



# Vivika Halapuu

# PIAAC uuringu analüütik

# PIAAC uuringu eesmärk

- Koguda infot 16–65-aastaste inimeste infotöötlusoskuste kohta selliselt, et oleks võimalik teha järeldusi uuringus osalenud riikide elanike oskuste kohta.
- Järeldused puhul pakuvad seejuures huvi **populatsiooni, mitte indiviidi** tasandil.

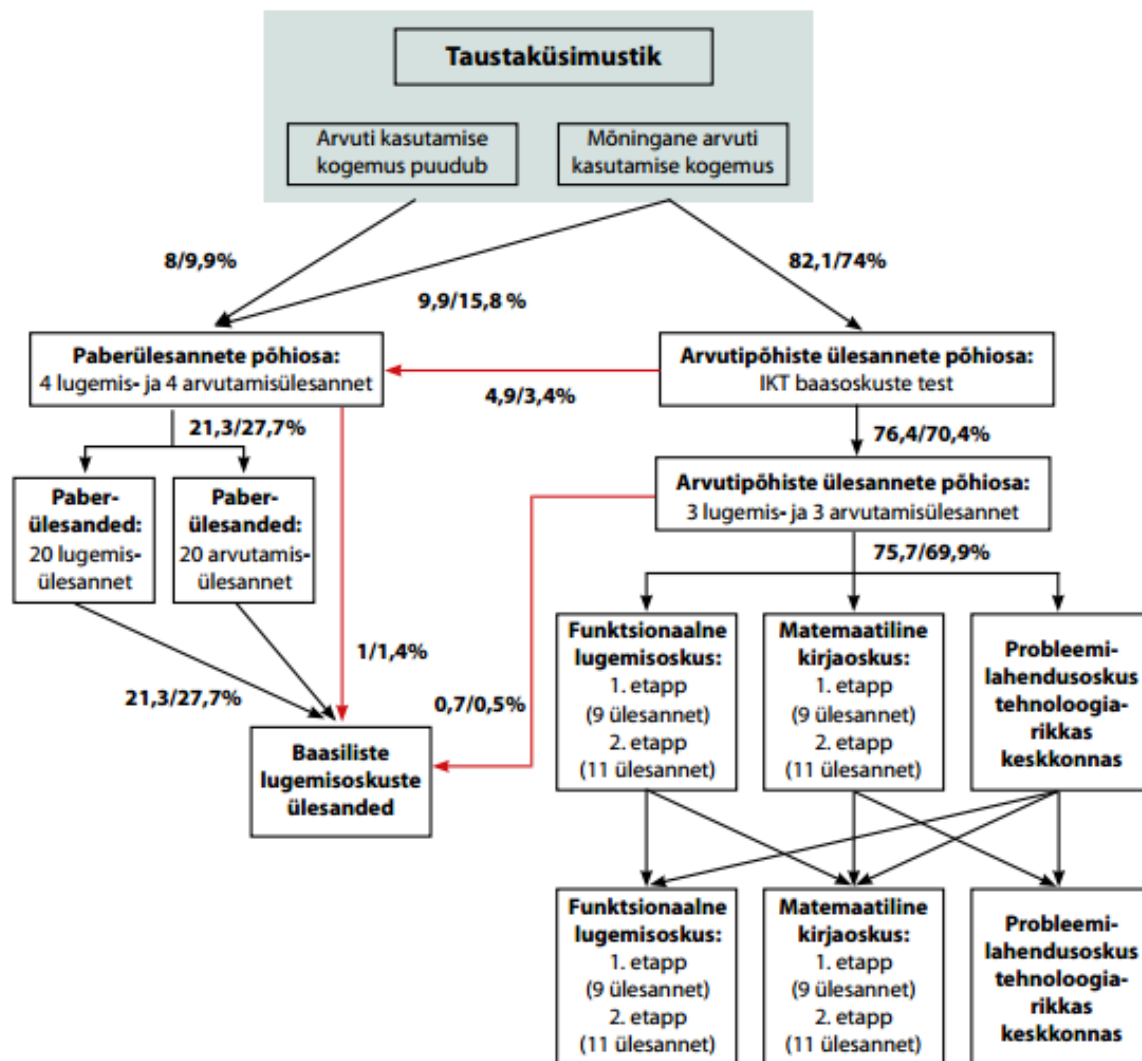
# Kõik ei vastanud kõigele

Nii taustaküsimustiku kui ülesannete lahendamise osas suunamised!

## Suunamised taustaküsimustikus:

- D blokile vastasid üksnes uuringu hetkel hõivatud
- E blokile need, kes olid olnud hõivatud viimase 5 aasta jooksul
- F ja G bloki küsimusi (töökohal kasutatavad oskused) küsiti ainult nendelt, kes olid olnud hõivatud 12 kuu jooksul enne uuringu toimumist

# Erinevad vastamisteed ülesannete lahendamise osas



# Miks kõik kõigele ei vasta?

- PIAAC uuringus mõõdeti kolme erinevat domeeni, lisaks baasilised lugemisoskused (58 FL ülesannet, 56 MK ja 14 PL ülesannet)

Keskmine ajakulu minutites:

	Funktsionaalne lugemisoskus	Matemaatiline kirjaoskus	Probleemilahendusoskus tehnoloogiarikas keskkonnas
Ülesanne	1	1,04	2,69
Moodul	19,9	20,8	18,8

- Lisaks mahukas taustaküsimustik
- Kõikidelt inimestelt kõige küsimine võtnuks liiga kaua aega

# PIAAC uuringu disain

- PIAACi disaini puhul on tegu mitmese maatriksdisainiga (*multiple matrix design*): populatsiooni esindavad teatud inimesed, kellest igaüks lahendab teatud osa ülesannete kogumist.
- Ülesanded, mida inimene lahendab, sõltuvad tema oskuste tasemest (aga ka vastustest teatud taustaküsimustiku küsimustele)

Kõik ei lahenda igat liiki ülesandeid. Küsimus, kuidas saada kõigile kõigi oskuste skoorid?

Vastus – IRT

# Üksikvastuste teooria (Item Response Theory – IRT)

IRT annab võimaluse hinnata mudeleid, mis seovad omavahel jälgitava vastamismustri ( $x_i$ ) ja mittejälgitava alustunnuse (võimekuse,  $\Theta$ ) ning ülesannete parameetrid (raskusaste  $\beta$ , diskrimineerimisparameeter  $\alpha$  jne).

$$P(x_i = k | \theta, \underline{\beta}), \begin{cases} k = 0 \text{ or } 1 \text{ for dichotomous items} \\ k = 0, 1, \dots, K \text{ for polytomous items} \end{cases}$$

# IRT peamised lähtekohad

- Inimese vastamismuster sõltub nii tema enda vahetult jälgitamatust võimekusest kui ülesannete parameetritest (eelkõige raskusaste).
- Kui inimesed lahendavad ülesandeid selliselt, et lahendatud ülesannetes eksisteerivad teatud kattuvused, võimaldab IRT erinevate inimeste vastustest tulenevat infot n.ö summeerida.
- Oluline on, et ülesanded, mida inimesed lahendavad, mõõdaks ühte ja sama nähtust (ühedimensionaalsus).



# IRT peamised lähtekohad

- Ülesannete raskusaste on defineeritud tõenäosuse kaudu, millega teatud võimekusega inimene on võimeline ülesannet korrektselt lahendama.
- Inimese võimekus on defineeritud tõenäosuse kaudu, et teatud raskusastmega ülesandele antakse õige vastus.
- Kui skaala on valitud, on võimalik välja arvutada seos võimekuse ja ülesannete raskusastme vahel.

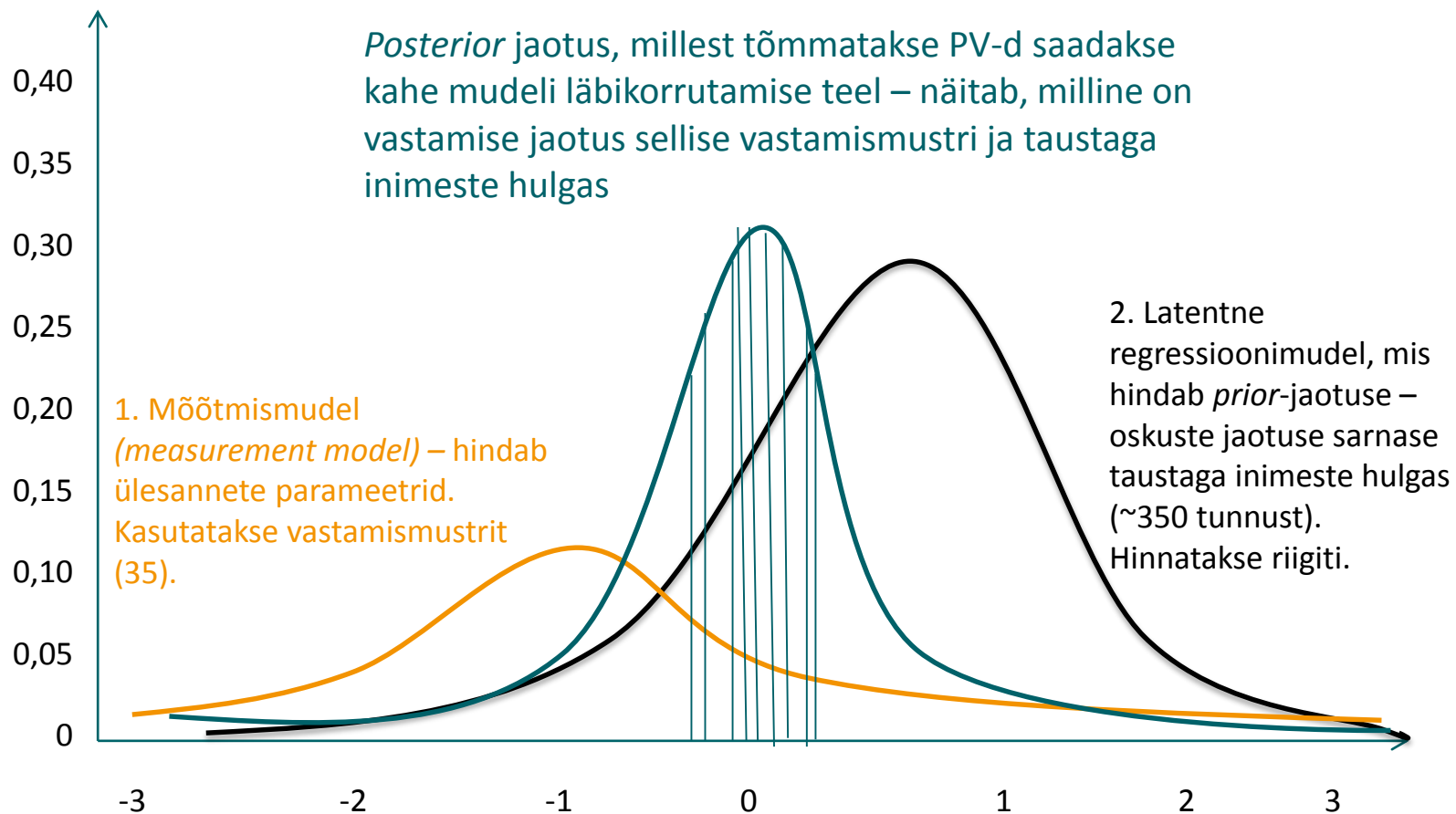
# Parameetrite hindamine (1)

- IRT puhul kasutatakse enamasti Raschi mudelit (PIAACis 2PL mudel ja *partial credit model*)
- Esimese hooga ei ole ei ülesannete parameetreid ega võimekuse kohta käivat infot. Need tuleb andmete põhjal hinnata.
- Enamlevinud meetod parameetrite hindamiseks on suurima tõepära meetod, mida praktikas rakendatakse läbi rohkearvuliste iteratsioonide.

# Parameetrite hindamine (2)

- Ülesannete parameetrite hindamisel maksimeeritakse tõenäosusjaotust, mis vastaks parimal viisil tegelikule vastamismustrile. Edaspidi käsitletakse parameetreid fikseeritud suurustena.
- Võimekuse jaotuse hindamisel kasutatakse *Expectations-Maximization* (E-M) algoritmi.

# 2-tasandiline IRT mudel



# Tuletatud tunnused ehk *plausible value'd*

- Nii funktsionaalse lugemisoskuse, matemaatilise kirjaoskuse kui tehnoloogiarikas keskkonnas probleemilahenduskuse tulemusi esitatakse skaalal 0...500
- Vastajad paigutuvad 500-punktilisele skaalale oma võimekuse, ülesanded oma raskusastme põhjal
- Iga domeeni kohta esitatud 10 tuletatud väärtust
- *Plausible value'd* on mõeldud gruppide ja mitte indiviidide kohta järelduste tegemiseks!

# Plausible value'de kasutamine arvutustes

- Korrektsete tulemuste saamiseks tuleks teha kõik arvutused kõigi replikatsioonikaalude ja kõigi PVdega ning võtta tulemustest keskmine.
- Kindlasti ei tohi kasutada PVde keskmist, kuna sel juhul luuakse ideaaljuht, mis ei ole reaalne.
- Lihtsustatud juhtude puhul kasutada pigem ühte PVd ja arvestada sellega, et standardvigu tõenäoliselt alahinnatakse.

# Oskuste tasemed

Funktsionaalne lugemisoskus/ matemaatiline kirjaoskus		Probleemilahendusoskus tehnoloogiarikkas keskkonnas	
0-175	Alla 1. taset	0-240	Alla 1. taset
176-225	1. tase	241-290	1. tase
226-275	2. tase	291-340	2. tase
276-325	3. tase	341-500	3. tase
326-375	4. tase		
376-500	5. tase		

# Plausible value'dest ja valimi üldisest iseloomust tulenev mõju andmeanalüüsile

- Kuna PV-d on juhuslikud suurused *posterior* jaotusest, käib nendega kaasas täiendava vea komponent.
- Valimi võtmisest tingitud vea arvutamiseks peame kasutama nii populatsioonikaalu kui replikatsioonikaale (kuna tegu pole juhuvalimiga, vaid kihistatud juhuvalimiga)
- Viga = valimi võtmisest tingitud viga (valimiviga) + andmete imputeerimisest tingitud viga (mõõtmisviga)

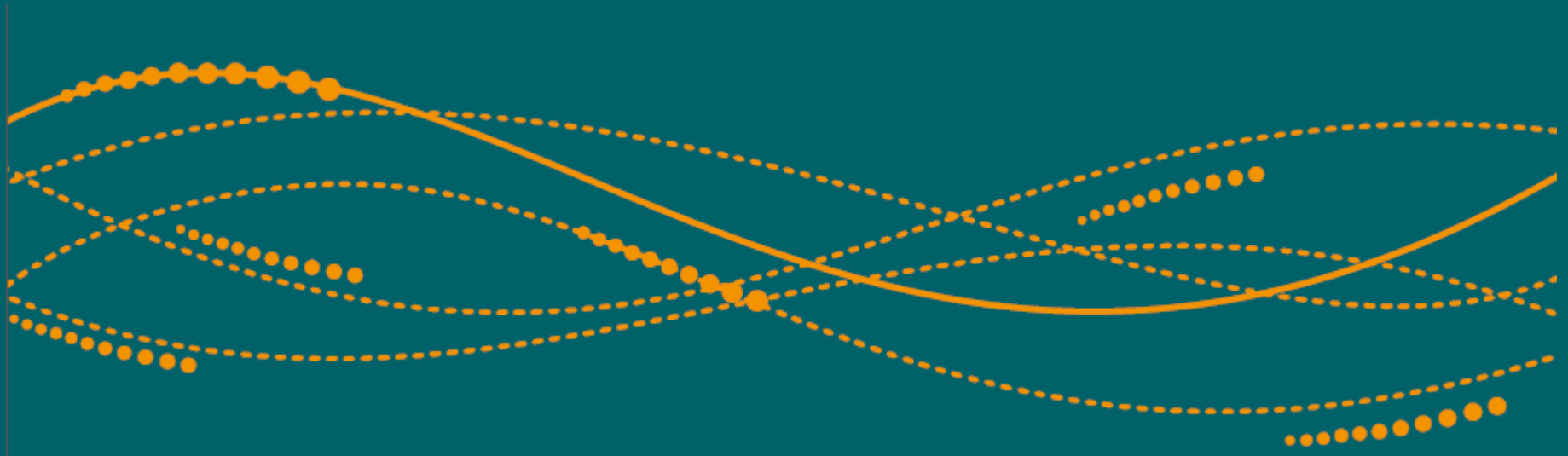


# Kas tavapärasel viisil andmete analüüsile lähenemine viiks valede järeldusteni?

	Mean	Standard Error	
		SRS	SMP + IMP
Overall Mathematics Score	491	1.5	5.2
Mathematics Score for Girls	495	2.0	5.5
Mathematics Score for Boys	487	2.1	7.6
Difference Between Girls & Boys	8	2.9	8.3

Märkus: tabelis Uus-Meremaa 1999. aasta TIMSS uuringu andmed E. Gonzaleze slaididelt

- PV-de kasutamisest tingitud vea ignoreerimise tulemusena alahindaksime kogu viga.
- Tulemuseks võivad olla valed järeldused nt erinevuste olulisuse kohta



Aitäh kuulamast!

[vivika.halapuu@hm.ee](mailto:vivika.halapuu@hm.ee)