

## Mjerenje u psihologiji

### Što je psihologijsko mjerenje?

- mjerenje psihičkih fenomena
- psihologijsko mjerenje je vrlo...
  - delikatno – kako izmjeriti 'psihičko'?
  - odgovorno ☐ donošenje važnih odluka
  - često, potrebno svakom psihologu

### Koje psihičke varijable mjerimo?

- inteligencija – verbalna, numerička, prostorna, opća (*g*)...
- ličnost – ekstraverzija, neuroticizam, psihoticizam, savjesnost, ugodnost...
- interesi
- vrijednosti
- stavovi, mišljenja
- emocije
- mentalno zdravlje, psihopatološko ponašanje

### Kako ih mjerimo?

- testovima
- intervjuom, fokus grupama
- portfolijima (zbirkama radova)
- metodom analize slučaja (*case study*)
- opažanjem ponašanja
- metodom igranja uloga

### Zbog čega ih mjerimo?

- u svrhu provjere teorije i razvijanja opće spoznaje o funkcioniranju čovjeka (znanstvena psihologija)
- u praktične svrhe (primijenjena psihologija)
  - odabir zaposlenika
  - zrelost za školu
  - klinička psihodijagnostika
  - status vojnih invalida
  - prijemni ispiti
  - predizborna istraživanja
  - profesionalna orijentacija ...

### Cattell, 1890

- "Psihologija ne može zadržati jasnoću i direktnost fizikalnih znanosti, osim ako se ne osloni na eksperiment i mjerenje. Korak u tom smjeru može biti primjena mentalnih testova i mjerenja na većem broju pojedinaca. Rezultati bi bili znanstveno vrlo važni i omogućili bi otkrivanje mentalnih procesa, njihove konstantnosti, međuzavisnosti, i promjene pod utjecajem raznih okolnosti. S druge strane, pojedinci bi takve testove smatrali zanimljivim i možda korisnim za bolje upoznavanje sebe ili kao indikaciju mogućih bolesti.
- Znanstvena i praktična vrijednosti ovakvih testova bila bi još veća kada bi se razvili ujednačeni sistemi, koji bi omogućili da se zaključci prikupljeni u različitim vremenima i mjestima mogu uspoređivati i kombinirati."

### Ψ vs. Φ

- što mjerimo metrom?
- ima li tu neke dvojbe?
- hoće li se nešto promijeniti...
  - ... ako ponovimo mjerenje?
  - ... ako mjerenje ponovi druga osoba?
  - ... ako uzmemo drugi metar?
- ... 'fizičko' je relativno jasno
- ... a 'psihičko'??

### Prema tome...

- na ovom kolegiju naučit ćete što je to psihologijski test, a što nije;
- upoznat ćete principe psihologijskog testiranja;
- upoznat ćete osobine koje se testovima mjere;
- kao i neke psihologijske testove koji se koriste u praksi...

### ...teme...

#### PSIHOLOGIJSKI TEST (7 predavanja)

- što je to psihologijski test? što ga određuje?
- kakvi mogu biti zadaci u testovima? što čini rezultat u testu?
- koje su osnovne karakteristike Ψ testova?
- kada primjenjujemo testove? & kako?

#### MJERENJE POJEDINIH OSOBINA (4 predavanja)

- što sve mjerimo psihologijskim testovima?
- kako mjerimo inteligenciju, ličnost, interese, vrijednosti, stavove?
- primjeri testova

### Literatura

Uglavnom nema knjiga na HR, ali ima mnogo dobre literature na ENG

- Anastasi, A. (1976). *Psychological Testing*. NY: Macmillan Publishing CO., Inc.
- Nunnally, J.C. (1959). *Tests and Measurements: Assessment and Prediction*. NY: McGraw-Hill Book Company.
- Cohen, R.J. & Swerdlik, M.E. (2002). *Psychological testing and assessment*. Boston: McGraw-Hill.
- Jackson, C. (2000). *Psihologijsko testiranje*. Jastrebarsko: Naklada Slap.

Mnogo informacija i na netu – pazite na kvalitetu izvora

## POVIJEST PSIHOLOGIJSKOG TESTIRANJA

- Astrologija, numerologija, grafologija, frenologija, čitanje iz karata, kave, graha...
- 20. stoljeće – razvoj mjerenja zahvaljujući...
  - Razvoju teorija ljudskog ponašanja
  - Razvoju statističkih metoda
  - Primjeni psihologije u sve širem polju djelatnosti
  - Većoj brizi za osobe sa problemima u razvoju

### Zašto nam je važna povijesna perspektiva?

- Genijalci
- Utjecaj na današnje stanje
- Učenje: prednosti i nedostaci; izbjegavamo greške

### Antički počeci

- Kina 2200 pr.n.e.
  - Prvi oblici testiranja
  - Svake tri godine provjera “spremnosti za javnu službu”
  - Iscrpljujuća višednevna mjerenja
  - Uspješnost u baratanju lukom i strijelom, jahanju konja, pisanju, računanju, obnašanju ceremonija i rituala, poznavanju zakona i oporezivanju, vojnim vještinama, poljoprivredi & zemljopisu, muzici

### Crni srednji vijek

- Srednji vijek: 5 st (propast Antičkih civilizacija) do kraj 15.st (Kolumbovo otkriće Amerike 1492.)
- Progon vještica
  - intelektualno nasljeđe srednjeg vijeka
  - kraj 15. st te 16. i 17. st
  - 300.000 žena (!!)
  - 9/10 u protestantskim zemljama
- Pitanja od javnog značaja:
  - Tko je vještica?
  - Tko je prodao dušu vragu?
- Indikatori
  - postojanje madeža,
  - crvene kose,
  - ‘čudno’ ponašanje...

### Termin psihologija ~16. st

- paralelno sa progonom vještica
- pojava termina ‘psihologija’; autori
  - Marko Marulić (1450 - 1524)
  - Filip Melanchthon (1497 – 1560)
  - Rudolf Gockel (1547 - 1628 )
  - Ollo Casmann (1562 - 1607)
- Marko Marulić
  - Prvo spominjanje termina *psihologija* u književnim izvorima

- Psihologija je “znanost o duši”
- Djelo *Psihologija o temelju ljudske duše*, ostano zapisan samo naslov, spominje ga Krstić, 1964 u *Acta Instituti Psychologici Universitatis Zagrebensis*

### Psihologija znanost ~18. st.

- Christian von Wolff
  - Definirao psihologiju kao znanost
  - Psihologijsko mjerenje – njena specijalnost
  - Djela:
- *Psychologia empirica* (1732)
- *Psychologia rationalis* (1734)

### Pojava moderne znanosti 19. st

- tehnološka otkrića, brzi razvoj prirodnih znanosti
- znanost pomalo počela poprimati obilježja znanosti 20-og stoljeća
- psihologiji naravno najbliža biologija
- Znanstvenici najzaslužniji za postavljanje temelja znanstvene psihologije → prirodnjaci Charles Darwin i Francis Galton

### Charles Darwin (1809-1882)

- Engleski prirodnjak
- Istraživačko putovanje - brod Beagle, 5 g., 1831-1836
  - prva sistematična znanstvena zapažanja
  - raznolikost – geologija, fosili, živi org. – otočje Galapagos
- Teorija evolucije  
*Porijeklo vrsta* - 1859.
  - Sva su se živa bića oblikovala od davnog zajedničkog pretka procesom prirodne selekcije kroz dugi vremenski period
  - Prirodna selekcija → razvoj osobina korisnih za opstanak vrste, postupno gubljenje manje korisnih ili nepovoljnih
  - I čovjek se razvio prirodnom selekcijom, njegovo porijeklo je također od zajedničkog pretka, njegov posljednji prethodnik bio je majmun
- Važnost teorije za psihologiju
  - Istakao postojanje individualnih razlika
  - Prirodna selekcija bazira se na činjenici da postoje individualne razlike u izraženosti pojedinih osobina
  - Čovjeka direktno povezao sa životinjama
- Utjecaj teorije na psihologiju
  - potakla je niz eksperimentalnih provjera na životinjama
  - postavila pitanje kako se ljudi i životinje razlikuju po svojoj svijesti
  - pobudila je znanstveni interes za individualne razlike

### Francis Galton (1822-1911)

- Engleski prirodnjak, antropolog, statističar, psihomteričar → ‘otac moderne psihometrije’
- Proučavao genetiku
- inspiriran Darwinom, proučavao biljke, pa ljude
- 1869. knjiga *Hereditary genius* – revolucionarno
- pokušao klasificirati ljude prema njihovom *prirodnom daru* i ustanoviti kako se oni razlikuju od prosjeka

- Osnovao antropometrijski laboratorij
- Mjerenje visine, raspona ruku, težine, kapaciteta pluća, snage stiska, jasnoće vida, pamćenja oblika, razlikovanja boje, mirnoće ruke... → komercijalno & znanstveno
- Prva baterija testova, prvi znanstveni pristup mjerenju
- Prvi podaci koji ukazuju na interindividualne razlike
- Vrlo napredno!
- Važnost za psihologiju
  - Smatrao je da senzomotorne funkcije ukazuju na stupanj intelektualnog funkcioniranja
  - Razvio objektivnost mjerenja – omogućio komparaciju rezultata
  - Postavio osnovnu ideju korelacije – kasnije su ju uobličili Pearson i Spearman
  - Najvažnije: potaknuo je interes za mjerenje psiholoških konstrukta → prvi pravi začetnik psihologijskog mjerenja

#### Kraj 20.og stoljeća

- Ulazak u 'pravu psihologiju'
- Prva istraživanja psihologijskih fenomena
- Prvi eksperimentalni psiholozi
  - Wilhelm Wundt
  - James McKeen Cattell
  - Charles Spearman
  - Victor Henri
  - Emil Kraepelin
  - Hermann Ebbinghaus
- Prethodnica razvoju psihologijskih testova

#### Wilhelm Max Wundt (1832-1920)

- Nijemac, medicinar, fiziolog, sveučilišni nastavnik
- Oslanjao se na metode iz prirodnih znanosti, povezivao ljudski mozak i um
- Predložio prvi kolegij znanstvene psihologije
- Objavio prve psihološke tekstove
  - predavanja *Lectures on the Mind of Humans and Animals*, 1863.
  - *Principles of physiological psychology* 1874., istraživanje svjesnoga introspekcijom – doživljaji, osjećaji, ideje
  - *Social psychology* (10 tomova) na kraju 65-godišnje karijere
  - Objavio 54.000 stranica rukopisa (knjiga, članaka)
- Predavao psihologiju mnogim studentima koji su kasnije osnovali studije psihologije u svijetu
- 1879. prvi eksperimentalni laboratorij za psihologiju na Sveučilištu u Leipzigu
- Sustavno laboratorijsko mjerenje
  - saznati ne kako se ljudi razlikuju, već po čemu su slični, što je to što ih čini *ljudima*
  - cilj → donijeti generalne zaključke o ljudskoj vrsti, a ne proučavati interindividualne razlike
- Razlike u ljudskom funkcioniranju – smetnja, pogreška mjerenja → ukloniti
- Standardizirani postupci za smanjenje nesistematskih varijabilnih faktora i dobivanje *pravog* rezultata mjerenja
- Eksperimentalna psihologija, za razliku od psihologije interindividualnih razlika (diferencijalna psihologija)

#### James McKeen Cattell (1860-1944)

- Američki psiholog, studirao kod Wundta, upoznao Galtona u Evropi, inspirirao se njegovim poimanjem *interindividualnih razlika*

- prvi doktor psihologije u SAD, osnivač studija, predavao 26 godina, educirao brojne buduće psihologe
- osnovao časopise *Psychological Review* i *Science*, koji izlaze i danas
- 1921. osnovao udruženje istaknutih psihologa *Psychological Corporation*, cilj "razvoj psihologije i promocije njezine praktične primjene", radi i danas
- 1890. prvi put spomenuo termin "mental test"
- Po uzoru na Galtona, bavio se mjerenjima
- "Mentalni testovi i mjerenje"
  - Snaga stiska šake mjerena dinamometrom
  - Brzina naizmjeničnog micanja ruke na udaljenosti od 50 cm
  - Razlike u osjetu dodira
    - Minimalna udaljenost na kojoj se dvije točke mogu doživjeti kao različite
  - Jakost pritiska potrebna da izazove osjet boli
    - Guma pritisnuta na čelo
  - Razlikovanje težine
    - Razlikovanje težine kutija koje isto izgledaju
  - Vrijeme reakcije na zvuk
  - Vrijeme potrebno za imenovanje boja
  - Podjela linije od 50-cm na dva jednaka dijela
  - Prosudba 10 sekundi vremena
    - Broj točno ponovljenih slova nakon jednog čitanja
- Kao i Galton, smatrao je kako su to sve odrazi mentalnih osobina

#### Charles Spearman (1863-1945)

- Engleski psiholog sa ranijom tehničkom karijerom, studirao kod Wundta, bio pod velikim utjecajem Galtona
- Poznat po doprinosu u statistici:
  - Spearmanov koeficijent korelacije *rho*,
  - korekcija zbog atenuacije,
  - osnove faktorske analize
- Pionirski rad u području modela ljudskog intelekta – teorija o postojanju generalnog faktora inteligencije ili "*g*"
- 1904. članak "*General intelligence, objectively determined and measured* – faktorski pristup teoriji intelekta – empirijski
- Ideja o pouzdanosti mjerenja – što je test duži, to je pouzdaniji

#### Ostali važni...

- Victor Henri
- Suradivao sa Binetom na tekstovima o tome kako mjeriti složene mentalne procese
- Emil Kraepelin
- Psihijatar, istraživač u području asocijacija
- Uobličio asocijativne tehnike u formu pravog testa
- 1892 – Test slobodnih asocijacija – na zadanu riječ osoba odgovara prvu koju se sjeti
- 1885 – Test računanja koji je mjerio utjecaj uvježbavanja, pamćenje i otpornost na umor i distrakciju
- Hermann Ebbinghaus
- Testovi matematičkih operacija, pamćenja i čitanja za djecu školske dobi – posljednji pokazao vezu sa školskim uspjehom

## 20. stoljeće

- Početak 20. st. - početak razvoja psihologijskih testova
- Postojao je veliki interes za testove kojima bi se mogle mjeriti mentalne osobine
- Dosad su se mjerile varijable koje su se smatrale psihologijskima u svojoj prirodi, no to su bile senzomotorne sposobnosti – vrijeme reakcije i sl.
- Sada dolazi vrijeme pravog psihologijskog mjerenja – Alfred Binet je bio začetnik te ideje
- Početak sa testovima inteligencije, kasnije slijede testovi ličnosti, pa interesa, stavova, vrijednosti

### Alfred Binet (1857-1911)

- Francuski psiholog; završio prirodne znanosti na Sorbonni, psihologiju naučio iz knjiga
- Autor prvog testa inteligencije
  - 1895. objavio niz članaka o važnosti mjerenja sposobnosti sa Victorom Henrijem
  - 1904. naručen da osmisli detekciju djece koji bi mogli imati potrebe za prilagođenim školskim programom
  - 1905., u suradnji sa Theodoreom Simoneom, sastavio je prvi instrument za mjerenje inteligencije: Binet-Simoneovu skalu
- Binet-Simoneova skala – 1905. g.
  - Mjeri intelektualne funkcije djece → razmišljanje, rasuđivanje, razumijevanje
  - Standardizirani mjerni instrument: primjena & norme
  - 30 zadataka
    - uglavnom bihevioralni, senzorni i perceptivni,
    - no prvi puta su se pojavili i verbalni

### Primjeri zadataka Binet-Simoneova skala

- Očima prati objekt koji se pomiče.
- Hvata mali predmet koji mu je stavljen u ruku.
- Pokazuje imenovane objekte koje poznaje, npr., “Pokaži mi šalicu.”
- Uspoređuje dvije linije vidno nejednake duljine.
- Uspoređuje dvije težine.
- Ponavlja rečenice od 15 riječi.
- Definira razliku između dva poznata predmeta, npr., “papir i karton.”
- Odgovara na apstrakta pitanja, npr. “Kada te osoba uvrijedi, pa se ispriča, što ćeš učiniti?”
- Okreće kazaljke sata i očitava tako zamišljeno vrijeme.
- Definira apstraktne riječi određujući razlike među njima, npr. “dosada i umor.”

### Binet-Simoneova skala

- 1908. druga revizija B-S skale
  - Veći broj zadataka
  - Složeni po grupama za djecu od 3-11 godina – za svaku dob 6 zadataka
  - Zadaci odabrani tako da ih rješava dio djece (65–75% djece)
  - Rezultat djeteta je njegova mentalna dob – odgovara dobnoj razini grupe zadataka koje je dijete uspjelo riješiti
- Daljnje revizije
  - 1911. – skala proširena i na starije ispitanike
  - 1912. – Kuhlmanova revizija – proširenje na djecu do 3 mjeseca
  - 1916. – Termanova revizija: Stanford-Binet skala – Kvocijent inteligencije – omjer mentalne i kronološke dobi –  $IQ = MD/KD$
  - 2003. posljednja revizija koja se i danas koristi

- B-S skala potakla
  - opći pokret mjerenja intelektualnih sposobnosti i
  - primjenu testova u kliničkoj praksi → npr. u školama, učeničkim domovima, sudovima za maloljetnike, popravnim domovima, zatvorima...

### David Wechsler (1896-1981)

- Američki psiholog, radio kao klinički psiholog u Bellevue bolnici, NY
- Interes za područje intelekta: Spearmanov student (tradicija – g)
  - Definicija inteligencije – opća sposobnost pojedinca da u svojoj okolini djeluje svrsishodno, razmišlja racionalno i ponaša se efikasno
- Razvio testove inteligencije WISC i WAIS koji se koriste i danas
  - WAIS – Wechsler Adult Intelligence Scale – prvo izdanje 1939. (tada Wechsler Bellevue Intelligence Scale), posljednja revizija 1997.
  - WISC – Wechsler Intelligence Scale for Children, prvo izdanje 1949., posljednja revizija 2003.
- Poznat je i njegov test pamćenja, zatim test inteligencije za predškolsku djecu, te neverbalni i skraćeni testovi inteligencije

### Prvi svjetski rat: pojava grupnog testiranja

- potreba: detektirati najposposobnije pojedince za različite funkcije u vojsci
  - ispitati mnogo ljudi od jednom, ubrzati proces
  - razvoj grupnih testova inteligencije
- Američka vlada & Udruženje američkih psihologa (APA) → zadatak psihologa – razviti grupne testove inteligencije
- Pojava “pravih testova inteligencije”, u onom obliku na koji pomislite kad kažemo test inteligencije:
  - zadaci višestrukog izbora i
  - različitih sadržaja – slikovni (figuralni), numerički, verbalni

### Robert Yerkes: Army Alfa i Army Beta testovi grupni testovi inteligencije za regrutaciju

- Army Alpha: 8 verbalnih subtestova
  - (1) praćenje verbalnih uputa
  - (2) numeričko rezoniranje
  - (3) praktično rasuđivanje
  - (4) sinonimi i antonimi
  - (5) uređivanje rečenica
  - (6) završavanje numeričkih nizova
  - (7) analogije
  - (8) informacije
- Army Beta: 7 subtestova vizulane percepcije i motorike

### Utjecaj testova Army Alpha i Army Beta

- Koriste se i danas
- Služili su kao modeli za razvijanje novih grupnih testova inteligencije
- Nastupa prava produkcija (hiperprodukcija!) testova – za različite dobne skupine, za različite svrhe
- Testira se svugdje – u školama, koledžima, bolnicama, zatvorima, pri zapošljavanju
- Često loša predikcija → nepovjerenje u testove i testiranje uopće

### Problem primjene testova u različitim kulturama

- Army Alfa i Army Beta su ukazali na taj problem
- Ranije su svi testovi bili standardizirani na bijelcima

- Primjena takvih testova nije mogla na adekvatan način opisati intelektualno funkcioniranje drugih rasa
- Trebalo je napraviti testove koji su u jednakoj mjeri prilagođeni za sve rase – boom 1960-1984
- Zovemo ih Culture fair testovi – testovi u kojima sve kulture imaju jednake “šanse”
- Raymond Cattell je jedan od autora CFIT – Culture Fair Intelligence Test – kristalizirana i fluidna int.

#### Mjerenje pojedinačnih sposobnosti

- Testovi Army Alfa i Army Beta
  - zamijećene razlike između različitih subtestova
  - visok rezultat na verbalnom testu, nizak na numeričkom i sl.
- Uočeno postojanje “specijalnih” sposobnosti
- Važno za prof. orijentaciju i selekciju zaposlenika i vojnika
- Potaklo razvoj baterija čiji subtestovi mjere različite sposobnosti i organizirani su u cjeline
- Preduvjet – razvoj statističkih metoda (korelacije i FA) → upoznavanje sa sadržajem testova, odnosno sa sadržajem zadataka

#### Mjerenje pojedinačnih sposobnosti

- 1945. zahvaljujući FA – razvoj višedimenzionalnih baterija testova – umjesto jednog ukupnog rezultata (IQ), formiraju se rezultati za više različitih sposobnosti
  - verbalno razumijevanje
  - numeričko rezoniranje
  - spacijalna vizualizacija
  - matematičko razmišljanje
  - perceptivna brzina
- Primjena u 2. svj.ratu – testovi za pilote, radio operatere – za svakog su potrebne drugačije sposobnosti

#### Mjerenje postignuća

- Testovi postignuća: znanja, vještina, uspješnosti u aktivnostima
- Školska testiranja: razvijene su skale za ispitivanje kvalitete rukopisa, testovi pisanja, matematičkog računanja
- 1923. Kelly, Rush i Terman – *Stanford Achievement Test* – baterija mjera uspjeha u različitim školskim predmetima
- 1930. uočen problem sa testovima esejskog tipa
  - loše slaganje ocjenjivača
  - zahtijevaju više vremena i od ispitanika i od ocjenjivača
  - manje su pouzdani od testova objektivnog tipa
- U SAD razvoj nacionalnih programa za školsko testiranje
  - CEEB College Entrance Examination Bord
  - ETS Educational Testing Service (od 1945)
  - ACTP American Collenge Testing Program (od 1959)
- Time se mjerenje znanja približilo kvaliteti i ozbiljnosti kojom je mjerena inteligencija

#### Mjerenje vještina - deksterimetri

- okulomotorna koordinacija, spretnost i mirnoću ruke i sl.
- Minnesota deksterimetar – *Minnesota Dexterity Test*
- O'Connorov deksterimetar – O'Connor Finger / Tweezer Dexterity Test
- Test koordinacije ruku - *Two Arm Coordination Test*
- Test zrcala - *Auto-Scoring Mirror Tracer*

#### Mjerenje ličnosti

- Počeci: Kraepelin, 1892, test slobodnih asocijacija
- Potaknuto izradom testova inteligencije
  - Već 8. godina nakon Binet-Simonove skale (1905.) psihologiji se predbacivalo da se previše oslanja na testove
  - Do kraja 30-ih u optičaju je bilo ~4000 testova → slična stvar se počela dešavati i u kliničkoj Ψ
- Boom → 1. svjetski rat → mogućnost emocionalne prilagodbe i emocionalna stabilnost vojnika

#### Robert S. Woodworth (1869-1962)

- Američki psiholog, bavio se emocionalnim zdravljem, na čelu Američkog odbora za emocionalno zdravlje
  - 1921. *Psychology: A study of mental life*
  - 1938. *Experimental Psychology*
- Naglašavao važnost “unutarnjih stanja” (osjećaja), te nagona
  - postavio osnove funkcionalističkog pristupa Stimulus-Organism-Response (SOR), za razliku od biheviorističkog SR
- 1.svj. rat – angažiran da izradi test kojim bi se mjerilo emocionalno zdravlje vojnika i koji se može primijeniti grupno
- *Personal Data Sheet* (ime skriva namjenu)
  - pretočio pitanja iz klasičnog intervjua u formu papir-olovka
  - ‘Smeta li vas kada vas ljudi na cesti promatraju? DA – NE’
  - ponuđeni odgovori indiciraju simptome psihopatološkog ponašanja i doživljavanja
  - zbroj prisutnih simptoma – razina psihopatološkog ponašanja
- Upitnik je malo ‘zakasnio’, rat je završio, pa je napravljena civilna verzija: *Woodworth psychoneurotic inventory*
  - verzije za djecu i za odrasle
  - pojam emocionalne prilagodljivosti razdijeljen na facete – prilagodljivost u školi, prilagodljivost u obitelji...
- Prvi upitnik *samoprocjene (self-report)*
  - ispitanik sam procjenjuje svoje osobine (ponašanje, doživljavanje)
- Samoprocjena postala standard u upitnicima ličnosti
- Prednosti:
  - nema bolje osobe od klijenta samoga da odgovori na pitanje
  - jednostavna primjena
- Nedostaci:
  - osjetljivost na iskrivljavanje odgovora (namjerno ili slučajno; preuveličavanje ili umanjivanje simptoma)
  - mogućnost da klijent ne može dobro procijeniti

#### Projektivni testovi

- Umjesto upitnika samoprocjene, postoje i drugačije metode: projektivne tehnike
- Ispitanici trebaju na nejasan podražaj projicirati svoja razmišljanja, stavove, potrebe, strahove, nade, želje, motivaciju – dakle: ličnost.
- Nejasan podražaj može biti mrlja tinte, crtež, fotografija ili nešto drugo
- Nema problema sa iskrivljavanjem odgovora
- Ali postoji problem interpretacije – ona uvelike ovisi o osobnom iskustvu ispitivača → test gubi objektivnost, potreban je izuzetno dugačak period edukacije za primjenu takvih testova

### Rorschach inkblot test (1921)

- Hermann Rorschach, švicarski psihijatar
- 10 simetričnih mrlja od tinte, neke c/b, neke u boji
- Zadatak ispitanika: reći na što podsjećaju, što u njima vidi
- KomPLICIRANA interpretacija – loša pouzdanost i objektivnost

### TAT Thematic Apperception Test (1935)

- Henry Murray i Christiana Morgan, Sveučilište Harvard
- 30 c/b slika sa 1-2 osobe (ponekad i nekoliko osoba u interakciji)
- Zadatak ispitanika: opisati što je do događaja dovelo, što se trenutno događa, kakve će biti posljedice, što likovi osjećaju, što misle, što bi mogli slijedeće napraviti...

### Rosenzweigov test frustracija

- Saul Rosenzweig, Harvard
- Test latentne hostilnosti, neprijateljstva
- Crtan poput stripa – dvije osobe u frustrirajućoj situaciji, oblačići sa tekstom, nedostaje odgovor ili reakcija frustrirane osobe

### Upitnici interesa i vrijednosti

- Interesi – sviđanja različitih aktivnosti, sklonosti, preferencije
- Prvi upitnik interesa Edward Strong – 1920.
  - 1920. *Strong Vocational Interest Blank SVIB*
  - 1974. *Strong-Campbell Interest Inventory SCII*
  - 1994. *Strong Interest Inventory*
- Vrijednosti – važni ciljevi u životu pojedinca, usmjeravaju naše ponašanje
- Prvi upitnik vrijednosti Allport i Vernon – 1932.
  - 1932. *Study of Values*
  - 2003. Kopelmanova revizija tog upitnika

## PSIHOLOGIJSKI TEST

- Psihometrija – psihologijska disciplina koja se bavi problemima mjerenja psihičkih procesa
- Dva aspekta
  - Metodološki → kako mjerimo
  - Supstancijalni → što mjerimo

### Metode psihometrije

- U psihometriju ubrajamo konstrukciju i provjeru instrumenata za mjerenje
- kognitivnih ili intelektualnih osobina (intelektualne sposobnosti, znanja, perceptivne sposobnosti)
- konativnih osobina ili osobina ličnosti (interesi, vrijednosti, ličnost u užem smislu, patološke osobine ličnosti, emocije)
- stavova i mišljenja (skale stavova prema raznim predmetima mjerenja)
- psihomotornih vještina i spretnosti
- skale za procjenjivanje osobina
- esteziometrijske metode za ispitivanje apsolutne i diferencijalne osjetljivosti

### Direktno i indirektno mjerenje

- Direktno mjerenje
  - predmet mjerenja i jedinica mjerenja iste vrste
  - dužina, širina, visina – dužina
- Indirektno mjerenje
  - predmet mjerenja i jedinica mjerenja različite vrste
  - temperatura – visina žive u termometru
  - logika → visina žive je posljedica djelovanja temperature
  - isti princip kod psihologijskog mjerenja npr. broj točno riješenih zadataka je posljedica intelektualnog funkcioniranja
- Sva psihologijska mjerenja (osim esteziometrijskih!) jesu mjerenja indirektnog tipa o psihičkom funkcioniranju sudimo na temelju reakcija na podražaje različite složenosti

### Mjerenje u psihologiji

- Indirektno mjerenje
- standardizirani postupak pomoću kojeg se neki psihički proces izaziva, pa se efekti tog procesa (reakcije ispitanika) registriraju, mere i vrednuju uspoređujući sa reakcijama sličnih ispitanika koje su dobivene u jednakoj situaciji (A. Krković).

### Standardiziranost mjerenja

- Ključno (definicija!) – obuhvaća:
  - specifikaciju je li mjerenje individualno ili grupno
  - specifikaciju fizikalnih uvjeta i potrebnog pribora (osvjetljenje, veličina radnog prostora, pregrade i sl.)
  - specifikaciju zadataka u testu koji moraju biti isti za sve ispitanike
  - preciznu uputu ispitanicima što je cilj ispitivanja i kako trebaju raditi na zadacima i na koji način trebaju davati odgovore
  - popis odgovora koji imaju simptomatsku valjanost (ključ za odgovore)
  - pravila o tome kako se formira i izračunava individualni rezultat u testu
  - podatke potrebne za vrednovanje individualnog rezultata (norme, veličinu pogreške mjerenja, metrijske karakteristike)

### Idealna standardizacija...?

- Na rezultate testiranja utječu i drugi faktori koje ne možemo u potpunosti kontrolirati
- Raspoloženje, motivacija, stres
- Testiranje za posao, ili za profesionalnu orijentaciju
- Testiranje male djece
- Problemi privatne prirode
- Posebne grupe
- Kulturalni background, SES...

### Objektivnost mjerenja

- Osnovna metrijska karakteristika: nezavisnost rezultata mjerenja od ispitivača-mjerioca
- Osnovni preduvjet – slijeđenje uputa iz priručnika, poštovanje standardizacijskih pravila
- Objektivnost ovisi o dvije stvari
  - O ispitivaču
  - O vrsti zadataka

- Mjerni postupak je objektivan samo ako različiti ispitivači dolaze do istih rezultata – edukacija
- Ključ za ocjenjivanje – popis odgovora koji imaju simptomatsku vrijednosti – točni ili indikativni odgovori

#### Što je psihologijski test?

- Test je sustavni postupak promatranja i opisivanja ponašanja uz pomoć brojčanih ljestvica ili utvrđenih kategorija (Cronbach)
- Psihologijski test je objektivna i standardizirana mjera uzorka ponašanja (Anastasi)
- Standardizirani postupak pomoću kojeg se izaziva neka određena aktivnost, a onda se učinak te aktivnosti mjeri i vrednuje tako da se individualni rezultat usporedi sa rezultatima koji su dobiveni kod drugih individua u jednakoj situaciji (Z. Bujas)

#### Što je psihologijski test?

Bitno iz definicija:

- Standardizirani oblik prikupljanja informacija
- Mjeri psihologijske osobine indirektnim putem
- Mjera uzorka ponašanja
- Omogućava kvantificiranje razvijenosti osobine

Svrha → detekcija, predikcija

Kvantificirati osobine i ponašanja

Prikazati osobine pojedinca na nekom konstrukt

Razlike u rezultatima trebaju predstavljati razlike u izraženosti konstrukta (usporedba pojedinaca)

#### Test

- Test nije fotokopirana stranica sa pitanjima
- Sastoji se od slijedećih elemenata:
  - Test, svezak sa zadacima (i formular za odgovore)
  - Priručnik za test
- Upute za provedbu ispitivanja
- Upute za ispitanika
- Primjeri zadataka i skala za odgovore
- Ključ za odgovore
- Norme za različite populacije
- Informacije o metrijskim karakteristikama

#### Vrste psihologijskih testova

Klasifikacija prema...

- ... sadržaju predmeta mjerenja
- ... načinu primjene testa
- ... bodovanju odgovora
- ... vremenskom ograničenju
- ... prirodi zadatka
- ... broju predmeta mjerenja
- ... broju čestica
- ...

#### Testovi prema sadržaju predmeta mjerenja

Testovi sposobnosti

- Opće inteligencije, pojedinačnih intelektualnih sposobnosti, razvojni testovi, specijalnih sposobnosti (umjetničkog prosuđivanja...)

Testovi postignuća

- znanja, vještina, informiranosti...

Upitnici ličnosti

- Ličnosti u užem smislu, pojedinačnih crta ličnosti (agresivnost, kreativnost...), psihopatoloških osobina, interesa, vrijednosti
- Analitički ('klasični') i sintetički (npr. projekтивni)

Neuropsihologijski testovi

#### Testovi prema načinu primjene

Individualni testovi

- 1 na 1
- tipično za razvojne testove, testove inteligencije za 'donji dio krivulje', neke testove ličnosti, projekтивne tehnike, deksterimetre, neuropsihološke testove
- varijabla ispitivača!

Grupni testovi

- grupa klijenata, 1-2 ispitivača
- često testovi inteligencije, ličnosti, interesa, vrijednosti
- uvjeti grupnog testiranja!

#### Testovi prema načinu bodovanja odgovora

Točni odgovori

- Klasičan točan odgovor:
  - Mjesec se vrti oko Zemlje      DA ili NE
- Testovi postignuća, sposobnosti, deksterimetri

Simptomatski odgovori

- Odgovor koji upućuje na razvijenost osobine
  - Bojim se mraka      DA ili NE
- Testovi ličnosti, interesa, vrijednosti...

→ Indikativni odgovori

#### Testovi prema vremenskom ograničenju

(Klasifikacija testova sa T ili N odgovorima)

Testovi snage

- Zadaci u testu od lakih prema teškim, nema vremenskog ograničenja, na razvijenosti osobine ukazuje težina najtežeg točno riješenog zadatka

Testovi brzine

- Svi zadaci lagani, vremensko ograničenje ne dozvoljava da se riješi do kraja, na razvijenosti osobine nam ukazuje do kud je ispitanik stigao

#### Testovi prema broju predmeta mjerenja

Jednodimenzionalni testovi

- Testovi koji mjere samo jedan konstrukt, npr. agresivnost, sramežljivost, opću inteligenciju, znanje iz kolegija Mjerenje u psihologiji, itd.

Višedimenzionalni testovi

- Testovi koji mjere više konstrukata, npr. osobine ličnosti (Big Five), interese (RIASEC), pojedinačne intelektualne sposobnosti (numerička, verbalna, specijalna...)

### Testovi prema prirodi zadataka

Testovi maksimalnog učinka

- Zadaci sa točnim i netočnim odgovorima, ispituju koliko je kandidat u nečemu uspješan

Testovi tipičnog ponašanja

- Zadaci bez točnih i netočnih rješenja, ispituju kako se kandidat uobičajeno ponaša

→ uobičajena i sveobuhvatna podjela

Testovi maksimalnog učinka	Testovi tipičnog ponašanja
Testovi	Upitnici samoprocjene
Sposobnosti, znanje...	Ličnost, interesi, vrijednosti, potrebe, stavovi, mišljenje...
Točno vs. Netočno	Simptomatično vs. nesimptomatično
Često ograničeni u vremenu	Uglavnom neograničeni u vremenu
Laganje rijetko, samo prema nižem rezultatu; ali problem pogađanja	Laganje - u svim smjerovima

### Rezultat ispitnika

- U testu
- U zadacima
- Reakcije oslikavaju razvijenost osobina koje mjerimo
- Kvantificiranje ponašanja – brojčani rezultati koji upućuju na razvijenost osobine
- O tome ćemo govoriti drugi put

### Drugi oblici procjena

- Uz testove koriste se i drugi oblici procjena psiholoških osobina
- Intervju, igranje uloga, fokus grupe, opažanje
- I te metode imaju svoja pravila, no one ne ulaze u psihometriju
- Distinkcija: kvantitativne i kvalitativne metode
- Kvalitativne metode mogu na prvi pogled djelovati 'humanije', ali njihova 'pravednost' ovisi o njihovoj specifikiranosti
- Je li postupak dobro osmišljen, pravilno proveden, jesu li rezultati točno interpretirani, jesu li na temelju njih donesene primjerene odluke

### Primjena testova

- U dijagnostičke i prognostičke svrhe
- Pomoć u odlučivanju

- Sam rezultat nam ništa ne znači, cilj je ustanoviti...
  - Ima li dijete sposobnosti da krene u školu?
  - Hoće li kandidat biti dobar na nekom poslu?
  - Koje zanimanje bi mi najviše odgovaralo?
  - Ima li osoba narušeno mentalno zdravlje do te mjere da nije u mogućnosti raditi?
  - Ima li 'nestašno dijete' hiperaktivni poremećaj (ADHD) ili je to samo trenutna faza?
- Primjena testova u selekciji, za samospoznavanje, za trijažu i dijagnozu, za analizu radnog mjesta

### Primjena testova u selekciji

- Najčešća primjena testova
- Rijetko kad jedini kriterij odabira kandidata – molbe, životopisi, preporuke, intervjui
- Koriste se i testovi sposobnosti i upitnici ličnosti
- Testovi sposobnosti – opće, pojedinačnih
  - povezanost inteligencije i radne uspješnosti – dobar prediktor
  - najčešće odličnih metrijskih karakteristika i pružaju točne procjene
  - paziti na eventualnu nepravdu prema određenim grupama ljudi
  - test ne smije biti niti pretežak, niti prelagan - osjetljivost
- Testovi posebnih sposobnosti i vještina
  - Vještine, spretnost, točnost, brzina, tremor → važno za neka zanimanja
- Upitnici ličnosti
  - Sporno – samoopažanje i samoprocjene
  - Upotreba u nekim posebnim okolnostima – emocionalna stabilnost

### Primjena testova samospoznavanja

- Novija primjena testova
- Posebno potaknuta razvojem kompjutoriziranog testiranja
- Upitnici interesa
  - Prof. orijentacija – nema problema sa samoprocjenama
- Upitnici ličnosti i testovi sposobnosti
  - Korisni, ali samoprimjena ne bi trebala biti omogućena

### Primjena testova za trijažu i dijagnostiku

- Ocjena pojedinčevog stanja
- Različite svrhe, različiti testovi
- Testovi sposobnosti
  - obrazovni programi, poslovna sposobnosti, sudjelovanje u sudskim procesima
- Testovi ličnosti – širi, uža, psihopatološke osobine
  - Bolnice, klinička psihologija, školstvo

### Primjena testova za analizu radnih mjesta

- Analiza radnih aktivnosti – važna za izbor i raspodjelu kadrova, planiranje obrazovanja, određivanje stimulacije i plaća, planiranje razvoja tvrtke
- Različite metode: opažanje, intervjuiranje, upitnici, check-liste
- McCormickov PAQ Position Analysis Questionnaire

### Sudionici mjerenja i njihova prava

- Autori i izdavači testova
  - pojedinci i izdavačke kuće - komercijalni i nekomercijalni testovi
- Ispitivači



- licencirani psiholozi – ograničenja
- Korisnici testovnih rezultata
  - različiti stručnjaci – kliničari (klinički psiholozi, psihijatri), rukovoditelji, razni šefovi, zaposlenici u kadrovskim odjelima (HR), profesori, učitelji, savjetovatelji, zaposlenici ministarstava, društvo u cjelini → izvještaji
- Ispitanici
  - Prava ispitanika: da bude informiran o načinu testiranja, da dobije povratnu informaciju o rezultatima, da mu se osigura povjerljivost nalaza

#### Kvaliteta testova

- Koji test je dobar test?
- Test...
  - ... koji ima razrađene upute za primjenu, bodovanje i interpretaciju rezultata
  - ... koji je jednostavan i brzo se primjenjuje
  - ... koji se sviđa ispitanicima
  - ... čija primjena poboljšava kvalitetu života ispitanika
  - ... koji nam ukazuje na razlike među ispitanicima
  - ... čije ocjenjivanje ne ovisi o ispitivaču
  - ... čiji rezultati nisu slučajni, već su odraz osobina ispitanika
  - ... koji mjeri ono što mi mislimo da mjeri
  - ... koji nam omogućava da usporedimo rezultate različitih ispitanika

#### 5 metrijskih karakteristika

- Test...
  - ... koji nam ukazuje na razlike među ispitanicima - OSJETLJIV
  - ... čije ocjenjivanje ne ovisi o ispitivaču - OBJEKTIVAN
  - ... čiji rezultati nisu slučajni, već su odraz osobina ispitanika - POUZDAN
  - ... koji mjeri ono što mi mislimo da mjeri - VALJAN
  - ... koji nam omogućava da usporedimo rezultate različitih ispitanika - BAŽDAREN

→ Samo je takve testove opravdano koristiti u praksi

#### ZADACI U TESTU

##### Osnovna klasifikacija testova

- testovi maksimalnog učinka
  - testovi znanja i sposobnosti
  - točni i netočni odgovori
  - zadaci: višestruki izbor, nadopunjavanje, T/N, uparivanje
  - ocjenjivanje 0 i 1 (ili drugačije)
- testovi tipičnog ponašanja
  - upitnici ličnosti, interesa i sl.
  - samoprocjena razvijenosti osobine
  - nema točnih i netočnih odgovora
  - zadaci: procjenjivanje, rangiranje
  - ocjenjivanje: rezultat na skali, hijerarhija

##### Zadaci u testovima maksimalnog učinka

- zadaci višestrukog izbora (abcd..)
- dihotomni zadaci (N,T)
- nadopunjavanja
- uparivanja

##### Zadaci višestrukog izbora

Statistička mjera koja pokazuje koliko se “gusto” grupiraju rezultati oko aritmetičke sredine je:

- A. standardna pogreška
- B. standard
- C. stanine skala
- D. standardna devijacija
- (E. ne znam)

##### Dihotomni zadaci

Projektivne tehnike predstavljaju sistem upotrebe projekcijskog sustava u svrhu eksperimentalnog ispitivanja.

- A. DA
- B. NE
- (C. ne znam)

##### Zadaci nadopunjavanja

Zajednički naziv koje objedinjuje testove sposobnosti, znanja i postignuća u kojima postoje točni i netočni odgovori je \_\_\_\_\_.

##### Zadaci uparivanja

Uparite tko predaje koji predmet:

- |                       |                               |
|-----------------------|-------------------------------|
| 1. Mlačić             | a. Psihologija rada           |
| 2. Brkljačić          | b. Deskriptivna statistika    |
| 3. Žebec              | c. Percepcija                 |
| 4. Kaliterna-Lipovčan | d. Diferencijalna psihologija |
| 5. Babarović          | e. Modeli analize varijance   |
| (6. Franc)            | (f. Motivacija)               |
- (dodavanje još jedne opcije otežava pogađanje)

### Zadaci u testovima maksimalnog učinka

- Vrste zadataka:
  - višestruki izbor, dihotomni T/N, nadopunjavanje, uparivanje
- Ocjenjivanje:
  - rješenje je ili točno ili netočno
  - to su sve dihotomni zadaci
  - bodovi 0 i 1
- Težina zadatka:
  - težina pitanja i težina distraktora
  - ponderiranje
  - bodovi npr. 0 i 2 ili 0 i 4... itd
- Problem: pogađanje
  - korekcija: odgovor 'ne znam' i negativni bodovi
  - bodovi npr. -1, 1

### Zadaci u testovima tipičnog ponašanja

- Rangiranje
  - Jednostavno
  - U parovima
  - U grupama (trijadama, tetradama...)
  - Metoda prisilnog izbora
- Procjenjivanje
  - Unipolarno (kontinuum opisan s jednim atributom)
  - Bipolarno (semantički diferencijal)

### Zadaci jednostavnog rangiranja

Slijedeće životne ciljeve rangirajte prema njihovoj važnosti za vas, počevši time što ćete najvažnijem cilju dodijeliti rang 1, sljedećem po važnosti rang 2... itd

Ugodan život	_____
Jednakost	_____
Uzbudljiv život	_____
Obiteljska sigurnost	_____
Sloboda	_____
Zdravlje	_____

### Zadaci rangiranja u parovima

Pročitajte svaki par ponuđenih tvrdnji i u svakom paru odaberite onu vrijednost koja je vama osobno važnija:

1. a) jednakost  
b) dobrohotnost

### Zadaci rangiranja u grupama

U svakoj pojedinoj grupi, rangirajte zanimanja prema tome koliko vas interesiraju:

1. a) bračni terapeut  
b) sociolog  
c) profesor engleskog jezika  
d) marketinški stručnjak

### Metoda prisilnog izbora

- Modifikacija metode rangiranja u parovima ili grupama gdje su sve ponuđene alternative jednako socijalno poželjne
- Unutar svakog para, odaberite tvrdnju koja vas točnije opisuje
  - 1a) Sviđa mi se biti vani i družiti se s raznim ljudima.
  - 1b) Zanima me učiti o umjetnosti.
  - 2a) Često se osjećam depresivno.
  - 2b) Često se osjećam anksiozno.

### Rangiranje – prednosti i mane

Rezultat ispitnika – hijerarhija, rang – ordinalna razina mjerenja

- nema info o *izraženosti* svojstva
- relativno mjerenje, a ne apsolutno (isti 1. rang kod dvije osobe)
- ipsativno mjerenje, intraindividualne razlike

Dva entiteta ne mogu se proglasiti jednako važnima

- smanjuje stabilnost hijerarhije, pogotovo u sredini

Zahtijevno kod velikog broja entiteta

Moguće cirkuarne pogreške (osim kod jednostavnog rangiranja)

Metoda prisilnog izbora – prednost: laganje

### Procjenjivanje

Ispitanik procjenjuje *stupanj* izraženosti svoje osobine

Vrlo uobičajena metoda, također

### Skala procjena

Koristeći skalu od 1 do 4, procijenite koliko vam se sviđa svako od ponuđenih zanimanja...

1 – uopće mi se ne sviđa

4 – izuzetno mi se sviđa

○ Soboslikar	1	2	3	4
○ Elektrotehničar	1	2	3	4
○ Profesor u srednjoj školi	1	2	3	4
○ Dizajner	1	2	3	4
○ Liječnik	1	2	3	4

Broj uoprišnih točaka može varirati (2,3,4,5,7,9,10...) - povećanje varijabiliteta, no pitanje efikasnosti

Postojanje središnje točke ili ne (neparan ili paran broj točaka)

Dihotomna skala procjena: Sviđaju li vam se navedena zanimanja... novinar... A NE

Likertova skala: svi stupnjevi opisani, najčešće broji 5 točaka, ima središnju točku

### Kontinuirane skale procjena

Ispitanik može odrediti bilo koji broj na skali (za razliku od *diskretnih*)

Označite kvačicom u kojoj mjeri osjećate nelagodu u slijedećim situacijama:

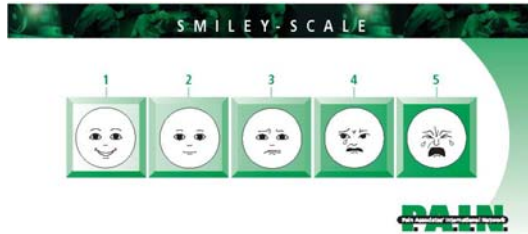
Ne osjećam nelagodu	Osjećam nelagodu
---------------------	------------------

U zubarskoj čekaonici -----

Prije ispita \_\_\_\_\_  
Na ispitu \_\_\_\_\_

### Grafičke skale procjena

Za djecu koja nemaju dovoljno razvijen pojmovni sustav i vokabular i koja nisu došla do potrebnog stupnja kognitivnog razvoja



### Skala semantičkog diferencijala

Bipolarna skala, 7 uporišnih točaka, Osgood, Suci, Tannenbaum (1957)

Mjerenje samopoimanja i ispitivanje mišljenja i stavova

Kako se osjećam kada sam vani u društvu drugih ljudi

Toplo	3	2	1	0	1	2	3	Hladno
Napeto	3	2	1	0	1	2	3	Opušteno
Slabo	3	2	1	0	1	2	3	Snažno

Usporedba različitih ispitanika, ali i različitih objekata (npr. "kako se osjećam kada sam u prirodi", "kako se osjećam kada sudjelujem na natjecanju"...)

### Procjenjivanje – prednosti i mane

Rezultat ispitanika – brojčana procjena izraženosti svakog svojstva

→ apsolutno mjerenje – **NORMATIVNO**

→ interindividualne i intraindividualne razlike

Mogućnost davanja istih procjena

Jednostavna primjena

Pseudointervalna skala → parametrijske analize

Bolje u kroskulturalnoj primjeni jer pojedine čestice ne ovise o drugima

Problemi

- nekad asimetrične i sužene distribucije (pogotovo kod skala vrijednosti)
- osjetljive na način odgovaranja (npr. davanje ekstremnih odgovora, slaganje sa ponuđenim tvrdnjama *akvisencija*)

### Procjenjivanje ili rangiranje

- Dva suprotstavljena pristupa
- Češće je procjenjivanje
- Parametrija, jednostavnost, iste procjene
- Odabir ovisi o predmetu mjerenja & vrsti testova

### BODOVANJE URATKA U ZADATKU

- Rezultat ispitanika u zadatku je broj
- U testovima max učinka ili dihotomnim skalama - broj 0 ili 1
  - 1 = točno rješenje ili indikativan odgovor
  - 0 = netočno rješenje ili odsutstvo indikativnog odgovora
- U skalama procjena – brojčani rezultat na skali (npr. 1-5)
- Kod rangiranja – rang svojstva koje je ispitanik ocijenio (npr. 4)
- Svaki zadatak je varijabla
- Nazivamo ga još i čestica ili item
- Kakve su njene karakteristike?
  - parametri zadatka
  - što se s njome događa ako promijenimo način bodovanja?

### Dihotomni zadaci

#### Značenje uratka u dihotomnom zadatku

- Razvrstava ispitanike u dvije kategorije
  - Ima svojstvo
  - Nema svojstvo
- Dihotoman = dvije kategorije
  - 1 – točan odgovor ili prisutstvo svojstva koje se mjeri
  - 0 – netočan odgovor ili odsutstvo svojstva koje se mjeri
- Jačina intervalne skale
- Parametri – M, sd, varijanca, distribucija

#### Dihotomni zadatak

isp	z <sub>1</sub>	z <sub>2</sub>	z <sub>3</sub>	z <sub>4</sub>
1	1	0	0	1
2	1	1	1	0
3	0	0	0	0
4	1	1	0	0
5	1	0	0	1
6	1	1	0	1

### Distribucija dihotomnog zadatka

Rezultat	Frekvencija	Proporcija
1	$f_1$	$p = f_1/N$
0	$f_0$	$q = f_0 / N$
$\Sigma$	N	1
	$(f_1 + f_0 = N)$	$(p + q = 1)$

### Aritmetička sredina i varijanca dihotomnog zadatka

$x_j$	f	$fx_j$		$d=X-M$	$d^2$	$fd^2$
1	$f_1 = pN$	$1 \cdot pN = pN$		$1-p = q$	$q^2$	$pN \cdot q^2$
0	$f_0 = qN$	$0 \cdot qN = 0$		$0-p = -p$	$p^2$	$qN \cdot p^2$
$\Sigma$	N	$pN$				$pqN (q + p) = pqN$
		<b>M = p</b>		<b>(d=X-p)</b>		<b>v = pq</b>

(detaljan opis izvoda možete naći u Krkoviću)

### Parametri dihotomnog zadatka

- $M = p$
- $v = p \cdot q$
- Test srednje težine  $\rightarrow p=0.5$
- Najveće razlikovanje ispitanika  $\rightarrow v_{\max} = 0.25$

**Primjer 1.** Neki zadatak rješavalo je 1000 ispitanika, od kojih je 600 dalo indikativan odgovor. Koliko iznose M i v zadatka?

**Primjer 2.** U tablici su navedeni rezultati ispitanika u nekom zadatku. Koliko iznose parametri zadatka?

isp	$z_1$	$z_2$	$z_3$
1	1	0	1
2	1	1	1
3	0	0	1
4	1	1	0
5	1	0	1
6	1	1	1
7	0	0	0
8	1	1	1
9	0	1	1
10	1	1	1
M=?			
V=?			

### TRANSFORMACIJA URATKA U ZADATKU

- Što se događa ako promijenimo bodovanje?
- Tipično bodovanje: T, N  $\rightarrow$  1, 0
- Promjena: T, N  $\rightarrow$  2, 0 ... ili ... T,N  $\rightarrow$  1, -1
- Svaka od tih promjena – transformacija

### Transformacije

- Dodavanje konstante  $a=1$   
 $1, 0 \rightarrow 2, 1$
- Oduzimanje konstante  $a=1$   
 $1, 0 \rightarrow 0, -1$
- Množenje s konstantom  $a=2$   
 $1, 0 \rightarrow 2, 0$
- Dijeljenje s konstantom  $a=2$   
 $1, 0 \rightarrow 0.5, 0$

**Primjer.** O kojoj se transformaciji radi?

- T,N = 3,0 ?
- T,N = 2,1?
- T,N = 1,-1?

### Kako se transformacije reflektiraju na parametre zadatka?

Dodavanje konstante

- $M' = M + a$
- $v' = v$
- $Sd' = Sd$

Oduzimanje konstante

- $M' = M - a$
- $v' = v$
- $Sd' = Sd$

Množenje konstantom

- $M' = a \cdot M$
- $v' = a^2 \cdot v$
- $Sd' = a \cdot Sd$

Dijeljenje konstantom

- $M' = M / a$
- $v' = v / a^2$
- $Sd' = Sd / a$

**Primjer 1.** U nekom zadatku ispitanici su mogli postići 1 bod za točno rješenje i 0 bodova za netočno. Utvrđeni parametri tog zadatka iznosili su  $M=.40$ ,  $v=.24$ . Kako će se promijeniti parametri zadatka, ako promijenimo bodovanje tako da točno rješenje donosi 2 boda, a netočno 0, te ako ga promijenimo tako da točno rješenje donosi 1 bod, a netočno -1. Prvo definirajte o kojoj se transformaciji radi, a zatim izračunajte parametre zadatka

Promjena bodovanja  $1, 0 \rightarrow 2, 0$

Promjena bodovanja  $1, 0 \rightarrow 1, -1$

**Primjer 2.** Zadatak u testu riješilo je 30 od 100 ispitanika. Točno rješenje nosi 2, a netočno 0 bodova. Odredite transformaciju, te potom izračunajte parametre netransformiranog i transformiranog zadatka.

## REZULTAT U TESTU

### Zadatak u testu

- Varijabla, čestica, item
- Odgovori ispitanika
  - Rezultat na skali (npr. 1-5)
  - Dihotomni odgovor (0 ili 1)
  - Rang nekih entiteta – izvan ove rasprave
- Parametri
  - distribucija, indeks lakoće p, M, sd, v,

### Dihotoman zadatak

- Dijeli ispitanike u 2 kategorije: 0 i 1
- Distribucija:  $f_0$  i  $f_1$
- Indeks lakoće  $p = f_1 / N$
- $M = p$
- $v = p * q = sd^2$
- Srednja težina  $p = .5$
- Maksimalna varijanca  $v = .25$

### Procjena na skali (npr. 1-5)

- Dijeli ispitanike u N kategorija: 1, 2, 3, 4 i 5
- Distribucija:  $f_1 \quad f_2 \quad f_3 \quad f_4 \quad f_5$
- $M = \sum x / N$
- $v = sd^2$
- Srednja težina  $p =$  na polovici skale odgovora (3 u ovom slučaju)

### Zadatak vs. test

- Rezultat u zadatku oslikava razvijenost osobine koju mjerimo, ali...
  - najčešće nije dovoljan da donesemo čvrst i pouzdan zaključak (npr. samo jedno pitanje na ispitu)
- Test se sastoji od niza zadataka
  - broj zadataka varira – skala zadovoljstva životom ( $k = 1$ ) ili MMPI ( $k = 567$ )
  - Pouzdanija mjera izvedena iz većeg broja mjerenja
  - I osjetljivija
  - I sadržajno valjanija – kvalitetnija

### Rezultat u testu

- Broj iskazanih indikativnih ponašanja u nekom nizu zadataka
  - Veći broj iskazanih indikativnih ponašanja → veći rezultat u testu → veća razvijenost mjerene osobine
- Matematički, rezultat u testu je zbroj odgovora ispitanika u pojedinim zadacima ili LINEARNA KOMBINACIJA (LK)

### Linearna kombinacija

$$LK = x_1 + x_2 + x_3 + \dots + x_k$$

isp	$z_1$	$z_2$	$z_3$	$z_4$	LK
1	1	0	0	1	2
2	1	1	1	0	3
3	0	0	0	0	0
4	1	1	0	0	2
...					
N-ti	1	0	0	0	1

isp	$z_1$	$z_2$	$z_3$	$z_4$	LK
1	5	4	4	5	18
2	2	1	2	3	8
3	1	3	3	2	9
4	4	3	4	4	15
...					
N-ti	2	3	3	5	13

### Paramteri bruto rezultata

- Distribucija
- Indeks težine / lakoće
- M
- Sd, v

### Distribucija bruto rezultata

- Raspon: min-max
  - min: najmanji rezultat koji ispitanik može postići
  - max: najveći rezultat
- Raspon varira ovisno o
  - broju zadataka (k) i
  - formatu skale odgovora (dihot. zadaci ili skale procjena)
  - veći broj zadataka i veći broj uporišnih točaka povećavaju teoretski raspon testa
- **Npr.** Teorijski raspon testa koji broji 10 zadataka bodovanih 0 i 1 je min= 0, max = 10
- **Npr.** Teorijski raspon testa koji broji 50 zadataka-skala procjene raspona 1-5 je min= 50, max = 250

### Izgled distribucije: težina testa

- Normalna distribucija
  - Primjeren test
- Negativno asimetrična
  - Razvučena na lijevom kraju, svi nagurani na desnom
  - Prelagan test
- Pozitivno asimetrična
  - Razvučena na desnom kraju, svi nagurani na lijevom
  - Pretežak test

### Indeks težine (lakoće) testa

- Intuitivno jasno: neko 'težište' rezultata
- poput aritmetičke sredine, ali izražava se u jedinicama skale
  - za testove sa dihotomnim zadacima kreće se u vrijednosti 0-1
  - za testove sa zadacima-skalama procjene, u vrijednosti raspona skale
- prosječni  $p = M / k$

### Aritmetička sredina bruto rezultata

- zbroj aritmetičkih sredina zadataka
$$M_{LK} = M_1 + M_2 + M_3 + \dots + M_k$$
- kod dihotomnih zadataka
$$M = p, \text{ pa je}$$
$$M_{LK} = p_1 + p_2 + p_3 + \dots + p_k$$

### Varianca bruto rezultata

- Komplikiranija od samog zbroja varijanci
- Varianca = zbroj matrice kovarijanci
- Kovarijanca – pokazuje kako dvije varijable variraju zajedno (varianca pokazuje kako varira jedna varijabla)
- Kovarijanca između dvije varijable iznosi
$$c_{12} = sd_1 * sd_2 * r_{12}$$
- Uključuje element varijabilneta i element korelacije među varijablama

### Matrica kovarijanci C

Matrica korelacija R pomnožena s pripadajućim standardnim devijacijama

	$z_1$	$z_2$	$z_3$
$z_1$	$c_{11} = r_{11}sd_1sd_1$	$c_{12} = r_{12}sd_1sd_2$	$c_{13} = r_{13}sd_1sd_3$
$z_2$	$c_{21} = r_{21}sd_2sd_1$	$c_{22} = r_{22}sd_2sd_2$	$c_{23} = r_{23}sd_2sd_3$
$z_3$	$c_{31} = r_{31}sd_3sd_1$	$c_{32} = r_{32}sd_3sd_2$	$c_{33} = r_{33}sd_3sd_3$

	$z_1$	$z_2$	$z_3$
$z_1$	$c_{11} = v_1$	$c_{12} = r_{12}sd_1sd_2$	$c_{13} = r_{13}sd_1sd_3$
$z_2$	$c_{21} = r_{21}sd_2sd_1$	$c_{22} = v_2$	$c_{23} = r_{23}sd_2sd_3$
$z_3$	$c_{31} = r_{31}sd_3sd_1$	$c_{32} = r_{32}sd_3sd_2$	$c_{33} = v_3$

### Varianca LK

- $v_{LK} = \sum v_i + 2\sum r_{ij}sd_i sd_j$ 
  - uključuje informaciju o varijabilitetu pojedinih čestica
  - i o korelacijama među njima

**Primjer.** Neki test sadrži 3 zadatka. Aritmetičke sredine i standardne devijacije te korelacije među pojedinim zadacima prikazane su u sljedećim tablicama. Izračunajte koliko iznosi M i v cijelog testa.

	$z_1$	$z_2$	$z_3$
M	0,8	0,4	0,4
Sd	0,2	0,3	0,2

	$z_1$	$z_2$	$z_3$
$z_1$	1	0,4	0,6
$z_2$	0,4	1	0,5
$z_3$	0,6	0,5	1

### Linearne transformacije LK

- Bodovanje zadataka može se promijeniti
- Ponderiranje
  - Jednako
  - Nejednako
- Kako se to odražava na ukupni rezultat u testu?

#### Dodavanje konstante

- $M' = M + a$
- $v' = v$
- $Sd' = Sd$

#### Oduzimanje konstante

- $M' = M - a$
- $v' = v$
- $Sd' = Sd$

#### Množenje konstantom

- $M' = a M$
- $v' = a^2 v$
- $Sd' = a Sd$

#### Dijeljenje konstantom

- $M' = M / a$
- $v' = v / a^2$
- $Sd' = Sd / a$

**Primjer.** Neki test broji 10 zadataka. Njegovi parametri su  $M = 1$ ,  $Sd = 1,1$ ,  $v = 1,19$ . Koliko bi iznosili M i Sd da su svi zadaci bodovani sa 2 i 0? O kojoj se transformaciji radi?

## PET METRIJSKIH KARAKTERISTIKA

- Pouzdanost
- Valjanost
- Osjetljivost
- Objektivnost
- Baždarenost

### Ključni statistički pojmovi

- varijabilitet
- korelacija

## OSJETLJIVOST (DISKRIMINATIVNOST)

- **svojstvo testa (i zadataka) da uspješno razlikuje ispitanike po izraženosti svojstva koje se mjeri**
- Razlika među ispitanicima važna – praktična primjena
- Osjetljivost testa & osjetljivost zadataka

### Uradak (rezultat) u testu

- Rezultat u testu je LK uratka u zadacima
- Varijabla – M, Sd, distribucija...
- Kada je rezultat u testu osjetljiv?
  - Kada se ispitanici dobro razlikuju
  - Idealno: kada svatko dobije različit rezultat → pravokutna distribucija

### Što čini osjetljivost testa?

- Osjetljivost zadataka
- Varijabilitet rezultata u testu (broj zadataka, Sd)
- Težina testa
- Oblik distribucije rezultata u testu

### Osjetljivost zadataka

- S obzirom da je uradak u testu zbroj uratka u zadacima...  
...što su zadaci osjetljiviji, to je i test osjetljiviji

### Varijabilitet testa

- Što veći sd, v, TR, to veća osjetljivost testa
  - Veći broj mogućih rezultata (zadataka)
  - Broj stupnjeva na skali procjena
  - Različito ponderiranje rezultata

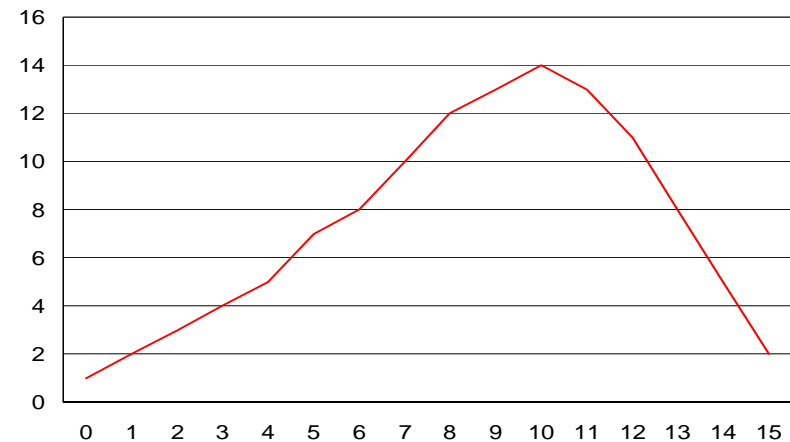
### Težina testa

- Najbolje osrednja

- Teški test → pozitivno asimetrična distribucija
- Lagani test → negativno asimetrična distribucija
- Prosječno težak test → normalna distribucija

### Distribucija testa

- Područje veće osjetljivosti – mjesto gdje je manja 'gužva' rezultata (na slici: lijevi kraj distribucije)
- Visoka osjetljivost: razvučena distribucija, te pravokutna distribucija



## POUZDANOST

- **konzistentnost, dosljednost ili stabilnost mjerenja**
- Najjednostavnije poimanje pouzdanosti: ako pri više mjerenja dobijemo iste rezultate, mjerenje je pouzdano
- To nije točnost mjerenja

### Primjer: mjerenje težine osobe od 60 kg

vaga 1: svaki put ista i točna težina: 60, 60, 60, 60

→ točno i pouzdano mjerenje

vaga 2: svaki put ista težina, ali 3 kg manja: 57, 57, 57, 57

→ pouzdano, ali nije točno

→ nije problem, matematički ispravimo

→ sistematski varijabilitet

vaga 3: svaki put druga težina: 61, 59, 62, 57

→ netočno & nepouzdan

→ problem, nema matematičke korekcije

→ nesistematski varijabilitet

Bruto rezultat ispitanika pod utjecajem je

1. pravog rezultata
2. sistematskih varijabilnih faktora
3. nesistematskih varijabilnih faktora

Zanima nas pravi rezultat i pokušavamo ga procijeniti

**Pravi rezultat** – stvarna veličina konstrukta kojeg mjerimo, npr.  
težina ispitanika (u primjeru 60 kg)  
stupanj intelektualnog funkcioniranja  
poznavanje sadržaja nekog kolegija

→ želimo da su rezultati mjerenja oslobođeni utjecaja SV i NSV

#### Sistematski varijabilitet

- stabilno djelovanje, u jednom smjeru
- matematički zahvatljivo, ako smo svjesni njegovog postojanja  
npr. vaga koja sustavno pokazuje nižu težinu ili verbalni test inteligencije na engleskom
- utjecaj na individualni rezultat → na ispitanike djeluje isto u različitim ispitnim situacijama
- ne narušava pouzdanost mjerenja

#### Nesistematski varijabilitet

- ne utječe prema nekom pravilu, već slučajno → nekad povećavaju, a nekad smanjuju bruto rezultat
- raspodjeljuje se normalno
- u beskonačnom broju mjerenja  $M = 0$  → poništava se
- ne korelira ni sa čime, jer je slučajna
- utjecaj na individualni rezultat → djeluje različito u različitim ispitnim situacijama
- narušava pouzdanost mjerenja

#### Izvori NSV

- Konstrukcija testa
- Primjena
- Bodovanje i interpretacija rezultata
- Drugi mogući izvori

#### Konstrukcija testa

- Odabir pitanja i indikativnih ponašanja – test znanja (nekad imam, nekad nemam sreće), testovi ličnosti, sposobnosti...

#### Primjena testa

- Ispitivanje - vrućina, zagušljivost, osvjetljenje prostorije, doba dana, olovka koja ne piše, klupa koja odvlači pažnju
- Ispitanik - umor, emocionalno stanje, djelovanje lijekova, krivo shvaćena pitanja (“što ne čini...”), krivi upis u formularu za odgovore (npr. zbog jednog preskočenog pitanja)
- Ispitivač - stil i način, profesionalnost, potpitanja (navođenje na odgovor na usmenom, klimanje glavom)

#### Bodovanje i interpretacija

- Kompjutorski testovi & objektivni testovi višestrukog izbora → najmanji utjecaj NSV pri bodovanju
- Neobjektivni testovi (kreativnost: cigla; ličnost: projektni testovi; opažanje ponašanja: reći “dobro

jutro”) – problem 2 opsežni priručnici, dugogodišnji trening za adekvatno bodovanje

#### Drugi izvori NSV

- Iskrivljavanje odgovora (namjerno ili nenamjerno) → npr. zlostavljanje u obitelji

#### Procjena pouzdanosti

- Cilj mjerenja je da rezultati odražavaju isključivo pravu veličinu mjerenja, tj. pravi rezultat, a ne utjecaj SV i NSV
- Međutim, izmjerimo težinu ispitanika i dobijemo: 61, 59, 62, 57
- Što činimo u tom slučaju? Što je prava težina tog ispitanika? Kako ju pokušavamo procijeniti?

Prosjeck → 59.75 kg (Pravi rezultat = 60 kg)

#### Procjena pouzdanosti

- NSV → u beskonačnom broju mjerenja  $M = 0$
- Intuitivno jasno: što veći broj mjerenja, to veća pouzdanost
- Utjecaj NSV se smanjuje što je broj mjerenja veći

#### Procjena pouzdanosti

- **Pouzdanost je nezavisnost rezultata mjerenja od utjecaja nesistematskih varijabilnih faktora**
- Analiza pouzdanosti jest razdvajanje nesistematskog (slučajnog) varijabiliteta od stabilnog
- Stabilni varijabilitet = pravi rezultat & sistematski varijabilitet
- Analiza pouzdanosti ne zahvaća razliku između varijabiliteta uzrokovanog pravim rezultatom i sistematskog varijabiliteta

#### Vrste pouzdanosti

stabilnost testovnih rezultata u vremenu → retest pouzdanost

ekvivalentnost testovnih rezultata u dvije forme testa → pouzdanost ekvivalentnih formi

homogenost (konzistentnost) testovnih rezultata → pouzdanost unutarnje konzistencije

#### Retest pouzdanost – stabilnost mjerenja

- Stabilnost osobine – dvije primjene – isti rezultat (OK za ispitivanje pouzdanosti mjera stabilnih osobina, npr. karakteristike ličnosti, ali ne raspoloženja)
- primjena jednog testa u dva navrata
- koeficijent pouzdanosti je korelacija između te dvije primjene
- razmak između dvije primjene – od 2 tjedna do 6 mjeseci, ovisi o predmetu mjerenja – 2 mjeseca

#### Pouzdanost ekvivalentnih formi

- ekvivalentne forme – jako slični testovi (slične M, v, interkorelacije, korelacije sa vanjskim varijablama)
- blaže definirane nego paralelne forme (iste M, v, interkorelacije, korelacije sa vanjskim varijablama)
- koeficijent pouzdanosti → korelacija između rezultata u dvije forme
- forma A i forma B → jedan test → povezanost testa sa samim sobom
- različiti zadaci (nema učenja), primjena odjednom (nema osipanja)

#### Homogenost testa

- Ujednačenost mjerenja svih zadataka u testu
- Može se ispitivati na razini pojedinačnih zadataka ili na razini dijelova testa
- Logika ispitivanja konzistentnosti
  - Ako svi zadaci mjere isto (visoke korelacije), znači da nisu interferirali NSV



- Ako mjere različite stvari, djelovali su NSF
- ispituje se povezanost među cjelinama testa ili među zadacima
  - visoka povezanost → test je homogen
- razlikujemo dvije osnovne metode za računanje homogenosti testa
  - split-half pouzdanost i
  - pouzdanost unutarnje konzistencije
 → razlika je matematička, a ne logička

#### Split-half pouzdanost

- test podijelimo na dvije polovice (transverzalno, longitudinalno)
- korelacija među njima → koeficijent pouzdanosti svake polovice → polovica je upola kraća nego cijeli test → dobiveni koeficijent treba korigirati Spearman-Brownovom formulom, kako bismo procijenili pouzdanost cijelog testa
- takva aproksimacija koeficijenta pouzdanosti naziva se split-half koeficijent pouzdanosti

#### Pouzdanost unutarnje konzistencije

- najčešće korištena metoda za procjenu pouzdanosti
- bazira se na prosječnim korelacijama svih zadataka u testu
- određuju se korelacije među svim zadacima, pa se prosječna korelacija korigira SB formulom
- za račun dovoljni paramteri testa i testovnih zadataka
- autori matematičkih formula:
  - Frederic Kuder i M.W. Richardson (1937) – KR-20 i KR-21
  - Lee Cronbach (1951) – Cronbachov alfa

#### Interpretacija koeficijenata pouzdanosti

Veličina koeficijenata (raspon 0-1)

0.7 → relativno loše, ali ovisi o

- vrsti testa,
- dužini testa,
- selekcioniranosti ispitanika i
- načinu računanja pouzdanosti

0.8 → dobar, pristojan, prihvatljiv, ali također ovisi o navedenom

0.9 → odlično

#### Stabilnost i homogenost

- koeficijenti homogenosti govore o djelovanju NSV u trenutku kada je test primijenjen
  - ukoliko postoji djelovanje NSV zadaci su u niskim korelacijama
  - govorimo o utjecaju NSV u toj jednoj primjeni testa
- visoka homogenost nije garancija da NSV neće djelovati pri nekoj budućoj primjeni testa (o tome nam govori koeficijent stabilnosti) ☹ pojava može biti stabilna u vremenu, ali zadaci ne moraju biti homogeni
- homogenost i stabilnost nisu povezani i na osnovu jedne karakteristike testa ne možemo zaključivati o drugoj

## OBJEKTIVNOST

### Objektivnost je neovisnost testovnih rezultata o mjeriocima.

Slaganje rezultata dobivenih od različitih ispitivača

- Ako više ispitivača primjenjuje *isti* test na *istim* ispitanicima, i ako oni dobiju i *iste* rezultate – tada test možemo smatrati objektivnim
- S druge strane, ukoliko postoje razlike među rezultatima dobivenih od različitih ispitivača – tada test *nije* objektivna, već su rezultati pod utjecajem ispitivača

→ vrlo slična pouzdanosti mjerenja, ali nije ista

→ neki ju udžbenici klasificiraju u tu kategoriju, no postoje jasni razlozi zašto ju treba tretirati kao samostalnu metrijsku karakteristiku

### Određivanje objektivnosti

- mjera slaganja različitih ispitivača
- primijenimo test, ocijeni ga više ocjenjivača → zanima nas koliko su ga oni slično ocijenili
- kvantitativna analiza objektivnosti: nekoliko pristupa
  - Korelacijski pristup
  - Razlike u ocjenama

### Korelacijski pristup

- Visoka korelacija između ocjenjivača, visoka objektivnost
  - stabilnost
  - isto kao u priči o pouzdanosti
- ali, korelacija pokazuje slučajne pogreške, ali ne i sistematske
- dva ocjenjivača mogu imati  $r=1$ , a da jedan uvijek daje 2 boda više zato nije dovoljan pokazatelj objektivnosti

### Osobna jednadžba ispitivača

- tendencija ispitivača da daje “bolje” ili “lošije” ocjene
- različiti ispitivači imaju različiti *kriterij ocjenjivanja*, tj. neki ispitivači imaju *blagi* kriterij, dok drugi imaju *stroži* kriterij ocjenjivanja
- sistematski varijabilitet

### Objektivnost vs. pouzdanost

- da bismo dobili objektivna test, moramo smanjiti utjecaj sistematskog i nesistematskog varijabiliteta
- upravo je sistematski varijabilitet ono što objektivnost mjerenja izdvaja iz pouzdanosti mjerenja → pouzdanost govori isključivo o nesistematskom varijabilitetu

### Faktori o kojima ovisi objektivnost

- *uvježbanost ocjenjivača* → neizvježbani ispitivač radi više grešaka i više djeluju i nesistematski i sistematski faktori
- *način odgovora na podražajnu situaciju* → najbolji su testovi sastavljeni od *zadataka višestrukog izbora* – takvi testovi su potpuno objektivni jer imaju jednoznačno određen ključ za ocjenjivanje
- *standardizacija uvjeta primjene testa* – sve specifikacije koje su navedene u priručniku za test koje opisuju uvjete u kojima test funkcionira najvažnije su odrednice → primjena testa treba biti standardizirana

### Ključ za ocjenjivanje

Popis točnih ili indikativnih odgovora ispitanika sa pridruženom kvantitativnom oznakom njihove simptomatske vrijednosti.

#### Potpuno objektivni test

- ima jednoznačno definiran ključ za ocjenjivanje
- svaki zadatak ima samo jedno točno rješenje (tada bi svaki ispitivač trebao dobiti isti rezultat)
- uzroci neobjektivnosti su banalne greške: npr. slučajno priznavanje netočnih odgovora i obrnuto, preskakanje pitanja itd. → tih grešaka nikako ne bi smjelo biti → da bi se te greške svele na najmanju moguću mjeru, bitno je svaki individualni učinak, radi provjere 2x korigirati i vrednovati

#### Nepotpuno objektivni test

- test s odgovorima otvorenog tipa ili projektivni testovi
- nema ponuđenih odgovora među kojima treba izvršiti izbor
- ocjenjivanje je podložno subjektivizmu i to ugrožava objektivnost

#### Standardiziranost testa

→ preduvjet za objektivnost (ali i za pouzdanost)

- preciziranje svih postupaka primjene nekog testa, te bodovanja i načina vrednovanja bruto rezultata
- standardizacija opisana u priručniku → neophodan pribor uz svako psihologijsko mjerenje, sastavni dio svakog testa

→ izjednačavanje uvjeta mjerenja za sve ispitanike

#### Standardizacija uključuje

NAČIN PRIMJENE TESTA → specifikacija testa kao individualnog ili grupnog

OPĆI MIKROKLIMATSKI UVJETI → osvijetljenje, temperatura, zračnost, veličina radnog prostora, potreban pribor i sl.

UPUTA → specifikacija upute ispitaniku – obrazloženje primjene testa, pojašnjenje načina rješavanja zadataka i načina davanja odgovora

ZADACI → svi ispitanici moraju rješavati iste zadatke

KLJUČ ZA OCJENJIVANJE → najvažniji pribor – popis simptomatski valjanih reakcija i točnih rješenja: popis točnih odgovora: testovi postignuća (znanja ili sposobnosti) popis indikativnih simptoma: testovi ličnosti

→ direktne upute kako interpretirati ispitanikove odgovore

FORMIRANJE UKUPNOG REZULTATA → određenje kako se formira ukupni rezultat u testu:

homogen test: zbroj, ponderirani ili ne

heterogen test: koji se zadaci zbrajaju u koji faktor, te ponderiraju li se neki od njih

VREDNOVANJE INDIVIDUALNIH REZULTATA : upute kako rezultate u testu prezentirati na način da budu međusobno usporedivi, proizlaze iz baždarenja testa

#### Poboljšavanje objektivnosti

Optički čitači

- nema banalnih "ljudskih" grešaka
- nemoguće kod zadataka otvorenog tipa ili kod projektivnih testova

Kompjuterizirana primjena testova

- nema ispitivača koji bi bio izvor pogrešaka.
- ne mogu se primjenjivati masovno, treba imati dvorane, neki ispitanici mogu imati problema sa radom na osobnom računalu

### Odnos objektivnosti, osjetljivosti & pouzdanosti

Objektivnost

- objektivnost testa osnovni je preduvjet za sve ostale metrijske karakteristike
- test mora biti objektivan da bi mogao biti pouzdan i valjan

Objektivnost – pouzdanost

- objektivnost nesistematski i sistematski varijabilitet
- pouzdanost nesistematski

Osjetljivost

- Preduvjet za pouzdanost, valjanost – niži varijabilitet, niža korelacija

### VALJANOST

Valjanost → svojstvo testa da mjeri ono što smatramo da mjeri

- problem indirektnih mjerenja u psihologiji

Mjerimo težinu

- vaga → nema dileme, mjeri težinu

Mjerimo inteligenciju

- test inteligencije → ?? znamo li mi da je to (upravo i isključivo) test inteligencije?
- pitamo se:
  - što taj test mjeri?
  - mjeri li baš inteligenciju?
- saznavanje odgovora na ta pitanja je ispitivanje valjanosti testa

#### Valjanost i pouzdanost

Valjanost → svojstvo testa da mjeri ono što smatramo da mjeri

Pouzdanost → svojstvo testa da daje stabilne i homogene rezultate određene pravim rezultatom, a ne pogreškom mjerenja

#### Odnos pouzdanosti i valjanosti testa

- rezultat = pravi rezultat + SV + NSV
- Samo u okviru pouzdane varijance možemo tražiti valjanu varijancu
- Logika: nešto možemo izmjeriti pouzdano, a da uporno mjerimo krivu stvar
- Test nužno mora biti pouzdan, kako bi mogao biti valjan
- Ne vrijedi obrnuto: test koji je pouzdan (stabilan, homogen), ne mora nužno biti valjan

#### Validacija instrumenta

Postupak utvrđivanja valjanosti tj. utvrđivanje što test mjeri

Potreban uvijek pri konstrukciji nekog testa, prijevodu (jezične finese, sinonimi...), adaptaciji (kulturološke razlike)

Bez njega nema opravdanja za primjenu testa

## Vrste valjanosti

Kad određujemo valjanost testa analiziramo 3 aspekta valjanosti

- o Sadržajna valjanost
- o kriterijska valjanost
- o konstruktna valjanost

### Sadržajna valjanost (content validity)

Aprioristička, teorijska

Pri konstrukciji testa procjenjujemo mjere li čestice testa zamišljeni konstrukt

To rade stručnjaci u određenom području

Najčešće prva faza u konstrukciji nekog testa

- polazi se od definicije, teorijskog okvira
- pobroje se sve manifestacije, simptomi
  - o jesu li zaista sve manifestacije simptomi konstrukta?
  - o jesu li svi simptomi nabrojani?
- biraju se važne manifestacije (važni simptomi)
- formiraju se čestice

→ sadržajna valjanost naziva se još i simptomatska

### Facijalna valjanost (prividna, pojavna, face validity)

- razlikuje se od sadržajne, ali joj je bliska
- sadržajna: stručnjaci, facijalna: ispitanici
- stupanj u kojem ispitanici vjeruju da test mjeri neku osobinu, da je relevantan
  - o test ekstraverzije – volite li više čitati ili biti u društvu → posjeduje facijalnu valjanost
  - o Rorschachov test mrlja → vjerojatno ne
  - o test iz Cosma → možda čak i da
- motivacijska uloga → *povjerenje* u test
- ako test ne posjeduje dobru facijalnu valjanost
  - o ispitanici ga mogu ocijeniti lošim, bezveznim, nepotrebnim
  - o negativan stav prema testiranju
  - o utjecaj na rezultate (čak iako je test dobar)
- dobar test bi najčešće trebao posjedovati i dobru facijalnu valjanost
- ali nekada to može biti i loše → npr. ako je evidentno da test mjeri neuroticizam, a mi ga primijenimo pri selekciji za posao → moguće laganje, iskrivljavanje odgovora

### Kriterijska valjanost (criterion-related validity)

- praktična valjanost
- ne zanimaju nas rezultati nekog testa sami zbog sebe, već zato da na osnovu tih rezultata nešto zaključimo
- zanima nas koliko je “dobar” neki test, možemo li ga upotrijebiti u praktične svrhe, tj.
  - o za prognozu ili
  - o za dijagnozu
- kriterijska valjanost je svojstvo testa da na osnovu njegovih rezultata možemo prognozirati ili dijagnosticirati rezultat ispitanika u nekoj kriterijskoj varijabli

### Analiza kriterijske valjanosti

- ispitujemo povezanost našeg testa sa vanjskim ponašanjem
- 2 skupa varijabli
  - o prediktori → testovi

- o kriteriji → realne mjere učinka izvan prediktora
- kriterij je standard prema kojem određujemo valjanost testa → objektivna mjera relevantnog ponašanja, nije test
- to može biti:
  - o broj izazvanih prometnih nesreća
  - o broj dana proveden u psihijatrijskoj ustanovi
  - o psihijatrijska dijagnoza
  - o dužina studiranja
  - o prosjek ocjena
  - o broj izostanaka s posla (u radnim danima)
  - o procjene uspješnosti na radu od nadređenih

### Dva aspekta kriterijske valjanosti

Dijagnostička valjanost (*concurrent validity*)

- možemo li na osnovi testa uspješno dijagnosticirati neku osobinu kod ispitanika
- često u kliničkoj situaciji

Prognostička valjanost (*predictive validity*)

- možemo li na osnovi testa uspješno predvidjeti pojavu, razvoj, neke osobine kod ispitanika
  - često u selekcijskoj situaciji
- osnovna razlika: vremenski interval

### Logika kriterijske valjanosti

- ako je ona niska – znači da rezultati našeg testa ne govore baš puno o onoj osobini koja nas zanima
- nas zanima – je li netko nadareni učenik, treba li netko prilagođeni školski program, ima li netko PTSP, hoće li netko biti dobar zaposlenik, hoće li netko biti dobar student, a ne kako je riješio test
- test je samo alat za spoznaju o razvijenosti te neke osobine ili svojstva
- ako je naš test povezan sa tom osobinom ili svojstvom, onda ok
- ako nije – test je neupotrebljiv

### Konstruktna valjanost (construct validity)

→ sud o tome mjeri li naš test zaista onu varijablu (onaj konstrukt) koji bi trebao mjeriti?

→ drugim riječima – konstruktna valjanost je dokaz o tome da naš test mjeri upravo željeni konstrukt

### Konstrukt

- konstrukt je pojmovna tvorevina u znanosti nastala sustavnom integracijom različitih podataka o nekoj pojavi na koju se odnosi
- ne može se neposredno opažati niti mjeriti
- ma realne i opipljive učinke koji upućuju na njegovo postojanje
- konstrukt je nastao (izmišljen) u okviru neke teorije da bi opisao neko konkretno ponašanje
- u psihologiji većina ključnih pojmova jesu konstrukti: inteligencija, učenje, anksioznost, zadovoljstvo poslom, ličnost, motivacija, samopoštovanje, samoefikasnost, interesi...

### Analiza konstruktna valjanosti

- provjera mijenjaju li se rezultati nekog testa u skladu sa teorijskim konceptom
- prema nekoj teoriji, postavimo hipoteze
- provjerimo ih i ako nismo dobili očekivane nalaze:
  - a) teoriju treba preformulirati
  - b) nacrt je upitan
  - b) test nije dobra mjera konstrukta

- ako znamo da je teorija ok i da nemamo metodološke greške u nacrtu, onda naš test nije dobra mjera konstrukta
- istovremeno provjera teorije i provjera testa

**Primjer.** Neke postavke Hollandove teorije izbor zanimanja koje mogu poslužiti za analizu konstruktne valjanosti upitnika interesa:

- o žene imaju izraženije interese prema društvenim zanimanjima/aktivnostima, a muškarci prema tehničkim
- o društveni tipovi više vrednuju socijalne vrijednosti, a manje individualističke, dok umjetnički tipovi više vrednuju kreativnost i samostalnost, a manje sigurnost, itd
- o interesi se strukturiraju tijekom srednje škole
- o ...

→ sve su to moguće hipoteze za provjeru konstruktne valjanosti

#### Provjeravanje hipoteza u analizi konstrukte valjanosti

- broj hipoteza koja se mogu postaviti na osnovu svake teorije je velik, baš kao i broj načina na koje se na njih može odgovoriti
- zato nema jednog jedinstvenog načina ispitivanja konstruktne valjanosti (isto tako nema niti jednog jedinog koeficijenta konstruktne valjanosti)
- ispitivanje konstruktne valjanosti je niz mogućih provjera koji pokazuju mijenja li se konstrukt u skladu sa teorijskim postavkama

#### Analiza konstruktne valjanosti

- različiti rezultati u testu za različite grupe ispitanika
- povezanost sa drugim testom koji mjeri isti konstrukt
- povezanost sa jednim varijablama, a nepovezanost sa drugim
- promjene rezultata u testu u funkciji dobi, trajanja treninga...
- adekvatna faktorska struktura testa

### BAŽDARENOST

Što nam znači broj bodova na nekom testu?

Ništa, ako ne znamo koliko su postigli drugi.

#### Norme

- norme testa su prosječne vrijednosti neke relevantne grupe ispitanika
- mogu biti izražene na više načina

#### Baždarenje testa

Baždarenje testa = proces određivanja normi

- o Test baždarimo na *normativnom uzorku* → to je grupa ispitanika čiji su rezultati korišteni kako bi se prema njima usporedili rezultati svih ostalih ispitanika
- o Normativni uzorak → određivanje normi
- o Standardizacijski uzorak → standardizacija testa

#### Standardizacija testa

Postupak primjene testa na reprezentativnom uzorku u svrhu definiranja postupka primjene testa te određivanja bodovanja i normi

#### Baždarenje - postupak

- o odabrati uzorak
- o primijeniti test
- o izračunati deskriptivnu statistiku
- o odrediti "vrijednost" svakog mogućeg bruto rezultata
- o to sve detaljno opisati u priručniku za test
- kvaliteta (preciznost, upotrebljivost) normi ovisi o tim opisima

#### Odabir uzorka

Treba definirati populaciju za koju je test namijenjen, pa odabrati uzorak

- o uzorak mora biti *reprezentativan* za tu populaciju: mora obuhvaćati sve važne subgrupe (spol, stupanj obrazovanja, regija stanovanja...) i dovoljno velik za zaključivanje npr.
- o reprezentativan uzorak odraslih građana RH
- o pacijenti sa dijagnosticiranom depresijom
- o radnici koji rade u smjenama
- o učenici 4-8. razreda osnovne škole
- o djeca-kandidati za upis u osnovnu školu

#### Primjena testa

- strogo određeni uvjeti
- lako ponovljivi
- u budućim primjenama uvijek isti

Postupak primjene testa vrlo detaljno opisati u priručniku za test – o tome ovisi usporedivost rezultata

#### Određivanje normi

- deskriptivna statistika: M i Sd i/ili medijan
- prikazati rezultate za sve relevantne subgrupe (što je relevantno, ovisi o namjeni testa), to mogu biti:
  - o spol
  - o dob
  - o stupanj obrazovanja (npr. završeni stupanj ili razred u školi)
  - o SES (socioekonomski status)
  - o red rođenja
  - o ljevorukost ... itd

#### Izražavanje rezultata

- o bruto rezultat  $X = 30 \rightarrow$  netko je postigao 30 bodova
- o norme za adekvatnu populaciju su  $M = 45$  i  $Sd = 7,5$
- o kako je on/ona riješio test?
  - nama je to razumljivo, ali drugi?

- prikazati i na drugačiji način → transformacija na druge skale da bude razumljivo i drugima

#### Transformacija rezultata na druge skale mjerenja

Prikaz rezultata u 'razredima', 'grupama', 'usporedivim bodovima' tako da oni sami po sebi nešto znače

- o Centili, decili
- o Z-vrijednosti
- o T-skala
- o C-skala...

### Decili

- Rezultati izraženi u grupama po 10
- Medijan je na polovici, od njega ima 50% viših i 50 % nižih rezultata
- Prvi decil – najnižih 10%, 90% viših rezultata
- Drugi decil –10% nižih rezultata, 80% viših rezultata
- Deseti decil – najviših 10% rezultata
- (ista logika za centile)

### Z-vrijednosti

$$z = (X - M) / sd$$

- raspon -3 do 3, M = 0, Sd = 1,  $\pm 1Sd = 68\%$ ,  $\pm 2Sd = 95\%$ ,  $\pm 3Sd = 99\%$
- vrlo razumljivo, nama
- ali kako to priopćiti klijentu (imate rezultat -0,7z) & rezultat 0, negativni rezultati
- zahtijeva normalnu distribuciju
- na osnovu z-vrijednosti proizašle su druge transformacijske skale

### T-skala

$$T = 10 z + 50$$

- raspon 1-100, M = 50, Sd = 10
- rezultat ispitanika T=1, 2, 3 ... 100
- prednosti: nema 0 i negativne rezultate, nema decimalne vrijednosti
- ograničenja: neosjetljiva na krajevima (npr.  $z = 3 \rightarrow T = 80$ ) & “prefina” – razlike od 1T najčešće nisu stvarne razlike
- često se koristi

### C-skala

$$C = 2 z + 5$$

- raspon 0-10, M = 5, Sd = 0,5
- rezultati ispitanika C=0,1,2...10 (11 intervala)
- otvoreni krajevi (npr.  $z = 3 \rightarrow C=10$ )
- prednosti: prihvatljiv raspon & razumna osjetljivost, nema decimalnih brojeva
- nedostaci: rezultat 0
- često se koristi

### Druge skale

- Stanine (Standard nine) – raspon 1 do 9, M = 5, Sd = 0,5
- Stan (Standard ten) – raspon 1 do 10, M = 5,5; Sd = 0,5
- Skala školskih ocjena – raspon 1 do 5, M = 3, Sd = 1,2
- Skala devijacionog kvocijenta inteligencije – raspon 55-145, M = 100, Sd = 15

### Sve skale za izražavanje rezultata

T, C, Stanine, Stan, Školske ocjene, i DQI

- zahtijevaju normalnu distribuciju
- nemaju decimalne vrijednosti
- najčešće korištene T i C