

Näited üldistava statistika meetodite kasutamisest.

Näited on kogutud PS eriala üliõpilaste poolt kevad/2013

SISUKORD

1. VAHEMIKHINNANG	2
2. VAHEMIKHINNANG	3
3. VAHEMIKHINNANG	4
4. T-TEST (SÕLTUVAD VALIMID)	4
5. T-TEST (SÕLTUVAD VALIMID)	5
6. T-TEST (SÕLTUMATUD VALIMID)	6
7. T-TEST (SÕLTUMATUD VALIMID)	6
8. DISPERSIOONANALÜÜS	7
9. DISPERSIOONANALÜÜS	8
10. HII RUUT TEST	9
11. HII RUUT TEST	9
12. KORRELATSIOONANALÜÜS	10

1. VAHEMIKHINNANG

Eesti tööjõu-uuring (Metoodika). Statistikaamet. Tallinn 2012

Uuringu eesmärk: Tööjõu-uuringut on vaja, et saada ülevaade Eesti elanike tööhõivest, töötusest, tööoludest ja tööturu muutustest.

Tööjõu-uuringu üldkogumis on kõik Eestis elavad tööealised isikud. Tööealised on inimesed, kes uuringunädalal olid 15–74-aastased.

Valimi jaotus aastate ja kvartalite kaupa, 2000–2011

	I	II	III	IV	Aastas kokku
2000	2186	2189	2277	2475	9127
2001	2463	2467	2480	2487	9897
2002	2425	2407	2520	2529	9881
2003	2431	2412	2516	2518	9877
2004	2530	2522	2492	2480	10 024
2005	2327	2196	2198	2178	8899
2006	2369	2583	2582	2523	10 057
2007	2812	3082	3032	2973	11 899
2008	3006	3034	2997	2940	11 977
2009	2957	2893	2874	2890	11 614
2010	2992	3026	2986	2955	11 959
2011	2902	3051	2981	2935	11 869

Näiteks oli 2011. aastal töötuse määra hinnang 12,5% ja standardviga 0,48. 95-protsendilise usaldusvahemiku otspunktid on $12,5 \pm 1,96 \times 0,48 \approx 12,5 \pm 0,9$. Teiste sõnadega, tõenäosusega 0,95 on 2011. aasta tegelik töötuse määr 11,6% ja 13,4% vahel.

2011. aasta põhinäitajate hinnangud koos usaldusvahemiku ja suhtelise standardveaga on esitatud tabelis 3, 2000.–2011. aasta töötuse määra hinnangud koos usaldusvahemiku ja suhtelise standardveaga tabelis 4.

Tabel 3 Hinnangud koos usaldusvahemiku ja suhtelise standardveaga, 2011

	Hinnang	Usaldusvahemik tõenäosusega 0,95	Suhteline standardveiga, %
Tööjõud, tuhat	695,9	±15,0	1,1
hõivatud, tuhat	609,1	±14,5	1,2
töötud, tuhat	86,8	±6,9	4,1
Mitteaktiivsed, tuhat	333,8	±10,4	1,6
15–74-aastased kokku, tuhat	1029,8	±12,0	0,6
Tööjõus osalemise määr, %	67,6	±1,0	0,8
Tööhõive määr, %	59,1	±1,1	1,0
Töötuse määr, %	12,5	±0,9	3,9

Tabel 4 Töötuse määra hinnang usaldusvahemiku ja suhtelise standardveaga, 2000–2011

	Hinnang, %	Usaldusvahemik tõenäosusega 0,95	Suhteline standardveiga, %
2000	13,6	±0,9	3,4
2001	12,6	±1,0	4,1
2002	10,3	±0,9	4,5
2003	10,0	±0,9	4,8
2004	9,7	±1,0	5,3
2005	7,9	±0,9	6,0
2006	5,9	±0,7	6,1
2007	4,7	±0,6	6,1
2008	5,5	±0,6	5,2
2009	13,8	±1,0	3,7
2010	16,9	±1,1	3,4
2011	12,5	±0,9	3,9

2. VAHEMIKHINNANG

Vaatlusalusteks oli 18 noorsuusatajat, kes olid suusatreeninguga tegeleunud vähemalt 4 aastat. Pedagoogilise eksperimendi läbiviimiseks moodustati 2 gruppi – eksperimentaal- ja kontrollgrupp. Vaatlusaluste gruppide üldandmed on toodud tabelis 1. Eksperimendi alustamisel gruppide vahelisi statistiliselt usaldusväärseid erinevusi uuritud näitajates ei olnud (Tabel 2,3).

Tabel 1. Vaatlusaluste gruppide üldandmed ($X \pm SD$) uuringu algul.

Grupp	n	Vanus (aastat)	Kasv (cm)	Kehamass (kg)
Eksperimentaalgrupp	9	15,5 ± 1,09	173,1 ± 7,4	61,8 ± 7,0
Kontrollgrupp	9	15,8 ± 1,87	171,7 ± 10,6	59,0 ± 10,3

3. VAHEMIKHINNANG

Üldistuste jaoks arvutatakse välja standardhälbe, kus p on antud vastuste osakaal valimis ning N - kõikide vastuste arv.

$$\sigma = \sqrt{\frac{p(1-p)}{N}}$$

Selleks, et teha üldistused 95% tõenäosusega, tuleb kasutada ligikaudu $\pm 2\sigma$, täpsemalt $\pm 1,96\sigma$. Näiteks, õpetajatelt küsiti, missuguste internetipõhiste õppematerjalide vajadus on nende arvates kõige suurem? Õpetajad leidsid, et kõige suurem puudus on videotest 52 õpetajat (86%).

$$p = 52/60 = 0,86(86\%) \quad \Delta p = t_{\beta} \sqrt{\frac{p(1-p)}{N}}$$

$$\Delta p = t_{\beta} \sqrt{\frac{0,86 \cdot (1-0,86)}{60}} = 0,011$$

Saame $86\% \pm 1\%$. Niisiis, võime teha üldistuse, et 95% tõenäosusega $86\% \pm 1\%$ õpetajatest tunnevad puudust lisamaterjali osas õppevideotest.

4. T-TEST (SÕLTUVAD VALIMID)

Gadi Rohtjärg – “Kehaliste võimete arengu dünaamika Tallinna koolide algklasside õpilastel”.

Magistritöö. Tallinn, 2012

Käesoleva magistritöö eesmärgiks oli välja selgitada Tallinna koolide algklasside õpilaste kehalise arengu ja kehalise võimekuse näitajad. Eesmärgi saavutamiseks viidi kolmeteistkümnes Tallinna koolis läbi 9-10-aastaste laste kehalise võimekuse testimine, mis koosnes paigalt kaugushüppe, käe jõu, 10x5 meetrit süstikjooksu, 40 meetrit sprintjooksu, vastupidavus-süstikjooksu, istest ettepainduvuse ning flamingo tasakaalu testidest. Student T-testi ning korrelatsioonanalüüsi abil uuriti, kuidas on muutunud kehaliste võimete omavahelised seosed ning seosed antropomeetriliste näitajate vahel.

Valimi moodustasid 459 7-8-aastast last esimesel mõõtmisel ja 459 9-10-aastast last teisel mõõtmisel, kellest 225 olid poisid ja 234 tüdrukud. Uuringus osalesid 2007/2008 õppeaastal esimese klassi 7-8-aastased õpilased ja 2009/2010 õppeaastal samad õpilased, kuid siis olid juba 9-10-aastased ja käisid kolmandas klassis.

Keskmi võrreldi Studenti T-testi põhjal, et teada saada kahe erineva aasta muutuste usaldusväärsus. Samuti kasutati Student T-testi ka poiste ja tüdrukute erinevuste hindamiseks ja seda mõlemal õppeaastal. Statistilise olulisuse nivoo võeti $p < 0,05$.

Antud töö põhjal võime teha järgmised järeldused:

- Esimeses ja kolmandas klassis mõõdetud testide keskmised tulemused paranesid statistiliselt usutavalt, välja arvatud painduvustest, kus tüdrukute keskmine näitaja jäi samaks, poiste tulemus aga halvenes statistiliselt usutavalt.

- Kui esimeses klassis kehakaal jõu ja kiiruse tulemustele mõju ei avaldanud, siis kolmandas klassis ilmnes nende näitajate vahel väga tugev usutav seos ($p < 0,001$).
- Kui esimeses klassis oli usutav seos pikkuse ja vastupidavuse ning tasakaalu näitajate vahel, siis kolmandas klassis enam sellist seost ei avaldunud.
- Kui jõu näitaja esimeses klassis keha pikkusest ei sõltunud, siis kolmandas klassis saavutasid pikemad õpilased selle näitaja puhul usutavalt paremaid tulemusi.
- Kolmandas klassis ilmnes statistiliselt usutav seos mõlema jõunäitaja vahel, mis esimeses klassis veel ei avaldunud.

5. T-TEST (SÕLTUVAD VALIMID)

Jaana Liigand – “Kuidas mõjutab teadvustamatult esitatud samasuunalist või vastassuunalist omadust kandev eelnev nägu järgneva näo isiksuseomaduste tajumist”.

Magistritöö. Tartu, 2012

Antud magistritöö eesmärk on uurida, kas ja kuidas mõjutab teadvustamatult esitatud samasuunalist või erisuunalist omadust kandev eelnev ehk praimiv nägu järgneva näo isiksuseomaduste tajumist.

Uuring viidi läbi katse kujul, kus 72 katseisikut hindasid 80 eelhinnatud näopilti, mis kandsid 5 isiksuseomadust: usaldusväarsust, mõjutatavust, heatahtlikkust, ausust ja alluvust. Iga nägu näidati katseisikule arvutiekraanilt 300 millisekundi jooksul ning igale näole eelnes 20 millisekundi jooksul samasuunalist või vastassuunalist isikuomadust kandev (näiteks usaldusväärne-usaldusväärne või mitteusaldusväärne-usaldusväärne) praimiv näopilt.

Kasutades paariviisilist t-testi selgus, et statistiliselt olulist erinevust samasuunalise ($M = 6,17$, $SD = 0,72$) ja vastassuunalise ($M = 6,14$, $SD = 0,78$) praimi tingimustes ei leitud ($t(71) = 0,62$, $p = 0,54$).

Järgnevalt analüüsiti eraldi kõiki isiksuseomaduste kategooriaid, mida katse puudutas ning võrreldi paariviisilise t-testiga nendevahelisi erinevusi. Statistiliselt olulist erinevust samasuunalise ja vastassuunalise praimimise tingimustes ei leitud ei usaldusväarsuse (samasuunalise praimimise korral $M = 5,95$, $SD = 1,07$ ning vastassuunalise korral $M = 5,86$, $SD = 1,10$, $t(71) = -0,65$, $p = 0,52$), mõjutatavuse (samasuunalise praimimise korral $M = 6,31$, $SD = 1,13$ ning vastassuunalise korral $M = 6,31$, $SD = 1,15$, $t(71) = 0,07$, $p = 0,95$), heatahtlikkuse (samasuunalise praimimise korral $M = 6,53$, $SD = 1,08$ ning vastassuunalise korral $M = 6,59$, $SD = 1,08$, $t(71) = 0,49$, $p = 0,63$), aususe (samasuunalise praimimise korral $M = 5,98$, $SD = 1,08$ ning vastassuunalise korral $M = 6,03$, $SD = 1,21$, $t(71) = 0,70$, $p = 0,49$) ega alluvuse (samasuunalise praimimise korral $M = 6,04$, $SD = 0,90$ ning vastassuunalise korral $M = 5,92$, $SD = 1,06$, $t(71) = 1,03$, $p = 0,31$) puhul.

6. T-TEST (SÕLTUMATUD VALIMID)

Statsionaarse ja mittestatsionaarse õppevormi üliõpilaste akadeemilise enesetõhususe võrdlus Tartu Ülikooli näitel.

Koostaja: Kristina Koidu

- a. Hüpotees: Neil mittestatsionaarse õppevormi üliõpilastel, kes ei osale akadeemilises tegevuses esmakordselt on teistega võrreldes kõrgem hinnang akadeemilisele enesetõhususele ja paremad akadeemilised tulemused. Läbiviidud sõltumatute valimitega t-test näitas, et akadeemilise enesetõhususe osas ei esine võrreldavatel gruppidel statistilist erinevust $t(df=115) = .797$, $p = .427$, $p > .05$. Samade gruppide võrdlemisel on statistiliselt oluline akadeemiliste tulemuste erinevus $t(df=126) = 2.434$, $p = .016$, ehk $p < .05$. Seega osutus see hüpotees osaliselt tõseks: nendel mittestatsionaarse õppevormi õpilastel, kes ei osale akadeemilises tegevuses esmakordselt, on teistega võrreldes paremad akadeemilised tulemused.
- b. Hüpotees: Mittestatsionaarse õppevormi üliõpilased ja tavaõppe üliõpilased erinevad üksteisest akadeemilise enesetõhususe osas. Sõltumatute valimitega t-test näitas, et akadeemilise enesetõhususe osas esineb statsionaarsetel üliõpilastel ($M = 66.6$, $SD = 8.72$) ja mittestatsionaarse õppevormi üliõpilastel ($M = 60.0$, $SD = 7.22$) statistiline erinevus $t(df=356) = 7.879$, $p = .000$, ehk $p < 0,05$, mis kinnitab püstitatud hüpoteesi. Statsionaarse õppe üliõpilastel on keskmiselt kõrgem akadeemilise enesetõhususe hinnang kui mittestatsionaarse õppevormi üliõpilastel.

7. T-TEST (SÕLTUMATUD VALIMID)

Kutse - ja tavakoolide õpilaste õpihoiakute erinevused

Kutsekoolides ja tavakoolides õppivate õpilaste enesetõhususe võrdlemiseks viisin läbi sõltumatute katsegruppidega t-testi, mis näitas, et statistiliselt olulist erinevust enesetõhususes ei leidu (M kutsekoolid = 2,75, $SD = 0,62$; M tavakoolid = 2,79, $SD = 0,64$; $t(259) = -0.52$; $p = .605$).

Sama testiga võrdlesin ka kutsekoolide ja tavakoolide akadeemilist enesetõhusust. Ilmnes, et kutsekooli õpilaste akadeemiline enesetõhusus ($M = 2,75$; $SD = 0,56$) on kõrgem kui tavakoolis õppivatel õpilastel ($M = 2,46$; $SD = 0,57$) ning see tulemus on statistiliselt oluline ($t(257) = 3.97$; $p < .001$).

8. DISPERSIOONANALÜÜS

SOOLISED JA VANUSELISED ERINEVUSED VIHATUNNETES, VIHA KONTROLLIS JA SELLE VÄLJENDAMISES (2012).

Käesolevas töös püstitas autor järgmise **hüpoteesi**:

1. Meestel esineb rohkem vihatundeid kui naistel, ning viimased on ka viha väljendamise osas tagasihoidlikumad ja suudavad oma vihaemotsioone suuremal määral kontrollida.

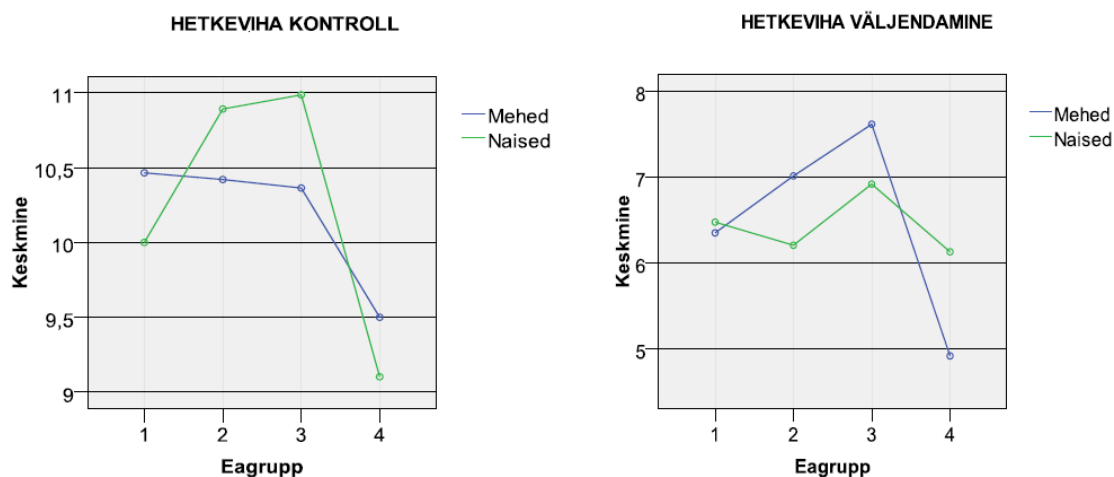
Autor võrdles mehi ja naisi ning erinevaid vanusegrupe ANOVA meetodil. Moodustati neli gruppi:

Grupp 1 : 11-13 aastased (81 vastanut)

Grupp 2: 14-16 aastased (170)

Grupp 3: 17-19 aastased (116)

Grupp4: 20 ja vanemad (99)



Joonis1. Vihaseisundi skaala hetkeviha kontrolli ja väljendamise alaskaalade keskmised tulemused eagruppide ja soo lõikes.

Dispersioonanalüüsi (ANOVA) põhjal selgub, et meeste ja naiste võrdluses on kontrollihinnangud keskmiselt kõrgemad naistel 2. (14-16 aastased) ja 3. (17-19 aastased) vanusegrupis. Naistel toimub hetkeviha kontrolli keskmiste hinnangute järsk langus 3. ja 4. (20 ja vanemad) vanusegrupi vahel. Nimelt on naistel 4. vanusegrupil keskmised kontrollihinnangud väga madalad. Niisamuti langeb ka 4. vanusegrupi meeste vihakontrolli hinnangute kõver järsult allapoole.

Hetkeviha väljendamisel on meeste ja naiste tulemused väga sarnased, kuigi meestel on veidi kõrgemad keskmised hetkeviha väljendamise hinnangud. Kolmandal vanusegrupil on nii meeste kui naiste keskmised hetkeviha väljendamise hinnangud kõige kõrgemad. Suurim sugudevaheline erinevus ilmneb aga taas vanimas eagrupis meeste ja naiste vahel, kus viimastel on kõrgemad keskmised hetkeviha väljendamise hinnangud.

9. DISPERSIOONANALÜÜS

ÕNNELIKKUS JA ELUGA RAHULOLU EESTI- JA VENEKEELSE ELANIKKONNA SEAS (2012).

Autori eesmärk on uurida, kas Eestis elavate eesti- ja vene keelt rääkivate inimeste vahel on erinevusi nende õnnelikkuse ja eluga rahulolu määras aastatel 2004-2010.

Autor püstitas järgmise **hüpoteesi**:

1. Eesti venekeelse elanikkonna õnnelikkus ja eluga rahulolu on madalam võrreldes eestikeelse elanikkonna rahulolu ja õnnelikkusega vahemikus aastatel 2004-2010.

Tulemuste võrdlemiseks kasutab autor dispersioonanalüüsi (one-way ANOVA) ning aastaarvud jagatakse gruppideks: 2004, 2006, 2008 ja 2010.

Tabel 12. Õnnelikkus eesti- ja venekeelse elanikkonna seas 2004-2010

	2004				2006				2008				2010			
	M	SD	df	p	M	SD	df	p	M	SD	df	p	M	SD	df	p
Eestik.	6,33	1,98	df=1		6,93	1,84	df=1		6,92	1,80	df=1		6,99	1,81	df=1	
Venek.	6,10	2,23	df=1972	p= .021	6,51	2,13	df=1475	p= .000	6,15	2,19	df=1606	p= .000	6,58	2,30	df=1785	p= .000

Tabel 10. Eluga rahulolu eesti- ja venekeelse elanikkonna seas 2004-2010

	2004				2006				2008				2010			
	M	SD	df	P	M	SD	df	p	M	SD	df	p	M	SD	df	p
Eestik.	6,05	2,18	df=1		6,61	2,01	df=1		6,59	2,03	df=1		6,79	2,09	df=1	
Venek.	5,46	2,34	df=1979	p= .000	5,94	2,49	df=1487	p= .000	5,21	2,36	df=1615	p= .000	5,50	2,68	df=1784	p= .000

Dispersioonanalüüs (ANOVA) näitab, et võrreldes omavahel eesti- ja venekeelse elanikkonna **eluga rahulolu** aastatel 2004-2010, selgub, et eesti keelt kõnelevad inimesed olid keskmiselt enam eluga rahul kui vene keelt kõnelevad inimesed.

Võrreldes omavahel eesti- ja venekeelse elanikkonna **õnnelikkust** aastatel 2004- 2010, siis selgub, et eestikeelne elanikkond oli keskmiselt enam õnnelik kui venekeelne elanikkond.

10. HII RUUT TEST

Uurimistöös otsiti rahvuste vahelisi erinevusi väärtushinnangute suhtes. Vene ja eesti koolide õpilaste tulemuste võrdlemiseks kasutati χ^2 -testi.

Uuriti missuguste inimestega koos olles on noored kõige õnnelikumad.

Hüpotees, et eesti koolide õpilased peavad vene koolide õpilastest tähtsamaks sotsiaalseid väärtusi ei leidnud kinnitust. Siis pigem leidis osaliselt kinnitust vastupidine (küsimuses „Missuguste inimestega koos olles oled sa kõige õnnelikum?“), et vene koolide õpilased väärtustavad sotsiaalseid väärtusi rohkem kui eesti koolide õpilased.

Tabel 2. χ^2 -testi tulemused väärtuste puhul, missuguste inimestega koos olles on noored kõige õnnelikumad.

väärtused	eestlasi %	venelasi %	χ^2	p
Heasoovlikus	71,8	70,5	0,38	0,85
Kõikehaaravus	0,0	2,6	2,67	0,10
Füüsilised väärtused	4,9	0,0	3,89	0,05
Hedoism	1,9	1,3	0,12	0,73
Röömsameelsus	21,4	6,4	8,95	0,01
Sotsiaalsed väärtused	24,3	56,4	19,44	0,00
Lemmikloom	1,0	1,3	0,04	0,84

11. HII RUUT TEST

Uurimistöös otsiti nakkuse (*Toxocara spp*) esinemist

Toksokariaasi riskitegurite selgitamiseks kasutati hii-ruut testi, kus seos loeti statistiliselt oluliseks, kui p väärtus jäi alla 0,05 (tabel 1).

Tabel 1. *Toxocara spp.* vastaste IgG antikehade esinemine olulisemate tunnuste järgi

Tunnus	Nakatunute arv	Nakatumus	p
Keskkooli lõpuni elukohaks eramaja	22	15,1%	0,0239
Keskkooli lõpuni elukohaks korter	13	7,3%	
Koer on	31	13,1%	0,0185
Koera ei ole	6	5,0%	
Lemmikloom ei saa anthelmintikumi	7	18,9%	0,0056
Lemmikloom saab anthelmintikumi	22	8,8%	

12. KORRELATSIOONANALÜÜS

Helen Puusemp

Hirmu seos seisundi- ja püsiärevusega teismelistel ja noortel. 2012

Autor on püstitanud kolm hüpoteesi.

1. **Eesti teismelistel ja noortel on nii seisundiärevusel kui ka püsiärevusel hirmudega tugev positiivne korrelatsioon.**
2. **Noorema kohordi hirmude ja ärevuse seose suurus ei erine märkimisväärselt vanema kohordi hirmude ja ärevuse seose suurusest.**
3. **Eesti neistel on nii hirmude kui ka ärevuse näitajad kõrgemad kui meestel, olenemata hirmuobjektist või sellest, kas on tegemist püsi- või seisundiärevusega.**

Pearson'i korrelatsiooni koefitsient r näitas, et nii koguhirmul kui ka üksikutel hirmuobjektidel oli nii püsiärevusega kui ka seisundiärevusega arvestatav positiivne korrelatsioon (tabel 1). Leitud seosed olid statistiliselt märkimisväärsed, välja arvatud hirmuobjekti hiired-rotid ja seisundiärevuse vaheline korrelatsioon, milles statistilist olulisust ei leitud, vastavate tunnuste vaheline seos on juba enam kui 5% tõenäosusega juhuslik.

Ilmnes, et koguhirmu ja püsiärevuse vaheline seos ($r=0,38$; $p < 0,001$) on tugevam kui koguhirmu ja seisundiärevuse vaheline seos ($r=0,25$; $p < 0,001$). Seisundi- ja püsiärevuse vaheline korrelatsioon oli märkimisväärne ($r=0,64$; $p < 0,001$).

Vaadates erinevate hirmuobjektide seoseid püsi- ja seisundiärevusega, selgus, et kõige kõrgem positiivne korrelatsioon kokkuvõtlikult püsi- ja seisundiärevuse näitajatega oli hirmuobjektidel surnud loomad, et keegi murrab korterisse sisse ja jääda sõpradeta. Hirmuobjekti seista või istuda avaliku ruumi keskel korrelatsioon püsiärevusega oli võrreldes teiste hirmuobjektidega kõrgeim, seisundiärevuse puhul ta kõrgeimate hulka ei kuulunud.

Korrelatsioon on 0,692 – ümardatult 0,7. Tulemus tõestab alternatiivset ja ühtlasti uurija poolt enne uuringu teostamist püstitatud hüpoteesi- meste haridustaseme ja tapeedi vahel on statistiliselt oluline seos.

<http://fideelia.future.ee/kunst/Meeste%20haridustaseme%20seos%20tapeediga.pdf>