasta küsimustik muutus oluliselt pärast Balti Uuringute Instituudi poolt läbiviidud lõimumismonitooringu rekontseptualiseerimise aruandes antud soovitusi. Sellegi poolest on monitooringutes läbivalt kajastatud sarnaseid teemasid nagu tööturg, haridus, kontaktid ja keeleoskus, meedia, ühiskondlik osalus, võrdsed võimalused ning üldised lõimumisalased hoiakud. Üksteisele sarnasemad paistsid algusaastate 2000-2004 küsimustikud ning alates 2015. aastast läbiviidud monitooringute küsimustikud.

Ankeediküsimuste sõnastuse analüüs näitas, et täiesti muutumatul kujul (sõnastus ja vastusevariandid) ei leidu ankeetides ühtegi küsimust, mida oleks läbivalt kõigis kaheksas monitooringus küsitud. See on huvitav tulemus, kuna sageli eeldatakse taoliste jätku-uuringute puhul ankeetide muutumatust ajas. Samas ei tähenda see, et andmeid ei oleks üldse võrreldavad

või puuduks täielikult võimalus ajalisi trende jälgida. Analüüs näitas, et nii mõnegi vaadeldud küsimuse puhul võis mõõdetavat konstrukti pidada samaks hoolimata sellest, et sõnastus või vastusevariandid on aastate jooksul muutunud. Sellistel juhtudel on sageli võimalik tunnuseid ümber kodeerida üldisemale tasemele ning sellega tagada andmete võrreldavus eri aastatel. Olulisim oli aga ankeediküsimuste analüüsist selgunud järeldus, et kuna küsimuse sõnastuse ja vastusevariandi muutmisel võib olla mõju ka saadud tulemustele, tuleks andmete kasutajal siiski ajaliste võrdluste tegemisel läheneda igale küsimusele individuaalselt ning eraldi hinnata küsimuses esinevate muutuste mõju andmete võrreldavusele.

Lüklikkuse analüüsi tulemused näitasid, et aastate võrdluses võib lüklikkust suuremaks probleemiks pidada aastatel 2000, 2015 ja 2017. Ulatusliku lüklikkusega küsimusi uurides selgus, et andmelünkade esinemine küsimustes paistis rohkem olevat tingitud küsimuste sisust kui keerulisest sõnastusest. 2020. aasta monitooringu andmete täpsem analüüs näitas, et osade tunnuste puhul lüklikkust ei saa pidada täiesti juhuslikuks (MCAR) ning analüüsist välja tulnud mustrite põhjal võiks seda eeldada ka vähemalt monitooringu aastate kohta kui lüklikkuse probleem oli suurem (ehk 2000, 2015 ja 2017). See tähendab, et monitooringu andmete analüüsimisel võib tekkida osalisest mittevastamisest tingitud viga, mis võib vähendada andmete esinduslikkust ja tulemuste täpsust. Tulemused näitasid, et ka andmete kaalumise uuringufirma poolt arvutatud kaaludega, ei kompenseeri osalisest mittevastamisest tingitud esinduslikkuse kadu.

Kokkuvõttes võib öelda, et monitooringute olemasoleva dokumentatsiooni põhjal on raske teha põhjanevaid järeldusi eri aastate andmete võrreldavuse kohta. Raporteid lugedes jääb mulje, et uurimisrühmad ise taotlevad andmete esinduslikkust (olenevalt aastast näiteks soo, vanuse, rahvuse ja regiooni suhtes), kuid käesoleva magistritöö analüüs seab selle siiski kahtluse alla. Seega on taoliste jätku-uuringute, nagu seda on lõimumismonitooring, planeerimisel ja läbiviimisel väga oluline meetodikaraamistiku kokku leppimine ja selle detailne dokumenteerimine, et järgnevatel aastatel oleks võimalik dokumentatsiooni põhjal uuringumetoodika ja kogutud andmete võrreldavus tagada.

Autor leiab, et magistritöös läbiviidud analüüs on väärtuslik mitte ainult lõimumismonitooringu andmete kasutajatele, vaid ka üldiselt erinevate jätku-uuringute andmeid analüüsivale teadlasele või uurijale. Magistritöö analüüs näitab selgelt, et isegi, kui uuringut on korraldatud ajalise võrdluse võimaldamise eesmärgiga, ei pruugi üks-ühene võrreldavus realiseeruda, ning toob välja

olulisemad aspektid, millele tähelepanu pöörata. Käesolev teema on kokkuvõttes uuringumetoodiliselt väga huvitav ning pakub mitmeid erinevaid võimalusi edaspidiseks uurimiseks. Autor leiab, et kindlasti tuleks analüüsida ka venekeelsete ankeetide samasust ajas (ning tõlgete adekvaatsust) ning täiendavat analüüsi saaks läbi viia näiteks ka andmete lünklikkuse osas, uurides lünklikkust põhjalikumalt kõigis monitooringu lainetes.

KASUTATUD KIRJANDUS

Andrews, F. M. (1984). Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach. *The Public Opinion Quarterly*, 48(2), 409–442. <http://www.jstor.org/stable/2749034>

Babbie, E. (2010). *The practice of social research*. Belmont: Wadsworth Publishing

Beilmann, M. (2020). Küsitlusuuringud. URL (kasutatud 20.06.2022) <https://samm.ut.ee/k%C3%BCsitlusuuringud>

Benjamini, Y. ja Y. Hochberg. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. <http://www.jstor.org/stable/2346101>

Bennett, D. A. (2001). How can I deal with missing data in my study?. *Australian and New Zealand journal of public health*, 25(5), 464-469

Biemer, P. P. (2010). TOTAL SURVEY ERROR: DESIGN, IMPLEMENTATION, AND EVALUATION. *The Public Opinion Quarterly*, 74(5), 817–848. <http://www.jstor.org/stable/40985407>

Biemer, P. P. ja L. E. Lyberg. (2008). Quality Assurance and Quality Control in Surveys. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 421-441). New Yourk: Routledge

Biemer, P. P. ja S. Christ. (2008). Weighting Survey Data. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 317-341). New Yourk: Routledge

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of public health*, 27(3), 281-291

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, 10(4), 417-451

Davern, M. E. (2008). Representative sample. P. J. Lavrakas (toim). *Encyclopedia of Survey Research Methods*. (lk 720–723). Washington DC: SAGE Publications, Inc

De Leeuw, E. D. (2008). Choosing Methods of Data Collection. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 113-135). New Yourk: Routledge

De Leeuw, E. D., Hox, J., ja Dillman, D. (2008). *International handbook of survey methodology*. New Yourk: Routledge

Documentation of ESS Post-Stratification Weights. (2014). URL (kasutatud 25.06.2022) https://www.europeansocialsurvey.org/docs/methodology/ESS_post_stratification_weights_documentation.pdf

Dong, Y., ja C. Y. Peng. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(1), 222. <https://doi.org/10.1186/2193-1801-2-222>

Eesti ühiskonna lõimumise monitooring 2020. (2020). SA Poliitikauuringute Keskus Praxis, Balti Uuringute Instituut, Tallinna Ülikool, Tartu Ülikool, Turu-uuringute AS. URL (kasutatud 25.06.2022) <https://www.kul.ee/media/3240/download>

ESS tutvustus. (2022).URL (kasutatud 24.06.2022) <https://uhiskond.ut.ee/et/sisu/ess-tutvustus>

Fowler, F. J. ja C. Cosenza. (2008). Writing effective questions. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 136-160). New Yourk: Routledge

Garbarski, D., Schaeffer, N. C. ja J. Dykema. (2016). The effect of response option order on self-rated health: a replication study. *Quality of Life Research*, 25(8), 2117–2121. <http://www.jstor.org/stable/44852503>

Gendall, P. ja J. Hoek. (1990). A question of wording. *Marketing Bulletin*, 1(5), 25-36. Chicago

Groves, R. M. ja L. Lyberg. (2010). TOTAL SURVEY ERROR: PAST, PRESENT, AND FUTURE. *The Public Opinion Quarterly*, 74(5), 849–879. <http://www.jstor.org/stable/40985408>

Hirsjärvi, S., Remes, P. ja P. Sajavaara. (2005). Uuri ja kirjuta. Tallinn: Medicina

James, G., Hastie, T., Tibshirani R. ja D. Witten. (2021). *An introduction to statistical learning: With applications in R*. Springer. https://hastie.su.domains/ISLR2/ISLRv2_website.pdf

Integratsiooni monitooring 2011. (2011). AS Emor, SA Poliitikauuringute Keskus Praxis, Tartu Ülikool. URL (kasutatud 25.06.2022) <https://www.kul.ee/media/389/download>

Kaldur, K., Kivistik, K. ja Pohla T. (2019). *Integratsiooni monitooringu meetodika (re)kontseptualiseerimine*. Balti Uuringute Instituut. URL (kasutatud 15.05.2022) <https://www.ibs.ee/wp-content/uploads/2022/01/EIM-aruanne-1.pdf>

Kaldur, K. (2021). Eesti Integratsiooni Monitooring(ud). Presentatsioon, *20 aastat lõimumismonitooringu andmestikku Eestis – tabamata seosed kaardistamata aladel*. Veebiseminar, 16. märts. Rahvusvaheliste Sotsiaaluuringute Keskus

Kaldur, K., Lauristin, M., Vetik, R., Vihalemm, T. ja K. Kivistik. (2021). 20 aastat uurimist. Lõimumise monitooring. *Paneelarutelu, Rahvusvaheline konverents: 30 aastat lõimumist*. Tallinn, 11.-12. november. Integratsiooni Sihtasutus. Youtube (kasutatud 20.06.2022) https://www.youtube.com/watch?v=YC64AgG0iyY&ab_channel=IntegratsiooniSihtasutus

Kalton, G. (1983). *Introduction*. In *Introduction to survey sampling*. SAGE Publications, Inc., 5-8. <https://dx.doi.org/10.4135/9781412984683.n1>

Kalton, G. ja H. Schuman. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society: Series A (General)*, 145(1), 42-57

Kamoen, N., Holleman, B., van den Bergh, H. ja T. Sanders. (2013). Positive, negative, and bipolar questions: The effect of question polarity on ratings of text readability. In *Survey Research Methods* (Vol. 7, No. 3, 181-189)

Kish, L. (1965). *Survey sampling*. New York: Wiley Pty Ltd

Krosnick, J. A., Malhotra, N., ja U. Mitta. (2014). Public misunderstanding of political facts: How question wording affected estimates of partisan differences in birtherism. *Public opinion quarterly*, 78(1), 147-165

Körner, T. ja I. Meyer. (2005). Harmonising socio-demographic information in household surveys of official statistics: experiences from the Federal Statistical Office Germany. J. H. P. Hoffmeyer-Zlotnik, ja J. Harkness (toim.), *Methodological aspects in cross-national research* (lk 149-162). Mannheim: GESIS-ZUMA

Lagerspetz, M. (2017). Ühiskonna uurimise meetodid. Sissejuhatus ja väljajuhatus. Tallinn: Tallinna Ülikooli Kirjastus

Lavallée, P. ja J. F. Beaumont. (2016). Weighting: Principles and practicalities. Wolf, C., Joye, D., Smith, T. E., Smith, T. W., ja Fu, Y. C. (toim). *The SAGE Handbook of Survey Methodology*. (lk 460-476.) SAGE Publications Ltd

Lepik, N. ja I. Traat. (2013). Valikuuringu teooria I. Õppematerjal. Tartu Ülikool. URL (kasutatud 25.06.2022) <http://dspace.ut.ee/bitstream/handle/10062/30680/ValiuurI.pdf>

Little, R. J. A. ja D. B. Rubin (2002). *Statistical analysis with missing data* (2nd ed). New York: Wiley

Lohr, L.(2008). Coverage and Sampling. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 97-112). New Yourk: Routledge

Lõimuv Eesti 2020. (i.a). *Kultuuriministeeriumi koduleht*. URL (kasutatud 23.06.2022) <https://www.kul.ee/kultuuriline-mitmekesisus-ja-loimumine/strateegilised-dokumendid/loimuv-eesti-2020>

Lyberg, L. (2012). Survey quality. *Survey Methodology*, 38(2), 107-130

Mohler, P. P., Pennel B.-E. ja F. Hubbard (2008). Survey Documentation: Toward Professional Knowledge Management in Sample Surveys. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 403-420). New Yourk: Routledge

Neuman, W. L. (2006). *Social Research Methods: qualitative and quantitative approaches* (Ed. 6th). Harlow: Pearson Education Limited

Ruxton, G. D. ja M. Neuhäuser. (2010). Good practice in testing for an association in contingency tables. *Behavioral Ecology and Sociobiology*, 64(9), 1505–1513. <http://www.jstor.org/stable/40863216>

RV021: rahvastik soo ja vanuserühma järgi, 1. Jaanuar. *Statistikaameti andmebaas*. (i.a). URL (kasutatud 15.08.2022) https://andmed.stat.ee/et/stat/rahvastik_rahvastikunaitajad-ja-koosseis_rahvaarv-ja-rahvastiku-koosseis/rv021/table/tableviewlayout2

RV0221: rahvastik soo, vanuserühma ja maakonna järgi, 1. Jaanuar. Eestlased , (1990-2017). *Statistikaameti andmebaas.* (i.a). URL (kasutatud 15.08.2022) https://andmed.stat.ee/et/stat/Lepetatud_tabelid_Rahvastik.Arhiiv_Rahvastikun%C3%A4itaja_d%20ja%20koosseis.%20Arhiiv/RV0221

RV0221U: rahvastik soo, vanuserühma ja maakonna järgi, 1. Jaanuar. Eestlased. Haldusjaotus seisuga 01.01.2018. (i.a). *Statistikaameti andmebaas.* URL (kasutatud 15.08.2022) https://andmed.stat.ee/et/stat/rahvastik_rahvastikunaitajad-ja-koosseis_rahvaarv-ja-rahvastiku-koosseis/RV0221U

Rässler S., Rubin D. B. ja N. Schenker. (2008). Incomplete Data: Diagnosis, Imputation and Estimation. E. D. De Leeuw, J. Hox, ja D. Dillman (toim). *International handbook of survey methodology* (lk 370-386). New Yourk: Routledge

Schaeffer, N. C. ja S. Presser. (2003). The science of asking questions. *Annual review of sociology*, 29(1), 65-88

Schuman, H. ja S. Presser. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context.* Sage

Sidusa Eesti arengukava 2021-2030. (i.a). *Kultuuriministeeriumi koduleht.* URL (kasutatud 23.06.2022) <https://www.kul.ee/siduseesti2030>

Spieß, M. (2016). Dealing with missing values. Wolf, C., Joye, D., Smith, T. E., Smith, T. W., ja Fu, Y. C. (toim). *The SAGE Handbook of Survey Methodology.* (lk 597-612). SAGE Publications Ltd

Tabachnick, B. G. ja L. S. Fidell. (2014). *Using multivariate statistics* (6th ed). Harlow: Pearson

Tiit, E-M. ja Tooding, L-M. (2019). *Statistikaleksikon.* Tartu: Tartu Ülikooli Kirjastus

Tooding, L.-M. (2015). *Andmete analüüs ja tõlgendamine sotsiaalteadustes.* Tartu: Tartu Ülikooli Kirjastus

Toepoel, V. (2016). Sampling. *Doing surveys online* (lk 53-67). SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781473967243.n4>

Tourangeau, R., Conrad, F. G., ja M. P. Couper. (2013). *The Science of Web Surveys*. Oxford University Press

Varasemad lõimumiskavad. (i.a). *Kultuuriministeeriumi koduleht*. URL (kasutatud 23.06.2022) <https://www.kul.ee/kultuuriline-mitmekesisus-ja-loimumine/strateegilised-dokumendid/varasemad-loimumiskavad>

Vehovar, V., Toepoel, V. ja Steinmetz, S. (2016). Non-probability sampling. E. Wolf, D. Joye, T. W. Smith ja Y-c. Fu (toim), *The SAGE Handbook of survey Methodology* (lk 329–345). London: SAGE.

Wolf, C., Schneider, S. L., Behr, D. Ja D. Joye. (2016). Harmonizing survey questions between cultures and over time. *The SAGE handbook of survey methodology*, (lk 502-524)

LISAD

LISA 1. Kodeerimisjuhised

| Tunnus | Selgitus | Tüüp | Kategooriad |
|--|---|-------------|--|
| Monitooringu aasta | | Number | 2000-2020 |
| Raporti pealkiri | | Vabatekst | |
| Uuringu teostaja(d) | Vastava aasta monitooringu teostaja ehk hanke võitja | Vabatekst | |
| Ankeetküsitluse läbiviija | Uuringusse kaasatud uuringufirma, kes teostas ankeetküsitluse andmekogumist | Vabatekst | |
| Raporti autorite nimed | | Vabatekst | |
| Kas monitooringu aruanne on veebis kättesaadav | | Kategooria | Jah/Ei |
| Link monitooringu aruandele veebis | | Vabatekst | |
| Kas monitooringu raportis on kirjeldatud uuringu meetodikat? | | Kategooria | Jah, meetodika kohta on eraldi peatükk Jah, kirjeldatud vähesel määral Ei |
| Uurimisviis | | Kategooria | Kvantitatiivne Kvalitatiivne Kombineeritud |
| Andmekogumise meetodid | Milliseid andmete kogumise meetodeid on monitooringu koostamisel kasutatud? | Kategooria | Silmast silma küsitlus Fookusgrupp Veebiküsitlus Süvaintervjuu Küsitlus (pole täpsustatud) |

| | | | |
|---|---|------------|---|
| Fookusgrupi sihtrühmad | | Vabatekst | |
| Ankeetküsitluse üldkogumi kirjeldus | | Vabatekst | |
| Kas üldkogumi suurus on monitooringu raportis välja toodud? | | Kategooria | Jah Ei |
| Valikuraami kirjeldus | | Vabatekst | |
| Ankeetküsitluse valimi suurus | Lõplikult kujunenud valimi suurus monitooringu kvantitatiivses osas | Number | |
| Rahvuse jaotus valimis | Ankeetküsitluse valimi jaotus rahvuse põhiselt (eestlased vs mitte-eestlased) | Number | |
| Ankeetküsitluse valimi tüüp | | Kategooria | Tõenäosuslik valim Mittetõenäosuslik valim Info puudub |
| Ankeetküsitluse valimi alatüüp | | Kategooria | Kõikne valim Lihtne juhuvalim Süsteemaatiline juhuvalim Stratifitseeritud juhuvalim Klastervalim Mitmeastmeline valim Lumepallivalim Mugavusvalim Iseselekterumine Strateegiline valim |
| Valimipunktide arv | | Number | Kolmekohaline arv Info puudub |
| Lähteadressi meetodi kasutamine | Kas andmekogumisel on kasutatud lähteadressi meetodit | | Jah Ei Info puudub |
| Nn noore mehe/noore naise reegli kasutamine | Kas andmekogumisel on kasutatud noore mehe/noore naise reeglit | Kategooria | Jah Ei Info puudub |
| Lisavalim | Lisavalim(ite) kirjeldus | Vabatekst | |

| | | | |
|---|--|------------|---|
| Kas järelkontroll on teostatud? | | Kategooria | Jah Ei Info puudub |
| Kas monitooringu aruandes on välja toodud valimiviga? | | Kategooria | Jah Ei |
| Kas monitooringu aruandes on välja toodud vastamismäär? | | Kategooria | Jah Ei |
| Kas monitooringu aruandes on infot toodud kaalude arvutamise kohta? | | Kategooria | Jah Ei |
| Mis tüüpi kaalud on arvutatud? | | Kategooria | Järelkihistuskaalud Disainikaalud Info puudub |
| Millised tunnused on kaalude arvutamise aluseks olnud | | Vabatekst | |

LISA 2. Integratsiooni monitooringu uuringuraportite ülevaade

| Aasta | Raporti pealkiri | Uuringu teostajad | Küsitluse läbiviija | Raporti autorid | Veebilink raportile |
|-------|--|--|---------------------|--|---|
| 2000 | Integratsioon Eesti ühiskonnas: Monitooring 2000 | TPÜ Rahvusvaheliste ja Sotsiaaluuringute Instituut | Saar Poll OÜ | Marju Lauristin, Iris Pettai, Marje Pavelson, Jüri Kruusvall, Raivo Vetik, Klara Hallik, Külliki Korts, Ragne Kõuts, Ivi Proos, Piia Tammpuu, Olga Peresild, Denis Trapido, Jelena Ilina, Valeria Jakobson, Triin Vihalemm | https://www.integratsioon.ee/integratsioon-eesti-uhiskonnas-monitooring-2000 |
| 2002 | Integratsioon Eesti ühiskonnas: Monitooring 2002 | TPÜ Rahvusvaheliste ja Sotsiaaluuringute Instituut | Saar Poll OÜ | Klara Hallik, Jüri Kruusvall, Marje Pavelson, Iris Pettai, Ivi Proos, Raivo Vetik | https://www.integratsioon.ee/integratsioon-eesti-uhiskonnas-monitooring-2002 |
| 2005 | Integratsiooni monitooring 2005 | TPÜ Rahvusvaheliste ja Sotsiaaluuringute Instituut | Saar Poll OÜ | Klara Hallik, Jüri Kruusvall, Marje Pavelson, Iris Pettai, Ivi Proos, Raivo Vetik | https://www.integratsioon.ee/en/integratsiooni-monitooring-2005-uuringu-aruanne |
| 2008 | Eesti ühiskonna integratsiooni monitooring | Tallinna Ülikool | Saar Poll OÜ | Raivo Vetik, Kairi Kasearu, Avo Trumm, Jelena Helemäe, Ellu Saar, Triin Vihalemm, Külliki Korts, Jüri Kruusvall, Peeter Vihalemm, Gerli Nimmerfeldt, Marju Lauristin | https://www.integratsioon.ee/eesti-uhiskonna-integratsiooni-monitooring-2008 |
| 2011 | Eesti ühiskonna integratsiooni monitooring | AS Emor, SA Poliitikauuringute Keskus Praxis, Tartu Ülikool | AS Emor | Marju Lauristin, Esta Kaal, Laura Kirss, Tanja Kriger, Anu Masso, Kirsti Nurmela, Külliki Seppel, Tiit Tammaru, Maiu Uus, Peeter Vihalemm, Triin Vihalemm | https://www.kul.ee/eesti-uhiskonna-integratsiooni-monitooring-2011 |
| 2015 | Eesti ühiskonna integratsiooni monitooring | Balti Uuringute Instituut, Tallinna Ülikool, SA Poliitikauuringute Keskus Praxis | Turu-uuringud AS | Kristina Kallas, Raivo Vetik, Jüri Kruusvall, Ellu Saar, Jelena Helemäe, Laura Kirss, Cenely Leppik, Külliki Seppel, Kats Kivistik, Pille Ubakivi-Hadachi | https://www.kul.ee/eesti-uhiskonna-integratsiooni-monitooring-2015 |
| 2017 | Eesti ühiskonna integratsiooni monitooring | Balti Uuringute Instituut, SA Poliitikauuringute Keskus Praxis | Turu-uuringud AS | Kristjan Kaldur, Raivo Vetik, Laura Kirss, Kats Kivistik, Külliki Seppel, Kristina Kallas, Märt Masso, Kristi Anniste | https://www.kul.ee/eesti-uhiskonna-integratsiooni-monitooring-2017 |
| 2020 | Eesti ühiskonna lõimumismonitooring | SA Poliitikauuringute Keskus Praxis, Balti Uuringute Instituut, Tallinna Ülikool, Tartu Ülikool, Turu-uuringute As | Turu-uuringud AS | Eve Mägi, Ivan Polynin, Katrina Koppel, Kats Kivistik, Kirsti Meesk, Kristi Anniste, Kristjan Kaldur, Külliki Seppel, Mari-Liis Sepper, Meeli Murasov, Märt Masso, Nawal Shaharyar, Nikolai Kunitsõn, Raivo Vetik, Triin Pohla | https://www.kul.ee/eesti-integratsiooni-monitooring-2020 |

LISA 3. Andmete lünklikkuse analüüsi kaasatud 2020. a lõimumismonitooringu küsimused

| Tunnus | N | Küsimus |
|---------|------|---|
| K3X1_4 | 1414 | Kuidas hindate eestlaste ja teisest rahvusest inimeste võimalusi, olukorda järgnevates valdkondades? Poliitilise tegevuse võimalused |
| K3X1_5 | 1414 | Kaasatus oma linna, valla elu mõjutavate otsuste tegemisse |
| K3X1_7 | 1414 | Võimalus saada juhtivale ametikohale riigi- ja omavalitsusasutustes |
| K3X1_9 | 1414 | Osalemine kodaniku-, vabäihendustes |
| K3X1_10 | 1414 | Võimalus saada endale vajadustele vastav eluase |
| K3X1_11 | 1414 | Võimalus saada usaldusväärset informatsiooni ühiskonna kohta |
| K3X3 | 1414 | Eesti tööturu statistika näitab, et venekeelsed elanikud saavad eestlastega võrreldes umbes 10-20% vähem palka ja nende tööpuudus on oluliselt suurem kui eestlastel. Kummast põhjusest tuleneb Teie arvates eesti- ja venekeelsete inimeste ebavõrdsus tööturul? |
| K3X4_4 | 1414 | Kui rahul olete järgmiste teenuste kättesaadavusega oma linnas/vallas? Lasteaiad ja lastehoid |
| K3X4_14 | 1414 | Tööturuteenused |
| K3X4_15 | 1414 | Eesti keele õpe |
| K3X5_1 | 920 | Mainisite, et Te ei ole rahul mõne teenuse kättesaadavusega oma linnas/vallas. Kui olulised on Teie jaoks järgmise põhjused, miks Te ei ole rahul teenuse kättesaadavusega? Teenuste koha ei ole piisavat informatsiooni mulle sobivas keeles |
| K3X5_2 | 920 | Teenuseid ei osutata mulle sobivas keeles |
| K3X5_3 | 920 | Teenindajad suhtuvad minusse tõrjuvalt, ebaviisakalt minu päritolu või keele tõttu |
| K3X5_4 | 920 | Teenindajad kohtlevad mind põhjendamatult ebavõrdselt minu päritolu või keele tõttu |
| K4X7_3 | 1414 | Kuidas Te suhtute sellesse, kui Te oleksite korteri- või majanaaber paljude ... ? uus-sisserändajatega ehk viimase 5 aasta jooksul Eestisse saabunud välismaalastega |
| K4X8_3 | 1414 | Kuidas Te suhtute sellesse, kui Te õpiksite või töotaksite kollektiivis, kus enamuses on ... ? uus-sisserändajatega ehk viimase 5 aasta jooksul Eestisse saabunud välismaalastega |
| K4X9_3 | 1414 | Kuidas Te suhtute sellesse, kui Teie lähedane sugulane (näiteks laps, lapselaps, õde või vend) oleks abiellus ... ? uus-sisserändajatega ehk viimase 5 aasta jooksul Eestisse saabunud välismaalastega |
| K5_4 | 1414 | Mil määral valdate järgnevaid keeli? Muu keel |
| K6X1_4 | 1414 | Kuivõrd nõustute järgmiste seisukohtadega? Teistest riikidest inimeste Eestisse elama asumine rikastab Eesti kultuurielu |
| K6X3_4 | 1053 | Kuivõrd nõustute järgmiste väidetega? Te ei tunne mingit survet oma rahvuskultuurist loobumiseks ja eestistumiseks |
| K6X3_5 | 1053 | Teie rahvuskultuuri säilimine Eestis on ohus |
| K6X6_2 | 1414 | Kuivõrd Te nõustute järgmiste väitega? Mitte-eestlaste kaasamine Eesti riigi juhtimisse on Eestile kasulik |
| K6X7 | 1414 | Kuivõrd edukas on Teie arvates seni üldiselt olnud teisest rahvusest inimeste lõimumine ja kaasamine ehk integratsioon Eesti ühiskonda? |

| | | |
|----------|------|--|
| K8X5_1 | 1414 | Kui ühes rühmas/klassis õpiksid koos erinevast rahvusest või eri emakeelt kõnelevad lapsed, siis kuidas hindaksite seda ...? lapse õpitulemuste seisukohalt |
| K8X5_3 | 1414 | eesti keele ja kultuuri säilitamise seisukohalt |
| K8X5_4 | 1414 | teiste rahvuste keele ja kultuuri säilitamise seisukohalt |
| K8X5_5 | 1414 | ühiskonna sidususe seisukohalt |
| K8X8_7 | 1414 | Järgnevalt on toodud mõned väited, mida inimesed kasutavad, kui räägivad muu emakeelega laste õppimisest eesti keeles. Kuivõrd Te nõustute järgmiste väidetega? Muu emakeelega lapse õppimine eesti keeles ... Takistab lapsel ainealaste teadmiste omandamist |
| K8X8_8 | 1414 | Tekitab õpilastel psühholoogilist stressi, pingeid |
| K8X8_10 | 1414 | Ohustab lapse kultuurilise identiteedi säilimist ja arenemist |
| K9X5_1 | 205 | Millised alljärgnevatest põhjustest võivad mõjutada Teie soovi mitte-omandada Eesti kodakondsust? Eesti kodakondsus ei ole mainekas |
| K9X5_3 | 205 | Kodakondsuseksami tegemine on liiga keeruline |
| K9X5_5 | 205 | Eesti ei luba topelt-kodakondsust, aga teil on juba teise riigi kodakondsus |
| K9X5_6 | 205 | Kodakondsuse puudumine ei takista Eestis elamist |
| K9X5_7 | 205 | Ei ole piisavalt infot Eesti kodakondsuse saamise võimaluste kohta |
| K9X5_8 | 205 | Ei ole infot selle kohta, mis õigusi ja kohustusi Eesti kodakondsus kaasa toob |
| K9X5_9 | 205 | Eesti kodanike kohustuslik kaitseväeteenistus |
| K9X7_13 | 1414 | Millist erakonda peate enda vaadetele ja huvidele kõige lähedasemaks? Ei oska / soovi vastata |
| K9X10_5 | 1414 | Kuivõrd Te usaldate järgnevaid institutsioone ja rühmi Eestis? Kohtusüsteem |
| K9X10_10 | 1414 | Kodanikuühiskond |
| T18 | 1414 | Millisesse vahemikku jääb Teie leibkonna eelmise kuu kogusissetulek, pärast maksude maha arvestamist? |

LISA 4. Hii-ruut-testi tulemused: Sugu

| | Küsimus | Test | Kaalutud rahvusega | | | | Kaalutud üldkaaluga | | | |
|----------|------------------------|------------------------|--------------------|---------|---------|----------|---------------------|---------|---------|----------|
| | | | N | Väärtus | df | p | N | Väärtus | df | p |
| Sugu | K3X1_4 | Pearsoni hii-ruut-test | 1414 | 5,12 | 1 | 0,024* | 1413 | 4,74 | 1 | 0,029* |
| | K3X1_5 | Pearsoni hii-ruut-test | 1413 | 0,51 | 1 | 0,475 | 1414 | 0,71 | 1 | 0,400 |
| | K3X1_7 | Pearsoni hii-ruut-test | 1414 | 3,98 | 1 | 0,046* | 1414 | 2,83 | 1 | 0,093 |
| | K3X1_9 | Pearsoni hii-ruut-test | 1414 | 9,71 | 1 | 0,002** | 1414 | 10,90 | 1 | 0,001* |
| | K3X1_10 | Pearsoni hii-ruut-test | 1414 | 1,18 | 1 | 0,278 | 1414 | 0,92 | 1 | 0,338 |
| | K3X1_11 | Pearsoni hii-ruut-test | 1413 | 5,36 | 1 | 0,021* | 1414 | 5,71 | 1 | 0,017* |
| | K3X3 | Pearsoni hii-ruut-test | 1414 | 2,66 | 1 | 0,103 | 1414 | 2,37 | 1 | 0,124 |
| | K3X4_4 | Pearsoni hii-ruut-test | 1414 | 0,05 | 1 | 0,818 | 1413 | 0,09 | 1 | 0,769 |
| | K3X4_14 | Pearsoni hii-ruut-test | 1415 | 3,76 | 1 | 0,052 | 1414 | 2,62 | 1 | 0,105 |
| | K3X4_15 | Pearsoni hii-ruut-test | 1415 | 1,17 | 1 | 0,279 | 1414 | 0,98 | 1 | 0,322 |
| | K3X5_1 | Pearsoni hii-ruut-test | 878 | 4,35 | 1 | 0,037* | 871 | 5,85 | 1 | 0,016* |
| | K3X5_2 | Pearsoni hii-ruut-test | 879 | 4,57 | 1 | 0,033* | 870 | 4,59 | 1 | 0,032* |
| | K3X5_3 | Pearsoni hii-ruut-test | 878 | 9,72 | 1 | 0,002** | 872 | 8,84 | 1 | 0,003** |
| | K3X5_4 | Pearsoni hii-ruut-test | 878 | 4,41 | 1 | 0,036* | 872 | 3,71 | 1 | 0,054 |
| | K4X7_3 | Pearsoni hii-ruut-test | 1415 | 7,67 | 1 | 0,006** | 1415 | 7,13 | 1 | 0,008** |
| | K4X8_3 | Pearsoni hii-ruut-test | 1415 | 8,72 | 1 | 0,003** | 1414 | 9,99 | 1 | 0,002** |
| | K4X9_3 | Pearsoni hii-ruut-test | 1413 | 1,92 | 1 | 0,166 | 1414 | 2,21 | 1 | 0,137 |
| | K5_4 | Pearsoni hii-ruut-test | 1414 | 0,24 | 1 | 0,623 | 1413 | 0,00 | 1 | 0,950 |
| | K6X1_4 | Pearsoni hii-ruut-test | 1414 | 0,00 | 1 | 0,978 | 1414 | 0,01 | 1 | 0,926 |
| | K6X3_4 | Pearsoni hii-ruut-test | 927 | 0,07 | 1 | 0,785 | 921 | 0,25 | 1 | 0,620 |
| | K6X3_5 | Pearsoni hii-ruut-test | 927 | 0,02 | 1 | 0,875 | 921 | 0,08 | 1 | 0,776 |
| | K6X6_2 | Pearsoni hii-ruut-test | 1413 | 0,23 | 1 | 0,631 | 1413 | 0,51 | 1 | 0,475 |
| | K6X7 | Pearsoni hii-ruut-test | 1414 | 15,33 | 1 | 0,000*** | 1413 | 16,81 | 1 | 0,000*** |
| | K8X5_1 | Pearsoni hii-ruut-test | 1413 | 0,03 | 1 | 0,859 | 1413 | 0,04 | 1 | 0,840 |
| | K8X5_3 | Pearsoni hii-ruut-test | 1415 | 0,27 | 1 | 0,600 | 1414 | 0,28 | 1 | 0,593 |
| | K8X5_4 | Pearsoni hii-ruut-test | 1414 | 0,64 | 1 | 0,424 | 1414 | 0,66 | 1 | 0,415 |
| | K8X5_5 | Pearsoni hii-ruut-test | 1415 | 0,12 | 1 | 0,730 | 1415 | 0,27 | 1 | 0,601 |
| | K8X8_7 | Pearsoni hii-ruut-test | 1413 | 0,04 | 1 | 0,833 | 1414 | 0,25 | 1 | 0,614 |
| | K8X8_8 | Pearsoni hii-ruut-test | 1415 | 0,43 | 1 | 0,513 | 1413 | 0,60 | 1 | 0,440 |
| | K8X8_10 | Pearsoni hii-ruut-test | 1414 | 0,11 | 1 | 0,739 | 1414 | 0,37 | 1 | 0,540 |
| | K9X5_1 | Pearsoni hii-ruut-test | 138 | 0,88 | 1 | 0,347 | 137 | 0,94 | 1 | 0,332 |
| | K9X5_3 | Pearsoni hii-ruut-test | 137 | 0,63 | 1 | 0,428 | 136 | 0,41 | 1 | 0,523 |
| | K9X5_5 | Pearsoni hii-ruut-test | 137 | 0,38 | 1 | 0,537 | 136 | 0,51 | 1 | 0,476 |
| | K9X5_6 | Pearsoni hii-ruut-test | 137 | 0,02 | 1 | 0,901 | 136 | 0,05 | 1 | 0,831 |
| K9X5_7 | Pearsoni hii-ruut-test | 138 | 0,49 | 1 | 0,483 | 136 | 0,26 | 1 | 0,613 | |
| K9X5_8 | Pearsoni hii-ruut-test | 137 | 0,97 | 1 | 0,325 | 135 | 0,63 | 1 | 0,429 | |
| K9X5_9 | Pearsoni hii-ruut-test | 136 | 0,37 | 1 | 0,543 | 136 | 0,28 | 1 | 0,597 | |
| K9X7_13 | Pearsoni hii-ruut-test | 1414 | 3,86 | 1 | 0,049* | 1415 | 4,09 | 1 | 0,043* | |
| K9X10_5 | Pearsoni hii-ruut-test | 1414 | 3,31 | 1 | 0,069 | 1414 | 3,17 | 1 | 0,075 | |
| K9X10_10 | Pearsoni hii-ruut-test | 1414 | 7,77 | 1 | 0,005** | 1414 | 6,92 | 1 | 0,009** | |
| T18 | Pearsoni hii-ruut-test | 1414 | 0,27 | 1 | 0,604 | 1414 | 0,58 | 1 | 0,447 | |

*** p < 0,001; ** p < 0,01; * p < 0,05
df – vabadusastmete arv

LISA 5. Hii-ruut-testi ja Fisheri täpse testi tulemused: Rahvus

| | Küsimus | Test | Kaalutud rahvusega | | | | Kaalutud üldkaaluga | | | |
|--------|------------------------|------------------------|--------------------|---------|-------|----------|---------------------|---------|-------|----------|
| | | | N | Väärtus | df | p | N | Väärtus | df | p |
| Rahvus | K3X1_4 | Pearsoni hii-ruut-test | 1414 | 8,18 | 1 | 0,004** | 1413 | 8,98 | 1 | 0,003** |
| | K3X1_5 | Pearsoni hii-ruut-test | 1414 | 15,05 | 1 | 0,000*** | 1414 | 16,00 | 1 | 0,000*** |
| | K3X1_7 | Pearsoni hii-ruut-test | 1414 | 15,65 | 1 | 0,000*** | 1414 | 17,77 | 1 | 0,000*** |
| | K3X1_9 | Pearsoni hii-ruut-test | 1414 | 2,66 | 1 | 0,103 | 1415 | 2,67 | 1 | 0,102 |
| | K3X1_10 | Pearsoni hii-ruut-test | 1414 | 8,70 | 1 | 0,003** | 1415 | 9,30 | 1 | 0,002** |
| | K3X1_11 | Pearsoni hii-ruut-test | 1413 | 0,77 | 1 | 0,381 | 1415 | 1,17 | 1 | 0,280 |
| | K3X3 | Pearsoni hii-ruut-test | 1415 | 0,71 | 1 | 0,401 | 1414 | 0,21 | 1 | 0,648 |
| | K3X4_4 | Pearsoni hii-ruut-test | 1414 | 0,88 | 1 | 0,348 | 1413 | 1,02 | 1 | 0,313 |
| | K3X4_14 | Pearsoni hii-ruut-test | 1413 | 0,07 | 1 | 0,794 | 1414 | 0,00 | 1 | 0,967 |
| | K3X4_15 | Pearsoni hii-ruut-test | 1415 | 6,60 | 1 | 0,01* | 1415 | 6,31 | 1 | 0,012* |
| | K3X5_1 | Pearsoni hii-ruut-test | 878 | 40,29 | 1 | 0,000*** | 871 | 42,23 | 1 | 0,000*** |
| | K3X5_2 | Pearsoni hii-ruut-test | 879 | 37,95 | 1 | 0,000*** | 870 | 40,35 | 1 | 0,000*** |
| | K3X5_3 | Pearsoni hii-ruut-test | 878 | 10,11 | 1 | 0,001** | 872 | 11,14 | 1 | 0,001** |
| | K3X5_4 | Pearsoni hii-ruut-test | 877 | 8,70 | 1 | 0,003** | 871 | 9,83 | 1 | 0,002** |
| | K4X7_3 | Pearsoni hii-ruut-test | 1413 | 1,17 | 1 | 0,280 | 1413 | 1,12 | 1 | 0,289 |
| | K4X8_3 | Pearsoni hii-ruut-test | 1414 | 0,02 | 1 | 0,894 | 1415 | 0,07 | 1 | 0,798 |
| | K4X9_3 | Pearsoni hii-ruut-test | 1413 | 3,57 | 1 | 0,059 | 1414 | 3,50 | 1 | 0,061 |
| | K5_4 | Pearsoni hii-ruut-test | 1413 | 0,04 | 1 | 0,846 | 1415 | 0,02 | 1 | 0,877 |
| | K6X1_4 | Pearsoni hii-ruut-test | 1414 | 4,99 | 1 | 0,025* | 1414 | 5,44 | 1 | 0,02* |
| | K6X3_4 | Pearsoni hii-ruut-test | 927 | 67,01 | 1 | 0,000*** | 921 | 66,29 | 1 | 0,000*** |
| | K6X3_5 | Pearsoni hii-ruut-test | 927 | 10,50 | 1 | 0,001** | 921 | 11,52 | 1 | 0,001** |
| | K6X6_2 | Pearsoni hii-ruut-test | 1415 | 3,11 | 1 | 0,078 | 1413 | 3,20 | 1 | 0,073 |
| | K6X7 | Pearsoni hii-ruut-test | 1414 | 1,10 | 1 | 0,294 | 1413 | 1,07 | 1 | 0,300 |
| | K8X5_1 | Pearsoni hii-ruut-test | 1414 | 3,03 | 1 | 0,082 | 1413 | 2,55 | 1 | 0,110 |
| | K8X5_3 | Pearsoni hii-ruut-test | 1415 | 17,54 | 1 | 0,000*** | 1415 | 16,20 | 1 | 0,000*** |
| | K8X5_4 | Pearsoni hii-ruut-test | 1414 | 0,00 | 1 | 0,951 | 1413 | 0,02 | 1 | 0,881 |
| | K8X5_5 | Pearsoni hii-ruut-test | 1414 | 1,45 | 1 | 0,229 | 1415 | 0,99 | 1 | 0,321 |
| | K8X8_7 | Pearsoni hii-ruut-test | 1413 | 9,49 | 1 | 0,002** | 1414 | 8,55 | 1 | 0,003** |
| | K8X8_8 | Pearsoni hii-ruut-test | 1414 | 0,71 | 1 | 0,400 | 1414 | 0,36 | 1 | 0,547 |
| | K8X8_10 | Pearsoni hii-ruut-test | 1413 | 8,50 | 1 | 0,004** | 1415 | 7,27 | 1 | 0,007** |
| | K9X5_1 | Fisherit täpne test | 137 | | | 0,035* | 136 | | | 0,035* |
| | K9X5_3 | Fisherit täpne test | 138 | | | 0,071 | 136 | | | 0,074 |
| | K9X5_5 | Fisherit täpne test | 137 | | | 0,041* | 136 | | | 0,042* |
| | K9X5_6 | Fisherit täpne test | 137 | | | 0,029* | 136 | | | 0,029* |
| | K9X5_7 | Fisherit täpne test | 138 | | | 0,028* | 136 | | | 0,029* |
| | K9X5_8 | Fisherit täpne test | 137 | | | 0,024* | 136 | | | 0,024* |
| | K9X5_9 | Fisherit täpne test | 138 | | | 0,112 | 136 | | | 0,116 |
| | K9X7_13 | Pearsoni hii-ruut-test | 1413 | 15,44 | 1 | 0,000*** | 1413 | 14,93 | 1 | 0,000*** |
| | K9X10_5 | Pearsoni hii-ruut-test | 1414 | 23,82 | 1 | 0,000*** | 1414 | 22,59 | 1 | 0,000*** |
| | K9X10_10 | Pearsoni hii-ruut-test | 1414 | 94,52 | 1 | 0,000*** | 1413 | 86,27 | 1 | 0,000*** |
| T18 | Pearsoni hii-ruut-test | 1414 | 0,03 | 1 | 0,854 | 1414 | 0,23 | 1 | 0,634 | |

*** p < 0,001; ** p < 0,01; * p < 0,05

df – vabadusastmete arv

LISA 6. Hii-ruut-testi ja Fisheri täpse testi tulemused: Regioon

| | Küsimus | Test | Kaalutud rahvusega | | | | Kaalutud üldkaaluga | | | |
|---------|------------------------|------------------------|--------------------|---------|--------|----------|---------------------|---------|---------|----------|
| | | | N | Väärtus | df | p | N | Väärtus | df | p |
| Regioon | K3X1_4 | Pearsoni hii-ruut-test | 1416 | 27,40 | 4 | 0,000*** | 1414 | 31,07 | 4 | 0,000*** |
| | K3X1_5 | Pearsoni hii-ruut-test | 1415 | 26,65 | 4 | 0,000*** | 1415 | 28,71 | 4 | 0,000*** |
| | K3X1_7 | Pearsoni hii-ruut-test | 1414 | 30,11 | 4 | 0,000*** | 1413 | 33,96 | 4 | 0,000*** |
| | K3X1_9 | Pearsoni hii-ruut-test | 1413 | 12,34 | 4 | 0,015* | 1414 | 13,17 | 4 | 0,01* |
| | K3X1_10 | Pearsoni hii-ruut-test | 1414 | 5,74 | 4 | 0,219 | 1414 | 8,83 | 4 | 0,065 |
| | K3X1_11 | Pearsoni hii-ruut-test | 1415 | 20,62 | 4 | 0,000*** | 1415 | 27,99 | 4 | 0,000*** |
| | K3X3 | Pearsoni hii-ruut-test | 1415 | 11,01 | 4 | 0,026* | 1415 | 14,63 | 4 | 0,006** |
| | K3X4_4 | Pearsoni hii-ruut-test | 1415 | 5,57 | 4 | 0,234 | 1414 | 6,93 | 4 | 0,140 |
| | K3X4_14 | Pearsoni hii-ruut-test | 1416 | 7,72 | 4 | 0,102 | 1413 | 9,04 | 4 | 0,060 |
| | K3X4_15 | Pearsoni hii-ruut-test | 1414 | 22,34 | 4 | 0,000*** | 1414 | 20,87 | 4 | 0,000*** |
| | K3X5_1 | Pearsoni hii-ruut-test | 877 | 29,62 | 4 | 0,000*** | 871 | 28,24 | 4 | 0,000*** |
| | K3X5_2 | Pearsoni hii-ruut-test | 879 | 40,35 | 4 | 0,000*** | 872 | 45,10 | 4 | 0,000*** |
| | K3X5_3 | Pearsoni hii-ruut-test | 880 | 33,79 | 4 | 0,000*** | 872 | 40,63 | 4 | 0,000*** |
| | K3X5_4 | Pearsoni hii-ruut-test | 877 | 27,87 | 4 | 0,000*** | 871 | 36,78 | 4 | 0,000*** |
| | K4X7_3 | Pearsoni hii-ruut-test | 1413 | 16,77 | 4 | 0,002** | 1414 | 17,71 | 4 | 0,001** |
| | K4X8_3 | Pearsoni hii-ruut-test | 1415 | 14,88 | 4 | 0,005** | 1415 | 14,90 | 4 | 0,005** |
| | K4X9_3 | Pearsoni hii-ruut-test | 1414 | 4,58 | 4 | 0,333 | 1414 | 4,69 | 4 | 0,320 |
| | K5_4 | Pearsoni hii-ruut-test | 1413 | 47,35 | 4 | 0,000*** | 1415 | 49,45 | 4 | 0,000*** |
| | K6X1_4 | Pearsoni hii-ruut-test | 1414 | 4,71 | 4 | 0,319 | 1414 | 4,33 | 4 | 0,364 |
| | K6X3_4 | Pearsoni hii-ruut-test | 928 | 21,42 | 4 | 0,000*** | 921 | 20,37 | 4 | 0,000*** |
| | K6X3_5 | Pearsoni hii-ruut-test | 926 | 9,79 | 4 | 0,044* | 921 | 8,78 | 4 | 0,067 |
| | K6X6_2 | Pearsoni hii-ruut-test | 1415 | 5,08 | 4 | 0,279 | 1413 | 5,58 | 4 | 0,233 |
| | K6X7 | Pearsoni hii-ruut-test | 1414 | 6,76 | 4 | 0,149 | 1414 | 7,75 | 4 | 0,101 |
| | K8X5_1 | Pearsoni hii-ruut-test | 1413 | 0,79 | 4 | 0,940 | 1415 | 1,84 | 4 | 0,766 |
| | K8X5_3 | Pearsoni hii-ruut-test | 1414 | 10,73 | 4 | 0,03* | 1415 | 11,50 | 4 | 0,021* |
| | K8X5_4 | Pearsoni hii-ruut-test | 1413 | 4,13 | 4 | 0,389 | 1414 | 4,94 | 4 | 0,293 |
| | K8X5_5 | Pearsoni hii-ruut-test | 1413 | 4,21 | 4 | 0,378 | 1415 | 4,74 | 4 | 0,315 |
| | K8X8_7 | Pearsoni hii-ruut-test | 1414 | 7,61 | 4 | 0,107 | 1414 | 7,71 | 4 | 0,103 |
| | K8X8_8 | Pearsoni hii-ruut-test | 1414 | 2,68 | 4 | 0,613 | 1413 | 2,41 | 4 | 0,662 |
| | K8X8_10 | Pearsoni hii-ruut-test | 1414 | 1,82 | 4 | 0,769 | 1413 | 3,43 | 4 | 0,488 |
| | K9X5_1 | Fisherit täpne test | 137 | 5,43 | | 0,192 | 137 | 5,51 | | 0,193 |
| | K9X5_3 | Fisherit täpne test | 139 | 4,00 | | 0,674 | 137 | 1,84 | | 0,745 |
| | K9X5_5 | Fisherit täpne test | 136 | 5,06 | | 0,210 | 136 | 5,14 | | 0,197 |
| | K9X5_6 | Fisherit täpne test | 136 | 7,52 | | 0,080 | 135 | 7,49 | | 0,082 |
| | K9X5_7 | Fisherit täpne test | 137 | 3,36 | | 0,444 | 136 | 3,39 | | 0,457 |
| | K9X5_8 | Fisherit täpne test | 137 | 5,34 | | 0,180 | 136 | 5,38 | | 0,178 |
| | K9X5_9 | Fisherit täpne test | 138 | 5,37 | | 0,205 | 136 | 3,24 | | 0,474 |
| | K9X7_13 | Pearsoni hii-ruut-test | 1414 | 27,21 | 4 | 0,000*** | 1415 | 29,38 | 4 | 0,000*** |
| | K9X10_5 | Pearsoni hii-ruut-test | 1413 | 8,18 | 4 | 0,085 | 1413 | 8,79 | 4 | 0,066 |
| | K9X10_10 | Pearsoni hii-ruut-test | 1413 | 37,67 | 4 | 0,000*** | 1413 | 37,42 | 4 | 0,000*** |
| T18 | Pearsoni hii-ruut-test | 1414 | 12,47 | 4 | 0,014* | 1415 | 14,60 | 4 | 0,006** | |

*** p < 0,001; ** p < 0,01; * p < 0,05
df – vabadusastmete arv

LISA 7. Hii-ruut-testi ja Fisheri täpse testi tulemused: K6X8

| | Küsimus | Test | Kaalutud rahvusega | | | | Kaalutud üldkaaluga | | | |
|---------------------------------------|------------------------|------------------------|--------------------|---------|----------|----------|---------------------|---------|----------|----------|
| | | | N | Väärtus | df | p | N | Väärtus | df | p |
| K6X8: Tulevik seotud Eestiga | K3X1_4 | Pearsoni hii-ruut-test | 1377 | 10,06 | 3 | 0,018* | 1376 | 10,04 | 3 | 0,018* |
| | K3X1_5 | Pearsoni hii-ruut-test | 1377 | 2,30 | 3 | 0,513 | 1375 | 3,13 | 3 | 0,372 |
| | K3X1_7 | Pearsoni hii-ruut-test | 1377 | 0,05 | 3 | 0,997 | 1374 | 0,18 | 3 | 0,981 |
| | K3X1_9 | Pearsoni hii-ruut-test | 1377 | 0,81 | 3 | 0,847 | 1377 | 0,98 | 3 | 0,806 |
| | K3X1_10 | Pearsoni hii-ruut-test | 1376 | 3,20 | 3 | 0,362 | 1376 | 2,24 | 3 | 0,524 |
| | K3X1_11 | Pearsoni hii-ruut-test | 1376 | 1,12 | 3 | 0,771 | 1375 | 0,87 | 3 | 0,833 |
| | K3X3 | Pearsoni hii-ruut-test | 1377 | 2,48 | 3 | 0,479 | 1375 | 2,78 | 3 | 0,427 |
| | K3X4_4 | Pearsoni hii-ruut-test | 1379 | 15,81 | 3 | 0,001** | 1376 | 17,16 | 3 | 0,001** |
| | K3X4_14 | Pearsoni hii-ruut-test | 1378 | 14,16 | 3 | 0,003** | 1376 | 16,89 | 3 | 0,001** |
| | K3X4_15 | Pearsoni hii-ruut-test | 1376 | 21,33 | 3 | 0,000*** | 1375 | 25,61 | 3 | 0,000*** |
| | K3X5_1 | Pearsoni hii-ruut-test | 856 | 9,47 | 3 | 0,024* | 850 | 8,19 | 3 | 0,042* |
| | K3X5_2 | Pearsoni hii-ruut-test | 857 | 5,70 | 3 | 0,127 | 850 | 5,65 | 3 | 0,130 |
| | K3X5_3 | Pearsoni hii-ruut-test | 857 | 3,97 | 3 | 0,264 | 848 | 4,66 | 3 | 0,199 |
| | K3X5_4 | Pearsoni hii-ruut-test | 857 | 3,03 | 3 | 0,387 | 850 | 3,49 | 3 | 0,322 |
| | K4X7_3 | Pearsoni hii-ruut-test | 1377 | 12,00 | 3 | 0,007** | 1376 | 12,24 | 3 | 0,007** |
| | K4X8_3 | Pearsoni hii-ruut-test | 1376 | 20,44 | 3 | 0,000*** | 1376 | 21,30 | 3 | 0,000*** |
| | K4X9_3 | Pearsoni hii-ruut-test | 1377 | 8,10 | 3 | 0,044* | 1375 | 7,03 | 3 | 0,071 |
| | K5_4 | Pearsoni hii-ruut-test | 1377 | 11,07 | 3 | 0,011* | 1374 | 11,89 | 3 | 0,008** |
| | K6X1_4 | Pearsoni hii-ruut-test | 1376 | 0,51 | 3 | 0,917 | 1376 | 0,91 | 3 | 0,822 |
| | K6X3_4 | Pearsoni hii-ruut-test | 895 | 5,69 | 3 | 0,127 | 888 | 7,36 | 3 | 0,061 |
| | K6X3_5 | Pearsoni hii-ruut-test | 894 | 1,55 | 3 | 0,672 | 887 | 2,11 | 3 | 0,550 |
| | K6X6_2 | Pearsoni hii-ruut-test | 1377 | 5,15 | 3 | 0,161 | 1376 | 7,40 | 3 | 0,060 |
| | K6X7 | Pearsoni hii-ruut-test | 1377 | 8,22 | 3 | 0,042* | 1375 | 9,19 | 3 | 0,027* |
| | K8X5_1 | Pearsoni hii-ruut-test | 1376 | 3,10 | 3 | 0,376 | 1376 | 2,10 | 3 | 0,552 |
| | K8X5_3 | Pearsoni hii-ruut-test | 1377 | 6,28 | 3 | 0,099 | 1375 | 6,85 | 3 | 0,077 |
| | K8X5_4 | Pearsoni hii-ruut-test | 1377 | 0,84 | 3 | 0,840 | 1375 | 0,69 | 3 | 0,876 |
| | K8X5_5 | Pearsoni hii-ruut-test | 1377 | 0,82 | 3 | 0,845 | 1376 | 0,62 | 3 | 0,891 |
| | K8X8_7 | Pearsoni hii-ruut-test | 1376 | 1,28 | 3 | 0,733 | 1376 | 1,56 | 3 | 0,668 |
| | K8X8_8 | Pearsoni hii-ruut-test | 1376 | 3,82 | 3 | 0,282 | 1376 | 4,53 | 3 | 0,210 |
| | K8X8_10 | Pearsoni hii-ruut-test | 1377 | 0,36 | 3 | 0,948 | 1374 | 0,75 | 3 | 0,863 |
| | K9X5_1 | Fisherit täpne test | 123 | 1,49 | | 0,672 | 122 | 2,07 | | 0,513 |
| | K9X5_3 | Fisherit täpne test | 122 | 0,50 | | 0,956 | 121 | 0,59 | | 0,918 |
| | K9X5_5 | Fisherit täpne test | 123 | 0,44 | | 0,943 | 121 | 0,47 | | 0,946 |
| | K9X5_6 | Fisherit täpne test | 123 | 1,43 | | 0,737 | 122 | 1,37 | | 0,707 |
| | K9X5_7 | Fisherit täpne test | 122 | 0,81 | | 0,833 | 121 | 0,94 | | 0,833 |
| K9X5_8 | Fisherit täpne test | 122 | 1,06 | | 0,811 | 122 | 1,21 | | 0,769 | |
| K9X5_9 | Fisherit täpne test | 122 | 1,22 | | 0,742 | 121 | 1,25 | | 0,730 | |
| K9X7_13 | Pearsoni hii-ruut-test | 1378 | 3,67 | 3 | 0,299 | 1375 | 5,30 | 3 | 0,151 | |
| K9X10_5 | Pearsoni hii-ruut-test | 1377 | 8,82 | 3 | 0,032* | 1376 | 7,58 | 3 | 0,055 | |
| K9X10_10 | Pearsoni hii-ruut-test | 1377 | 3,40 | 3 | 0,334 | 1375 | 3,96 | 3 | 0,266 | |
| T18 | Pearsoni hii-ruut-test | 1377 | 32,81 | 3 | 0,000*** | 1377 | 36,38 | 3 | 0,000*** | |

*** p < 0,001; ** p < 0,01; * p < 0,05

df – vabadusastmete arv

LISA 8. Hii-ruut-testi ja Fisheri täpse testi tulemused: K6X5_4

| | Küsimus | Test | Kaalutud rahvusega | | | | Kaalutud üldkaaluga | | | |
|---|------------------------|------------------------|--------------------|---------|-------|----------|---------------------|---------|-------|----------|
| | | | N | Väärtus | df | p | N | Väärtus | df | p |
| K6X5_4: Tunneb end osana Eesti ühiskonnast | K3X1_4 | Pearsoni hii-ruut-test | 1379 | 2,41 | 3 | 0,491 | 1376 | 1,80 | 3 | 0,615 |
| | K3X1_5 | Pearsoni hii-ruut-test | 1377 | 4,98 | 3 | 0,173 | 1376 | 5,22 | 3 | 0,157 |
| | K3X1_7 | Pearsoni hii-ruut-test | 1379 | 6,77 | 3 | 0,080 | 1376 | 7,18 | 3 | 0,067 |
| | K3X1_9 | Pearsoni hii-ruut-test | 1379 | 1,21 | 3 | 0,750 | 1376 | 1,03 | 3 | 0,794 |
| | K3X1_10 | Pearsoni hii-ruut-test | 1378 | 15,60 | 3 | 0,001** | 1375 | 14,92 | 3 | 0,002** |
| | K3X1_11 | Pearsoni hii-ruut-test | 1378 | 2,42 | 3 | 0,489 | 1376 | 1,73 | 3 | 0,630 |
| | K3X3 | Pearsoni hii-ruut-test | 1379 | 0,28 | 3 | 0,964 | 1376 | 0,30 | 3 | 0,960 |
| | K3X4_4 | Pearsoni hii-ruut-test | 1378 | 6,54 | 3 | 0,088 | 1377 | 6,60 | 3 | 0,086 |
| | K3X4_14 | Pearsoni hii-ruut-test | 1377 | 0,84 | 3 | 0,840 | 1376 | 0,22 | 3 | 0,975 |
| | K3X4_15 | Pearsoni hii-ruut-test | 1380 | 3,81 | 3 | 0,283 | 1376 | 3,62 | 3 | 0,306 |
| | K3X5_1 | Pearsoni hii-ruut-test | 856 | 25,00 | 3 | 0,000*** | 847 | 22,27 | 3 | 0,000*** |
| | K3X5_2 | Pearsoni hii-ruut-test | 857 | 27,46 | 3 | 0,000*** | 846 | 26,32 | 3 | 0,000*** |
| | K3X5_3 | Pearsoni hii-ruut-test | 855 | 28,79 | 3 | 0,000*** | 846 | 28,19 | 3 | 0,000*** |
| | K3X5_4 | Pearsoni hii-ruut-test | 856 | 22,74 | 3 | 0,000*** | 845 | 21,71 | 3 | 0,000*** |
| | K4X7_3 | Pearsoni hii-ruut-test | 1379 | 1,81 | 3 | 0,613 | 1376 | 2,18 | 3 | 0,536 |
| | K4X8_3 | Pearsoni hii-ruut-test | 1380 | 2,88 | 3 | 0,410 | 1376 | 3,62 | 3 | 0,305 |
| | K4X9_3 | Pearsoni hii-ruut-test | 1379 | 2,48 | 3 | 0,480 | 1377 | 4,59 | 3 | 0,204 |
| | K5_4 | Pearsoni hii-ruut-test | 1378 | 2,89 | 3 | 0,409 | 1377 | 4,57 | 3 | 0,206 |
| | K6X1_4 | Pearsoni hii-ruut-test | 1380 | 3,98 | 3 | 0,264 | 1377 | 4,10 | 3 | 0,251 |
| | K6X3_4 | Pearsoni hii-ruut-test | 895 | 8,05 | 3 | 0,045* | 887 | 6,63 | 3 | 0,085 |
| | K6X3_5 | Pearsoni hii-ruut-test | 897 | 0,69 | 3 | 0,875 | 889 | 0,71 | 3 | 0,872 |
| | K6X6_2 | Pearsoni hii-ruut-test | 1379 | 3,91 | 3 | 0,272 | 1377 | 3,08 | 3 | 0,379 |
| | K6X7 | Pearsoni hii-ruut-test | 1379 | 4,59 | 3 | 0,205 | 1377 | 3,90 | 3 | 0,273 |
| | K8X5_1 | Pearsoni hii-ruut-test | 1377 | 3,82 | 3 | 0,281 | 1376 | 4,38 | 3 | 0,223 |
| | K8X5_3 | Pearsoni hii-ruut-test | 1380 | 10,31 | 3 | 0,016* | 1377 | 11,71 | 3 | 0,008** |
| | K8X5_4 | Pearsoni hii-ruut-test | 1378 | 4,06 | 3 | 0,255 | 1376 | 5,30 | 3 | 0,151 |
| | K8X5_5 | Pearsoni hii-ruut-test | 1379 | 6,69 | 3 | 0,082 | 1377 | 7,77 | 3 | 0,051 |
| | K8X8_7 | Pearsoni hii-ruut-test | 1379 | 1,83 | 3 | 0,609 | 1377 | 1,90 | 3 | 0,593 |
| | K8X8_8 | Pearsoni hii-ruut-test | 1380 | 6,13 | 3 | 0,106 | 1377 | 5,48 | 3 | 0,140 |
| | K8X8_10 | Pearsoni hii-ruut-test | 1379 | 2,51 | 3 | 0,473 | 1376 | 2,30 | 3 | 0,513 |
| | K9X5_1 | Fisherit täpne test | 129 | 0,81 | | 0,959 | 126 | 1,15 | | 0,812 |
| | K9X5_3 | Fisherit täpne test | 130 | 0,80 | | 0,911 | 126 | 1,19 | | 0,818 |
| | K9X5_5 | Fisherit täpne test | 127 | 0,77 | | 0,865 | 126 | 0,72 | | 0,916 |
| | K9X5_6 | Fisherit täpne test | 129 | 1,20 | | 0,744 | 126 | 1,26 | | 0,735 |
| | K9X5_7 | Fisherit täpne test | 128 | 1,11 | | 0,801 | 127 | 0,58 | | 1,000 |
| | K9X5_8 | Fisherit täpne test | 127 | 0,50 | | 1,000 | 127 | 0,47 | | 1,000 |
| | K9X5_9 | Fisherit täpne test | 128 | 1,30 | | 0,783 | 127 | 1,39 | | 0,765 |
| | K9X7_13 | Pearsoni hii-ruut-test | 1379 | 7,01 | 3 | 0,072 | 1375 | 7,34 | 3 | 0,062 |
| | K9X10_5 | Pearsoni hii-ruut-test | 1378 | 0,58 | 3 | 0,901 | 1377 | 0,06 | 3 | 0,996 |
| | K9X10_10 | Pearsoni hii-ruut-test | 1379 | 36,48 | 3 | 0,000*** | 1377 | 32,87 | 3 | 0,000*** |
| T18 | Pearsoni hii-ruut-test | 1380 | 3,88 | 3 | 0,275 | 1377 | 3,69 | 3 | 0,298 | |

*** p < 0,001; ** p < 0,01; * p < 0,05
df – vabadusastmete arv

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Liina Eamets

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose Integratsiooni monitooringu eri aastate andmete võrreldavus - uuringumetoodika, ankeetküsimuste ja andmete esinduslikkuse analüüs,

mille juhendaja on Indrek Soidla, kaasjuhendaja Maarja Vollmer

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Liina Eamets
18.08.2022