

Mihajlo Mijanović  
Filozofski fakultet  
Nikšić

## MJERENJE I TESTIRANJE, SKALIRANJE I NORMIRANJE

U samom naslovu navedena su četiri pojma koji su metodološki u kazualnosti, a etimološki, epistemološki, suštinski i empirijski se bitno razlikuju. Navedeni pojmovi, kao i mnogi drugi su višeznačni. Značenje navedenih pojmova zavisi ko ih i u kom kontekstu upotrebljava. Ovom prilikom neće biti detaljnog i posebnog osvrta na same pojmove, mada su oni bitni zbog lakše, preciznije i bolje komunikacije.

Svi oni koji se bave antropološkim, posebno sociološkim, empirijskim istraživanjima pouzdano znaju da nema normiranja bez skaliranja, nema skaliranja bez testiranja, nema testiranja bez mjerenja.

Mjerenje je teško definisati ako se ne odnosi na mjerenje neke fizičke karakteristike ili obilježja. Postoji više definicija mjerenja pa je teško odabrati pravu ili najbolju. *Između ostalih, moglo bi se reći da je mjerenje poređenje nekog svojstva sa nekim dijelom ili veličinom tog svojstva koje je precizno definisano.*

Istorija mjerenja je veoma duga. Prvi materijalni dokazi o postupku mjerenja odnose se na mjerenje dužine, težine, brzine ne samo čovjeka. Poznate su prastare mjere za dužinu kao što su: palac, pedalj, lakat, korak itd. Primjer jeste egipatski lakat, mjera koja datira 3000 godina prije Isusa. To je mjera koja definiše dužinu ruke od lakta do vrha prstiju. Pošto ta antropometrijska dimenzija varira između 43 i 53 cm, carskim ukazom napravljen je standard, tj. *mjerni štap* koji je vrijedio u čitavom Carstvu. Evidentna su svakodnevna mjerenja, posebno je prisutno mjerenje dužine i težine u trgovini, građevini, industriji itd. Nije teško zaključiti da je Keopsova piramida, i ostala prastara zdanja, sagrađena i opstala do današnjih dana zahvaljujući preciznim mjerenjima, testiranjima, statičkim i dinamičkim proračunima. Neobjašnjivo ali istinito, te stare civilizacije poznavale su odlično mjerenje, raspolagale su sa nepoznatim mjernim instrumentima, posebno u građevini, orijentaciji na zemlji, moru i kosmosu. Mjerenje vremena, pritiska, sile, brzine, čvrstine materijala, takođe je bilo izuzetno važno, praktično kao i danas – svakako, razlikovale su se jedinice mjere koje su vjerovatno bile veće. Kako tada tako i danas, na mjerenje i testiranje se oslanjamo prije donošenja važnih odluka.

Teorijski pristup mjerenju i definisanju veličine kao apstraktne kategorije, odnosno broja kao načina komuniciranja, pretpostavlja se da je počelo sa Euklidom, aproksimativno u IV vijeku p. n. e. Vremenom se potreba za mjerenjem širila, tako da se mjerenje definisalo na apstraktan, principijelan i univerzalan

način. Mjerenje je nemoguće zamisliti bez pojave koja prati svako mjerenje, a to je greška mjerenja. Mjerenje – testiranje, skaliranje – normiranje jeste sastavni dio empirijskih nauka. Dakle, riječ je o jednoj širokoj i važnoj oblasti koja se oslanja na matematiku, fiziku, formalnu logiku, teoriju brojeva, vjerovatnoću sa jedne strane, te na prirodne zakone i objekte, sa druge strane. Sa matematičke tačke gledišta, teorija nije valjana ako nije aksiomatizovana. Dobra teorija mora da bude izvedena iz formalnih aksioma deduktivnog tipa, gdje se predviđena svojstva utvrđuju induktivnim putem. Ovakav pristup je pravilo svih formalnih sistema i ne podrazumijeva da je teorijski model tačan. Tačnost modela provjerava se kroz analizu unutrašnje usaglašenosti i kroz analizu usaglašenosti sa realnim svijetom. Usaglašenost sa realnim svijetom jeste hipotetska, a provjerava se pomoću podataka dobivenih mjerenjem. Moderna teorija mjerenja datira od kraja XIX vijeka, a vezana je za ime Normana Cempbella, fizičara, i Smith Stevensa, psihologa, preciznije psihofizičara.

Prema Cempbellu mjerenje je pridavanje brojki pojavama ili svojstvima objekata na osnovu određenih pravila. Pojam broj je višeznačan, u navedenom kontekstu pojmovno „*broj*“ odnosi se na kvantitativan opis objekata. Iz navedene konstatacije implicitno slijedi da nema mjerenja ako nema kvantitativnih pokazatelja. Jedna od definicija mjerenja koja datira od psihometričara glasi: ***Mjerenje je utvrđivanje individualnih razlika.*** Ovakva definicija primjenjena je problemima empirijskih istraživanja u sociologiji, psihologiji, kineziologiji, medicini i ostalim antropološkim naukama. Svakako, postoji evidentna razlika između mjerenja fizičkih obilježja i socioloških i psiholoških osobina. Fizičko obilježje neke pojave stabilnije je, postojanije, sa manje variranja, za razliku od socioloških i psihičkih osobina, koje su podložne velikom variranju od mjerenja do mjerenja. Ako se mjeri težina i visina čovjeka više puta u toku jednog dana, kvantitativne vrijednosti daleko su homogenije nego ako se mjeri neka psihološka osobina. Evidentno je da prethodni psihotest utiče na naredni. Iz navedenih razloga vrlo su bitna pravila mjerenja, ne samo kada je riječ o psihotestovima, već svim testovima i svim mjerenjima. Osnovno pravilo jeste da postoje tačna pravila mjerenja koja se moraju striktno poštovati kod svih mjerenja. Mjerilac ne može biti bilo ko, mjerioci jesu stručna, specijalno obučena lica, koja su upućena u detalje, pravila i proceduru mjerenja, kojih se strogo pridržavaju. Bez obzira na rutinu poznavanja procedure mjerenja, mjerioci se moraju sa njom upoznati prije pristupanja mjerenju. Posebnu pažnju mjerioci obraćaju na redovno baždarenje fizičkih mjernih instrumenata, tj. na njihovu ispravnost u pogledu valjanosti, pouzdanost i osjetljivosti.

### Nivoi mjerenja

Mjerenje fizičkih obilježja odnosi se na ***neposredno*** mjerenje, dok se mjerenje socioloških, psiholoških i drugih osobina i sposobnosti realizuje putem raznih testova koji su ***posredni*** pokazatelji osobine, svojstva ili sposobnosti.

Posredna mjerenja svode se na testove koji mjere latentne osobine. Latentnu osobinu generišu manifestne vrijednosti dobivene putem testa. Poseban

problem mjerenja predstavljaju upravo latentne osobine. Do njih se ne dolazi lako, proceduralno je riječ o više složenih operacija, logičke, tehničke i matematičko-statističke prirode. N. Cempbell je naveo devet karakteristika koje su preslikane u brojeve pomoću kojih se opisuju osobine ili svojstva. S obzirom na to da su sociološke i psihološke karakteristike generalno latentne, do njih se dolazi putem posrednih manifestnih testova. Zbog pomenutog, Cempbellovo preslikavanje u navedenim i mnogim drugim slučajevima ima neprocjenljivu vrijednost.

U sljedećoj tabeli prikazano je Cempbellovo preslikavanje devet osobina u brojeve.

### Cempbellovo preslikavanje

1.	Ili je	$a = b$ ili $a \neq b$
2.	Ako je	$a = b$ tada je i $b = a$
3.	Ako je	$a = b$ a $b = c$ tada je i $a = c$
4.	Ako je	$a > b$ tada je $b < a$
5.	Ako je	$a > b$ a $b > c$ tada je $a > c$
6.	Ako je	$a = b$ a $b > 0$ tada je $a + b > 0$
7.	Vrijedi da je	$a + b = b + a$
8.	Ako je	$a = d$ a $b = e$ tada je $a + b = d + e$
9.	Vrijedi da je	$(a + b) + c = a + (b + c)$

- Prva tri preslikavanja ukazuju da svaka brojka ima svoj identitet, tj. da je svojerodna. U prvom pravilu tvrdi se da brojku možemo pripisati nekoj osobini ili ne. Drugo pravilo jednakosti odnosi se na jednakost osobina bez obzira na permutaciju. Treće pravilo odnosi se na tvrdnju: ako su dvije osobine jednake trećoj, samim tim međusobno su jednake sve tri.
- Četvrto pravilo ukazuje na relaciju asimetričnosti: ako je osobina  $a$  veća od osobine  $b$ , obrnuto vrijedi da je osobina  $b$  manja od osobine  $a$ .
- Peto pravilo odnosi se na relaciju tranzitivnost – prelaznost. Ako je osobina  $a$  veća od osobine  $b$ , a osobina  $b$  veća od osobine  $c$ , slijedi da je osobina  $a$  veća od osobine  $c$ .
- Posljednja četiri pravila odnose se na brojeve koji imaju značenje količine.
- Šesto pravilo ukazuje na to da se dodavanjem nule na neku količinu, prvobitna količina ne mijenja.
- Sedmo pravilo govori da redosljed sumiranja nije bitan, vrijedi pravilo permutacije.
- Osmo pravilo odnosi se na to da više identičnih osobina pri sumiranju mogu zamijeniti jedna drugu.
- Deveto pravilo odnosi se na pravilo asocijacije, tj. da redosljed sumiranja ne utiče na konačnu sumu.

Na bazi Cembellovog preslikavanja osobina u brojke, američki psihometričar **S. Stevens** je mjerenje osobina preveo na četiri mjerna nivoa. Navedeni mjerni nivoi posebno su bitni zbog dalje statističke obrade, tj. kvantifikacije. U svakom slučaju treba voditi računa koje su statističke procedure primjerene datom nivou. Ovom prilikom neće biti eksplicirani Stevensovi nivoi koji se često u literaturi imenuju kao skale, sem da je riječ o četiri standardna nivoa i to: **Nominalni nivo, Ordinalni nivo, Intervalni nivo, Razmjerni nivo.**

Kada je riječ o primjeni statističkih testova i metoda, treba istaći da je nominalni nivo vrlo ograničen izborom statističkih metoda, za razliku od razmjernog nivoa, gdje u teorijskom smislu ne postoji ograničenje.

### Šta je test

Semantika testa razlikuje se u zavisnosti od toga koju karakteristiku, obilježje, sposobnost testom mjerimo. Sve definicije koje se ne mogu svesti na matematičku simboliku – varijabilne, kvalitativne, proizvoljne, spekulativne – jesu nepotpune.

- Test je objektivni postupak za izazivanje određene reakcije i utvrđivanje nivoa te reakcije.
- Test je sredstvo za utvrđivanje individualnih razlika.
- Test je ustaljeni standardizovani postupak za izazivanje određene aktivnosti itd.

Kad je riječ o testovima koji se primjenjuju u empirijskim sociološkim istraživanjima, pomoću testa se uglavnom indirektno, tj. posredno dolazi do željenih vrijednosti. Rezultati na samom testu jesu uistinu manifestni, npr. broj članova porodice, starosna struktura, zaposlenost ili nezaposlenost, visina primanja, posjedovanje kuće ili stana itd. U navedenom primjeru riječ je o manifestnim varijablama i konkretnim vrijednostima. Vrijednosti dobivene putem testa, manje ili više uspješno, determinišu znatno složeniju dimenziju kao što je socijalni status. Dakle, putem mjerenja manifestnih varijabli posredno dolazimo do latentnih vrijednosti. Rezultat na testu nikada se ne može poistovjetiti sa dimenzijom koju taj test mjeri. Ukoliko je korelacija između manifestnih testovnih vrijednosti i dimenzije koju test mjeri veća, tim je test „**čistiji – valjaniji**“.

S obzirom na to da nema potpuno čistih testova, ne samo u sociologiji, psihologiji, antropologiji, nego ni u ostalim naukama, neminovno implicira potrebu da neku latentnu – složenu dimenziju, koju uglavnom imenujemo kao **faktor**, determinišemo uz pomoć tri ili više manifestnih varijabli, tj. testova. Svakako, bilo bi vrlo pragmatično kad bismo, npr. socijalni status, psihološko ili zdravstveno stanje mogli izmjeriti pomoću samo jedne manifestne varijable, jednog testa. Sa žaljem, tako čistih, valjanih, pouzdanih, objektivnih testova u sociološkim, medicinskim i antropološkim naukama za sada nema. Vjerovatno, nikad ih neće ni biti. Problem konstrukcije čistih testova je generalnog tipa. Posebno je pri-

sutan u antropološkim naukama. Psiholozi, odnosno psihometričari, posebno su opsjednuti konstrukcijom „čistih“ testova za mjerenje inteligencije i crta ličnosti (*kognitivnih sposobnosti i konativnih karakteristika*). Bez pretjerivanja, psihometričari su dali najveći doprinos razvoju metodologije konstrukcije testova uopšte. Metodološki pristup konstrukciji testova u sociometriji uglavnom slijedi algoritam psihometrije.

***Osnovni atributi nekog testa jesu: valjanost, pouzdanost – relijabilnost, diskriminativnost – osjetljivost i objektivnost.***

Da bi neki test imao upotrebnu vrijednost, mora ispuniti određeni nivo prema svim atributima. Procedura konstrukcije testa izuzetno je složena ne samo sa matematičko-statističkog pristupa. Upotrebna vrijednost testa moguća je nakon duge empirije, odnosno provjere u praksi. Sva naučna saznanja, pa i saznanja koja se dobivaju putem testova u sociologiji, ne vrijede ako ih empirija svakodnevno ne potvrđuje.

### **Skaliranje i normiranje testova**

Problem i postupak skaliranja i normiranja uslovljen je problemom i postupkom testiranja i mjerenja. Još prije rođenja, u utrobi majke, počinju razna testiranja, mjerenja i dijagnosticiranja. Rođenjem, tj. dolaskom na svijet, novorođenčetu se, između ostalog, mjeri visina i težina. Mjerenje i testiranje ne jenjavaju tokom života, naprotiv, ona se intenziviraju bez obzira na zanimanje i profesiju. Testiranje i mjerenje nastavljaju se nakon života, poznata su brojna mjerenja i testiranja na leševima.

Zbog uporedivosti sa drugim testovima, odnosno skorovima, originalne sirove skorove neophodno je transformisati u relativne vrijednosti. Transformacija je potrebna i u slučajevima u kojima su izvorni skorovi dati u istoj jedinici mjere, prvenstveno zbog različitih aritmetičkih sredina i standardnih devijacija.

***Transformacija sirovih skorova na zajedničku mjernu skalu naziva se skaliranje testova.***

Semantika skaliranja i normiranja testova bitno se razlikuje. Po prirodi stvari, mjerenje i testiranje prethodi skaliranju, a skaliranje prethodi normiranju. Postupak mjerenja, testiranja i skaliranja jeste unaprijed poznat.

Normiranje je postupak nakon skaliranja i vrijedi da je norma promjenljiva vrijednost. Norma se utvrđuje na osnovu postignutih rezultata u nekom testu. Norma nije vrijednost kojoj treba težiti, već realna vrijednost koja je uspostavljena na osnovu ukupnih rezultata u nekom testu. Norma se dobija na uzorku i treba da bude standard za populaciju iz koje je izvučen uzorak.

▪ ***Uzorak na kome je napravljen standard naziva se normirani uzorak, a standard koji je dobiven na takvom testu naziva se normirani test.***

Normirani uzorak mora da ispunjava uslove kako bi imao attribute normativnog uzorka. Ako je riječ o uzorku stanovnika nekog planinskog sela, stanovnici moraju da pripadaju određenom planinskom selu, tj. prostoru (*državi, regiji, oblasti*), moraju da pripadaju određenom polu, određenom uzrastu (*sta-*

*rosnoj grupi*), naobrazbi, zanimanju itd. Ovo su osnovna relevantna obilježja kod uspostavljanja norme uzorka. Relevantna obilježja vrlo su različita, nekada je bitan pol, nekada su bitne godine starosti, nekada naobrazba, zanimanje, pripadnost državi, regiji, kantonu, opštini, nekada je bitno sve skupa.

Normativi testa prave se na osnovu normativnog uzorka, a normativni uzorak definiše se na osnovu relevantnih obilježja. Vrijednost normativnog testa nije univerzalni standard kome treba težiti, to je realna vrijednost koja vrijedi za normativni uzorak i populaciju iz koje je uzorak izvučen.

▪ **Skala nastaje transformacijom originalnih skorova i ona je u biti norma testa.**

Normiranjem testa moguće je izvršiti izvjesne korekcije originalnih vrijednosti. Te korekcije odnose se na prilagođavanje originalnih vrijednosti normalnoj raspodjeli. Dakle, ako originalne vrijednosti značajno odstupaju od normalne raspodjele, statističkim postupcima moguće je izvršiti njihovo prilagođavanje normalnoj raspodjeli.

Ovaj postupak naziva se **normalizacija originalnih vrijednosti** ili prilagođavanje originalnih vrijednosti normalnoj raspodjeli.

Dakle, treba razlikovati normiranje testa od normalizacije testovnih vrijednosti.

– **Normiranje testa odnosi se na transformaciju originalnih vrijednosti na neku skalu.**

– **Normalizacija testa odnosi se na prilagođavanje vrijednosti testa normalnoj raspodjeli.**

Postupkom normiranja, tj. zamjenom originalnih vrijednosti novim vrijednostima, definiše se ukupni raspon variranja i mogućnost transformacije negativnih skorova u pozitivne, te transformacija decimalnih vrijednosti u cijele brojeve.

Postupkom normalizacije vrši se ipak izvjesno narušavanje odnosa između originalnih vrijednosti i novih normaliziranih vrijednosti. Ako se radi samo o normiranju testa, dakle, bez normalizacije testa, odnosi sirovih i transformisanih skorova ne mijenjaju se.

### Vrste skala

U današnje vrijeme postoji više skala koje se primjenjuju u zavisnosti od vrste mjerenih varijabli i dobivenih skorova. Razvojem empirijskih nauka pomoću metodologije, statistike mjerenja i testiranja, sociometrije, psihometrije i kineziometrije, razvijalo se skaliranje i normiranje.

U današnje vrijeme najčešće su u upotrebi: **Percentilna skala, Z-skala, T-skala, C-skala, S-skala, IQ-skala, Likertova skala.**

▪ **Percentilna skala** jeste relativna skala, njen raspon je od 1 do 100. Osnovni nedostatak Percentilne skale jeste u tome što se sirovi podaci raspoređuju na jednake dijelove duž cijele skale, dakle u svakoj grupi – klasi ili razredu ima isti

broj slučajeva. Veliki broj varijabli nema linearnu već normalnu raspodjelu, tako da ova skala ne odražava realno stanje. Kod varijabli sa normalnom raspodjelom evidentna je koncentracija podataka oko aritmetičke sredine, sa udaljavanjem od centra koncentracija postepeno opada. Dakle, normalna raspodjela podrazumijeva najviše slučajeva u centru, a najmanje ekstremno visokih i ekstremno niskih vrijednosti. Percentilna skala podrazumijeva linearnan potpuno isti broj slučajeva duž cijele skale.

▪ **Z-skala** otklanja osnovni nedostatak Percentilne skale. Postupkom standardizacije originalni skorovi transformišu se u relativne, gdje se aritmetička sredina centrirana na nulu, a standardna devijacija normira se na jedinicu. **Z-skala** je univerzalna i smatra se osnovom svih skala.

– Obrazac sa standardizaciju sirovih skorova jeste:  $Z_{(xi)} = (x_i - \mu) \sigma^{-1}$

– Dokaz da je aritmetička sredina 0:

$$\mu_{(z)} = 1/n \sum [(x_i - \mu) \sigma^{-1}] = 1/n \cdot 0/\sigma = 0/\mu\sigma = 0$$

– Dokaz da je standardna devijacija jednaka 1:

$$\sigma_{(z)} = [1/n \sum (x_i - \mu)^2 \sigma^{-2}]^{1/2} = [\sum (x_i - \mu)^2 (n\sigma^2)^{-1}]^{1/2} = [\sigma^2/\sigma^2]^{1/2} = 1^{1/2} = 1$$

Svi standardizirani skorovi veći od nule veći su od aritmetičke sredine; podrazumijeva se da je aritmetička sredina nula, odnosno standardizirani skorovi sa negativnim predznakom manji su od aritmetičke sredine. Suma svih standardiziranih skorova je uvijek nula. Transformisani standardizirani skorovi i sirovi skorovi zadržavaju linearnan odnos, prema tome vrijedi pravilo da je korelacija potpuna i iznosi 1. Slijedi pravilo da je distribucija sirovih skorova i standardiziranih skorova potpuno identična. Standardizirani skorovi ispunjavaju intervalni nivo odnosa, dakle moguće je primijeniti sve aritmetičke operacije, a odnosi između skorova neće se promijeniti, što je bio osnovni cilj standardizacije.

U skladu sa Gaussovom raspodjelom, standardizirani skorovi kreću se u rasponu od -3 do 3. Vjerovatnoća da će se slučajno promjenljiva  $Z_i$  naći u rasponu od -3 do 3, prema Gaussovoj raspodjeli jeste 0,9973, približno 1. Već je istaknuto da je **Z-skala** univerzalna, te da je osnova svih skala, što je sasvim tačno. **T-skala**, **C-skala**, **S-skala**, **IQ-skala**, izvedene su iz **Z-skale**. Modifikacija se odnosi na dodavanje konstanti na aritmetičku sredinu koja iznosi nula i standardnu devijaciju koja iznosi jedan. Svrha modifikacije jeste izbjegavanje decimalnih vrijednosti i negativnih predznaka, odnosno približavanje skaliranih vrijednosti široj populaciji.

### Umjesto sažetka

Saznanje realnog svijeta moguće je zahvaljujući čulima, razumu i empiriji. Od nivoa razvijenosti čula, razuma i empirije zavisi kvantitet i kvalitet saznanja. Ipak, kada treba donijeti važnu, bitnu i odgovornu odluku, pored čula, zdravog razuma i empirije, po mogućnosti, oslanjamo se uvijek na rezultate mjerenja i testiranja, koji se prema potrebi skaliraju i normiraju.

### L i t e r a t u r a

- Bukvić, A. (1996), *Načela izrade psiholoških testova*, Zavod za udžbenike i nastavna sredstva, Beograd.
- Fajgelj, S. (2003), *Psihometrija (Metod i teorija psihološkog mjerenja)*, Centar za primenjenu psihologiju, Beograd.
- Gilford, J. P. (1968), *Osnovi psihološke i pedagoške statistike*, Savremena administracija, Beograd.
- Mijanović, M. (2002), „Valjanost (Validnost) testiranja i mjerenja“, *Savremeni sport*, god. I, br. 1–2, Banja Luka.
- Mijanović, M., M. Vojvodić (2005), „Kauzalnost među pojavama“, *Glasnik Fakulteta fizičkog vaspitanja i sporta*, br. 1.
- Mijanović, M. (2005), *Metrijske karakteristike mjernih instrumenata. Nove tehnologije u sportu (NTS)*, Međunarodni naučni skup, Sarajevo.
- Mijanović, M. (2005), „Diskriminativnost (Osjetljivost testova)“, *Savremeni sport*, god. IV, br. 5–6, Banja Luka.