

Tartu Ülikool

Sotsiaalteaduste valdkond

Ühiskonnateaduste instituut

Sotsioloogia õppekava

Liisi Lillipuu

Andmete kvaliteet veebipaneelis tõenäosusliku valimiga veebipaneeli CRONOS näitel

Magistritöö

Juhendaja: MA, I. Soidla

Juhendaja allkiri:_____

Tartu 2019

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

Liisi Lillipuu, 24.05.2019

Abstract

Data Quality in Probability-Based Online Panel CRONOS

The purpose of the present master thesis is to analyse what are the main factors influencing data quality in CRONOS, concentrating on data from Estonian respondents. CRONOS is the first pilot of cross-national probability-based web panel. Data collection took place in December 2016 until February 2018 and was conducted according to the same harmonised methodology in all participating countries including Estonia. As the result of higher Internet penetration and computer literacy, online panel has been considered as more efficient data collection method than traditional methods mostly in market research, but also more and more in probability-based whole population surveys. Therefore, the methodological research is required as a prerequisite for wider use of online panels. Author analyses main factors influencing data quality in the frame of common data quality indicators such as unit and item nonresponse, survey break-off, completion time, non-differentiation and survey evaluation. Berzelak, Weber and Revilla (2018) have analysed panellist behaviour in the welcome wave of the CRONOS panel and their paper has been considered as baseline for the master thesis.

As analysis show unit nonresponse may have influenced the overall data quality in CRONOS since unit nonresponse appears to be higher in groups of younger, male and individuals who have general secondary education or lower. According to Estonian panellist responding behaviour, survey break-off is also predicted by lower education and younger age. Item nonresponse is more likely for respondents with general secondary education or lower, respondents aged 31-45 and over 60 years and individuals not so interested in politics, which indicates lower individual capacity and lower interest in topics.

In addition, analysis of non-differentiation, completion time and survey evaluation show that younger respondents are quicker in answering, thus are less likely to provide non-differentiated responses. Although younger respondents evaluated the effort to provide accurate answers higher than older respondents, older respondents decided more often not to answer questions during the survey.

Additionally, experimental manipulations where respondents received accuracy emphasis and commitment request and negative information about the consequences of careless responding

have shown an increasing effect of engagement in answering. Furthermore, it was found that respondents who received two or more reminders before answering have shown a moderate indication of lower data quality in CRONOS, especially they appear to have a remarkably higher risk of survey break-off.

Keywords: online panel, data quality, nonresponse, survey break-off, measurement error, non-differentiation, completion time, survey evaluation

SISUKORD

Abstract	3
Sissejuhatus	8
1 Mõistete määratlemine ja seotud uurimistulemuste ülevaade	10
1.1 Veebipaneeli definitsioon ja liigitus	10
1.2 Paraandmed veebiuuringus	12
1.3 Uuringu koguvea hindamine	12
1.3.1 Vaatlusest sõltumatu viga	14
1.3.1.1 Alakaetus	14
1.3.1.2 Kutsele mittevastamine	16
1.3.1.3 Vastamise katkestamine	20
1.3.1.4 Küsimusele mittevastamine	22
1.3.2 Mõõtmisviga	23
1.3.2.1 Joonvastamine	23
1.3.2.2 Vastamiseks kuluv aeg	25
1.3.2.3 Hinnang uuringus osalemise kogemusele	26
1.4 Uurimuse eesmärk ja uurimisküsimused	27
1.5 Varasemad tõenäosuslikul valimil põhinevad veebipaneelid	27
2 Metoodika	29
2.1 Andmed	29
2.2 Valimi kirjeldus	30
2.3 Tunnused	32
2.4 Analüüsimeetodid	35
3 Tulemused	38

3.1	Ülevaade andmekvaliteedist	38
3.1.1	Interneti mittekasutajad	38
3.1.2	Kutsele mittevastamine	39
3.1.3	Vastamise katkestamine	41
3.1.4	Küsimusele mittevastamine.....	44
3.1.5	Joonvastamine	47
3.1.6	Vastamisele kuluv aeg.....	50
3.1.7	Hinnang uuringus osalemise kogemusele	54
3.2	Andmekvaliteeti mõjutavad tegurid: prognoosimudelid	56
3.2.1	Kutsele mittevastamine	56
3.2.2	Vastamise katkestamine	59
3.2.3	Küsimusele mittevastamine.....	63
3.2.4	Joonvastamine	65
3.2.5	Vastamisele kuluv aeg.....	69
3.2.6	Hinnang uuringus osalemise kogemusele	71
4	Arutelu.....	77
4.1	Mittevastamise viga CRONOSes	77
4.1.1	Interneti mittekasutajad	77
4.1.2	Kutsele mittevastamine	78
4.1.3	Vastamise katkestamine	79
4.1.4	Küsimusele mittevastamine.....	81
4.1.5	Hinnang CRONOSe mittevastamise vea ulatusele	81
4.2	Mõõtmisviga CRONOSes	82
4.2.1	Joonvastamine	82

4.2.2	Vastamisele kuluv aeg.....	85
4.2.3	Hinnang uuringus osalemise kogemusele	86
4.2.4	Hinnang CRONOSe mõõtmisvea ulatusele	87
	Kokkuvõte	89
	Kasutatud allikad.....	91
	Lisad.....	101

Sissejuhatus

Interneti ja digitaalse kirjaoskuse levik maailmas võimaldavad veebipõhise andmekogumisviisi populaarsuse kasvu nii mittetõenäosuslike kui ka tõenäosuslike valimitega uuringutes. Veebipaneelid, mis moodustavad ühe osa veebiuuringutest, on omased peamiselt mittetõenäosuslikel valimitel põhinevatele turu-uuringutele. Tõenäosuslikul valimil põhinevad esinduslikud veebipaneelid on suhteliselt uus ja vähelevinud nähtus ning seetõttu alles metodoloogia väljatöötamise faasis. Käesolev magistritöö baseerub esimese rahvusvahelise tõenäosuslikul valimil põhineva ühtlustatud metodoloogiaga veebipaneeli pilootprojekti CRONOS andmetele, keskendudes Eesti andmetele, mida koguti ajavahemikul detsember 2016 – veebruar 2018. CRONOS viidi läbi eesmärgiga hinnata võrdleva tõenäosuslikul valimil põhineva rahvusvahelise veebipaneeli teostatavust ning välja töötada meetodika uute ja tõhusate veebipõhiste uuringute infrastruktuuride loomiseks Euroopas (*A survey future...*, 2019). CRONOSe andmed on kättesaadavad ESS koduleheküljel (CRONOS, i.a).

Käesolev töö baseerub Berzelaki, Weberi ja Revilla (2018) CRONOSe veebipaneeli sissejuhatava laine paneeliliikmete vastamiskäitumise analüüsil, milles autorid esitlesid võimalikku analüütilist lähenemist paneeli järgnevate lainete vastamiskäitumise uurimiseks. Antud analüüs oli sissejuhatav CRONOSe andmekvaliteedi uurimus ning seni veel täiendavaid analüüse avaldatud ei ole. Veebipaneelides nähakse ressursisäästlikumat alternatiivi traditsioonilistele andmekogumismeetoditele. Eesmärgiga hinnata sarnaste tõenäosuslikel valimitel põhinevate veebipaneelide rakendatavust, on oluline välja selgitada, millist mõju avaldavad erinevad tegurid veebipaneeli andmekvaliteedile.

Käesoleva magistritöö eesmärgiks on välja selgitada erinevate tegurite mõju CRONOSe andmekvaliteedile. Töös käsitletakse peamiselt sotsiaal-demograafiliste tunnuste, vastamiseks kasutatud seadme (nutitelefon, tahvelarvuti või arvuti), edastatud meeldetuletuste ja eksperimendi raames vastajatele kuvatud täpsete vastuste olulisust rõhutavate teavituste mõju andmekvaliteedile. Nimetatud tegurite mõju andmekvaliteedile on töös uuritud järgnevate andmekvaliteedi indikaatorite abil: uuringukutsele mittevastamine, vastamise katkestamine, küsimusele mittevastamine, joonvastamine, vastamiseks kuluv aeg ning hinnang vastamiskogemusele.

Töö teoreetilises osas on esitatud ülevaate veebipaneeli olemusest, võimalikest uuringus tekkivatest vigadest, levinud andmekvaliteedi mõõdikutest ja andmekvaliteediga seotud uurimistulemustest. Töö empiirilises osas analüüsitakse andmekvaliteeti mõjutavaid tegureid. Uurimistulemused CRONOSe andmekvaliteedi ja seda mõjutavate tegurite kohta võimaldavad hinnata veebipaneeli mittevastamisest tuleneva vea ja mõõtmisvea ulatust. Andmekvaliteeti mõjutavate tegurite väljaselgitamiseks on töös kasutatud kirjeldavat statistikat, mitteparameetrilisi teste, lineaarset, binaarset logistilist ning Coxi regressiooni.

1 Mõistete määratlemine ja seotud uurimistulemuste ülevaade

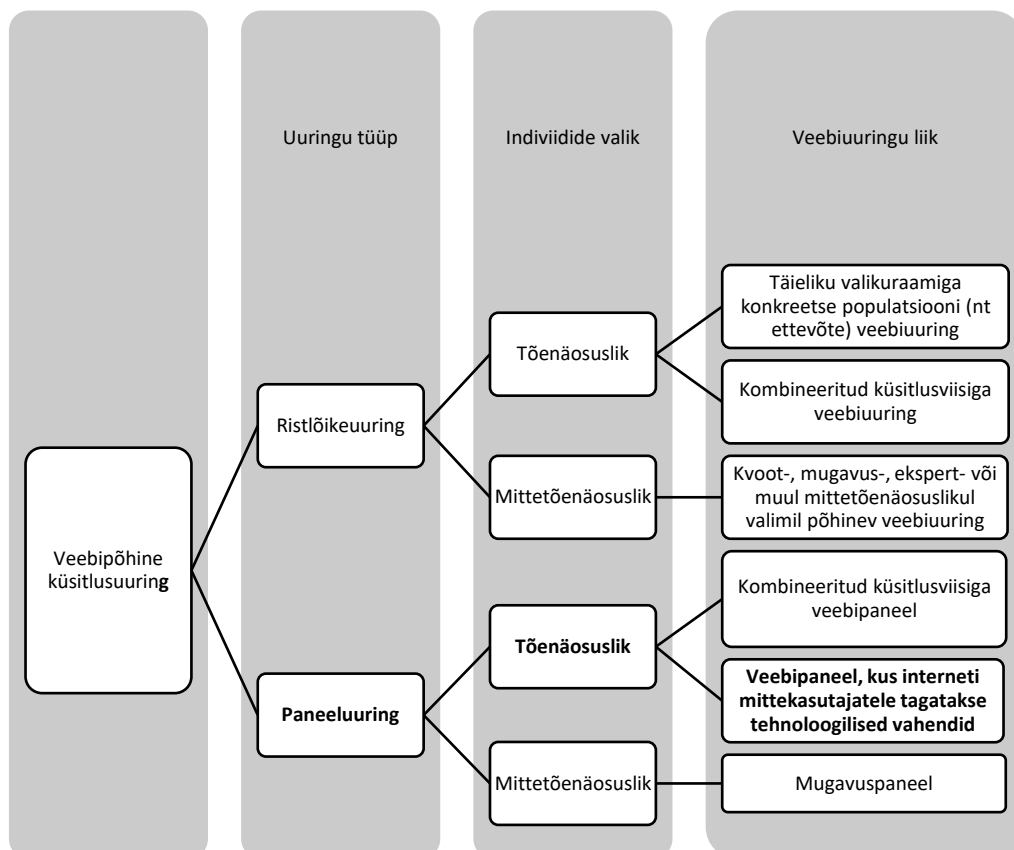
1.1 Veebipaneeli definitsioon ja liigitus

Uuringuid, mille andmekogumise etapis kasutatakse veebipõhist andmekogumisviisi, milles vastajad täidavad uuringu küsimustikku veebis ja küsimustikule ligipääsuks jagatakse respondendiga veebilinki, võib üldistatult nimetada veebiuuringuteks (Bethlehem ja Biffignandi, 2011: 32). Ühe osa veebiuuringutest moodustavad veebipaneelid. Veebipaneel on veebiuuring, milles samu objekte (indiviidid, leibkonnad, ettevõtted jne) intervjueritakse korduvalt erinevatel ajahetkedel (paneeli lainetel) (Bethlehem ja Biffignandi, 2011: 419, 422). Veebipaneeli andmed on oma olemuselt longituudandmed (põhinevad kordusmõõtmistel). Longituuduuringus võib laineti korduvate küsimuste osakaal (olenevalt uuringu iseloomust) olla üsna suur, samas kui veebipaneelis ei ole see enamasti eesmärk (Hansen, 2008: 330; Toepoel, 2016: 69). Veebipaneelid, mille igal lainel uuritakse erinevaid teemasid, kuuluvad ristlõikeuuringute hulka. Samas erineb veebipaneel traditsioonilisest ristlõikeuuringust peamiselt seetõttu, et ristlõikeuuringu kordumisel moodustatakse uus valim, kuid veebipaneel säilitab sama valimit (Bethlehem ja Biffignandi, 2011: 419).

Valimi moodustamise põhimõtte alusel jaotuvad veebipaneelid kahte rühma: tõenäosuslikel ja mittetõenäosuslikel valimitel põhinevateks veebipaneelideks (Couper, 2000; Toepoel, 2016: 70-74; Tourangeau, Conrad ja Couper, 2013: 11-35). Mittetõenäosuslikul valimil (tihti sisuliselt mugavusvalimil) põhinevad uuringud moodustavad suure turuosa teostatavatest veebiuuringutest (Toepoel, 2016: 70-74). Mugavusvalimi rakendamise eelisteks on odavus ja rakendatavuse lihtsus. Mittetõenäosuslikul valimil põhinevaid veebipaneele võib nimetada mugavuspaneelideks, mille all võib mõista nii paneele, kus liikmeks astumisel puuduvad igasugused piirangud kui ka paneele, mis on piiratud ligipääsuga, kuid põhinevad nt kvootvalimil (st mittetõenäosuslikul valimil). Üheks levinumaks mittetõenäosusliku valimiga veebipaneeliks on nn ligipääsupaneel (*access panel* või ka *opt-in panel*), mida kasutatakse eelkõige turu-uuringutes. Ligipääsupaneeliks loetakse veebipaneeli, millega on indiviidid liitunud vabatahtlikult (Callegaro jt, 2015: 67, ISO 26362, 2009). Vabatahtliku liitumisega ligipääsupaneelid hõlmavad erinevaid teemasid, kuid on üles ehitatud samal andmebaasil rajanevale valimile. Eestis võib selliste paneelide näidetena esile tuua Turu-uuringute AS loodud veebipaneeli (<https://veebiuuring.ee/>) ja Norstat AS hallatavat keskkonda Panel.ee.

Tõenäosusliku valimiga veebipaneelide hulka kuuluvad: paneelid, kus interneti mittekasutajatele võimaldatakse uuringus osalemiseks tehnoloogiline varustus (tehnoloogiline vahend, nt tahvelarvuti koos ligipääsuga internetti); kombineeritud küsitlusmeetodiga veebipaneelid, kus interneti mittekasutajatele võimaldatakse uuringus osaleda muul viisil, nt telefoni teel (CATI) (Couper, 2000; Toepoel, 2016: 73). Veebipaneeli, milles tagatakse interneti mittekasutajatele tehnoloogiline varustus uuringus osalemiseks, iseloomustab vähene alakaetus, kuid madal vastamismäär (Couper, 2000). Kombineeritud meetod võimaldab kaetust parandada ja samaaegselt tõsta vastamismäära, kuid lisandub meetodite erinevusest tulenev mõõtmisviga (Callegaro jt 2015: 368 – 374; Couper, 2000; Toepoel, 2016: 94-100).

Veebiuuringute liikide kokkuvõtlik ülevaade on esitatud joonisel 1. Käesoleva töö empiiriliseks aluseks on tõenäosusliku valimiga veebipaneel CRONOS, mille raames võimaldati interneti mittekasutajatele uuringus osalemiseks tehnoloogiline varustus. Tõenäosuslik valim CRONOSes võimaldab laiendada uurimistulemusi üldkogumile.



Joonis 1. Veebiuuringute liigitus uuringu tüübi ja valimi moodustamise põhimõtte järgi

1.2 Paraandmed veebiuuringus

Käesolev töö on üles ehitatud osaliselt CRONOSe veebipaneelis kogutud paraandmetele. Paraandmetena käsitletakse andmeid, mida luuakse veebiuuringu toimumise käigus (Bethlehem ja Biffignandi, 2011: 217). Callegaro jt (2015: 206 - 211) jagavad paraandmeid otsesteks ja kaudseteks, millest esimene hõlmab automaatselt kogutud andmeid (kontaktiloomise, kasutatava seadme ja küsimustikul navigeerimise info) ja teine paraandmeid, mida on kogutud lisavarustuse abil (nt silmade või ajutegevuse jälgimine). Kontaktiloomise andmete hulka kuuluvad andmed kontaktiloomise protsessi, kutsete kättesaamise, küsimustiku veebilehele ligipääsu kohta ja viimase vastatud küsimuse kohta. Kasutatava seadme andmed annavad infot kasutatava seadme tüübi, ekraani suuruse, brauseri jpm kohta. Küsimustikul navigeerimise andmed sisaldavad infot küsimustele vastamiseks kulutatud aja, hiire liikumise, saadud veateadete ja vastuste muutumise kohta. Veebipaneelid loovad lisaks ühekordse veebiuuringu käigus loodavatele paraandmetele veel lisateavet kõikide lainete soorituse kohta, nt viimati osaletud laine ja vastatud teemade kohta (Callegaro, 2013).

Paraandmeid kasutatakse eelkõige küsimustike testimiseks ja vastamiskäitumise analüüsiks eesmärgiga parandada andmete kvaliteeti tulevastes uuringutes ning kaalude arvutamiseks andmestikus (Callegaro jt, 2015: 211-212; *National Research Council*, 2013). Eesti paraandmetel baseeruvatest töödest võiks esile tuua Kivistiku (2012) magistritöö intervjuu kestuse ja andmekvaliteedi seostest ESS 2010 baasil ning Undi, Solvaku, Vassili (2017) e-valimise uuringu. Võimalusi paraandmete kasutamiseks uuringu koguvea ja andmekvaliteedi mõõtmisel on kirjeldanud erinevad autorid (Jans, Sirkis ja Morgan, 2013: 197, 199; Kreuter ja Olson, 2013:13-42).

Käesolevas töös esitatud CRONOSe andmete analüüs baseerub osaliselt veebiküsitlusele vastamise ajal tekkivatele otsestele paraandmetele. Paraandmetest on töös kasutatud andmeid kutsete ja meeldetuletuste edastamise ning kohalejõudmise, vastamiseks kasutatud seadme tüübi, vastamiseks kulunud aja ja vastamise katkestamise kohta.

1.3 Uuringu koguvea hindamine

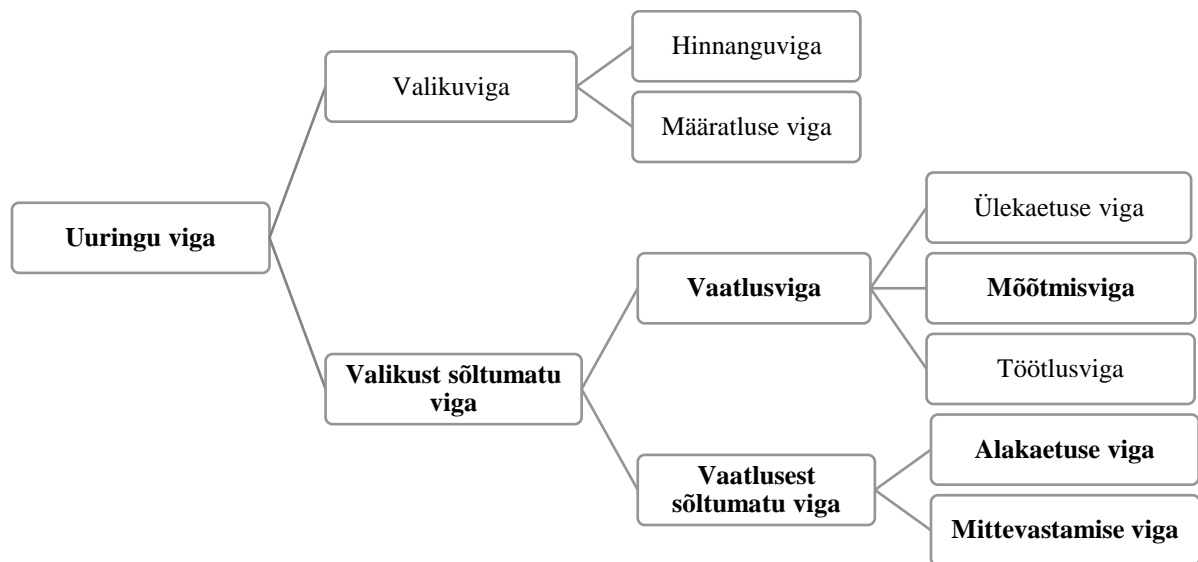
Uuringu läbiviijatel võib olla hästi planeeritud uuringudisain, kuid ikka tuleb ette asjaolusid, mille teostumist polnud võimalik vältida võimalik vältida ja mis võivad tekitada vea tulemuste üldistamisel. Joonisel 2 on kujutatud uuringu koguvea komponente, mille esinemisel ei pruugi

saadud tulemused kajastada uuritava tõelist väärtust (Bethlehem ja Biffignandi, 2011: 100). Uuringu koguvea moodustavad valikuvead ja valikust sõltumatud vead. Valikuvigadeks võib lugeda vigu, mis tekivad valimi võtmisest üldkogumist (Bethlehem ja Biffignandi, 2011: 100; De Leeuw, Hox ja Dillman, 2008: 9-10). Valikuviga võib esineda igal valimi moodustamisel, st iga uue valimi moodustamisega ühe ja sama valikuraami alusel saadakse mõnel määral erinev tulemus (Couper, 2000: 467). Valikuraamina käsitletakse uurimisaluste indiviidide koondnimekirja, mis on moodustatud eelnevalt defineeritud üldkogumi kõikidest indiviididest (Sue ja Ritter, 2012: 34). Valikuraami kuuluvate indiviidide teatud tingimustega väljavõtu tulemiks on valim, kellele esitatakse kutse uuringus osalemiseks (*Ibid.*).

Käesolev töö keskendub valikust sõltumatutele vigadele, mis jagunevad vaatlusvigadeks ja vaatlusest sõltumatuteks vigadeks. Vaatlusvead võivad tekkida andmekogumise või andmete töötlemise etapis (kui vaatlus on toimunud) (Bethlehem ja Biffignandi, 2011: 101). Vaatlusvigadest käsitletakse töös mõõtmisviga. Mõõtmisviga tekib, kui küsimus uuringus ei mõõda seda, mida see mõõtma oleks pidanud (Andres, 2016:112). Mõõtmisviga on uuringus osalejate vastuste kõrvalekalle tõelisest väärtusest, mis võib tuleneva mitmesugustest teguritest (Andres, 2016: 112; Callegaro jt 2015: 368 – 374; Couper, 2000: 475 – 477; De Leeuw, Hox ja Dillman, 2008: 11 - 13). Töö empiirilises osas uuritakse võimalikku mõõtmisviga CRONOSes ja sellest mõjutatud andmekvaliteeti eelkõige joonvastamise näol. Täiendavalt uuritakse võimalikku mõõtmisviga ka vastamisele kulunud aja ja respondendi subjektiivse vastamiskogemuse kaudu.

Vaatlusest sõltumatute vigade alla loetakse vigu järeldustes seoses vaatluse (mõõtmise) ära jäämisega (Bethlehem ja Biffignandi, 2011: 102). Vaatlusest sõltumatutest vigadest käsitletakse töös nii alakaetuse kui ka mittevastamise viga, keskendudes rohkem mittevastamise veale. Kaetuse veaks nimetatakse tunnuse väärtuse lahknevust valikuraamiga kaetud ja üldkogumi indiviidide vahel (Groves, 1989: 85). Alakaetuse veana käsitletakse juhtu, mil uuringutulemuste alusel tehakse järeldusi üldkogumile, arvestamata valikuraamiga katmata jäänud osa üldkogumis (Cooper 2000: 467; Lyberg ja Weisberg, 2016: 30). Töös ei uurita otseselt CRONOSe alakaetuse viga, vaid käsitletakse laiemalt veebiuuringute alakaetuse probleemi seoses indiviididega, kellele võimaldati CRONOSes osalemiseks tahvelarvuti koos internetiga (interneti mittekasutajad).

Mittevastamise veaks loetakse viga hinnangutes, mis on põhjustatud valimisse kuulunud indiviidide vastuste puudumisest või kui antud vastuseid ei ole võimalik kasutada (Bethlehem ja Biffignandi, 2011: 102, Toepoel, 2016: 9). Mittevastamise viga uuritakse töös CRONOSe andmetel peamiselt kutsele mittevastamisena. Täiendavalt uuritakse vastamise katkestamist ja küsimusele mittevastamist, mille esinemisel on individid andnud vastuseid küsitluslaine lõikes osaliselt.



Joonis 2. Uuringu vigade klassifikatsioon, töös käsitletavad uuringu vea komponendid on märgitud tumedas kirjas. Allikas: Bethlehem ja Biffignandi (2011: 100).

1.3.1 Vaatlusest sõltumatu viga

1.3.1.1 Alakaetus

Cooper (2000:467) toob veebiuuringute puhul olulisima probleemina esile kaetuse temaatika. Kaetuse all peetakse silmas valikuraami ja üldkogumi vastavust. Alakaetus esineb siis, kui valikuraam ei hõlma kõiki üldkogumisse kuuluvaid indiviide (Bethlehem ja Biffignandi, 2011: 66-67). Juhul kui valikuraami kuulub rohkem indiviide kui on üldkogumis (või on topelt), siis esineb ülekaetus. Veebipaneelide puhul on takistused kaetuse saavutamisel tugevalt seotud erinevustega internetile ligipääsus ja arvutikasutusoskustes üldkogumis (Callegaro ja Krosnick, 2014: 56). Mõnel juhul on kaetuse saavutamine võimalik olulise vaevata, nt uuringud üliõpilaste või konkreetse veebilehe külastajate kohta, kuid suuremaks väljakutseks on kaetuse saavutamine kogu rahvastiku uuringutes (Tourangeau jt, 2013: 11). Valimi raamiga mittekaetud populatsiooni osa moodustavate indiviidide tõenäosus valimisse sattuda on 0, mis on vastuolus

lihtsa juhuvaliku põhimõttega, et kõikidel indiviididel on võrdne võimalus uuringusse sattuda, ning seetõttu pole võimalik teha üldistusi üldkogumile (Callegaro jt, 2015: 111).

Interneti leviku ja arvutikasutusoskuse hindamisel üldkogumis on mitmeid olulisi defineerimisest tulenevaid väljakutseid, sh kuidas määratleda internetile ligipääsu või kasutusoskust (interneti kasutamine kodus, avalikes ruumides või töökohas). Internetile ligipääs on Eurostati metodoloogia dokumendile tuginedes defineeritud kui vähemalt ühe leibkonnaliikme võimalus internetti kasutada ükskõik millise seadme vahendusel, olenemata, kas seda kasutatakse või mitte (*Methodological Manual...*, 2018). Eestis omas 2009. aastal ligipääsu internetile 62% leibkondadest (valimis 16-74 aastased elanikud), 2018. aastal juba 90% leibkondadest (*IKT usage...*, 2019). Samuti on kasvanud kümne aasta jooksul keskmine regulaarsete arvutikasutajate (viimati kasutanud viimase kolme kuu jooksul) osakaal. Eestis kasutas 2009. aastal regulaarselt arvutit 71% indiviididest ning 2017. aastal juba 88% (*IKT usage...*, 2019.).

Seoses interneti leviku ja digitaalse kirjaoskuse kasvuga eelkõige arenenud riikides, on kaetuse probleem veebiuuringutes tasapisi vähenemas, vähemalt selles osas, mis puudutab internetiga kaetust (Callegaro jt, 2015: 111). Kaetuse probleemi teravdab aga asjaolu, et internetile ligipääsu ja digitaalse kirjaoskuse poolest ei jagune indiviidid kaheks osaks juhuslikult, vaid teatava suundumusega. Callegaro jt (2015: 112, 227) toovad esile, et suurema tõenäosusega kuuluvad interneti mittekasutajate hulka eakad, pigem naised, teatud vähemused, vähem haritud ja väiksema sissetulekuga indiviidid. Kõnealust eristumist interneti kasutuses ja arvuti kasutussageduses, laiemalt kõikides info-ja kommunikatsioonitehnoloogiaga seotud oskustes, on nimetatud ka lihtsaks digitaalseks lõhestumiseks (Vehovar jt, 2006: 281).

Alakaetuse probleemi on töö empiirilises osas käsitletud seoses interneti mittekasutajatega, kellele võimaldati uuringus osalemiseks tehnoloogilised vahendid. Tahvelarvutit pakuti indiviididele, kes CRONOSe eelnenud 2016. aasta Euroopa Sotsiaaluuringu (ESS, i.a) 8. vooru intervjuus vastasid, et nad ei kasuta internetti. Seega tahvelarvuti saanud respondendid esindavad seda osa ühiskonnast, kes veebiuuringust muidu välja jääksid. Antud teadmiste tuginedes sõnastan käesolevaks tööks esimese uurimisküsimuse. Kas ja mille poolest CRONOSe uurimisrühma poolt jagatud tahvelarvutitega respondendid erinevad teistest vastanutest? Eelnevalt esitatud kirjanduse ülevaatele tuginedes keskendun erisuste

väljaselgitamisel järgnevatele teguritele: indiviidi vanus, lõpetatud haridusaastate arv, sugu ja keel, milles küsitlusele vastati.

1.3.1.2 Kutsele mittevastamine

Tõenäosuslikul valimil põhineva veebiuuringu puhul eelneb uuringukutse edastamisele enamasti indiviidide uuringusse värbamine. Tõenäosusliku valimiga veebiuuringute puhul eelistatakse indiviidi värbamist näost näkku, telefoni või kirja teel. Eelistust tingib peamiselt asjaolu, et enamasti ei ole olemas täielikku valikuraami koos indiviidide meiliaadressidega. Selline kontaktiloomine indiviidiga võimaldab tõsta vastamismäära tänu personaalsuse ja legitiimsuse efektile, kuid viimastel aastatel on populaarsemaks muutunud ka värbamine e-maili teel (Toepoel, 2016: 75-76). Kui indiviid on värbamisel andnud oma nõusoleku uuringus osalemiseks, siis edastatakse talle uuringukutse. Vastamise alustamise eeltingimuseks on aga uuringukutse kättesaamine. Peamiste kutse kohalejõudmist takistavate teguritena on käsitletud vigaseid või aegunud e-maili aadresse ja internetiühendusest tulenevad tehnilised probleemid (Lozar Manfreda ja Vehovar, 2002; Callegaro ja DiSogra, 2008). Indiviidid, kelle puhul on teada, et kutse kättetoimetamine pole võimalik (nt õige meiliaadress pole selgunud ka teisi kontaktvahendeid kasutades), arvatakse küsitluseks sobivatest indiviididest välja. Veebiuuringute puhul tuuakse kutse kättetoimetamise takistustena esile ka kutse filtreerimist rämpsposti hulka, täitunud postkasti, pealiskaudset ja valikulist e-kirjade lugemist ning viiruste ja küberrünnakute kartust, mistõttu ei julgeta edastatud lingile klikkida (Bethlehem ja Biffignandi, 2011: 438 -439; Callegaro jt, 2015: 252; Toepoel, 2016: 75 - 76). Viimasena loetletud takistuste tõttu uuringust välja jäänud respondente pole aga võimalik kindlaks teha.

Indiviididega kontakti loomisel rakendatakse kolme tüüpi teavitusi: eelteavitusi, uuringu kutse edastamist ja meeldetuletusi (Fan ja Yan, 2010: 134 - 135). Üldiselt kõrgem kontaktiloomiste arv tõstab vastamismäära (Callegaro jt, 2015: 253 – 255; Götitz ja Grutzen, 2012). Laias plaanis on võimalik saavutada kõrgem vastamismäär siis, kui kontaktiloomise ajal on uuritav sihtrühm kõige vähem hõivatud, arvestades aastaaegu, vaheaegu, elustiili jne (Callegaro jt, 2015: 251 - 256). Vastamismäära võimaldavad suurendada ka korduvad meeldetuletused, kuid juba teise meeldetuletuse järgselt vastanute andmete kvaliteet ja lisandväärtus uuringu tulemustele on madal (Callegaro jt, 2015: 244-245). Samas on leitud, et hilisemad liitujad võivad anda uuringutulemustele olulist lisandväärtust, näiteks riskikäitumise või faktiliste näitajate uurimisel, mõjutades oluliselt vastuste jaotust (Kypri, 2004; Kypri jt, 2011). Veebipaneeli puhul

võib olla korduvate meeldetuletuste saatmisel aga pikas perspektiivis vastupidine efekt, sest paneeliliikmete jaoks võib pidev meeldetuletuskirjade saamine osutada tüütavaks (Göriz ja Grutzen, 2012: 248).

Uuringus osalemise soovi on selgitatud erinevate teoreetiliste lähenemiste kaudu. Callegaro jt (2015: 222) on mitmete autorite töid koondades esitanud nimekirja erinevatest võimalikest teooriatest, mida vastamise ja mittevastamise selgitamisel on kasutatud. Üheks uuringule vastamist selgitavaks teooriaks on sotsiaalse vahetuse teooria, mille keskseks ideeks on, et indiviid vastab teise inimese palvele suurema tõenäosusega siis, kui ta usaldab ja usub, et saadav kasu ületab tegevusega kaasnevat kahju. Sotsiaalse vahetuse teooriale on üles ehitatud Dillmani jt (2014) kohandatud disaini meetod (Callegaro jt, 2015: 222). Veel lähtutakse vastamise analüüsil ja selgitamisel Ajzeni planeeritud käitumise teooriast. Ajzeni (1991) planeeritud käitumise teooria selgitab inimekäitumist eelkõige kavatsuse kaudu. Indiviidi planeeritud käitumine sõltub otseselt indiviidi soovist käituda. Soovi mõjutab aga see, kuidas antud käitumisse või tegevusse ühiskonnas suhtutakse ning kuidas indiviid ise sellesse suhtub. Planeeritud käitumise teostumise määrab see, kas antud tegevus on ka päriselt teostatav, st kas on olemas vajalikud vahendid. Ajzeni planeeritud käitumise ideid on võimalik üle kanda ka vastamiskäitumisele. Bosnjak, Tuten ja Wittmann (2005: 494) on veebiuuringute kontekstis laiendanud Ajzeni teooriat, lisades olulise vastamise mõjutajana moraalse kohustuse. Eelmainitutele lisaks võiks veel esile tuua võimenduse – olulisuse teooria, mille peamine idee seisneb selles, et uuringus osalejatele mõjuvad uuringu atribuudid (kaasatus, motivaatorid, teema olulisus) erinevalt, mistõttu mõjutavad atribuutide rakendamine ja muutmine uuringutulemusi erinevalt, st võimendumine toimub individipõhiselt (Groves, Singer ja Corning, 2000). Lisaks mainitud teooriatele on veel mitmeid teooriaid, mida on uuringutele vastamise ja mittevastamise selgitamisel kasutatud (Callegaro jt, 2015: 222).

Soovi uuringus osaleda võivad mõjutada individuaalsed omadused, sotsiaalsed tegurid, näiteks üldine meelestatus uurimiskultuuri suhtes ühiskonnas, ning uuringu disainist tulenevad tegurid (Fan ja Yan, 2010). On leitud, et veebiuuringutes soovivad osaleda pigem nooremad, mehed, kõrgema haridustasemega ja kõrgema sotsiaalse staatusega inividid (Callegaro jt, 2015: 227). Antud omadused seostuvad ka kõrgemate IKT oskustega. Seega veebiuuringutes osalevad pigem kõrgema digitaalse kirjaoskusega inividid, kes kas soovivad rohkem veebiküsitlusele vastata või lihtsalt tulevad sellega paremini toime (*Ibid.*).

Vastamise keskseks mõõdikuks loetakse vastamismäär. Madal vastamismäär on oluliseks ohuallikaks uurimistulemuste usaldusväärsusele (Bethlehem ja Biffignandi, 2011: 432). Vastamismäärana käsitletakse üldiselt uuringule vastanud indiviidide ja küsitlusele vastamiseks sobivate indiviidide arvude suhet (Callegaro jt, 2015: 228). Küsitluseks sobivate indiviididena on käsitletud indiviide, kes kuuluvad valimisse ja esindavad uuritavat üldkogumit ning kelle puhul ei ole teadaolevaid takistusi uuringukutse kättetoimetamisel. Vastamiseks sobivate indiviidide määratlemine uuringus võib osutuda aga keeruliseks, mida on võimalik veebiuuringute puhul hõlbustada erinevate meetmete abil, nt automaatvastuste sisu arvesse võtmisega või keeldumise põhjuste küsimisega (*Ibid.*: 223).

Veebiuuringute vastamismäärad on selgelt madalamad kui teiste küsitlusmeetodite puhul (vastamismäär parimal juhul 30-40%) ning nende vastamismäär on langustrendis (Callegaro jt, 2015: 60, 231; Fan ja Yan, 2010, Lozar Manfreda jt, 2008; Toepoel, 2016: 114). Madalamat vastamismäärat võivad esile kutsuda digitaalse kirjaoskuse vähesus ja sellest tulenev lisanduv koormus vastamisele, tehnilised takistused, turvatunde puudumine veebis andmete edastamiseks, intervjuerija puudumine ning üleküsitletus ühiskonnas (Callegaro jt, 2015; Lozar Manfreda jt, 2008). Dillman, Smyth ja Christian (2014: 37) rõhutavad paneeluuringu eripära, et paneeluuring võimaldab hoida vastamismäärat stabiilsena, sest kui respondent on huvi uuringus osalemise vastu üles näidanud ühe korra, teeb ta seda tõenäoliselt ka järgmine kord.

Vastamismäärat aitab tõsta motivaatorite kasutamine. Nendeks võivad olla rahalised või mitterahalised, tingimuslikud või tingimusteta, enne või pärast uuringut jagatud motivaatorid. (Callegaro jt, 2015: 248 – 251; Toepoel, 2016: 108 - 111). Kõik motivaatorid ei mõju ühetaoliselt, mida võib selgitada võimenduse – olulisuse teooriale tuginedes individipõhise motivaatori olulisusega, st mõnele indiviidile võib mitterahaline preemia mõjuda motiveerivalt ning mõnele mitte. Toepoel (2016: 111) on välja toonud, et vastamismäärat tõstmiseks võivad töötada tingimusteta rahalised või mitterahalised ja tingimuslikud rahalised motivaatorid. Edukaimaks on osutunud tingimusteta rahaline motivaator (Bethlehem ja Biffignandi, 2011: 149). Veebiuuringute puhul on tingimusteta motivaatoritel olnud väiksem efekt võrreldes teiste küsitlusviisidega (Callegaro jt, 2015: 248). Motivaatori suuruse määrab enamasti uuringu eelarve. Kui uuringu üldised kulud on kõrged, on mõistlik rakendada suurema väärtusega motivaatoreid (Toepoel, 2016: 112 -113). Lisaks on täheldatud, et motivaatori väärtuse

suurendamine ei tõsta vastamismäära lineaarselt (Fan ja Yan 2010: 135; Toepoel, 2016:113). Veebipaneelides rakendatakse ühe motivaatorina ka punktide kogumist, mille eest teatud arvule küsimustikele vastates on võimalik soetada kinkekaart. Üldiselt parandavad motivaatorid andmete kvaliteeti eelkõige vastamismäära tõstmise abil, vähendavad küsimustele mittevastamist, parandavad avatud küsimuste vastuste sisukust ning vähendavad piisavalt rahuldavate vastuste andmist (Callegaro jt, 2015: 250).

Kutsele mittevastamisena käsitletakse vastuste puudumist küsitluseks sobivalt indiviidilt, st uuringukutse saanud indiviid pole vastanud ühelegi küsimusele (Callegaro jt, 2015:220). Kutsele mittevastamine võib tuleneda kontakti mittesaavutamisest, vastamisest keeldumisest ja teistest asjaoludest (nt kolmandate isikute sekkumine, tervislikud, keelelised, tehnilised või administratiivsed takistused) (Andres, 2012: 113; Callegaro jt, 2015: 220; De Leeuw jt, 2008). Kontakti saavutamine võib osutada keerulisemaks hõivatamate, eakamate ja jõukamate indiviididega (Callegaro, 2015: 220). Mittejhuslik mittevastamine suurendab mittevastamise viga uurimistulemustes (Andres, 2012: 112; Toepoel, 2016: 9).

Kutsele mittevastamisest tuleneva vea vähendamiseks on kaks võimalust. Üheks võimaluseks on vähendada mittekontaktsete ja keeldujate arvu. Mittekontaktsete ja keeldujate arvu vähendamist on võimalik saavutada korduvate kontaktikatsetega, meeldetuletusteadete edastamisega, kontaktikatsete aja (päev, kellaaeg) varieerimisega ning rakendades motivaatoreid (De Leeuw jt, 2008: 10 – 11, Fan ja Yan, 2010; Göritz ja Grutzen, 2012: 247; Holbrook, Krosnick & Pfent; 2008). Vastamismäära tõstmine ei tähenda tingimata aga andmete kvaliteedi tõusu (Bethlehem ja Biffignandi, 2011: 432; Holbrook, Krosnick & Pfent, 2008; Göritz ja Grutzen, 2012; Peytchev, 2013). Teise strateegiana kasutatakse kohandamise strateegiat, mida on võimalik rakendada pärast andmekogumise etappi, nt rakendada kaalumist (Biemer ja Christ, 2008; Berinsky, 2008).

Seega võib öelda, et kutsele mittevastamine ja madal vastamismäär on ohuallikaks uurimistulemuste laiendamisel üldkogumile. On leitud, et kutsele mittevastamist esineb enam vanemate, naiste, madalama haridustasemega ja madalama sotsiaalse positsiooniga indiviidide seas, keda iseloomustab üldiselt ka madalam digitaalne kirjaoskus. Ajzeni planeeritud käitumise teooriale tuginedes võib antud trendi seostada tegutsemiseks vajalike vahendite ja oskuste vähesusega. Vastamismäära tõstmine korduvate meeldetuletuste abil ei pruugi paraku uuringu andmekvaliteeti parandada. Samas just meeldetuletustega kaasatud indiviidid võivad

esindada mingit olulist osa ühiskonnast, kelle väljajäämine uuringust suurendaks oluliselt mittevastamise viga. Antud tulemustele toetudes sõnastan andmekvaliteedi ja mittevastamise vea uurimiseks teise uurimisküsimuse: millised tegurid mõjutavad CRONOSe veebipaneelis kutsele mittevastamist? Varasemate uurimuste kogemustele toetudes uuritakse käesolevas töös, kas ja kuidas mõjutavad kutsele mittevastamist indiviidi sugu, vanus, haridustase, keel (2016. a ESS intervjuus kasutatud keel), huvi poliitika vastu, interneti kasutussagedus ning eelmisel lainel vastamiseks kasutatud seade. Vastamiseks kasutatud seadme lisan uuritavate tegurite juurde seoses võimaliku erineva vastamiskogemusega eelneval lainel sõltuvalt sellest, millist seadet vastamiseks kasutati (toetudes peatükkidele 1.3.2.1 - 1.3.2.3).

1.3.1.3 Vastamise katkestamine

Vastamise katkestamisena käsitletakse juhtu, kus uuringus osaleja alustab veebis küsimustikule vastamist, kuid katkestab vastamise enne lõppu (Callegaro, 2015: 221). Steinbrecher, Roßmann ja Blumenstiel (2015) on poliitikateemaliste uuringute puhul leidnud, et katkestamist kutsuvad esile peamiselt motivatsiooni langus, poliitiline apaatsus, tehnilised probleemid vastamisel, ülesande keerukus ja individuaalne võimekus.

Motivatsiooni võib vähendada küsimustiku oodatav pikkus ning see, kuivõrd see ühtib tegelikkusega (Fan ja Yan, 2010; Tourangeau jt, 2009). Küsitluse pikkus mõjutab nii küsimustele vastamise katkestamist, vastamise alustamist kui ka vastamise katkestamist (Callegaro jt, 2015: 175-177; Galesic ja Bosnjak, 2009). Motivatsiooni langus võib tuleneda huvi puudumisest uuritava teema vastu, aga ka küsimuste kehvast sõnastusest või disainivalikutest. Motivatsiooni langust võib esile kutsuda kasutatud skaalatüübid (nt liugurskaala) või ühel leheküljel mitme küsimuse kuvamine (Bethlehem ja Biffignandi, 2011: 442; Peytchev, 2009: 93). Lisaks võib voogmeedia või animatsioonide kasutamine esile kutsuda tehnilisi takistusi vastamise jätkamisel (viibib küsimuste kuvamine või ei lase tehniliselt vastamist jätkata) (Waltson jt, 2006). Katkestamine võib tuleneda ka veebilehe ekslikust sulgemisest vähese arvutikasutusoskuse tõttu.

Katkestamist kutsub esile ülesande keerukus, mis seostub eelkõige küsimustiku sisu või disainiga. Vastamine on keerukam avatud küsimuste või arvulist vastust eeldavate küsimuste puhul andmist (Peytchev, 2009: 93). Vastamist muudab raskemaks, kui kuvatud küsimused on pikad, eeldavad küsimuse täpsemat mõtestamist või hinnangu andmist (*Ibid.*).

Individuaalset võimekust küsimustele vastata mõjutavad kognitiivne võimekus, uuritava teemaga tuttav olemine ning kindlakskujunenud arvamus uuritava teema suhtes (Steinbrecher jt, 2015: 291). Individuaalseid sotsiaal-demograafilisi näitajaid ei seostata otseselt vastamiskäitumist prognoosivate teguritena, kuid need viitavad teisele mõju omavatele faktoritele. Näiteks on Peytchev (2009) seostanud hariduse mõju katkestamisele eelkõige kognitiivse võimekuse kaudu, st kõrgema haridustasemega indiviidil on kõrgem kognitiivne võimekus. Kui kõrgem vanus võiks samuti prognoosida katkestamist, siis autori analüüsi tulemustes avaldus just noorematel indiviididel kõrgem risk katkestada, mida autor seostas teema vastu huvi puudumisega (tervis). Võiks arvata, et küsimustele vastamise katkestanud indiviidid suhtuvad küsimustele vastamisse tähelepanematult või hooletumalt, ent autor leidis, et katkestaja keskmine küsimusele vastamise aeg on hoopis pikem. See näitab, et katkestamist mõjutab pigem madal individuaalne võimekus, mis kutsus esile vastamisele kulunud aja pikenemist ning seejärel vastamise katkestamist ülemäärase pingutusvajaduse tõttu.

Uurides hiliste küsitlusele vastajate vastamiskäitumist, on Göritz ja Grutzen (2012) leidnud, et rohkete meeldetuletustega uuringusse kaasatud respondentide puhul esineb sageli vastamise katkestamist. Autorid on pakkunud, et meeldetuletuse järgsetel vastajatel võib olla madalam digitaalne kirjaoskus, sest harvema e-kirjade lugejaid seostatakse ka madalama digitaalse kirjaoskusega, mida võib lugeda ka üheks võimalikuks katkestamiste sageduse põhjuseks. Lisaks on veebiuuringute kontekstis oluline, et katkestamine ühel lainel prognoosib mittevastamist järgmistel lainetel (Göritz, Wolff ja Goldstein, 2008).

Vastamise katkestamisel võivad seega olulist prognoosivat mõju omada respondendi motivatsioon, huvi poliitika vastu, küsimustiku sisu ja disainiaspektid, individuaalne võimekus ja meeldetuletuste arv. On leitud, et katkestamisi esineb enam madalama haridustasemega ja noorte respondentide seas. Lisaks võib katkestamist esile kutsuda see, kui respondent ei ole uuritava teemaga tuttav või kui puudub teema suhtes kindlakskujunenud arvamus. Antud tulemustele toetudes sõnastan andmekvaliteedi ja mittevastamise vea uurimiseks kolmanda uurimisküsimuse: millised tegurid mõjutavad CRONOSe veebipaneelis vastamise katkestamist? Eelnevalt esitatud varasemate uurimistulemuste põhjal uuritakse, kas ja kuidas mõjutab vastamise katkestamist respondendi sugu, vanus, haridustase, vastamiseks kasutatud keel, interneti kasutussagedus, huvi poliitika vastu ning edastatud meeldetuletuste arv.

1.3.1.4 Küsimusele mittevastamine

Küsimusele mittevastamisena käsitletakse konkreetsele küsimusele vastamata jätmist respondentide poolt, kes enne vastava küsimuseni jõudmist pole vastamist katkestanud (Callegaro jt, 2015: 220). Küsimusele mittevastamise hulka loetakse arvamuse puudumised (sh „ei oska öelda“) ja vastamisest keeldumised. Küsimusele mittevastamist mõjutavad üldised tegurid, mis tulenevad individuaalsetest omadustest (haritus, avatus, huvi uuritava teema vastu, kognitiivsed võimed ja kognitiivse pingutuse vajadus), küsimuste sõnastusest (vaimset pingutust nõudvad küsimused), küsimuste tundlikkusest (nt sissetulek, religioosne kuuluvus) ja intervjuerija mõjust intervjueeritavale (positiivne ootus vastuse saamisele) (Berinsky, 2008; Callegaro jt, 2015: 60; De Leeuw, 2012). Erinevad sotsiaaldemograafilised tunnused omavad mõju küsimusele mittevastamisele pigem kaudselt, näiteks vanuse kasvades suurenevad küsimusele mittevastamist seostatakse vastamise ülesande keerukuse ning vananemisega kaasneva kognitiivse võimekuse langusega (Peytcher, 2009).

Küsimusele mittevastamist võib mõjutada küsimustiku disain. Näiteks suurendab avatud küsimuse vastuse lahtri suurus vastuse pikkust, kuid suurendab riski mittevastamiseks (Callegaro jt, 2015:128-129). Küsimuse vastusevariantide esitusviis võib samuti mõjutada küsimusele mittevastamist. Healey (2007: 123) on leidnud, et vastuste formaadi ja küsimusele mittesisuliste vastuste andmise vahel otsene seos puudub, kuid on täheldanud veidi suuremat küsimusele vastamata jätmist küsimuste puhul, mille vastus on kuvatud rippmenüuna.

Küsimusele mittevastamisega toimetulekuks on laias plaanis kaks strateegiat: ennetamine ja kohanemine. Ennetamise etapis on võimalik minimeerida võimalike vastamata jäävate küsimuste arvu, seda näiteks sobiva andmekogumise viisi valikuga või küsimustiku disaini abil. Küsimusele mittevastamisest tuleneva vea vähendamise teiseks võimaluseks on rakendada puuduvate väärtuste asendamiseks mitmesuguseid tehnikaid, nt imputeerimist (Berinsky, 2008).

Küsimusele mittevastamist võib kokkuvõtvalt seostada peamiselt individuaalsete omaduste, vastamise keerukuse ja küsimuste sisuga. Vastamine on vaimset pingutust nõudvate küsimuste ja eelkõige avatud küsimuste puhul keerulisem, kutsudes esile vastamata jätmist. Samuti võib esineda küsimusele mittevastamist olenevalt küsimuse vastusevariantide formaadist. Individuaalseid omadusi arvesse võttes võib küsimusele mittevastamist esineda enam

madalama kognitiivse võimekuse tõttu vanemate ja madalama haridustasemega, lisaks teemast vähem huvitatud indiviidide seas. Antud uurimistulemustele toetudes sõnastan andmekvaliteedi ja mittevastamise vea uurimiseks neljanda uurimisküsimuse: millised tegurid mõjutavad CRONOSe veebipaneelis küsimusele mittevastamist? Varasematele uurimistulemustele tuginedes käsitletakse võimalikku mõju omavate teguritena käesolevas töös küsimuse teemat, vastusevariantide formaati ja skaalatüüpi, lisaks respondendi sugu, vanust, haridustaset, keelt, milles küsitlusele vastati, interneti kasutussagedust, huvi poliitika vastu, mitme asjaga vastamisega samaaegset tegelemist, meeldetuletuste arvu ning vastuste täpsuse rõhutamist. Viimane tegur tuleneb esimese, kolmanda ja viienda laine eksperimendist, mille raames kuvati vastajatele erinevaid küsitluse sissejuhatusi (täpsemalt ptk 2.3).

1.3.2 Mõõtmisviga

1.3.2.1 Joonvastamine

Ebatäpsete vastuste andmisega seoses kasutatakse laiemat mõistet *satisficing* (Tourangeau jt, 2013: 4-5). Mõistet kasutatakse piisavalt rahuldavate vastuste andmise tähenduses (Simon, 1956). Kitsamalt on mõistet kasutatud joonvastamisena, s.o süstemaatiline või juhuslik vastamine küsimuste sisusse keskendumata (Anduiza ja Galais, 2017; Callegaro jt, 2015: 177 - 178). Käesolevas töös käsitletakse joonvastamist samasuguse skaalaga küsimustes süstemaatilise ühetaoliste vastuste andmisena. Joonvastamise tõenäosust suurendab üldiselt vastamise keerukus, vähene motivatsioon ja madal individuaalne võimekus (Krosnick, 1991; Hamby ja Taylor, 2016).

Küsimustiku disainivalikud võivad suurendada vastamise keerukust ja seeläbi mõjutada joonvastamise esinemist uuringus. Joonvastamist esineb enam avatud vastusega ja maatriksina kuvatud küsimuste puhul (Steinbrecher, Roßmann ja Blumenstiel, 2015: 290; Liu ja Cernat, 2018: 701). Disainivalikud võivad mõjuda vastamiseks kasutatud seadmest sõltuvalt erinevalt. Näiteks pole maatrikstüüpi küsimused parim valik nutitelefonidega vastajate jaoks, sest skaalad ei pruugi ekraanile ära mahtuda ja vastamine eeldab ekraanipildi pidevat liigutamist, mis kutsub esile enam joonvastamist (Liu ja Cernat, 2018: 701). Lisaks on uuritud vastuste esitusviisi mõju andmekvaliteedile veebiuuringutes ning on leitud, et rippmenüüga valikuvариandid kutsuvad esile vastuste muutmist ja pikemat vastamisele kuluvat aega, kuid vähendavad primaarefkti esinemist võrreldes raadionuppudega (Healey, 2007: 123). Primaarefektina käsitletakse veebiuuringu visuaalsest esitlusest tulenevat nähtust, mille puhul respondendid kalduvad

valima visuaalselt eespool kuvatud vastusevariante (Callegaro jt 2015: 78 – 79; Schwarz jt, 2008: 33).

Joonvastamist mõjutavad ka individuaalne võimekus ja motiveeritus. On leitud, et joonvastamist esineb enam madalamalt haritute, vanemate, poliitilisest osalusest vähem huvitatute ja nn. professionaalsete vastajate seas (Anduiza ja Galais, 2017; Peytchev, 2009: 80). Täiendavalt on uuritud hiliste küsitlusele vastajate andmekvaliteeti veebiuuringus ning Göritz ja Grutzen (2012: 247) on leidnud, et indiviidile saadetud meeldetuletuste arv ei prognoosi joonvastamist ega ka küsimusele mittevastamist.

Veebiuuringule on võimalik vastata erinevaid tehnoloogilisi seadmeid kasutades, seetõttu on uuritud ka vastamiseks kasutatud seadme mõju andmekvaliteedile. Antoun, Couper ja Conrad (2017) on leidnud, et nutitelefonide kasutamine ei suurenda joonvastamise tõenäosust ja üllatuslikult pikendab avatud küsimustele antavaid vastuseid. Nutitelefonides vastuste andmist seostatakse võimalusega vastata küsitlusele väljaspool kodu ja rahvarohketes avalikes kohtades, mis võib esile kutsuda ebatäpsete vastuste andmist. Revilla (2016) on aga leidnud, et kuigi nutitelefonide kasutamisel küsitlusele vastamiseks võib olla lähedal palju võõraid, ei ole andmekvaliteedis olulisi erisusi, sh tundlikele küsimustele vastamises.

Üldiselt võib öelda, et joonvastamist kutsub esile vastamise keerulisus, madal individuaalne võimekus ja vähene motivatsioon. Vastamise keerukus võib tuleneda küsimuste sisust, aga ka küsimuste disainivalikutest. On leitud, et joonvastamist esineb enam indiviidide seas, kellel on madal haridustase, kes ei ole huvitatud poliitilisest osalusest ja nn. professionaalsete vastajate seas. Lisaks on leitud, et vastamiseks kasutatud seade pigem ei mõjuta andmekvaliteeti. Siinkohal sõnastan andmekvaliteedi ja mõõtmisvea uurimiseks järgmise, s.o viienda uurimisküsimuse: millised tegurid mõjutavad CRONOSe veebipaneelis joonvastamist? Eelnevalt esitatud uurimistulemuste valguses uuritakse täpsemalt, kas ja kuidas mõjutavad joonvastamist järgmised tegurid: respondendi sugu, vanus, haridustase, interneti kasutussagedus, huvi poliitika vastu, keel, milles küsitlusele vastati, mitme asjaga tegelemine vastamisega samal ajal, edastatud meeldetuletuste arv, vastamiseks kasutatud seade, vastamisele kulunud aeg ning eksperimentidest tulenev vastuste täpsuse olulisuse ja hooletu või hoolika vastamise tagajärgede rõhutamine. Viimane tegur tuleneb teise, neljanda ja kuuenda laine eksperimentidest, mille raames kuvati küsimustiku jooksul vastajatele erinevaid teateid vastuste hoolsusest tulenevate tagajärgede kohta (täpsemalt ptk 2.3).

1.3.2.2 Vastamiseks kuluv aeg

Veebiuuringute üheks oluliseks eeliseks on, et küsimustele vastamise aega on võimalik väga täpselt mõõta. Respondendi vastamiseks kulunud aeg võib anda täiendavat informatsiooni andmekvaliteedi kohta. Vastamisaega võivad pikendada keeruline uurimisteema, disainivalikud, individuaalne võimekus või tehnoloogilised faktorid nagu internetiühenduse või seadme eripärad. Antoun ja Cernat (2019) on leidnud, et nutitelefoni vastamise aega pikendas eelkõige see, kui vastaja polnud sage nutitelefoni kasutaja, kulutades seetõttu vastamiseks oluliselt enam aega kui arvutis vastajad. Nutitelefoni puhul täheldati ka suuremat mitme tegevusega samaaegselt tegelemist, mis pikendab üldiselt vastamiseks kuluvat aega (*Ibid.*, Antoun jt, 2017). Vastamisele kuluvat aega võib pikendada ka pühendumine vastamisele või vastamiseks antud panus, kuid Revilla ja Ochoa (2015: 109) on leidnud, et vastamisele kuluv aeg on küll seotud vastamiseks tehtud pingutuse hinnanguga, kuid mõju on nõrk.

Respondendi omaduste mõju uurimisel vastamisajale on leitud, et vastamiseks kulub vähem aega noorematel ja kõrgemalt haritud respondentidel (Antoun ja Cernat, 2019). Samas võib osutada problemaatiliseks ka liigne kiirustamine vastamisel, mis viitab nn. professionaalsetele vastajatele, kes eristuvad teistest vastajatest kiirustamisega vastuste andmises (Callegaro jt, 2015: 343 – 344; Tourangeau jt, 2013: 5). Kuigi ka ebatäpsete vastuste andmisele kuluv aeg varieerub, võib liialt kiire küsimustele vastamise aeg viidata võimalikule mõõtmisveale (Callegaro jt, 2015: 209; Revilla ja Ochoa, 2015). Callegaro jt (2015: 208 - 210) järgi on lihtsaimaks viisiks kiirustajate väljaselgitamisel lugeda normaalseks lugemiskiiruseks kaks kuni kolm standardhälvet keskmisest või mediaanist, millest kiiremaid vastajaid võib lugeda kiirustajateks.

Üldiselt võib vastamisele kulunud aeg olla mõjutatud uurimisteemast, küsimustiku disainivalikutest, respondendi võimekusest, muudest tegevustest vastamisega samal ajal või erinevatest tehnoloogilistest faktoritest. Sellest lähtuvalt sõnastan andmekvaliteedi ja mõõtmisvea uurimiseks kuuenda uurimisküsimuse: millised tegurid mõjutavad CRONOSe veebipaneelis vastamiseks kuluvat aega? Varasemate uuringute tulemustele tuginedes käsitletakse töös võimalike mõjuteguritena respondendi sugu, vanust, haridustaset, vastamiseks kasutatud keelt, interneti kasutussagedust, huvi poliitika vastu, mitme asjaga tegelemist vastamisega samal ajal, edastatud meeldetuletuste arvu ning eksperimentidest tulenevat vastuste täpsuse olulisuse ja hoolika või hooletu vastamise tagajärgede rõhutamist.

1.3.2.3 Hinnang uuringus osalemise kogemusele

Küsitlusele vastamise soovi võib mõjutada see, milline on olnud indiviidi eelnev uuringutes osalemise kogemus (Callegaro jt, 2015: 227). Veebipaneeli kontekstis võib hinnang uuringus osalemise kogemusele ühel lainel anda täiendavat informatsiooni prognoosimaks edasist vastamiskäitumist.

Hinnang vastamiskogemusele võib olla näidata, kuivõrd pühendunult on respondent küsitlusele vastanud (Callegaro jt, 2015: 170-173; McMahon ja Stamp, 2009). Pühendumise indikaatoritena kasutatakse lisaks tehnilistele indikaatoritele (vastamiseks kuluv aeg, silma liikumine jne), küsimustele mittevastamist, joonvastamist, aga ka küsitluse lõpuosas uuritud hinnanguid vastamiskogemusele. Hinnang vastamiskogemusele võib peegeldada pühendumist vastamisele ja seeläbi ka andmekvaliteeti, näiteks võib madal hinnang pingutusele vastamisel tähendada, et vastamine oli respondendi jaoks kerge, aga ka võimalikku madalat andmekvaliteeti vähese pühendumise tõttu. Revilla ja Ochoa (2015) on aga leidnud, et madal hinnang küsimustikule vastamiseks antud panusele ei ole madala andmekvaliteediga otseselt seotud. Autorid põhjendavad seose puudumist sotsiaalselt soovitatavate vastuste andmisega hinnanguküsimustes. Samas on Callegaro jt (2015: 171-172) pidanud respondendi hinnanguid, sh hinnangut vastamise keerukusele ja vastamise nautimisele, olulisteks kaudseteks pühendumist peegeldavateks näitajateks.

Seega hinnang vastamiskogemusele võib anda täiendavat informatsiooni pühendumise ja edasise vastamiskäitumise kohta. Varasematele uurimistulemustele tuginedes on teada, et vastamiseks antud panuse ja andmekvaliteedi vahel küll otsene seos puudub, kuid võib oletada, et kaudne seos hinnangute ja andmekvaliteedi vahel on siiski olemas. Subjektiivsed hinnangud võivad anda täiendavat informatsiooni teisi andmekvaliteedi mõõdikuid mõjutavate tegurite kohta. Siit tuleneb ka viimane uurimisküsimus andmekvaliteedi ja mõõtmisvea uurimiseks: millised tegurid mõjutavad CRONOSe veebipaneelis respondendi hinnangut uuringus osalemise kogemusele? Varasemate uuringute kogemustest lähtudes uuritakse töös respondendi soo, vanuse, haridustaseme, vastamiseks kasutatud keele, interneti kasutussageduse, poliitika huvi, vastamisega samal ajal mitme asjaga tegelemise, meeldetuletuste arvu ning eksperimentidest tuleneva täpsete vastuste olulisuse ning hoolika või hooletu vastamise tagajärgede rõhutamise mõju vastamiskogemuse hinnangule.

1.4 Uurimuse eesmärk ja uurimisküsimused

Käesoleva magistritöö eesmärgiks on välja selgitada erinevate tegurite mõju veebipaneeli andmekvaliteedile ning seeläbi hinnata CRONOSe mittevastamisest tuleneva vea ja mõõtmisvea ulatust. Töö empiiriline osa keskendub CRONOSe Eesti paneeliliikmete kõikide paneeli lainete andmetele ning rakendab analüüsis Berzelaki jt (2018) poolt eeskujuks seatud metodoloogilist lähenemist.

Varasemate andmekvaliteediga seotud uurimistulemuste põhjal on CRONOSe andmetel andmekvaliteedi, mittevastamise vea ja mõõtmisvea uurimiseks sõnastatud seitse uurimisküsimust. Siinkohal esitan kokkuvõtlikult eespool sõnastatud uurimisküsimused.

1. Kas ja mille poolest CRONOSe uurimisrühma poolt jagatud tahvelarvutitega respondendid erinevad teistest vastanutest? (ptk 1.3.1.1)
2. Millised tegurid mõjutavad kutsele mittevastamist? (ptk 1.3.1.2)
3. Millised tegurid mõjutavad vastamise katkestamist? (ptk 1.3.1.3)
4. Millised tegurid mõjutavad küsimusele mittevastamist? (ptk 1.3.1.4)
5. Millised tegurid mõjutavad joonvastamist? (ptk 1.3.2.1)
6. Millised tegurid mõjutavad vastamiseks kuluvat aega? (ptk 1.3.2.2)
7. Millised tegurid mõjutavad respondendi hinnangut uuringus osalemise kogemusele? (ptk 1.3.2.3).

1.5 Varasemad tõenäosuslikul valimil põhinevad veebipaneelid

Tõenäosuslikul valimil põhinevaid kogu rahvastiku veebipaneele on läbi viidud seni küllaltki vähe, kusjuures kõik sellised paneelid on läbi viidud ühe riigi piires. Antud peatüki lõpetuseks ja käesoleva töö taustaks esitan ülevaate varasemalt läbi viidud tõenäosuslikul valimil põhinevatest olulisematest kogu rahvastiku veebipaneelidest. Blom jt (2016) on võrrelnud nelja Euroopa ning Sommer (2017) seitset ülemaailmset tõenäosuslikul valimil põhinevat ja avalikult kättesaadava dokumentatsiooniga veebipaneeli:

1. kõige tuntum on 2007. aastal alustatud LISS (*Longitudinal Internet Studies for the Social sciences*, Holland), mis on ulatuslik leibkonnapõhine veebipaneel, mille valikuraam moodustati riikliku rahvastikuregistri põhjal, paneeli lained on kuised ja intervjuu

- kestuseks on 30 minutit, rakendatakse motivaatoreid ja indiviididele, kellel on piiratud ligipääs internetile võimaldatakse uuringus osalemiseks tehnoloogilised vahendid;
2. 2012. aastal algatatud GIP (*German Internet Panel*, Saksamaa), mis on klastervalimil põhinev isiku-uuring, paneeli lainete sagedusega kord kahe kuu jooksul ja kestusega 20-25 minutit, rakendatakse rahalisi motivaatoreid ja indiviididele piiratud ligipääsuga internetile tagatakse tehnoloogilised vahendid;
 3. GESIS Panel 2013-2014. a (Saksamaa), mis on kohalikel rahvastikuregistritel põhinev kombineeritud meetodiga paneeluuring, kus interneti mittekasutajatele võimaldatakse vastamine traditsioonilistel meetoditel, paneeli lained toimuvad iga kahe kuu järel ja kestuseks on 20-25 minutit;
 4. ELIPSS (*Étude Longitudinale par Internet Pour les Sciences Sociales*, algatatud 2012. aastal Prantsusmaal), mis on klastervalimil põhinev leibkonnapaneeel, mille kõikidele respondentidele võimaldatakse tehnoloogilised vahendid uuringus osalemiseks eesmärgiga pakkuda võrdseid stiimuleid ja vältides erinevatest tehnoloogiate vahendite kasutamise eripärasid, võimaldada maksimaalset mõõtmistäpsust;
 5. 2013-2015. aastal läbiviidud NCP (*Norwegian Citizen Panel*, Norra), mille valimi freim põhineb Norra riiklikul rahvastikuregistril, arvesse võttes Norra kõrget interneti levikut, ei pakutud interneti mittekasutajatele uuringus osalemiseks alternatiive, paneeli lained leiavad aset kolm korda aastas ja ühe küsitluse kestuseks on 15 minutit;
 6. 2014. aastal algatatud ATP (*American Trends Panel*, USA) on kombineeritud meetodiga kakskeelne (inglise ja hispaania keel) paneeluuring, mida algselt korraldati kuisena, seejärel 2-3 kuu järel, valim moodustati telefoninumbrite juhusliku valiku põhjal ning respondentidele piiratud ligipääsuga internetile võimaldatakse traditsioonilisi uuringus osalemise meetodeid;
 7. 2008-2009 toimunud FFRISP (*Face-to-Face Recruited Internet Survey Platform*, USA), mille värbamine toimus mitmes etapis, paneeli lained toimusid kuisena 12 kuu jooksul kestusega 25-30 minutit, kõikidele potentsiaalsetele respondentidele pakuti sülearvutit ja rahalisi motivaatoreid, selle paneeli jätkuks kutsuti liikmeid paneeli ALP (*The RAND American Life Panel*).

2 Metoodika

2.1 Andmed

Käesoleva töö empiirilises osas analüüsitakse veebipaneeli *CROss-National Online Survey* (CRONOS) andmeid. CRONOS viidi läbi 2016. aasta Euroopa Sotsiaaluuringu (ESS, i.a) 8. vooru andmekogumise järgselt, kus ESS silmast silma intervjuude lõpus kutsuti vähemalt 18-aastaseid Eesti, Sloveenia ja Suurbritannia (v.a Põhja-Iirimaa) intervjuueeritavaid osalema kuues 20-minutilises veebiuuringus järgneva 12 kuu jooksul (CRONOS, i.a). CRONOSe andmekogumine jäi ajaperioodi detsember 2016 kuni veebruar 2018. Indiviididele, kes märkisid ESS 8. voorus, et nad ei kasuta internetti, pakuti võimalust kasutada veebiuuringus osalemiseks vajalikke tehnoloogilisi vahendeid, st tahvelarvutit koos ligipääsuga internetile. Lisaks õpetati tahvelarvuti kasutamist uuringus osalemise eesmärgil (Villar jt, 2018). CRONOSe andmed on avalikult kättesaadavad ESS koduleheküljelt. CRONOSe paneeli disain ja läbiviimine oli juhitud tsentraalselt eesotsas uurijatega ESS Euroopa Teaduse Infrastruktuuri ERIC peakontoris Londonis, Norra Teadusandmete Keskuses (*Norwegian Centre for Research Data - NSD*), Ljubljana Ülikoolis Sloveenias ja Hollandi uuringutekeskuses CentERdata. Keskse tiimiga tegid koostööd riiklikud meeskonnad, mille Eestis moodustasid Mare Ainsaar ja Indrek Soidla Tartu Ülikoolist (Villar jt, 2018). Pilootprojekt on osa suuremast projektist *Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS)*, mida rahastas Euroopa Liidu teadusuuringute ja innovatsiooni rahastamisprogramm Horisont 2020 (CRONOS, i.a).

CRONOSe paneel on esimene rahvusvaheline tõenäosuslikul valimil põhinev ühtlustatud metodoloogiaga veebipaneeli pilootprojekt, mille eesmärgiks on hinnata veebipaneeli värbamise efektiivsust olemasolevale rahvusvahelisele uuringule baseerudes, võttes arvesse kulusid, valimi esinduslikkust, osalust ja paneelist lahkumist ning seeläbi hinnata andmekvaliteeti (Villar jt 2018). CRONOSes osalejate motiveerimiseks rakendati kõikides osalenud riikides rahalisi motivaatoreid, mis riigiti erinesid. Eesti puhul kasutati motivaatorina 10-euroseid Partner e-kinkekaarte, mida jagati vastajatele üle ühe paneeli laine (*Ibid.*). Tahvelarvutiga respondentidele edastati e-kinkekaart prindituna koos uuringukutsega posti teel.

Käesolevas töös kasutan analüüsiks ESS koduleheküljel avalikult kättesaadavaid CRONOSe andmeid. Andmete analüüsiks on moodustatud avalikult kättesaadavate andmestike baasil ühendandmestik, kuhu on kaasatud:

- integreeritud andmestik kõikide paneeli lainete ja ESS 8. vooru 2016. aasta andmed paneelis osalenute kohta (CROss-National Online Survey panel [CRONOS0-6 ESS8 integrated data file, edition 1.0, round 8], 2018);
- andmestik kontaktiloomiste kohta (CROss-National Online Survey panel [CRONOS Correspondence data, edition 1.0 round 8], 2018);
- administratiivsed andmed (CROss-National Online Survey panel [CRONOS Administrative data, edition 1.0 round 8], 2018);
- paraandmed (CROss-National Online Survey panel [CRONOS Paradata, edition 1.0 round 8], 2018).

Loetletud andmestikud on ühendatud tunnuste “IDNO” ja “Country” järgi vastavalt andmete kasutusjuhendile (Villar jt, 2018:12). Koondandmestik sisaldab ühendatud informatsiooni ESS 8. voorus antud vastuste kohta ja CRONOSe kõikide lainete vastuseid. Lisaks sisaldab koondandmestik individuaalseid andmeid selle kohta, kuidas ja millal toimus indiviidiga kontaktiloomine, üldiseid administratiivseid andmeid nagu tahvelarvuti saamine ning paraandmeid kasutatud seadme tüübi ja teiste tehniliste parameetrite kohta, küsimustele vastamise aja, vastamise staatuse ja viimasena avatud küsimuse kohta.

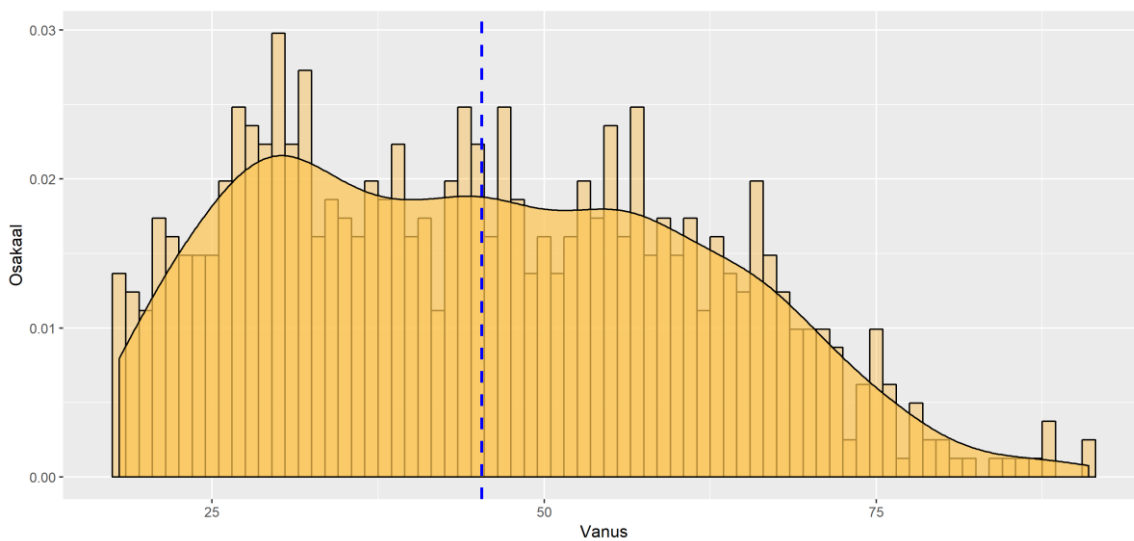
2.2 Valimi kirjeldus

CRONOSe valimisse kuuluvad ESS 8. vooru valimisse kuulunud vähemalt 18-aastased Eesti, Sloveenia ja Suurbritannia (v.a Põhja-Iirimaa) respondendid. ESS 8. vooru valimisse kuulus ka 15-17-aastaseid respondente, kuid CRONOSest arvati nooremad kui 18-aastased respondendid valimist välja (Villar jt, 2018: 4). Ülemist vanusepiiri uuringus osalemiseks ei seatud. Villar jt (2018: 8-9) on esitanud veebipaneelis CRONOS kõikide paneeli lainete vastamismäärad ja osalusmäärad osalenud riikide lõikes, millest tõstaksin esile, et Eesti madalaim osalusmäär (täielike ja osaliste vastuste osakaal kutsututest) oli sissejuhataval paneeli lainel (72%) ja kõrgeim teisel lainel (82%), vastamismäärad (täielike ja osaliste vastuste osakaal koguvälimi küsitluseks sobivatest indiviididest) jäid Eesti puhul kõikidel lainetel 20% ja 25% vahele.

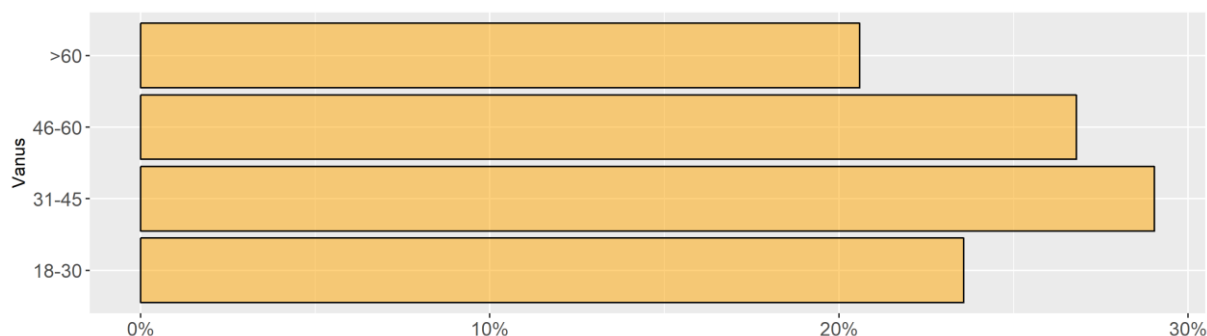
Võttes aluseks töö empiiriliseks osaks moodustatud ühendatud andmestikku kõikide paneeli lainete kohta, sisaldab ühendandmestik andmeid 2437 paneelis osalenud respondendi kohta, kellest 806 olid Eestist, 926 Suurbritanniast ja 705 Sloveeniast. Ühendandmestik koondab andmeid kõikide paneelis osalenute kohta, kes on vähemalt ühel paneeli lainel vastuseid

andnud, kuid ei sisalda andmeid kõikide küsitluseks sobivate indiviidide kohta. Käesolevas töös keskendun Eesti andmetele.

Eesti respondentidest 81,6% vastasid eestikeelsele, 18,4% venekeelsele küsimustikule (keel tulenes ESS 8. voorus intervjuus kasutatud keelest). Eesti respondentidest 43,3% moodustasid mehed ja 56,7% naised. Joonisel 3 on kujutatud Eesti vastajate vanuseline jaotus, millelt on eristatav suurem vastajate arv vanusevahemikus 25-35 eluaastat ning vastanute arvu langus alates 55. eluaastast. Noorim vastaja oli Eestis 18-aastane, vanim 91-aastane, keskmiseks vanuseks 45,3 aastat ja mediaanvanuseks oli 44 aastat. Käesolevast uurimisteemast ja teemaga seotud teoreetilisest taustast lähtudes võib oletada peamiselt kõige nooremate indiviidide ja kõige vanemate indiviidide lahknevust andmekvaliteedi seisukohalt. Seetõttu jagan invidiidid nelja vanuserühma ning respondentide jaotus vanuserühmadesse on esitatud joonisel 4. Suurima vanuserühma Eesti vastajate seas moodustavad 31-45-aastased respondendid (29,0%) ja kõige vähem on respondente, kelle vanus ületab 60. aastat (20,6%).

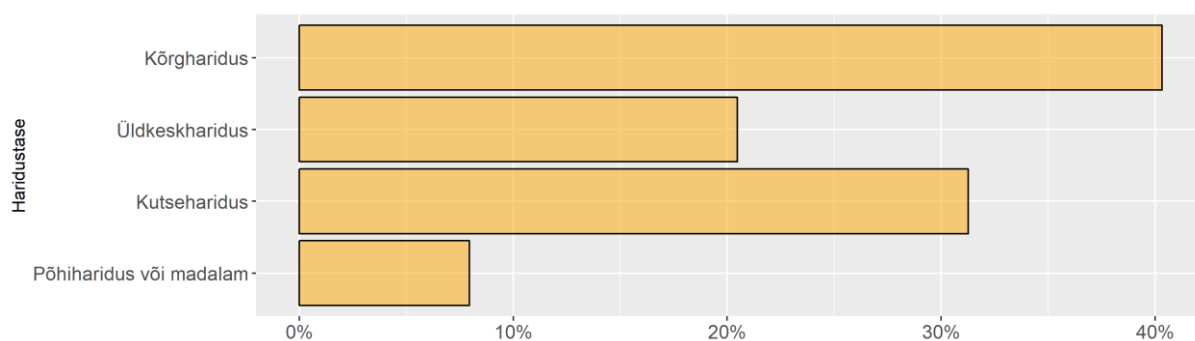


Joonis 3. Eesti vastajate vanuseline jaotus, lisatud keskmine vanus($N=806$)



Joonis 4. Eesti vastajate vanuseline jaotus vanuserühmade järgi (N=806)

Joonisel 5 on esitatud respondentide jaotus haridustaseme järgi. Haridustaseme rühmade selgitus on esitatud peatükis 2.3. Eesti vastanutest suurima rühma (40,3%) moodustasid kõrgharidusega indiviidid. Kõige väiksema rühma moodustasid põhiharidusega või madalama haridustasemega indiviidid (7,9%).



Joonis 5. Eesti respondentide jaotus haridustaseme järgi (N=806)

2.3 Tunnused

- **Vastamise staatus:** kasutatud kõikidel lainetel tunnust $wXdisp_status$, kus X tähistab laine numbrit. **Kutsele mittevastamisena** käsitletakse vastamise mitte alustamist (uuringukutse saanud). Alustamisena käsitletakse variante: katkestamine, kohene täielik vastamine, vastamine pärast pausi.
- **Vastuste variatiivsus:** individuaalne vastuste dispersioon konkreetses küsimuste ploki pideva tunnuseks
- **Küsimusele mittevastamine:** regressioonimudelil dihhotoomne tunnus, kus 1 – vähemalt üks küsimusele mittevastamine („ei oska öelda“, „eelistan mitte vastata“ või vastuse puudumine) küsimustiku jooksul; 0 – küsimusele mittevastamisi ei esine.

Taustakategooria: küsimusele mittevastamist küsimustiku jooksul ei esine. Mitteparameetrilises testis kasutatud pideva tunnusega.

- **Vastamisele kulunud aeg** (*wxtotal_time*): vastamisele kulunud aja tunnustes eristan sissejuhatavat ja esimest lainet teistest järgnevatest lainetest, sest esimese kahe laine puhul polnud tehnilistel põhjustel võimalik mõõta täpset vastamisele kulutatud aega, st oli võimalik eristada küsitluse keskkonda sisenemise ja väljumise ajalist vahet, mis ei mõõda küsimustele vastamiseks kulunud aega. Alates teisest paneeli lainest oli kokku arvatud täpne aeg igale küsimusele vastamiseks. Vastamisele kulunud aeg on teisendatud minutiteks. Vastamisele kulunud aja väärtused, mis ületavad ülemist ja alumist protsentiili, on asendatud ülemise ja alumise protsentiili väärtusega (Yan ja Tourangeau, 2008).
- **Raske vastata:** kui raske oli vastajal küsitluse küsimustest aru saada ja neile vastata – dihhotoomne tunnus, kus 0 on „Üldse mitte raske“, 1 koondab teisi võimalikke vastuseid alates valikust „Vähesel määral raske“. Taustakategooriaks on „Üldse mitte raske“. Ei ole saadaval 5. laine puhul.
- **Pingutus täpsete vastuste andmiseks:** taustakategooriaks on pingutus mõõdukal määral või vähem. Tunnus on ümber kodeeritud dihhotoomseks tunnuseks, kus 1 koondab väärtusi „Väga palju“ ja „Palju“; 0 aga väärtusi „Mõõdukal määral“, „Natuke“, „Üldse mitte“.
- **Nautimine:** taustakategooriaks on vastamise nautimine mõõdukal määral või vähem. Tunnus on ümber kodeeritud dihhotoomseks tunnuseks, kus 1 koondab väärtusi „Väga palju“ ja „Palju“; 0 aga väärtusi „Mõõdukal määral“, „Natuke“, „Üldse mitte“. Ei ole saadaval 5. laine puhul.
- **Sugu:** Mees, taustakategooriaks on naine.
- **Vanus:** kodeeritud nelja vanuserühma 18-30, 31- 45, 46-60, 61 aastat ja vanemad.
- **Haridustase:** taustakategooriaks on põhiharidus või madalam. ESS 8. voorus küsitud kõrgeim omandatud haridustase on ISCED 2011 baasil ümber kodeeritud nelja haridustaseme rühma:
 - Põhiharidus või madalam: alghariduseta (alla 4 klassi); lõpetatud algharidus (4-6 klassi), lõpetatud põhiharidus (7-9 klassi);
 - Kutseharidus: kutseharidus koos põhihariduse omandamisega, kutseharidus põhihariduse baasil kutseharidus põhihariduse baasil (õppekavaga alla 2 aasta või 2 aastat või enam), kutsekeskharidus, kutseharidus keskhariduse baasil, keskeriharidus keskhariduse baasil;

- Üldkeskharidus: üldkeskharidus;
- Kõrgharidus: kutsekõrgkooli või rakenduskõrgkooli diplomiõpe (kuni 2 aastat õpinguid, kuid mitte bakalaureuse kraad), kutsekõrgharidus, rakenduskõrghariduse diplomiõpe või bakalaureus (3-4 aastat õpinguid), ülikooli bakalaureusekraad (3-4 aastat õpinguid), magistrikraad rakenduskõrgkoolist, kutsekõrgkoolist, magistrikraad (3+2, 4+2 või 5+4 süsteemi järgi, sh integreeritud bakalaureuse- ja magistriõpe), enne 1992. aastat alustatud kõrgharidus (diplomeeritud spetsialistiõpe), arstiõpe, doktorikraad (sh kandidaadikraad).
- **Keel:** vene keel, taustakategooriaks eesti keel. Tunnuse näitab, millises keeles ESS 2016 8.vooru intervjuu anti ning millises keeles CRONOSe küsimustikku täideti.
- **Rööprähklus:** taustakategooriaks on ühe asjaga, st vastamisega, tegelemine. Vastamisaegsete tegevuste tunnus on ümber kodeeritud nii, et kui vastaja märkis vähemalt ühe tegevuse, millega ta veel vastamise ajal tegeles, on tunnuse väärtuseks 1, taustakategooriaks on ainult küsimustikule vastamine.
- **Kasutatud seade:** nutitelefon, tahvelarvuti, taustakategooriaks on arvuti. Vastamiseks peamiselt kasutatud seade on saadaval vaid 2.-6. laine puhul.
- **Sage interneti kasutus:** taustakategooriaks harvemini kui iga päev. Interneti kasutussageduse tunnus on ümber kodeeritud binaarseks tunnuseks, kus 1 – iga päev, 0- enamikel päevadel või harvem.
- **Poliitikahuvi:** taustakategooriaks on „Vähe huvitatud“ või „Ei tunne üldse huvi“. Poliitikahuvi tunnus on kodeeritud dihhotoomseks tunnuseks, kus 1 koondab väärtusi „Väga huvitatud“ või „Üsna huvitatud“.
- **Kaks või kolm meeldetuletust:** taustakategooria 0-1 meeldetuletust. Tunnus ümber kodeeritud dihhotoomseks tunnuseks, kus 1 – respondentile edastati kaks või kolm meeldetuletust uuringus osalemiseks; 0 – respondentile edastati 0-1 meeldetuletust.
- **w1exp0/ w3exp0/ w5exp0:** 1., 3. ja 5. laine eksperiment. Respondendid on jagatud kolme rühma, kellele juhiti sissejuhatuseks erinevalt tähelepanu täpse vastamise olulisusele. 2 – täpsuse vastamise olulisuse rõhutamine ja täpsuse kinnituse küsimine; 1 – täpse vastamise olulisuse rõhutamine; 0 – standardsissejuhatuse kuvamine.
- **w2exp0/ w4exp0/ w6exp0:** 2., 4. ja 6. laine eksperiment. Respondendid jagatud kolme rühma, kellele juhiti erinevalt tähelepanu hoolika/hooletu vastamise tagajärgedele. 2 –

hooletu vastamise negatiivsed tagajärjed; 1 – hoolika vastamise positiivsed tagajärjed; 0 – informatsiooni tagajärgede osas ei kuvatud.

2.4 Analüüsimeetodid

Käesolevas uuritakse CRONOSe Eesti paneeliliikmete vastamiskäitumist esmalt kirjeldava statistikana ja sagedusjaotuste abil, seejärel regressioonanalüüsi rakendades. Töös rakendatakse ristlõikelist analüüsi paneeli lainete kaupa, mis võimaldab uurida vastamiskäitumist igal lainel ja seda mõjutavaid tegureid. Ristlõikeline analüüs võimaldab analüüsi kaasata kõiki indiviide, kes konkreetsel lainel vastasid ning see, et individid kõikidel lainetel ei osalenud, ei sea analüüsile piiranguid. Ristlõikelise analüüsi valik ei võimalda aga individuaalsete vastamiskäitumise muustrite uurimist terve paneeli vältel, mida võimaldaks longituudandmete analüüsimeetodid.

Töös kasutatakse erinevate andmekvaliteeti mõjutavate tegurite väljaselgitamiseks erinevaid regressioonanalüüsi mudeleid.

- Kutsele mittevastamist prognoosivate tegurite välja selgitamiseks on koostatud binaarne logistiline regressioon paneeli kõikidel lainetel, sõltuvaks tunnuseks on mudelis dihhotoomne mittevastamise tunnus, kus 1 – ei vastanud ühelegi küsimusele, 0 – alustas vastamist.
- Eesmärgiga välja selgitada paneeli erinevatel lainetel vastamise katkestamise põhjuseid, on läbi viidud binaarne logistiline regressioon. Väikse katkestamiste arvu tõttu on katkestamise uurimiseks andmestik üle viidud pikkformaadile, st ühe indiviidi kõikide lainete andmed ei ole enam ühel real, vaid ühe indiviidi kohta on kokku seitse andmerida (seitse lainet). Regressioonis on sõltuvaks tunnuseks kõikide indiviidide kõikide lainete ülene dihhotoomne tunnus, kus 1 esindab katkestamist, 0 täielikult vastamist.
- Täiendavalt on uuritud katkestamist küsimustiku jooksul. Katkestamise uurimine ajas võimaldab hinnata katkestamist prognoosivate tegurite mõju küsimustiku vältel ja anda mõjuteguritele täiendavat tõlgenduslikku lisandväärtust. Sündmuse toimumise prognoosimiseks ajas erinevate faktortunnuste kaudu on rakendatud sündmuspõhiste andmete analüüsimeetodit - Coxi regressioonimudelit, mis võimaldab uurida erinevate faktortunnuste abil sündmuse toimumist või selle „üleelamist“ ajas (Aalen, 2008). Lisaks on tulemuste hindamisel kasutatud Kaplani-Meieri hinnanguid sündmuse „üleelamise“

tõenäosusele ja kumulatiivsetele riskidele sündmuse toimumiseks (*Ibid.*). Töös on sündmusena käsitletud vastamise katkestamist (esimest katkestamist paneelis), sündmuse toimumise ajana viimasena avatud küsimuse numbrit (millel katkestamise toimus). CRONOSe üks läbivaid eksperimente oli seotud küsimuste järjekorraga, mida eri rühmades varieeriti. Individuaalsete küsimuse numbrite arvestamisel on arvesse võetud individuaalset eksperimentidest tulenevat küsimuse kuvamise järjekorda (ja seda, kui eksperimendi raames mingitele küsimustele vastama ei pidanud). Arvesse pole võetud üksikuid küsimusi, mida teatud vastuste korral möödasuunamise tulemusel vahele jäeti.

- Küsimusele mittevastamist prognoosivate tunnuste uurimiseks on koostatud binaarse logistilise regressiooni mudel, kus sõltuvaks tunnuseks on dihhotoomne vähemalt ühele küsimusele mittevastamine, nii vastamata jätmised kui ka mittesisulised vastused (1 – vähemalt üks küsimusele mittevastamine, 0 – kõik küsimused vastatud). Täiendavalt on uuritud küsimusele mittevastamiste arvu (pideva tunnusena) küsimustiku jooksul Kruskali-Wallise mitteparameetrilise testi abil, millega autor selgitab, kas küsimusele mittevastamise variatiivsus on tunnuse tüübi või küsimuse teemaploki lõikes erinev. Mitteparameetrilise testi rakendamine osutus vajalikuks seoses erindite esinemisega küsimusele mittevastamise jaotuses. Kruskali-Wallise test võimaldab võrrelda üksteisest sõltumatuid kogumeid keskmiste astakute kaudu (Kruskal ja Wallis, 1952).
- Joonvastamise tuvastamiseks CRONOSe andmetes on välja valitud 17 küsimuste plokki üle kõikide lainete, kus vähemalt viiel järjestikusel küsimusel on sama vähemalt 5-punktiline vastuste skaala ning on arvutatud individuaalne variatiivsus küsimuste ploki lõikes. Antud variatiivsust mõjutavate tegurite väljaselgitamiseks on koostatud lineaarsete regressioonimudelite võrdlus.
- Küsimustikele vastamiseks kulunud aega (äärmised protsentiilid tasandatud) prognoosivate tegurite leidmiseks on koostatud kõikide lainete lineaarse regressiooni mudelid, kus sõltuvaks tunnuseks on logaritmi vastamisele kulunud ajast positiivse asümmeeria esinemise tõttu.
- Küsimustiku lõpuosas uuriti respondentidelt ka nende hinnanguid uuringus osalemise kogemusele. Kogemuse hindamiseks küsiti respondentidelt nende hinnangut raskusele küsimuste mõistmisel ja vastamisel, hinnangut pingutusele täpsete vastuste andmisel ning hinnangut, kui võrd respondentid vastamist nautisid. Et välja selgitada küsimustiku täitmise protsessi tajumist mõjutavaid tegureid ning seeläbi anda andmekvaliteeti mõjutavatele teguritele täiendavaid selgitusi, on koostatud kõikide vastamiskogemust käsitlevate tunnuste

kohta binaarse logistilise regressiooni mudelid. Sõltuvate tunnuste ümberkodeerimine on selgitatud peatükis 2.3.

Binaarse logistilise regressiooni mudelites, kus on võrreldud erinevaid paneeli laineid, esitatakse mudelite võrdleva tõlgendamise eesmärgil keskmised piirefektid (*average marginal effect* - AME), sest regressioonikordajad peegeldavad logistilises regressioonimudelis varjatud heterogeensust (Mood, 2010). Mudeli puhul, kus on ühte mudelisse kaasatud korraga kõik lained, on esitatud regressioonikordajad ja nende eksponendid mugavama tõlgendamise eesmärgil (B , $\exp(B)$).

Töö empiirilise osa analüüs on teostatud analüüsitarkvaraga R, lisaks on täiendavalt andmete analüüsiks ettevalmistamisel kasutatud tabelitöötlusprogrammi MS Office Excel.

3 Tulemused

3.1 Ülevaade andmekvaliteedist

3.1.1 Interneti mittekasutajad

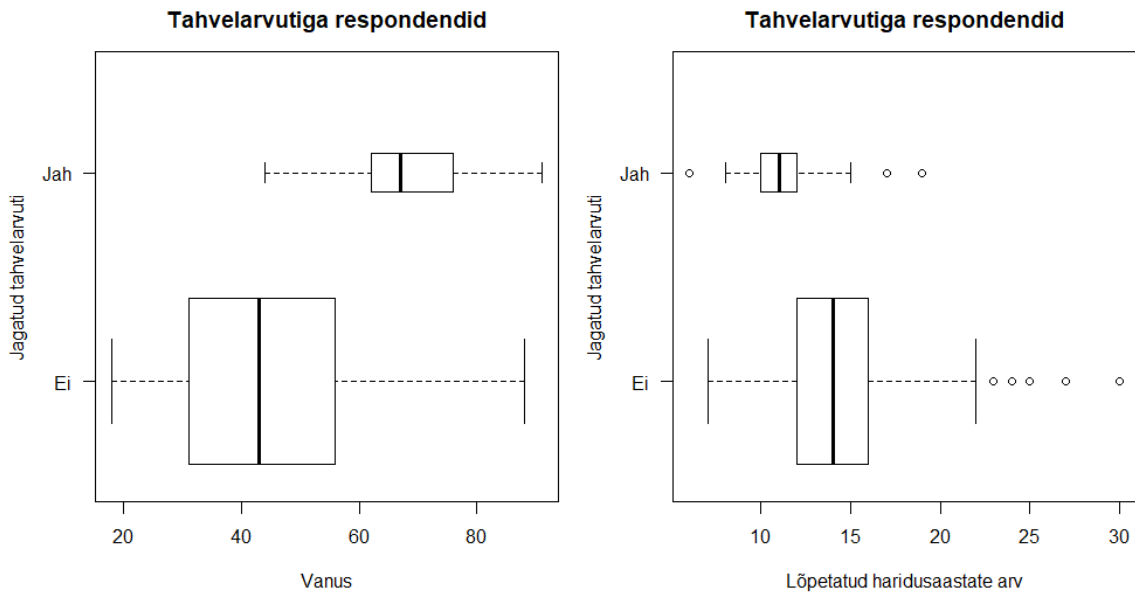
Internetile piiratud ligipääsuga indiviide, kellele jagati uuringus osalemiseks tahvelarvuti koos internetiühendusega, oli Eestis kokku 39. Tabelis 1 on esitatud tahvelarvuti saanud respondentide jaotus vastamise staatuse järgi paneeli kõikidel lainetel. Tabelist selgub, et kõnealuste indiviidide seas esines katkestamisi enim paneeli sissejuhataval lainel (4).

Tabel 1. Uurimisrühma poolt võimaldatud tahvelarvutiga indiviidide jaotus absoluutarvudes vastamise staatuse järgi kõikidel paneeli lainetel

	<i>Laine number</i>						
	0	1	2	3	4	5	6
<i>Vastamist alustanud indiviidide arv</i>	39	35	32	33	33	33	30
<i>Täielik vastamine ühes osas</i>	32	31	28	29	31	30	28
<i>Täielik vastamine pärast pausi</i>	3	3	4	4	2	3	1
<i>Katkestamine</i>	4	1	0	0	0	0	1

Jagatud tahvelarvutite jaotusest haridusaastate arvu ja vanuse järgi (joonis 6) on näha, et internetile ligipääsu puudumise tõttu võimaldati tehnoloogilised vahendid uuringus osalemiseks eelkõige madalama lõpetatud haridusaastate arvuga ning vanematele indiviididele. Kruskali-Wallise testid kinnitavad mõlema tunnuse alusel erisusi tahvelarvutite kasutuses: lõpetatud haridusaastate arvu puhul on hii-ruudu väärtuseks 99,36 (vabadusastmete arvu 21 korral, $p < 0,001$), vanuse puhul 249,06 (vabadusastmete arvu 70 korral, $p < 0,001$). Vene või eesti keeles vastanud indiviidide seas oli tahvelarvuti saanud respondentide osakaal oma keelerühmas ühetaoline (Kruskali-Wallise testi hii-ruut on 0,24, $p = 0,623$). Ka naiste ja meeste vahel puudusid erisused tahvelarvutite jaotuses (Kruskali-Wallise testi hii-ruut on 0,60, $p = 0,440$).

Seega CRONOSes osalemiseks võimaldati tehnoloogilisi vahendeid kokku 39 indiviidile, keda iseloomustas väiksem lõpetatud haridusaastate arv ja kõrgem vanus. Jagatud tahvelarvutiga respondendid vastasid suures osas küsimustikule täielikult ning kuundaks laineks olid paneelis veel vastamas 30 tahvelarvuti saanud respondenti.

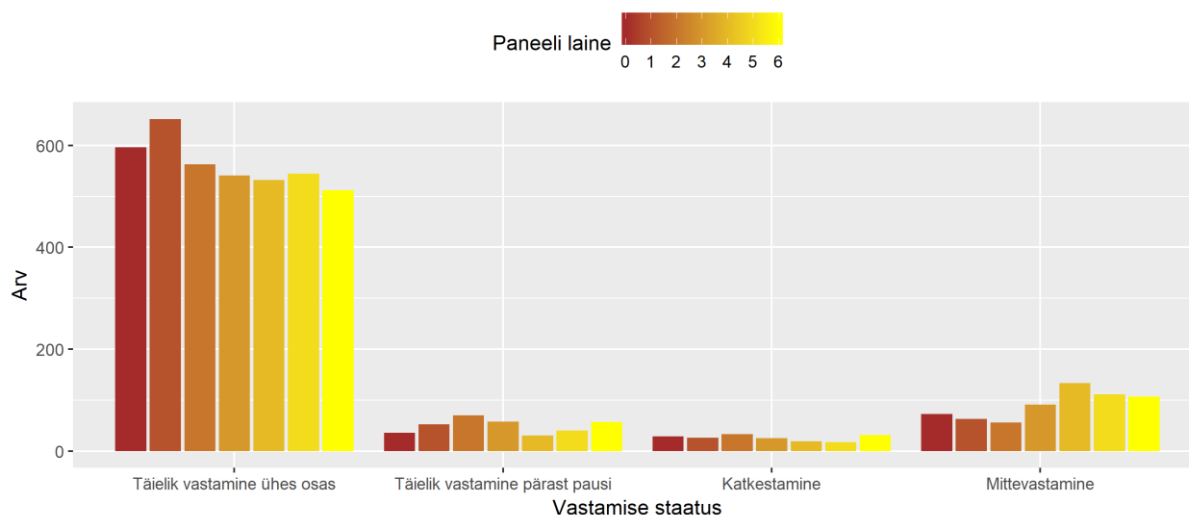


Joonis 6. Tahvelarvuti saanud respondentide arvu karpdiagrammid lõpetatud haridusaastate arvu ja vanusegrupi järgi ($N = 39$)

3.1.2 Kutsele mittevastamine

Üheks uuringust väljajäämise põhjuseks on ebaedu kontaktiloomisel potentsiaalse respondentiga. Eesti andmetel selgub, et CRONOSe paneeli valimisse kuulunutega kontakti loomine oli väga edukas, sest vaid üksikutel juhtudel ei osutunud uuringu käigus indiviidi e-mail aktiivseks või ei õnnestunud kontakti luua (Vt täpsemalt lisa 1). Edukus kontaktiloomisel tuleb asjaolust, et enne uuringu algust võeti kõigi valimisse kuulunutega ühendust, et küsida neilt luba meiliaadresside saatmiseks Norra Teadusandmete Keskusesse, mille käigus oli võimalik vigased ja mittetoimivad meiliaadressid tuvastada ning parandada.

Joonisel 7 on esitatud kõikide kutse saanute jaotus absoluutarvudes vastamise staatuse järgi kõikidel paneeli lainetel. Jooniselt selgub, et küsitlusele vastamine oli üle paneeli lainete kahaneva trendiga, kusjuures märkimisväärne langus toimus pärast paneeli esimest lainet. Reeglina (keskmiselt 92% andmestikku kuulunud paneeli lainesse kutsututest) vastati küsimustikule ühes osas. Kõige enam oli ühes osas antud täielikke vastuseid paneeli neljandal lainel (94,7% täielikest vastustest) ja kõige enam täielikke vastamisi pärast pausi oli paneeli teisel lainel (11,1% täielikest vastustest). Kutsele mittevastamine oli kõrgem veebipaneeli teises pooles, olles kõrgeim paneeli neljandal lainel (18,6%), millele järgnes viies laine, kus vastamist mitte alustanud individid moodustasid 15,6% kutse saanutest.



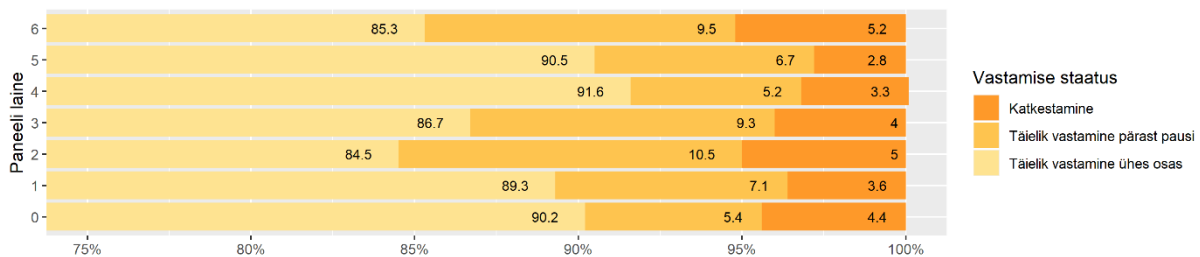
Joonis 7. Kutsele vastamine ja mittevastamine kõikidel paneeli lainetel

Eesti andmestikku kuulunud respondentidest alustasid vastamist sissejuhataval lainel 661, esimesel lainel 730, teisel lainel 666, kolmandal lainel 624, neljandal lainel 581, viiendal lainel 601 ja kuuendal lainel 600 indiviidi. Kõrgeima osakaalu vastamist alustanud indiviidide seas moodustas täielike vastuste andmine ühes sessioonis neljandal (91,6%), seejärel viiendal lainel (90,5%) (vt joonis 7). Kõige vähem anti täielikke vastuseid ühes osas proportsionaalselt paneeli teisel lainel (84,5%). Kõrgeima osakaalu moodustas täielik vastamine pärast pausi teisel (10,5%), seejärel kuuendal lainel (9,5%). Osalise vastamise osakaal kõikidest vastamist alustanud respondentidest oli suurim paneeli kuuendal lainel (5,2%), millele järgnes paneeli teine laine 5%-ga.

Eesti andmestikku kuulub 30 indiviidi, kes vastasid sissejuhataval lainel küsimustikule täielikult või osaliselt ning järgnevatel lainetel ei osalenud. Nendest neljale ei edastatud esimesel lainel uurimiskutset, kaks indiviidi ei saanud kutset kätte seoses mitteaktiivse e-maili aadressiga ning ühe indiviidiga ei saadud kontakti. 52 indiviidi ei võtnud enam paneelist osa pärast esimest paneeli lainet, kellest 15 indiviidi olid esimese laine katkestajad (esimesel lainel oli katkestajaid kokku 26), mis tähendab, et üle poole indiviididest, kes esimesel lainel katkestasid, ei jätkanud enam paneelis. Pärast teist lainet paneelis mitte osalejaid oli kokku 9, kes kõik olid ka teise laine küsimustele vastamise katkestanud. Pärast kolmandat lainet ei alustanud küsimustikele vastamist enam 25 indiviidi, kellest 9 olid kolmanda laine katkestajad. Alates viiendast lainest jäi kõrvale 14 indiviidi, kes neljandal lainel vastasid küsimustikule

täielikult või osaliselt (6 indiviidi). Kuuendal lainel ei alustanud küsimustiku täitmist 46 indiviidi, kes olid vastasid viiendal lainel täielikult või osaliselt (9 indiviidi).

CRONOSe andmetele tuginedes võib öelda, et keskel läbi üheksa respondenti kümnest vastasid küsitlusele ühes osas. Indiviididega kontaktloomine oli CRONOses väga edukas, kuid kutsele mittevastamine suurenes paneeli teises pooles. Lisaks on oluline, et suur osa küsitlusele vastamise katkestanutest ei jätkanud järgmisel lainel enam uuringus osalemist.



Joonis 8. Vastamist alustanute protsentjaotus vastamise staatuse järgi

3.1.3 Vastamise katkestamine

Vastamist alustanud respondentide seas esines osalist vastamist katkestamise tõttu kõikidel paneeli lainetel üldiselt harva (joonis 8, täpsemalt lisa 1), moodustades keskmiselt 4% paneeli lainel vastamist alustanutest. Katkestanute osakaal kõikidest küsitlusele vastamise alustanutest jäi 2,8% ja 5,2% vahele, olles kõrgeim paneeli kuuendal, seejärel teisel lainel. Vastamise katkestamist esines paneeli kõikide lainete sissejuhatavatel lehekülgedel ja nõusoleku küsimise leheküljel. Katkestamine sissejuhatavatel lehekülgedel oli arvukas paneeli sissejuhataval lainel, kus 17 indiviidi ei jõudnud küsimusteni. Järgnevatel lainetel katkestamine sissejuhatavatel lehekülgedel nii arvukas polnud, järgnes viies laine kaheksa katkestamisega sissejuhataval leheküljel, millest kolm moodustasid indiviidid, kellelt küsiti eksperimendi korras täiendavat kinnitust täpsete vastuste andmiseks uuringus.

Jättes välja vastamise katkestamised sissejuhatavatel lehekülgedel, oli enim katkestamisi teisel (28) ja kolmandal (27), seejärel kuuendal paneeli lainel (24). Joonisel 9 on kujutatud vastamise katkestamist küsimustiku jooksul (sissejuhatavad leheküljed on välja jäetud). Enamikel lainetel iseloomustas katkestamist suurem esinemissagedus küsimustiku esimeses pooles, erandiks vaid teine laine, kus katkestamiste arv küsimustiku teises pooles ületas esimest poolt. Paneeli lainete võrdlusest on märgata katkestamisi esimese ja viimase küsimuse juures, millest viimane võib

olla tulenenud mittetahtlikust enneaegsest küsimustiku sulgemisest. Kõikide paneeli lainete puhul on üldiselt iseloomulik katkestamiste suhteliselt ühtlane jaotus üle terve küsimustiku ning kõigil lainetel esines maksimaalselt kolm katkestamist ühel küsimusel. Teise laine maksimaalne katkestamiste arv oli küsimuste juures, mis juhatasid sisse järgmist küsimuste plokki, käsitledes esimesel juhul vaateid „hea“ ja „halva“ laste kasvatamise osas, teisel juhul arvamust selle kohta, milline on lastevanemate elu tänapäeval. Kolmandal lainel esines katkestamisi enim küsimuse juures, milles küsiti, kas erinevused kvalifikatsioonitasemetes, mis inimesed saavutavad Eestis, on ebaõiglaselt väikesed, õiglasel või ebaõiglaselt suured. Neljandal ja kuuendal lainel katkestati vastamine enim esimesel küsimusel, käsitledes esimesel juhul enesetunde teemat, teisel juhul hinnanguid sissetulekute õiglusele Eestis.

Eesti andmestikku kuulunud indiviidide seas ei esinenud juhtu, kus üks indiviid oleks katkestanud vastamise kõikidel paneeli lainetel. Respondente, kes katkestasid kogu paneeli jooksul vähemalt ühel lainel küsimustikule vastamise (k.a sissejuhatavatel lehekülgedel), oli kokku 143, kellest 5 olid jagatud tahvelarvutitega respondendid. Eesti keeles vastanute seas oli katkestajaid 18,2%, vene keeles vastanute seas 15,5%. Vähemalt ühel lainel vastamise katkestanud indiviidide seas oli enim kahe noorema vanuserühma esindajaid (18-30.a 51 indiviidi ja 31-45.a samuti 51 indiviidi). Vähemalt ühel lainel vastamise katkestanute seas oli 72 meest ja 71 naist.

CRONOSes esines katkestamisi üldiselt vähe. Esmene katkestamine toimus küsimustiku sissejuhatavatel lehekülgedel ning seejärel üldiselt rohkem küsimustiku esimeses pooles. Lisaks võib täheldada, et vastamist on katkestatud näiteks järgmist küsimuste plokki alustaval küsimusel ja küsimustiku viimasel küsimusel, mis võib tuleneda küsimustiku enneaegsest sulgemisest. Vähemalt ühel lainel katkestanud indiviidide seas oli enim 18-45-aastaseid respondente.

3.1.4 Küsimusele mittevastamine

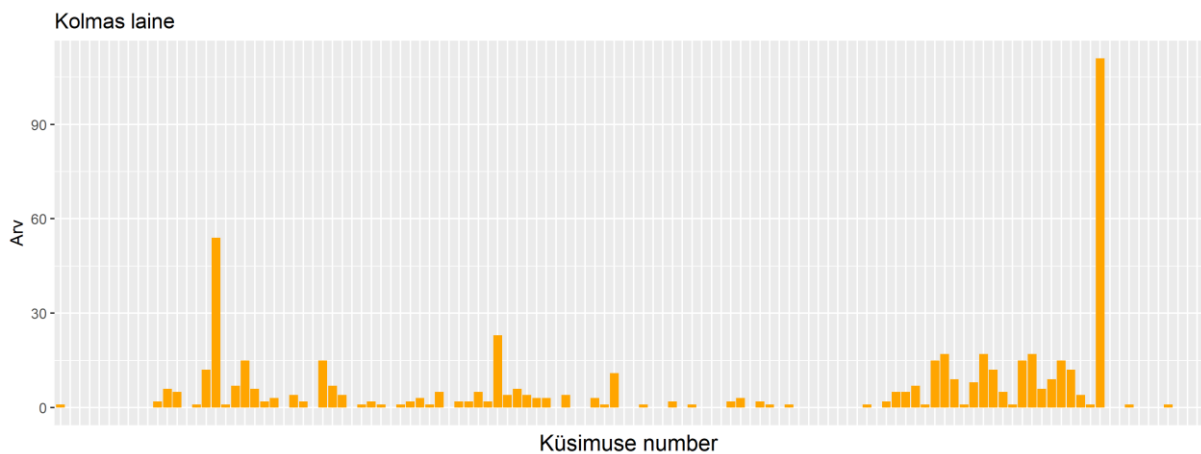
Küsimusele mittevastamise alla on loetud vastamisest keeldumised ja mittesisuliste vastuste andmine („Eelistan mitte vastata“, „Ei oska öelda“ ja vastuse puudumine). Küsimusele mittevastamist esines sissejuhataval, esimesel, kolmandal, viiendal ja kuuendal paneeli lainel. Sissejuhataval lainel esines kaks küsimust, millel mittevastamist esines. Üks neist oli küsimus, kus küsiti milliseid seadmeid vastaja internetis käimiseks tavaliselt kasutab (seitse mittevastamist, 1,1% vastustest). Antud küsimus oli jätkuküsimuseks eelmisele, milles küsiti, millal vastaja viimati internetti kasutas ning üheks võimalikuks vastusevariandiks oli „Pole enne tänast kunagi kasutanud“, mille valis vastuseks 22 indiviidi. Samas puudus jätkuküsimuses sobilik vastusevariant, mistõttu seitse indiviidi otsustasid küsimusele mittevastamise kasuks. Teine küsimus, millel sissejuhatavas laines mittevastamist esines, oli mitme võimaliku vastusevariandiga küsimus küsitluse liikide kohta, milles vastaja varem oli osalenud (kolm mittevastamist, 0,5% vastustest).

Paneeli esimesel lainel oli vaid üks küsimus, millel esines mittevastamist. Küsimus käsitles sotsiaalmeedia aspekte, mida vastaja usaldab ja milliseid aspekte ei usalda. Laine ainsa avatud küsimusena oli võimalik avaldada pikemalt arvamust, kuid esines ka arvukalt mittevastamisi (239), moodustades 33,8% kõikidest vastustest. Lisaks esines arvukalt sõnalisi väljendusi ka vastustena väljendamaks sama mõtet nagu „ei oska öelda“, „ei oska vastata“ jne.

Paneeli kolmandal lainel esines küsimusele mittevastamist 74 küsimuse puhul. Küsimusele mittevastamist esines kolmandal lainel kokku 574 juhul, millest 62,5% moodustasid vastused „Ei oska öelda“ ja 9,9% „Eelistan mitte vastata“ ning 27,5% jäeti vastuskoht tühjaks. Joonisel 10 on kujutatud küsimusele mittevastamise sagedust kolmanda laine küsimustiku jooksul. Kõikide küsimuste keskmiseks mittevastamiste arvuks oli 4,86 standardhälbe 11,9 ning maksimaalse väärtusega 111. Maksimaalse mittevastamiste arvuga küsimus oli seotud eelnevalt vastatud küsimuste ploki, kus vastaja pidi ette kujutama Eesti valitsuse hariduse edendamise programmi ja erinevaid ühekordse toetuse taotlejaid ning hindama, kui palju iga taotleja raha peaks saama. Antud küsimuste plokk eristub ka joonisel 9 tihedalt koondunud arvukamate mittevastamistega. Maksimaalse mittevastamisega küsimuses küsiti tagasivaatavalt, kui palju kokku oli võimalik kandidaatidele raha eraldada. Küsimuse eesmärgiks oli mõõta, kuivõrd ja kuidas olid respondendid eelnevatest keerukatest küsimustest aru saanud. Kõrge küsimusele

mittevastamise arv võis tuleneda eelnevate küsimuste või konkreetse tagasivaatava küsimuse keerukusest.

Joonisel 10 on näha ka küsimustiku algusepoole üht rohkemate mittevastamistega küsimuste plokki, milles käsitleti õiglust ühiskonnas. Antud küsimuste plokki iseloomustab eelnevalt esitatud küsimusele tagasivaatav iseloom, eeldades eelnevalt antud vastuse mõtestamist. Kõige enam oli mittevastamisi antud kõnealusel küsimuste plokkis avatud küsimusele, kus küsiti, kuidas mõistab vastaja eelnevalt esitatud küsimuses „kvalifikatsioonitaseme“ mõistet.



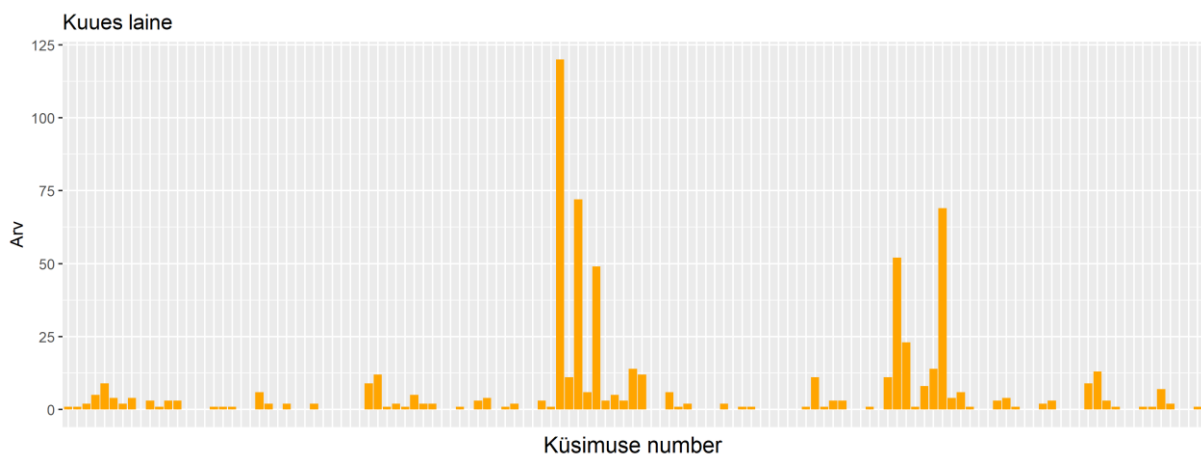
Joonis 10. Küsimusele mittevastamine kolmanda paneeli laine küsimustiku jooksul (N=574)

Viienda laine küsimusele mittevastamiste koguarvuks oli 319, millest 98,4% moodustasid tühjad vastused. Mittevastamiste suur arv kujunes kahe avatud religiooniteemalise küsimuse arvelt, milles uuriti, kuidas mõistis vastaja neile eelnenud küsimuses mõisteid „religioon“ (124 mittevastamist) ja „denominatsioon“ (152 mittevastamist).

Paneeli kuuendal lainel esines sarnaselt kolmanda lainega arvukalt küsimusele mittevastamisi. Kokku esines kõikide küsimuste jooksul 688 mittevastamist, millest 60,9% moodustasid vastused „Ei oska öelda“, 7,4% „Eelistan mitte vastata“ ja 31,8% juhtudest jäeti vastus andmata. Mittevastamisi esines 80 küsimusel. Kuuenda laine keskmine mittevastamiste arv küsimuse kohta oli 5,5, standardhälbeks oli 15,1 ning maksimaalseks väärtuseks 120. Enim mittevastamisi esines küsimusel, mis käsitles kurjategijate kohtlemist kohtute poolt, olles ka üks eksperimendi küsimus, kus testgrupile näidati vastusevarianti „Ei oska öelda“ kohe koos põhivastustega ning juhiti sellele ka küsimuse sõnastuses kohe tähelepanu. 120 mittevastamist olid just sellisele küsimuse sõnastusele, samas kontrollgrupile kuvatud küsimusel esines vaid

11 mittevastamise juhtu. Sama tendents esines ka teiste kõnealuse eksperimendi küsimustele vastamisel. Joonis 11 illustreerib küsimustele mittevastamist kuuenda laine küsimustiku jooksul, kus kõnealune küsimuste plokk keskel üle ühe küsimuse eristub (küsimused testgrupile ja kontrollgrupile on näidatud kõrvuti).

Joonisel 11 eristub selgelt ka teine grupp küsimusi seoses kõrgema küsimusele mittevastamise arvuga. Kaks küsimust, milles kuuendal lainel palju mittevastamisi esines, olid avatud küsimused kõrgeima haridustaseme omandamise kuu (52, 9,0% vastustest) ja omandatud kvalifikatsiooni kohta (69, 11,9% vastustest). Kõrgeima haridustaseme omandamise kuu puhul jagati respondendid juhuslikult kolme rühma, kellest kaks moodustasid testrühmad ja üks kontrollrühm. Testrühmadele kuvati pärast küsimuse teksti ja enne vastust täpsete vastuste andmise meeldetuletus, ühel juhul positiivse, teisel juhul negatiivse sõnastusega. Kõrgeima haridustaseme omandamise kuu küsimus sai enim mittevastamisi kontrollrühmalt, seejärel negatiivse sõnastusega testrühmalt ning kõige vähem positiivse sõnastusega testrühmalt. Kõrgeima haridustaseme omandamise aasta küsimuses oli mittevastamisi poole vähem (23, 4,0% vastustest), neid oli enim positiivse sõnastusega testrühmal, seejärel kontrollrühmal.



Joonis 11. Küsimusele mittevastamine kuuenda paneeli laine küsimustiku jooksul (N=688)

Indiviide, kes jätsid vähemalt ühele küsimusele vastamata (keeldusid või andsid mittesisulise vastuse), oli sissejuhataval lainel 10, esimesel 239, kolmandal 212, viiendal 138 ja kuuendal lainel 172 indiviidi. Kõige enam küsimusele mittevastamisi ühe respondendi kohta oli paneeli 6. lainel (22 küsimusele mittevastamine), seejärel kolmandal lainel (18 küsimusele) (tabel 2). Keskmiselt kõrgeim küsimusele mittevastamine oli 3. lainel, mil 5% küsimusele mittevastamistest ühe indiviidi kohta ületas 5 mittevastamist küsimustiku jooksul. Kolmandal

lainel oli keskmiselt 0,91 ning kuuendal lainel keskmiselt 0,77 küsimusele mittevastamist ühe indiviidi kohta. Tabelisse 2 on kaasatud individid, kes vastasid vähemalt ühele küsimusele, seega on tabelis ka need individid, kes 3. või 6. lainel katkestasid, kuid eelnevalt võisid anda mittesisulisi vastuseid või vastamisest keelduda.

Tabel 2. Küsimusele mittevastamine ühe indiviidi kohta CRONOSe veebipaneeli sissejuhataval, 1., 3., 5. ja 6. lainel

	<i>0. laine</i>	<i>1. laine</i>	<i>3. laine</i>	<i>5. laine</i>	<i>6.laine</i>
<i>Vähemalt ühe vastusega indiviidide arv</i>	<i>641</i>	<i>715</i>	<i>622</i>	<i>591</i>	<i>589</i>
<i>sh. küsimustikule täielikult vastajaid (%)</i>	<i>98,6</i>	<i>98,5</i>	<i>96,3</i>	<i>98,7</i>	<i>96,6</i>
<i>Osakaal alustanutest</i>	<i>97,0</i>	<i>98,0</i>	<i>99,7</i>	<i>98,3</i>	<i>98,2</i>
<i>Küsimusele mittevastamine</i>					
<i>Vähemalt ühe mittevastamisega: indiviidide arv (N)</i>	<i>10</i>	<i>239</i>	<i>212</i>	<i>138</i>	<i>172</i>
<i>Osakaal vähemalt ühe vastusega indiviididest (%)</i>	<i>1,6</i>	<i>33,4</i>	<i>34,1</i>	<i>23,4</i>	<i>29,2</i>
<i>Katkestajate osakaal küsimusele mitte vastanud indiviididest (%)</i>	<i>0,0</i>	<i>1,3</i>	<i>5,7</i>	<i>0,7</i>	<i>2,9</i>
<i>Keskmine mittevastatud küsimuste arv indiviidi kohta</i>	<i>0,02</i>	<i>0,3</i>	<i>0,9</i>	<i>0,5</i>	<i>0,8</i>
<i>Standardhälve</i>	<i>0,1</i>	<i>0,5</i>	<i>2,1</i>	<i>0,9</i>	<i>2,0</i>
<i>Mediaan</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>
<i>0,95 kvantiil</i>	<i>0,0</i>	<i>1,0</i>	<i>5,0</i>	<i>2,0</i>	<i>4,0</i>
<i>0,99 kvantiil</i>	<i>1,0</i>	<i>1,0</i>	<i>9,8</i>	<i>3,0</i>	<i>9,0</i>
<i>Maksimaalne mittevastatud küsimuste arv indiviidi kohta</i>	<i>1</i>	<i>1</i>	<i>18</i>	<i>3</i>	<i>22</i>

Küsimusele mittevastamise esmasest ülevaatest selgus, et küsimusele mittevastamist (nii vastamata jätmist kui mittesisulisi vastuseid) esines arvukalt CRONOSe kolmandal ja kuuendal lainel. Küsimusele mittevastamist iseloomustab koondumine peamiselt küsimustele, mis eeldavad suuremat eelnevate küsimuste või antud vastuste mõtestamist (on tagasivaatava ja interpreteeriva iseloomuga) ning peamiselt avatud küsimustele. Avatud küsimuste puhul selgus, et küsimusele mittevastamist esines rohkelt nii sõnalist arvamust või selgitust kui täpset fakti küsivate küsimuste puhul. Enim mittevastamisi ühe indiviidi kohta esines samuti kuuendal ja seejärel kolmandal lainel.

3.1.5 Joonvastamine

Joonvastamise tuvastamiseks on paneeli igal lainel valitud välja küsimuste plokid (1-3 igal lainel), kuhu kuuluvatel vähemalt viiel järjestikusel küsimusel on sama vähemalt viie

vastusevariandiga skaala. Sissejuhataval lainel oli selliseid küsimuste plokkide kolm, esimesel lainel kaks, teisel lainel kaks, kolmandal lainel üks, neljandal lainel kolm, viiendal lainel kolm ning kuuendal lainel kaks. Indiviidid, kellel on küsimuste komplektis puuduvad väärtusi, ei ole analüüsist välja arvatud, sest vastasel juhul ei oleks pikemates küsimuste plokkides vaatlusi. Küsimustikud on leitavad CRONOSe kodulehel dokumentatsiooni plokis (CRONOS, i.a).

Tabelis 3 on esitatud iga valitud küsimuste komplekti kohta individuaalsete vastuste variatiivsuste keskmine, standardhälve, maksimaalne variatiivsus ning täielike joonvastajate (valitud täpselt samad vastusevariandid) arv. Individuaalne variatiivsus on arvatud küsimuse komplektis antud vastuste dispersioonina (keskmise ruuthälvena). Joonvastamist uuritakse dispersiooni abil eelkõige seetõttu, et täielike joonvastajate arv CRONOSe eri lainetel oli üldiselt väike. Minimaalne variatiivsus on tabelist välja jäetud, sest see on praktiliselt kõigis küsimuste komplektides 0.

Tabelist 3 selgub, et kõige väiksem vastuste variatiivsus esines viienda laine kolmandas küsimuste komplektis, milles uuriti, mil määral vastaja tunneb muret erinevate inimeste ja ühiskonnagruppide eluolu pärast, mõõdetud 5-punktilisel skaalal (1 – „Väga palju“; 5 – „Üldse mitte“). Madala vastuste variatiivsusega paistis silma ka sissejuhatava laine kolmas ja neljanda laine kolmas küsimuste komplekt, kus uuriti mõlemal juhul respondendi vajadust kognitiivsetele ülesannetele ja seisukohtade väljendamisele, mõõtmisaskaalaks oli 5-punktiline nõusolekaskaala (1 – „Täielikult nõustun“; 5 – „Üldse ei nõustu“). Kuna loetletud küsimuste plokkidesse kuuluvad küsimused on oma sisult loogiliselt samasuunalised, siis võib joonvastamine antud küsimuste puhul olla loomulik. Lisaks on tabelis 3 esile tõstetud ka esimese laine esimest küsimuste komplekti, milles uuriti respondendi nõusolekut erinevate tööd puudutavate väidetega 5-punktilisel skaalal (1 – „Täiesti nõus“; 5 – „Üldse ei ole nõus“).

Kõikide küsimuste puhul, mille vastuste variatiivsus on üle kolme ühiku, on tegemist pika vastusskaalaga (10- või 11-punktilise skaalaga) küsimustega ja annab seeläbi võimaluse suuremaks vastuste hajuvuseks. Samas võiks esile tuua sissejuhatava laine esimest ja teise laine esimest küsimuste komplekti, mille vastuste variatiivsus on vastavalt 1,16 ja 1,18 ühikut, on mõlemad 5-punktilise skaalaga ning sisaldavad küsimusi, mis on sisult loogiliselt vastupidised, suurendades seega oodatult vastuste variatiivsust. Kõrgemate täielike joonvastajate arvudega küsimuste plokkid olid teise laine teine küsimuste komplekt ning kuuenda laine teine küsimuste komplekt.

Tabel 3. Vastuste variatiivsus sama vastusskaalaga järjestikuste küsimuste plokkides

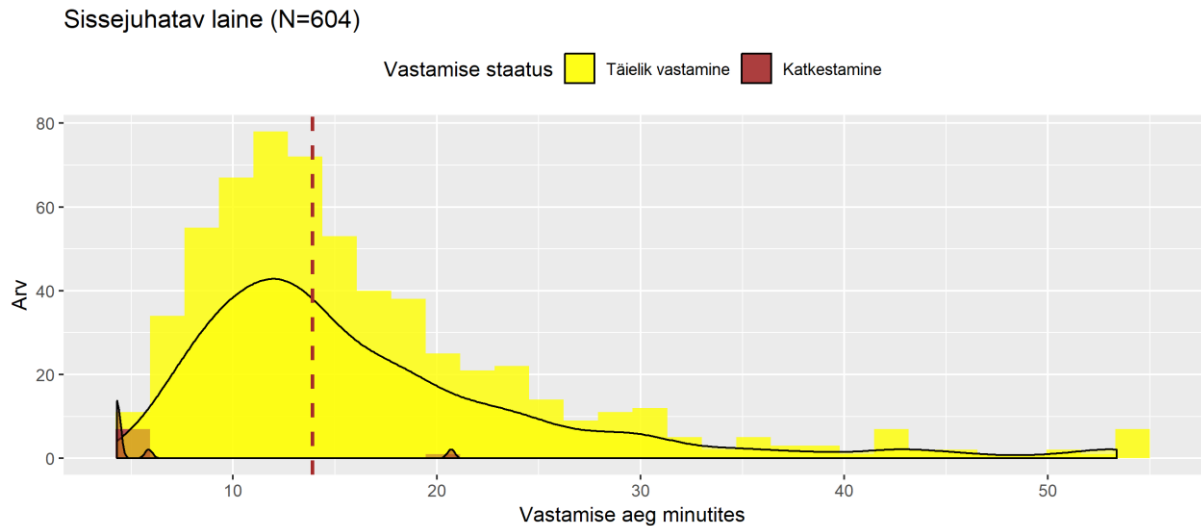
<i>Laine</i>	<i>Küsimuste komplekti number</i>	<i>Küsimused</i>	<i>Valdkond</i>	<i>Keskmine variatiivsus</i>	<i>Keskmise variatiivsuse standardhälve</i>	<i>Maksimaalne variatiivsus</i>	<i>Täielikke joonvastajaid (N)</i>
0	1	w0q10-w0q15	<i>Teadus ja tehnoloogia</i>	1,16	1,02	3,36	12
0	2	w0q17-w0q28	<i>Isikuomadused</i>	1,16	0,64	3,36	1
0	3	w0q29-w0q34	<i>Isikuomadused</i>	0,78	0,68	4,17	23
1	1	w1q31-w1q35	<i>Töö</i>	0,82	0,66	4,20	31
1	2	w1q68-w1q82	<i>Väärtused</i>	7,43	3,90	19,84	3
2	1	w2q37-w2q41	<i>Ühiskonna ja perekonna ülesanded</i>	1,18	1,01	4,80	38
2	2	w2q55-w2q78	<i>Lapsekasvatus</i>	1,06	1,05	8,00	159
3	1	w3q55-w3q65	<i>Ühiskondlik õiglus</i>	1,18	0,63	4,02	3
4	1	w4q5-w4q35	<i>Kultuurilised normid</i>	4,37	2,55	15,13	2
4	2	w4q36-w4q47	<i>Loodus ja keskkond</i>	1,36	0,72	3,79	4
4	3	w4q54-w4q59	<i>Isikuomadused</i>	0,64	0,56	3,20	31
5	1	w5q1-w5q7	<i>Soorollid</i>	1,40	0,81	4,80	4
5	2	w5q44-w5q49	<i>Suhtumine immigratsiooni</i>	8,58	6,24	26,67	12
5	3	w5q67-w5q72	<i>Eluolu</i>	0,61	0,53	3,47	43
6	1	w6q5-w6q14	<i>Sotsiaalne, poliitiline usaldus</i>	3,36	2,27	15,62	5
6	2	w6q48-w6q56	<i>Internet</i>	1,38	1,18	8,00	120

Vastuste üldist variatiivsust valitud küsimuste komplektides võrreldes selgub, et vastused on väiksema variatiivsusega küsimuste puhul, mis on oma sisult loogiliselt samasuunalised, st et joonvastamine võib esineda ka täpsete vastuste andmisel ilma, et tekiks loogilisi vastukäivusi vastustes. Küsimuste puhul, mis sisaldavad selgelt loogiliselt vastupidiseid küsimusi, esineb ka suuremat variatiivsust. Pika vastusskaalaga küsimustel on ka suurem vastuste variatiivsus, kuid see ei piira antud näitaja kasutamist analüüsis, sest töös ei võrrelda omavahel variatiivsuse näitajaid eri küsimuste plokkides.

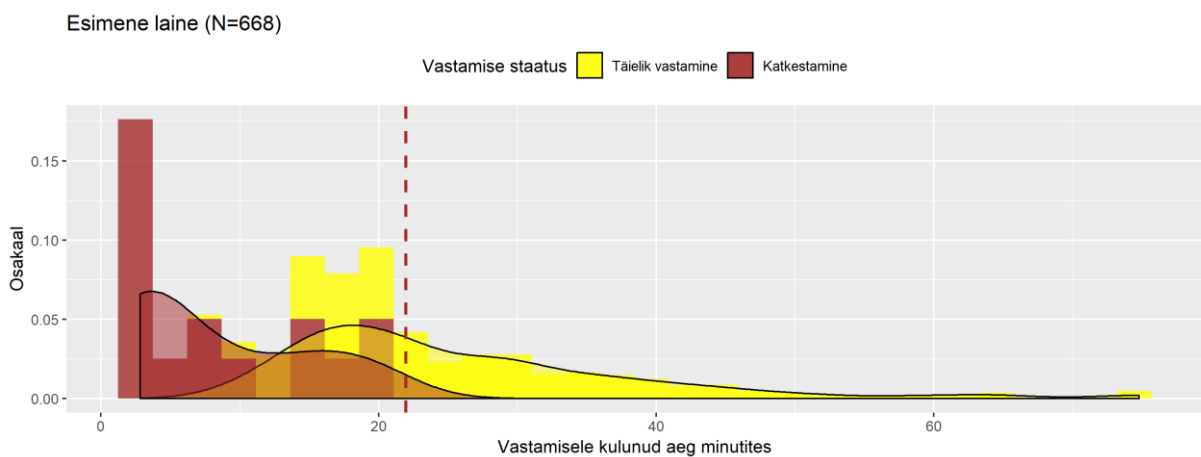
3.1.6 Vastamisele kuluv aeg

Küsimustikule vastamise eeldatavaks ajaliseks pikkuseks oli sissejuhataval lainel 10-15 minutit, teistel lainetel 20 minutit. Neljanda laine küsimustiku eelduslikku ajalist pikkust lühendati 15 minuti peale kolmanda laine küsimustiku keerulisuse ja pikenenud vastamisaja tõttu (Villar jt, 2018). Vastamisele kulunud aja mõõtmine erines sissejuhataval ja esimesel lainel võrreldes järgnevate lainetega, seetõttu on sissejuhatava ja esimese laine puhul võrdlusest välja jäetud pärast pausi vastanute vastamiseks kulunud ajad. Kuna täielikke vastuseid andnud ja katkestanud indiviidide jaotuses (vt joonis 7) on märkimisväärseid erinevusi, esitan vastamisele kulunud aja illustreerimiseks protsentjaotused (v.a sissejuhatav paneel, mille puhul esitan absoluutarvud, sest katkestamise kontsentreerumine vastamise algusminutitele muudab jaotusdiagrammi halvasti vaadeldavaks).

Vastamise aja jaotust iseloomustab üldiselt positiivne asümmeetria kõikidel lainetel. Sissejuhataval lainel kulus küsimustikule vastamiseks keskmiselt 16,7 minutit (mediaan 13,9 minutit, vt joonis 12), standardhälbega 9,3 minutit. Esimese laine keskmiseks vastamise ajaks oli 25,6 minutit (mediaanväärtuseks 22,0 minutit, vt joonis 13), standardhälbeks oli 12,9 minutit. Paneeli sissejuhataval ja esimesel lainel katkestati vastamine enim esimestel minutitel, mis on seotud peamiselt vastamise katkestamisega sissejuhataval leheküljel. Esimesel lainel toimusid kõik katkestamised 20 minuti jooksul.



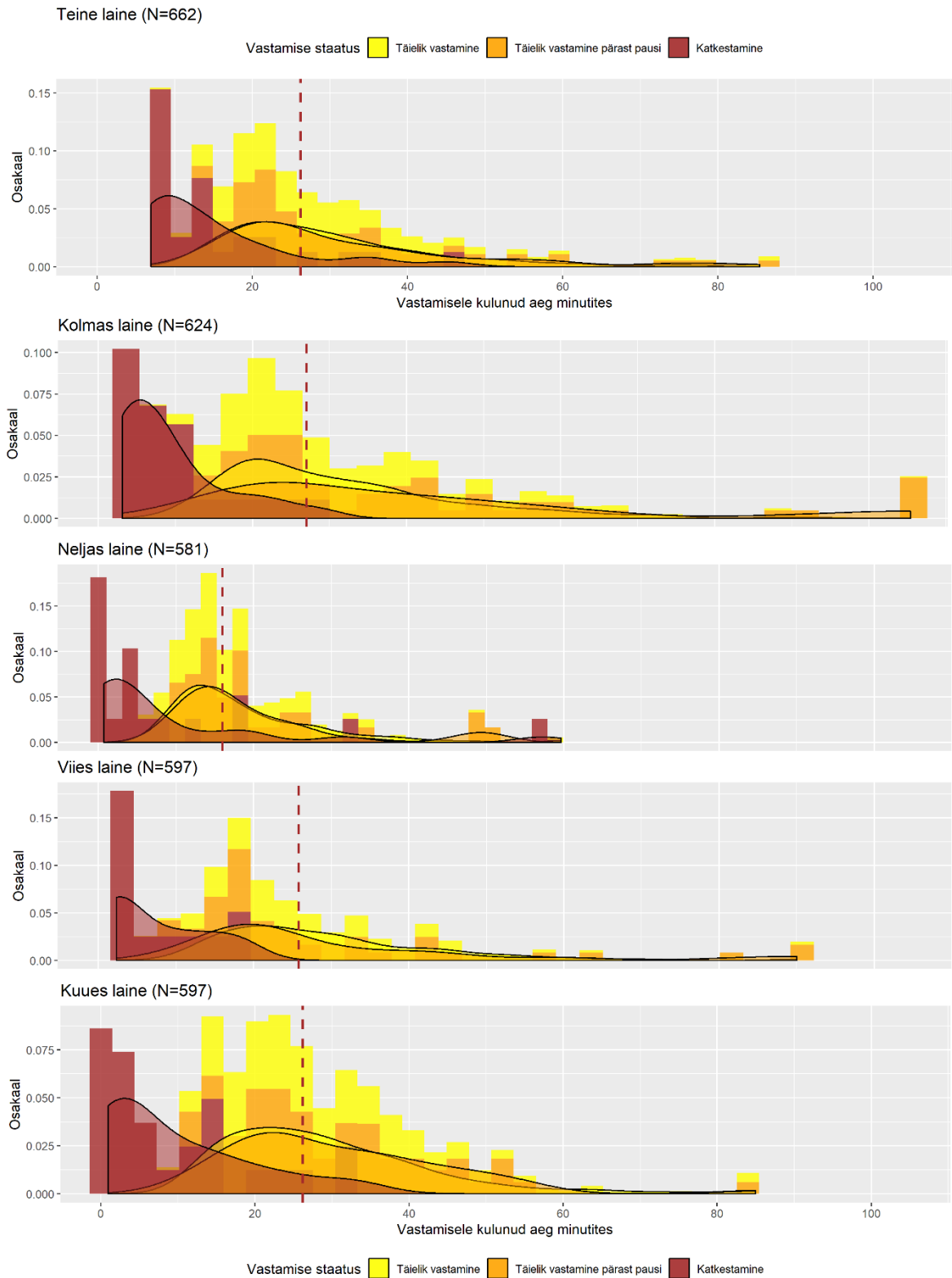
Joonis 12. Küsimustikule vastamiseks kulunud aeg minutites koos mediaanväärtusega vastamise staatuse järgi



Joonis 13. Küsimustikule vastamiseks kulunud aeg minutites koos mediaanväärtusega vastamise staatuse järgi

Paneeli 2. – 6. laine vastamisele kulunud aegade jaotuse võrdlus on esitatud joonisel 14, millelt selgub, et keskmiselt kõige vähem aega kulus küsimustikule vastamiseks paneeli neljandal lainel (keskmine 18,7, mediaan 16,0 , standardhälve 10,0 minutit), mis tulenes küsimustiku lühendamisest. Kõige pikemaid küsimustele vastamiste aegu esines paneeli kolmandal lainel (keskmine 31,4, mediaan 27,0 minutit, standardhälve 17,5 minutit). 2.- 6. laine katkestajate vastamisele kulunud aega iseloomustab samuti koondumine pigem vastamise esimestele minutitele. Neljandal lainel on näha ka oluliselt hilisemat katkestamist, kuid peatükis 3.1.3 selgus, et viimasel küsimusel katkestamine võis olla tingitud küsitluse enneaegsest sulgemisest neljandal lainel.

CRONOSe küsimustikele vastamisel osutusid ajakulukamateks kolmanda ja kuuenda laine küsimustikud. Lühemana planeeritud sissejuhatava ja neljanda laine küsimustikud tõid eeldatud tulemuse, st lühema vastamisele kulunud aja. Samas on oluline, et vastamiseks kulus mõnel juhul aega kauem kui tund, isegi üle 1,5 tunni, kusjuures paraandmetele tuginedes on arvestatud summaarset igale küsimusele vastamiseks kulunud aega, mitte küsimustiku avamise ja sulgemise ajalisi vahet.



Joonis 14. Küsimustikule vastamiseks kulunud aeg minutites koos mediaanväärtusega vastamise staatus järgi 2. – 6. paneeli lainel Eestis

3.1.7 Hinnang uuringus osalemise kogemusele

Paneeli erinevatel lainetel uuriti küsimustiku lõpuosas respondentidelt hinnangut vastamise kogemusele, mille küsimused olid igal lainel samad, v.a viiendal lainel, milles küsimused erinesid ning kus keskenduti respondentide hinnangule paneelis osalemise kogemusele. Esimese laineti korduva küsimusena uuriti respondentide hinnangut küsimustest arusaamisele ja vastamise raskusele. Tabelist 4 selgub, et enamasti hinnati küsimustest arusaamist ja vastamist üldse mitte raskeks või vähesel määral raskeks (keskmiselt 82,6% üle lainete, standardhälve 14,9%). Kõige raskemaks hinnati kolmanda laine küsimustikku, mille puhul 50% vastanutest hindasid vastamist mõõdukal määral, väga või äärmiselt raskeks. Kõige kergemini mõistetavaks osutus sissejuhatava laine küsimustik.

Tabel 4. Hinnangud küsimustest arusaamisele ja vastamise keerukusele 0.-4. ja 6. paneeli lainel

		Laine number					
		0 (N=631)	1 (N=704)	2 (N=632)	3 (N=598)	4 (N=562)	6 (N=568)
Üldse mitte raske	N	467	352	373	142	235	276
	%	74,0	50,0	59,0	23,8	41,8	48,6
Vähesel määral raske	N	127	290	206	160	220	218
	%	20,1	41,2	32,6	26,8	39,2	38,4
Mõõdukal määral raske	N	32	56	45	186	99	69
	%	5,1	8,0	7,1	31,1	17,6	12,2
Väga raske	N	5	4	5	76	7	5
	%	0,8	0,6	0,8	12,7	1,3	0,9
Äärmiselt raske	N	0	2	3	34	1	0
	%	0,0	0,3	0,5	5,7	0,2	0,0

Teiseks küsimuseks, mille abil uuriti respondentide hinnangut vastamise kogemusele, oli küsimus, kui palju vastaja pingutas, et anda küsimustele võimalikult täpsed vastused. Tabel 5 kajastab respondentide hinnanguid pingutusele täpsete vastuste andmiseks. Kõige suurema osakaalu moodustasid vastusevalikud „palju“ ja „väga palju“ kõikidest sisulistest vastustest viiendal (29,4%) ja kuuendal lainel (27,6%). Kõige vähem tajuti vastamisel pingutust täpsete vastuste andmiseks paneeli sissejuhataval, teisel ja neljandal lainel, kus vastusevalikute „palju“ ja „väga palju“ osakaalud jäid vastavalt 14,6%, 17,6% ja 16,9% juurde.

Tabel 5. Hinnang pingutusele täpsete vastuste andmiseks kõikidel paneeli lainetel

		Laine number						
		0 (N=630)	1 (N=704)	2 (N=632)	3 (N=599)	4 (N=562)	5 (N=582)	6 (N=569)
Üldse mitte	N	149	95	97	34	59	39	45
	%	23,7	13,5	15,4	5,7	10,5	6,7	7,9
Natuke	N	226	255	222	179	211	120	143
	%	35,9	36,2	35,1	29,9	37,5	20,6	25,1
Mõõdukal määral	N	165	220	202	255	197	252	224
	%	26,2	31,3	32,0	42,6	35,1	43,3	39,4
Palju	N	66	112	89	107	75	129	127
	%	10,5	15,9	14,1	17,9	13,4	22,2	22,3
Väga palju	N	26	22	22	24	20	42	30
	%	4,1	3,1	3,5	4,0	3,6	7,2	5,3

Kolmandaks uuriti respondentide hinnangut, kuivõrd vastaja nautis küsimustele vastamist konkreetse küsimustiku raames (vt tabel 6). Kõige enam nautisid respondendid vastamist teise ja neljanda laine küsimustikele. Teisel lainel moodustasid 23,3% kõikidest sisulistest vastustest valikud „palju“ ja „väga palju“ ning neljandal lainel 22,6%. Kõige madalamaks hinnati vastamise naudingut paneeli kolmandal lainel, kus vaid 15,2% vastustest näitasid suuremat naudingut („palju“, „väga palju“).

Tabel 6. Hinnang vastamise nautimisele Eestis 0.-4. ja 6. paneeli lainel

		Laine number					
		0 (N=630)	1 (N=704)	2 (N=630)	3 (N=598)	4 (N=562)	6 (N=568)
Üldse mitte	N	56	44	37	72	49	54
	%	8,9	6,3	5,9	12,0	8,7	9,5
Natuke	N	132	147	128	187	122	107
	%	21,0	20,9	20,3	31,3	21,7	18,8
Mõõdukal määral	N	312	354	318	248	264	285
	%	49,5	50,3	50,5	41,5	47,0	50,9
Palju	N	106	135	120	78	113	99
	%	16,8	19,2	19,1	13,0	20,1	17,4
Väga palju	N	24	24	27	13	14	23
	%	3,8	3,4	4,3	2,2	2,5	4,1

Üldiselt kinnitavad respondentide hinnangut ka eespool esitatud tulemused. Nimelt osutusid ka respondentide hinnangul raskemateks kolmanda, kuuenda ja viienda laine küsimustikud, teistel lainetel nii raske vastata polnud ning seda näitas ka hinnang vastamise nautimisele.

3.2 Andmekvaliteeti mõjutavad tegurid: prognoosimudelid

3.2.1 Kutsele mittevastamine

Kutsele mittevastamist prognoosivate tegurite uurimiseks on koostatud binaarse logistilise regressiooni mudelid, mille sõltuvaks tunnuseks on dihhotoomne tunnus, kas esines kutsele mittevastamine või mitte. Tabelis 7 (täpsemalt lisa 3) on esitatud keskmised piirefektid, et võrrelda mudelitesse kaasatud sõltumatute tunnuste keskmist mõju sõltumatule tunnusele. Keskmised piirefektid (nagu ka regressioonikordajad) on tinglikud teiste sõltumatute tunnuste suhtes, st näitavad keskmist mõju teiste tingimuste samaks jäädes. Kuna eelmisel lainel vastamiseks kasutatud seadme andmed on saadaval alates 3. lainest, on mudelid esitatud kahes grupis – 0.-6. laine ilma eelmisel lainel vastamiseks kasutatud seadme mõjuta ja 3.-6. laine mudelid arvestades seadme mõju. McFaddeni kordaja järgi jääb mudelite kirjeldusaste 0,19 ja 0,69 vahele.

Kutsele mittevastamist prognoosib oluliselt potentsiaalse respondendi sugu. Tõenäosus mittevastamiseks on meestel keskmiselt 4-8 protsendipunkti suurem kui naistel. Noorimast vanuserühmast eristuvad kutsele mittevastamises kaks vanemat vanuserühma, kusjuures vanemate respondentide risk kutsele mittevastamiseks on väiksem kui noorematel (keskmiselt 5-20 protsendipunkti, $p < 0,05$). Teise, neljanda ja kuuenda laine mudelite põhjal selgub eelkõige kõrgharidusega indiviidide keskmiselt 12-17 protsendipunkti väiksem tõenäosus mittevastamiseks kui põhiharidusega või madalama haridustasemega indiviididel olulisustõenäosusega alla 0,05.

Eelmisel lainel vastamiseks kasutatud seade muudab mudeli selektiivseks, kuid üldiselt tulevad ka nendest mudelitest esile sarnased mõjud. Teise mudelite komplekti põhjal ilmnes viienda ja kuuenda laine mudelites eelmisel lainel kasutatud seadme mõju. Indiviididel, kes vastasid eelmisel lainel tahvelarvutiga, on keskmiselt 6 protsendipunkti väiksem tõenäosus kutsele mittevastamiseks võrreldes arvutiga vastanutega ($p < 0,001$). Kuuenda laine mudelist selgub eelmisel lainel vastamiseks nutitelefoniga kasutanud indiviidide veidi suurem risk mittevastamiseks võrreldes arvutiga vastanutega ($p < 0,1$).

Esitatud prognoosimudelitest selgus, et kutsele mittevastamist risk on kõrgem meeste, 18-30-aastaste ja põhi- või sellest madalama haridustasemega indiviidide seas. Kui uuritud tunnustele on lisatud eelmisel lainel kasutatud seadme tüübi mõju, jäid mõjud üldises plaanis samaks ning

selgus, et eelmisel lainel tahvelarvutiga vastanute risk kutsele mitte vastata on väiksem kui arvutiga vastanutel ning kuuenda laine puhul selgus nutitelefonide kasutajate veidi suurem risk kutsele mittevastamiseks võrreldes indiviididega, kes kasutasid eelmisel lainel vastamiseks arvutit.

Table 7. Binaarne logistiline regressioon: kutsele mittevastamine

	Laine 0	Laine 1	Laine 2	Laine 3	Laine 4	Laine 5	Laine 6	Laine 3	Laine 4	Laine 5	Laine 6
	AME	AME	AME	AME	AME	AME	AME	AME	AME	AME	AME
<i>Sugu: mees¹</i>	0,028	0,042 *	0,052 *	0,063	0,079 **	0,043	0,051 .	0,029	0,034	0,000	0,039 .
<i>Vanus: 31-45²</i>	-0,009	-0,049	-0,028	-0,076	-0,043	-0,004	-0,088 *	-0,088 *	-0,043	-0,013	-0,032
<i>Vanus: 46-60²</i>	-0,017	-0,086 **	-0,033	-0,064	-0,088 .	-0,052	-0,097 *	-0,085 *	-0,093 *	-0,041	-0,038
<i>Vanus: >60²</i>	-0,047	-0,084 *	-0,083 **	-0,108	-0,196 ***	-0,137 ***	-0,167 ***	-0,121 **	-0,151 ***	-0,048	-0,060 .
<i>Kutseharidus³</i>	0,054	0,001	-0,122 *	-0,084	-0,003	-0,024	-0,103	-0,054	0,047	-0,069	-0,043
<i>Kõrgharidus³</i>	0,025	-0,027	-0,148 **	-0,136	-0,118 *	-0,076	-0,172 **	-0,080	-0,012	-0,055	-0,100 .
<i>Üldkeskharidus³</i>	0,019	-0,057	-0,161 **	-0,110	-0,051	-0,025	-0,133 *	-0,065	0,037	-0,061	-0,089 .
<i>Vene keel⁴</i>	0,062 .	0,035	-0,008	0,004 .	0,058	0,056	0,036	0,006	0,078 *	0,010	-0,004
<i>Eelmisel lainel kasutatud seade: tahvelarvuti⁵</i>								0,036	0,077	-0,063 ***	-0,056 ***
<i>Eelmisel lainel kasutatud seade: nutitelefon⁵</i>								-0,009	-0,021	-0,028	0,048 .
N	734	793	722	715	714	712	707	658	623	578	596
AIC	485,750	438,580	385,010	538,740	659,130	607,310	581,980	403,378	411,486	257,459	291,931
Pseudo R ² : McFadden	0,025	0,053	0,191	0,052	0,072	0,050	0,068	0,308	0,439	0,624	0,558

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1
 Mudelisse veel kaasatud: poliitikahuvi, sage interneti kasutus
¹ Taust: naine
² Taust: 18-30 aastat
³ Taust: põhiharidus või madalam
⁴ Taust: eesti keel
⁵ Taust: arvuti

3.2.2 Vastamise katkestamine

Eesmärgiga välja selgitada erinevate sõltumatute tegurite mõju vastamise katkestamisele, on läbi viidud binaarne logistiline regressioon, mille tulemused on esitatud tabelis 8. Ühte mudelisse on kaasatud kõikide lainete andmed. Regressioonimudeli sõltuvaks tunnuseks on dihhotoomne tunnus, kas indiviid on vähemalt ühel lainel katkestanud või mitte. Prognoosimudeli tulemustena on esitatud regressioonikordajad, standardvead, olulisustõenäosused ja regressioonikordajate eksponendid tulemuste arusaadavama tõlgendamise huvides. Regressioonikordaja eksponent näitab, kui mitu korda muutub vaadeldava indiviidide grupi katkestamise risk teiste tunnuste samaks jäädes.

Tabel 8. Binaarne logistiline regressioon: katkestamine

	<i>B</i>	<i>Exp(B)</i>	<i>SE (B)</i>	<i>P</i>
<i>(vabaliige)</i>	-3,270	0,380	-8,670	***
<i>Sugu: mees¹</i>	0,300	0,170	1,800	
<i>Vanus: 31-45²</i>	-0,220	0,200	-1,120	
<i>Vanus: 46-60²</i>	-0,910	0,260	-3,440	***
<i>Vanus: >60²</i>	-1,310	0,330	-3,920	***
<i>Kutseharidus³</i>	-0,500	0,280	-1,780	.
<i>Kõrgharidus³</i>	-0,690	0,280	-2,440	*
<i>Üldkeskharidus³</i>	-0,580	0,300	-1,950	*
<i>Vene keel⁴</i>	-0,020	0,220	-0,100	
<i>Sage interneti kasutus</i>	-0,230	0,260	-0,900	
<i>Poliitikahuvi</i>	-0,130	0,170	-0,740	
<i>Kaks või kolm meeldetuletust</i>	2,960	0,180	16,140	***
N = 3460				
$\chi^2(11) = 381,79, p = 0,00$				
Pseudo-R ² (McFadden) = 0,26				
AIC = 1113,11				
*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1				
¹ Taust: naine				
² Taust: 18-30				
³ Taust: põhiharidus või madalam				
⁴ Taust: eesti keel				

Tabelist 8 selgub, et kahte vanemasse vanuserühma kuulunutel on enam kui kolm korda väiksem risk vastamine katkestada kui kõige nooremal vanuserühmal teiste tunnuste samaks jäädes ($p < 0,05$). Haridustasemeid võrreldes on põhiharidusega ja madalama haridusega indiviididel risk katkestada kõrgeim. Teiste haridustasemete esindajate risk katkestada on üle kolme korra väiksem kui põhi- või sellest madalama haridustasemega indiviididel

kutsehariduse puhul olulisustõenäosusega alla 10%, teiste haridustasemetega puhul alla 5%. Lisaks selgub, et katkestamise risk on märkimisväärselt suurem indiviidide puhul, kellele on konkreetse laine küsimustikule vastamiseks edastatud kaks või kolm meeldetuletust võrreldes indiviididega, kellele edastati kuni üks meeldetuletus. Esimene meeldetuletus CRONOSes edastati respondentile 4-5 päeva pärast esimest kutset, teine umbes kaks nädalat pärast kutset ja kolmas nädal enne andmekogumise perioodi lõppu (Villar jt, 2018: 6).

Lisaks on läbi viidud Coxi regressioonanalüüs eesmärgiga välja selgitada, kas katkestamise hetke (küsimuse järjekorranumber, millel vastamine katkestati) aluseks võttes esineb seaduspärasusi küsimustikule vastamise katkestamises. Regressioonimudeli sõltuvaks tunnuseks on katkestamise individuaalne küsimuse number. Katkestamisi oli paneelis üldiselt vähe, seetõttu on analüüsis vaatluse all respondendi esimene katkestamine paneeli vältel, mis annab võimaluse uurida kõiki paneeli laineid koos. Tulemustena on esitatud tabelis 9 regressioonikordaja, kordaja eksponent, standardviga ja olulisustõenäosus.

Tabel 9. Coxi regressioon: katkestamine

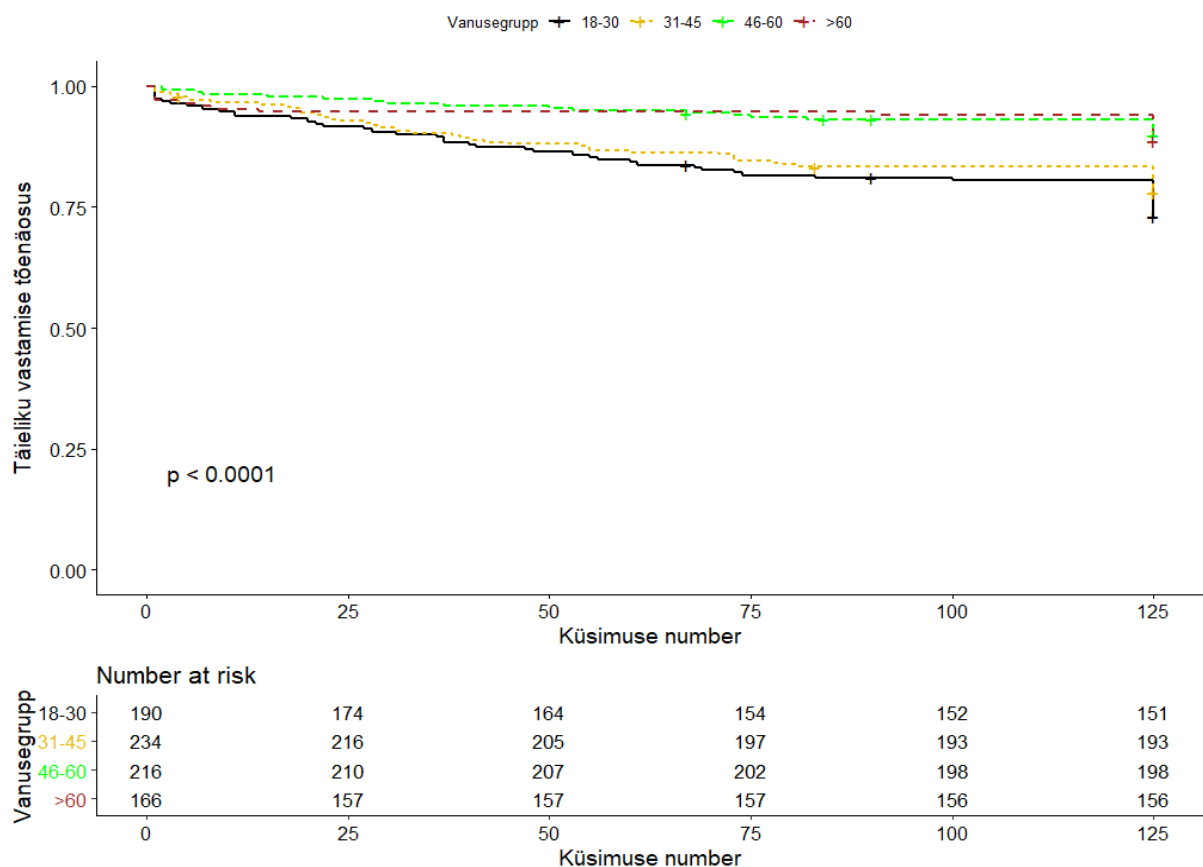
	<i>B</i>	<i>exp(B)</i>	<i>SE(B)</i>	<i>p</i>	
<i>Sugu: mees¹</i>	-0,250	0,779	0,170	0,141	
<i>Vanus: 31-45²</i>	-0,185	0,831	0,203	0,361	
<i>Vanus: 46-60²</i>	-1,025	0,359	0,268	0,000	***
<i>Vanus: >60²</i>	-1,044	0,352	0,301	0,001	***
<i>Kutseharidus³</i>	-0,244	0,784	0,277	0,379	
<i>Kõrgharidus³</i>	-0,501	0,606	0,281	0,075	.
<i>Üldkeskharidus³</i>	-0,379	0,685	0,297	0,202	
<i>Sage interneti kasutus</i>	-0,332	0,717	0,229	0,146	
<i>Poliitika huvi</i>	-0,064	0,938	0,175	0,715	
<i>Vene keel⁴</i>	-0,151	0,860	0,229	0,511	

N= 806, sündmusi= 143
Waldi test = 34,52 on 10 df, p=0,0002
R²= 0,044
*** p < 0,001; ** p < 0,01; * p < 0,05; . p<0,1
¹ Taust: naine
² Taust: 18-30
³ Taust: põhiharidus või madalam
⁴ Taust: eesti keel

Tabelist 9 selgub, et risk küsimustiku jooksul vastamine katkestada on enam kui kaks korda väiksem 46-60-aastaste ja vanemate respondentide seas võrreldes noorima vanuserühmaga teiste tunnuste samaks jäädes (p<0,001), kuid statistiliselt oluline erinevus 31-45-aastaste

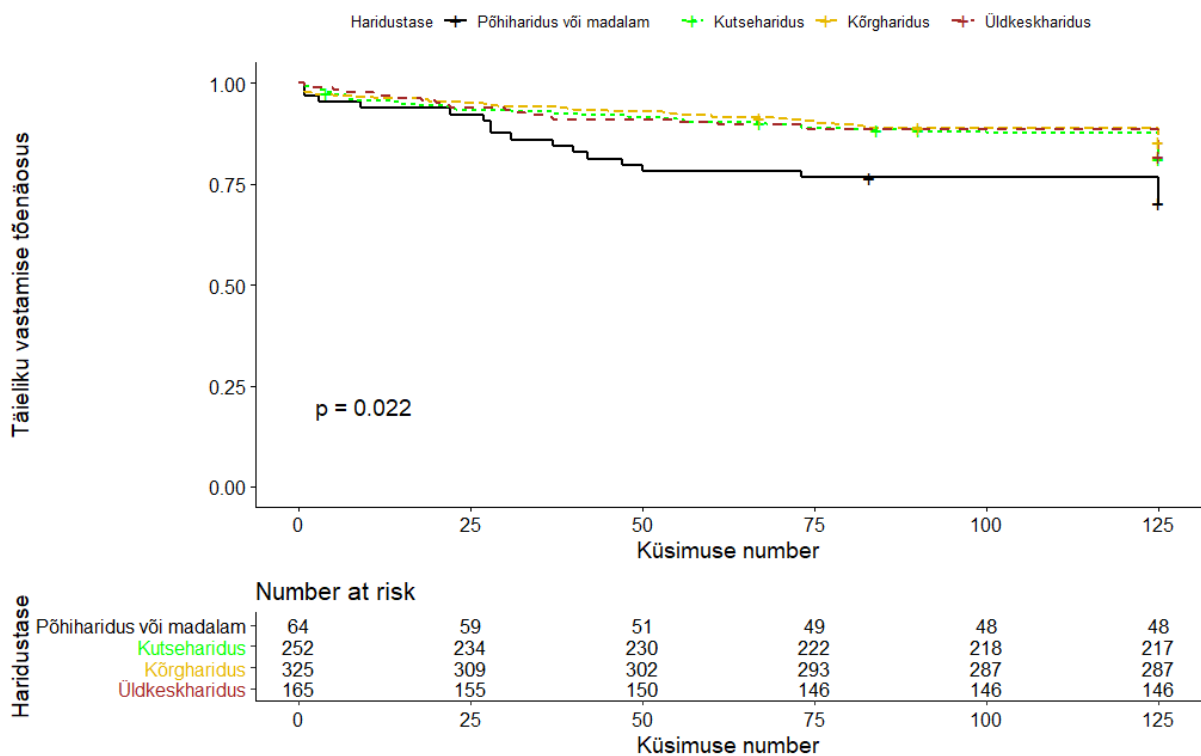
respondentide riskiga katkestada puudub. Lisaks eristub kõrgharidusega indiviidide väiksem risk küsimustiku jooksul katkestada võrreldes põhi- ja madalama haridustasemega, olles 1,65 korda väiksem kui madalama haridustaseme esindajatel olulisustõenäosusega alla 10%.

Tulemuste illustreerimiseks on joonisel 15 esitatud küsimuse n katkestuseta vastamise tõenäosushinnangud Kaplani-Meieri meetodil. Jooniselt on visuaalselt näha, et kahel nooremal vanuserühmal on oluliselt suurem risk vastamine katkestada. Kuigi vanima vanuserühma enamus katkestamisi toimus küsimustiku esimese kümne küsimuse juures, siis üldiselt oli katkestajaid viimases vanuserühmas väga vähe. Katkestajaid oli vähe ka 46-60-aastaste seas, kuid katkestamised jaotusid küsimustiku vältel hajusalt. Nooremaid vanuserühmi iseloomustas suur katkestajate hulk kohe küsimustiku algul, kuid siis esines katkestamist hajusalt terve küsimustiku vältel. Nooremates vanuserühmades katkestas vähemalt ühel korral vastamise 20% vanusegrupi esindajatest.



Joonis 15. Täielike vastuste andmise tõenäosuse hinnangud Kaplani-Meieri meetodil vanuserühmade järgi (N=806)

Coxi regressioonist selgus veel, et haridustasemel on statistiliselt oluline prognoosiv mõju katkestamise toimumisele olulisusnivoool 10%. Uurides Kaplani-Meieri hinnanguid küsimuse n täieliku vastamise toimumisele, selgub, et põhi- ja sellest madalama haridustasemega indiviidide väiksema arvu ning suurema katkestajate osakaalu tõttu eristuvad nad teistest haridustasemetest kiirema ja arvukama katkestamise poolest teiste tingimuste samaks jäädes ning seos on statistiliselt oluline olulisustõenäosusega alla 5%. (vt joonis 16). Lisaks on jooniselt näha, et alates 25. küsimusest toimub küsimuse n vastamise tõenäosuse järsk langus põhi- ja sellest madalama haridustasemega indiviidide puhul, mis võib viidata liigsele pingutusvajadusele madalama haridustasemega indiviidide puhul tulenevalt küsimustiku pikkusest.



Joonis 16. Täielike vastuste andmise tõenäosuse hinnangud Kaplani-Meieri meetodil vanuserühmade järgi (N=806)

Prognoosimudelitest selgus, et vastamise katkestamise risk on oluliselt suurem kahes nooremas vanuserühmas, põhi- ja sellest madalama haridustasemega indiviidide ning kaks või kolm meeldetuletust saanud indiviidide seas. Lisaks selgus, et kui nooremate respondentide katkestamised olid hajutatud üle terve küsimustiku, siis põhi- ja madalama haridustasemega indiviide risk katkestada suurenes alates küsimustiku 25. küsimusest.

3.2.3 Küsimusele mittevastamine

Küsimusele mittevastamise uurimiseks on koostatud binaarse logistilise regressiooni mudel, mille sõltumatuks tunnuseks on dihhotoomne tunnus, kas indiviidil esineb vähemalt üks küsimusele mittevastamine vastava laine küsimustikus või mitte. Tulemustena on tabelis 10 esitatud tunnuste keskmine mõju keskmiste piirefektidena.

Tabel 10. Binaarne logistiline regressioon: vähemalt üks küsimusele mittevastamine lainel (1., 3. ja 5.laine)

	Laine 1 AME	Laine 3 AME	Laine 5 AME	
<i>w1w3w5exp0: vastamise täpsuse rõhutamine¹</i>	0,002	-0,024	0,012	
<i>w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine¹</i>	-0,062	-0,045	0,078	.
<i>Sage interneti kasutus</i>	0,074	.	-0,084	0,038
<i>Poliitikahuvi</i>	-0,082	*	-0,116	** 0,047
<i>Sugu: mees²</i>	0,081	*	-0,067	-0,009
<i>Vanus: 31-45³</i>	0,112	*	0,047	-0,044
<i>Vanus: 46-60³</i>	0,061		0,089	-0,049
<i>Vanus: >60³</i>	0,139	*	0,178	* -0,056
<i>Kutseharidus⁴</i>	0,012		-0,253	** -0,263 **
<i>Kõrgharidus⁴</i>	-0,148	.	-0,294	** -0,330 ***
<i>Üldkeskharidus⁴</i>	0,014		-0,260	** -0,302 ***
<i>Vene keel⁵</i>	-0,059	-0,006	0,100	*
N	704	489	581	
AIC	887,860	626,210	632,190	
Pseudo R ² : McFadden	0,125	0,358	0,184	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1
Mudeliselle veel kaasatud: rööprähklus, kaks või kolm meeldetuletust
¹ Taust: standardne sissejuhatus
² Taust: naine
³ Taust: 18-30 aastat
⁴ Taust: põhiharidus või madalam
⁵ Taust: eesti keel

Mudelitest selgub, et täpsete vastuste andmise rõhutamine küsimustiku algul üldiselt ei prognoosi küsimusele mittevastamist. Poliitikast huvitatutel oli küsimusele mittevastamiseks keskmiselt üle 8 protsendipunkti väiksem risk küsimusele mittevastamiseks (p<0,05). Esimese laine mudeli järgi oli meestel kõrgem risk küsimusele mittevastamiseks võrreldes meestega teiste tunnuste samaks jäädes. Esimese ja kolmanda laine mudelitest selgub ka vanuse mõju küsimusele mittevastamisele: 31-45- ja vanematel kui 60-aastastel indiviididel on keskmiselt suurem risk anda vähemalt ühele küsimusele mittesisulisi vastuseid või jätta vastamata võrreldes noorima vanuserühmaga teiste tunnuste samaks jäämisel (p<0,05).

Kolmanda ja viienda laine mudelitele tuginedes võib öelda, et põhi- ja sellest madalama haridustasemega indiviididel on suurem risk küsimusele mittevastamiseks võrreldes teiste haridustasemete esindajatega, kui teised omadused on samad. Väiksem risk küsimusele mittevastamiseks on kõrgharidusega indiviididel võrreldes kuni põhiharidusega indiviididega. Kuuenda laine andmetel omas küsimusele mittevastamisel prognoosivat mõju ka keel, milles küsimustele vastati, olles vene keeles vastanud indiviidide puhul keskmiselt 10 protsendipunkti võrra suurem kui eesti keeles vastanutel olulisustõenäosusega alla 5%. Meeldetuletuste arv ja mitme asjaga vastamise ajal tegelemine küsimusele mittevastamisele prognoosivat mõju ei oma.

Küsimusele mittevastamist küsimustiku vältel on uuritud mitteparameetriliste testide abil. Loodud on andmestik, kus käsitletakse kirjena üht küsimust ning mis sisaldab andmeid küsimusele mittevastamise kohta. Mitteparameetriliste testide rakendamine osutus vajalikuks erindite esinemise tõttu küsimusele mittevastamise jaotuses ning arvesse on võetud küsimusele mittevastamist pideva tunnuseks. Mitteparameetriliste testide aluseks on võetud küsimusele mittevastamine kolmandal ja kuuendal lainel, mil esines enim mittesisulisi vastuseid ja vastamata jätmist. Mitteparameetrilise testi abil on uuritud küsimuse teemavaldkonna, tunnuse tüübi ja vastusevariantide formaadi (raadionupud, mitmene valik) erisusi küsimusele mittevastamises. Teemavaldkonnana on käsitletud uurimisrühma poolt kasutatud küsimuste temaatilist jaotust.

Kruskali-Wallise testi tulemusel võib kinnitada erinevusi küsimusele mittevastamises küsimuse teemavaldkonna järgi (hii-ruut = 40.72, df = 11, $p < 0,001$). Enim küsimusele mittevastamisi koondusid hariduse, jaotava õigluse ja MSQD (*Multinational Study of Questionnaire Design*) küsimuste teemaplokki, milles käsitleti erinevaid teemasid nagu keskkond, ebavõrdus, poliitika ja kuritegevus. Lisaks esines olulisi erisusi küsimusele mittevastamises tunnuse tüübi järgi (Kruskali-Wallise hii-ruut = 33.06, $p < 0,001$), olles ootuspäraselt kõrgeim sõnalist vastust eeldavate avatud küsimuste puhul võrreldes ordinaalsete, dihhotoomsete, nominaalsete ja arvuliste tunnustega. Vastuste formaadi uurimine tulemust ei andnud, sest peamiselt kasutati CRONOSes valikvastustes raadionuppe ja mitmest valikut (linnukestena) esines vaid kuue küsimuse puhul. Mitmese valiku puhul küsimusele mittevastamist ei esinenud.

3.2.4 Joonvastamine

Joonvastamise tuvastamiseks CRONOSe andmestikus on küsimustikest välja valitud küsimuste komplektid, milles on vähemalt viis järjestikust vähemalt viie sama vastusevariandiga küsimust. Vastuste variatiivsuse ja variatiivsust mõjutavate tegurite uurimiseks on arvatud komplekti kuulunud küsimuste individuaalne vastuste dispersioon. Kõrgem variatiivsus tähendab erinevate vastusevariantide valimist ja väiksem variatiivsus võib tähendada süstemaatilist ebatäpsete vastuste andmist, st joonvastamist. Kui tegemist on suhteliselt ühetaoliste ja ühesuunaliste küsimustega, ei pruugi joonvastamine ebatäpsete vastuste andmist tähendada. Variatiivsust prognoosivate tegurite (ehk joonvastamist pärssivate tegurite mõju) välja selgitamiseks on koostatud lineaarse regressiooni mudel, mille sõltuvaks tunnuseks on individuaalne vastuste variatiivsus konkreetsetes küsimuste plokis pideva tunnuseks. Kuna mudelisse on kaasa arvatud prognoosiva tunnuseks eksperimendi tunnus (vastuste täpsuse ja hoolsuse rõhutamine) ja ka lainel vastamiseks kasutatud seade, siis on mudelite tulemused esitatud koos kolmanda ja viienda laine kohta ning teise, neljanda ja kuuenda laine kohta. Mudelite tulemused on esitatud regressioonikordajatenä tabelis 11.

Viienda laine mudelite alusel võib öelda, et vanemate vanuserühmade esindajatel esineb joonvastamist rohkem kui noorimal vanuserühmal teiste tunnuste samaks jäädes ($p < 0,05$). Lisaks prognoosib viienda laine kolmanda mudeli alusel joonvastamist indiviidi haridustase. Joonvastamise tõenäosus on kõrgem kõrg- ja kutseharidusega indiviididel võrreldes põhi- ja sellest madalama haridustasemega indiviididega ($p < 0,05$). Lisaks võib esile tuua, et viiendal lainel eristub vene keeles vastanute suurem vastuste variatiivsus, seega väiksem joonvastamise tõenäosus ($p < 0,001$). Ilma eksperimendi tunnusega mudelis on ülejäänud sõltumatute tunnuste prognoosid samad, esinevad vaid väiksed erinevused kordajate väärtustes. Mudelite üldine kirjeldusvõime jääb 2,1% ja 11,9% vahele.

Tabelis 12 esitatud teise ja neljanda laine vastuste variatiivsuse mudelitest selgub, et venekeelsele küsimustikule vastanute vastuste variatiivsus on suurem kui eesti keeles vastanud respondentidel olulisusnivool 5% ($p < 0,05$), mis tähendab eesti keeles vastanute suuremat joonvastamist. Indiviidi vanuse mõju vastuste variatiivsusele avaldus teise laine teise ja neljanda laine teise küsimuste komplekti puhul ($p < 0,05$), kusjuures vanematel vanuserühmal on vastuste variatiivsus väiksem kui noorimal vanuserühmal teiste tunnuste samaks jäädes.

Mõju osas vastupidine suund on avaldunud kuuenda laine esimese küsimuste ploki andmetel, kus 31-45-aastaste vastuste variatiivsus on olnud suurem kui noorimal vanuserühmal ($p < 0,05$).

Üheks oluliseks vastuste variatiivsust prognoosivaks tunnuseks on respondendi sugu: meeste vastuste variatiivsus on üldiselt väiksem kui naistel teiste tunnuste samaks jäädes ($p < 0,05$). Vastupidine tulemus avaldus kuuenda laine esimese küsimuste komplekti puhul. Lisaks omab vastuste variatiivsusele teatud tingimustel negatiivset mõju sage interneti kasutamine ja kahe või kolme meeldetuletuse saamine ning ilma eksperimenttunnuseta mudelis kaotab meeldetuletuste arvu tunnus statistilise olulisuse. Tabelis 12 esitatud mudelite üldine kirjeldusvõime jääb 2,7% ja 7,9% vahele.

Lineaarse regressiooni mudelitest selgus, et joonvastamist võib enam esineda vanemate, kutse- või kõrgharidusega ja eesti keeles vastanud indiviidide seas. Teatud tingimustel võib joonvastamist enam esineda ka hiliste vastajate ja nende seas, kes kasutavad igapäevaselt internetti.

Tabel 11. Lineaarne regressioon: vastuste variatiivsus 3. ja 5. lainel

	Laine 3 1. kmpl B		Laine 5 1. kmpl B		Laine 5 2. kmpl B		Laine 5 3. kmpl B	
(Vabaliige)	1,021	***	1,527	***	12,585	***	0,762	***
Sage interneti kasutus	-0,082		0,034		-1,493	*	-0,060	
Vanus: 31-45 ¹	0,022		-0,098		-0,790		0,116	
Vanus: 46-60 ¹	0,083		-0,330	**	-1,880	*	0,009	
Vanus: >60 ¹	0,056		-0,385	**	-0,906		-0,015	
Kutseharidus ²	0,059		-0,076		-2,200		-0,194	*
Kõrgharidus ²	0,108		0,037		-2,040		-0,285	**
Üldkeskharidus ²	0,089		-0,024		-2,293		-0,184	
Vene keel ³	0,073		0,188	*	2,438	***	0,373	***
Kaks või kolm meeldetuletust	0,059		0,028		-1,090		-0,193	*
N	489		581		581		581	
R ²	0,021		0,051		0,056		0,119	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

Mudelisse veel kaasatud: vastuste täpsuse rõhutamise eksperimenttunnus, rööprähklus, poliitikahuvi, sugu, vastamiseks kasutatud seade, vastamisele kulunud aeg

¹ Taust: 18-30

² Taust: põhiharidus või madalam

³ Taust: eesti keel

Tabel 12. Lineaarne regressioon: vastuste variatiivsus 2., 4. ja 6. lainel

	Laine 2 1.kmpl B	Laine 2 2.kmpl B	Laine 4 1.kmpl B	Laine 4 2.kmpl B	Laine 4 3.kmpl B	Laine 6 1.kmpl B	Laine 6 2.kmpl B
(Vabaliige)	1,394 ***	1,053 ***	5,888 ***	1,574 ***	0,910 ***	4,054 ***	0,878 **
Sage interneti kasutus	-0,109	-0,115	-0,221	-0,126	-0,137 *	0,041	0,018
Sugu: mees ¹	0,109	-0,192 *	-0,645 **	-0,112	-0,057	0,394 *	-0,239 *
Vanus: 31-45 ²	0,056	-0,215	0,121	-0,223 *	0,007	0,668 *	-0,006
Vanus: 46-60 ²	-0,090	-0,319 *	0,211	-0,400 ***	-0,061	0,572	-0,137
Vanus: >60 ²	-0,079	-0,367 *	-0,117	-0,568 ***	0,022	0,490	-0,347
Kutseharidus ³	-0,055	0,054	-0,905	0,102	-0,064	-1,154 *	0,451
Kõrgharidus ³	-0,194	0,320	-1,270 *	0,124	-0,096	-1,019 *	0,614 *
Üldkeskharidus ³	-0,226	0,184	-1,003 *	0,021	-0,124	-1,415 **	0,508
Vene keel ⁴	0,208 *	0,268 *	1,206 ***	-0,033	0,007	-0,359	0,107
Kasutatud seade: tahvelarvuti ⁵	-0,261	-0,435	0,390	0,024	-0,043	0,773	0,011
Kasutatud seade: nutitelefon ⁵	0,078	-0,264 *	0,333	-0,127	-0,039	-0,377	-0,020
Vastamisele kulunud aeg	-0,003	0,009 **	-0,006	0,003	0,001	0,000	0,001
Kaks või kolm meeldetuletust	0,165	-0,310 *	-0,399	0,027	-0,055	0,040	-0,008
N	632	594	562	562	562	567	534
R ²	0,040	0,067	0,072	0,079	0,027	0,048	0,039

*** p < 0,001; ** p < 0,01; * p < 0,05; . p<0,1

Mudelisse veel kaasatud: vastuste hoolikuse eksperimenttunnus, rööprähklus, poliitikahuvi

¹ Taust: naine
² Taust: 18-30
³ Taust: põhiharidus või madalam
⁴ Taust: eesti keel
⁵ Taust: arvuti

3.2.5 Vastamisele kuluv aeg

Vastamisele kulunud aega mõjutavate tegurite välja selgitamiseks on läbi viidud lineaarne regressioon, mille sõltuvaks tunnuseks on logaritmitud vastamise aeg minutites. Tabelis 13 on esitatud mudelite tulemused. Tabelis on esitatud eksperimentitunnuse ja vastamisel kasutatud seadme tunnuste kaasamise tõttu kolmanda ja viienda ning teise, neljanda ja kuuenda laine mudelid. Tulemuste tabelis on esitatud regressioonikordajad.

Vastamiseks kuluvat aega prognoosib oluliselt respondendi vanus, olles väikseim noorima vanuserühma esindajatel. Vastamisele kulub kõige rohkem aega vanima vanuserühma esindajatel ($p < 0,001$). Haridustaseme mõju avaldus vaid kõrghariduse ja põhi- ja sellest madalama haridustasemega indiviidide eristumises. Kõrgharidusega indiviididel kulus vastamiseks vähem aega kui madalaima haridustasemega indiviididel ($p < 0,5$). Lisaks selgub, et mitme asjaga korraga tegelemine pikendab ootuspäraselt vastamiseks kuluvat aega ($p < 0,001$) ning sage interneti kasutus prognoosib kiiremat vastamist ($p < 0,05$). Vastamisel täpsuse rõhutamine ja täpsuseks kinnituse küsimine ning hooletu vastamise negatiivsete tagajärgede rõhutamine pikendab vastamiseks kuluvat aega olulisustõenäosusega alla 5%. Täpsuse ja hoolsuse teadete mõju avaldus kõikides mudelites.

Üldiselt võib öelda, et vastamisele kulub enam aega vanematel, põhi- ja sellest madalama haridustasemega indiviididel võrreldes kõrgharidusega respondentidega, mitut asja samaaegselt tegevatel respondentidel. Lisaks kulus enam aega vastamiseks indiviididel, kellele rõhutati täpsuse vastuste olulisust (koos kinnitusega) ja hooletu vastamise negatiivseid tagajärgi.

Tabel 13. Lineaarne regressioon: vastamiseks kulunud aeg 2.-6. lainel (logaritm)

	Laine 3	Laine 5	Laine 2	Laine 4	Laine 6
	B	B	B	B	B
(Vabaliige)	3,325 ***	3,064 ***	3,179 ***	2,725 ***	3,154 ***
<i>w1w3w5exp0</i> : vastamise täpsuse rõhutamine ¹	0,055	0,030			
<i>w1w3w5exp0</i> : vastamise täpsuse rõhutamine ja lubaduse küsimine ¹	0,104 *	0,175 ***			
<i>w2w4w6exp0</i> : hoolika vastamise positiivsed tagajärjed ²			0,056	0,056	0,080 *
<i>w2w4w6exp0</i> : hooletu vastamise negatiivsed tagajärjed ²			0,109 **	0,090 *	0,128 **
Rööprähklus	0,075	0,141 ***	0,049	0,069	0,152 ***
Sage interneti kasutus	-0,102	-0,141 **	-0,163 ***	-0,178 ***	-0,145 **
Vanus: 31-45 ³	0,160 **	0,219 ***	0,155 ***	0,160 **	0,184 ***
Vanus: 46-60 ³	0,267 ***	0,312 ***	0,273 ***	0,308 ***	0,323 ***
Vanus: >60 ³	0,530 ***	0,559 ***	0,458 ***	0,504 ***	0,520 ***
Kutseharidus ⁴	-0,078	0,020	0,023	-0,005	-0,057
Kõrgharidus ⁴	-0,224 **	-0,121	-0,106	-0,123	-0,166 *
Üldkeskharidus ⁴	-0,087	-0,006	-0,024	-0,028	-0,100
N	489	581	632	562	569
R ²	0,220	0,282	0,242	0,241	0,261

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

Mudelisse veel kaasatud: sugu, keel, kaks või kolm meeldetuletust, vastamiseks kasutatud seade

¹ Taust: Standardne sissejuhatus

² Taust: informatsiooni tagajärgede osas ei kuvatud

³ Taust: 18-30

⁴ Taust: põhiharidus või madalam

3.2.6 Hinnang uuringus osalemise kogemusele

Kõikidel lainetel (v.a viies laine) küsiti küsimustiku lõpuosas respondentidelt tagasisidet küsimustikule vastamise kogemuse kohta. Esmalt paluti hinnata küsimustest arusaamise ja vastamise raskust, mida mõjutavate tegurite välja selgitamiseks on koostatud binaarse logistilise regressiooni mudelid (vt tabel 14). Sõltuvaks tunnuseks on dihhotoomne tunnus, kus 1 tähendab, et indiviid koges raskust vähesel määral või enam, 0 et ei tajunud raskust vastamisel. Tulemustena on esitatud sõltumatute tunnuste keskmised piirefektid.

Mudelist selgub, et poliitikast huvitatutel on oluliselt väiksem tõenäosus hinnata vastamist keerukamaks (olulisustõenäosus olenevalt lainest alla 1-10%). Lisaks osutus oluliseks kõrgharidust omavate respondentide väiksem tõenäosus (üle 20 protsendipunkti) hinnanguks, et küsimustest arusaamine ja vastamine on rasked ($p < 0,05$) võrreldes põhi- ja sellest madalama haridustasemega indiviididega. Esimese ja kolmanda laine mudeli põhjal hindavad vastamist vähem keerulisemaks vene keeles vastanud indiviidid, teistel lainetel mõju oluliseks ei osutunud. Respondentide hinnangud vastamise keerukusele olid nõrgalt seotud ka eksperimentitunnusega kuuendal lainel, olles keskmiselt umbes 10 protsendipunkti kõrgemad, kui respondentile kuvati hooletu/ hoolika vastamise tagajärgede infot ($p < 0,1$). Lisaks omas negatiivset mõju see, kui indiviidid on poliitikast huvitatud, mis tähendab, et need indiviidid tajusid vähem keerukust küsimusest arusaamisel ja vastamisel. Lisaks suurendab riski tajuda vastamisel raskust indiviidi internetikasutuse sagedus ($p < 0,1$). Kuuenda laine andmetel omas nõrka prognoosivat mõju raskuse kogemisele vastamisel ka meeldetuletuste arv. Kaks või kolm meeldetuletust saanud indiviididel oli keskmiselt 8 protsendipunkti kõrgem risk hinnata vastamist keeruliseks kui kuni ühe meeldetuletuse saanud indiviidid olulisustõenäosusega alla 10% teiste tunnuste samaks jäädes.

Seega võib öelda, et vastamine on tundunud raskem eesti keeles vastanute, põhi- ja sellest madalama haridustasemega indiviidide, poliitikast vähem huvitatute, harvem interneti kasutavate indiviidide jaoks ning teatud tingimustel ka nende jaoks, kes on alustanud vastamist pärast mitut meeldetuletust.

Teiseks küsiti vastajatelt hinnangut küsimustiku jooksul panustatud pingutusele täpsete vastuste andmiseks, mida mõjutavate tegurite uurimiseks on koostatud binaarse logistilise regressiooni mudelid (vt tabel 15). Mudeli sõltuvaks tunnuseks on dihhotoomne tunnus, kus pingutusena

käsitletakse antud vastuste väärtusi „Väga palju“ ja „Palju“. Mudeli tulemused on esitatud keskmiste piirefektidena.

Prognoosimudelist selgub, et vanusel on prognoosiv mõju tajutud pingutusele täpsete vastuste andmiseks. Kahe vanema vanuserühma respondendid tajuvad oluliselt väiksemat pingutust täpsete vastuste andmiseks kui noorim vanuserühm teiste tunnuste samaks jäädes ($p < 0,001$). 31-45-aastaste erinevus noorimast vanuserühmast osutus oluliseks esimese ja neljanda laine puhul, olles samuti negatiivne, st tajuti vähem pingutust täpsete vastuste andmiseks ($p < 0,05$). Rööprähklus vastamisel prognoosib väiksemat tajutud pingutust. Täpsete vastuste rõhutamine ja lubaduse küsimine ning hooletu vastamise negatiivsete tagajärgede rõhutamine suurendab täpsete vastuste andmisel pingutuse tajumiseks olulisustõenäosusega alla 5%. Oluliselt eristub ka vene keeles vastanud respondentide tajutud pingutuse hinnang eesti keeles vastanud indiviidide hinnangust, olles märkimisväärselt madalam, st hinnati pingutust oluliselt väiksemaks võrreldes eesti keeles vastanud indiviididega ($p < 0,001$). Vastupidine tulemus selgus kolmanda laine mudelist, kus venekeelsele küsimustikule vastanud tajusid suuremat pingutust võrreldes eesti keeles vastanutega ($p < 0,001$).

Haridustase omab samuti prognoosivat mõju pingutuse tajumisel täpsete vastuste andmiseks. Kõrgharidusega indiviididel on keskmiselt 14-18 protsendipunkti väiksem risk tajuda vastamisel pingutust võrreldes põhi- või madalama haridustasemega indiviididega ($p < 0,05$). Lisaks osutus neljanda ja viienda laine põhjal oluliseks kahe või kolme meeldetuletusega indiviidide väiksem risk pingutuse tajumiseks vastamisel võrreldes kuni ühe meeldetuletuse saanud indiviididega ($p < 0,05$).

Pingutust täpsete vastuste andmisel hindavad kõrgemaks nooremad, põhi- ja sellest madalama haridustasemega, eesti keelt kõnelevad ning indiviidid, kellele kuvati informatsiooni vastuste täpsuse olulise kohta (koos lubadusega) või hooletu vastamise negatiivseid tagajärgi. Oluline on, et vastamisel tajusid väiksemat pingutust täpsete vastuste andmisel mitut asja vastamisega samaaegselt teinud ja mitu meeldetuletust saanud indiviidid.

Eraldi küsimusena uuriti respondentidelt, kuivõrd nad nautisid küsimustele vastamist. Uurimaks vastamise nautimist prognoosivaid tegureid, on koostatud binaarse logistilise regressiooni mudelid kõikide lainete, v.a viienda laine, andmete põhjal (vt tabel 16). Sõltuvaks

tunnuseks dihhotoomne tunnus, vastamise nautimine tähendab, et inividid nautis vastamist palju või väga palju. Tulemused on esitatud keskmiste piirefektidena tabelis 16.

Mudelitest selgub, et enim nautisid vastamist põhiharidusega ja madalama haridusega ning vene keeles vastanud respondendid. Ka sage internetikasutus prognoosib vastamise nautimist. Kutse-, üldkesk- ja kõrgharidusega inividide tõenäosus vastamist nautida on märkimisväärselt madalamad kui põhiharidusega või madalama haridustasemega inivididel ($p < 0,05$). Esimese laine mudeli põhjal nautisid vastamist enam need, kellele rõhutati vastuste täpsust ja küsiti lubadust täpsete vastuste andmiseks, võrreldes inivididega, kellele täpsust ei rõhutatud ($p < 0,05$). Lisaks avaldus tagasihoidlik positiivne mõju ka hoolika vastamise positiivsete tagajärgede rõhutamisel vastamise nautimisele ($p < 0,1$). Teise ja kuuenda laine mudelite alusel võib öelda, et hilised küsitlusele vastajad ei nautinud küsitlusele vastamist ($p < 0,1$).

Prognoosimudelitele toetudes võib seega öelda, et vastamisprotsessi nautisid enam põhi- ja sellest madalama haridustasemega, vene keeles vastanud, sagedasti internetti kasutavad ja poliitikast enam huvitatud inividid. Teatud tingimustel nautisid vastamist enam inividid, kellele kuvati vastuste täpsuse olulisust ja küsiti lubadust täpsuseks ja hoolika vastamise positiivseid tagajärgi. Vastamist nautisid vähem mitu meeldetuletust saanud inividid võrreldes kuni ühe meeldetuletuse saanud respondentidega.

Tabel 14. Binaarne logistiline regressioon: hinnang küsimustest arusaamise ja vastamise keerukusele

	<i>Laine 1</i>		<i>Laine 3</i>		<i>Laine 2</i>		<i>Laine 4</i>		<i>Laine 6</i>
	AME		AME		AME		AME		AME
<i>w1wexp0: vastamise täpsuse rõhutamine¹</i>	0,012		0,084	*					
<i>w1w3exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine¹</i>	0,027		0,013						
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed²</i>					-0,010		0,034		0,113
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed²</i>					0,049		0,092	.	0,090
<i>Sage interneti kasutus</i>	-0,101	*	-0,073		-0,105	.	-0,010		-0,100
<i>Poliitika huvi</i>	-0,140	***	-0,118	**	-0,068	.	-0,059		-0,078
<i>Sugu: mees³</i>	-0,064	.	-0,019		0,014		-0,061		-0,051
<i>Kutseharidus⁵</i>	-0,012		-0,114	*	0,060		-0,024		-0,028
<i>Kõrgharidus⁵</i>	-0,196	*	-0,180	**	0,092		-0,082		-0,022
<i>Üldkeskharidus⁵</i>	-0,108		-0,134	*	0,104		-0,006		-0,009
<i>Vene keel⁶</i>	-0,083	.	-0,091	.	-0,075		-0,054		-0,050
<i>Kaks või kolm meeldetuletust</i>	0,053		-0,073		-0,038		0,032		0,086
N	704		489		632		561		568
AIC	945,890		485,230		868,550		777,710		795,470
Pseudo R ² : McFadden	0,062		0,248		0,020		0,021		0,027
<p>*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1 Mudelisse veel kaasatud: vanus, rööprähklus ¹ Taust: Standardne sissejuhatus ² Taust: informatsiooni tagajärgede osas ei kuvatud ³ Taust: naine ⁴ Taust: 18-30 ⁵ Taust: põhiharidus või madalam ⁶ Taust: eesti keel</p>									

Tabel 15. Binaarne logistiline regressioon: hinnang pingutusele täpsete vastuste andmiseks

	Laine 1	Laine 3	Laine 5	Laine 2	Laine 4	Laine 6
	AME	AME	AME	AME	AME	AME
<i>w1w3w5exp0: vastamise täpsuse rõhutamine¹</i>	0,034	0,005	0,027			
<i>w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine¹</i>	0,112 ***	0,015	0,076 .			
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed²</i>				0,021	0,070 *	0,071
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed²</i>				0,102 **	0,087 *	0,090 *
<i>Rööprähklus</i>	0,003	-0,066 *	-0,113 **	-0,086 **	-0,058 .	-0,071 .
<i>Vanus: 31-45³</i>	-0,105 *	0,021	-0,079	-0,061	-0,175 **	-0,011
<i>Vanus: 46-60³</i>	-0,284 ***	-0,047	-0,251 ***	-0,241 ***	-0,331 ***	-0,183 **
<i>Vanus: >60³</i>	-0,262 ***	-0,006	-0,297 ***	-0,267 ***	-0,331 ***	-0,257 ***
<i>Kutseharidus⁴</i>	-0,042	-0,126	-0,090	-0,049	-0,066	-0,111
<i>Kõrgharidus⁴</i>	-0,042	-0,159 *	-0,179 *	-0,103	-0,139 .	-0,135
<i>Üldkeskharidus⁴</i>	-0,080	-0,137 .	-0,103	-0,092	-0,081	-0,108
<i>Vene keel⁵</i>	-0,218 ***	0,329 ***	-0,275 ***	-0,176 ***	-0,171 ***	-0,270 ***
<i>Kaks või kolm meeldetuletust</i>	0,047	-0,014	-0,177 **	0,042	-0,072 *	-0,023
N	704	489	581	632	562	569
AIC	584,200	454,590	641,900	500,790	436,600	618,440
Pseudo R ² : McFadden	0,191	0,325	0,132	0,199	0,204	0,122

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

Mudelisse veel kaasatud: sugu, poliitikaarvi, interneti kasutussagedus

¹ Taust: standardne sissejuhatus

² Taust: informatsiooni tagajärgede osas ei kuvatud

³ Taust: 18-30

⁴ Taust: põhiharidus või madalam

⁵ Taust: eesti keel

Tabel 16. Binaarne logistiline regressioon: hinnang vastamise nautimisele

	Laine 1	Laine 3	Laine 2	Laine 4	Laine 6
	AME	AME	AME	AME	AME
<i>w1w3exp0: vastamise täpsuse rõhutamine¹</i>	0,042	0,005			
<i>w1w3exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine¹</i>	0,107	**	0,015		
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed²</i>			0,064	.	0,032
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed²</i>			0,041		0,060
<i>Sage interneti kasutus</i>	0,057	0,053	0,076	.	0,035
<i>Politiikahuvi</i>	-0,022	0,004	0,007		0,005
<i>Vanus: 31-45³</i>	-0,026	0,021	0,007		0,005
<i>Vanus: 46-60³</i>	-0,045	-0,047	-0,028		0,006
<i>Vanus: >60³</i>	-0,042	-0,006	-0,064		-0,022
<i>Kutseharidus⁴</i>	-0,155	*	-0,126	**	-0,361
<i>Kõrgharidus⁴</i>	-0,159	*	-0,159	*	-0,308
<i>Üldkeskharidus⁴</i>	-0,127	.	-0,137	.	-0,291
<i>Vene keel⁵</i>	0,295	***	0,329	***	0,356
<i>Kaks või kolm meeldetuletust</i>	-0,029		-0,014		-0,084
N	704	489	630	562	568
AIC	719,170	379,400	622,610	555,140	535,810
Pseudo R ² : McFadden	0,084	0,315	0,134	0,126	0,144

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1
Mudelisse veel kaasatud: sugu, rööprähklus
¹ Taust: Standardne sissejuhatus
² Taust: informatsiooni tagajärgede osas ei kuvatud
³ Taust: 18-30
⁴ Taust: põhiharidus või madalam
⁵ Taust: eesti keel

4 Arutelu

4.1 Mittevastamise viga CRONOSes

4.1.1 Interneti mittekasutajad

Tõenäosusliku valimiga veebipaneeli puhul on oluline interneti mittekasutajatele võimaldada uuringus osalemist teisiti. CRONOSes pakuti interneti mittekasutajatele tehnoloogilisi vahendeid. Tööst selgus, et tehnoloogilised vahendid võtsid uuringus osalemiseks kasutada eakamad ja madalama lõpetatud haridusaastate arvuga inividid võrreldes teiste respondentidega. Antud tulemus vastab üldiselt lihtsa digitaalse lõhestumise teesile, võttes arvesse vanuse ja hariduse mõju (Vehovar jt, 2006:281). Lisaks vanuse ja haridustaseme selekterivale mõjule veebiuuringus osalemiseks on esile toodud ka vähemuste piiratud ligipääsu internetile (Callegaro jt, 2015: 112), kuid näiteks vene ja eesti keeles vastanud inividide seas oli tahvelarvutite osakaalude erisus olematu. Vähemuste all ei peeta silmas vaid keelilisi või rahvusest tulenevaid vähemusi, aga ka näiteks elamispiirkonda või religioosset kuuluvust, mida käesoleva töö raames ei uuritud.

On oluline, et tahvelarvuti vastu võtnud inividid ei esinda kõiki CRONOSE valimisse kuulunud interneti mittekasutajaid, moodustades seega vaid osa nendest inivididest, kes muidu uuringust välja jääksid. Sellele vaatamata on tahvelarvuti vastu võtnud inividide kaasamine uuringusse võimaldanud tõsta vanemate ja madalama haridustasemega inividide esindatust CRONOSes, mõjutades seega oluliselt üldist vanus- ja haridusjaotust. CRONOSE tahvelarvuti saanud respondentide andmeid on analüüsinud ka Ainsaar ja Soidla (2018), kes on leidnud, et võrreldes teiste interneti mittekasutajatega võtsid tahvelarvutid vastu pigem nooremad ja inividid, kelle kodukeeleks on eesti keel. Autorid leidsid, et tahvelarvutite jagamine parandas esinduslikkust märkimisväärselt küll kuni 75-aastaste vastajate seas, kuid sellest vanemaid vastajaid oli jätkuvalt oluliselt vähem kui esinduslikkuse saavutamiseks tarvis oleks olnud.

Seega ei võimaldanud tahvelarvutite jagamine CRONOSes muidu veebiuuringust väljajäävat osa ühiskonnast kõikides vanuserühmades esinduslikkuse saavutamiseks piisaval määral siiski uuringusse kaasata (vähemalt vanust arvesse võttes). Sellest lähtuvalt pole veebiuuringute peamine kaetuse probleem, mis seostub eelkõige lihtsa digitaalse lõhestumisega ühiskonnas, Eestis kuhugi kadunud ning tehnoloogiliste vahendite võimaldamine veebiuuringus ei pruugi täielikult tagada soovitud tulemust – valimi esinduslikkust (vastavust üldkogumile oluliste

sotsiaaldemograafiliste tunnuste lõikes). Samas võimaldab tehnoloogiliste vahendite jagamine veebiuuringus siiski peamiste interneti mittekasutajate rühmade, eakate ja madalama haridustasemega indiviidide, esindatust veebiuuringus olulisel määral parandada.

4.1.2 Kutsele mittevastamine

Callegaro jt (2015: 220) on esitanud väite, et kutsele mittevastamist kutsub esile suurem hõivatus, kõrgem vanus ja suurem sissetulek. CRONOSe andmetele tuginedes võib aga öelda, et suurimateks mittevastajateks on eelkõige nooremad, meessoost, põhi- või sellest madalama haridustasemega indiviidid. Seega võib arvata, et indiviide, kes vastamist ei alustanud, võib iseloomustada eeltoodud omadustest vaid suurem hõivatus. Lisaks on leitud, et veebiuuringutes soovivad osaleda pigem nooremad, mehed, kõrgema haridustasemega ja kõrgema sotsiaalse staatusega indiviidid (seostub ka kõrgema digitaalse kirjaoskusega) (Callegaro jt, 2015: 227). CRONOSe andmetele tuginedes ei saa aga antud tulemusi kinnitada. Kuigi kõrgemaid IKT oskusi seostatakse nooremate ja meessoost indiviididega, siis just nende indiviidide kutsele mittevastamise risk (soov mitte osaleda) on kõrge. Vastuolu Callegaro jt (*Ibid.*) tulemustega võib tuleneda sellest, et käesolevas töös on uuritud vastamiskäitumist paneeli jooksul, mitte esialgset nõusolekut uuringus osalemiseks. Ainsaar ja Soidla (2018: 185-186) on leidnud, et kuigi noorte ja meeste seas oli esialgne nõusolek uuringus osalemises kõrge, oli nende kadu uuringu jooksul suur. Järelikult tuleks eristada osalemisega nõustumist ja veebipaneeli jooksul vastamist. Üldiselt saa CRONOSe andmetele toetudes öelda, et kutsele mittevastamine paneeli jooksul oleks seotud madalama digitaalse kirjaoskusega, pigem võib küsitlustele mittevastamist mõjutada indiviidi vähene püsivus, hõivatus või üldine suhtumine uuringutesse.

CRONOSe kutsele mittevastamise analüüsi tulemustele võis mõju avaldada andmestiku ülesehitus, st osa mittevastamisi (u 20 indiviidi, kes ei vastanud ühelgi lainel) on kasutamiseks vabalt saadaolevast andmestikust puudu. Tulemusi võis mõjutada ka asjaolu, et CRONOSe paneeli läbiviimine oli üles ehitatud 2016. aasta Euroopa Sotsiaaluuringu 8. vooru andmekogumisele, mis tähendas, et kontaktiloomine CRONOSe respondentiga ja ka kontaktandmete õigsuse kontroll toimus juba ESS küsitluse raames, suurendades nii tõenäosust uuringukutse kättetoimetamiseks ja uuringus osalemiseks. Dillmani, Smythi ja Christiani (2014: 37) üks väide on, et kui respondent on huvi uuringus osalemise vastu üles näidanud ühe korra, teeb ta seda tõenäoliselt ka järgmine kord, mis võib olla mõjutanud CRONOSe paneelis vastamist. ESS mõju CRONOSe mittevastamisele võis olla ka vastupidine. Intervjueerija mõjul

võis respondent nõustuda uuringus osalema, kuid võis hiljem ka tunda, et on teinud ESSi intervjuule vastates andmekogujate heaks juba piisavalt. Mõju esinemisel avaldus see ilmselt paneeli värbamisel, aga ei saa välistada ka hilisemat mõju.

Lisaks selgus, et eelmisel lainel tahvelarvutiga vastanute risk kutsele mitte vastata on väiksem kui arvutiga vastanutel ning et nutitelefoni kasutajatel esines kuuendal lainel veidi suurem risk kutsele mittevastamiseks võrreldes indiviididega, kes kasutasid eelmisel lainel vastamiseks arvutit. Antud asjaolu võib selgitada see, et tahvelarvutiga vastanute sekka kuulusid jagatud tahvelarvutiga respondendid, kelle puhul võis täheldada väga kohusetundlikku vastamist. Antud tendents võib tuleneda laiendatud Ajzeni teooriale tuginedes moraalsest kohustusest (Bosnjak jt, 2005: 494). Nutitelefoni kasutajate suurema kutsele mittevastamise taga võib olla kõrgem nooremate indiviidide mittevastamine paneelis, sest nutitelefoni kasutajateks on pigem nooremad indiviidid.

4.1.3 Vastamise katkestamine

CRONOSe veebipaneelis katkestanute arv oli kõikidel lainetel suhteliselt madal, moodustades keskmiselt 4% vastamist alustanud respondentidest. Esmane katkestamine toimub kohe sissejuhatavatel lehekülgedel. Kirjeldavale statistikale, Kaplani-Meieri riskihinnangute joonistele 15 ja 16 tuginedes iseloomustab edasist katkestamist koondumine suures osas küsimustiku esimesse poole. Töös esitatud regressioonianalüüsist selgub, et katkestamise risk on suurem 18-30-aastastel, põhiharidusega või madalama haridustasemega ning hilistel vastajatel, kinnitades CRONOSe andmetel ka Görizti ja Grutzeni (2012) uurimistulemusi. Autorid on põhjendanud seost madala digitaalse kirjaoskusega (meeldetuletusi saanud indiviidid võivad vaadata lihtsalt harvemini e-maili), mis võib esile kutsuda ka katkestamist. CRONOSe andmetele tuginedes võib öelda, et meeldetuletustega vastamist alustanud indiviide suure tõenäosusega madala digitaalse kirjaoskusega seostada ei saa, sest katkestamist prognoosib lisaks meeldetuletuste arvule pigem noorimasse vanuserühma kuulumine ning sage interneti kasutus prognoosivat mõju katkestamisele ei oma, mistõttu harvema e-mailide vaatamisega seostatud madalat digitaalset kirjaoskust ei saa veenvalt lugeda katkestamise põhjuseks. Mis täpselt meeldetuletuste arvu katkestamist prognoosiva mõju taga on, tööst ei selgu, kuid võib oletada, et seost võib põhjustada madalam kognitiivne võimekus või üldine suhtumine uuringutesse.

Steinbrecher, Roßmann ja Blumenstiel (2015) on leidnud, et katkestamist kutsuvad esile peamiselt motivatsiooni langus, poliitiline apaatsus, tehnilised probleemid vastamisel, ülesande keerukus ja individuaalne võimekus. Tehnilisi probleeme vastamisel CRONOSe andmetest ei selgunud. Motivatsiooni langusele CRONOses võib viidata indiviidide katkestamise esinemine küsimustiku uut plokki sissejuhatavatel küsimustel või ka sissejuhatavatel ja nõusoleku küsimise lehekülgedel. Lisaks võib motivatsiooni langus peegelduda ka põhi- ja sellest madalama haridustasemega indiviidide katkestamise trendis, et katkestamise risk kasvas oluliselt pärast 25. küsimust, mis võib viidata motivatsiooni langusele tulenevalt küsimuse pikkusest. Täpsemalt võib aimata, et küsimustikule vastamise ajal ületas see oodatud pikkust, kutsudes esile katkestamist.

Huvi poliitika vastu või selle puudumine CRONOSe andmetel katkestamist ei prognoosi, seega võib aimata, et Steinbrecheri, Roßmanni ja Blumenstieli (2015) leid on pigem seotud poliitikateemaliste uuringutega, kus poliitika vastu huvi mittetundvatel indiviididel ongi väga raske küsimustikule täielikke vastuseid anda ja vastamisega ka küsimustiku lõpuni jõuda. CRONOS hõlmas väga laia teemaderingi, mis võib olla tinginud selle tunnuse mõju puudumise.

Uurides katkestanute individuaalset võimekust, selgub, et katkestanuid iseloomustab võrreldes täielikke vastuseid andnud indiviididega kuulumine noorimasse vanuserühma, mida seostatakse üldiselt just kõrgema kognitiivse võimekuse ja digitaalse kirjaoskusega. Kõrgem katkestamise risk on aga põhi- või sellest madalama haridustasemega indiviididel, viidates madalamale kognitiivsele võimekusele. Peytchev (2009: 95) on väitnud, et katkestamist võivad prognoosida uuritava teema suhtes võõras olemine või madalamad kognitiivsed võimed, mis pikendavad vastamiseks kuluvat aega ja seejärel ülemäärase pingutusvajaduse tõttu katkestamist. Katkestanute vastamiseks kuluvat aega töös eraldi ei käsitletud, mistõttu ei ole võimalik täpselt öelda, kas katkestanute vastamisele kulunud aeg ühe küsimuse kohta pikenes, kuid üldiselt kulus noorimal vanuserühmal vastamiseks just kõige vähem aega. Samas võib nooremate ja madalalt haritud indiviidide puhul oletada, et nende jaoks on osad uuritud teemad tõesti võõrad ning neil võivad puududa kindlaks kujunenud arvamused teatud küsimustes, mis võib ka katkestamist esile kutsuda. Katkestamine ülemäärase pingutusvajaduse tõttu on samuti väga tõenäoline madala haridustasemega katkestatute taustal.

4.1.4 Küsimusele mittevastamine

Küsimusele mittevastamist iseloomustab CRONOSe andmetele tuginedes mittevastamise koondumine teatud küsimuste plokkidesse või teatud küsimusele, koondudes sõnalist selgitust või faktilist vastust eeldavatele avatud küsimustele. On oluline, et küsimusele mittevastamiste arv oli märkimisväärselt kõrgem küsimuste puhul, mis olid kognitiivset pingutust nõudvamad, st eelmistele küsimustele tagasivaatavad või enam süvenemist ja küsimuse tõlgendamist nõudvate küsimuste puhul. Mittevastamiste arvu tõstis märgatavalt „ei oska öelda“ vastusevariandi kuvamine koos küsimusega. Sellisel juhul võib mittesisulise vastusevariandi valimist pidada ülesande kergema lahenduse valimisena. Järelikult võib öelda, et ülesande keerukus tõstab märkimisväärselt küsimusele mittevastamist ning kui keerulise küsimuse puhul on võimalik valida kergem tee, siis seda ka tehakse.

Lisaks on töös uuritud tegureid, mis mõjutavad vähemalt ühe küsimusele mittevastamise esinemist indiviidi kohta. Berinsky (2008) järgi tingivad küsimusele mittevastamist peamiselt individuaalsed isikuomadused nagu madal kognitiivne võimekus ja vähene huvi uuritava teema vastu. Madalam haridustase ja kõrgem vanus, mis CRONOSe andmete järgi prognoosivad küsimusele mittevastamist, võivad peegeldada madalamat kognitiivset võimekust. Vanemate respondentide suuremat vaimset pingutust küsimustele vastamisel peegeldab ka pikem vastamisele kulunud aeg võrreldes noorima vanuserühmaga, kuigi nad ei eristunud noorematest respondentidest hinnangutes vastamise keerukusele ning hindasid pingutust täpsete vastuste andmiseks noortest respondentidest madalamaks. Prognoosimudelitest selgus ka, et küsimusele mittevastamist kutsus esile vähene huvi poliitika vastu, mis samuti Berinsky (2008) väitega ühtib. Poliitikahuvi võib peegeldada laiemat huvi CRONOSe raames käsitletud päevakajaliste või ühiskondlike teemade suhtes. Seega võib öelda, et individuaalne võimekus ja huvi uuritava teema vastu omavad positiivset mõju küsimustele sisuliste vastuste andmisele.

4.1.5 Hinnang CRONOSe mittevastamise vea ulatusele

CRONOSe andmetest selgus, et kutsele mittevastamise risk on kõrgem nooremate, meessoost ja põhi- või sellest madalama haridustasemega indiviidide seas. Antud tendents avaldus olukorras, kus interneti mittekasutajatele võimaldati uuringus osalemiseks tehnoloogilised vahendid. Katkestamise risk on samuti kõrgem just nooremate ja põhi- või sellest madalama haridustasemega indiviidide seas. Tundub, et süstemaatiliselt kukuvad paneeli jooksul respondentide seast välja just nooremad ja madala haridustasemega indiviidid. Lisaks on

oluline, et indiviidide seas, kes esimestele uuringukutsetele kohe ei reageerinud ja viivitasid vastamise alustamisega, esines ka märkimisväärselt enam katkestamisi võrreldes nendega, kes said kuni ühe meeldetuletuse. See tähendab, et hiliseid liitujaid võib iseloomustada üldiselt madalam motiveeritus vastamiseks.

Kypri (2004; Kypri jt, 2011) on leidnud, et hilisemad uuringusse kaasatud indiviidid võivad lisada uuringutulemustele olulist lisandväärtust, kui uurimise all on riskikäitumine või faktilised näitajad. CRONOS-st selgunud potentsiaalsed vaatluse alt väljajäävad indiviidid võivad samuti anda olulist lisandväärtust üldkogumi väärtuse hinnangutele, sest süstemaatiliselt jääb paneeli käigus mõnevõrra rohkem vaatluse alt välja (ei vasta üldse või vastavad osaliselt) pigem põhi- ja sellest madalama haridustasemega ning nooremaid indiviide. Seega mittevastamise viga üldkogumi väärtuse hinnangutes on problemaatiline eelkõige faktiliste näitajate ja näiteks hälbiva käitumise küsimustes, mis võib puudutada just väljajäävat osa ühiskonnast.

CRONOSe andmetel on küsimusele mittevastamise risk suurem põhi- või sellest madalama haridustasemega, kuid pigem kõrgema vanusega indiviidide seas. Antud suundumuse tõttu võib mittevastamise viga olla suurem küsimuste puhul, mis nõuavad suurt vaimset pingutust, jättes näiteks faktiliste küsimuste puhul vaatluse alt välja just need ühiskonnagrupid. CRONOSe näitel võib pidada mittevastamise viga suureks nende küsimuste puhul, milles käsitleti haridustaseme omandamise aega ja täpset kvalifikatsiooni, mille meenutamisel eelkõige vanematel respondentidel selgelt raskusi esineb. Samade küsimuste puhul on vastamisel raskusi ka madala haridustasemega indiviididel, kes tõenäoliselt ei saa või ei oska antud küsimustele vastata, otsustades jätta küsimusele vastamata.

4.2 Mõõtmisviga CRONOSes

4.2.1 Joonvastamine

Regressioonanalüüsist selgus, et tõenäosus joonvastamiseks on suurem kõrgema vanuse ja kõrgema haridustasemega indiviidide seas. Antud tulemus kinnitab Peytchevi (2009: 80) väidet, et eakamate indiviidide seas vastuste variatiivsus väheneb. Autor põhjendab suundumust vananemisega kaasneva kognitiivse võimekuse langusega, mis võib ka käesoleva töö tulemustele tuginedes olla variatiivsuse erinevuste mehhanismiks. Samas ei haaku see põhjendus vähenenud vastuste variatiivsusega kõrg- või kutseharidusega indiviidide seas. Göritz ja Grutzen (2012) on aga pakkunud, et madalat andmekvaliteeti võib seostada ka

madalama digitaalse kirjaoskusega. Analüüsist selgunud vanemate respondentide kalduvust joonvastamisele võib teoreetiliselt seostada ka madalamate IKT oskustega, kuid võttes arvesse, et haridustasemel oli käesolevas töös pigem vastupidine mõju, ei saa joonvastamist madalama digitaalse kirjaoskusega seostada.

Kõrgema haridustasemega indiviide iseloomustas võrreldes põhi- või sellest madalama haridustasemega indiviididega ka üldiselt madalam vastamisprotsessi nautimine, mida võib seostada joonvastamisele kaldumisega. Võib oletada, et eakamad ja kõrgema haridustasemega indiviidid võivad tajuda üleküsitletust, andes võimaliku tüdimuse tõttu suhteliselt ebatäpsemaid vastuseid, sest antud indiviidid on näidanud püsivat osalust paneelis ning võib oletada, et antud indiviidide rühmad osalevad tõenäoliselt palju ka teistes veebiuuringutes (toetudes vastamise alustamise mustritele CRONOSes).

Anduiza ja Galais (2017) on leidnud, et joonvastamist esineb enam madalamalt haritute, poliitilisest osalusest vähem huvitatute ja kiirete vastajate seas, kes võivad olla nn. professionaalsed küsitlustele vastajad. CRONOSe andmetel antud väidet kinnitada ei saa, sest joonvastamist esineb peamiselt just kõrgemalt haritute seas ning vähene huvi poliitika vastu ega suur vastamiskiirus joonvastamist ei prognoosi. Läbiviidud regressioonanalüüsi põhjal ei ole võimalik teada saada, millest haridustaseme mõju otseselt tuleneb. Subjektiivselt võib oletada, et haridustaseme mõju viienda laine teises ja kolmandas küsimuste komplektis võib tuleneda kaudselt kõrgemalt haritud indiviidide laiemast silmaringist või tolerantsemast suhtumisest immigratsiooni (D'Hombres ja Nunziata, 2016), kuid erinevust esile kutsuvaid kaudsemaid mehhanisme võib olla mitmeid. Lisaks küsimuste sisust tulenevatele erinevustele võib vastuste variatiivsuse erinevusi esile kutsuda ka vastusskaalast tulenevad eripärad või muud tegurid. Kuuenda laine teises küsimuste komplektis ilmneb haridustaseme vastassuunaline mõju joonvastamisele. Nimelt on kõrgema haridustasemega indiviididel suurem vastuste variatiivsus küsimuste juures, mis käsitlevad suhtumist internetti, mis võib peegeldada kindlakskujunenud arvamusi interneti osas või laiemat silmaringi. Antud küsimuste komplekt sisaldas ka sisu poolest suhteliselt vastuolulisi küsimusi, mille puhul peakski vastuste variatiivsus pigem suurem olema.

Teise, neljanda ja kuuenda laine mudelitest selgub meeste vastuste väiksem variatiivsus küsimuste plokkides, mis käsitlesid lapsekasvatuse, kultuuriliste normide ning interneti teemat. Võib öelda, et nende küsimuste plokkide puhul suhestuvad mehed ja naised küsimustele

vastamise erinevalt, täpsem põhjus analüüsist ei selgu. Meeldetuletuste arvu prognoosiv mõju vastuste variatiivsusele avaldus kahe küsimuste komplekti puhul (lapsekasvatuse ja eluolu küsimused). Võib öelda, et teatud tingimustel on hiliste vastajate tõenäosus joonvastamiseks kõrgem võrreldes kuni ühe meeldetuletuse saanud indiviididega. Saadud tulemus peegeldab antud indiviidide üldisemat madalamat motivatsiooni küsitlusele vastamisel.

Oluliseks tulemuseks võib pidada ka seda, et täpsete vastuste ja hoolsuse rõhutamine ei mõjuta joonvastamist. Samuti on oluline, et vastamiseks kasutataval seadmel puudub enamiku mudelite puhul prognoosiv mõju vastuste variatiivsusele. Mõju ilmnis nutitelefonide kasutajate suurema joonvastamisena ainult laste kasvatamise küsimuste puhul, mida võib seostada vajadusega suhestuda võõra teemaga, sest nutitelefonide kasutajateks olid just nooremad respondendid. Lastekasvatamise küsimuste plokis esines kõikidest analüüsitud küsimuste plokkidest ka enim täielikke joonvastajaid (teine oli internetiga seotud küsimuste plokk). Kuigi ühe mudeli puhul seadme mõju esines vastuste variatiivsuse prognoosimisel, võib pigem öelda, et vastamiseks kasutatud seade ei mõjuta joonvastamist: esinenud mõju võib pigem seostada teatud seadme kasutajate teistest (sotsiaal-demograafilistest) tunnustest põhjustatud erinevustega küsimustega suhestumisel. Samade järeldusteni on jõudnud ka teised autorid (Antoun jt, 2017; Revilla, 2016).

Ühe huvitava asjaoluna selgus, et vene keeles vastanud indiviidide vastustes esineb suuremat variatiivsust kui eesti keeles vastanud indiviididel, nad ise on hinnanud oma pingutust täpsete vastuste andmisel väheseks ning on hinnanud vastamisprotsessi enam nauditavaks kui eesti keeles vastanud indiviidid. On ilmne, et vene keeles vastanud indiviidid on paneeli üldiselt suhestunud erinevalt, on võtnud vastamist vabamalt ja nauditavamalt. Antud tendentsi täpset mehhanismi käesoleva töö raames ei uuritud, kuid võib subjektiivselt arvata, et mõju taga võib olla erinev suhestumine uuringu läbiviijasse, küsitluse disaini ja sisusse või laiemalt uuringukultuuri.

Üldiselt võib kinnitada, et joonvastamist kutsub esile vastamise keerulisus, madal individuaalne võimekus ja vähene motivatsioon. Vastamine on keerukam küsimuste puhul, mille teema on vastajale võõras või puuduvad kindlaskujunenud arvamused. Motivatsiooni langust võib esile kutsuda üleküsitletus, sest joonvastamine on kõrgem indiviidide seas, kes on püsivad vastajad CRONOSes. Vastamiseks kasutatud seadme valik joonvastamist ei mõjuta. Samuti ei mõjuta

joonvastamist vastuste täpsuse või hoolsuse rõhutamine. Küll aga esineb teatud tingimustel joonvastamist enam hilistel vastajatel.

4.2.2 Vastamisele kuluv aeg

Vastamisele kulunud aja regressioonanalüüsist selgub, et pikemad vastamisajad on olnud kõrgema vanusega, põhi- ja sellest madalama haridustasemega ning mitut asja vastamisega samaaegselt teinud indiviididel. Hariduse mõju esines kahel lainel, kus põhi- ja sellest madalama haridustasemega indiviidide vastamiseks kuluv aeg oli pikem kui kõrgharidusega indiviididel, kusjuures need küsimustikud olid ka kõige raskema (kolmas ja kuues laine). Antud tulemused kinnitavad Antouni ja Cernati (2019) uurimistulemusi, kuigi haridustaseme mõju nii selgelt ei eristunud. Vastamisele kulunud aja pikenedmist võib seega esile kutsuda individuaalne kognitiivne võimekus, seda eriti keeruliste küsimustike puhul.

Vastamiseks kasutatud seade aga vastamiseks kulunud aega ei prognoosi. Kuigi Antouni ja Cernati (2019) tulemused näitasid erisusi vastamisajas kasutatud seadme osas, siis võib oletada, et CRONOSe küsimustikele kasutasid vastamiseks nutitelefoni indiviidid, kes end nutitelefonis mugavalt tundsid. Lisaks võib tulemus olla seotud CRONOSe küsimustiku disainist tuleneva efektiga, sest vastuste formaatide varieeruvus oli väike ning ei kasutatud problemaatilisi formaate (nt liugurskaala) või maatriksküsimusi (Antoun jt, 2017; Liu ja Cernat, 2018: 701).

Vastamisele kuluv aeg võib viidata respondendi pühendumisele küsitlusele vastamisel. Vastamisele kuluva aja regressioonimudelitesse olid kaasatud ka eksperimendi tunnused, mis rõhutasid või ei rõhutanud olenevalt kontroll- või testrühmast kas täpsuse olulisust vastuste andmisel või hoolika/hooletu vastamise tagajärgi. Regressioonanalüüsist selgus, et vastuste täpsuse rõhutamine ja selleks ka lubaduse küsimine ning hooletu vastamise negatiivsete tagajärgede kuvamine pikendavad vastamiseks kuluvad aega. Antud tulemust võib pidada oluliseks, sest täpsuse ja hoolsuse rõhutamine suurendavad tõenäoliselt pühendumist küsimustele vastamisel, pikendades seeläbi vastamiseks kuluvat aega. Sama kinnitab ka antud eksperimendi testrühma kuuluvate respondentide kõrgem hinnang täpsete vastuste andmiseks tehtud pingutusele, kuid hinnang võib olla mõjutatud ka sotsiaalselt soovitud vastuse andmisest.

Analüüsist selgus, et vastamiseks kulus vähem aega noorematel ja kõrgema haridustasemega indiviididel, mis viitab kõrgemale individuaalsele võimekusele. Vastamiseks kasutatud seade vastamiseks kuluvat aega ei prognoosi. Vastamisaeg võib siiski peegeldada ka andmekvaliteeti,

sest vastamisel täpsete vastuste ja hoolsuse rõhutamine pikendas vastamisele kuluvat aega ning seetõttu võimaldab tõenäoliselt vähendada mõõtmisviga.

4.2.3 Hinnang uuringus osalemise kogemusele

Käesolevas töös on andmekvaliteedi uurimiseks käsitletud ka respondentide subjektiivseid hinnanguid uuringus osalemise kogemusele: küsimustest arusaamise ja vastamise keerukusele, pingutusele täpsete vastuste andmiseks ning vastamise nautimisele. Callegaro jt (2015: 170-173) toovad esile, et küsitluse lõpuosas küsitud hinnangud vastamiskogemusele võivad peegeldada respondendi pühendumist vastamisele.

Tööst selgub, et kui küsimustest arusaamist ja vastamist hindasid keerukamaks poliitikast vähem huvitatud, põhi- ja sellest madalama haridustasemega invidiidid, siis pingutust täpsete vastust andmiseks hindasid kõrgemaks nooremad, põhiharidusega või sellest madalama haridustasemega invidiidid ning need, kellele kuvati vastuste täpsust või hoolsust rõhutavat teavet. On näha mustrit, et vastamisel kogesid raskust eelkõige madalama kognitiivse võimekusega respondendid. Täpsete vastuste andmisel on lisandunud teatud lisanüanss, et nooremad respondendid just pingutust kogesid. Antud tendents võib olla seotud uuritava suhestumisega, st noored võivad olla uuritava teema suhtes võõrad või ka vähese elukogemusega, mis võimaldaks anda küsimustele kaalutletud ja täpseid vastuseid. Samas on huvitav, et just haritumad on hinnanud vastamist vähem nauditavaks, kuigi nad raskusi vastamisel ei kogunud. Haritud invidiidide vähene vastamise nautimine võib tuleneda juba eespool nimetatud üleküsitletuse fenomenist. Vastamise nautimise hinnangud avaldusid ka joonvastamise muustrites. Vastamist nautisid peamiselt põhiharidusega või sellest madalama haridustasemega, vene keeles vastanud respondendid ning need, kellele kuvati täpsust või hoolsust rõhutavat teavet. Vastamist nautisid vähem vastamisega viivitanud invidiidid, kes on näidanud välja üldiselt madalamat motivatsiooni uuringus osalemiseks.

Hinnangud uuringus osalemise kogemusele võivad siiski peegeldada andmete kvaliteeti. Revilla ja Ochoa (2015) on leidnud, et madal hinnang küsimustikule vastamiseks antud panusele ei ole madala andmekvaliteediga otseselt seotud, põhjendades seose puudumist sotsiaalselt soovitatavate vastuste andmisega hinnangute küsimustele. Käesolev töö kinnitab aga, et teatud seosed andmekvaliteedi ja hinnangute vahel eksisteerivad. Töös seoseid empiiriliselt ei uuritud, seetõttu jääb seoste olulisus ja prognoosiva mõju suurus selgusetu, kuid on võimalikuks uurimisteenaks tulevastes uuringutes.

4.2.4 Hinnang CRONOSe mõõtmisvea ulatusele

CRONOSe andmetele tuginedes selgub, et kuigi vastuste variatiivsuse prognoosimisel eksperimendi korras vastuste täpsuse või hoolsuse rõhutamine statistiliselt olulist mõju ei oma, prognoosib see vastamisele kulunud aja pikenemist ning hinnanguid vastamise keerukusele ja täpsete vastuste andmiseks antud panusele. Mõju on oluline just vastuste täpsuse rõhutamise ja selleks ka lubaduse küsimisel ning hooletute vastuste negatiivsete tagajärgede rõhutamisel. Antud tulemus võib olla märk nimetatud testrühmadesse kuulunud indiviidide suuremast pühendumisest vastamisele ning seeläbi väiksemast mõõtmisveast võrreldes teise testrühma ja kontrollrühma tulemustega. Samas tuleb meeles pidada ka Revilla ja Ochoa (2015) väidet, et subjektiivsed hinnangud võivad esindada sotsiaalselt soovitavaid vastuseid.

Kiired vastajad olid eelkõige noorema vanuserühma esindajad, kes andsid üldiselt suurema variatiivsusega vastuseid. Antud tulemust võib seostada kõrge kognitiivse võimekusega, kuid küsimustele vastamise keerukuse hinnangutes ealised erisused puudusid ning pingutust täpsete vastuste andmiseks hindasid kõrgemaks just nooremad invidiidid. Seega nooremad invidiidid on tõenäoliselt lihtsalt kiiremad lugejad ja vastajad, arvestades ka asjaolu, et täpsete vastuste andmisele võisid noored enda hinnangul oodatust enam aega kulutada. Lisaks selgus, et noorematel respondentidel, kes hindasid täpsete vastuste andmise pingutust vanematest respondentidest suuremaks, esines vähem küsimusele mittevastamisi. Antud tulemus võib viidata suuremale mõõtmisveale 18-30-aastaste respondentide seas seoses võimalusega, et noored võivad olla andnud suhteliselt juhuslikke vastuseid, samas pole selline suhteline juhuslikkus avaldunud joonvastamises.

Vastamiseks kasutatud seade ei mõjutanud vastamiseks kuluvat aega. Seadme mõju vastuste variatiivsusele avaldus vaid ühe küsimuste komplekti puhul, kus nutitelefoni kasutajad eristusid arvutiga vastanutest. Saadud tulemust võiks seostada pigem teistest teguritest tuleneva kaudse mõjuga, sest nutitelefoni aktiivsemat kasutust võib seostada nooremate invidiididega, kellele võivad lapsekasvatusega seotud küsimused olla võõrad. Seega võib öelda, et küsitlusele vastamise vahendite varieeruvus üldiselt ei mõjutanud CRONOSe mõõtmisviga.

Vastamisega viivitanud respondentide puhul ilmnnes teatud tingimustel kõrgem joonvastamise risk. Meeldetuletuste arv ei prognoosi vastamisele kuluvat aega, nõrk mõju raskusele küsimustele vastamisel avaldus vaid ühe laine puhul ning pingutust täpsete vastuste andmisel hindasid hilised vastajad madalamalt. Hiliste vastajate puhul võib siiski tähendada vastamise

väiksemat nautimist. Antud tulemuste valguses saab öelda, et hiliste vastajate kaasamine uuringusse mõjutab mõõtmisviga vähesel määral.

Kokkuvõte

Veebipaneelide võimaliku rakendatavuse hindamiseks tõenäosuslikul valimil põhinevates kogu rahvastiku uuringutes on vajalik metodoloogiline eeltöö, milline on veebipaneelide andmekvaliteet. Eesti osalus CRONOSes, esimeses rahvusvahelises tõenäosuslikul valimil põhinevas ühtlustatud metodoloogiaga veebipaneeli pilootprojektis, on andnud erakordse võimaluse uurida, millised tegurid mõjutavad andmekvaliteeti taolises veebipaneelis Eesti andmetel, mis oli ka käesoleva magistritöö eesmärgiks.

Töö empiirilisest osast selgus, et peamiselt ei alustanud küsitlustele vastamist pigem nooremad, meessoost ning põhi- ja sellest madalama haridustasemega invidiidid. Mittejhuslik invidiidide väljajäämine küsitlusest mõjutab aga uurimistulemusi. Eelkõige võivad uurimistulemused olla mõjutatud küsimuste puhul, mis puudutab enam just uuringust väljajäävaid invidiide, nt faktilised või hälbiva käitumise küsimused (Kypri, 2004; Kypri jt, 2011). Lisaks selgus, et küsimusele vastamata jätmine mõjutab tulemusi eelkõige küsimuste puhul, mis nõuavad suurt vaimset pingutust. Vastamist alustanute seas jätsid küsimustele enam vastamata või andsid mittesisulisi vastuseid vanemad ning põhi- ja sellest madalama haridustasemega invidiidid, kellel selgelt teatud küsimustele vastamine raskusi tekitas. Põhi- ja sellest madalama haridustasemega invidiidide puhul oli väljakutseks ka püsivus vastamisel seoses küsimustiku pikkusega, sest antud invidiidide katkestamise risk kasvas oluliselt alates 25. küsimusest (kokku oli küsimusi ühes laines keskel läbi 85).

Vastamiseks kuluv aeg ei mõjuta joonvastamist. Vastamiseks kulus vähem aega noorematel ja kõrgema haridustasemega invidiididel. Kui kõrgharidusega invidiidide seas esines ka enam joonvastamist, siis nooremate invidiidide seas esines joonvastamist oluliselt vähem kui vanematel respondentidel. Kuigi noorematel vastajatel kulus vastamiseks vähem aega, hindasid nad pingutust täpsete vastuste andmiseks üldiselt kõrgemaks kui vanemad vastajad. Vastamisele kuluv aeg võib siiski peegeldada andmekvaliteeti. Uurimistulemustest selgus, et respondentile küsitluses täpsete vastuste olulisuse rõhutamine ja täpsuseks lubaduse küsimine ning hooletu vastamise negatiivsete tagajärgede rõhutamine suurendavad pühendumist, mis väljendub eelkõige pikenenud vastamisele kulvas ajas ning subjektiivses hinnangus vastamisel täpsete vastuste andmiseks antud pingutusele.

Üheks töös käsitletud teemaks oli mitme meeldetuletuse järgselt vastanute andmekvaliteet. CRONOSe andmetele tuginedes võib öelda, et hiliste vastajate andmekvaliteet on mõnevõrra madalam võrreldes indiviididega, kes said kuni ühe meeldetuletuse. Erinevatele andmekvaliteedi mõõdikutele tuginedes selgus vastamisega viivitanute madal motivatsioon, millele viitab ka kõnealuste indiviidide märkimisväärne risk vastamine katkestada. Kõrgem katkestamise risk kahandab aga oluliselt andmestiku üldist kvaliteeti.

Veebipaneele on tõenäosuslikul valimil põhinevates kogu rahvastiku uuringutes rakendatud vaid loetud näitetel ning rakendatavuse hindamiseks on tarvis täiendavat metodoloogilist tööd. Käesoleva magistritöö aluseks võetud CRONOSe andmestik on mahukas, mistõttu on töös käsitletud vaid väikest osa võimalikest andmekvaliteeti mõjutavatest teguritest. Tulevastes uurimustes on oluline uurida CRONOSe andmekvaliteeti kõikidel lainetel rahvusvahelises võrdluses, mis annaks võimaluse hinnata paremini andmekvaliteeti mõjutavate tegurite olulisust ja universaalsust.

Kasutatud allikad

A survey future online. (2019) WP7: A survey future online, SERISS kodulehekülj. Kasutatud 22.05.2019, <https://seriss.eu/about-seriss/work-packages/wp7-a-survey-future-online/>

Aalen, O., Borgan, Ø., Gjessing, S., & Gjessing, H. (2008) Survival and event history analysis: A process point of view. Retrieved from <https://ebookcentral-proquest-com.ezproxy.utlib.ut.ee>

Ainsaar, M., Soidla, I. (2018) Mis aitab saada digiühiskonna liikmeks? Riigikogu Toimetised 37/2018. Kasutatud 18.05.2019, <https://rito.riigikogu.ee/wordpress/wp-content/uploads/2018/06/Ainsaar-Soidla.pdf>

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179 – 211.

Andres, L. (2016) *Designing & Doing Survey Research*. Sage.

Anduiza, E., Galais, C. (2017) Answering Without Reading: IMCs and Strong Satisficing in Online Surveys, *International Journal of Public Opinion Research*, Volume 29, Issue 3, 1 September 2017, Pages 497–519, <https://doi-org.ezproxy.utlib.ut.ee/10.1093/ijpor/edw007>

Antoun, C., & Cernat, A. (2019) Factors Affecting Completion Times: A Comparative Analysis of Smartphone and PC Web Surveys. *Social Science Computer Review*. <https://doi.org/10.1177/0894439318823703>

Antoun, C., Couper, M.P. & Conrad, F.G. (2017) Effects of Mobile versus PC Web on Survey Response Quality: A Crossover Experiment in a Probability Web Panel, *Public Opinion Quarterly*, Volume 81, Issue S1, 1 Pages 280–306, <https://doi.org/10.1093/poq/nfw088>

Berinsky, A. (2008) Survey non-response. In Donsbach, W., & Traugott, M. W. *The SAGE handbook of public opinion research* (pp. 309-321). London: SAGE Publications Ltd doi: 10.4135/9781848607910

Berzelak, N., Weber, W., and M. Revilla (2018) Analysis of panellist responding behaviour in the CRONOS welcome survey. Deliverable 7.11 of the SERISS project funded under the

European Union's Horizon 2020 research and innovation programme GA No: 654221.
Available at: www.seriss.eu/resources/deliverables

Berzelak, N., Weber, W., M. Revilla (2018) Analysis of panellist responding behaviour in the CRONOS welcome survey. Deliverable 7.11 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables

Bethlehem, J., & Biffignandi, S. (2011) Handbook of web surveys (Vol. 567). John Wiley & Sons. <http://onlinelibrary.wiley.com/book/10.1002/9781118121757>

Biemer, P.P.& Christ, S.L. (2008) Weighting Survey Data. E.D. De Leeuw, J.J. Hox, D.A. Dillman,. (Ed.) International Handbook of Survey Methodology. CRC Press, Taylor & Francis Group, pp 317 – 341. <http://joophox.net/papers/SurveyHandbookCRC.pdf>

Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A.-S., Das, M., Douhou, S., & Krieger, U. (2016) A Comparison of Four Probability-Based Online and Mixed-Mode Panels in Europe. *Social Science Computer Review*, 34(1), 8–25. <https://doi.org/10.1177/0894439315574825>

Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A.-S., Das, M., Douhou, S., & Krieger, U. (2016) A Comparison of Four Probability-Based Online and Mixed-Mode Panels in Europe. *Social Science Computer Review*, 34(1), 8–25. <https://doi.org/10.1177/0894439315574825>

Bosnjak, M. , Tuten, T. L. and Wittmann, W. W. (2005) Unit (non)response in Web-based access panel surveys: An extended planned-behavior approach. *Psychology & Marketing*, 22: 489-505. doi:10.1002/mar.20070

Callegaro, M. (2013) Paradata in Web Surveys. Kreuter, F. (Ed.). Improving surveys with paradata : analytic uses of process information. John Wiley & Sons.

Callegaro, M.& Krosnick, J. A. (2014) Introduction to part I. In M. Callegaro et al. (Eds), Online panel research: A data quality perspective , pp.56-60. John Wiley & Sons, Incorporated, 2014. ProQuest Ebook Central, <https://ebookcentral-proquest-com.ezproxy.utlib.ut.ee/lib/tartu-ebooks/detail.action?docID=1662761>

Callegaro, M., Baker, R. P., & Bethlehem, J. (Eds.). (2014) Online panel research: a data quality perspective. Retrieved from <https://ebookcentral-proquest-com.ezproxy.utlib.ut.ee>

Callegaro, M., DiSogra, C. (2008) Computing Response Metrics for Online Panels, *Public Opinion Quarterly*, Volume 72, Issue 5, December Pages 1008–1032, <https://doi.org/10.1093/poq/nfn065>

Callegaro, M., Lozar Manfreda, K., & Vehovar, V. (2015) Web survey methodology. [E-raamat]. Sage. Recieved from <https://play.google.com/store/books?hl=en>

Couper, M. P. (2000) Web surveys. A REVIEW OF ISSUES AND APPROACHES. *Public Opinion Quarterly* Volume 64: 464–494. Kasutatud 12.02.2019, <https://pdfs.semanticscholar.org/e008/8a042873136464bd5b993feeb47f67301d3.pdf>

CRONOS (i.a) CROss-National Online Survey (CRONOS) panel. ESS kodulehekülj. Kasutatud 13.03.2019, https://www.europeansocialsurvey.org/methodology/methodological_research/modes_of_data_collection/cronos.html

CROss-National Online Survey panel [CRONOS Administrative data, edition 1.0 round 8] (2018) NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of CRONOS data for ESS ERIC

CROss-National Online Survey panel [CRONOS Correspondence data, edition 1.0 round 8] (2018) NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of CRONOS data for ESS ERIC

CROss-National Online Survey panel [CRONOS Paradata, edition 1.0 round 8] (2018) NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of CRONOS data for ESS ERIC

CROss-National Online Survey panel [CRONOS0-6 ESS8 integrated data file, edition 1.0, round 8] (2018) NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of CRONOS data for ESS ERIC

D'Hombres, B, Nunziata, L. (2016) Wish you were here? Quasi-experimental evidence on the effect of education on self-reported attitude toward immigrants, *European Economic Review*, Volume 90, Pages 201-224, <https://doi.org/10.1016/j.eurocorev.2016.02.007>

De Leeuw, E. D. (2012) Counting and Measuring Online: The Quality of Internet Surveys. *BULLETIN DE METHODOLOGIE SOCIOLOGIQUE*, 68. Retrieved from <http://ezproxy.utlib.ut.ee/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbl&AN=RN343320991&site=eds-live>

De Leeuw, E., Hox, J.J. & Dillman, D. A. (2008) *Cornerstones of Survey Research*. E.D. De Leeuw, J.J. Hox, D.A. Dillman, (Ed.) International Handbook of Survey Methodology. CRC Press, Taylor & Francis Group, pp 1 – 17 . Kasutatud 20.04.2019, <http://joophox.net/papers/SurveyHandbookCRC.pdf>

Dillman, D.A., Smyth, J.D, Christian, L.M. (2014) Internet, Phone, Mail, and Mixed-Mode Surveys : The Tailored Design Method. John Wiley & Sons. ProQuest Ebook Central.

DiSogra, C. & Callegaro, M. (2009) Computing Response Rates for Probability-Based Web Panels. Kasutatud 20.04.2019 https://www.researchgate.net/publication/253456344_Computing_Response_Rates_for_Probability-Based_Web_Panels

Eckman, S. (2016) Does the Inclusion of Non-Internet Households in a Web Panel Reduce Coverage Bias? *Social Science Computer Review*, 34(1), 41–58. <https://doi.org/10.1177/0894439315572985>

ESS (i.a.) European Social Survey kodulehekül. Kasutatud 13.05.2019, <https://www.europeansocialsurvey.org/>

Fielding, N., Lee, R., & Blank, G. (2017) The SAGE Handbook of online research methods. 55 City Road, London: SAGE Publications Ltd doi: 10.4135/9781473957992

Galesic, M., Bosnjak, M. (2009) Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey, *Public Opinion Quarterly*, Volume 73, Issue 2, 1 January 2009, Pages 349–360, <https://doi-org.ezproxy.utlib.ut.ee/10.1093/poq/nfp031>

Groves R.M., Fowler F.J., Couper M.P., Lepkowski J.M., Singer E. and Tourangeau R. (2009) *Survey Methodology*, 2nd Edition. Hoboken: Wiley.

Groves, R. M. (1989) *Survey Errors and Survey Costs*. New York: Wiley.

Groves, R., Singer, E., & Corning, A. (2000) Leverage-Saliency Theory of Survey Participation: Description and an Illustration. *The Public Opinion Quarterly*, 64(3), 299-308. Retrieved from <http://www.jstor.org.ezproxy.utlib.ut.ee/stable/3078721>

Groves, R., Singer, E., & Corning, A. (2000) Leverage-Saliency Theory of Survey Participation: Description and an Illustration. *The Public Opinion Quarterly*, 64(3), 299-308. Retrieved from <http://www.jstor.org.ezproxy.utlib.ut.ee/stable/3078721>

Görizt, A. S., & Crutzen, R. (2012) Reminders in Web-Based Data Collection: Increasing Response at the Price of Retention? *American Journal of Evaluation*, 33(2), 240–250. Retrieved from <http://ezproxy.utlib.ut.ee/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ963544&site=eds-live>

Görizt, A. S., Wolff, H.-G., & Goldstein, D. G. (2008) Individual payments as a longer-term incentive in online panels. *Behavior Research Methods*, (4), 1144. Retrieved from <http://ezproxy.utlib.ut.ee/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbl&AN=RN240771728&site=eds-live>

Hamby, T., & Taylor, W. (2016) Survey Satisficing Inflates Reliability and Validity Measures: An Experimental Comparison of College and Amazon Mechanical Turk Samples. *Educational and Psychological Measurement*, 76(6), 912–932.

Hansen, J. (2008) Panel surveys. In Donsbach, W., & Traugott, M. W. *The SAGE handbook of public opinion research* (pp. 330-339). London: SAGE Publications Ltd doi: 10.4135/9781848607910

Healey, B. (2007) Drop Downs and Scroll Mice: The Effect of Response Option Format and Input Mechanism Employed on Data Quality in Web Surveys. *Social Science Computer Review*, 25(1), 111–128. <https://doi.org/10.1177/0894439306293888>

Holbrook, A. L., Krosnick, J. A., & Pfent, A. (2008) The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In J. M. Lepkowski et al (Eds.), *Advances in telephone survey methodology*, 499-528. Hoboken: John Wiley & Sons.

ISO 26362 (2009) International Organization for Standardization. ISO 26362 - Access panels in market, opinion and social research — Vocabulary and service requirements Kasutatud 11.02.2019, <https://www.iso.org/obp/ui/#iso:std:iso:20252:ed-2:v1:en>

ITK usage in Households and by individuals (2019) Connection to the Internet and Computer Use. Eurostat and mebaas. Kasutatud 12.02.2019, <https://ec.europa.eu/eurostat/web/digital-economy-and-society/data/database>

Jans, M., Sirkis, R., Morgan, D. (2013) *Managing Data Quality Indicators with Paradata Based Statistical Quality Control Tools: The Keys to Survey Performance*. Kreuter, F. (Ed.). *Improving surveys with paradata : analytic uses of process information*. John Wiley & Sons

Kivistik, M. (2012) *Intervjuu kestus ja andmete kvaliteet*. Magistritöö. Tartu Ülikool, Sotsiaal- ja Haridusteaduskond.

Kreuter, F., Olson, K. (2013) *Paradata for Nonresponse Error Investigation*. Kreuter, F. (Ed.). *Improving surveys with paradata : analytic uses of process information*. John Wiley & Sons

Krosnick, J. A. (1991) Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi-org.ezproxy.utlib.ut.ee/10.1002/acp.2350050305>

Kruskal, W., & Wallis, W. (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583-621. doi:10.2307/2280779

Kypri, K. (2004) Assessment of Nonresponse Bias in an Internet Survey of Alcohol Use. *Alcoholism: Clinical & Experimental Research*, 28(4), 630–634. <https://doi-org.ezproxy.utlib.ut.ee/10.1097/01.ALC.0000121654.99277.26>

Kypri, K., Samaranayaka, A., Connor, J., Langley, J. D., & MacLennan, B. (2011) Non-response bias in a web-based health behaviour survey of New Zealand tertiary students. *Preventive*

Medicine, (4–5), 274. Retrieved from <http://search.ebscohost.com.ezproxy.utlib.ut.ee/login.aspx?direct=true&db=edsbl&AN=RN301552566&site=eds-live>

Liu, M., & Cernat, A. (2018) Item-by-item Versus Matrix Questions: A Web Survey Experiment. *Social Science Computer Review*, 36(6), 690–706. <https://doi.org/10.1177/0894439316674459>

Lohr, S.L. (2008) *Coverage and Sampling*. E. D. De Leeuw, J.J. Hox, D.A. Dillman,. (Ed.) International Handbook of Survey Methodology. CRC Press, Taylor & Francis Group. <http://joophox.net/papers/SurveyHandbookCRC.pdf>

Lozar Manfreda K., Vehovar, V. (2002) Survey Design Features Influencing Response Rates in Web Surveys, 2002 International Conference on Improving Surveys, August 25–28, Copenhagen, Denmark. Kasutatud 03.05.2019. <https://pdfs.semanticscholar.org/7b40/4cab3ebd5575ed69e6323dd0a0601836c059.pdf>

Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008) Web surveys versus other survey modes: a meta-analysis comparing response rates. *International Journal of Market Research*, (1), 79. Retrieved from <http://ezproxy.utlib.ut.ee/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbl&AN=RN222586196&site=eds-live>

Lozar Manfreda, K., Vehovar, V. (2008) E.D. De Leeuw, J. J. Hox, D. A. Dillman,. (Ed.) International Handbook of Survey Methodology. CRC Press, Taylor & Francis Group, pp. 264 – 284.

Lyberg, L. & Weisberg, H. (2016) Total survey error: a paradigm for survey methodology. In Wolf, C., Joye, D., Smith, T. W., & Fu, Y. The SAGE Handbook of survey Methodology (pp. 27-40). 55 City Road, London: SAGE Publications Ltd doi: 10.4135/9781473957893

McMahon, L. & Stamp, R. (2009) Questionnaire intelligence: New rules of engagement for online survey design . Presented at the Worldwide Readership Research Symposia , Valencia, Spain. Kasutatud 22.05.2019, <https://www.pdf.net/wp-content/uploads/2013/03/965.pdf>

Methodological manual 2018 Part II Household survey (2018) CIRCABC, Euroopa Komisjoni kodulehekülj. Kasutatud 12.02.2019, <https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp>

Mood, C. (2010) Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. *European Sociological Review*, 26 (1), 67-82.

National Research Council. (2013) Nonresponse in Social Science Surveys: A Research Agenda. Roger Tourangeau and Thomas J. Plewes, Editors. Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, <https://doi-org.ezproxy.utlib.ut.ee/10.17226/18293>

Peytchev, A. (2009) Survey Breakoff. *The Public Opinion Quarterly*, 73(1), 74-97. Retrieved from <http://www.jstor.org.ezproxy.utlib.ut.ee/stable/25548063>

Peytchev, A. (2013) Consequences of Survey Nonresponse. *The ANNALS of the American Academy of Political and Social Science*, 645(1), 88–111. <https://doi.org/10.1177/0002716212461748>

Ravilla, T. (2016) Smartphones vs PCs: Does the Device Affect the Web Survey Experience and the Measurement Error for Sensitive Topics? A Replication of the Mavletova & Couper's 2013 Experiment. *Survey Research Methods*, Vol. 10, No. 2, pp. 153-169 [doi:10.18148/srm/2016.v10i2.6274](https://doi.org/10.18148/srm/2016.v10i2.6274)

Revilla, M., & Ochoa, C. (2015) What are the Links in a Web Survey Among Response Time, Quality, and Auto-Evaluation of the Efforts Done? *Social Science Computer Review*, 33(1), 97–114. <https://doi.org/10.1177/0894439314531214>

Schwarz, N., Knäuper, B., Oyserman, D. & Stich, C. (2008) The psychology of asking questions E.D. De Leeuw, J.J. Hox, D.A. Dillman., (Ed.) International Handbook of Survey Methodology. CRC Press, Taylor & Francis Group, pp 317 – 341. <http://joophox.net/papers/SurveyHandbookCRC.pdf>

Simon, H. A. (1956) Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129–138. <https://doi-org.ezproxy.utlib.ut.ee/10.1037/h0042769>

Sommer, E. (2017) Summary of existing web panel strategies. Deliverable 7.1 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables

Steinbrecher, M., Roßmann, J., Blumenstiel, J. E. (2015) Why Do Respondents Break Off Web Surveys and Does It Matter? Results From Four Follow-up Surveys, *International Journal of Public Opinion Research*, Volume 27, Issue 2, pp 289–302, <https://doi-org.ezproxy.utlib.ut.ee/10.1093/ijpor/edu025>

Sturgis, P., Allum, N. & Brunton-Smith, I. (2009) Attitudes Over Time: The Psychology of Panel Conditioning. Lynn, P. (Ed.). *Methodology of longitudinal surveys*, pp. 113 – 126.

Sue, V. M., & Ritter, L. A. (2012) *Conducting online surveys*. SAGE publications.

Toepoel, V. (2016) *Doing surveys online*. Sage.

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013) *The Science of Web Surveys*. Oxford University Press. Kasutatud 27.02.2019, https://books.google.ee/books?id=kqzUCE6KI74C&printsec=frontcover&hl=et&source=gbs_ViewAPI&redir_esc=y#v=onepage&q&f=false

Tourangeau, R., Groves, R. M., Kennedy, C. & Yan, T. (2009) The Presentation of a Web Survey, Nonresponse and Measurement Error among Members of Web Panel. *Journal of Official Statistics*, Vol. 25, No. 3, 2009, pp. 299–321. <https://www.scb.se/contentassets/f6bcee6f397c4fd68db6452fc9643e68/the-presentation-of-a-web-survey-nonresponse-and-measurement-error-among-members-of-web-panel.pdf>

Unt T, Solvak M, Vassil K. (2017) Does Internet voting make elections less social? Group voting patterns in Estonian e-voting log files (2013–2015). *PLoS ONE* 12(5): e0177864. <https://doi.org/10.1371/journal.pone.0177864>

Walston, J. T. , Lissitz, R. W. & Rudner, L. M. (2006) The Influence of Web-based Questionnaire Presentation Variations on Survey Cooperation and Perceptions of Survey Quality. *Journal of Official Statistics*, Vol. 22, No. 2, 2006, pp. 271–291.

Vehovar, V., Sicherl, P., Hüsing, T. & Dolnicar, V. (2006) Methodological Challenges of Digital Divide Measurements, *The Information Society*, 22:5, 279-290, DOI: 10.1080/01972240600904076

Villar, A., Sommer, E., Finnøy, D., Gaia, A., Berzelak, N., & Bottoni, G. (2018) CROss-National Online Survey (CRONOS) panel data and documentation user guide. London: ESS ERIC.

Yan, T. ja Tourangeau, R. (2008) Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times. *Applied Cognitive Psychology* 22(1): 51–68. <http://doi.wiley.com/10.1002/acp.1331>

Lisad

Lisa 1. CRONOSe Eesti andmestikku kuulunud indiviidide vastamine ja mittevastamine kõikidel paneeli lainetel (N=806)

<i>Staat</i>	<i>0</i>		<i>1</i>		<i>2</i>		<i>3</i>		<i>4</i>		<i>5</i>		<i>6</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<i>Täielik vastamine ühes osas</i>	596	94,3	652	92,6	563	88,9	541	90,3	532	94,7	544	93,2	512	90,0
<i>Täielik vastamine pärast pausi</i>	36	5,7	52	7,4	70	11,1	58	9,7	30	5,3	40	6,8	57	10,0
<i>Täielik vastamine. Osakaal alustanutest.</i>	632	95,6	704	96,4	633	95,0	599	96,0	562	96,7	584	97,2	569	94,8
<i>Katkestamine. Osakaal alustanutest.</i>	29	4,4	26	3,6	33	5,0	25	4,0	19	3,3	17	2,8	31	5,2
<i>Alustanuud. Osakaal kutse saanutest.</i>	661	90,1	730	92,1	666	92,2	624	87,3	581	81,4	601	84,4	600	84,9
<i>Katkestamine. Osakaal kutse saanutest.</i>	29	4,0	26	3,3	33	4,6	25	3,5	19	2,7	17	2,4	31	4,4
<i>Kutsele mittevastamine. Osakaal kutse saanutest.</i>	73	9,9	63	7,9	56	7,8	91	12,7	133	18,6	111	15,6	107	15,1
<i>Kutse saanuud. Osakaal kutsututest ja mitte kutsututest.</i>	734	91,6	793	98,9	722	99,7	715	99,7	714	99,9	712	99,9	707	99,2
<i>Mitteaktiivne e-mail.</i>	3	0,4	2	0,2	0	0,0	1	0,1	0	0,0	0	0,0	0	0,0
<i>Ei õnnestunud kontakti luua</i>	0	0,0	1	0,1	0	0,0	0	0,0	0	0,0	0	0,0	2	0,3
<i>Ei kutsutud. Osakaal kutsututest ja mitte kutsututest.</i>	64	8,0	6	0,7	2	0,3	1	0,1	1	0,1	1	0,1	4	0,6
<i>Kokku: Kutsutud ja mitte kutsutud</i>	801	99,4	802	99,5	724	89,8	717	89,0	715	88,7	713	88,5	713	88,5
<i>Puuduvad väärtused</i>	5	0,6	4	0,5	82	10,2	89	11,0	91	11,3	93	11,5	93	11,5
<i>Kokku:</i>	806	100,0	806	100,0	806	100,0	806	100,0	806	100,0	806	100,0	806	100,0

Lisa 2. Küsimusele mittevastamine CRONOS veebipaneeli kõikidel lainetel

	Küsimusele mittevastamise tüüp	N	Keskmine	Mediaan	Std. hälve	95% protsentiil	99% protsentiil	Min	Max
0. laine	"Ei oska öelda"	6	0,10	0	0,53	0,00	3,00	0	3
	"Eelistan mitte vastata"	4	0,06	0	0,50	0,00	1,56	0	4
	Puuduv väärtus	4	0,06	0	0,50	0,00	1,56	0	4
	Küsimusele mittevastamine kokku	0	0,00	0	0,00	0,00	0,00	0	0
1. laine	"Ei oska öelda"	0	0,00	0	0,00	0,00	0,00	0	0
	"Eelistan mitte vastata"	0	0,00	0	0,00	0,00	0,00	0	0
	Puuduv väärtus	239	2,06	0	22,09	0,00	0,00	0	239
	Küsimusele mittevastamine kokku	239	2,03	0	21,91	0,00	0,00	0	239
2. laine	"Ei oska öelda"	0	0,00	0	0,00	0,00	0,00	0	0
	"Eelistan mitte vastata"	0	0,00	0	0,00	0,00	0,00	0	0
	Puuduv väärtus	0	0,00	0	0,00	0,00	0,00	0	0
	Küsimusele mittevastamine kokku	0	0,00	0	0,00	0,00	0,00	0	0
3. laine	"Ei oska öelda"	359	3,15	1	10,27	9,35	19,83	0	106
	"Eelistan mitte vastata"	57	0,50	0	0,98	3,00	3,88	0	5
	Puuduv väärtus	158	1,39	0	5,15	5,35	13,83	0	52
	Küsimusele mittevastamine kokku	574	4,86	1	11,90	15,30	48,73	0	111
4. laine	"Ei oska öelda"	0	0,00	0	0,00	0,00	0,00	0	0
	"Eelistan mitte vastata"	0	0,00	0	0,00	0,00	0,00	0	0
	Puuduv väärtus	0	0,00	0	0,00	0,00	0,00	0	0
	Küsimusele mittevastamine kokku	0	0,00	0	0,00	0,00	0,00	0	0
5. laine	"Ei oska öelda"	5	0,06	0	0,28	0,00	1,14	0	2
	"Eelistan mitte vastata"	0	0,00	0	0,00	0,00	0,00	0	0
	Puuduv väärtus	314	3,57	0	20,88	0,00	127,38	0	150
	Küsimusele mittevastamine kokku	319	3,54	0	20,80	1,00	127,08	0	152
6.laine	"Ei oska öelda"	419	3,64	1	13,47	9,90	66,20	0	119
	"Eelistan mitte vastata"	51	0,55	0	1,19	3,40	5,08	0	6
	Puuduv väärtus	218	1,77	0	8,09	5,90	45,62	0	69
	Küsimusele mittevastamine kokku	688	5,50	1	15,08	14,00	71,28	0	120

Lisa 3. Binaarse logistilise regressiooni mudelid kõikidel lainetel: kutsele mittevastamine

	Laine 0	Laine 1	Laine 2	Laine 3	Laine 4	Laine 5	Laine 6	Laine 3	Laine 4	Laine 5	Laine 6				
	AME	AME	AME	AME	AME	AME	AME	AME	AME	AME	AME				
	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE				
<i>Sugu: mees</i> ¹	0,028	0,042 *	0,052 *	0,063	0,079	**	0,043	0,051	.	0,029	0,034	0,000	0,039	.	
<i>Vanus: 31-45</i> ²	0,023	0,020	0,021	0,026	0,030		0,028	0,028	*	0,024	0,026	0,020	0,022		
	-0,009	-0,049	-0,028	-0,076	-0,043	-0,049	-0,004	-0,088	*	-0,088	*	-0,043	-0,013	-0,032	
	0,033	0,032	0,031	0,037	0,045		0,041	0,041		0,038		0,044	0,032	0,030	
<i>Vanus: 46-60</i> ²	-0,017	-0,086 **	-0,033	-0,064	-0,088	.	-0,052	-0,097	*	-0,085	*	-0,093	*	-0,041	-0,038
	0,035	0,032	0,034	0,040	0,046		0,042	0,043		0,040		0,044	0,031	0,034	
<i>Vanus: >60</i> ²	-0,047	-0,084 *	-0,083 **	-0,108	-0,196 ***		-0,137 ***	-0,167 ***		-0,121 **	**	-0,151 ***	-0,048	-0,060	.
	0,036	0,035	0,031	0,041	0,043		0,040	0,043		0,040		0,041	0,032	0,034	
<i>Kutseharidus</i> ³	0,054	0,001	-0,122 *	-0,084	-0,003		-0,024	-0,103		-0,054	0,047	-0,069	-0,043		
	0,041	0,040	0,057	0,061	0,063		0,058	0,065		0,060	0,050	0,057	0,054		
<i>Kõrgharidus</i> ³	0,025	-0,027	-0,148 **	-0,136	-0,118 *		-0,076	-0,172 **		-0,080	-0,012	-0,055	-0,100	.	
	0,038	0,038	0,055	0,059	0,060		0,056	0,063		0,059	0,047	0,057	0,051	.	
<i>Üldkeskharidus</i> ³	0,019	-0,057	-0,161 **	-0,110	-0,051		-0,025	-0,133 *		-0,065	0,037	-0,061	-0,089	.	
	0,040	0,038	0,055	0,061	0,064		0,060	0,065		0,060	0,052	0,058	0,052	.	
<i>Vene keel</i> ⁴	0,062	0,035	-0,008	0,004	0,058		0,056	0,036		0,006	0,078 *	0,010	-0,004		
	0,033	0,027	0,024	0,032	0,039		0,038	0,036		0,030	0,037	*	0,026	0,026	
<i>Sage interneti kasutus</i>	0,024	-0,014	-0,031	0,010	-0,015		0,014	0,001		0,006	-0,023	0,004	0,039	.	
	0,028	0,029	0,033	0,035	0,043		0,040	0,039		0,033	0,041	0,028	0,024		
<i>Politiikahuvi</i>	0,014	0,000	-0,003	-0,013	-0,001		-0,026	-0,017		0,012	0,007	0,004	0,003		
	0,023	0,020	0,020	0,026	0,029		0,028	0,027		0,023	0,025	0,020	0,022		
<i>Eelmisel lainel kasutatud seade: tahvelarvuti</i> ⁵										0,036	0,077	-0,063 ***	-0,056 ***	***	
										0,083	0,081	0,012	0,011		
<i>Eelmisel lainel kasutatud seade: nutitelefon</i> ⁵										-0,009	-0,021	-0,028	0,048	.	
										0,026	0,026	0,020	0,027		
N	734	793	722	715	714		712	707		658	623	578	596		
AIC	485,750	438,580	385,010	538,740	659,130		607,310	581,980		403,378	411,486	257,459	291,931		
Pseudo R ² : McFadden	0,025	0,053	0,191	0,052	0,072		0,050	0,068		0,308	0,439	0,624	0,558		

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

¹ Taust: naine
² Taust: 18-30 aastat
³ Taust: põhiharidus või madalam
⁴ Taust: eesti keel
⁵ Taust: arvuti

Lisa 4. Binaarne logistiline regressioon: küsimusele mittevastamine

	<i>Laine 1</i>	<i>Laine 3</i>	<i>Laine 5</i>	
	<i>AME</i>	<i>AME</i>	<i>AME</i>	
	<i>SE</i>	<i>SE</i>	<i>SE</i>	
<i>w1w3w5exp0: vastamise täpsuse rõhutamine</i> ¹	0,002	-0,024	0,012	
	0,043	0,052	0,041	
<i>w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine</i> ¹	-0,062	-0,045	0,078	.
	0,042	0,051	0,042	
<i>Rööprähhklus</i>	-0,027	0,041	-0,011	
	0,041	0,052	0,042	
<i>Sage interneti kasutus</i>	0,074	.	-0,084	
	0,045	0,062	0,045	
<i>Poliitika huvi</i>	-0,082	*	-0,116	**
	0,037	0,044	0,036	
<i>Sugu: mees</i> ²	0,081	*	-0,067	
	0,036	0,044	0,036	
<i>Vanus: 31-45</i> ³	0,112	*	0,047	
	0,048	0,059	0,054	
<i>Vanus: 46-60</i> ³	0,061	0,089	-0,049	
	0,049	0,063	0,056	
<i>Vanus: >60</i> ³	0,139	*	0,178	*
	0,059	0,073	0,062	
<i>Kutseharidus</i> ⁴	0,012	-0,253	**	-0,263
	0,078	0,094	0,090	**
<i>Kõrgharidus</i> ⁴	-0,148	.	-0,294	**
	0,076	0,093	0,087	***
<i>Üldkeskharidus</i> ⁴	0,014	-0,260	**	-0,302
	0,080	0,097	0,090	***
<i>Vene keel</i> ⁵	-0,059	-0,006	0,100	*
	0,045	0,054	0,050	
<i>Kaks või kolm meeldetuletust</i>	-0,035	-0,058	-0,018	
	0,048	0,066	0,075	
N	704	489	581	
AIC	887,860	626,210	632,190	
Pseudo R ² : McFadden	0,125	0,358	0,184	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

¹ Taust: standardne sissejuhatus

² Taust: naine

³ Taust: 18-30 aastat

⁴ Taust: põhiharidus või madalam

⁵ Taust: eesti keel

Lisa 5. Lineaarne regressioon: joonvastamine 3. ja 5. lainel

	Laine 3	Laine 5	Laine 5	Laine 5
	1. kmpl	1.kmpl	2.kmpl	3.kmpl
	B SE	B SE	B SE	B SE
(Vabaliige)	1,021 0,174	*** 1,527 0,199	*** 12,585 1,524	*** 0,762 0,125
w1w3w5exp0: vastamise täpsuse rõhutamine ¹	-0,059 0,072	0,090 0,083	0,149 0,640	0,028 0,053
w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine ¹	-0,003 0,071	0,030 0,083	0,105 0,633	0,066 0,052
Rööprähklus	0,043 0,071	-0,101 0,086	0,188 0,657	0,006 0,054
Sage interneti kasutus	-0,082 0,084	0,034 0,094	-1,493 0,719	* 0,059
Politiikahuvi	0,080 0,060	-0,022 0,071	-0,449 0,543	0,029 0,045
Sugu: mees ²	-0,003 0,062	-0,059 0,071	0,274 0,542	-0,060 0,045
Vanus: 31-45 ³	0,022 0,087	-0,098 0,103	-0,790 0,787	0,116 0,065
Vanus: 46-60 ³	0,083 0,093	-0,330 0,110	** -1,880	* 0,009
Vanus: >60 ³	0,056 0,107	-0,385 0,126	** 0,970	-0,015 0,080
Kutseharidus ⁴	0,059 0,127	-0,076 0,149	-2,200 1,141	-0,194 0,094
Kõrgharidus ⁴	0,108 0,128	0,037 0,147	-2,040 1,132	-0,285 0,093
Üldkeskharidus ⁴	0,089 0,132	-0,024 0,153	-2,293 1,177	-0,184 0,097
Vene keel ⁵	0,073 0,075	0,188 0,090	* 2,438 0,693	*** 0,373 0,057
Kasutatud seade: tahvelarvuti ⁶	0,129 0,148	-0,193 0,207	2,858 1,586	0,178 0,131
Kasutatud seade: nutitelefon ⁶	0,048 0,070	0,027 0,083	0,184 0,640	0,092 0,053
Vastamisele kulunud aeg	0,000	0,002	-0,010	-0,001

	<i>Laine 3</i> <i>1. kmpl</i>	<i>Laine 5</i> <i>1.kmpl</i>	<i>Laine 5</i> <i>2.kmpl</i>	<i>Laine 5</i> <i>3.kmpl</i>	
<i>Kaks või kolm meeldetuletust</i>	0,002	0,003	0,020	0,002	
	0,059	0,028	-1,090	-0,193	*
	0,093	0,149	1,144	0,094	
N	489	581	581	581	
R2	0,021	0,051	0,056	0,119	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1
¹ Taust: Standardne sissejuhatus
² Taust: naine
³ Taust: 18-30
⁴ Taust: põhiharidus või madalam
⁵ Taust: eesti keel
⁶ Taust: arvuti

Lisa 6. Lineaarne regressioon: joonvastamine 2., 4. ja 6. lainel

	<i>Laine 2</i> <i>1.kmpl</i>	<i>Laine 2</i> <i>2.kmpl</i>	<i>Laine 4</i> <i>1.kmpl</i>	<i>Laine 4</i> <i>2.kmpl</i>	<i>Laine 4</i> <i>3.kmpl</i>	<i>Laine 6</i> <i>1.kmpl</i>	<i>Laine 6</i> <i>2.kmpl</i>
	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>
	<i>SE</i>	<i>SE</i>	<i>SE</i>	<i>SE</i>	<i>SE</i>	<i>SE</i>	<i>SE</i>
<i>(Vabaliige)</i>	1,394 ***	1,053 ***	5,888 ***	1,574 ***	0,910 ***	4,054 ***	0,878 **
	0,244	0,263	0,670	0,187	0,150	0,603	0,327
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed¹</i>	0,150	-0,024	0,102	0,015	-0,034	-0,254	0,017
	0,098	0,105	0,264	0,074	0,059	0,236	0,128
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed¹</i>	-0,006	0,014	0,097	0,131	-0,025	-0,069	0,140
	0,100	0,106	0,264	0,074	0,059	0,237	0,129
<i>Rööprähklus</i>	-0,030	0,086	-0,126	-0,127	0,058	0,001	0,069
	0,096	0,103	0,263	0,073	0,059	0,239	0,128
<i>Sage interneti kasutus</i>	-0,109	-0,115	-0,221	-0,126	-0,137 *	0,041	0,018
	0,111	0,118	0,300	0,084	0,067	0,265	0,146
<i>Poliitika huvi</i>	-0,057	-0,016	0,255	0,197 **	-0,057	-0,106	0,114
	0,084	0,089	0,225	0,063	0,050	0,200	0,108
<i>Sugu: mees²</i>	0,109	-0,192 *	-0,645 **	-0,112	-0,057	0,394 *	-0,239 *
	0,085	0,090	0,226	0,063	0,051	0,200	0,108
<i>Vanus: 31-45³</i>	0,056	-0,215	0,121	-0,223 *	0,007	0,668 *	-0,006
	0,119	0,127	0,317	0,088	0,071	0,288	0,155
<i>Vanus: 46-60³</i>	-0,090	-0,319 *	0,211	-0,400 ***	-0,061	0,572	-0,137
	0,125	0,134	0,342	0,096	0,077	0,307	0,165
<i>Vanus: >60³</i>	-0,079	-0,367 *	-0,117	-0,568 ***	0,022	0,490	-0,347

	Laine 2 1.kmpl	Laine 2 2.kmpl	Laine 4 1.kmpl	Laine 4 2.kmpl	Laine 4 3.kmpl	Laine 6 1.kmpl	Laine 6 2.kmpl
Kutseharidus ⁴	0,143	0,154	0,387	0,108	0,087	0,351	0,193
	-0,055	0,054	-0,905	0,102	-0,064	-1,154 *	0,451
Kõrgharidus ⁴	0,184	0,198	0,496	0,138	0,111	0,458	0,254
	-0,194	0,320	-1,270 *	0,124	-0,096	-1,019 *	0,614 *
Üldkeskharidus ⁴	0,183	0,196	0,491	0,137	0,110	0,453	0,251
	-0,226	0,184	-1,003 *	0,021	-0,124	-1,415 **	0,508
Vene keel ⁵	0,189	0,203	0,510	0,142	0,114	0,472	0,259
	0,208 *	0,268 *	1,206 ***	-0,033	0,007	-0,359	0,107
Kasutatud seade: tahvelarvuti ⁶	0,103	0,113	0,284	0,079	0,064	0,254	0,138
	-0,261	-0,435	0,390	0,024	-0,043	0,773	0,011
Kasutatud seade: nutitelefon ⁶	0,235	0,244	0,612	0,171	0,137	0,499	0,275
	0,078	-0,264 *	0,333	-0,127	-0,039	-0,377	-0,020
Vastamisele kulunud aeg	0,103	0,112	0,267	0,074	0,060	0,249	0,135
	-0,003	0,009 **	-0,006	0,003	0,001	0,000	0,001
Kaks või kolm meeldetuletust	0,003	0,003	0,012	0,003	0,003	0,008	0,004
	0,165	-0,310 *	-0,399	0,027	-0,055	0,040	-0,008
	0,128	0,133	0,325	0,091	0,073	0,211	0,114
N	632	594	562	562	562	567	534
R2	0,040	0,067	0,072	0,079	0,027	0,048	0,039

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

¹ Taust: informatsiooni tagajärgede osas ei kuvatud

² Taust: naine

³ Taust: 18-30

⁴ Taust: põhiharidus või madalam

⁵ Taust: eesti keel

⁶ Taust: arvuti

Lisa 7. Lineaarne regressioon: vastamiseks kulunud aeg

	Laine 3 B SE	Laine 5 B SE	Laine 2 B SE	Laine 4 B SE	Laine 6 B SE
(Vabaliige)	3,325 ***	3,064 ***	3,179 ***	2,725 ***	3,154 ***
	0,109	0,091	0,086	0,101	0,095
w1w3w5exp0: vastamise täpsuse rõhutamine ¹	0,055	0,030			
	0,048	0,040			
w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine ¹	0,104 *	0,175 ***			
	0,048	0,039			

	Laine 3	Laine 5	Laine 2	Laine 4	Laine 6		
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed²</i>			0,056	0,056	0,080	*	
			0,037	0,042	0,040		
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed²</i>			0,109	**	0,090	*	
			0,037		0,042	0,039	
<i>Rööprähklus</i>	0,075	0,141	***	0,049	0,069	0,152	***
	0,047	0,041		0,036	0,042	0,040	
<i>Sage interneti kasutus</i>	-0,102	-0,141	**	-0,163	***	-0,178	***
	0,056	0,044		0,041		0,047	0,044
<i>Sugu: mees³</i>	-0,052	0,021		0,039	-0,011	-0,037	
	0,041	0,033		0,032	0,036	0,033	
<i>Vanus: 31-45⁴</i>	0,160	**	0,219	***	0,155	***	**
	0,058		0,049		0,045	0,050	0,048
<i>Vanus: 46-60⁴</i>	0,267	***	0,312	***	0,273	***	0,308
	0,061		0,051		0,046	0,053	0,050
<i>Vanus: >60⁴</i>	0,530	***	0,559	***	0,458	***	0,504
	0,068		0,056		0,051	0,058	0,055
<i>Kutseharidus⁵</i>	-0,078	0,020		0,023	-0,005	-0,057	
	0,086	0,071		0,070	0,079	0,077	
<i>Kõrgharidus⁵</i>	-0,224	**	-0,121	-0,106	-0,123	-0,166	*
	0,085		0,070	0,069	0,078	0,076	
<i>Üldkeskharidus⁵</i>	-0,087	-0,006		-0,024	-0,028	-0,100	
	0,089	0,074		0,071	0,081	0,079	
<i>Vene keel⁶</i>	-0,044	0,026		0,012	-0,026	0,044	
	0,050	0,043		0,039	0,045	0,043	
<i>Kaks või kolm meeldetuletust</i>	-0,067	0,012		-0,059	0,075	0,008	
	0,062	0,072		0,048	0,052	0,036	
<i>Kasutatud seade: tahvelarvuti⁷</i>	-0,060	0,131		0,036	-0,135	0,077	
	0,100	0,099		0,089	0,098	0,084	
<i>Kasutatud seade: nutitelefon⁷</i>	-0,044	-0,056		0,026	0,021	0,016	
	0,047	0,040		0,039	0,043	0,042	
N	489	581		632	562	569	
R2	0,220	0,282		0,242	0,241	0,261	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

¹ Taust: standardne sissejuhatus

² Taust: informatsiooni tagajärgede osas ei kuvatud

³ Taust: naine

⁴ Taust: 18-30

⁵ Taust: põhiharidus või madalam

⁶ Taust: eesti keel

⁷ Taust: arvuti

Lisa 8. Binaarne logistiline regressioon: küsimusest arusaamise ja vastamise keerukus

	Laine 1 AME SE	Laine 3 AME SE	Laine 2 AME SE	Laine 4 AME SE	Laine 6 AME SE	
<i>w1wexp0: vastamise täpsuse rõhutamine</i> ¹	0,012	0,084 *				
	0,045	0,043				
<i>w1w3exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine</i> ¹	0,027	0,013				
	0,044	0,045				
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed</i> ²			-0,010	0,034	0,113 *	
			0,048	0,052	0,051	
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed</i> ²			0,049	0,092 .	0,090 .	
			0,048	0,051	0,051	
<i>Rööprähklus</i>	0,034	-0,001	-0,018	-0,022	-0,004	
	0,044	0,043	0,047	0,051	0,052	
<i>Sage interneti kasutus</i>	-0,101 *	-0,073	-0,105 .	-0,010	-0,100 .	
	0,050	0,045	0,054	0,057	0,055	
<i>Poliitika huvi</i>	-0,140 ***	-0,118 **	-0,068 .	-0,059	-0,078 .	
	0,039	0,036	0,041	0,044	0,044	
<i>Sugu: mees</i> ³	-0,064 .	-0,019	0,014	-0,061	-0,051	
	0,038	0,037	0,041	0,044	0,043	
<i>Vanus: 31-45</i> ⁴	0,043	0,026	-0,001	0,012	-0,036	
	0,053	0,052	0,057	0,062	0,062	
<i>Vanus: 46-60</i> ⁴	-0,025	-0,077	-0,005	0,027	-0,086	
	0,055	0,058	0,059	0,065	0,064	
<i>Vanus: >60</i> ⁴	0,038	0,046	0,067	0,078	-0,065	
	0,063	0,058	0,067	0,070	0,072	
<i>Kutseharidus</i> ⁵	-0,012	-0,114 *	0,060	-0,024	-0,028	
	0,079	0,056	0,085	0,094	0,098	
<i>Kõrgharidus</i> ⁵	-0,196 *	-0,180 **	0,092	-0,082	-0,022	
	0,078	0,055	0,083	0,091	0,096	
<i>Üldkeskharidus</i> ⁵	-0,108	-0,134 *	0,104	-0,006	-0,009	
	0,081	0,060	0,087	0,095	0,101	
<i>Vene keel</i> ⁶	-0,083 .	-0,091 .	-0,075	-0,054	-0,050	
	0,048	0,050	0,049	0,056	0,055	
<i>Kaks või kolm meeldetuletust</i>	0,053	-0,073	-0,038	0,032	0,086 .	
	0,052	0,061	0,062	0,062	0,046	
N	704	489	632	561	568	
AIC	945,890	485,230	868,550	777,710	795,470	
Pseudo R ² : McFadden	0,062	0,248	0,020	0,021	0,027	

*** p < 0.001; ** p < 0.01; * p < 0.05; . p < 0.1

¹ Taust: standardne sissejuhatus ² Taust: informatsiooni tagajärgede osas ei kuvatud ³ Taust: naine ⁴ Taust: 18-30 ⁵ Taust: põhiharidus või madalam ⁶ Taust: eesti keel	<i>Laine 1</i>	<i>Laine 3</i>	<i>Laine 2</i>	<i>Laine 4</i>	<i>Laine 6</i>
--	----------------	----------------	----------------	----------------	----------------

Lisa 9. Binaarne logistiline regressioon: hinnang pingutusele täpsete vastuste andmisel

	<i>Laine 1</i>		<i>Laine 3</i>		<i>Laine 5</i>		<i>Laine 2</i>		<i>Laine 4</i>		<i>Laine 6</i>
	AME		AME		AME		AME		AME		AME
	SE		SE		SE		SE		SE		SE
<i>w1w3w5exp0: vastamise täpsuse rõhutamine¹</i>	0,034		0,005		0,027						
	0,032		0,037		0,043						
<i>w1w3w5exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine¹</i>	0,112	***	0,015		0,076	.					
	0,033		0,037		0,043						
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed²</i>							0,021		0,070	*	0,071
							0,033		0,034		0,043
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed²</i>							0,102	**	0,087	*	0,090
							0,034		0,034		0,043
<i>Rööprähklus</i>	0,003		-0,066	*	-0,113	**	-0,086	**	-0,058	.	-0,071
	0,032		0,032		0,039		0,029		0,030		0,041
<i>Sage interneti kasutus</i>	-0,012		0,053		-0,023		0,041		-0,005		-0,045
	0,044		0,038		0,051		0,042		0,046		0,051
<i>Poliitikahuvi</i>	-0,030		0,004		0,044		0,006		0,027		-0,010
	0,029		0,031		0,036		0,029		0,030		0,037
<i>Sugu: mees³</i>	-0,019		0,024		-0,033		-0,012		-0,065	*	-0,008
	0,028		0,032		0,037		0,029		0,029		0,037
<i>Vanus: 31-45⁴</i>	-0,105	*	0,021		-0,079		-0,061		-0,175	**	-0,011
	0,048		0,047		0,059		0,051		0,055		0,059
<i>Vanus: 46-60⁴</i>	-0,284	***	-0,047		-0,251	***	-0,241	***	-0,331	***	-0,183
	0,044		0,046		0,057		0,046		0,050		0,057
<i>Vanus: >60⁴</i>	-0,262	***	-0,006		-0,297	***	-0,267	***	-0,331	***	-0,257
	0,050		0,054		0,060		0,048		0,053		0,058
<i>Kutseharidus⁵</i>	-0,042		-0,126		-0,090		-0,049		-0,066		-0,111
	0,057		0,081		0,082		0,068		0,075		0,089
<i>Kõrgharidus⁵</i>	-0,042		-0,159	*	-0,179	*	-0,103		-0,139	.	-0,135

	<i>Laine 1</i>		<i>Laine 3</i>		<i>Laine 5</i>		<i>Laine 2</i>		<i>Laine 4</i>		<i>Laine 6</i>	
<i>Üldkeskharidus</i> ⁵	0,056		0,079		0,080		0,065		0,071		0,087	
	-0,080		-0,137	.	-0,103		-0,092		-0,081		-0,108	
<i>Vene keel</i> ⁶	0,056		0,081		0,084		0,067		0,074		0,090	
	-0,218	***	0,329	***	-0,275	***	-0,176	***	-0,171	***	-0,270	***
	0,018		0,054		0,033		0,023		0,024		0,031	
<i>Kaks või kolm meeldetuletust</i>	0,047		-0,014		-0,177	**	0,042		-0,072	*	-0,023	
	0,042		0,044		0,060		0,046		0,034		0,039	
N	704		489		581		632		562		569	
AIC	584,200		454,590		641,900		500,790		436,600		618,440	
Pseudo R ² : McFadden	0,191		0,325		0,132		0,199		0,204		0,122	

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1
¹ Taust: standardne sissejuhatus
² Taust: informatsiooni tagajärgede osas ei kuvatud
³ Taust: naine
⁴ Taust: 18-30
⁵ Taust: põhiharidus või madalam
⁶ Taust: eesti keel

Lisa 10. Binaarne logistiline regressioon: hinnang vastamise nautimisele

	<i>Laine 1</i>		<i>Laine 3</i>		<i>Laine 2</i>		<i>Laine 4</i>		<i>Laine 6</i>	
	AME		AME		AME		AME		AME	
	SE		SE		SE		SE		SE	
<i>w1w3exp0: vastamise täpsuse rõhutamine</i> ¹	0,042		0,005							
	0,036		0,037							
<i>w1w3exp0: vastamise täpsuse rõhutamine ja lubaduse küsimine</i> ¹	0,107	**	0,015							
	0,037		0,037							
<i>w2w4w6exp0: hoolika vastamise positiivsed tagajärjed</i> ²					0,064	.	0,032		0,015	
					0,038		0,039		0,038	
<i>w2w4w6exp0: hooletu vastamise negatiivsed tagajärjed</i> ²					0,041		0,060		0,060	
					0,038		0,041		0,040	
<i>Rööprätklus</i>	-0,008		-0,066	*	-0,025		0,059		0,027	
	0,035		0,032		0,036		0,043		0,041	
<i>Sage interneti kasutus</i>	0,057		0,053		0,076	.	0,035	***	0,073	**
	0,039		0,038		0,039		0,043		0,039	
<i>Poliitikahuvi</i>	-0,022		0,004		0,007		0,005	***	-0,009	***
	0,032		0,031		0,033		0,035		0,033	
<i>Sugu: mees</i> ³	-0,032		0,024		-0,072	*	-0,111		-0,086	
	0,031		0,032		0,032		0,033		0,032	

	<i>Laine 1</i>		<i>Laine 3</i>		<i>Laine 2</i>		<i>Laine 4</i>		<i>Laine 6</i>
<i>Vanus: 31-45⁴</i>	-0,026		0,021		0,007		0,005		-0,007
	0,044		0,047		0,047		0,049		0,046
<i>Vanus: 46-60⁴</i>	-0,045		-0,047		-0,028		0,006		-0,029
	0,046		0,046		0,048		0,052		0,048
<i>Vanus: >60⁴</i>	-0,042		-0,006		-0,064		-0,022	***	0,032
	0,054		0,054		0,053		0,056		0,057
<i>Kutseharidus⁵</i>	-0,155	*	-0,126		-0,237	**	-0,361	***	-0,300
	0,073		0,081		0,084		0,089		0,091
<i>Kõrgharidus⁵</i>	-0,159	*	-0,159	*	-0,308	***	-0,417	***	-0,354
	0,072		0,079		0,082		0,086		0,089
<i>Üldkeskharidus⁵</i>	-0,127	.	-0,137	.	-0,291	***	-0,366	***	-0,258
	0,074		0,081		0,084		0,090		0,093
<i>Vene keel⁶</i>	0,295	***	0,329	***	0,356	***	0,311		0,365
	0,047		0,054		0,048		0,052		0,052
<i>Kaks või kolm meeldetuletust</i>	-0,029		-0,014		-0,084	.	-0,026		-0,065
	0,041		0,044		0,044		0,048		0,033
N	704		489		630		562		568
AIC	719,170		379,400		622,610		555,140		535,810
Pseudo R ² : McFadden	0,084		0,315		0,134		0,126		0,144

*** p < 0,001; ** p < 0,01; * p < 0,05; . p < 0,1

¹ Taust: standardne sissejuhatus

² Taust: informatsiooni tagajärgede osas ei kuvatud

³ Taust: naine

⁴ Taust: 18-30

⁵ Taust: põhiharidus või madalam

⁶ Taust: eesti keel

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Liisi Lillipuu,

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

Andmete kvaliteet veebipaneelis tõenäosusliku valimiga veebipaneeli CRONOS näitel,

mille juhendaja on Indrek Soidla,

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.

Liisi Lillipuu
24.05.2019