

Gospodu Zupanu Ljubljanskega  
dr. Adlešiču poklanja avtor.

DR. ZAJC DRAGO:

Dr. Zajc  
Lj. 8.3. 1940.

# **Problematika intelligentnostnih testov**

**Poizkus merjenja same razumnosti**

V LJUBLJANI 1940.

Založila knjigarna Kleinmayr in Bamberg

68645



0300 13132

Natisnila tiskarna Slatnar d. z o. z. (Vodnik in Knez) v Kamniku

## **Predgovor**

Vprašanje testov se je posebno s svojim pedagoškim pomenom uvrstilo v slovenski strokovni literaturi med najaktualnejša vprašanja.

Naši prosvetni in kulturni delavci sledijo nekaterim tujim testom in jih hočejo uporabljati za slovenskega otroka! Nov primer, kako se vsiljuje praksa pred teorijo.

Za smiselno uporabljanje testov in njih razumevanje namreč še nimamo primerne literature. Ta pa je potrebna, če naj pričakujemo, da bo za slovenske razmere standardiziran test prišel v pripravljene roke, kar je za uspeh *conditio sine qua non*.

Pričujoča razprava je poglavje iz večjega dela. Morem pa trditi, da bodo vprav misli tega poglavja tvorile prve temelje za razumevanje in razlaganje celotnega gradiva.

Predvsem gre za poznavanje inteligentnostnega testa, obenem pa še za razumevanje testiranja sploh. Iz tega razloga ima pričujoče delo še ta poglobitni namen, da poda tudi značilne psihološke osnove in pogoje za teoretično umevanje in praktično uporabljanje testološke metode. Imam namreč upanje, da bo po taki poti postal še tem vidnejši in učinkovitejši tudi pravi pedagoški pomen testološkega dela.

## Uvod

Že Platon je spoznal važnost individualnih sposobnosti in misel o individualnih razlikah je važen člen v njegovi veličastni zamisli o ureditvi in organizaciji države. Vprav v tem razumevamo tudi pomen psihologije kot prve pomožne vede pedagogike. Od nje prvenstveno zahtevamo odgovor na vprašanje, kakšno je človeško bitje, ki ga je treba vzgajati. Razumeli bomo zato, da vzbuja vsako prizadevanje v psihologiji, ki skuša z znanstvenimi metodami ugotavljati različne sposobnosti poedinca, največje zanimanje pedagoškega svetá.

Med najvažnejše probleme vzgojeslovja spada vprašanje o možnostih in mejah obrazovanja doraščajočega človeškega bitja. Upoštevati moramo tukaj dva činitelja: pedagoški milje v najširšem pomenu besede in sposobnosti poedinca. Ne da bi se podrobneje spuščali v probleme nativizma in empirizma, lahko trdimo, da velja sicer na eni strani načelo o vplivu okolja na izoblikovanost poedinca, da pa moramo v procesu vzgoje in obrazovanja človekove osebnosti upoštevati tudi individualne sposobnosti.

Zgodovina pedagogike nam kaže, da so se tega dejstva v večji ali manjši meri zavedali pedagogi vseh dob. Metoda psihološkega opazovanja je bila pri tem najvažnejša metoda za določevanje individualnih sposobnosti poedinca.

Razvoj otrokovih sposobnosti in stopnja njegovega celotnega duševnega razvoja se da namreč največkrat uspešno ugotoviti že s samim »opazovanjem«. Otrokovo vedenje in zunanji izraz ob najrazličnejših prilikah je predmet opazovanja, ki zavzame včasih tudi obliko preizkušanja. Zavedati pa se moramo, da bo opazovanje in preizkušanje poedinčevih spo-



sobnosti le tedaj uspešno, če poznamo poedinca že delj časa in če imamo pri tem tudi dar praktičnega opazovanja in intuitivnega pogleda v tujo duševnost, razen tega pa še tudi teoretično psihološko znanje. Opazovanje tuje duševnosti nas lahko tudi vara. Razumljivo je tedaj, da je človek že od nekdaj skušal izločiti subjektivnost opazovanja, in sicer z objektivnimi merili.

V drugi polovici 19. stoletja je vsled velikanskega napredka eksaktnih prirodoslovnih ved, ki so omogočile silni razmah tehnike, zavzel eksperiment vidno mesto tudi v psihologiji. Vse to so storili v upanju, da bomo preko merjenj, ki jih omogoča preizkus, prišli tudi v psihologiji do eksaktnih spoznanj o duševnih pojavih.

Vzporedno s tem se javlja stremljenje, omogočiti uporabo psiholoških izsledkov v praksi. Spočetka ni bilo dosti izgledov za uspeh. Psihološka veda se je proti koncu 19. stoletja ukvarjala s preveč abstraktnimi vprašanji in tudi eksperiment je služil predvsem za razjasnitev teoretičnih vprašanj. Vendar so se kmalu pojavili poskusi, kako bi rešili na eksperimentalni način tudi psihološke probleme iz praktičnega življenja.

Med prve take izrazito sistematske poskuse spada B i n e t o v o prizadevanje, da bi našel način za zanesljivo ločitev duševno normalno razvitih šolarjev od zaostalih, za katere bi bili potrebni pomožni razredi. Stavil si je nalogo, da s preizkusi določi za vsako starostno dobo seznam vprašanj, ki naj jih duševno normalno razvit otrok te starosti reši.

Ogromen razmah je ta praktična smer psihologije doživela v Ameriki. V ospredje je stopilo vprašanje, kakó lahko praktično preizkušamo in merimo poedinčeve sposobnosti.

Sestava in uporaba teh preizkušenj je dobila vedno bolj enotno obliko in končno se je tudi ustalilo ime za take preizkušnje — »test«, kar pomeni v tej rabi toliko kakor »preizkus«. Razvila se je mogočna veja uporabne psihologije — testologija. Njen praktični smoter je, da z raznimi preizkusi in to n a č r t n o ugotovi tiste poedince iz neke skupine, ki so za določen študij ali poklic (tehnični poklic, vojaško službo, šoferski poklic, strojevodjo itd.) najbolj sposobni.

Minilo je 40 let, odkar je A. Binet postavil metodi testov vogelni kamen. Ta veliki mož je odprl vrata novi možnosti ugotavljanja individualnih razlik. Začetek njegovega dela je bil skromen, zato pa tembolj temeljit. Binet je bil res na pravi poti in zato lahko razumemo, zakaj je v razmeroma kratkem času dosegel tako lepe uspehe. Ne smemo pa prezreti dejstva, da je Binetovo delovanje ustrezalo smeri celotnega razvoja dobe, v kateri je živel. Zamisel, preizkušati in meriti poedinčeve sposobnosti, izhaja namreč že iz prve eksperimentalno psihološke šole W. W u n d t a v Leipzigu.

V času, ko se je Binet vedno bolj uveljavljal, so nekateri njegovi vrstniki tudi mnogo pripomogli k razvoju psihološke metode testov. Med njimi moramo zlasti upoštevati E b b i n g h a u s a, ustanovitelja dopolnjevalnega testa (Ergänzungstest; completion test) in Amerikanca J. Mc. C a t t e l l a, kateremu pripisujejo zaslugo, da je leta 1890 prvi uvedel besedo »test« kot strokovni izraz, ki prej ni imel današnjega pomena. Končno moramo imenovati še Angleža Ch. S p e a r m a n a, ki je posegel v razvoj testov posebno s svojo razpravo »General Intelligence«, objavljeno v American Journal of Psychology 1904. Z njo je precej ostro začrtal meje in možnosti metode testov ter obenem opozoril na njih statistično zgradbo. Kasneje je zavzel Ch. Spearman vidno mesto pri preučevanju psihološkega ozadja inteligentnosti in s svojo znano teorijo »dveh činiteljev« je utrdil veljavnostno načelo teh testov.

Prvo večjo pozornost so vzbudili Binetovi testi, ko jih je objavil že leta 1899 pod naslovom »Attention et adaptation« (»Pozornost in prilagajanje«). Sicer sami záse niso pokazali večje vrednosti, so pa izpričali njegovo veliko praktično usmerjenost, na drugi strani pa tudi njegovo pedagoško izhodišče. Bolj izpopolnjeni so bili testi, ki jih je objavil leta 1905. Tu že nahajamo postopico od lažjega do težjega vprašanja; za enotno izračunavanje pa je bil prvič nakazan osnovni princip tkzv. »umske starosti« (âge mental). Učenci enake starosti, a z različno stopnjo inteligentnosti naj bi se po različno rešenih testih tako razlikovali, kot da bi bili različno stari. V tej zvezi je Binet skušal določiti, katere sposobnosti so na pr. v 3., 5. ali 7. letu starosti že toliko razvite, da jih lahko tudi uspešno



preizkušamo. Na osnovi tega je sestavil lažja vprašanja, ki naj sestavljajo test. Uredil jih je v progresivno postopico, ki jo je imenoval lestvico.

Vse te zamisli in celotni osnutek je izpeljal skoro do kraja v svoji razpravi, ki jo je objavil leta 1908. Lestvica testov je segala od 3. do 13. leta, število vprašanj in nalog v vsakem testu pa je bilo zelo različno. Binetova objava je na mah prodrla v znanstvenem svetu in oglašale so se že prve kritike, katerim gre zasluga, da je svojo lestvico revidiral. V predelani obliki jo je objavil tik pred svojo smrtjo leta 1911. V bistvu ta revizija ni spremenila prvotne zamisli, temveč jo je le poglobila z zamenjavo nekaterih neprimernih testov in z izenačenjem števila vprašanj v testih.

Binetovo delo je dobilo po vsem kulturnem svetu popolno priznanje. Njegove teste so uporabili mnogi psihologi ter jih priredili za svoje potrebe in razmere.

V Binetovem delu samem pa so bili položeni temelji za uspešen nadaljnji razvoj takó inteligentnostnih kakor tudi drugih vrst testov.

Poznamo mnogo Binetovih testov, ki zahtevajo za svojo rešitev predvsem »znanje« testiranca. Tudi sam se je tega že zavedal in skušal je ta moment izločiti, vendar se je to vprašanje v polni meri rešilo šele v teku razvoja. V tem prizadevanju so osnove za nastanek posebnih učnih testov (občih za obče znanje — educational test; in posebnih za določen predmet — achievement test).

Razvoj točkovne lestvice (point scale), ki je prvotno vsebovala vse Binetove teste, moramo prav tako prišteti med plodove njegovega dela. Točkovna lestvica, ki ni preosnova Binetove koncepcije v principu, ampak le v metodi, ne postavlja za normo tako ali drugačno umsko starost, temveč le določeno število točk, ki naj bi bilo značilno za to ali ono starostno stopnjo.

Pri nadaljnjem razvoju v tej smeri so nastali skupinski testi (group test), ki so se izkazali kot uspešna in nenadomestljiva vrsta istočasnega preizkušanja inteligentnosti večjih ali manjših skupin poedincev.

Metode za merjenje inteligentnosti so se vedno bolj izpopolnjevale ter so bile skupno z razvojem oblik teh testov najboljše izhodišče za razvoj še drugih vrst testov. Tako so polagoma nastali testi za temperament, dispozicijo, značaj, za strukturo osebnosti in drugi.

Za celokupni razvoj testne psihologije ima s teoretičnega in praktičnega vidika mnogo zaslug na eni strani razvojna psihologija, na drugi pa statistika. V tej zvezi naj omenimo ime Slovenca Franca Žižka, svetovno znanega statistika, ki je do nedavnega deloval na univerzi v Frankfurtu in se je proslavil zlasti z delom »Grundriss der Statistik« (ki bi se z vso upravičenostjo lahko nazivalo »Teorija statistike«).

Ob inteligentnostnih testih se še danes razvija vsa glavna problematika testne psihologije. Načelna vprašanja, ki zadevajo preizkušanje inteligentnosti, so najboljša osnova za razumevanje metodičnih osnov preizkušanja s testi. V naslednjih izvajanjih bomo skušali prikazati osnove testologije, kakor tudi njene možnosti in meje. Ob problemih, ki se nam postavljajo ob testiranju inteligentnosti, bomo tako dobili vpogled v praktično in teoretično veljavo testiranja.



## I. DEL.

### **Inteligentnost z vida testne psihologije**

#### *§ 1. Pojem inteligentnosti.*

Trije značilni načini človeške dejavnosti izvirajo iz instinkta, dresure in intelekta. Instinktivna dejanja označuje nekak avtomatizem. Živo bitje, žival ali človek, jih prinese s seboj na svet in v določeni situaciji potekajo vedno enako; n. pr. komaj zagleda mlado pišče luč svetá, že kljuva po hrani; rojenček »zna« že takoj po rojstvu sesalne gibe z ustnicami itd. Čim nižje gremo po razvojni lestvici živalskih vrst, tem večji pomen in razmah dobivajo instinktivna dejanja. Pri žuželkah dobimo naravnost čudovite sisteme instinktov.

Pri višje razvitih živalskih rodovih opazamo možnost, da se prirojeni načini dejavnosti prilagajajo raznim življenjskim situacijam. Nastopa že določena mera učenja — dresura, ki jo mnogokje tudi človek izkorišča v svoje namene. Takó izkoriščamo na pr. lovski instinkt psa za to, da nam pomaga pri lovu. Z vajo smo prvotno instinktivno dejanje izoblikovali dalje in ga prilagodili novim okoliščinam. Vendar se tudi v dresuri javlja samo nek avtomatizem.

Šele razvojna stopnja intelekta ali uma omogoča živemu bitju, da obvlada novo situacijo na izviren način. Na pr. kadar človek v življenjski situaciji, ki je zanj popolnoma nova, na osnovi »razmišljanja« najde primeren izhod.

Za človeško dejavnost je značilno predvsem delovanje intelekta, ki ga pa tudi tu mogočno podpirata še instinkt in dresura (navada). Vprav zato je razvojna stopnja intelekta tista, ki daje vsemu človeškemu dejstvomvanju še poseben

»človeški« pečat. Sposobnost, ki prihaja do izraza v dejavnosti intelekta, je inteligentnost.

Inteligentnost prihaja do izraza v vseh umskih storitvah in je osrednja zmogljivost za umske funkcije. Točneje jo lahko določimo z dveh vidikov, z vidika introspektivne psihologije in z vidika ameriške behavioristične struje.

Po prvem vidiku so psihologi označili inteligentnost kot sposobnost prilagajanja na nove življenjske prilike in zahteve; označili so jo nadalje za »obrazljivost«, ker so jo smatrali za ekvivalent umske zmogljivosti itd. V vsakdanjem življenju pomeni inteligentnost posebno visoko stopnjo umske zmogljivosti, ki se označuje tudi z izrazi, kakor: »bistrost«, »brihtnost« in dr.

Take označbe inteligentnosti morejo biti sicer psihološko pravilne, a ne zadostujejo, da bi jo mogli praktično raziskovati in meriti. Praktično določevanje in merjenje inteligentnosti zahteva drugačen vidik.

Zasluga behaviorizma je, da je omogočil teoretično utemeljeno praktično preučevanje inteligentnosti in s tem našel nove vidike tudi za introspektivno psihologijo, ne glede na to, da jo radikalni behaviorizem v celoti odklanja<sup>1</sup> (na primer Watson).

Umska zmogljivost nam je praktično dostopna le po svojem javljanju v konkretnih umskih storitvah. Kar res lahko opazujemo in celo merimo, to je le javljanje inteligentnosti v določenih storitvah, ne moremo pa neposredno meriti inteligentnosti same.

Behaviorizem se dosledno omejuje na raziskovanje tega, kar se da »od zunaj« opazovati, ugotavljati in meriti. Zato je

---

<sup>1</sup> Že K. Bühler je opozoril v svojem delu »Die Krise der Psychologie«, da se morejo različne psihološke struje med seboj koristno izpopolnjevati, čeprav so včasih zgrajene na teoretičnih osnovah, ki si že načelno nasprotujejo. V tej zvezi naj še dodam, da je znani introspektivni vidik tudi za testologijo iz načelnih razlogov neobhodno potreben. Zakaj brez tega vidika bi nam samo (vnanje) »testiranje« nikdar ne moglo pomagati do ugotavljanja notranjih (psihičnih) pojavov in sposobnosti.



razumljivo, da preučuje vsako sposobnost in tudi inteligentnost le toliko, kolikor se nam prikazuje kot aktualni realni činitelj v okviru nekega konkretnega dejstvovanja. »Sposobnost« poedinca ugotavlja in meri omenjena struja samo po konkretni storitvi.

Ugotovitev tega posebnega gledanja na človeške sposobnosti je važna radi tega, ker je to tudi teoretična osnova ameriške testne psihologije, čeprav v najnovjšem času že opažamo vplive introspektivnega stališča.

Ta osnovni vidik testne psihologije nam obenem takoj pokaže glavno pomanjkljivost preučevanja inteligentnosti po njeni storitvi. Zaprta nam je pot do vseh tistih subjektivnih činiteljev, ki nastopajo v naši duševnosti kot gibajoče in usmerjajoče sile in ki šele omogočajo, da se obča umska zmogljivost sploh »vpreže« in aktualizira v konkretni dejavnosti in storitvi.

Praktično določevanje inteligentnosti po metodah testne psihologije nikakor ni enostavno. Največje teoretične težkoče nastajajo vsled dejstva, da ne moremo natančno povedati, ali in koliko je neka storitev res izraz obče umske zmogljivosti. V konkretnih umskih storitvah, kakršne so sklepanje, kombiniranje, presojevanje itd., se javlja obča zmogljivost vedno v okviru čisto določene situacije, ki dovoljuje razmah obče zmogljivosti le v večji ali manjši meri.

Višina obče zmogljivosti, kolikor jo lahko izmerimo po njenem javljanju v konkretni storitvi, je zavisna torej tudi od posebnih zmogljivosti, ki so potrebne za premagovanje vprav te in te konkretne situacije.

Po psihološkem spoznanju sklepamo, da se javlja inteligentnost v teh in teh določenih umskih storitvah. Če naj govorimo o merjenju višine teh storitev, tedaj se dobro zavedamo, da iz tega še ne moremo neposredno sklepati na višino obče umske zmogljivosti. Preko izmerjene višine storitve, n. pr. v sklepanju, kombiniranju, presojevanju itd., lahko pač neposredno pridemo do zmogljivosti za sklepanje, kombiniranje in presojevanje v tej ali oni določeni situaciji, ne pridemo pa neposredno do višine inteligentnosti same.

## § 2. Teorija »dveh činiteljev« — Spearman.

Sprevideli bomo, da nam je za določevanje višine obče umske zmogljivosti ali inteligentnosti odprta le posredna pot. Vemo, da je za konkretno umsko storitev potrebna neka obča zmogljivost, prav tako pa tudi posebna zmogljivost, ki ustreza premagovanju vprav te konkretne situacije. Zato lahko ob ugotovitvi matematičnih korelacij med številčno izraženimi posameznimi umskimi storitvami tudi čisto matematično določimo v vsaki posamezni storitvi neko mero občega in neko mero posebnega činitelja.

Ako poznamo številčne izraze za višino večjega števila posameznih umskih storitev, lahko torej z matematičnimi operacijami številčno določimo, koliko so te posamezne storitve izraz obče zmogljivosti in koliko posebne.

S praktičnega vidika se je približal temu problemu že Anglež Ch. Spearman, profesor na londonski univerzi, s svojo teorijo »dveh činiteljev« (The Two-Factor Theory). To teorijo moramo razlagati tako, da je vsaka naša umska storitev izraz dveh vrst zmogljivosti, »obče« in »posebne«. S tem pa dobi ta teorija tudi drug okvir. Ob številčno izraženi storitvi lahko javljanje te obče in posebne zmogljivosti matematično izrazimo z dvema posebnima činiteljema, ki jih je Spearman imenoval »general factor« (obči činitelj) in »specific factor« (posebni činitelj). Kratica za prvega je »g«, za drugega pa »s«; tako ju uporabljam tudi v pričujoči razpravi.

Če hočemo pravilno razumeti Spearmanovo teorijo »dveh činiteljev«, moramo upoštevati teoretično osnovo njegovega psihološkega gledanja, ki je zasidrana v behaviorizmu. Ta struja se že načelno ne ukvarja s pojavi doživljanja, sposobnosti, dispozicij, z vprašanji, s katerimi se peča introspektivna psihologija, temveč se zavestno strogo omejuje na to, kar lahko ob dejavnosti živega bitja opazujemo »od zunaj«, na njegovem »obnašanju«. Zato tudi višina činitelja »g« v Spearmanovem smislu še ne pomeni višine inteligentnosti, ki jo ima ta ali oni poedinec, ampak pomeni samo višino ali količino tiste zmogljivosti, ki je dejansko aktualizirana v dejavnosti, katero smo preizkusili.



Nastane sedaj vprašanje, kaj sta prav za prav ta dva činitelja »g« in »s« in v kakšnem razmerju je činitelj »g« do inteligentnosti.

Inteligentnost smatramo za občo umsko zmogljivost, ki je potrebna za uspešne umske storitve. Do pojma take obče zmogljivosti nas dovede psihološka analiza (introspektivni vidik!). Višine te zmogljivosti neposredno seveda ne moremo »meriti«.

Činitelj »g« pa moramo smatrati za neko računsko dobljeno količino, ki smo jo dobili iz rezultatov preizkušnje, to se pravi, po neki storitvi. Pove nam, v koliki meri je udeležena obča umska zmogljivost pri vsaki posamezni od večjega števila umskih storitev, ki jih med seboj primerjamo. Činitelj »g« je vedno izražen s številom manjšim kot 1, kar pomeni, da preizkušnje konkretnih umskih funkcij ne bi smeli v celoti smatrati za preizkušnjo obče umske zmogljivosti — inteligentnosti. To število nam obenem pove, s kolikšno verjetnostjo merimo pri preizkušnji neke določene umske funkcije (na pr. sklepanja) občo umsko zmogljivost.<sup>2</sup> To pomeni: če smo ugotovili pri dveh različnih poedincih razliko v storitvi ob neki umski preizkušnji, velja samo z določeno verjetnostjo, ki je enaka »g« (občemu činitelju), da ugotovljena razlika res ustreza razliki v inteligentnosti.

Razen obče umske zmogljivosti so namreč za vsako konkretno umsko storitev potrebne še najrazličnejše posebne sposobnosti, ki sploh šele omogočajo razmah obče umske zmogljivosti v konkretni situaciji. Te posebne sposobnosti je določil Spearman v okviru storitve s činiteljem »s«. Razlagal si ga je z elementarnimi sposobnostmi, ki slonijo na fiziološki podlagi ter so odvisni od čutnih in gibalnih organov, in potem še z raznimi specialnimi sposobnostmi, ne gledé na to, ali jih lahko točno definiramo ali ne.

Činitelj »g« je torej računsko dosegljiv korelat inteligentnosti. Pri praktičnem delu smatramo, da zavisí storitev toliko od obče umske zmogljivosti, kolikor lahko v njej ugotovimo

<sup>2</sup> F. C. Thomas: Ability and Knowledge, London, 1935, str. 290.  
C. S. Spearman: Abilities of Man, London, 1927, str. 147.

činitelj »g«. Obči činitelj neke storitve nam torej v najboljšem primeru pove, v koliki meri se je pri tej storitvi izrazila inteligentnost. S tem, da smo ugotovili višino činitelja »g« v neki dejavnosti, smo na nek način izmerili samo aktualno občo umsko zmogljivost, kolikor se je izrazila v tej konkretni dejavnosti.

McDougall zato pravi, da je obči činitelj neke storitve, činitelj »g«, predvsem izmerjena sposobnost poedinca, ki je potrebna, da uspešno zberemo razpoložljivo energijo za določeno storitev, »it is rather the power of individual to concentrate his available energy effectively upon the task in hand«.<sup>3</sup>

Intelligentnosti torej nikakor ne smemo istovetiti s činiteljem »g«, vendar lahko rečemo, da so ugotovitve, ki veljajo za činitelj »g«, praktično v veliki meri veljavne tudi za inteligentnost, kakor jo pojmujejo z vidika introspektivne psihologije.

Iz tega je razvidno, da je inteligentnost praktično merljiva le toliko, kolikor prihaja do aktualnega izraza. Druga omejitev pa je v tem, da lahko izmerimo inteligentnost poedinca le v odnosu do skupine, kateri pripada. Zato je potrebno, da izvršimo večje število preizkusov že za samo sestavljanje testa. Izgotovljen test pa znova preizkusimo z ozirom na to, ali ustrezajo njegovi rezultati zakonom statistike. Od preizkušanja sploh ne pričakujemo absolutnih izsledkov, temveč samo relativne; skušamo namreč ugotoviti le mesto poedinca v določeni skupini. Tembolj se moramo zavedati, da »merjenje« v tem primeru sploh ni pravo merjenje, ampak samo neke vrste klasifikacija in razporeditev. Nimamo nobenih enot inteligentnosti, s katerimi bi lahko neposredno izmerili množino poedinčeve inteligentnosti.

### § 3. Razmerje med občim in posebnim činiteljem »g« in »s«.

Činitelj »g« je pri različnih poedincih kvantitativno različen, saj je razumljivo, da imajo eni več občega umskega talenta, drugi pa manj. Če nastopa »g« v visoki meri, tedaj je

<sup>3</sup> »Character and Personality« Dec. 1934, str. 134.



s tem dana najboljša podlaga za razvoj umskih sposobnosti, ki omogočajo uspešno delovanje v posameznih umskih storitvah. Enako velja obratno, da bomo zapazili pri duševni zaostalosti predvsem zelo neznatno mero »g«.

S tem, da je Spearmanova šola tudi matematično (s statistiko) z veliko verjetnostjo dokazala existenco činitelja »g« in »s«, je bila na osnovi statističnega gradiva odprta pot k psihološki razlagi prirode činitelja »g«.

Skrbna opazovanja in preučevanja kažejo doslej na to, da nobena vaja ne doprinese empirično zaznatnih pozitivnih sprememb k razvoju in povečanju činitelja »g«. Poizkusi so namreč pokazali, da je višina činitelja »g« prirojene narave in je z vajo ne moremo povečati.<sup>4</sup> Iz tega sledi, da inteligentnosti same ne moremo z vajo izboljšati.

Odziv ali odgovor na situacijo občega umskega testa bi torej v smislu povedanega določevali z dvema činiteljema ali dvema skupinama činiteljev. Eden izmed teh činiteljev je skupen vsem storitvam in tega bi po Spearmanu označili s činiteljem »g«. Činitelj »s« pa je bolj ali manj specializiran ter se razlikuje pri vsaki testni situaciji. Različne testne situacije zahtevajo različne stopnje »g«, oziroma različne stopnje aktualizirane obče umske zmogljivosti. Pri nekaterih funkcijah lahko ugotovimo, da so zavisne v večji meri od činitelja »g«, druge pa zopet od činitelja »s«.

Nastane vprašanje, v kakšnem medsebojnem razmerju se ta dva činitelja javljata v neki storitvi. To razmerje moremo s testi in statističnimi metodami bolj ali manj točno ugotoviti. Spearman piše v svojem delu »Abilities of Man« (str. 75), da je na enem koncu lestvice talent za duhoslovne vede (talent for classics), ki zahteva največjo mero činitelja »g«. To se pravi, ta činitelj »g« je za to sposobnost mnogo važnejši, vplivnejši in ga ne more nadomestiti še tako izpopolnjeni in njemu ustrezajoči činitelj »s«. Razmerje med tema dvema činiteljema je Spearman izrazil s številčnim razmerjem 15 : 1, kar pomeni, da je za duhoslovne vede potrebno 15krat več obče inteligentnosti kot pa posebnih sposobnosti. Na drugem koncu

<sup>4</sup> F. C. Thomas: Ability and Knowledge, London, 1935, str. 37.

lestvice, ki kaže manjšo vplivnost činitelja »g« pri neki sposobnosti, se nahaja sposobnost za glasbo (reproduktivno!). Tukaj, je Spearman dobil razmerje 4:1, kar pomeni, da je za reproduktivno glasbo potrebno le 4krat več obče umske zmogljivosti, kakor pa posebnih zmogljivosti.

Ta razmerja seveda ne pomenijo, da je človek, ki se ukvarja z reproduktivno glasbo, približno 4krat manj inteligenten od nekoga, ki se ukvarja z duhoslovnimi vedami. Z razmerji hočemo le povedati, da normalne dejanske storitve človeka, ki se peča z duhoslovnimi vedami, zajamejo (»vprežejo«) inteligentnost mnogo bolj, kakor normalne dejanske storitve reproduktivnega glasbenika. O dejanski inteligentnosti stopnji prvega ali drugega s tem še ni nič povedano.

Pri vsem tem moramo opozoriti, da so naša spoznanja na tem področju še v razvoju in da jih ne moremo smatrati za dokončna. Iz gornjega posnemamo le, da nastopata činitelja »g« in »s« v posameznih storitvah v zelo različnih medsebojnih razmerjih. Če hočemo te odnose točneje določiti, je potrebno, da izvršimo veliko število preizkusov. Šele tedaj dobimo tako ali drugačno razmerje, ki pa je le povprečno razmerje in velja zato za vso skupino, ki smo jo preizkusili, ne pa za določenega poedinca.

Akoravno velja, da moremo smatrati konkretno umsko storitev za zavisno od činitelja »g« in »s«, se lahko glede dejanske količine enega ali drugega činitelja motimo. Poedinec z visoko stopnjo inteligentnosti je verjetno bolj sposoben reševati probleme iz matematike, kakor drug z manjšo stopnjo; doseže pa lahko enako storitev, če ima zelo dober spomin za matematične simbole. Na drugi strani pa visoka inteligentnost zopet ne more mnogo pomagati pri kompenzaciji nerazvitih posebnih sposobnosti, kot na pr. v glasbi.

Če sta dva poedinca dosegla pri testiranju z inteligentnostnim testom enak rezultat, s tem še ni rečeno, da sta činitelja »g« in »s« pri obeh enaka. Pri enem je lahko »g« večji, pri drugem manjši in sorazmerno se razlikujeta tudi oba činitelja »s«. Možno je tudi, da sta dva poedinca dosegla različne rezultate in je pri obeh storitvah enaka količina činitelja »g«. Rezultat se tedaj razlikuje vsled različnosti činiteljev »s«.



Za primer vzemimo dopolnjevalni test (Ergänzungstest). Rezultati te preizkušnje zavisijo deloma od inteligentnosti, deloma pa tudi od razumevanja in obvladanja jezika. Prvo se kaže v storitvi kot činitelj »g«, drugo pa kot činitelj »s«. Če dosežeta dva poedinca enak rezultat (n. pr. enako število točk), moreta biti oba kljub temu po stopnji svoje inteligentnosti različna. Dejstvo, da sta doseгла enak rezultat, lahko izvira iz tega, da manj inteligentni bolje obvlada jezik. Različni rezultat dveh poedincev pa si moremo razlagati različno: tako, da sta

- a) različna činitelja »g« in različna »s«, ali da sta
- b) različna činitelja »g« in enaka »s«, ali pa da sta
- c) enaka činitelja »g« in »različna »s«.

#### § 4. Teorija skupinskih činiteljev.

Spearmanova teorija, ki je uvedla pojem generalnega činitelja v testno psihologijo, ni ostala brez nasprotnikov. Če namreč med posameznimi konkretnimi umskimi storitvami ne bi ugotovili neke smiselne zveze in s tem tako rekoč že naprej postavili domnevo o neki obči umski zmogljivosti, tedaj tudi po matematični poti ne bi mogli ugotoviti nekega občega činitelja v posameznih umskih storitvah.

Zato je n. pr. ameriški psiholog Thorndike še do nedavna zanikal obstoj neke obče umske zmogljivosti, češ da je naš razum skupek (množica) visoko razčlenjenih (diferenciranih) in nezavisnih sposobnosti (»a host of highly particularised and independent faculties«)<sup>5</sup>. Vendar pa je smatral, da vsaka od teh posebnih sposobnosti zavisí od lahkote učenja.

To pojmovanje inteligentnosti pa ima svojo omejitve. Če namreč res ne bi bilo občega činitelja, ki naj bi bil skupen vsem umskim storitvam, tedaj bi morali kljub temu še pri vsaki posebni vrsti storitve ločiti činitelje »s« in njim ustrezajoči činitelj »g«. Vprav ta misel je pokazala novo pot k nadaljnjemu preučevanju naših umskih sposobnosti.

Anglež G. d. Thomson je bil prvi, ki je ugotovil med posebnim in občim činiteljem še tako imenovani skupinski čini-

<sup>5</sup> Thorndike E. L. Educational Psychology, str. 31. New York 1903.

telj (group factor). Ta naj bi bil nekak »obči« činitelj v ožjem pomenu besede, ki je skupen samo nekaterim skupinam storitev posebne vrste. Nadalje naj bi bil ta skupinski činitelj matematičen izraz za občo sposobnost v ožjem pomenu, ki je potrebna za kompleksno storitev ali za storitev, ki je sestavljena iz več delnih posebnih storitev.

Ko je Thomson šele razvijal svojo teorijo skupinskih činiteljev, je naglašal, da si moramo naše sposobnosti razlagati samo po posebnih in skupinskih činiteljih. Podoba je bila, da se bo moral Spearman tej teoriji ukloniti, toda preizkusi so potrdili, da številni skupinski činitelji ne morejo imeti pravega občega značaja in zato ne morejo zavzeti povsem istega mesta kakor obči činitelj. Skupinski činitelji, ki bi jih še drugače lahko imenovali »zastopnike« nekaterih skupin posebnih sposobnosti, torej načelno še ne nasprotujejo zamisli obče sposobnosti.

Profesor Hamley je ugotovil sedem takih skupinskih činiteljev, ki nastopajo kot nekaka obča zmogljivost v ožjem pomenu, to je v okviru nekaterih skupin storitev. Ti skupinski činitelji so: verbalni (verbal), številčni (number), spominski (memory), prostorno dojemalni (spatial concept), mehanični (mechanical), hitrostni (speed) in sklepalni (reasoning).<sup>6</sup>

N. pr. skupinski činitelj za spominske storitve lahko ugotovimo v vseh vrstah spominskih storitev, kjer gre za to, da si neko gradivo ohranimo v spominu. Te storitve se sicer med seboj razlikujejo z ozirom na gradivo — ker je potrebna nekoliko drugačna usposobljenost, da si ohranimo v spominu števila ali matematične simbole, in drugačna, da si ohranimo melodijo, vendar imajo nekaj skupnega, kolikor gre pač za usposobljenost, da si sploh kaj obdržimo v spominu.

Na osnovi dosedanjih izvajanj lahko torej ugotovimo, da moramo pri vsaki konkretni umski storitvi upoštevati troje (najmanj pa dvojje) vrst činiteljev, obči činitelj, skupinske in posebne činitelje. Ti činitelji nam pomenijo v konkretni umski storitvi: aktualizirano občo umsko zmogljivost ali inteligentnost, aktualiziramo skupinsko zmogljivost, od katere je zavisen

<sup>6</sup> Hamley H. R. Testing Intelligence, London, 1936.



določen kompleks storitev, in končno aktualizirane specialne zmogljivosti. Kolikor so te sposobnosti res aktualizirane, jih lahko tudi merimo v posebnih preizkušnjah.

### § 5. O prirodi in funkciji inteligentnosti.

Spoznanje, da je inteligentnost obča in osrednja umska zmogljivost, ki prihaja do izraza v vseh umskih storitvah, nam še ne pove ničesar o prirodi inteligentnosti. Prihajamo do vprašanja, kaj so tipične storitve obče umske zmogljivosti. Z vida testne psihologije se lahko približamo temu problemu na ta način, da skušamo določiti, katere vrste storitev so odvisne v največji meri od samega činitelja »g«.

Spearmanova šola je ugotovila, da je zavisen uspeh toliko od činitelja »g«, kolikor je zavisna neka storitev od izvajanja ali edukcije (ki jo skušam spodaj pojasniti). Dejavnost pri izvajanju določenih odnosov se nam torej prikazuje kot tipična storitev, ki je zavisna predvsem od obče umske zmogljivosti ali inteligentnosti. Z analizo te dejavnosti bomo torej dobili obenem vpogled v prirodo inteligentnosti.

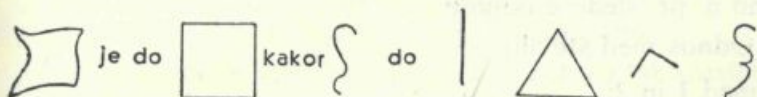
Pri inteligentnostnih testih zapazimo za preizkušanje sposobnosti izvajanja (ali edukcije) naslednje naloge:

1. naloge, kjer gre za razumevanje odnosne edukcije:

a) v vrsti petih likov moramo prečrtati lik, ki ne spada v to vrsto.



b) pokazati je treba, po katerem izmed zadnjih štirih likov je možno dobiti analogijo.



2. Navedemo lahko tudi naloge, kjer je treba izvajati odnose iz reprodukcije po primerjavi:

a) ali sta pojma »močan« in »šibek« enaka ali različna po pomenu?

- b) kakor je »belo« v razmerju s »črnim«, tako je »dan« v razmerju z jutrom, večerom, nočjo? (Pravilni odgovor je treba podčrtati).

$$A > B, B > C, A : C?$$

Rešitev teh nalog mora biti sama po sebi jasna in razvidna in izvajanje (ali edukacija) ne sme izvirati iz pridobljenega izkustva, oziroma pridobljenega znanja. Preizkušanec mora priti do rešitve popolnoma sam, tako rekoč čisto »na novo«, in mora biti o pravilnosti svojega izvajanja tudi prepričan.

Pri izvajanju odnosov v danih nalogah igra važno vlogo različna značilnost (ali pomembnost) mogočih odnosov. Z ozirom na postavljeno nalogo za pravilno rešitev niso enako značilni ali pomembni vsi mogoči odnosi, ki bi se dali iz danega gradiva izvajati. Pri izvajanju gre za odkrivanje značilnih odnosov, ki nam omogočajo pravilno rešitev dane naloge.

Kaj pa je prav za prav značilnost odnosov in kako prihajamo do značilnih odnosov?

Oglejmo si ta problem ob sledečem primeru!

Naloga se glasi:

Dodaj številčni postopici 1, 2, 4, 7, 11, 16, 22 naslednji dve števili, ki sledita po notranji zakonitosti te postopice!

Med posameznimi členi navedene postopice moremo izvajati najrazličnejše odnose, gre nam pa za to, da najdemo vprav te, na osnovi katerih lahko dopolnimo postopico z nadaljnji členi. Ker vemo, da mora biti med posameznimi členi postopice neka zakonitost, in sicer zakonitost postopice, bomo skušali določiti tiste odnose med posameznimi členi, ki ustrezajo gornjim zahtevam. Med posameznimi členi lahko izvajamo n. pr. sledeče odnose:

Odnos med števili:

I.) med 1 in 2:

- a) da sta si med seboj kot polovica do celote,
- b) da drugo število neposredno sledi prvemu v naravni številčni vrsti,
- c) da je razlika med obema 1.



II.) med 2 in 4:

- d) da sta si med seboj kakor koren in kvadratna vrednost števila ( $2^2 = 4$ ;  $\sqrt{4} = 2$ ),
- e) da je prvo število po vrednosti polovica drugega,
- f) da drugo število ne sledi neposredno prvemu po naravni številčni vrsti, kakor pri prvih dveh številih (1 in 2),
- g) da je razlika med njima enaka 2.

III.) med 4 in 7:

- h) da sta si med seboj manj kakor polovica do celote,
- i) da drugo število ne sledi neposredno prvemu v smislu naravne številčne vrste,
- j) da se razlikujeta med seboj za 3.

Iz teh mogočih odnosov moramo sedaj odkriti tiste, ki so v odnosu do zakonitosti postopice in ki ustrezajo zahtevam naloge. Samo ti odnosi so za nas v danem stanju pomembni in značilni. Odnose, ki jih navajamo pod c), g) in j), lahko postavimo še v posebne medsebojne odnose. Tako bo naslednji odnos v tej navedeni številčni vrsti tisti, katerega razlika bo znašala 4. Medsebojni odnos nastopa sicer tudi med odnosoma pod a) in pa pod e), to je med številoma 1 in 2 ter 2 in 4; seveda odpade tak odnos med 4 in 7 (pod h) in zato po tej poti ne pridemo do nadaljnjih členov pravilne postopice.

Pri izvajanju gre torej za odkrivanje takih odnosov, ki so pomembni z ozirom na usmerjenost ali nakazanost našega iskanja, kakor je to pač določeno po nalogi, ki jo moramo rešiti. Da sta si dva pojavi v nekem specifičnem odnosu, ne bi mogli razbrati, če ne bi bilo pri našem iskanju tega odnosa neke nakazanosti.

Z dano nalogo je smer našega izvajanja odnosov že določena in naperjena na iskanje čisto določenih odnosov.

Dejavnost dejanskega izvajanja pa ne poteka tako sistematično, kakor bi se dalo sklepati iz razloženega primera postopice. Konkretno izvajanje se ne vrši na ta način, da bi ugotavljali najprej vse možne odnose in nato izmed njih izbrali tiste, ki so v kakšni zvezi z dano nalogo. Tudi se nam ne zdijo na prvi pogled značilni, da bi končno iz njih izluščili tiste odnose,

ki šele omogočajo rešitev problema ali naloge. Tak postopek uporabljamo le za kontrolo in dokazovanje.

Praktično se vrši izvajanje tako, da posvetimo pozornost že takoj spočetka radi usmerjenosti našega iskanja odnosov samo značilnim odnosom. Že od vsega začetka vršimo neko selekcijo med odnosi, tako da jih ostane za končno izbiro samo nekaj res značilnih in pomembnih. Značilni in pomembni odnosi so tisti, ki v nečem ustrezajo osnovni usmerjenosti našega izvajanja. Iskani odnosi pa so tisti, ki popolnoma ustrezajo smeri izvajanja. Ti odnosi so s postavljeno nalogo že docela nakazani.

Nakazanost smotra našega izvajanja je najznačilnejši pojav v vsej dejavnosti izvajanja. Od nje zavisi izbor značilnih odnosov in končno tudi pravilna rešitev naloge.

Čim težja je naloga ali vprašanje, tem bolj sestavljena bo dejavnost pri izvajanju, tem več zmogljivosti bo zahtevala od poedinca in tem bolj bomo lahko v tej dejavnosti ugotovili momente, kakor sta nakazanost in značilnost odnosov. V enostavnih primerih, kjer je rešitev takoj dana in jasna, pa seveda sploh ne bo prišlo do še posebej opazljive dejavnosti izvajanja.

Kakor vidimo, moramo razlikovati pri dejavnosti izvajanja troje najosnovnejših momentov, in sicer: odnosnost, značilnost in nakazanost. V tem smislu se bo tudi funkcija inteligentnosti izražala predvsem le v izvajanju ali odkrivanju odnosov, potem v izvajanju značilnih odnosov in končno v odkrivanju z nalogo nakazanih odnosov, in sicer tako, da bodo vse te tri funkcijske komponente nastopale enotno kot ena sama rezultanta. Funkcija izvajanja pa je le aktualni izraz obče umske sposobnosti, kateri sami se ne moremo približati neposredno niti z najbolj dovršenim inteligentnostnim testom.

Vedno se javlja pri dejavnosti izvajanja večja ali manjša mera te ali one posebne sposobnosti in pa še večja ali manjša mera že pridobljenega izkustva (ali celo umskega znanja).

Kadar nam je rešiti nalogo ali vprašanje, kjer gre za odkrivanje z nalogo nakazanih odnosov, tedaj se izraža tista obča sposobnost, ki jo imenujemo inteligentnost. Inteligenten je



tisti, ki bo ob vprašanju ugotovil za ta problem značilne odnose in ki bo odkril tudi ustrežajoče, namreč že po sami nalogi nakazane (»dane«) odnose. V pravih inteligentnostnih testih morajo biti tudi posamezna vprašanja tako sestavljena, da sledijo omenjenim trem glavnim principom, to je 1.) principu odnosnosti, 2.) značilnosti ali pomembnosti in 3.) nakazanosti ali usmerjenosti.

V dosedanjih izvajanjih smo skušali pokazati, kako gleda testna psihologija na pojav inteligentnosti. Videli smo, da je ugotavljanje višine inteligentnosti praktično omejeno delno zato, ker moremo meriti inteligentnost samo po dejanski storitvi, delno pa zato, ker nimamo absolutnega merila za ocenjevanje teh storitev; s številkami ne moremo adekvatno izraziti kvalitetnih razlik v inteligentnosti. Posebno pa moramo poudariti, da s praktičnimi merjenji ne moremo zajeti vseh tistih subjektivnih momentov, ki odločajo o tem, da se naša umska zmogljivost sploh aktualizira in da se aktualizira zdaj tako, drugič pa drugače. Za poznavanje vse problematike inteligentnostnih testov pa je potrebno, da si podrobneje ogledamo tudi znanstvene osnove same metode testiranja.

## II. DEL.

### **Test kot svojevrstna eksperimentalna metoda**

#### *§ 6. Obseg pojma.*

Beseda test je latinskega izvora (testari — pričati), ki je dobila v angleščini pomen »preizkusa«. Odtod je prišel pojem v sodobno psihologijo, kjer pomeni test preizkus nekih sposobnosti.

V strokovni literaturi beremo najpogosteje to-le definicijo testa: test is a critical examination or trial of person's or things qualities, kar pomeni, da je test kritičen izpit ali preizkus osebnih ali stvarnih lastnosti (kvalitet).

Opozoriti moramo, da je ta definicija najširša, ker obsega i občo ljudsko rabo i svojevrstni strokovni pomen, kar moramo strogo ločiti. Tako je n. pr. test v kemiji lahko tudi reagent ali substanca, s katero ugotavljamo (preizkušamo) sestavne dele neke spojine (n. pr. modro galico in železo).

V naše pedopsihološko področje spadajoče teste moramo imenovati psihološke teste. S tem smo jih ločili od drugih vrst testov (n. pr. od medicinskih, kemičnih itd.) in besedi test smo dali že določen pomen.

#### *§ 7. Razlika od šolskega izpita.*

V naslednjem bomo prikazali razliko med psihološkim testom in šolskim izpitom. Najvidnejša značilnost pismenega ali ustmenega šolskega izpita je, da obsega vprašanja, ki se nanašajo na neko predelano učno stvarino, ki se je obravnavala v določeni učni dobi (šolskem letu, semestru, mesecu, tednu,



dnevu ali uri). Pri testiranju pa preizkusna oseba ne pozna in tudi ne sme poznati določene vsebine testa in razen morebitnega kratkega navodila ni zanjo nobenega pripravljanja in vežbanja.

Značilna razlika med šolskim izpitom in testom je torej v predelani oziroma nepredelani tvarini. To velja seveda s posebno omejitvijo za tisti specialni primer testa, s katerim merimo vprav zmogljivost vežbanja ali učenja, oziroma stopnjo doseženega znanja. Vendar velja tudi tukaj, da ni nobenega neposrednega pripravljanja za testiranje.

Šolski izpit po svojih konkretnih zahtevah ni namenjen kateri koli poljubni skupini, temveč le določeni skupini poedincev, ki se je za ta izpit na poseben način pripravljala. Psihološki test pa lahko predložimo tudi več skupinam poedincev, ki jih nato primerjamo po njih starosti, socialni pripadnosti, spolu, krajevnih razlikah itd.

Glede odgovorov je treba poudariti, da gre pri psihološkem testu za najkrajšo rešitev, ki je sploh mogoča. Reševanje pri pismenih testih se večinoma vrši s podčrtavanjem ali prečrtavanjem že nakazanih možnosti rešitve.

O ocenjevanju in izračunavanju uspeha pa velja, da je šolski izpit ponajveč zelo podvržen osebnemu, subjektivnemu gledanju, medtem ko ta moment pri psihološkem testu skoro ne prihaja do izraza. Rezultat nekega poedinca se pri psihološkem testu primerja s povprečnim rezultatom, ki ga pokaže vsa skupina in šele na tej osnovi določimo poedinčevo mesto v tej skupini. Tako lahko primerjamo višino, oziroma napredek sposobnosti in tudi znanja s povprečnimi storitvami vseh ostalih poedincev.

Vrednost vsakega testnega rezultata bomo torej smatrali za nekaj relativnega. Rezultat, ki ga je dosegel poedinec, dobi svoj določeni pomen šele tedaj, ko označimo stopnjo ali razred v skupini, v katero naj se poedinec uvršča ali s katero naj se primerja. Absolutni rezultat, ki bi ga dosegel neki poedinec, bi bil sam zase brez pomena. Metoda pa, s katero izražamo sorazmerje med rezultatom nekega poedinca in rezultati skupine, je najbolj občutljivo in najvažnejše vprašanje v tehniki testiranja sploh.

### § 8. Razlika med testom in eksperimentom.

Test se razlikuje tudi od psihološkega eksperimenta. Oba se sicer poslužujeta istih poti, vendar pa je njun cilj precej različen. Medtem ko služi preučevalni eksperiment teoretičnemu raziskovanju psihičnih pojavov, ima test kot preizkuševalna metoda predvsem praktični smoter.

Naloga eksperimenta je, ugotavljati, raziskovati (Forschungsexperiment), preizkušati (Prüfungsexperiment) in prikazovati, oziroma razlagati (Demonstrationsexperiment) duševne pojave sploh; pokazati, kako ti delujejo, kako se razvijajo, kaj so njihovi vzroki in učinki.

Čeprav bi test po svojem značaju še najbolj ustrezal preizkuševalnemu eksperimentu (Prüfungsexperiment), se vendar od njega razlikuje še po tem, da skuša meriti predvsem obsežnost, hitrost, točnost in učinkovitost trenutnega delovanja nekega čisto določenega duševnega pojava. Smoter testa ni v določevanju dejavnosti. Ne gre za obče vprašanje, kako se je ta dejavnost razvila, iz česa izvira, kateri so njeni sestavni elementi, temveč gre le za ugotavljanje stopnje individualne sposobnosti ali individualnega duševnega pojava, kakršen je v trenutku preizkušanja.

Vzemimo za primer zaznavanje! Smoter testa ni v tem, da to zaznavanje analizira, temveč da n. pr. določi sposobnost nekega poedinca, kako razlikuje med takimi posameznimi čutnimi dražljaji, ki se med seboj zelo malo razlikujejo. Pri aktu učenja n. pr. pa skušamo s testom dognati hitrost, točnost, obseg in druge podobne značilnosti učenja, ki jih zmore poedinec in nato primerjamo sposobnosti učenja tega poedinca s sposobnostjo drugih.

Iz povedanega razberemo, da testiranje psihološko znanje že predpostavlja. Tu nas tudi ne sme motiti sam zgodovinski razvoj testov, ki še posebej priča, kako je teoretična utemeljitev postopka, kakor tudi znanstvena določitev psiholoških pojmov vedno šepala za prakso.

Umljivo je, da nam prvi testologi niso dali natančnih definicij in analiz tistih sposobnosti, ki so jih preizkušali (na pr. Binet). Sicer pa moramo sploh poudariti, da nam testologija



kot uporabna veda (angewandte Wissenschaft; applied science) ne more odgovoriti na ta vprašanja, ker to ni njena naloga, temveč naloga teoretičnih psiholoških disciplin (deskriptivne psihologije, karakterologije itd.).

Do določitve razlike med eksperimentom in testom prihajamo še po sledeči ugotovitvi. Z eksperimentom ugotavljamo in preizkušamo skupino kot celoto z namenom, da določimo njene skupne znake. S testom pa preizkušamo poedinca ali tudi skupino s posebnim ozirom na poedinca samega. Eksperiment ima kolektivni značaj, test pa individualnega. To dejstvo izvira že iz omenjenega teoretičnega pomena eksperimenta in praktičnega pomena testa.

Na podlagi povedanega lahko trdimo, da je test svojevrstna eksperimentalna psihološka operacija; z njo izzovemo to ali ono duševno aktivnost, katere zunanji izraz potem izmerimo, oziroma klasificiramo. Ta definicija obsega tudi »neme« teste. To so take preizkušnje, kjer preizkusna oseba ne ve, da jo preizkušamo in jo le opazujemo v točno izbranih situacijah. Tudi tu gre namreč za enako vrsto opazovanja. (Popolnejša definicija bi se utegnila nanašati na drugačno izzivanje duševne aktivnosti.)

### § 9. Očrt inteligentnostnega testa vobče.

Med prve teste sploh spadajo umski testi, preizkušnje posameznih ali več umskih funkcij, iz katerih so se razvili inteligentnostni testi, ki so še danes za psihologa najhvaležnejše polje v celotni testni psihologiji.

Izmed vseh vrst testov so dosegli inteligentnostni najvišji razvoj. Spočetka se je test uporabljal kot sredstvo določevanja inteligentnosti posameznika in šele pozneje tudi za skupine. Pri tem so smatrali za inteligentnostne teste tudi take, ki so dejansko preizkušali več kakor to, kar danes imenujemo inteligentnost.

Inteligentnostni test naj bi primerjal poedinčeve storitve iz aktualnega področja hkrati med seboj in s storitvami drugih poedincev. Razen tega pa naj bi primerjal storitve ene starostne dobe s storitvami druge.

Tako bi lahko rekli, da nam podaja Binetov test diagnozo o intelektualnem razvoju posameznika v nekem določenem razvojnem stadiju a o Gesellovem testu bi rekli, da je pokazal diagnozo o občem duševnem nivoju posameznika v določenem razvojnem stadiju.

Teoretično jasnejši pojem inteligentnostnega testa je podal ameriški psiholog Ben D. Wood (Columbia univerza, New York).<sup>7</sup> Po njem merimo z inteligentnostnim testom dejanski samo del celotne umske zmogljivosti. Torej ne merimo neposredno inteligentnosti same, temveč »nekaj drugega«, kar pa v določenih pogojih skoro ustreza inteligentnosti. Wood poudarja obenem, da test sam ne meri prirojene sposobnosti, temveč njen zunanji izraz — storitev.

Kar se tiče razmerja med inteligentnostjo in uspehom v življenju, velja, da inteligentnost sama o uspehu še ne odločuje in da moramo upoštevati še druge činitelje. Tako n. pr. ne gre postavljati prognoze o uspešnem ali neuspešnem učnem, oziroma šolskem napredovanju na osnovi rezultatov, ki jih je dalo testiranje z inteligentnostnimi testi.

Enako velja Woodova ugotovitev tudi za vse ostale teste. S pomočjo testa moremo namreč meriti vedno samo »en« činitelj ali le del vseh psiholoških činiteljev. Test preizkuša le poseben primer storitve. Tudi ta trditev soglaša z individualno noto testa. V psihologiji gre n. pr. za jezo katerega koli poedinca, v testologiji pa gre za določeno doživljanje, to je za jezo tega in tega poedinca.

Vedeti moramo, da meri test neko sposobnost preko konkretne merljive storitve v neki konkretni situaciji in da pri tem nikoli ne prihaja neposredno do izraza sposobnost v celoti, niti ni prav ta sposobnost edini činitelj, ki odloča o višini storitve v testu.

Praktično moramo seveda računati s tem, da se nam sposobnost javlja pri testiranju v svoji maksimalni zmogljivosti. Kadar s testi preizkušamo neko zmogljivost, tedaj računamo,

<sup>7</sup> Ben D. Wood: *Measurement in Higher Education*, New York, 1923.

J. K. Kandel: *Examinations and their Substitutes in the United States*, New York 1936.



da se ta javlja v testu v svoji maksimalni stopnji. To maksimalno stopnjo imenujemo *optimalno stopnjo* (optimum) ter tvori ta osnova nujno izhodišče vsakemu določevanju zanesljivosti preizkušanja.

Poleg inteligentnostnih testov poznamo tudi še teste posebnih sposobnosti. S temi preizkušamo n. pr. kombinacijo, razne spretnosti itd. in uporabljamo jih posebno v vzgojni in poklicni svetovalnici. Razen testov, s katerimi lahko preizkušamo temperament, značaj in njegove kvalitete, moramo zlasti omeniti še učne teste (educational, achievement test). Z njimi skušamo meriti rezultate nekega vežbanja, oziroma učenja in s tem zopet določujemo samo zmogljivost, da človek doseže določeno višino znanja.

### III. DEL.

## **Možnosti in meje testiranja**

### *§ 10. Sposobnost in izvežbanost.*

Zunanji izraz, ki ga s testom merimo, je sicer zavisen od sposobnosti, ima pa še nekatere posebne pogoje. Pri tem mislimo na zavisnost od prodhodnega znanja in izkustva v najširšem pomenu besede.

Nastane vprašanje, v kakšnem razmerju sta zmogljivost in izvežbanost. Rekli bomo, da zna ta ali oni igrati na klavir ali pisati na stroj, ne samo zato, ker ima sposobnosti za učenje in naučenje tega instrumenta, temveč tudi zato, ker se je toliko in toliko časa vežbal. Razumljivo je, da je njegovo obvladanje instrumenta zavisno tudi od samega vežbanja. Vemo nadalje, da je ob manjši sposobnosti potrebno več vaje, da bi pa brez najmanjše mere sposobnosti tudi vaja nič ne hasnila.

Razumljivo je tudi, da obvlada nekdo tuj jezik zato, ker je prišel v stvarni odnos z njim in se je naučil izgovarjati tuje besede in spoznavati njih odnos in pomen, ne pa samo zato, ker ima sposobnost za naučenje tujih jezikov. Tako pravimo tudi, da poznamo in razlikujemo barve, ker smo se temu priučili ob primerjanju različnih barv.

V vseh teh primerih je treba razlikovati sposobnost, ki jo ima poedinec na prvi stopnji svojega razvoja, od rezultatov prizadevanja in vežbanja, ko je tako dosegel še neko posebno stopnjo usposobljenosti.

Izkustvo je pokazalo, da se razlikujejo poedinci bolj v svojih sposobnostih, kot pa v svoji pridobljeni izvežbanosti, vendar skušamo pri testiranju izločiti tudi razlike v izvežba-



nosti. Posebne dejavnosti, ki jih zahtevamo od poedinca pri testiranju, izberemo iz mnogih drugih, ki so skupne izkustvu vseh poedincev, s katerimi ga primerjamo. Zato smatramo, da je v dejavnostih, ki naj jih testiramo, stopnja izvežbanosti ali izkustva enaka ali vsaj praktično približno enaka pri vseh poedincih. Jasno pa je, da n. pr. izvežbano pisanje na stroj ali igranje na klavir ne more biti merodajno za merjenje osnovne motorne ali manuelne spretnosti. Če bo treba testirati take sposobnosti, bomo izbrali dejavnosti, kjer izvežbanost ne prihaja do izraza.

Merjenje sposobnosti za razlikovanje intervalov ali intenzitete dveh barv je — v nasprotju s preizkušnjo v pisanju na stroj ali igranju na klavir — primer testa za ugotovitev prirojene sposobnosti, kajti poedinci lahko kljub majhni izvežbanosti pokažejo svoje sposobnosti in izmerjena sposobnost ustreza vedno njihovi maksimalni zmogljivosti.

Ob teh ugotovitvah nastaja vprašanje, kako se pri testiranju izraža različna stopnja izvežbanosti posameznih sposobnosti.

Če hočemo meriti sposobnosti, pri katerih ima vežbanje zelo majhen učinek, bomo videli, da se bo izvežbanost izkazala za naša merjenja kot zelo malo pomembna. Ob preizkušnji dejavnosti, ki izvirajo iz kompleksa raznih sposobnosti, pa izvežbanost poedinih sposobnosti iz tega kompleksa v celotni storitvi ne prihaja v znatni meri do izraza.

Pri testiranju prirojenih sposobnosti pa moramo vendar upoštevati pridobljeno izvežbanost in znanje. Popolnoma bo ustrezal test samo tistim skupinam preizkušencev, ki so imele približno enako priliko za vežbanje in pa za testiranje umskih funkcij, kjer so razlike v izvežbanosti zelo neznatne. Vendar pa je težko določiti ostro mejo, do kod segajo testi, ki merijo izvežbanost, in do kod segajo testi, s katerimi preizkušamo prirojeno sposobnost!

Za pravo razumevanje testa je važno, da vemo, ali smo s testom dognali samo prirojeno sposobnost poedinca ali pa tudi njegovo izvežbanost. To razliko nam najlepše osvetli primer iz glasbe, ko naj nekoga nazivamo »mojstra«, drugega pa

»virtuoza«. Prvemu pripisujemo večje sposobnosti, globino in občutenost, drugemu pa bolj tehnično dovršenost. Pri elementarnih primerih nam sposobnosti ni tako težko spoznati, kakor pri višjih ali kompleksnih storitvah.

### § 11. Osnovni vidiki za sestavo testa.

Gradivo in metode preizkušanja niso zgolj teoretična konstrukcija psihologa, ki se s takim študijem ukvarja, temveč se ugotovijo in določijo šele po daljšem preučevanju in preizkušanju. Test velja šele tedaj, ko je že pokazal praktični uspeh pri preizkušanju in zadostil vsem zakonom statistike. Vemo tudi, da je merjenje relativno, kajti rezultat, ki ga pokaže poedinec, je treba razlagati v primerjavi z rezultati, ki so jih pokazali drugi poedinci.

Pri vsebini testa se moramo najprej vprašati, kako bomo ločili teste posebnega umskega procesa ali posebne sposobnosti na eni strani in teste obče sposobnosti ali inteligentnosti na drugi. Testi vobče sposobnosti so vedno skupina testov nekih določenih posebnih sposobnosti. Treba je le ugotoviti, katera skupina posameznih sposobnosti zadostuje za merilo obče inteligentnosti ali obče sposobnosti.

Prav tako je treba ugotoviti, v kakšni meri prihaja obča umska zmogljivost do izraza v določeni skupini umskih testov. Zato moramo res točno razlikovati občo sposobnost od posebnih sposobnosti in ne smemo nekaterih posebnih sposobnosti identificirati z občo. V tej zvezi pa nastaja tudi vprašanje, katere posebne sposobnosti imajo za nas vendar še »obči« značaj.

Predno moremo določiti vsebino posebnega ali specialnega testa, moramo preučiti sposobnost, ki jo hočemo preizkusiti, akoravno to ni ne prva, ne zadnja naloga testologa; to je bolj naloga teoretske psihologije.

Večinoma to ni preprosta zadeva. Tako je n. pr. preučeval Amerikanec *Seashore* sposobnost za glasbo in ugotovil na tej sposobnosti 30 komponent, kot n. pr. sposobnost za razlikovanje intervalov, razlikovanje intenzitete zvoka (dinamika), smisel za ritem, sposobnost spomina za melodije, motorno spretnost v izvajanju, glasbeno ocenjevanje, glasbeni smisel



itd. S tem je dan jasen primer, da se naše sposobnosti ne dajo tako enostavno analizirati, oziroma da ne moremo smatrati teh ali onih sposobnosti, ki jih merimo, za elementarne enote.

Poglejmo, kako je s sposobnostjo razlikovanja med različnimi stopnjami intenzitete tonov! Če delamo take poskuse s toni različnih glasbenih instrumentov, bomo dobili različne rezultate. To pomeni, da je »jakostno« glasovno razlikovanje v zvezi tudi z »barvenim«. Tudi testi, ki merijo take sposobnosti, ne merijo enostavne in enotne sposobnosti.

Upravičeno lahko trdimo, da je vsak specialni ali posebni test zopet bolj ali manj tudi obči ali generalni test. To pravilo velja tudi obratno in zato lahko merimo do neke mere s testom, ki preizkuša kompleksne storitve, tudi specialno stran teh storitev. Ako n. pr. ne moremo jasno označiti in definirati spomina, s tem še ni rečeno, da ne moremo meriti lahkote in hitrosti, s katero se nekdo nauči sestavka, imen, števil in podobno.

Iz vsega povedanega sledi, da ne moremo nikoli sestaviti testov, ki bi merili samo občo sposobnost, ne da bi pri tem vsaj v manjši meri ne merili tudi posebne sposobnosti, kar velja tudi obratno. S same načelne, teoretične strani je ta problem precej težaven in naši pozitivni izsledki so v marsikaterem oziru še nepopolni. Praktično pa je to vprašanje nekoliko lažje rešiti.

Ko smo napravili prvi korak in označili sposobnost, ki jo nameravamo meriti, je še potrebno, da dobimo tudi merske pripomočke. V ta namen se je nekaj časa uporabljala metoda deskriptivne analize. Psiholog je najprej sam zase definiral vso prirodo določene sposobnosti s psihološkimi izrazi in je potem skušal zbrati in sestaviti več testov, o katerih je mislil, da ustrezajo merjenju te sposobnosti. Nato je prišlo na vrsto praktično preizkušanje testov. Pri tem je uporabil vse teste, ki jih je sestavil, ter obdržal le tiste, ki so najbolj ustrezali zakonom statistike. Šele ta preizkušnja ali utrjevanje (Eichung; étalonner) določa uporabnost testa. Ako ugotovimo, da je bilo pri prvem preizkušanju uporabnih več testov, jih združimo v zaključeno obliko ali tudi lestvico (scale; l'échelle). Vidimo torej, da je šel postopek čisto po empirični poti.

Oglejmo si sedaj načela, ki odločajo pri izbiranju vsebine za teste obče intelektualne sposobnosti! Preden preidemo k razpravljanju o vrstah testov, katerih naloga je merjenje obče sposobnosti, je treba odgovoriti na vprašanje, ali sploh obstoja in koliko obstoja obča sposobnost. Eksistenca obče intelektualne sposobnosti je šele v zadnjem času tudi v ameriški psihologiji splošno priznana.

Kratke definicije, kakor tudi preprosti splošni opisi pojma inteligentnosti, so bili vedno merodajni pri sestavljanju, oziroma pri razvoju inteligentnostnih testov. Seveda marsikatera definicije niso bile znanstveno fundirane.

Hipoteze o eksistenci neke obče sposobnosti vodijo do zaključka, da morajo vsa prava merjenja inteligentnosti pokazati čim večjo medsebojno skladnost. Seveda so ta merjenja vedno le merjenja posameznih umskih storitev, kakor to drugače tudi biti ne more. Od sestave testa je pa zavisno, v koliki meri smo izmerili ob posamezni umski storitvi občo umsko sposobnost. Občo umsko sposobnost lahko najboljše merimo na storitvah, ki so izraz kompleksnih, ne enostavnih sposobnosti. Zato ne moremo reči, da bi bili najboljši inteligentnostni testi obenem tudi merila specifičnih in določenih umskih funkcij.

Storitev, ki so značilen izraz inteligentnosti, ne moremo smatrati obenem za specifične storitve spomina, zaznavanja, sklepanja itd. Vsekakor pa lahko sestavimo tudi teste take vrste.

Oglejmo si na kratko teste, s katerimi merimo različne umske procese! Psihologija upravičuje domneve, da zahtevajo ti testi obenem tudi pojav obče zmogljivosti, ki v visoki meri nastopa v vsaki kompleksni intelektualni dejavnosti. To domnevo potrjuje dejstvo o medsebojni skladnosti rezultatov inteligentnostnih testov, ki jo lahko matematično izrazimo s korelacijskim koeficijentom.<sup>8</sup> Iz tega sledi, da so le tisti testi najboljše merilo inteligentnosti, ki pokažejo najvišjo stopnjo te medsebojne skladnosti (t. j. interkorelacije).

<sup>8</sup> To je številčni indeks, ki nam kaže stopnjo odnosa med dvema vrstama rezultatov,



Ta ugotovitev ima praktičen pomen za sestavo testa. Praksa je na testnih rezultatih pokazala, da najbolj ustrezajo tisti inteligentnostni testi, ki so pokazali visoko interkorelacijo in pa še skladnost s kriterijem, to je s »sodbo« o višini inteligentnosti poedinca. Skladnost rezultatov nekega testa s kriterijem izražamo prav tako s številčnim indeksom po raznih statističnih metodah. Taka dejstva ponovno potrjujejo hipotezo o obči umski sposobnosti. Principu o nadvladanju neke obče zmogljivosti bomo torej ustregli, če sestavimo test, ki bo imel najvišjo interkorelacijo in najvišjo korelacijo s kriterijem.

### § 12. *Gradivo testa.*

Izkustvo je pokazalo, da skupina testov daje točnejše rezultate kakor kateri koli posamezni test. Dejstvo, da noben posamezni test ni dovršena mera in da je skupina testov boljša, moramo razumeti v tem smislu, da vsebuje vsak test še nekaj postranskih činiteljev, ki jih merimo tako, kakor poskušamo meriti tudi centralni činitelj sam. Ali drugače rečeno: vsak test meri sposobnosti, oziroma inteligentnost, z druge strani.

Inteligentnost se vedno javlja ob nekem gradivu in pri obvladovanju tega gradiva prihaja v poštev tudi naše izkustvo. Efekt mišljenja v določenem primeru ni zavisen samo od sposobnosti mišljenja v abstraktnem smislu, temveč še posebej od našega izkustva. Saj poznamo obče znani latinski izrek: *Nihil est in intellectu, quod non fuerit in sensu*. Misel se da izraziti z besedami, lahko pa se izraža tudi v obliki matematičnih simbolov ali mehaničnih opravil. Storitve sama, ki se tako različno prikazuje, pa je lahko povsem ista. Hitrost in lahkota, s katero poteka pri poedincu določen niz mišljenja, je zavisna torej deloma tudi od prirode gradiva in poedinčevega prilagajanja temu gradivu, ne pa samo od inteligentnosti.

Če upoštevamo razliko med gradivom in obliko mišljenja, tedaj smo upravičeni uporabljati najraznovrstnejše teste. Ta raznovrstnost ni toliko v tem, da se merijo z njimi različni umski procesi, marveč v tem, da vsak izmed njih meri umsko dejavnost vedno takó, kakor se javlja v posameznih konkretnih primerih mišljenja. Različnost teh konkretnih storitev je

zavisna delno od gradiva, delno od oblike ustrezajočih operacij, v vsakem primeru pa tudi od dane mere obče umske sposobnosti.

Če gremo še korak dalje, lahko postavimo vprašanje o veljavnosti naših inteligentnostnih testov. Ali bodo imeli testi isto veljavnost, če jih bomo uporabljali ob različnih prilikah, n. pr. v šoli ali izven šole? Ali lahko razlikujemo neko šolsko inteligentnost in neko »življenjsko« inteligentnost?

Brez dvoma moremo odgovoriti na to vprašanje, da »šolska« inteligentnost ni identična z občo usposobljenostjo, ki jo zahteva vsakdanje življenje. Ako trdimo, da ti dve usposobljenosti nista identični, ne pomeni, da ni nobenega medsebojnega razmerja. Vendar za sedaj lahko rečemo samo, da je skladnost rezultatov inteligentnostnih testov s šolskim uspehom tesnejša kakor skladnost z uspehom v praktičnem življenju. Iz tega dejstva vidimo, da omogoča testiranje z dosedanjimi inteligentnostnimi testi zanesljivejše prognoze za šolski uspeh kakor za uspeh v praktičnem življenju.

Kjub vsem razlagam imajo zato današnji inteligentnostni testi svoje omejitve, na kar pa mnogi ne polagajo dovolj važnosti. Če pokažejo n. pr. rezultati vojaških rekrutov po Stanford-Binetovem testu<sup>9</sup> isto vrednost kakor rezultati, doseženi po tem testu pri šolskih otrocih, tedaj je jasno, da tema enakima vrednostima ne smemo pripisovati enakega pomena.

Iz vsega povedanega lahko zaključimo sledeče: možnosti inteligentnostnih testov niso omejene samo po prirojeni intelektualni sposobnosti, temveč deloma tudi po pridobljenem izkustvu (znanju). Zgradba teh testov mora radi tega stremeti po čim večji izločitvi g r a d i v a , v katerem prihaja do izraza že pridobljeno izkustvo, kar je v današnjih inteligentnostnih testih že v precejšnji meri doseženo. Storitve, ki jih zahtevamo v inteligentnostnih testih, niso odvisne samo od inteligentnosti, temveč tudi od določene mere priučenega znanja. To posebno dejstvo je zelo podrobno obdelal Owen Bishop v svojem delu »What is measured by Intelligence Tests« (Journal of Educational Research, vol. 9., str. 29, 1924).

<sup>9</sup> L. M. Terman: The Measurement of Intelligence, New York, 1916.



### § 13. Razvrstitev gradiva.

Razlaga gornjih dejstev ima svoj praktični pomen za merjenje obče inteligentnosti, in sicer zato, ker nam da zanesljivost, da moremo meriti inteligentnost do neke mere (stopnje) dokaj natančno s pomočjo testov. Nikdar seveda ne moremo uporabljati v inteligentnostnem testu samo občega ali abstraktnega gradiva. Naše gradivo je vedno konkretno in zahteva čisto določene storitve. Radi konkretnosti gradiva, oziroma storitev, ki jih v testu zahtevamo, prihaja do neke mere vedno do izraza na eni strani izkustvo, na drugi pa določena mera posebne sposobnosti. Inteligentnostni testi zato ne izgubijo svojega pomena. S tem so samo nakazane meje in možnosti preizkušanja, zlasti pa vidimo, kakšno važnost ima razlaga rezultatov.

Če bi hoteli izraziti te omejitve z enim stavkom, bi rekli, da rezultate inteligentnostnega testa obravnavamo vedno kot neke približne, ne pa perfektne ugotovitve o meri inteligentnosti posameznika ali skupine. Pravo vrednost namreč dobi rezultat šele ob temeljiti znanstveni psihološki razlagi. Zato je razumljivo, kdo bo kos taki nalogi, in na osnovi tega dejstva bomo razumeli marsikatero neuspehe pri testiranju. Kolikor pa gre za praktično razlaganje testnih rezultatov, pa je končno potrebno tudi široko znanje o zakonih statistike.

Razmotrivanja o vsebini in obliki testa nam dajejo napotke, po kakšnih principih moramo pri sestavljanju testa izbirati posamezne naloge, vprašanja itd. Često se uporablja testno gradivo, ki ga je že kdo drugi nabral ali nakazal. V nekaterih primerih se gradivo modificira, v drugih pa je treba določiti novo gradivo ob primerjavi s starim. Končno lahko izberemo gradivo tudi popolnoma samostojno. Psiholog, ki je sestavil teste, je moral v začetku izbrati veliko število nalog in vprašanj in zanima nas, kateri principi so bili vodilni pri izbiranju in urejevanju tega gradiva.

Osnovno vprašanje je, kako urediti naloge, da si bodo postopno sledile od lažje do težje, in kako jih povezati med seboj ter vnesti v test. Vrstni red posameznih nalog v testu mora biti v nekem razmerju do zaključene, celotne oblike

testa. Test mora biti za tiste, katerim je namenjen, dovolj težak, vendar ne pretežak. Če je namenjen mlajšim otrokom, morajo biti skoki med lažjimi in težjimi nalogami drugačni, kakor pa tedaj, če je test namenjen odraslim.

Skoki med lažjo in težjo nalogo ne smejo biti niti premajhni, niti preveliki. Ne smejo biti tako majhni, da bi lahko smatrali vse naloge za enako težke, pa tudi ne tako veliki, da bi težje naloge že presegale zmogljivost testirancev.

V tem je načelo občutljivosti, kateremu je zadoščeno šele tedaj, če moremo s testom uspešno razlikovati tako podpovprečne in nadpovprečne poedince. Lažje in težje naloge so urejene po določeni postopici, ki mora biti taka, da je poedini test uporaben samo za določeno starostno stopnjo in za določene namene. Vprašanje izbiranja vsebine testov zahteva, da upoštevamo razmerje v težavnosti posameznih nalog ali delov testa.

Test lahko sestavljamo ne samo z vidika postopno naraščajočih, t. j. od lažjih do težjih nalog, temveč tudi po principu enako težkih vprašanj in nalog. Pri vsem je najvažnejše, da test ustreza glavnemu pogoju, to je: test mora biti uporabljiv za razlikovanje stopenj sposobnosti poedincev ali skupin.

#### § 14. Hitrostni in sposobnostni moment v testu.

Če je test sestavljen po omenjenih principih, tedaj lahko posameznike na osnovi doseženih rezultatov pravilno razvrstimo po stopnjah njih zmogljivosti. Prihajamo do vprašanja, kako je z razvrščanjem posameznikov v primeru, če so vprašanja v testu vsa enako težka. Ako pri takih testih reši kdo prvo nalogo, bo lahko rešil vse, ali vsaj večino od njih. Če pa so naloge tako težke, da jih nekateri od preizkušancev sploh ne morejo rešiti, potem teh poedincev ne bo mogoče razvrstiti med seboj.

V primeru, da je test lahek za vsakogar, določimo razvrstitev v glavnem po hitrosti, s katero posameznik reši test. Hitrost reševanja lahkih nalog pa je le delno merilo za individualno sposobnost pri reševanju težjih nalog. Zato bo test v tem primeru predvsem merilo hitrosti. Ako pa so naloge urejene



postopno od lažje do težje, tedaj jih brezdvomno moremo uporabljati bolj za merilo sposobnosti nego za merilo hitrosti.

Ker nam razlika v hitrosti reševanja še ne da prave slike o razlikah v sposobnosti, bomo pri sestavljanju sposobnostnih testov pazili na to, da nam sploh nikoli ne bo treba razlikovati poedincev po hitrosti reševanja. V test bomo postavili toliko nalog, da v odmerjenem času sploh nihče ne bo mogel rešiti vseh.

Vprašanje je, ali sploh moremo sestaviti docela hitrostni test ali celo docela sposobnostni test. Razlike v hitrosti si razlagamo deloma z razlikami v sposobnosti, na drugi strani pa je možno, da preizkušamo hkrati i hitrost i sposobnost, in sicer tedaj, če ima preizkušanec odmerjeno pravo mero časa.

Zato je pri testih, ki so sestavljeni v postopici od lažje do težje naloge, običajno odmerjen tudi čas za reševanje, in to bolj glede na sposobnostni moment nego hitrosten.

Najbolj poznan in mogoče najbolj dovršen poskus take vrste so izvedli ob priliki sestavljanja ameriških vojaških testov (Army Scale Alpha; podrobno glej »Army Report, Psychological Examining in the United States Army«, ki ga je izdal R. M. Yerkes; Washington, National Academy of Sciences 1921). Preučevali so rezultate, ki so jih dosegli, ko so podaljšali odmerjeni čas za reševanje testa.

Učinek, ki so ga dosegli s podvojenim časom, so preučevali na dva načina: z določevanjem korelacije med rezultati enojnega in podvojenega časa, in pa z ugotavljanjem odstotka tistih posameznikov različnih starosti, ki so v podvojenem času dosegli boljši rezultat. Rezultat tega preučevanja je bil, da je korelacija med testi z enojnim in testi s podvojenim časom razmeroma visoka pri hitrostnem testu. Nasprotno pa je bilo tako razmerje majhno pri sposobnostnem testu.

To se pravi: pri hitrostnem testu bomo v podvojenem času dosegli dvakrat boljše rezultate; pri pravem sposobnostnem testu pa bo razlika med rezultati, ki smo jih dobili pri enojnem in pri podvojenem času, znatno manjša. Lahko torej rečemo, da je vsak test, ki je pokazal v tem smislu visoko korelacijo, *nujno tudi hitrostni test*. Korelacija, ki jo je v tem smislu pokazala vprav Army lestvica med enojnim in podvojenim

časom, je znašala 0,965, kar nam takoj jasno pove, da je Army Alpha test skoro samo hitrostni test.

Razmerje med hitrostjo in sposobnostjo pa osvetljujejo še naslednja dejstva. Če ima poedinec, ki hitro rešuje test, tudi visoko stopnjo sposobnosti, tedaj bo pri reševanju testa (bodisi pri sposobnostnem ali hitrostnem testu) pokazal tudi visoko korelacijo med rezultati v enojnem in podvojenem času. To pa radi tega, ker je tu hitrostni test obenem sposobnostni test.

Če se pokaže med našimi rezultati hitrostnega in sposobnostnega testa nizka korelacija, tedaj lahko pričakujemo, da bo pri sposobnostnem testu pokazal podvojeni čas rezultate, ki so v zelo nizki korelaciji z rezultati, doseženimi v enojnem času. Pri hitrostnem testu pa bo pokazal podvojeni čas rezultate, ki so v zelo visoki korelaciji z rezultati, doseženimi v enojnem času. Sploh lahko trdimo, da je zelo visoka korelacija med rezultati, doseženimi v enojnem in podvojenem času zavisna predvsem od momenta hitrosti, mnogo manj pa od same sposobnosti.

Profesor pedagogike na univerzi v Chicagu, Frank Freeman, je skušal dokazati to domnevo z enostavnim eksperimentom. Sestavil je izrazito hitrostni test in izrazito sposobnostni test, ki sta bila precej enake vsebine ter sta imela enako število vprašanj in nalog. V prvem testu so bile naloge približno enako težke, v drugem pa so bila vprašanja razporejena po stopnjah od lažjega do težjega. Oba je preizkusil na visokošolcih in dobil je zelo značilne rezultate. Ugotovil je korelacijo med rezultati enojnega in podvojenega časa pri obeh testih in tudi korelacije med rezultati obeh testov samih, ki je znašala:

<sup>10</sup> Korelacija med rezultati v enojnem in podvojenem času pri hitrostnem testu . . . . .	0.87 ± 0.03
Korelacija med rezultati v enojnem in podvojenem času pri hitrostnem testu . . . . .	0.87 ± 0.04
Korelacija med hitrostnim in sposobnostnim testom . . . . .	0.63 ± 0.06

<sup>10</sup> Frank Freeman: Mental Tests, Chicago, 1926.



Manjša korelacija med hitrostnim in sposobnostnim testom kaže, da obstaja razlika med hitrostjo in sposobnostjo. Nižja korelacija med enojnim in podvojenim časom v sposobnostnem testu dokazuje, da so testi le tedaj hitrostni, če pokažejo v primerjavi s sposobnostnim testom zelo visoko korelacijo med enojnim in podvojenim časom. V tem imamo torej metodo, s katero ugotavljamo, da je vsak tak test hitrostni test. (Sem spada tudi Army Alpha Test, česar se avtorji niso zavedali.) Ti podatki obenem tudi kažejo, da je namen sposobnostnega testa merjenje sposobnosti in da je nujno treba odmeriti več časa za njegovo rešitev, ker bodo pri podaljšanem času razlike v sposobnostih posameznikov še jasnejše. Vedeti pa moramo, da odmera časa ne more biti poljubna, kajti omejitev nam narekujejo gradivo testa in namen preizkušanja.

Vendar je kvalifikacija hitrosti reševanja zelo važen moment, vsaj v nekaterih posebnih primerih.

Hitrost je zelo važen faktor n. pr. pri strojepisju ali steno grafiji. V teh primerih pa moramo obravnavati hitrost, kolikor lahko kontrastira s točnostjo. Točnosti tu ne moremo identificirati s sposobnostjo. Ko smo obravnavali hitrost, je bila mišljena prava efektivna hitrost ali pa kombinacija hitrosti in točnosti. To moramo dobro razlikovati, ker ima n. pr. za znanstveno izsledovanje ali tudi za umetniško ustvarjanje sposobnost in točnost relativno mnogo važnejšo vlogo kot hitrost.

Na podlagi tega moremo zaključiti, da ni mogoče meriti vseh intelektualnih storitev z isto dimenzijo. Nekateri zahtevajo, da jih merimo predvsem po hitrostni dimenziji, druge po kvalitetni dimenziji ali dimenziji sposobnosti, večina od njih pa zahteva, da jih merimo s kombinacijo obeh. Značaj dosedanjih testov kaže, da je bolj poudarjena njih hitrostna plat. Zdi se, da bi mogli najprimerneje dognati individualne sposobnosti s testom, ki bi vseboval naloge, nanašajoče se na hitrost izvajanja in posebej še na sposobnosti. Zgradba hitrostnega testa je drugačna od zgradbe sposobnostnega. Pri prvem gre za razporeditev nalog v vrste enako težkih vprašanj, pri drugem gre za razporeditev nalog v take vrste, kjer se naloge vrstijo od lažje do težje.

### § 15. Postopica nalog in zanesljivost testa.

Če so testna vprašanja urejena po stopnjah od lažjih do težjih, tedaj običajno odloča o vrstnem redu odstotek poedincev, ki vprašanj niso rešili. To se pravi, da je težje tisto vprašanje, ob katerem pokaže več reševalcev neuspeh.

Vprašanja izberemo tako, da moremo iz njih sestaviti enakomerno postopico, v kateri se vrstijo vprašanja od lažjih do težjih.

O določevanju teh stopenj prevladuje naziranje, da naj bodo v skladu z zakoni znane verjetnostne (Gaussove) krivulje. Vprašanja, ki so na sredini postopice, mora rešiti večina poedincev (v idealnem primeru 68 odstotkov). Najtežja vprašanja reši samo malo število in prav tako jih je samo nekaj, ki komaj in komaj rešijo le najlažja vprašanja. Vsaka stopnja v postopici vprašanj predstavlja obenem določen odstotek poedincev, ki dosežejo pri reševanju testa ustrežajočo višino. Ta odstotek bo najvišji na srednjih stopnjah postopice, najnižji pa na obeh koncih.

Ker so razlike pri umski rasti v mlajših letih (n. pr. med 5. in 6. letom) večje kot razlike v starejših letih (n. pr. med 15. in 16. letom), zato morajo biti testi za prvo od omenjenih starostnih stopenj bolj simptomatični (občutljivi) nego za drugo. Terman in C. Burt sta n. pr. dognala, da mora biti odstotek poedincev, ki naj rešijo srednje težke naloge v testu za 5. leto, znatno višji (okoli 81 %) kot odstotek poedincev, ki rešijo srednje težka vprašanja v testu za 16. leto (okoli 62 %), če naj bo test za dotično starostno dobo res simptomatičen.

Naslednja metoda urejevanja lestvice postopno lažjih in težjih vprašanj je metoda umske starosti. Sloni na principu, da je test mnogo težji, če ga reši določen odstotek starejših učencev, nego test, ki ga reši enak odstotek mlajših učencev. Teste po starostnem principu torej lahko uredimo po vrstnem redu težavnosti, ker računamo, da so najtežji oni testi, ki jih je rešil le manjši odstotek starejših otrok (mlajši pa sploh ne).

Pri sestavljanju testov po metodi umske starosti pa moramo upoštevati, da rast obče umske sposobnosti ni enako-



merna v vseh starostnih dobah. Rast obče umske sposobnosti izražamo grafično s progresivno starostno krivuljo (mental growth curve). Najnovejše ugotovitve, do katerih sta prišla Angleža Richardson in Stokes, nam povedo, da je ta rast najhitrejša približno od 5. do 15. leta ter se potem do 24. leta polagoma ustavlja.

Končno zadenemo še na vprašanje, koliko nalog naj bo v testu, oziroma kako dolg naj bo test. Povedali smo že, da se kaže daljši test kot mnogo zanesljivejše sredstvo od krajšega, ker se pri njem jasneje izražajo individualne razlike med posamezniki. Tega principa so se posluževali mnogi sestavljalci v skladu z odmerjenim časom, ki je določen za reševanje testa (tako n. pr. zahteva Thorndikeov test 3 ure časa). Nastane pa vprašanje, ali sploh moremo določiti razmerje med dolžino in zanesljivostjo testa in tako njegovo dolžino.

Določitev odnosa med dolžino in zanesljivostjo testa nam je podal Spearman s svojo formulo, katere se je praktično poslužil Holzinger (Karl J. Holzinger and Blythe Cleyton »Further Eksperiments in the Psychology«, vol. 16, str. 289, 1925). Če imamo vrsto podobnih testov enake dolžine in zanesljivosti in če poznamo računsko dobljeni matematični izraz te zanesljivosti, moremo na osnovi Spearmanove formule napovedati, kakšna bo zanesljivost posameznih komponent nekega testa. Razumljivo je, da se posamezni deli testa med seboj razlikujejo, da niso enaki niti po dolžini, niti po zanesljivosti in zato s formulo ne moremo dobiti absolutnih vrednosti. Holzinger pa je dognal, da so dobljeni rezultati o zanesljivosti testa precej točni.

Iz obravnavanih vprašanj razvidimo, da je sestavljanje uporabljivega inteligentnostnega testa težavna zadeva. Treba je upoštevati mnogo momentov, če naj test ustreza zahtevam praktične uporabljivosti in zanesljivosti. Vendar pa je treba tudi najskrbneje sestavljen test še praktično preizkusiti in prilagoditi danim razmeram, predno ga lahko uporabljamo.

#### IV. DEL.

### **Osnove standardizacije**

#### *§ 16. Nastanek in razvoj norme za testiranje.*

Rezultat inteligentnostnega testa je sam zase brez pomena in ne moremo se še izraziti o učenčevi bistrosti, dokler določenega rezultata ne primerjamo z drugimi rezultati, posebno pa s tistim rezultatom, ki velja za normalnega. Prvi je načel problem normalnih rezultatov A. Binet.

S svojimi testi, ki jih je uredil v lestvico, je skušal predvsem odkriti zaostale učence, da bi jih premestil v pomožne razrede. V ta namen je predložil svoje teste večjemu številu učencev različne starosti. Tako je lahko ugotovil za vsak test posebej, za katero starost je prikladen, upoštevajoč pri tem seveda uspeh povprečnih učencev.

Na tej osnovi je zgradil Binet svojo postopno lestvico od lažjega do težjega vprašanja, ki mu je šele omogočala uspešno preizkušati i slabše i boljše poedince neke starosti. Če je tako n. pr. 4 letni otrok rešil test, ki ga praviloma reši šele povprečen 5 letni otrok, pomeni to, da je za eno leto nadpovprečen ali za eno leto nad normalo.

V tej zvezi moramo opozoriti, da je bila sicer ta Binetova zamisel velikega pomena za nadaljnji razvoj testov, vendar pa netočna v svoji metodični (tehnični) izpeljavi.

Anglež C. Burt je namreč pokazal, da je treba ugotoviti, v kateri starosti reši test vsaj polovica poedincev. To se pravi, da gre za ugotovitev tiste starosti, katere »srednji« poedinec bo test uspešno rešil. Srednjega poedinca predstavlja v statistiki medijan (Mittelwert), kar pa ni isto kot povprečje,



s katerim je računal Binet. V vrsti več rezultatov je medijan tisti rezultat, nad katerim in pod katerim je še polovica rezultatov. Zato imenujemo medijan tudi srednji rezultat.

Na osnovi povprečnega rezultata je Binet vpeljal »umsko starost« (âge mental). Ta starost bo pri povprečnem, oziroma srednjem poedincu popolnoma enaka z njegovo »kronološko starostjo« (âge réel; tudi koledarska starost). Povprečni poedinec pa bo test svoje starosti vedno tudi uspešno rešil. Zamisel umske starosti nam šele omogoča, da obravnavamo rezultate inteligentnostnih testov z vida norme.

Uspeh, ki so ga pokazali povprečni ali srednji poedinci, nam daje neko odnosno točko, ki tvori izhodišče za določevanje uspeha vseh drugih posameznikov te starosti. Tako lahko določimo, kdo je za to ali ono starostno dobo povprečen, pod- ali nadpovprečen.

V začetkih razvoja so izražali rezultate z diferenco med umsko in kronološko starostjo. Kmalu pa so ugotovili, da je zaostalost v mlajših letih sorazmerno večja nego v višjih letih. Tega z diferenco seveda ni mogoče izraziti in tako so uvedli pojem »umskega razmerja« (mental ratio). Razlaga rezultatov je postala s tem lažja in točnejša. Če je n. pr. imel 10 letni otrok umsko starost 12, tedaj je znašalo njegovo razmerje 1:2. Če pa je na pr. imel 5 letni umsko starost 6 let, je znašalo njegovo razmerje prav tako 1:2. Na predlog W. Sterna so uvedli mesto umskega razmerja inteligentnostni količnik ali krajše I. K. (Intelligence Quotient), ki ga je Amerikanec L. M. Terman (prof. Standfordske univerze) populariziral s svojo Standfordsko revizijo Binet-ovih testov (l. 1916). Z uvedbo I. K. se vrednost »umskega razmerja« ni spremenila, kajti pomnožili smo ga le s 100. V tem smislu je umska starost povprečnega poedinca enaka njegovi kronološki starosti, ali drugače izraženo: povprečni ali srednji I. K. = 100.

Vedno večji razmah testov je zahteval točnejše in temeljitejše razlaganje rezultatov. Norma umske starosti ni več zadoščala, saj je bila njena uporaba omejena le na krog povprečnih poedincev te ali one starosti. Nastala je potreba, da se določi metoda, s katero bo mogoče postaviti norme za katero koli skupino poedincev. Rešitev tega problema se je pokazala

v »standardizaciji« testov, ki je še danes najaktualnejše vprašanje v testologiji.

Rezultate, ki jih je dosegla skupina poedincev, si razlagamo s posebnim ozirom na tisti rezultat, ki kaže sredino, oziroma srednjo ali tudi povprečno zmogljivost vseh poedincev. V tem je začetek standardizacije, brez katere si testov sploh ne moremo misliti.

Test standardizirati se pravi, očrtati mu okvir tistih razmer in zahtev, za katere naj velja. Iz tega tudi sledi, da se standardizacija toliko spreminja, kolikor se spreminjajo te zahteve in razmere. Z ugotavljanjem razmer in zahtev nekega kraja določamo njegov »življenjski standard«. Za standardizacijo testa moramo najprej s predhodnimi preizkusi ugotoviti take ali drugačne vplive krajevnih razlik kakor tudi socialno-ekonomskih, starostnih, spolnih in drugih razlik.

Upoštevajoč vse te zahteve moramo končno ugotoviti med njimi neko skupnost ali osrednjo os, od katere se bolj ali manj odmikajo tako bolj kot slabše klasificirani poedinci. Zaradi točnosti vseh teh dejstev se moramo zateči na statistično polje in z zakoni verjetnosti u t r d i t i (eichen; étalonneur) dobljene ugotovitve.

### § 17. Distribucija rezultatov.

Problem o sestavi testov in problem o tolmačenju doseženih rezultatov sta med seboj v najtesnejši zvezi. Pri sestavi testov moramo razen prvega in odločilnega psihološkega vprašanja upoštevati še druge zahteve.

Ko test preizkušamo z velikim številom poedincev, tedaj morajo rezultati ustrezati statističnim zakonom, če naj smatramo, da je test dobro sestavljen in sploh uporaben. Najprej moramo ugotoviti odstotek poedincev, ki more pravilno rešiti vsako vprašanje ali nalogo iz testa. Tako dobimo prvi zunanji kriterij testa, ki se nanaša na samo sestavo testa in ki se opira na statistično obravnavo rezultatov. To je d i s t r i b u c i j a ali r a z t r e s e n j e r e z u l t a t o v.

Če ima test tako razvrstitev lažjih in težjih vprašanj, da ustrezajo različnim stopnjam sposobnosti tiste skupine, ki jo preizkušamo, potem bo r a z t r e s e n j e ali d i s t r i b u -



cija rezultatov v velikem skladu z normalno verjetnostno krivuljo. To se pravi: največje število rešitev bo zavzelo mesto okoli povprečja in število rešitev bo strmo padalo v precej enaki meri tako nad kot pod tem povprečjem. Test bo ustrezal temu kriteriju, če se bo normalni distribuciji čim bolj približal.

Za pravilno razumevanje tega vprašanja bi bilo pač potrebno, da bi se natančneje seznanili z normalno verjetnostno krivuljo, kakor jo je že matematik Gauss utemeljil. Že beseda sama nam pove, da je izhodišče temu obravnavanju le verjetnost, kar vrže včasih svojevrstno senco na razlago vseh dobljenih rezultatov. Pravilo verjetnosti namreč pravi, da je veljavnost verjetnostnih zakonov tem večja, čim večje bo število preizkusov in nasprotno, namreč da bo veljavnost tem šibkejša, čim manj bo preizkusov. Ugotovimo lahko dejstvo, da se tudi rezultati inteligentnostnih testov ravnaajo po verjetnostnih zakonih, to je po normalni verjetnostni krivulji. V grafičnem prikazu je številnost rezultatov ordinata, višina doseženega rezultata pa je abscisa. Distribucija rezultatov pri inteligentnostnih testih bo torej sestavljena iz abcise, na katero bodo nanešeni I. K. (inteligentnostni količniki), oziroma število doseženih točk, in iz ordinate, na katero nanesemo število poedincev, ki so dosegli take in take količnike.

### § 18. Razvojna krivulja. (*Growth-curve.*)

Test ustreza statističnemu zakonu distribucije ali raztresenja tedaj, če se razvrstitev rezultatov čim bolj približuje idealni verjetnostni krivulji. Po tej zahtevi naj se ravnaajo posamezne starostne stopnje, kakor tudi večje ali manjše starostno obdobje, ki ga obsega lestvica.

Drugi kriterij za sestavo testa, ki se tudi opira na statistično obravnavo rezultatov, sloni na progresiji povprečnega ali srednjega rezultata v postopno višjih starostnih dobah. Ta kriterij se nanaša samo na sestavo testov obče umske sposobnosti ali inteligentnostnih testov. Pri posebnih sposobnostih namreč ne gre toliko za vprašanje, ali je preizkušanec dosegel

v določeni starostni dobi določeno stopnjo razvoja, temveč gre večinoma le za vprašanje, kakšna je sploh ta ali ona poedinceva posebna sposobnost, ne gledé na starostno stopnjo.

Če vemo, kako rastejo povprečni rezultati, lahko to ugotovitev uporabimo za kriterij testa. N. pr. za leta pred adolescenco dobo velja, da rastejo povprečni rezultati enakomerno. V tem smislu bi bil test neprimeren, če povprečni rezultati posameznih letnikov prehitro ali prepočasi rastejo. To se pravi: leta pred adolescenco dobo morajo imeti proporcionalen in enakomeren porast povprečnega rezultata, od najnižje do najvišje sem spadajoče starostne stopnje. Iz te ugotovitve sledi, da bo progresivna krivulja sestavljena iz porasta povprečnih rezultatov v zaporednih letih zelo podobna ravni črti. In če bo distribucija rezultatov od leta do leta enaka, tedaj bo tudi oblika progresivne krivulje obenem potrjevala pravilnost in veljavnost inteligentnostnih količnikov. Nastaja pa vprašanje, ali je rast povprečnih rezultatov enakomerna v vseh starostnih stopnjah.

Tako smo zadeli na problem, ki govori o stopnjah umske rasti pri različnih starostnih, oziroma različnih razvojnih stopnjah. Stopnje te rasti grafično izražamo s krivuljo umske rasti (the mental growth curve). Ta krivulja naj pokaže, v kateri je razvoj najhitrejši, kdaj zopet počasnejši, in tudi še to, ali so te stopnje »navzgor« enakomerne in strme.

Drugo vprašanje, ki ga je treba rešiti, je, kako se razlikujejo krivulje umske rasti pri poedincih. Ugotoviti je treba, ali so paralelne, ali se med seboj križajo in ali bodo te razvojne krivulje kulminirale vse na isti stopnji.

Z razvojno krivuljo ugotavljamo tudi raznoličnost ali variabilnost sposobnosti poedincev v različni starosti. To je važno zlasti pri preučevanju, kako nagla je rast pri poedincih različne umske višine (povprečnih, podpovprečnih), kajti če je variabilnost večja pri starejših poedincih, tedaj sledi, da poteka umska rast pri podpovprečnem počasneje, pri nadpovprečnem pa hitreje. Drugače rečeno: krivulje rasti se bodo v tem primeru odmikale druga od druge in ne bodo vzporedne. Končno pa moremo z razvojno krivuljo ugotoviti, kdaj doseže umska rast svojo stopnjo zrelosti.



Lahko tudi doženemo, kdaj se ustavi rast inteligentnosti. Terman je tako določil, da dozori inteligentnost pri povprečnih s 14 leti, pri povprečnih s 15 leti in pri nadpovprečnih med 16. in 18. letom. Mnogo šibkeje je podprta trditev, da bi ta zrelost nastopila že s 13 in pol leti (umske starosti!), kakor so bili standardizirani testi za vojaške rekrute v Ameriki (Army Alpha Test).

Določevanje inteligentnostne zrelosti precej zavisi od prirode in vsebine testa, s katerim hočemo to preizkusiti. S tem dejstvom si moramo tudi razlagati, zakaj smo ugotovili pri tem testu zrelost na tej stopnji, z onim testom pa na višji ali na nižji stopnji.

Poseben moment, ki vpliva na obliko razvojne krivulje, je tudi izbira poedincev, s katerimi nameravamo preizkusiti test ter na osnovi doseženih rezultatov sestaviti to krivuljo. Če bi n. pr. preizkusili v enem razredu vse poedince, bi zasledili, da so prav za prav mlajši z ozirom na svojo kronološko starost bistrejši in starejši.

Razvojne krivulje pa se pri poedincih zelo razlikujejo med seboj, če jih posebej določamo za vsako posebno ali specialno sposobnost, kajti različnost krivulj zavisi od prirode ter od stopenj rasti poedinih sposobnosti. Kulminacija rasti je zelo različna, vendar pa je oblika krivulje v veliki meri zavisna od vprašanj, ki sestavljajo test.

Prav tako se med seboj razlikujejo razvojne krivulje raznih posebnih sposobnosti istega poedinca. Zato se tudi pri sestavi testa običajno ne oziramo na te razvojne krivulje, temveč sestavimo test samo na osnovi povprečnih zahtev, ki jih n. pr. nek poklic zahteva od posameznika.

Poudariti moramo, da zapazimo očitno razliko med krivuljo, ki jo začrtajo rezultati občega umskega testa, in med krivuljo, ki jo dajejo rezultati testov za posebne sposobnosti. Razumljivo bo, da bodo srednji rezultati napredovali od leta do leta bolj enakomerno v običih testih, katerih razmah med lažjimi in težjimi vprašanji je zelo obširen, manj pa v specialnih testih. Zato bodo tudi sposobnosti, ki se nanašajo na prav diferencirane storitve, kulminirale v različnih starostnih stopnjah in bodo krivulje poedincev zelo različne.

Kakšne so razvojne krivulje pri različnih poedincih, to določamo tudi z variabilnostno mero. Tako bomo razumeli, da se variabilnost veča, če se razvojne krivulje poedincev postopoma odmikajo druga od druge; ako bi tekle vzporedno, bi bila variabilnost v vseh starostnih dobah enaka. Preizkusi so do sedaj pokazali, da se starejši poedinci (mladostniki) med seboj bolj razlikujejo in da so razlike večje kot one razlike, ki jih moremo ugotoviti med mlajšimi.

Vemo tudi, da je eden izmed pogojev za veljavnost I. K. ta, da je porast variabilnosti od starosti do starosti z ozirom na vsebino testa pravilno določen. To variabilnost lahko določimo šele po daljšem preizkušanju izbrane skupine poedincev, kar igra važno vlogo pri standardizacijskem postopku.

Rekli smo že, da se razvojne krivulje (mental growth curve) razlikujejo, če smo jih sestavljali na onovi rezultatov različnih testov. Enako velja ta ugotovitev, če bomo govorili o mejni starosti, pri kateri se umska rast ustavi. Tako začno pri Presseyevi lestvici (Cross-Out Scale) povprečni rezultati padati že pri 9. letu, pri Otisovi lestvici pa šele pri 17. letu. Določitev mejne starosti je naloga standardizacije, pri kateri bo zlasti odločala izbira gradiva, ki ga bomo v ta namen uporabili. N. pr. Otisova lestvica (Otis Intelligence Scale) bi bila za tako določevanje zelo primerna. Presseyeva lestvica nam pa s svojo omejitvijo ne bo dokaz, da se umska rast ustavlja že tako zgodaj, temveč nas opozarja le na svojevrstno omejitev postopice lažjih in težjih vprašanj tega testa.<sup>11</sup>

Vprašanje umske rasti, njene variabilnosti in mejne starosti sta temeljito preučila Angleža Richardson in Stokes,<sup>12</sup> ki sta se poslužila tako imenovane absolutne lestvice, kakršno je prvi sestavil L. Thurstone. Njuni rezultati so pokazali, da se razvojne krivulje sposobnosti, ki smo jih preizkušali z različnimi vprašanji, med seboj razlikujejo, vendar pa se vse zelo približujejo povprečni krivulji.

Zanimivo je, da je povprečna razvojna krivulja skoro docela skladna s tako zvano »G« krivuljo, ki je po obliki po-

<sup>11</sup> Frank Freeman: *Mental Tests*, Chicago, 1926, str. 252.

<sup>12</sup> Richardson C. A. and Stokes C. W.: *The Growth and Variability of Intelligence*, Cambridge, 1933.



dobna splošno znani magnetični krivulji iz fizike,<sup>13</sup> in katere vrednost se da izraziti s formulo:  $Y = a^{bx}$

Na osnovi statistične analize dobljenih rezultatov je prišel Richardson do zaključka, da absolutna variabilnost inteligentnosti narašča s starostjo in da je obseg inteligentnosti pri vsaki starosti sorazmeren s srednjo višino inteligentnosti te starosti.

Po analogiji starostne krivulje, ki je identična z »G« krivuljo, sklepamo, da delujeta na polju dve tendenci, in sicer anabolična ali konstruktivna in pa katabolična ali razdiralna. Umska rast se prikazuje kot rezultanta teh dveh komponent. Sklicujoč se na to krivuljo, bi tudi lahko sklepali, da ima poedinec že pri rojstvu  $6\frac{1}{2}\%$  vse svoje sposobnosti. Svojo najhitrejšo rast doseže sposobnost okoli 4. leta, ko dovrši nekako  $\frac{1}{3}$  svojega razvoja. Že pri 6. letu pa je razvite okoli 50 % vse sposobnosti, medtem ko doseže 96 % svoje zrelosti šele pri 18. letih in 99 % pri 24. letu. Krivulja je tudi pokazala, da predstavlja razdobje med 10. in 14. letom dobo stabilizacije, ki je zelo važna za otrokov duševni razvoj.

### § 19. Norme in standardi.

S tem, da smo ugotovili za poedini test njega distribucijo ali raztresenje rezultatov, potem pa še temu odgovarjajočo razvojno krivuljo, smo izsledili in očrtali prve standardizacijske temelje. Obenem smo tudi pripravili potrebno gradivo za določitev norm in standardov.

Norme imajo velik pomen za objektivno izračunavanje in tolmačenje doseženih rezultatov. Če ocenjevanje ne bi bilo objektivno, bi bilo preveč prepuščeno subjektivnemu naziranjju, razen tega pa tudi rezultatov ne bi mogli med seboj primerjati.

Že ob prvih začetkih razvoja inteligentnostnega testa je bilo vprašanje norme centralno vprašanje, in Binet je bil prvi, ki je to nalogo uspešno rešil. Kakor smo že povedali, je Binet osnoval svojo lestvico po starostnem principu ter jo zato imenujemo starostno lestvico.

<sup>13</sup> The Year Book of Education, 1935.

Če govorimo, da je test v tem smislu standardiziran, tedaj vemo, da se bo pri povprečnem poedincu (natančneje: srednjem) umska starost skladala s kronološko starostjo. V tem primeru mora tudi računsko operacija inteligentnostnega količnika pokazati rezultat 100 (t. j. I. K. = 100).

Pri nekaterih lokalnih priredbah testov se je pokazalo, da so povprečni poedinci dosegli boljše uspehe nego samo I. K. = 100. Namesto da bi se spuščali v psihološko in statistično restandardizacijo, ki naj naravna I. K. zopet nazaj na 100 s tem, da postavimo primerno težja vprašanja, lahko pomaknemo normo toliko više.

Norma je pri testu merilo v obliki srednje vrednosti (central tendency), ki smo jo določili na osnovi preizkušanja večje skupine poedincev. Poznamo pa tri vrste osrednje vrednosti in zato lahko norma predstavlja vrednost sredine (aritmetične sredine), medijana ali modusa. Pogosto se poslužujemo tudi 50. odstotka (percentila), ki pa je identičen z medijanom, le da je metoda izračunavanja drugačna.

Omenjene statistične vrednosti pridejo do izraza zlasti pri točkovni lestvici, kjer izračunavanja uspeha ne izrazimo z inteligentnostnim količnikom, ampak z določenim številom točk. V tem je bila tudi preosnova Binetove lestvice, ki se ne nanaša na njen princip, temveč je le njena metodična sprememba.

Pri točkovni lestvici ne govorimo več o umski starosti; to vlogo prevzema norma, ki znaša za to starost neko število točk, za drugo pa drugo število točk. Zavedati se moramo, da ima vsako psihološko merjenje le toliko vrednost, kolikor primerjamo posamezen rezultat z rezultati, ki so jih dosegli drugi poedinci. Srednja mera ali medijan pa kot računsko izpeljanka predstavlja izhodišče vsej razlagi rezultatov. Izhodišče razlagi rezultatov ali norma je za vsako starostno stopnjo drugačna in zato prikazujemo norme v obliki tabel, ki jih imenujemo informacijske table (information table).

V tej zvezi je potrebno razjasniti razliko med normo in standardi, kajti ti dve besedi se često istovetita. Standard nekega testa za določeno skupino poedincev je fisto območje v lestvici dosegljivih točk, v katerem se gibljejo rezultati reše-



vanja tega testa. Meji tega območja sta števili točk, ki izražata najboljše in najslabšo rešitev. To območje razdelimo v posamezne stopnje in tako nazorno prikazujemo, za katerimi stopnjami moramo stremeti. Če je test že standardiziran, lahko za poedinca tiste skupine takoj ugotovimo, ali je njegov rezultat povprečen ali nadpovprečen. Posebno pa nam standardi pokažejo, za koliko je poedinec podpovprečen ali nadpovprečen.

Standarde izražamo s 25. ali s 75. odstotkom (percentilom) ali s skrajno mejo, to je s 100. ali s 1. odstotkom. Na ta način bomo videli, kako nizko ali kako visoko segajo rezultati. Dve različni standardizaciji istega testa se v tem smislu lahko med seboj po standardu zelo razlikujeta, čeprav bi bila norma ista. V tem primeru je srednji rezultat pri obeh standardizacijah enak, mejni rezultati pa so različni.

Norma je torej srednji rezultat standardiziranega testa. Prav za prav je tudi norma neka standardna stopnja; je standard srednjega rezultata. Zato velja, da je vsaka norma obenem tudi standardna stopnja, ni pa vsaka standardna stopnja tudi norma. Skratka, v tej primerjavi so norme tiste stopnje, ki so tipične za poedince te in te starosti, ali tudi tega in tega razreda, kraja, socialnega sloja itd. Tako bomo razlikovali starostne norme (age norm), razredne norme (grade norm), kombinacijo teh dveh, to je starostno-razredne norme (age and grade norm), in druge.

Kakor smo rekli, izražamo norme s številčnim indeksom, ki ga izračunavamo s pomočjo statističnih metod. V tem okviru je medijan tisti rezultat v večji skupini rezultatov, nad katerim in pod katerim se nahaja še 50 % ostalih rezultatov iz te skupine. Na pr. če imamo 65 rezultatov urejenih v naraščajoči in padajoči postopici, bo padel medijan na 33. rezultat.

Pri določevanju norm se pojavlja eno najosnovnejših vprašanj, namreč, za koga naj te norme sploh veljajo. V smislu teorije se glasi odgovor, da bodo veljale norme samo za tiste poedince, ki so bili všteti s svojimi rezultati. Ta zahteva pa je seveda praktično neizvedljiva. Rešitev tega problema se je omejila na izbrane poedince, ki naj bodo verodostojni zastopniki svojih sovrstnikov.

Z drugimi besedami se to pravi, da prav za prav ni mogoče določiti neke perfektne norme, ki naj bi veljala za vse poedince brez izjeme. Tej normi se lahko z večjo verjetnostjo približamo, če upoštevamo zlasti krajevne ali miljejske razlike, socialno-ekonomske, rasne, spolne itd.

Kako izbrati verodostojne zastopnike, to je bila naloga, ki so si jo zastavljali psihologi že od vsega začetka. Dolgo so si pomagali z dvema metodama. Po eni izmed teh so preučili vse zgoraj navedene momente ter se omejili na okolico, ki so jo smatrali za najbolj tipično za vse širše področje. Po tej metodi je bila standardizirana tudi Standfordska revizija leta 1916. Druga metoda, ki je bila primernejša za skupinski test<sup>14</sup> (group test), pa je bila ta, da so preizkusili veliko število poedincev iz različnih okrajev, slojev, razredov itd., ki so bili skupaj v enakem medsebojnem razmerju, kakršno je veljalo za ves kraj, kateremu so bili testi namenjeni.

Najbolj dovršeno je rešil ta problem L. Terman<sup>15</sup> v svoji drugi standardizaciji Binetovih testov iz leta 1937. Pri tem postopku se je ravnal po točno določenih statističnih podatkih, ki jih je dobil na osnovi standardiziranih metod (n. pr. za določevanje socialno-ekonomskega položaja, krajevnih ali miljejskih razlik itd.).

Vprašanje krajevnih razlik, ki ima svoje zahteve tudi v standardizaciji testov, si razlagamo v tem smislu: tudi poedinci, od katerih bi sicer pričakovali enako stopnjo inteligentnosti, so različni, ker so pod zelo močnim vplivom različnega miljeja. Vse podatke, ki jih tako naberemo, uredimo v tako imenovano krajevno distribucijo (geographical distribution), ki naj pokaže podobnost z normalno verjetnostno krivuljo.

Vprašanje krajevnih razlik, oziroma miljeja, pa je tesno povezano z vprašanjem o socialno-ekonomskem položaju, ki ga prav za prav pri standardizaciji sploh ne izvajamo ločeno.

<sup>14</sup> Skupinski test je v nasprotju z individualnim oni test, ki nam omogoča, da preizkušamo poedince istočasno tudi v skupinah. To so po večini pismeni testi.

<sup>15</sup> L. M. Terman and M. Merrill: *Measuring Intelligence*, 1937, New York & London.



Za določevanje socialno-ekonomskega položaja (socio-economic status) otrok, oziroma njih staršev poznamo že več lestvic, sestavljenih in standardiziranih prav v ta namen. Terman je v svoji drugi reviziji leta 1937 uporabil 4 take lestvice. Ena od teh je n. pr. Simsova lestvica, ki razdeljuje na osnovi izpolnjenih vprašalnih pol vse poedince v pet skupin:

- a) Višji poklici, strokovnjaki in drugi; kratka ljudje z visokošolsko naobrazbo. To so zdravniki, inženirji, profesorji i. t. d.
- b) Trgovski, uradniški položaji in drugi ljudje s srednjo naobrazbo.
- c) Nižji uradniki, obrtniki, srednji trgovci.
- d) Kvalificirani delavci, mali trgovci in mali obrtniki, trgovski pomočniki, natararji itd.
- e) Navadni delavci, nosači, vozniki, dninarji itd.

Razdelitev poedincev v te skupine se ne vrši že nekako v naprej, temveč šele na podlagi izpolnjene vprašalne pole. V tem je tudi vrednost in objektivnost te metode. Vprašanja se n. pr. nanašajo na število sob doma, na šolsko izobrazbo staršev, bratov in sester, na zanimanje za koncerte, gledališče, na učenje glasbe, plesa, na športno udejstvovanje, na velikost domače knjižnice, na članstvo v družabnih organizacijah itd. Na osnovi te metode bo opredeljen marsikateri obrtniški otrok v višji položaj, lahko pa se tudi otrok z »višjega« položaja kvalificira za nižjega.

Terman se je poslužil še drugih takih lestvic. Klasifikacijo otrok po poklicu njih očetov je izvršil po Taussigovi 5 razredni metodi, ki jo je še po svoje predelal v smislu statističnih podatkov ljudskega štetja iz leta 1930. Obe ti metodi sta sicer zelo temeljiti in priporočljivi za standardizacijo, vendar puščata odprto vprašanje glede razporeditve kmetjskih otrok, ki bi ga zlasti pri slovenski restandardizaciji morali nujno upoštevati.

Tudi Termanovi psihologi so se tega zavedali ter so vršili večmesečne specialne preizkuse v izrazito kmetjskih krajih, kot n. pr. v Indiani, Kentuckyju in hribovskih krajih Virginije.

Pri takih preizkusih je treba upoštevati tudi dejstvo, da so skupnosti med nižjimi položaji mnogo ožje in tesnejše kot med višjimi in lahko rečemo, da so položaji toliko bolj diferencirani, kolikor višji so. Razdelitev po socialno-ekonomskem položaju zahteva končno tudi odstotno razvrstitev ter tako določitev distribucije, ki bo primerna za standardizacijo testov. To se pravi, povprečni položaj bo izhodišče pri utrjevanju naše lestvice tako za boljše kot za slabše otroke.

Pri vsem tem standardizacijskem postopku je treba upoštevati tudi šolsko vprašanje. Iz teh krajev je izbrati tiste šole, ki so po naših ugotovitvah v povprečnem socialnem položaju. Razumljivo je, da bomo dobili najlepšo homogenost učencev v obdobju starosti od 6. do 14. leta. Najvišje starostne stopnje ne kažejo take homogenosti, kajti višje šole, kamor prištevamo učence od 15. do 18. leta, se mnogo bolj razločujejo (diferencirajo) med seboj kakor nižje in zato je pri njih tudi težje ugotoviti šolo s pravim povprečnim ekonomskim položajem. Upoštevati je treba n. pr. tudi šolsko obveznost krajev, obči značaj tamkajšnje šolske mladine in pa število šolo obiskujočih po 14. letu.



## V. DEL.

**Načela za izbiro testov**

Pravilno distribucijo rezultatov in odgovarjajočo razvojno krivuljo prištevamo med zunanje kriterije testa. Z ozirom na različne namene testov pa je treba tudi določenih načel za izbiro testov in razen tega moramo upoštevati še razna druga merila, kakor n. pr. občo uporabnost testa, lahko interpretacijo itd.

Načela, po katerih izbiramo teste, so predvsem veljavnost (validity), zanesljivost (reliability), objektivnost (objectivity) in občutljivost (discriminative power) testa.

*§ 20. Veljavnost.*

Večkrat se vprašamo, ali ti in ti testi res merijo določene kvalitete ali sposobnosti, in tedaj nam gre za vprašanje, koliko je test veljaven. Veljavnost nekega testa docela zavisi od njegove vsebine. Izražamo jo tudi statistično in sicer po korelaciji z ugotovljenimi praktičnimi zahtevami. Veljavnost testa bo tem večja, v čim večji meri test res preizkuša tisto storitev, ki jo želimo preizkusiti. N. pr. poklicni testi (vocational test) imajo tem večjo veljavnost, čim bolj preizkušajo zmogljivost, ki je potrebna za ta ali oni poklic.

Test, ki naj meri sposobnost za nek poklic, ustreza načelu veljavnosti, če bodo poedinci, ki so se v tem poklicu izkazali z nadpovprečnimi storitvami, pokazali tudi pri testiranju nadpovprečen uspeh. Enako je tudi test, ki meri sposobnost za razumevanja čtiva, le tedaj veljaven, če pokažejo poedinci, ki

veljajo za nadpovprečne, ob tem testu izreden uspeh. To dejstvo igra važno vlogo, ker pri določevanju veljavnosti služi za osnovno izhodišče.

Za reševanje tega problema se poslužujemo različnih metod. Omejimo se n. pr. na test iz razumevanja pri čitanju. Najprej posebej preizkušamo, oziroma ocenjujemo skupino poedincev v tem smislu, da dobi najboljši 90 točk, najslabši pa 50 točk.

S tem smo v smislu šolskih ocen od odličnega do nezadostnega (1, 2, 3, 4, 5) uvrstili učence v postopico po stopnjah od 50—60—70—80—90. Ta razdelitev omogoča tudi uporabo vmesnih točk, kot 56, 69 ali 92 itd. Te številčne stopnje so poljubne in se ponajveč ravnaajo po srednjem I. K. ali po drugih merskih enotah. Če je srednji in povprečni I. K. = 100, potem bomo teh 5 stopenj postavili proporcionalno višje: 80—90—100—110—120. Predmet preizkušanja, oziroma ocenjevanja na eni strani, kakor tudi testiranje na drugi, je bil seveda popolnoma isti. To je »metoda postopnega razvrščanja«,<sup>16</sup> ki omogoča uporabo Pearsonovega korelacijskega koeficienta.

Točke, ki smo jih tako dobili, razvrstimo v padajočo vrsto, vzporedno z njimi pa razvrstimo še rezultate, ki so jih dosegli poedinci pri testni preizkušnji. Korelacija med ocenami in testnimi rezultati nam v neki meri služi za indikacijo testne veljavnosti, in sicer, kakor smo že omenili, z uporabo Pearsonovega korelacijskega koeficienta. Za pravilnost in zanesljivost računske operacije je potrebno, da šteje preizkušnja testa dovoljno število rezultatov, ker sicer ne bi mogli postaviti trdnega pravila, saj bi bila naša veljavnost mnogo prešibka.

Za ugotavljanje veljavnosti je poleg gornje metode razširjena še druga, ki je tej zelo podobna. Najprej določimo po edinim učencem njih skupni uspeh pri 5 šolskih predmetih, ki jih smatramo izmed ostalih za izrazito inteligentnostne predmete. V najboljšem primeru pokaže ta rezultat oceno 25, a najslabšem pa oceno 15, ako predstavlja red dobro (3) zadnji

---

<sup>16</sup> H. A. Greene: *Work-Book in Educational Measurement*, New York, 1937.



pozitivni red in izhodišče našemu primerjanju. Tako dobljene ocene primerjamo na osnovi korelacijske metode z izsledki pri inteligentnostnem testu in izračunani koeficient pokaže, kolika je stopnja iskane veljavnosti.

Iz navedenega primera, ki smo ga razložili in ki velja bolj za učne teste, pa lahko razberemo, da je veljavnost statistično izražena z večjim ali manjšim razmerjem med rezultati testa in med nekimi določenimi podatki, ki smo jih dobili na drug način.

Po omenjenih dveh metodah so psihologi dolgo časa preizkušali tudi veljavnost inteligentnostnih testov, dokler ni dokončno prodrla Spearmanova interkorelacijska metoda, ki je slonela na njegovi teoriji »dveh činiteljev«. Upravičeno je Spearman opozoril, da gornji metodi nista dovolj zanesljivi za ugotavljanje veljavnosti inteligentnostnim testom, ker da jeta pri ocenjevanju preveč prednosti znanju, ki se nikakor ne sme zahtevati pri inteligentnostnih testih. Poleg tega pa tudi ni zaupati objektivnosti šolskih ocen, ki bi v tem oziru lahko pravemu inteligentnostnemu testu znatno zamajale njegovo veljavnost.

Gornji dve metodi (kakor še nekatere podobne) pa sta obveljali za najboljše metodi, s katerima ugotavljamo veljavnost občnih učnih testov (educational test).

Za ugotavljanje testne veljavnosti se poleg omenjenih metod raje poslužujemo drugih, pri katerih gre za stopnjo odnosa med enim in drugim testom ter na tej osnovi odkrivamo skupne merske momente. Sorodnost ali podobnost med testi iste vrste se kaže ravno v tem, koliko vsebujejo eni in drugi tiste skupne elemente, ki se pri testiranju javljajo v enaki obliki.

Veljavnost nekega testa bomo določali neposredno na ta način, da bomo iskali korelacijo med tem testom in pa med nekim drugim merilom. Kot merilo za veljavnost testa morejo nastopati za učne teste zlasti objektivne šolske ocene, za inteligentnostne teste pa poljubne metode o kvalifikaciji inteligentnosti poedincev, predvsem pa, kakor smo že naglasili, drugi že priznani inteligentnostni testi. Rezultati že priznanega testa in onega, kateremu veljavnost šele ugotavljamo, morajo biti čim bolj podobni ali skladni med seboj. Tako s korelacijsko

metodo ugotovimo tem večji koeficient, čim večja je dejanska skladnost.

Predno pa bi dokončno priznali te korelacije, moramo vedeti, kolikšna je zanesljivost testa in pa kolikšna je zanesljivost kriterija, oziroma tistega priznanega testa, ki nam služi za ugotavljanje veljavnosti drugih testov. Na zanesljivost je zlasti treba paziti, ako vzamemo za merilo šolske ocene.

Če se izkaže, da kriterij ni zanesljiv, tedaj nam za preizkušanje veljavnosti nekega testa služijo še posredne ali indirektno metode. Princip tega posrednega ugotavljanja sloni na statističnem dokazu občega činitelja »g« in posebnega činitelja »s«. Iz tega postopka razvidimo, kolika mera činitelja »g« nastopa pri inteligentnostnem testu, ali celo pri posameznem vprašanju iz tega testa. Ta metoda se je pri statističnem ugotavljanju rezultatov izkazala sama po sebi, medtem, ko je bilo glavno težišče v tisti metodi, ki se poslužuje izračunavanja korelacij pri testih za posebne sposobnosti. Postopek, ki ugotavlja medsebojne skupne činitelje med testi, imenujemo medsebojno korelacijo ali interkorelacijo (inter-correlation).

V tem smislu bomo določali veljavnost nekemu inteligentnostnemu testu, če ga bomo preizkusili z drugim že priznanim testom. V ta namen vzoredimo in primerjamo rezultate, dobljene ob prvem testu, z rezultati, ki so jih dosegli poedinci ob drugem testu, ter na ta način določimo stopnjo njih medsebojnega odnosa, izraženega s številčnim indeksom (koeficientom).

### § 21. Zanesljivost.

Drugo nič manj važno vprašanje poleg veljavnosti je vprašanje zanesljivosti (reliability) testov. Ugotoviti je namreč treba, v koliki meri ta ali drugi test konstantno ali trajno meri tiste sposobnosti (ali znanje), za katere je določen.

Zanesljivost je po pomenu sicer tehničen izraz, zavisi pa predvsem od večjega ali manjšega obsega vsebine testa, o kateri smo podrobneje govorili pri zgradbi testa. Tako bi lahko rekli, da so bolj ali manj pomanjkljivi v zanesljivosti oni testi, ki ne vključujejo v svoji vsebini momentov kot so: nervoznost, časovne ali fizične kondicije in drugo, oziroma sploh ne



računajo s tem. Porast zanesljivosti se pokaže vedno, kadar vsebuje test čim večje število nalog, ali bolje rečeno, čim večje število nalog je poedinec mogel tudi rešiti.

Zanesljivost izražamo v statistiki s korelacijsko metodo. Korelacijske konficiente izračunavamo v primerjavi z istim, enakim ali podobnim testom. Razmerje (korelacijo) med rezultati lahko dobimo na osnovi sukcesivnega preizkušanja istega testa, lahko pa tudi na osnovi rezultatov dveh podobnih testov. Zanesljivost naj torej prikazuje nek konstantni indeks v reševanju, oziroma izvajanju nalog.

Običajno preizkušamo zanesljivost nekega testa na ta način, da ga po 4, 7 ali 9 mesecih ponovno predložimo isti ali enaki skupini poedincev. Obe vrsti rezultatov uredimo v korelacijsko razpredelnico ter s pomočjo Pearsonovega koeficienta ugotavljamo večjo ali manjšo stopnjo zanesljivosti. Ta postopek imenujemo samo-korelacijo (self-correlation) in z njo dosežemo določeno srednjo ali povprečno zanesljivost.

V statistiki je najbolj znana Spearman-Brownova formula (Spearman-Brown Prophecy Formula),<sup>17</sup> ki nam omogoča dovolj točno izmeriti zanesljivost. Ta formula se glasi:

$$r_x = \frac{n \cdot r_p}{1 + (n - 1) r_p}$$

Posamezni simboli pomenijo:  $r_x$  iskani koeficient zanesljivosti;  $r_p$  prvotni koeficient zanesljivosti, namreč tisti koeficient, ki ga dobimo od dveh različnih vrst rezultatov istega testa (lahko pa tudi, kot smo že uvodoma to omenili, med dvema zelo podobnima testoma);  $n$  pomeni v tej formuli število, s katerim bomo pomnožili dosedanjo dolžino testa.

To formulo uporabljamo v takih primerih, kadar hočemo vedeti, kolika je stopnja zanesljivosti, če bi bil test na pr. 3 krat daljši po svojem gradivu (vsebini). Recimo, da znaša koeficient ( $r_p$ ) (coefficient of reliability) 0,6 in se vprašamo,

<sup>17</sup> Garret H. E.: Statistics in Psychology and Education, New York, 1938. Objavljam s posebnim dovoljenjem avtorja, ki je profesor na Columbia univerzi v New Yorku.

kolika bo zanesljivost testa, če 3 krat povečamo njegovo vsebino, ne da bi se ta bistveno izpremenila.

V tem primeru bomo uporabili našo formulo takole:

$$r_x = \frac{3 (0.6)}{1 + (3 - 1) 0.6} = \frac{1.8}{1 + 2 (0.6)} = 0.82.$$

Rezultat te formule nam pove, da se je prvotna zanesljivost ( $r_p = 0.6$ ) povečala od 0.6 na 0.82 ( $= r_x$ ). To se pravi, da je pridobil test na svoji zanesljivosti, ko smo mu 3 krat povečali njegovo vsebino.

Vraščanje pa lahko zastavimo tudi še drugače. Test, ki smo ga ponovno uporabili, nam je pokazal koeficient zanesljivosti 0.72 (test traja na pr. 5 minut). Za koliko moramo povečati test po vsebini, da bo pokazal njegov koeficient zanesljivosti 0.90? (Opozoriti moramo, da ostane pri tem njegoova priroda in veljavnost ista.)

V tem primeru postavimo našo formulo takole:

$$n = \frac{r_x (1 - r_p)}{r_p (1 - r_x)}$$

ter zamenjamo neznanke z znanimi podatki:

$$n = \frac{0.90 (1 - 0.72)}{0.72 (1 - 0.90)} = 3.5$$

to se pravi, da moramo test povečati 3 in pol krat, ali pa, da moramo povečati časovno dolžino, ki je odmerjena za reševanje od 5 minut na 17 in pol minute, če hočemo povečati testno zanesljivost od 0.72 na 0.90. Ugotavljanje testne zanesljivosti s pomočjo te formule ima svoj pomen samo tedaj, kadar so naloge, s katerimi povečavamo test, enako težke in nudijo obenem isto zanesljivost (t. j. prvotno  $r_p$ ) kot prvotne naloge, ki so v testu.

Odveč bi bilo skoraj povdarjati, da ima tudi to razširjanje svojo točno odmerjeno mejo, preko katere ne moremo posegati brez škode. Pri tem mislimo na to, da ni mogoče povečati dolžine testa od 3 do 10 ali celo 15 krat, oziroma koeficient zanesljivosti od 0.2 na 0.8. Pri 4 ali 5 kratnem podaljšanju testa se namreč lahko dogodi, da nam koeficient zanesljivosti na-



pove mnogo več, kakor nam sploh more test pokazati. Zato bomo vedno iskali vzroke za nizke korelacijske koeficiente v drugi smeri, in sicer v predelavi vsebine, oblike itd.

Navedemo formulo s pridom uporabljajo tudi na drugih znanstvenih področjih.

### § 22. Objektivnost.

Tretji moment, ki odloča pri izbiri testov, je objektivnost (objectivity). Večkrat se moramo vprašati, ali je zgradba testa taka, da se nam pri ocenjevanju in izračunavanju rezultatov ne bodo pojavili osebni ali subjektivni vplivi. Objektivnost testa je neka posebna kvaliteta, ki se kaže v enakosti ali vsaj podobnosti rezultatov posameznih preizkušenj, ki so jih ugotovili različni testatorji in ki smo jih ugotovili v različnih vnanjih okoliščinah. To razmerje ima zato tudi svojo stopnjo, ki jo v statistiki izračunamo s korelacijsko metodo.

Eden izmed načinov določevanja te korelacije je ta, da določimo korelacijo med rezultati, ki smo jih dobili pri prvem in pri ponovnem testiranju.

S tem preizkusimo, koliko nastopa subjektivni moment pri nas samih. Razen te pa poznamo še bolj razširjeno metodo, po kateri računamo korelacijo med dvema vrstama rezultatov, ki sta jo ugotovila dva testatorja v različnih okoliščinah. Te različne okoliščine so n. pr. sprememba kraja in časa (n. pr. dopoldne, popoldne, poleti in pozimi in podobno).

Perfektna objektivnost bi v teoretičnem smislu nastopila samo tedaj, kadar se s statistično metodo izkaže, da je pri istem testu v dveh ali več vrstah rezultatov, ki so jih dosegli različni testatorji v različnih prilikah, med vsako dvojico posebej perfektna pozitivna korelacija.

Zbiranje posameznih vrst rezultatov za ugotovitev objektivnosti zahteva posebno razpredelnico, kakor nam to kaže tabela.

Zap. štev.	Testator A	Testator B	Testator B
	I	I	II
1	22	22	19
2	21	21	20
3	24	21	21
4	24	24	25
5	24	24	24
6	24	23	24
7	23	22	23
8	21	26	27
9	24	24	25
10	26	25	26
11	27	26	27
12	22	22	22
itd.	itd.	itd.	itd.

V drugem stolpcu so navedeni rezultati, ki jih je prvič ugotovil testator A; v naslednjem pa rezultati, ki jih je ugotovil testator B prvič.

Ko smo že rešili vprašanje veljavnosti in zanesljivosti, tedaj lahko šele uspešno preizkušamo objektivnost. Ta objektivnost se bo pokazala s korelacijo v večji ali manjši stopnji, in sicer v našem primeru med prvo in drugo vrsto rezultatov. Po drugi metodi, ki smo jo navedli, prav za prav preizkuša testator samega sebe, kajti tedaj ugotavlja korelacijo med dvema vrstama rezultatov, ki jih je sam ugotovil v različnih okoliščinah. V tem primeru se ugotavlja korelacija med vrstama rezultatov iz drugega in tretjega stolpca naše tabele. Če bi bila ta korelacija zelo visoka, potem bi lahko rekli, da test ni sprejemljiv za razne subjektivne vplive, niti za razne spremljajoče okoliščine. Za pravilnost in zanesljivost računa samega pa je potrebno, da obsega preizkušnja testa dovoljno število rezultatov.

### § 23. Občutljivost.

K osnovnim zahtevam, ki jih stavljamo na vsako merjenje s pomočjo testov, spada tudi zahteva, da moremo na njegovi osnovi uvrstiti poedinca v neko postopico sposobnosti ali tudi



znanja. Ta postopica naj pokaže, koliko se poedinci skladajo ali ne skladajo po svoji dejanski sposobnosti (ali znanju). To razlikovalno svojstvo mora imeti vsak test, če lahko o njem rečemo, da je zadostil načelu občutljivosti (discriminative power). Ker pa je test sestavljen iz večjega ali manjšega števila vprašanj, bo zadostil temu načelu šele tedaj, če je vsako vprašanje posebej občutljivo do skrajne možnosti. Občutljivost testa moramo zato preučevati najprej ob vsakem posameznem vprašanju ali nalogi.

Če govorimo o občutljivosti nekega testa, tedaj mislimo na različne kvalitete (ali veličino) odzivov ali odgovorov, ki naj jih pokaže testiranje poedincev z različno stopnjo sposobnosti. Poedinec z nižjo stopnjo sposobnosti ne bo smel uspeti pri tistih nalogah, ki jih zmore le poedinec z višjo stopnjo, akoravno spada sposobnost obeh še vedno v tisti razmah sposobnosti, ki ga smatramo za normalnega.

Koliko je občutljivo vsako vprašanje v testu, lahko določimo z raznimi statističnimi metodami. Običajno se v ta namen poslužujemo kontingenčnega koeficienta (coefficient of contingency), včasih pa tudi bolj enostavnih metod.

Mislimo si, da moramo preučiti občutljivost nekega testa. Za skupino poedincev, ki smo jih izbrali v to preizkušnjo, naj velja, da obstajajo v okviru normalnega razmaha sposobnosti razlike med slabšimi in boljšimi. Vseh poedincev je 100. Na osnovi pokazanih rezultatov jih razdelimo v dve skupini. Prvo skupino bodo tvorili oni, ki so dosegli boljši rezultat od srednjega rezultata ali medijana, drugo skupino pa oni, ki so dosegli slabši rezultat od medijana.

Nato ugotavljamo odstotek poedincev, ki niso uspeli pri prvem vprašanju iz (prve) boljše skupine, in koliko odstotkov poedincev ni uspelo iz druge, slabše skupine. To ugotavljamo za vsako vprašanje, odstotnost med obema vrstama rezultatov pa primerjamo. Spodnja tabela naj ponazori navedeni primer!<sup>18</sup>

---

<sup>18</sup> Greene and Jorgensen: *The Use and Interpretation of Elementary School Tests*, New York, 1935. (Objavljeno s posebnim dovoljenjem avtorjev.)

Občutljivost vprašanj iz testa bomo izrazili v % neuspeha iz boljše ali slabše skupine:

Vprašanje: boljša sk.: slabša sk.:

itd.	itd.	itd.
1	2%	5%
2	4	5
3	7	13
4	10	9
5	12	18
6	18	30
7	22	36

Iz tabele je razvidno, da pri prvem vprašanju iz testa ni uspelo iz boljše skupine 2% poedincev, iz slabše skupine pa 5%.

Razmerje med obema skupinama je premajhno, zato je občutljivost slaba. Še slabše je to razmerje pri 4. vprašanju, ko pokaže slabša skupina boljši uspeh nego boljša skupina. Da zadostimo načelu občutljivosti, moramo to vprašanje ponovno preučiti. Najboljšo občutljivost pa pokaže 7. vprašanje, kjer je razlika med uspehom boljše in slabše skupine najbolj očitna.

Pogosto je potrebno, da ugotovljamo občutljivost testa z ozirom na razlikovanje med slabšimi, srednjimi in boljšimi poedinci. Rešljivost vprašanj teh treh skupin, izraženo v odstotkih, primerjamo med seboj, pri čemer moramo upoštevati, do so odstotki v primerno velikih razdaljah in da postopno naraščajo pri vsakem nadaljnjem vprašanju ali nalogi v testu.

Razdelitev v tri skupine, med slabše, srednje in boljše poedince se nanaša na normalni razmah sposobnosti. Lahko pa preučujemo rešljivost tudi v še večjem razmahu. To storimo tedaj, če hočemo ugotoviti razmerje med podpovprečnimi, povprečnimi in nadpovprečnimi poedinci.



## Sklepne besede

Testologiji gre zasluga, da je utrla pot praktičnemu preučevanju poedinčevih sposobnosti. Pokazala nam je, kako lahko po storitvi določimo razlike in stopnje človekove zmogljivosti. Test nam služi kot praktično sredstvo za diagnozo sposobnosti pri tem ali onem poedincu in v tem je tudi njegov pomen za pedagogiko.

Preučevanje poedinčevih sposobnosti na osnovi testnih rezultatov pa je omejeno. S testom merimo samo količino zmogljivosti, ki je v storitvi dejansko aktualizirana. Tako nam je pot zaprta do vseh tistih gibal v človekovi duševnosti, ki odločajo o tem, da sploh prihajamo od zmogljivosti do dejavnosti. Vprav te sile pa vplivajo tudi na to, da se včasih postavljajo izražanju naših zmogljivosti nepremostljive ovire ali pa odpirajo človekovemu dejstvovanju najsvobodnejši razmah. Z vida testne psihologije radi tega ne moremo priti do razumevanja smiselne strukture poedinčevih osebnosti.

V tem je omejitev uporabnosti testnih rezultatov v praktičnem življenju, pa tudi v pedagogiki. O kvaliteti in kvantiteti storitve namreč naposled odloča le format celotne osebnosti, ne pa višina te ali one zmogljivosti. Življenje je to že nešteto-krat dokazalo. Tudi vzgojniku še ne zadostuje, če ve, kaj ta ali oni doraščajoči človek v danem trenutku zmore, saj si je stavil vprav nalogo, izoblikovati ga takó, da bo v življenju čim bolj uspeval. Rezultati testov mu sicer dajejo marsikak napotek, toda test ne bo mogel nikoli nadomestiti temeljitega preučevanja, oprtega na visoko teoretično izobrazbo, združeno z intuitivnim pogledom na značilne odtenke duševnosti poedinca.

Kljub vsem omejitvam ne smemo zmanjševati velikega pomena, ki ga ima test za pedagogiko. Z veliko verjetnostjo moremo v neki skupini določiti bolj ali manj nadarjene, dobimo vpogled, kako vplivajo na razvoj zmogljivosti razni tipi

miljeja (mesto, dežela, industrijski kraj), socialna pripadnost, poklic staršev itd.; končno pa iz rezultatov, ki jih je dalo testiranje nekega individua več ali manj tudi lahko črpamo vidike za usmerjenje celotnega izobraževalnega procesa, ki mora biti v skladu s poedinčevo individualno strukturo. Skratka, lahko rečemo, kjer koli nastaja potreba za poznavanje ali preučevanje poedinčevih sposobnosti, se moremo s pridom poslužiti testov. Kjer gre za določevanje specialnih zmogljivosti, kakor so potrebne n. pr. v raznih tehniških poklicih, pa testov sploh ne moremo več pogrešati.

Pomen testov bomo najbolje spoznali, če določimo mesto testne psihologije tudi gledé na teoretične osnove psihologije. Testna psihologija ni samo uporabna psihologija, temveč ima tudi teoretični pomen. Preučevanje poedinčeve zmogljivosti na osnovi statistično urejenih rezultatov ne služi samo praksi, temveč tudi teoretičnemu spoznanju.

Tudi behaviorizem si je v primeri z dosedanjim psihološkim gledanjem zgradil še svoje posebno izhodišče, brez katerega si sodobnega psihološkega prizadevanja ne moremo več misliti. Testna psihologija je na tej osnovi pokazala nov način gledanja na poedinčeve sposobnosti. Introspektivna psihologija razlikuje sposobnosti z vidika introspektivne metode; duhoslovna psihologija skuša razumeti duševno strukturo poedinca v smiselnem odnosu do objektivnega duha njegovih storitev; testna psihologija pa skuša meriti poedinčevo zmogljivost po količini storitve. Ta poslednji vidik lepo izpopolnjuje prvega ter drugega in rezultati testne psihologije dopolnjujejo spoznanja introspektivne in duhoslovne psihologije.

Res je, da nam more samo introspektivna psihologija povedati, kaj je n. pr. inteligentnost; da nam more samo duhoslovna psihologija prikazati strukturo in smisel duhovnih tvorb, kakor tudi poedinčevega duha; vendar pa nam šele testna psihologija s svojimi statističnimi metodami more dati odgovor na vprašanje, v koliki meri se pri poedinca ali pri neki skupini javlja ta ali ona zmogljivost v tej ali oni storitvi, kako so posamezne duševne kvalitete razdeljene v okviru nekega kolektiva in slednjič, kako se te kvalitete količinsko razvijajo.



## Seznam uporabljenе literature

1. Allen P. E. and Smith P.: Selection of Skilled Apprentices for the Engineering Trades, Birmingham, 1931.
2. American Council on Education: The Educational Record, Jan. 1939, New York.
3. American Council on Education: The Educational Record, Jan. 1938, New York.
4. Anschütz G. u. W. J. Ruttman: Alfred Binet, Die neuen Gedanken über das Schulkind, Leipzig, 1912.
5. Bakers H. J.: Manual of Directions. Detroit Advanced Intelligence Test, Detroit, 1924.
6. Baker H. J. and Leland B.: Detroit Test of Learning Aptitudes, 1935.
7. Ballard: Mental Tests, London, 1930.
8. Baumgarten F.: Die Berufseignungsprüfungen, München, 1928.
9. Baumgarten F.: Arbeitswissenschaft und Psychotechnik in Russland, Berlin, 1924.
10. Beers F. S.: The Human Side of this Testing Business, Washington, 1936.
11. Benrubi J.: Les sources et les courants de la philosophie contemporaine en France, Paris, 1933.
12. Bilimovič A.: Nauk o konjunkturah, Ljubljana, 1931.
13. Binet A.: L'étude expérimentale de l'Intelligence, Paris, 1922.
14. Binet A. et Simon Th.: La mesure du développement de l'intelligence chez les jeunes enfants, Paris, 1908.
15. Binet A. et Simon Th.: Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux, Paris, 1905.
16. Binet A. et V. Henri: La psychologie individuelle, Année psychologique, Paris, 1895.
17. Board of Education: Mental and Scholastic tests among retarded Children, London, 1923.
18. Bobertag O.: Über Intelligenzprüfungen, Leipzig, 1920.
19. Brueckner L. J. and E. O. Melby: Diagnostic and remedial teaching, New York, 1931.

20. Bureau of Educational Research and Service: Manual for Administration and Interpretation of 1939 Iowa Every-Pupil Test of Basic Skills, U. S. A., 1939, Iowa.
21. Bühler Ch.: Praktische Kinderpsychologie, Leipzig 1938.
22. Bühler Ch.: Kindheit und Jugend, Leipzig, 1928.
23. Bühler Ch. u. H. Hetzer: Kleinkindertests, Leipzig, 1932.
24. Bühler Karl: Abriss der geistigen Entwicklung des Kindes, 1928.
25. Bühler Karl: Die Krise der Psychologie, 1927.
26. Cattell R. B.: A Guide to Mental Testing, London, 1936.
27. Conn Mc M.: Educational Guidance is now possible, New York, 1931.
28. Conn Mc M.: History of the Co-operative Test Service of the American Council and the Proposed Nation Wide Testing of College Sophomores; »The Journal of Higher Education«, May 1931, New York.
29. Danzinger L.: Schulleistungstest, Leipzig, 1933.
30. Decroly O. et R. Buyse: La Pratique des Test Mentaux, Paris, 1928.
31. Decroly O. et R. Buyse: La Pratique des Test Mentaux, Atlas, Paris, 1928.
32. Ebbinghaus H.: Über eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern. 1897.
33. Eckert Ruth E.: Realism in Higher Education, Washington, 1938.
34. Ezekiel Mordecai: Methods of Correlation Analysis, New York, 1930.
35. Feder D. D.: Suggestions for Testing Programs at the College level, University of Iowa, 1939.
36. Freeman F. N.: Mental Tests, New York, 1926.
37. Garrett H. E.: Statistics in Psychology and Education, New York, 1938.
38. Gesell A.: Körperseelische Entwicklung in der frühen Kindheit, Halle, 1931.
39. Giese F.: Psychische Normen für Grundschule und Berufsberatung, Langensalza, 1920.
40. Giese Fr.: Objekts-Psychotechnik, Halle, 1930.
41. Giese Fr.: Handbuch psychotechnischer Eignungsprüfungen, 1928.
42. Gray W. S.: Summary of Reading Investigations, Journal of Educational Research, March 1939, Nr. 7.
43. Greene H. A. and Jorgensen A. N.: The use and interpretation of elementary school tests, New York, 1935.
44. Greene H. A.: Work-Book in Educational Measurement, New York, 1937.
45. Günther Just: Praktische Übungen zur Vererbungslehre, Berlin 1935.
46. Hamley H. R.: Testing Intelligence, London, 1936.
47. Hanselmann H.: Erziehungsberatung, Leipzig, 1937.
48. Hetzer H.: Kindheit und Armut, Leipzig, 1937.
49. Hetzer H.: Psychologische Untersuchungen der Konstitution des Kindes, 1937.
50. Hetzer H.: Soziologische und psychologische Studien, 1927.
51. Huizinga J.: Im Schatten von morgen, Leipzig, 1937.



52. Hunt E. P. A. and Smith P.: A Guide to Intelligence and other Psychological Testing, London, 1937.
53. Kandel I. L.: Examinations and Their Substitutes in the United States, New York, 1936.
54. Kammel W.: Beitrag zu einer Geschichte der Testforschung im XVIII. und XIX. Jahrhundert; 17. Jahrb. des Vereins für christliche Erziehungswissenschaft, München, 1926.
55. Knight Rex: Intelligence and Intelligence Tests, London, 1933.
56. Kramer E. E.: A first course in educational statistics, New York, 1935.
57. Krueger F. u. Volkelt H.: Experimentelle Kinderpsychologie, München, 1938.
58. Kuhlman F.: A Handbook of Mental Tests, New York, 1936.
59. Künkel F.: Die Arbeit am Charakter, Schwering, 1935.
60. Kwalwasser J.: Tests and Measurements in Music, New York, 1927.
61. Lay A.: Experimentelle Pädagogik, 1918.
62. Lazarsfeld P.: Statistisches Praktikum für Psychologen und Lehrer, Jena, 1929.
63. Learned W. S. and Wood Ben D.: The Student and his Knowledge, New York, 1938.
64. Lipmann Otto: Psychologische Berufsberatung, 1919.
65. Masters H. V. and C. C. Upshall: Intelligence test scores; Journal of Educational Research, 1934, No. 6.
66. Maudry M.: Die Entwicklung der Testpsychologie; »Kinderärztliche Praxis«, No. 1932, Heft 11, 12.
67. Meili R.: Psychologische Diagnostik, München, 1937.
68. Meumann E.: Vorlesungen zur Einführung in die experimentelle Pädagogik und ihre psychologischen Grundlagen, Leipzig, 1920, I, II, III.
69. Miletić Momčilo: Binet A. i Simon T., Ispitivanje i merenje inteligencije kod dece, II. izd., 1934.
70. Moede W.: Experimentelle Massenpsychologie, 1920.
71. Monroe W. S. and Buckingham R. B.: Teachers Handbook for the Illinois Examination, 1920, Bloomington.
72. Morris E. H.: Teachers Handbook for Morris Trait Index L; Bloomington, 1932.
73. Müller-Freienfels R.: Kindheit und Jugend, Leipzig, 1937.
74. Münsterberg H.: Grundzüge der Psychotechnik, Leipzig, 1920.
75. Münsterberg H.: Psychologie und Wirtschaftsleben, Leipzig, 1919.
76. National Institute of Industrial Psychology: A general explanation of Group test 33, with instruction for giving and scoring, London, 1939.
77. Nihard R.: La Méthode des Tests, Paris, 1932.
78. Osburn W. J.: Testing Thinking, Journal of Educational Research, 1934, No. 6.
79. Ozvald K.: Kaj rodi uspehe in kaj neuspehe v šoli? »Roditeljski list«, I. 1., 1937.

80. Ozvald K.: Možnosti in meje poklicnih svetovalnic. »Kronika mestne občine ljubljanske«, letnik I., 1934.
81. Piéron H.: Psychologie experimentale, Paris, 1934.
82. Pintner R. and Cumingham B. V.: Primary Mental Test for Kindergarten and first and second grades, New York, 1923.
83. Ramadanović: Psihognostički metodi u davanju saveta, Zemun.
84. Richardson C. A. and Stokes C. W.: The Growth and Variability of Intelligence, Cambridge, 1933.
85. Rorschach F.: Psychodiagnostik, Bern, 1921.
86. Rossolimo G.: Das psychologische Profil, Halle, 1926.
87. Shruballsall F. C. and Williams A. C.: Mental deficiency practice, London 1932.
88. Sims V. M.: The Measurement of Socio-Economic Status, Bloomington, 1928.
89. Sombart W.: Vom Menschen, Berlin-Charlottenburg, 1938.
90. Spearman C.: Abilities of Man, London, 1927.
91. Spranger Eduard: Lebensformen, Berlin, 1921.
92. Stern W.: The Psychological Methods of Testing Intelligence, Baltimore, Warwick & York, 1914.
93. Stern W.: Über Psychologie der individuellen Differenzen, 1900.
94. Stern W.: Die Intelligenz der Kinder und Jugendlichen, Leipzig, 1920.
95. Stern W. und O. Wiegemann: Methodensammlung zur Intelligenzprüfung, Leipzig, 1926.
96. Stevanović U.: Razviće dečje inteligencije i beogradska revizija Binet-Simonove skale, Beograd, 1934.
97. Stevanović B.: Merenje inteligencije, Beograd, 1937.
98. Symonds P. M.: Diagnosing personality and conduct, New York, 1931.
99. Terman L. M. and M. Merrill: Measuring Intelligence, London, 1937.
100. Terman L. M. and Others. The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence, Baltimore, Warwick & York, 1917.
101. Thirty-fourth Yearbook, National Society for the Study of Education, Bloomington, 1935.
102. Thomas F. C.: Ability and Knowledge, London, 1935.
103. Thorndike E. L.: The Measurement of Intelligence, New York, 1926.
104. Veber Fr.: Analitična psihologija, Ljubljana, 1924.
105. Veber Fr.: Očrt psihologije, Ljubljana, 1924.
106. Veber Fr.: Vprašanje stvarnosti; dejstva in analize, Ljubljana, 1939.
107. Wallon H.: Principes de Psychologie appliquée, Paris, 1930.
108. Wartegg E.: Gefühl und Phantasiebild, Leipzig, 1937.
108. Wood Ben D.: Measurement in Higher Education, New York, 1923.
109. Wood Ben D.: Criteria of Individualized Education; Vocational Guidance Magazin, New York, May 1936.
110. Wood Ben D.: Measurement in Higher Education, New York, 1923.



111. Wood Ben D.: The Major Strategy of Guidance; The Educational Record, Washington, Oct. 1934.
112. Year Book of Education, London, 1935.
113. Year Book of Education, London, 1936.
114. Year Book of Education, London, 1937.
115. Year Book of Education, London, 1938.
116. Year Book of Education, London, 1939.
117. Yerkes R.: Psychological Examining in the United States Army, National Academy of Sciences, Vol. XV, Washington, 1931.
118. Yule G. U. and Kendall U. G.: An Introduction to the Theory of Statistics, London, 1937.
119. Žižek dr. Fr.: Fünf Hauptprobleme der statistischen Methodenlehre, München, 1922.
120. Žižek dr. Fr.: Grundriss der Statistik, München, 1933.
121. Žižek dr. Fr.: Wie statistische Zahlen entstehen, Leipzig, 1937.
122. Žižek dr. Fr.: Meinen Kritikern; Erläuterungen und Ergänzungen zu »Grundriss der Statistik«, Leipzig, 1924.

## POPRAVEK

Na strani 70. spada 4. vrsta (kratice itd.) pod številčne kolone.

## Vsebina

	Stran
PREDGOVOR . . . . .	3
UVOD: KRATEK ZGODOVINSKI OČRT . . . . .	5
I. DEL: INTELIGENTNOST Z VIDA TESTNE PSIHOLOGIJE	
§ 1. Pojem inteligentnosti . . . . .	11
§ 2. Teorija »dveh činiteljev« — Spearman . . . . .	14
§ 3. Razmerje med občim in posebnim činiteljem »g« in »s« . . . . .	16
§ 4. Teorija skupinskih činiteljev . . . . .	19
§ 5. O prirodi in funkciji inteligentnosti . . . . .	21
II. DEL: TEST KOT SVOJEVRSTNA EKSPERIMENTALNA METODA	
§ 6. Obseg pojma . . . . .	27
§ 7. Razlika od šolskega izpita . . . . .	27
§ 8. Razlika med testom in eksperimentom . . . . .	29
§ 9. Očrt inteligentnostnega testa vobče . . . . .	30
III. DEL: MOŽNOSTI IN MEJE TESTIRANJA	
§ 10. Sposobnost in izvežbanost . . . . .	33
§ 11. Osnovni vidiki za sestavo testa . . . . .	35
§ 12. Gradivo testa . . . . .	38
§ 13. Razvrstitev gradiva . . . . .	40
§ 14. Hitrostni in sposobnostni moment v testu . . . . .	41
§ 15. Postopica nalog in zanesljivost testa . . . . .	45
IV. DEL: OSNOVE STANDARDIZACIJE	
§ 16. Nastanek in razvoj norme za testiranje . . . . .	47
§ 17. Distribucija rezultatov . . . . .	49
§ 18. Razvojna krivulja . . . . .	50
§ 19. Norme in standardi . . . . .	54
V. DEL: NAČELA ZA IZBIRO TESTOV	
§ 20. Veljavnost . . . . .	61
§ 21. Zanesljivost . . . . .	64
§ 22. Objektivnost . . . . .	67
§ 23. Občutljivost . . . . .	68
SKLEPNE BESEDE . . . . .	71
SEZNAM UPORABLJENE LITERATURE . . . . .	73



NARODNA IN UNIVERZITETNA  
KNJIŽNICA



00000420628

