

Tartu Ülikool  
sotsiaalteaduste valdkond  
psühholoogia instituut

Kevin Mändmets  
ERINEVUSED VEEBI- JA LABORIKESKKONNAS LÄBI VIIDUD  
EKSPERIMENDI VAHEL  
Uurimistöö

Juhendaja: Annegrete Palu

Jooksev pealkiri: KATSEKESKKONDADE TULEMUSTE ERINEVUS

Tartu 2021

**Erinevused veebi- ja laborikeskkonnas läbi viidud eksperimendi vahel****Kokkuvõte**

Käesoleva uuringu eesmärgiks oli võrrelda erinevusi labori- ning veebieksperimendi tulemuste vahel, viies läbi näotuvastuse eksperimendi nii laboris (n = 665) kui ka veebis (n = 1073). Veebiuuringud viidi läbi, kasutades MTurk keskkonda. Mõlemas uuringus pidid katseisikud vaatama kolme kuni viit videot, milles oli kujutatud vargust. Peale igat videot pidid nad äratundmisreast identifitseerima videos nähtud naised. Tulemused näitasid, et võrreldes laboritingimusega, oli äratundmistäpsus veebis 15% võrra madalam. Lisaks leiti, et kuigi äratundmistäpsus erines katsekeskkondade vahel, oli nii veebi- kui laborikeskkonnas ainsaks äratundmistäpsust statistiliselt oluliselt mõjutavaks muutujaks kurjategijate arv videos. Kindlushinnang enda otsusele oli veebis keskmiselt 14.45% võrra madalam. Antud tulemuste erinevust võis põhjustada katseisikute tähelepanematus katse ajal. Nendest tulemustest hoolimata oli veebis osalenute hinnang enda keskendumisele kõrgem, kuigi teised tulemused seda ei kinnita. Seega võib järeldada, et tähelepanu nõudvate ülesannete puhul on veebitingimuse puhul tulemused võrreldes laboritingimusega halvemad.

*Märksõnad:* Veebieksperiment, laborieksperiment, näotuvastus, kurjategijate tuvastamine

**Differences between the results of an experiment conducted  
on the web and in the lab**

**Abstract**

This paper aimed to compare the results of web and lab experiments by conducting an eyewitness identification experiment both in the lab ( $n = 665$ ) and on the web ( $n = 1073$ ). The web experiment was conducted using the MTurk platform. Participants in both experiments had to watch videos, in which a theft took place. After each video, the participants had to identify the women seen in the video from a lineup. The results showed that compared to the lab experiment, the identification accuracy on the web was 15% lower. Additionally, it was found that although the identification accuracy differed in the lab and on the web, the only factor that had a statistically significant effect on the identification accuracy was the number of women seen in the video. The self-reported confidence in their identification decisions that participants made was on average 14.45 points lower on the web. This might have been caused by participant inattention on the web. Despite this, the self-reported concentration during the experiment was higher on the web. This means that web experiments yield lower results, as far as experiments that require attention go.

*Keywords:* Web experiment, lab experiment, facial recognition, eyewitness identification

## Sissejuhatus

Eksperimentide läbiviimine veebis on tehnoloogia arenedes muutunud aina populaarsemaks. Veebikatsete loomist peetakse laboritestimise valdkonnas revolutsiooniliseks, sest see pakub eksperimendi läbiviimisel võimalusi, mida enne ei eksisteerinud (Musch ja Reips, 2000), võimaldades näiteks eksperimentaatoril automatiseerida eksperimendi protsess (Reips, 2002). Veebiuuringute nii suhtelise kui ka absoluutse hulga suurenemisel on aga vajalik uurida, kas nendest saadud andmed on usaldusväärsed. Käesoleva uuringu eesmärk on võrrelda veebi- ning laboritingimustes saadud andmeid, kasutades sama eksperimenti.

Veebis eksperimentide läbiviimise puhul tõi Reips (2002) enda töös välja 18 positiivset aspekti, sealhulgas eksperimentaatori kallutatuse vähenemise. Kuigi ka veebiuuringutes esineb eksperimentaatori kallutus, mis on tingitud peamiselt uurija nimest (White, Strezhnev, Lucas, Kruszewska ja Huff, 2018), siis väheneb eksperimentaatori kallutatust põhjustavate võimalike faktorite arv, sest katseisik ei puutu eksperimentaatoriga otseselt kokku. Samuti tunnevad katseisikud ennast mugavamalt tuttavas keskkonnas katset sooritades (Reips, 2002).

Nagu eelpool mainitud, on eksperimentaatoril veebikeskkonna eelistamiseks mitmeid põhjused. Musch ja Reips (2000) leidsid enda töös läbi viidud küsimustikus, et peamisteks motiivideks, miks eksperimentaatorid otsustavad veebikatse kasuks, on kõrge katseisikute hulk, kõrge statistiline võim, katse läbiviimiseks kuluv vähene aeg ning võimalus kaasata katseisikuid teistest riikidest. See tähendab, et eksperimentaatoritel on nüüd võimalik hõlpsasti koguda andmeid rahvusvaheliselt valimilt, mis muudab järeldused üldistatavamaks. Ühe suurima eelisena veebieksperimentide puhul tõi Musch ja Reips (2002) oma töös välja võimaluse koguda vähese ajaga suur kogus andmeid. Näiteks Pauszek, Sztybel ja Gibson (2017) kogusid 30 katseisiku andmed vaid 18 tunniga, mis oleks nende sõnul laboritingimustes võtnud aega nädalaid ning ka Birnbaum (2000) on öelnud, et tema jaoks oli meeldiv koguda ligi 2000 katseisiku andmeid vaid viie kuuga. Ka käesolevas uurimistöös kasutatava veebieksperimenti puhul koguti 1800 katseisiku andmed umbes nädala ajaga.

Veel üheks positiivseks omaduseks on veebiuuringute puhul see, et katsed on avatud ööpäevaringselt ning katseisik saab seda läbida endale sobival ajal endale ja sobivas keskkonnas. Lisaks leidsid Moss, Rosenzweig, Robinson ja Litman (2020), et näiteks *Amazon Mechanical Turk* (MTurk) keskkonnas eksperimentides osalevad inimeste majanduslik seisund esindab ka üldise populatsiooni majanduslikku seisut. Seega võiks arvata,

et selles või ka teistes sarnastes keskkondades leiduvad katseisikud esindavad ka üldist populatsiooni, muutes seeläbi andmed valiidsmaks.

Katseisikute seisukohast näivad veebiuuringud olevat paremad, kuigi võiks esialgu teisiti arvata. Levinud on uskumused ebainimlikult madala tasu ning katseisikute kuritarvitamise kohta. Need müüdid on ümber lükatud ning vastupidiselt on leitud, et inimesed võivad veebiuuringute täitmist eelistada isegi traditsioonilisele tööle, tuues ühe suurima plussina välja paindlikkuse, mida sellises keskkonnas töötamine pakub (Moss, Rosenzweig, Robinson ja Litman, 2020). Sealjuures on huvitav märkida, et need inimesed oleksid nõus eksperimentide tegemise välja vahetama vaid kordades suurema tasuga töö vastu (Moss, Rosenzweig, Robinson, Litman, 2020). Võib järeldada, et inimestele meeldib veebikeskkondades raha teenimise võimalus niivõrd, et on nõus selle nimel teistest hüvedest lahti ütleva. Seega võib tunduda, et veebieksperimentide suhtelise koguse suurenemine on ainult positiivne nii katse läbiviijale kui ka selles osalejale. Samas on oluline mõista nii protseduuri kui ka tulemuste erinevust kahe katsekeskkonna vahel, et andmeid siiski õigesti interpreteerida.

Veebis eksperimentide läbiviimisel on ka negatiivseid külgi. Esiteks on valim siiski kallutatud hoolimata asjaolust, et katset on võimalus sooritada inimestel mitmetes riikides ning kultuuriruumides, jättes valimist välja näiteks inimesed, kellel ei ole juurdepääsu internetile (Bethlehem, 2010). Teiseks on katsekeskkond varieeruv ning katse tulemusi võivad mõjutada mitmed sekkuvad muutujad, mida laboritingimustes püütakse kontrolli all hoida (Pauszek, Sztybel ja Gibson, 2017). See tähendab, et eksperimentaatoril on veebieksperimenti korral keerulisem garanteerida, et katseisik näeb stiimulit selliselt, nagu eksperimentaator seda plaanis näidata või kas ta seda üldse näeb (Krantz, 2000). Sekkuvateks muutujateks võivad olla katsevahendid - näiteks erineva kvaliteediga kõrvaklapid või erineva suurusega kuvarid, mille läbi inimesele stiimuleid esitatakse - või segavad stiimulid katse sooritamisel (Pauszek, Sztybel ja Gibson, 2017). Kolmandaks puuduseks on asjaolu, et katseisikud kasutavad õigust katset igal hetkel pooleli jätta rohkem veebi- kui laboritingimustes läbiviidud katse puhul (Bamert, 2002, viidatud Reips, 2002 kaudu). Seda võib vaadelda nii positiivsena - katset on sooritanud ainult motiveeritud isikud ning seega võivad katsetulemused olla usaldusväärsemad - kui ka negatiivsena - valim muutub selle võrra vähem juhuslikuks. Samuti võib oluliseks erinevuseks kujuneda katseisiku võimetus eksperimentaatori käest segaduse korral küsimusi küsida. Ainsad juhised, mida katseisik kasutada saab, on need, mida eksperimentaator katsesse kirjalikult pani. See võib viia

ebakorrektelt sooritatud katseni, kuid katse soorituse korrektsust ei pruugi olla võimalik katse tulemustest hinnata.

Üks populaarsemaid platvorme veebiuuringute läbiviimiseks on Amazon MTurk, mida ka antud uurimuse käigus kasutati katse veebiversiooni läbiviimiseks. MTurk on veebiplatvorm, mis toob kokku töö pakkujad (*requesters*) ning töötajad (*workers*). Platvorm on kiiresti kasutusele võetud ka teadustöö tegemiseks, kuna võimaldab uurijatel ammutada suurt hulka andmeid (Sheeran, 2018) ilma liigselt ressursse kasutamata. Näiteks tekitasid Buhrmester, Kwang, ja Gosling (2016) enda uurimistöös küsimustiku, milles otsisid 25 katseisikut ning pakkusid 30 minutit kestva küsimustiku eest \$0.02. Soovitud andmete hulk saavutati umbes 5 tunniga. See-eest jõuti sama suure andmete hulga juurde vaid kahe tunniga, kui küsimustiku läbimise eest pakuti \$0.50. Seega võib järeldada, et töötajad on rohkem motiveeritud eksperimentides osalema, kui selle eest on suuremat tasu oodata. Oma intervjuus ühe inimesega, kes veedab keskmiselt 30 tundi nädalas veebis MTurk küsimustike täites, sai Semules (2018) teada, et sellise tegevusega teenib tema intervjuueeritav umbes \$1 tunnis, kuid Moss, Rosenzweig, Robinson ja Litman (2020) leidsid enda uurimuses, et inimestel on võimalik teenida kuni \$9 tunnis. Katse läbiviimiseks vajaminevate kulude vähendamiseks kasutavad tööpakkujad vahel alatuid võtteid. Näiteks võivad eksperimentaatorid katse või küsimustiku läbimiseks kuluvat eeldatavat aega märkida vähemaks, kui tegelikult on. See toob kaasa töötajate jaoks oodatust väiksema tasu ühes ajaühikus.

Kogutud andmete kvaliteedi tagamiseks pakub MTurk mitmeid võimalusi katsetulemuste usaldusväärsemaks muutmiseks. Näiteks on eksperimentaatoril võimalus keelduda sellise katseisiku andmetest, kes kukkus liiga palju tähelepanu kontrolli küsimustest läbi ning see tähendab ka, et katseisikule ei maksta tehtud töö eest tasu (Sheeran, 2018). Sellega väheneb tõenäosus, et inimesed täidavad küsimustikke ning sooritavad katseid ainult tasu saamise eesmärgi. Sellegipoolest tuleb arvesse võtta, et MTurk töötajatele makstakse sisuliselt tunnitasa ning seega on nad motiveeritud katseid ning eksperimente võimalikult kiiresti läbima (Pauszek, Sztybel ja Gibson, 2017). See asjaolu võib mõjutada andmete kvaliteeti, sest vastuseid ei pruugita enne nende esitamist piisavalt palju läbi mõelda.

Tänaseks on juba läbi viidud mõningaid uuringuid, mis on võrrelnud veebi- ja laborikeskkonna tulemuste erinevusi. Sarnaselt käesolevale tööle võrdlesid Pauszek, Sztybel ning Gibson (2017) andmete erinevusi laboritingimustes ning veebikeskkonnas läbitud eksperimendi tähelepanukatse puhul, kus kiirus mängis olulist rolli. Nad leidsid, et veebis läbi viidud katse puhul olid inimeste reaktsiooniajad märgatavalt kiiremad, mis võib nende

sõnul tuleneda töötajate motivatsioonist ülesandeid võimalikult kiiresti täita. Teiseks võimalikuks põhjuseks toovad autorid välja asjaolu, et MTurk keskkonnas katset teinud inimesed pidid saavutama väiksema veamäära kui 20%, et tehtud töö eest tasu saada. Laboritingimustes aga sellist nõuet ei olnud.

Lisaks on keskkondade vahelisi tulemusi võrrelnud näiteks Clifford ja Jerit (2014). Nende töö sarnanes ülesehituselt käesoleva uurimisele selle poolest, et mõlemad katse versioonid olid võimalikult sarnased üksteisele, kuid käesoleva uurimuse eksperiment ning Clifford ja Jerit (2014) poolt läbi viidud uurimus erinevad katse sisu poolest. Clifford ja Jerit (2014) töös kasutatud eksperiment kontrollis inimeste poliitilisi teadmisi ning nad arvasid, et veebis katset sooritavad inimesed kasutavad väliseid materjale ning allikaid, et küsimustele vastuseid leida. Nad leidsid, et esines oodatust vähem erinevusi veebi- ning labori keskkonnas sooritatud katsete vahel ning järeldasid, et veebikatsed võivad olla, arvestades sellega kaasnevaid eeliseid ning väheseid negatiivseid külgi, sobiv alternatiiv laborikatsetele.

Ka õiguspsühholoogia valdkonnas on veebikeskkonnas eksperimente korraldatud. Näiteks Levett ning Kovera (2007) viisid veebis läbi õiguspsühholoogia alase eksperimendi, milles püüdsid uurida, kuidas mõjutab vastanduv ekspertarvamus vandekogu otsuseid ning üldist vaadet ekspertarvamusele. Lowrey, Maquire ja Bennett (2016) koostasid hiljuti uurimuse, mille käigus viisid veebis läbi eksperimendi, millega uuriti inimeste üldist suhtumist politseisse ning suhtluspüüdlikkuse (*accommodation*) mõju olukorra hindamisele. Katses paluti inimestel vaadata lühikest videot ning hiljem selle video ja enda kohta mõningatele küsimustele vastata. Enne aga, kui veebis läbi viidud õiguspsühholoogia alaste uurimuste tulemusi praktikasse rakendada, peab siiski mõistma, kas sellised läbi viidud katsed või eksperimendid võimaldavad valiidsid andmeid ning kuidas erinevad need laboritingimustes läbi viidud katsetest.

### **Käesolev uurimistöö**

Käesoleva uurimuse eesmärgiks on välja selgitada, kas ning kuidas erinevad katsetulemused, kui katset on läbi viidud labori- või veebikeskkonnas. Uurimuses kasutatud katse toetub väga tugevalt inimese tähelepanule. Tähelepanu ning mälu on omavahel tihedalt seotud (Oberauer, 2019) ning seega on mälu mõjutatav kõikide võimalike sekkuvate muutujate poolt, mis võivad inimese tähelepanu hajutada või tähelepanuvõimet vähendada. Antud uuringus kasutatavas katses olid laboritingimustes tähelepanu hajutavad faktorid võimalikult palju kontrolli all hoitud. See tähendab, et katset viidi läbi vaikes eraldatud ruumis, mida keegi katse ajal ei külastanud, ruumi ei pääsenud kõrvalist valgust ning

eksperimentaator sai katse ajal katseisiku kõrval viibides veenduda, et ta ei tegeleks katse ajal kõrvaliste tegevustega. Veebikatses neid ning teisi võimalikke tähelepanu hajutavaid faktoreid kontrollida ei olnud võimalik. Eelnevad uuringud on aga näidanud, et veebikatse tulemused ei erine olulisel määral laborikatse tulemustest. Seega on käesoleva töö hüpoteesiks, et inimeste äratundmisvõime on veebikeskkonnas sooritatud katse puhul sarnane laborikeskkonnas sooritatud katsega. Lisaks uuritakse, kuidas erinevad uuringute peamised tulemused laborikatses võrreldes veebiuuringuga.

## Meetod

Antud uurimistöö kasutab laboritingimuses läbi viidud eksperimendi andmeid uuringust “Silmailiigutused kui markerid eristamiseks varem nähtud ja mittenähtud nägusid kurjategijate äratundmisel” ning veebieksperimendi andmeid samanimelisest uuringust. Käesolevas uurimistöös on sõltumatuks muutujaks katsekeskkond ning sõltuvateks muutujateks on äratundmistäpsus, kindlushinnang ning hinnang keskendumisele katse ajal.

## Valim

### *Laboriuuring*

Eksperimendi laboriuuringus osales kokku 667 katseisikut. Lõppvalimist jäi välja kaks katseisikut, kelle puhul oli esimeses kolmes näidatud videos keegi näitlejatest tuttav. Laboriuuringu lõppvalimisse jäi 665 katseisikut, kellest 416 (62.6%) olid naised. Kõige noorem osaleja oli 18-aastane ning kõige vanem 54-aastane ning katseisikute keskmine vanus oli 26.15 ( $SD = 7.59$ ). Uurimistöö autori panuseks oli laboriuuringu eksperiment läbi viia 30 katseisikuga.

### *Veebiuuring*

Eksperimendi veebiuuringu valimi moodustavad 1814 katseisikut. Algvalimi keskmine vanus oli 37.27 ( $SD = 12.51$ ) ning 55% sellest moodustasid mehed, 44.4% naised ning 0.6% märkis enda sooks “muu”. Valimist jäid välja inimesed, kes raporteerisid tehniliste tõrgetekogemist katse ajal ( $n = 1$ ), kes vastasid tähelepanukontrolli küsimuse valesti ( $n = 11$ ), kes märkisid oma vanuseks 0 ( $n = 1$ ), kes jätsid uuringu pooleli ( $n = 473$ ) ning kes olid uuringus juba varasemalt osalenud ( $n = 5$ ). Kõik veebiuuringus osalenud ja katse lõpetanud said katses osalemise kompensatsiooniks \$1.



Kuna veebiuuring viidi läbi USA valimi peal, tuli veebiuuringu valimit kitsendada, et see oleks sarnasem laboriuuringu valimile. Lõppvalimisse kaasati valge etnilise päritoluga katseisikud, kelle sugu oli märgitud olema mees või naine ning kelle vanus oli vahemikus 18-54 aastat. Seega jäi veebiuuringu lõppvalimisse 1073 katseisikut, kellest 476 (44.4%) olid naised. Kõige noorem osaleja oli 18-aastane ning kõige vanem 54-aastane. Veebiuuringus osalenud katseisikute keskmine vanus oli 34.85 ( $SD = 8.66$ ).

### ***Uurimistöö***

Lõplik valim koosneb nii labori- kui ka veebiuuringu valimitest, kus oli kokku 1738 katseisikut, kellest 1013 (58.3%) olid naised. Lõpliku valimi keskmine vanus oli 31.52 ( $SD = 9.29$ ).

### **Katsematerjal ja -aparatuur**

#### ***Aparatuur***

**Laboriuuring.** Laboriuuringu läbiviimiseks kasutati arvutit Dell Precision M6500. Katseisikule esitati katse 23 tollisel LG Flatroni ekraanil värskendussagedusega 60 Hz. Katse ajal oli arvuti helitugevus 60%. Katse viidi läbi, kasutades Tobii Studio programmi. Silmaliigutusi ja pupilli suurust mõõtis Tobii X120 Eye Tracker.

**Veebiuuring.** Veebiuuringu läbiviimiseks kasutati *Amazon Mechanical Turk* (MTurk) keskkonda ning eksperiment loodi *LabVanced* programmi kasutades, mille abil salvestati ka tulemused. Eksperimenti lubati sooritada laua- või tahvelarvutit kasutades.

#### ***Katsematerjal***

Äratundmisrea esitamiseks oli katses kaks võimalust. Samaaegse äratundmise rea korral näidati katseisikule kõiki äratundmisreas olevaid inimesi korraga. Inimeste fotosid kuvati selliselt, et need asusid ringjoonel ning iga pildi juures oli number, mis tähistas selle foto järjekorranumbrit äratundmisreas. Järjestikuse äratundmisrea korral näidati katseisikule äratundmisreas osalejate fotosid ükshaaval ekraani keskel. Äratundmisriidades kasutatud fotod kujutasid isiku büsti, nende näoilme oli neutraalne ning nähtavad seljas olevad riided olid mustad. Fotode laius oli 285 ning pikkus 315 pikslit. Eksperimendis kasutati 20 eelnevalt valminud videot keskmise pikkusega 30 sekundit, mis kujutasid ühte viiest võimalikust varastamise stsenaariumist. Videod kujutasid varastamist ühe või kahe kurjategijaga ning seda ohvri või pealtnägija vaatepunktist. Ühes videos esines maksimaalselt kolm inimest, kellest üks oli kujutatava varguse ohver. Üks viiest videost oli kontrollvideo, milles vargust selle toimumise võimalusest hoolimata ei toimunud.

**Protseduur**

Nii labori- kui veebiuuringu katseplaani oli 2 (stiimulvideote nägemise vaatepunkt: ohver või pealtnägija) x 2 (kurjategijate arv stiimulvideotes: üks või kaks kurjategijat) x 2 (äratundmisrea esitamise viis: samaaegne või järjestikune) x 2 (sihtmärgi olemasolu või puudumine) faktoriseeritud segakatseplaani.

Laboriuuring koosnes kolmest osast, kuid antud uurimistöös raames kasutati andmeid ainult selle esimesest osast. Veebiuuring oli sarnane laboriuuringu esimese osaga.

**Laboriuuring.** Enne katse algust fikseeriti katseisiku silmad ning nende liikumine ning katseisik allkirjastas informeeritud nõusoleku lehe. Enne katsega alustamist pidid nad ka enda enesetunnet, keskendumis- ja nägude äratundmisvõimet hindama. Katseisikud jagati juhuslikult ühte kaheksast katsegupist, jälgides jooksvalt katsegruppide võrdset täituvust ja soolist jaotust. Stiimulvideote nägemise järjekorda varieeriti katseisikute seas juhuslikult.

**I osa.** Katseisikule näidati viite videot. Videod esitati ükshaaval ning katseisikul paluti peale iga video vaatamist vastata selle video kohta käivatele küsimustele. Iga video vaatamise järel näidati katseisikule videos olnud naisterahva kohta kas samaaegset või järjestikust äratundmisrida ning katseisiku ülesandeks oli tuvastada videos nähtud naisterahvas või naisterahvad äratundmisreast.

Järjestikuse äratundmisrea puhul näidati katseisikule äratundmisreas olevaid fotosid ükshaaval ning kõiki vaid ühe korra. Iga foto juures pidi katseisik langetama otsuse, kas sellel kujutatud isik oli videos nähtud inimene või mitte. Enda otsusest pidi katseisik teada andma, vajutades vastavat hiireklahvi ning enda otsuse ka välja öeldes. Kui katseisik otsustas, et fotol kujutatud isik oli ka nähtud videos, paluti tal anda hinnang enda kindlusele otsuse suhtes. Kindlushinnanguks kasutati skaalat 0-100 (0 – Sa oled kindel, et tegid vale valiku, 50 – Sa lihtsalt pakkusid vastuse, 100 – Sa oled täiesti kindel, et tegid õige valiku). Kui katseisik valis äratundmisreast kellegi välja, näidati talle siiski ka kõiki ülejäänud fotosid. Fotode vaatamise järel paluti katseisikul uuesti anda enda otsusele kindlushinnang. Kui katseisik kedagi fotode hulgast välja ei valinud, oli tema otsuseks, et videos nähtud isikuid fotodel ei olnud. Ka siis paluti katseisikul anda hinnang enda kindlusele oma otsuse suhtes.

Samaaegse äratundmisrea puhul kuvati katseisikule kõik fotod samaaegselt selliselt, et need asusid ringjoonel, kusjuures iga pildi juures oli number, mis tähistas selle foto järjekorranumbrit antud äratundmisreast. Katseisik pidi langetama otsuse selle osas, kas mõni foto kujutas videos nähtud isikuid ning otsusest valjuhäälselt teada andma. Otsuse

langetamise järel paluti katseisikul hinnata enda kindlust oma otsuse suhtes.

Kahe kurjategijaga video korral nägi katseisik veel ühte äratundmisrida, mis käis teise videos nähtud naisterahva kohta. Äratundmisrea vaatamise järel küsiti katseisikult, kas talle oli keegi esitatud fotodel isiklikult tuttav. Kahe kurjategijaga video puhul küsiti ka, kumba videos nähtud isiku kohta katseisiku arvates äratundmisrida käis. Lisaks paluti katseisikul oma sõnadega kirjeldada videos toimunut. Kui katseisik mainis enda video tegevuse kirjelduse käigus eseme võtmist või vargust, küsiti tema käest, kas ta nägi, kui vargus aset leidis. Kui katseisiku kirjeldus sellele ei viidanud, siis seda küsimust ka ei esitatud.

**II osa.** Eksperimendi teine osa sarnanes ülesehituselt esimesele osale, kuid nüüd näidati katseisikule väljavõtet esimeses osas nähtud videos olevast naisterahvast ning foto järel vastavat äratundmisrida. Äratundmisread ning nende esitamise järjekord oli sama, mis esimeses osas. Sarnaselt esimesele osale pidi katseisik äratundmisreast välja valima algul näidatud inimese ning samuti hindama kindlust enda otsuses.

**III osa.** Eksperimendi kolmas osa koosnes õppimis- ja testifaasist. Õppimisfaasis näidati katseisikule 32 inimese fotot kaheksas neljases plokis. Katseisik suunati nähtud fotosid kodeerima kas pindmiselt või sügavalt. Pindmise kodeerimise käigus pidi katseisik määrama, mitmes nähtud isik antud plokis oli. Sügava kodeerimise korral pidi katseisik pakkuma, millisel erialal võiks fotol nähtav isik töötada. Vastusevariantideks oli jurist, näitleja, geenitehnoloog, hambaarst ja õpetaja. Ühes plokis kasutati ühte kodeerimise meetodit ning kodeerimismeetodid vaheldusid plokkide lõikes. Testifaasis näidati katseisikule juhuslikus järjekorras 64 fotot, millest pooled olid õppimisfaasis nähtud isikud ning pooled olid uued isikud. Katseisiku ülesandeks oli määrata, kas ta oli katse kolmandas osas seda isikud näinud või mitte. Kui tema vastuseks oli, et ta oli seda isikut juba näinud, siis pidi ta ka määrama, kas ta teadis seda või mäletas.

**Katse lõpetamine.** Eksperimendi lõppedes paluti katseisikul veel kirjalikult vastata mõnele katse kohta käivale küsimusele. Samuti informeeriti teda uuringu eesmärkidest ning sisust. Katseisikul oli võimalus küsida katse tegeliku sisu ning eesmärkide kohta küsimusi, kuid katse sisu ning tegelikku eesmärki paluti hoida salajas. Kui katseisikul oli soov saada katse kohta tagasisidet, siis võeti ka tema kontaktandmed.

**Veebiuuring.** Veebiuuringus pidid katseisikud enda informeeritud nõusolekust teada andma elektrooniliselt. Enne katsega alustamist pidid nad ka enda enesetunnet, keskendumis- ja nägude äratundmisvõimet hindama. Eksperimendi veebiuuringu puhul erines protseduur

laboriuuringust selle poolest, et seal kasutati ainult katse modifitseeritud esimest osa. Modifikatsioonid tähendavad, et katseisikule näidati ainult kolme videot. Samuti oli otsuse langetamise protseduur pisut erinev. Kui samaaegse äratundmisrea katsetingimuses olnud katseisik arvas, et kuvatud pildil olev isik oli videos nähtud inimene, pidi ta tema fotol vajutama ning siis enda kindlushinnangu andma. Kui ta aga arvas, et pildil olev isik ei olnud videos nähtud inimene, pidi ta vajutama edasi minemise nupul ning andma kindlushinnangu enda otsuse kohta, et fotodel ei olnud ühtegi videos nähtud isikut kujutatud. Kui järjestikuse äratundmisrea katsetingimuses olnud katseisik arvas, et fotol kujutatud isik oli videos nähtud inimene, pidi ta fotol vajutama ning andma siis kindlushinnangu enda otsuse kohta. Kui aga katseisik otsustas, et fotol kujutatud inimest ei olnud nähtud videos, pidi ta vajutama edasi minemise nupul ning seejärel kuvati talle järgmist fotot. Kui katseisik vaatas kõik äratundmisreas olevad fotod ära ilma kedagi välja valimata, paluti tal anda enda otsusele kindlushinnang. Katse lõpus paluti katseisikul jagada enda demograafilisi andmeid, vastata küsimusele, kas nad on kunagi näinud kuritegu pealt ja vaadanud äratundmisrida ning vastama ka tähelepanu kontrolli küsimusele. Samuti paluti katseisikutel hinnata enda võimet nägusid tuvastada ning anda hinnang videote realistlikkusele. Katseisikud pidid eksperimendi lõpus hindama ka enda keskendumise taset katse ajal. Viimaks jagati katseisikule infot uuringu eesmärgi kohta ning anti neile ka unikaalne kood, mille nad pidid MTurk keskkonda sisestama, et katses osalemise eest kompensatsiooni saada.

**Andmete valideerimine.** Veebiversioonis rakendati mitmeid meetodeid, et veenduda andmete kvaliteedis. Selleks kasutati kontrollküsimust, mis oli oma sisult väga lihtne: “Mis on palli geomeetiline kujud?” Kontrollküsimuse eesmärk oli välistada katseisikud, kes küsimustele ja vastustele mõtlemata püüdsid katset läbida. Järgmiseks meetodiks andmete kvaliteetsuse tagamiseks ning katseisikute tähelepanu turgutamiseks olid mitte-intuitiivsed juhised. Esimene juhis katseisikule palus tal jätkamiseks vajutada ekraanile all vasakul nurgas, hoolimata edasiminemisele viitavale noolele all paremal nurgas. Sellega püüti veenduda, et katset läbima hakkavad isikud pööravad juhistele piisavalt tähelepanu. Katse algul pidi katseisikud sisestama enda isikliku MTurk koodi ning katse lõppedes said katseisikud unikaalse koodi, mille pidid sisestama MTurk keskkonda. Tulemuste analüüsist jäeti välja katseisikute andmed, kelle unikaalne kood ei klappinud tema MTurk koodiga.

### **Uuringu eetiline külg**

Ekspirimendi laboriuuring on saanud nõusoleku Tartu Ülikooli inimuuringute eetika

komiteelt (nr. 302/M-28) ning veebiuuring kiideti heaks Beloit College'i (Wisconsin, USA) eetikakomiteele vastava institutsiooni poolt (*The Institutional Review Board at Beloit College*).

Laboriuuringus osalejaid informeeriti uuringu sisust nii palju kui võimalik, toomata seejuures välja eksperimendi konkreetseid hüpoteese ning eesmärke, et vältida katseisiku käitumise muutumist. Seega kasutati nii veebi- kui laboriuuringu puhul mittetäielikku pealkirja. Uuringu käigus kogutud andmeid analüüsiti anonüümsel kujul. Laboriuuringu katseisikute eristamiseks määrati igale katseisikule katse alguses kood, mis märgiti ka nõusolekulehele. Isikustamist võimaldavad andmed hävitatakse, kui antud uuring lõpeb.

Suure valimi tõttu viisid laboriuuringu katset läbi ka üliõpilased Tartu ning Tallinna ülikoolist. Eksperimendi kvaliteedi tagamiseks said need üliõpilased hakata iseseisvalt katseid läbi viima alles siis, kui olid läbi viinud harjutuskatse vastutava uurijaga. Esimese läbiviidava katse juures viibis vaatejana ka vastutav uurija.

### **Andmete analüüs**

Andmete analüüsimiseks kasutati RStudiot, versioon 4.0.3 (R Core Team, 2020). Andmeanalüüsiks kasutati ka *readr* (Wickhan ja Hester, 2000), *dplyr* (Wickham, François, Henry ja Müller, 2020), *psych* (Revelle, 2020), *tidyr* (Wickham, 2020) ning *gmodels* (Warnes, Bolker, Lumley, Johnson, 2018) meetmepaketti (*package*).

Uuringus olid olulisteks muutujateks keskmine äratundmistäpsus (0-1), keskmine kindlushinnang (0-100) ning keskmine hinnang enda keskendumisele (0-100). Keskmise äratundmistäpsuse saamiseks kodeeriti vastused binaarsele skaalale, kus 0 tähistas valet otsust ning 1 õiget otsust ning keskmistati kolme video äratundmise tulemused iga katseisiku kohta. Keskmise kindlushinnangu saamiseks keskmistati katseisiku poolt antud kindlushinnangud enda otsuste õigsuse kohta. Kõik uuringus olulised muutujad olid pideval skaalal.

Selleks, et leida katsekeskkonna mõju keskmisele äratundmistäpsusele ja kindlushinnangule, kasutati mitmest regressiooni. Lisaks kasutati mitmest regressiooni, et uurida, milline mõju oli äratundmistäpsusele uuringus olulistel muutujatel. Keskendumishinnangu erinevust katsekeskkondade vahel uuriti t-testiga.

## Tulemused

### Äratundmistäpsus

Katsekeskkonna mõju keskmisele äratundmistäpsusele uuriti mitmese lineaarse regressiooniga. Pärast veebiuuringu valimi kitsendamist uuriti kõigepealt valimi erinevusi veebiuuringu lõppvalimi ja laboriuuringu valimi vahel. Katses osalenute keskmine vanus erines statistiliselt oluliselt katsekeskkonna lõikes,  $t(1545.3) = -21.98$ ,  $p < 0.001$ . Laboriuuringus osalenud katseisikute keskmine vanus oli madalam ( $M = 26.15$ ,  $SD = 7.59$ ) kui veebiuuringus osalenute keskmine vanus ( $M = 34.85$ ,  $SD = 8.66$ ). Lisaks erines katsekeskkonna lõikes ka katseisikute sooline jaotus,  $X^2(1, N = 1738) = 7.8$ ,  $p = 0.0052$ . Seega võeti äratundmistäpsuse ennustamiseks regressioonanalüüsi lisaks katsekeskkonnale kaasa ka sugu ja vanus ning nende interaktsioonid. Leiti, et regressioonimudel on statistiliselt oluline,  $F(7, 1730) = 5.02$ ,  $p < 0.001$ ,  $R^2 = 0.02$ . Tulemused on esitatud tabelis 1. Tulemustest selgus, et sugu ega ükski interaktsioon muutujaga sugu ei ennustanud statistiliselt oluliselt äratundmistäpsust.

Tabel 1

*Katsekeskkonna, soo ja vanuse mõju äratundmistäpsusele*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>CI</i>	<i>p</i>
(Vabaliige)	.59	.05	12.92	(.50; .68)	< 0.001***
Katsekeskkond	-.19	.06	-2.93	(-.31; -.06)	0.004**
Vanus	-0.004	.002	-2.16	(-.01; -.00)	0.03*
Sugu	-.14	.08	-1.76	(-.30; .02)	0.08
Katsekeskkond*Vanus	.004	.002	2.09	(.00; .01)	0.04*
Katsekeskkond*Sugu	.10	.11	0.95	(-.11; .31)	0.34
Vanus*Sugu	.002	.003	0.76	(-.00; .01)	0.45
Katsekeskkond*Sugu*Vanus	-.002	.004	-0.50	(-.01; .01)	0.61

*Märkus.* *B* – fikseeritud mõjude hinnang, *SE* – standardviga, *t* – *t*-statistik, *CI* – 95% usaldusvahemik, *p* – olulisustõenäosus, \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < 0.001$

Kuna sugu ei mõjutanud äratundmistäpsust statistiliselt oluliselt, otsustati täpsemate tulemuste saamiseks viia läbi uus mitmene regressioonanalüüs, kust jäeti välja sugu. Lõplikusse regressioonimudelisse kaasati lisaks katsekeskkonnale ka katseisikute vanus ja nende kahe muutuja interaktsioon. Leiti statistiliselt oluline regressioonivõrrand,  $F(3, 1734) = 5.93$ ,  $p < 0.001$ ,  $R^2 = 0.01$ . Tulemused on esitatud tabelis 2. Tulemustest selgus, et nii vanus, katsekeskkond kui ka nende interaktsioon omasid statistiliselt olulist mõju äratundmistäpsusele. Vanuse suurenedes ühe aasta võrra, langeb äratundmistäpsus 0.004 võrra. Võrreldes laboriga, oli veebis äratundmistäpsus 0.15 võrra madalam.

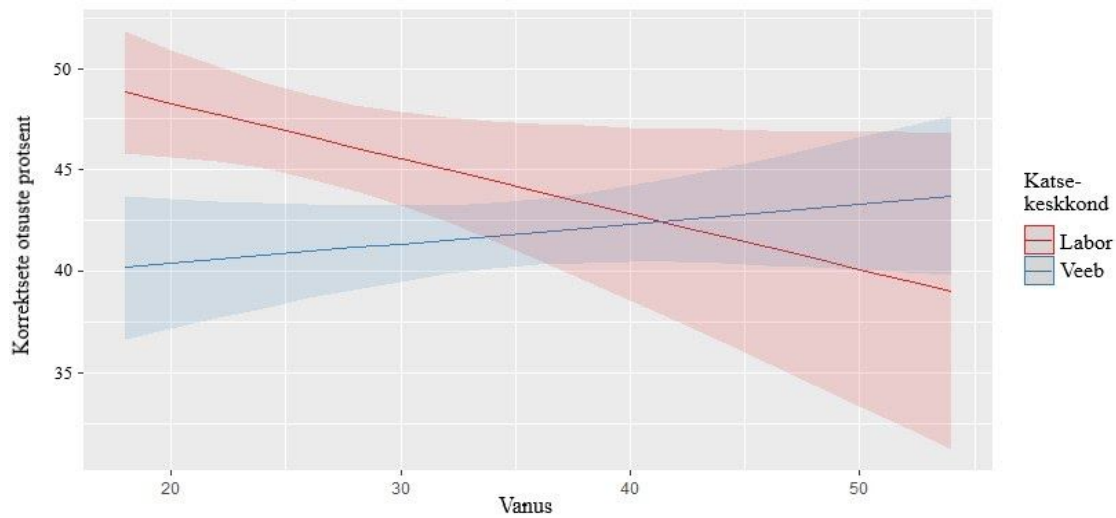
Tabel 2

*Vanuse ja katsekeskkonna mõju äratundmistäpsusele*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>CI</i>	<i>p</i>
(Vabaliige)	.54	.04	14.32	(.46; .61)	< .001***
Vanus	-.003	.001	-1.98	(-.01; -.00)	.048*
Katsekeskkond	-.15	.05	-3.02	(-.25; -.05)	.003**
Vanus*Katsekeskkond	.004	.002	2.21	(-.00; .01)	.03*

*Märkus.* *B* – fikseeritud mõjude hinnang, *SE* – standardviga, *t* – *t*-statistik, *CI* – 95% usaldusvahemik  
*p* – olulisustõenäosus, \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < 0.001$

Interaktsiooni tõlgendamiseks koostati joonis. Jooniselt 1 on näha, et laboris vähenes vanuse suurenedes äratundmistäpsus. Veebis seevastu suurenes äratundmistäpsus vanuse suurenedes.



Joonis 1. Vanuse mõju äratundmistäpsusele katsekeskkonna lõikes.

### Kindlushinnang

Järgmisena kasutati regressioonanalüüsi, et vaadata, kuidas ennustavad katsekeskkond, vanus, sugu ning nende interaktsioonid kindlushinnangut. Leiti statistiliselt oluline regressioonivõrrand,  $F(7, 1730) = 4.75$ ,  $p < 0.001$ ,  $R^2 = 0.01$ . Tulemused on esitatud tabelis 3. Veebis sooritanute kindlushinnang oli keskmiselt 14.45 võrra väiksem kui laboris sooritanute kindlushinnang. Vanuse suurenedes ühe aasta võrra langes katseisikute kindlushinnang 0.4 punkti võrra. Naiste kindlushinnang oli võrreldes meestega keskmiselt 18.98 madalam.



Tabel 3

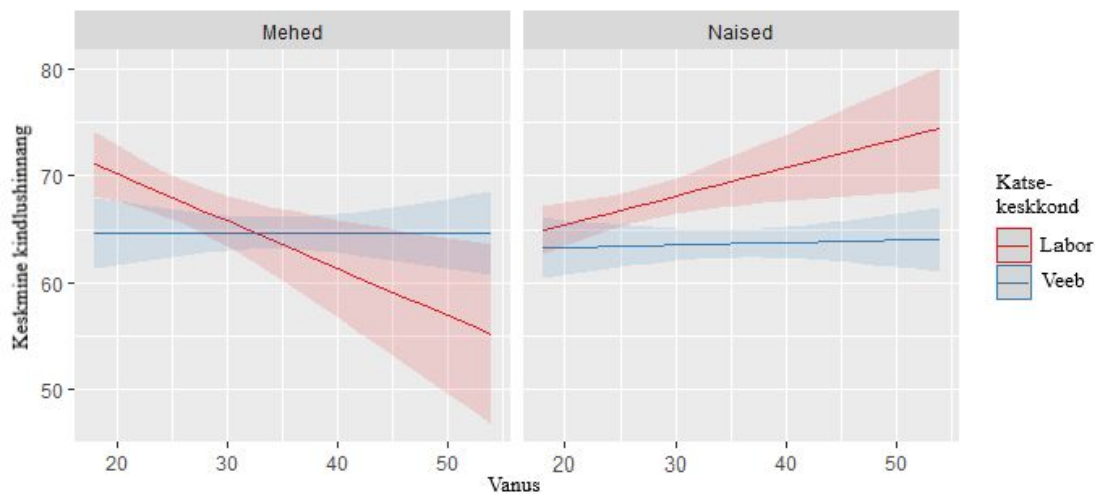
*Vanuse, soo ning katsekeskkonna mõju kindlushinnangule*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>CI</i>	<i>p</i>
(Vabaliige)	79.12	4.00	19.95	(71.34; 86.90)	< .001***
Katsekeskkond	-14.45	5.15	-2.81	(-24.54; -4.35)	.005**
Vanus	-.44	.15	-3.00	(-.73; -.15)	.003**
Sugu	-18.98	4.84	-3.92	(-28.47; -9.48)	< .001***
Katsekeskkond*Vanus	.44	.18	2.53	(.10; .79)	.012*
Katsekeskkond*Sugu	17.18	6.43	2.67	(4.56; 29.79)	.008**
Vanus*Sugu	.71	.18	3.96	(.36; 1.06)	< .001***
Katsekeskkond*Vanus*Sugu	-.69	.22	-3.19	(-1.11; -.26)	.001**

u

*Märkus.* *B* – fikseeritud mõjude hinnang, *SE* – standardviga, *t* – *t*-statistik, *CI* – 95% usaldusvahemik, *p* – olulisustõenäosus, \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < 0.001$

Interaktsioonide tõlgendamiseks kasutati joonist. Jooniselt 2 tuleb välja, et laborikeskkonnas vähenes meeste puhul vanuse kasvades kindlushinnang ning naiste puhul suurenes kindlushinnang koos vanusega. Veebikeskkonnas püsis kindlushinnang nii naiste kui meeste puhul vanuse kasvades stabiilsena.



Joonis 2. Keskmine kindlushinnang soo, vanuse ja kateskeskkonna lõikes.

### Keskendumishinnang

Lisaks uuriti, millised olid erinevused katseisikute keskendumishinnangute vahel katsekeskkondade lõikes. Laboriuuringus osalenud katseisikute keskmine hinnang enda keskendumisele katse ajal oli madalam ( $M = 84.39$ ,  $SD = 13.3$ ) kui veebiuuringus osalenute keskmine hinnang enda keskendumisele katse ajal ( $M = 94.37$ ,  $SD = 10.12$ ). Keskendumise hinnangud erinesid katsekeskkondade vahel teineteisest statistiliselt oluliselt,  $t(1128.9) = -16.57$ ,  $p < 0.001$ .

### Olulised muutujad

Käesolevas töös kasutatud eksperimendis olid olulisteks sõltumatuteks muutujateks keskmine tunnistaja roll, videos nähtud isikute arv ja äratundmisrea esitusviis. Järgmisena vaadati, kuidas need uuringus olulised muutujad mõjutasid äratundmistäpsust veebis ning laboris. Selleks koostati nii labori kui ka veebiandmete põhjal regressioonimudelid, millesse olid kaasatud uuringus olulised muutujad. Labori puhul leiti statistiliselt oluline regressioonivõrrand,  $F(7, 657) = 8.24$ ,  $p < 0.001$ ,  $R^2 = 0.07$ . Samuti oli veebiuuringu kohta koostatud regressioonivõrrand statistiliselt oluline,  $F(7, 1065) = 12.57$ ,  $p < 0.001$ ,  $R^2 = 0.07$ . Tulemused labori- ning veebieksperimendi mudelite kohta on toodud välja vastavalt tabelites 4 ning 5. Mõlema mudeli puhul oli ainsaks statistiliselt oluliseks muutujaks kurjategijate arv videos.

Tabel 4

*Äratundmistäpsust mõjutavad muutujad laboriuuringus*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>CI</i>	<i>p</i>
(Vabaliige)	.48	.03	16.45	(.42; .54)	< .001 ***
Tunnistaja roll	.08	.04	1.82	(-.01; .16)	.07
Osaliste arv	-.14	.04	-3.41	(-.22; -.06)	< .001 ***
Äratundmisrea esitusviis	-.00	.04	-.01	(-.08; -.06)	1
Roll*Osaliste arv	-.01	.06	-.21	(-.13; .10)	.84
Roll*Esitusviis	.04	.06	.71	(-.07; .16)	.48
Osaliste arv*Esitusviis	.07	.06	1.17	(-.04; .19)	.24
Roll*Osaliste arv*Esitusviis	-.06	.08	-.75	(-.22; .10)	.45

*Märkus.* Osaliste arv - kurjategijate arv videos, *B* – fikseeritud mõjude hinnang, *SE* – standardviga, *t* – *t*-statistik, *CI* – 95% usaldusvahemik, *p* – olulisustõenäosus, \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < 0.001$

Tabel 5

*Äratundmistäpsust mõjutavad muutujad veebiuuringus*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>CI</i>	<i>p</i>
(Vabaliige)	.44	.02	19.81	(.40; .49)	< .001***
Tunnistaja roll	.05	.03	1.62	(-.01; .11)	.11
Osaliste arv	-.11	.03	-3.40	(-.17; -.05)	< .001***
Esitusviis	-.02	.03	-.60	(-.08; .05)	.55
Roll*Osalised	-.05	.04	-1.15	(-.14; .04)	0.25
Roll*Esitusviis	.07	.04	1.63	(-.01; .16)	.10
Osalised*Esitusviis	.02	.04	.49	(-.07; .11)	.63
Roll*Osaliste arv*Esitusviis	.01	.06	.16	(-.11; .13)	.872

*Märkus.* Osaliste arv - kurjategijate arv videos, *B* – fikseeritud mõjude hinnang, *SE* – standardviga, *t* – *t*-statistik, *CI* – 95% usaldusvahemik, *p* – olulisustõenäosus, \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < 0.001$

### Arutelu

Antud uuringu eesmärgiks oli selgitada, kas esineb erinevusi veebi- ja laborikeskkonnas sooritatud eksperimentide tulemuste vahel. Leiti, et veebikeskkonnas eksperimenti sooritanute äratundmistäpsus ning kindlushinnang olid madalamad kui laborikeskkonnas sooritanute äratundmistäpsus. Seega ei leidnud püstitatud hüpoteesi kinnitust.

### Katsekeskkonna mõju äratundmistäpsusele

Keskmise äratundmistäpsuse erinevus katsekeskkondade vahel võib tuleneda eksperimendi nõudlikkusest. Finley ning Penningroth (2015) uurisid, millised on veebis läbi viidud kognitiivselt nõudliku eksperimendi tulemused. Nad järeldasid, et viies läbi nõudlike eksperimente veebis, kannatab andmekvaliteet osaleja tähelepanematuse ning unustamise tõttu. Seda seetõttu, et veebieksperimendi puhul võivad katseisiku keskendumist mõjutada

segavad stiimulid katse ajal (Pauszek, Sztybel ja Gibson, 2017). Käesolevas töös kasutatud eksperiment eeldas katseisikult tähelepanu juhtimist ülesandele ning mälu aktiivset kasutamist, mistõttu võiski veebieksperimendi puhul olla keskmine äratundmistäpsus madalam.

Leitud erinevused äratundmistäpsuses katsekeskkondade vahel ei ühti Clifford ja Jerit (2014) töö tulemustega, milles ei leitud märgatavaid erinevusi veebi- ning laboriuuringu tulemuste vahel. Nende uuring keskendus aga varasemalt omandatud teadmiste kontrollile ning oli küsimustiku vormis. Käesoleva uuringu eksperiment nõudis katseisikult tähelepanu, keskendumist ning katse ajal saadud teadmiste meenutamist, mis võis kaasa tuua erinevuse veebi- ning laborikatse tulemuste vahel (Finley ja Penningroth, 2015). Seega võib uuringute lahknevusi selgitada nende protseduuride ning eesmärkide erinevusega.

### **Katsekeskkonna mõju kindlushinnangule**

Tulemused näitavad, et lisaks keskmisele äratundmistäpsusele erines katsekeskkondade vahel ka keskmine kindlushinnang, olles veebikeskkonnas madalam kui laboris. See võis tuleneda sellest, et katseisikud pööravad veebieksperimentides vähem tähelepanu (Finley ja Penningroth, 2015), mistõttu ei saanud nad ka oma otsuses kindlad olla.

Keskmise kindlushinnangu erinevus katsekeskkondade vahel võis tuleneda ka katseisikute erinevast keskendumise määrast katse ajal. Veebi- ning laborieksperimendi meetodid erinesid aja poolest, mis oli videote nägemise ning äratundmisrea esitamise vahel. Laborikatses luges eksperimentaator katseisikule peale iga video nägemist ette samad juhised otsuse langetamiseks, pikendades sellega aega video ning äratundmisrea nägemise vahel. Veebieksperimendis ei pidanud katseisikud peale video nägemist instruksioonidel peatuma, vaid võisid kohe edasi liikuda äratundmisreani, mis viis väiksema ajani videote nägemise ning äratundmisrea vahel. Viimase põhjal võiks eeldada, et veebis osalenutel olid äratundmisrida vaadates videos nähtud isikud paremini meeles, sest meenutamise ajal oli stiimulvideost vähem aega möödas (Ebbinghaus, 2013). Selle aga lükkavad ümber nii madalam äratundmistäpsus kui ka madalam kindlushinnang veebis, mis tähendabki, et veebis osalenud katseisikud ei pööranud mingil põhjusel katse ajal nii palju tähelepanu, kui laboris osalenud katseisikud.

Tulemused näitavad, et veebis püsis nii soo kui vanuse lõikes kindlushinnang ühel tasemel, kuid laboris langes meeste kindlushinnang vanuse kasvades ning naiste keskmine kindlushinnang tõusis vanuse suurenedes. On teada, et katseisikud pööravad veebis eksperimendile vähem tähelepanu kui laborieksperimendi puhul (Finley ja Penningroth,

2015), mis võib tuleneda suuremast hulgast sekkuvatest muutujatest (Pauszek, Sztybel ja Gibsonn, 2017). See tähendab, et veebis eksperimenti sooritanud võisid tähelepanematus e tõttu juhustest valesti aru saada või polnud nad piisavalt motiveeritud eksperimenti süvenema, mistõttu ei avaldunud soo ega vanuse mõju kindlushinnangule selliselt nagu laboris.

### **Hinnang keskendumisele**

Võrreldes laboriga, oli veebis osalenud katseisikute hinnang enda keskendumisele katse ajal kõrgem. Ühest küljest võis laboris osalenute hinnang olla madalam, sest katse kestus oli pikem kui veebis (katses kolm osa võrreldes ühega veebis) ning hinnangut oma keskendumisele katse ajal küsiti alles katse lõpul. Pikem katse nõudis katseisikutelt rohkem ning nad olid katse lõpuks väsinumad.

Veebikatses osalenute kõrgem hinnang enda keskendumisele katse ajal võis tuleneda ka hirmust, et nad jäävad katses osalemise eest mõeldud tasust ilma, kui nad märgivad liiga madala keskendumishinnangu. Viimast kinnitab ka madalam äratundmistäpsus ja kindlushinnang veebis. Madalam kindlushinnang ja äratundmistäpsus võiksid viidata ka pööratud tähelepanu madalamale tasemele.

See tähendab, et kuigi veebis osalenud katseisikud hindasid enda keskendumist katse ajal kõrgemaks, võisid nad seda teha teistel ajenditel. Kõrgem hinnang keskendumisele ei ühti teiste katses leitud tulemuste erinevustega.

### **Katses olulised muutujad**

Tulemustest selgus veel, et kuigi kahe katsekeskkonna valimid ning äratundmistäpsus erinesid, olid oluliste muutujatega regressioonmudelid mõlemal juhul sarnased ning nii veebis kui laboris ennustas äratundmistäpsust statistiliselt oluliselt ainult kurjategijate arv nähtud videos. See järeldus kattub Birnbaum (1999) ning Birnbaum (2000) veebi- ning laborikatseid võrdlevas töödes tehtud järeldustega, milles leiti, et hoolimata tulemuste ning valimi erinevustest, olid saadud andmete põhjal tehtud järeldused mõlema katsekeskkonna puhul samad.

### **Uurimistöö piirangud ning edasised uurimisvõimalused**

Võib öelda, et kuigi äratundmistäpsus oli kahe katsekeskkonna vahel erinev, mõjutasid seda siiski samad faktorid. Ei saa öelda, et need tulemused oleksid ammendavad, sest valimite soolises ja vanuselises jaotuses esines märgatav vahe. Lisaks erines katse protseduur katsekeskkondade lõikes, mis võis põhjustada erinevusi kindlushinnangutes.

Edasised uuringud peaksid sekkuvate muutujate minimaliseerimiseks püüdma ühtlustada valimite jaotuse ning hoida võimalikult sarnasena katse meetodid ning protseduuri. Lisaks võiks ausate tulemuste saamiseks veebiuuringutes katse algul katseisikut informeerida sellest, et antud vastuste õigsus ei mõjuta nende õigust saada katses osalemise eest mõeldud tasu.

### **Kokkuvõte**

Käesoleva uuringu eesmärgiks oli leida, millised on erinevused labori- ning veebieksperimendi tulemuste vahel. Selgus, et veebieksperimendis olid katseisikud eksperimendi ajal vähem tähelepanelikud ning see võis viia ka madalamate tulemusteni. Katseisikute enda antud hinnang keskendumisele katse ajal aga ei olnud kooskõlas teiste leitud tulemustega. Kokkuvõtteks võib järeldada, et veebieksperimente võib edukalt kasutada erinevate uuringute läbiviimiseks (näiteks Clifford ja Jerit, 2014), kuid sekkuvate muutujate suurema hulga tõttu võivad veebis läbi viidud tähelepanu nõudvad eksperimendid põhjustada laboritulemustega võrreldes erinevaid tulemusi.

**Kasutatud kirjandus**

- Bamert, T. (2002). Integration von Wahrscheinlichkeiten: Verarbeitung von zwei Wahrscheinlichkeitsformationen (Doctoral dissertation).
- Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161-188.
- Birnbaum, M. H. (2000). Decision making in the lab and on the Web. In *Psychological experiments on the Internet* (pp. 3-34). Academic Press.
- Birnbaum, M. H. (1999). Testing critical properties of decision making on the Internet. *Psychological Science*, 10(5), 399-407.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2016). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data?. *Perspectives on Psychological Science*, 6(1), 3-5.
- Clifford, S., & Jerit, J. (2014). Is there a cost to convenience? An experimental comparison of data quality in laboratory and online studies. *Journal of Experimental Political Science*, 1(2), 120-131.
- Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of neurosciences*, 20(4), 155.
- Finley, A., & Penningroth, S. (2015). Online versus in-lab: Pros and cons of an online prospective memory experiment. *Advances in psychology research*, 135-161.
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic bulletin & review*, 19(5), 847-857.
- Krantz, J. H. (2000). Tell me, what did you see? The stimulus on computers. *Behavior research methods, instruments, & computers*, 32(2), 221-229.
- Levett, L. M., & Kovera, M. B. (2008). The effectiveness of opposing expert witnesses for educating jurors about unreliable expert evidence. *Law and Human Behavior*, 32(4), 363-374.
- Lowrey, B. V., Maguire, E. R., & Bennett, R. R. (2016). Testing the effects of procedural justice and overaccommodation in traffic stops: A randomized experiment. *Criminal justice and behavior*, 43(10), 1430-1449.
- Moss, A. J., Rosenzweig, C., Robinson, J., & Litman, L. (2020). Is it Ethical to Use Mechanical Turk for Behavioral Research? Relevant Data from a Representative Survey of MTurk Participants and Wages.



- Musch, J., & Reips, U. D. (2000). A brief history of Web experimenting. In *Psychological experiments on the Internet* (pp. 61-87). Academic Press.
- Oberauer, K. (2019). Working memory and attention—A conceptual analysis and review. *Journal of cognition*, 2(1).
- Pauszek, J. R., Szybel, P., & Gibson, B. S. (2017). Evaluating Amazon's Mechanical Turk for psychological research on the symbolic control of attention. *Behavior research methods*, 49(6), 1969-1983.
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental psychology*, 49(4), 243.
- Revelle, W. (2020) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych>
- Robinson, J., Rosenzweig, C., Moss, A. J., & Litman, L. (2019). Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool. *PloS one*, 14(12).
- Semules, A. (2018). The Internet Is Enabling a New Kind of Poorly Paid Hell. <https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-tuk/551192/>
- Sheehan, K. B. (2018). Crowdsourcing research: data collection with Amazon's Mechanical Turk. *Communication Monographs*, 85(1), 140-156.
- Warnes, G. R., Bolker, B., Lumley, T., & Johnson, R. C. (2015). gmodels: Various R programming tools for model fitting. R package version, 2(3).
- White, A., Strezhnev, A., Lucas, C., Kruszewska, D., & Huff, C. (2018). Investigator characteristics and respondent behavior in online surveys. *Journal of Experimental Political Science*, 5(1), 56-67.
- Wickham, H. (2020). tidyr: Tidy Messy Data. R package version 1.1.2. <https://CRAN.R-project.org/package=tidyr>
- Wickham, H., François, R., Henry, L., Müller, K. (2020). dplyr: A Grammar of Data Manipulation. R package version 1.0.2. <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., & Hester, J. (2020). readr: Read Rectangular Text Data. R package version 1.4.0. <https://CRAN.R-project.org/package=readr>

*Käesolevaga kinnitan, et olen korrekselt viidanud kõigile oma töös kasutatud teiste autorite poolt loodud kirjalikele töödele, lausetele, mõtetele, ideedele või andmetele.*

*Olen nõus oma töö avaldamisega Tartu Ülikooli digitaalarhiivis DSpace alates 01.01.2023.*

*/Kevin Mändmets/*