

# LESK A BIEDA MATEMATICKÝCH RIEŠENÍ V PSYCHODIAGNOSTIKE<sup>1</sup>

VLADIMÍR DOČKAL<sup>2,3</sup>

**Abstrakt:** *Keď pred sto rokmi začala psychológia kvantifikovať svoje zistenia, získala si rešpekt a možnosť zaradiť sa do sveta vedy – tak ako sa veda v tých časoch chápala. Kvantitatívny prístup je charakteristický najmä pre psychometriku, ktorá sa v duchu pozitivizmu stala keď nie úplným synonymom psychodiagnostiky, tak aspoň prevládajúcim smerom v nej. Aj dnes má pozitivistická paradigma v psychológii silnú pozíciu, no čoraz viac sa uplatňujú rôzne nové (kvalitatívne) pohľady. Je preto namieste zamyslieť sa nad úlohou matematiky v psychologickom poznaní, osobitne v psychodiagnostike. Príspevok kriticky hodnotí možnosti matematiky pri koncipovaní teórie, na ktorej je psychologický test založený, pri výbere položiek a overovaní psychometrických charakteristík testu a pri tvorbe testových noriem. Dokladá užitočnosť matematických prístupov, ale upozorňuje aj na ich limity. Matematika je v psychodiagnostike nezastupiteľná, je však iba pomocníkom, ktorý nemôže nahradiť kvalifikované psychologické usudzovanie.*

**Kľúčové slová:** *psychometrika, matematická štatistika, teória testu, položky, normy*

Možnosť kvantifikovať psychické procesy urobila z pôvodne špekulatívnej psychológie „vedu“, ako bola chápaná na konci 19. a temer v celom 20. storočí (pozri napr. Gould, 1998). Kvantitatívny prístup sa uplatnil vo výskumoch aj v psychodiagnostike – vznikla psychometrika (Urbánek a kol., 2011). Tá sa však, žiaľ, dlho vyvíjala nie ako teoretická disciplína naviazaná na vedeckú psychológiu, ale plnila predovšetkým inštrumentálnu úlohu opierajúc sa iba o matematickú štatistiku (tamtiež). V tejto podobe sa na dlhý čas stala temer synonymom psychodiagnostiky. Hoci je pojem psychodiagnostiky širší – zahŕňa aj kvalitatívny klinický prístup (Halama, 2011), psychologické meranie je v psychodiagnostike cenené a bezpochyby užitočné. Ide však o jeho podobu. O nej sa síce v poslednom období v literatúre dosť diskutuje, no hlavný prúd uvažovania nie je zatiaľ novými myšlienkami zasiahnutý (Urbánek a kol., 2011).

Kvantitatívna psychológia a psychometrika vyvíjajú od šesťdesiatych rokov minulého storočia čoraz prepracovanejšie a sofistikovanejšie matematické postupy (nové podoby

---

<sup>1</sup> Podkladom tohto textu bol referát prednesený na konferencii Psychologická diagnostika detí a dospelých v Brne 30.6.2011.

<sup>2</sup> Výskumný ústav detskej psychológie a patopsychológie, Cyprichova 42, 831 05 Bratislava

<sup>3</sup> Ústav aplikovanej psychológie, Fakulta sociálnych a ekonomických vied UK, Mlynské luhy 4, 821 05 Bratislava

faktorovej analýzy či teória odpovede na položku – pozri napr. Halama, 2011; Urbánek a kol., 2011; Žitný, 2014). Avšak nástup či skôr znovuobjavenie kvalitatívnej metodológie a jej rozvoj v poslednom desaťročí minulého storočia kvantitatívne prístupy v psychológii spochybnil (pozri napr. Bačová, 2009). Aj to je dôvod zamýšľať sa nad úlohou matematiky v psychologickom poznaní a v psychodiagnostike zvlášť.

Matematika sa v psychodiagnostike uplatňuje

- pri tvorbe teórie, na ktorej je psychologický test založený;
- pri výbere testových položiek;
- pri overovaní psychometrických charakteristík testu;
- pri tvorbe testových noriem.

Každej z týchto oblastí venujem teraz stručnú pozornosť.

## Psychologická teória

Prvé testy v psychológii vznikali na základe implicitných teórií ich autorov; opierali sa skôr o ich klinickú skúsenosť a intuíciu (pozri napr. Nedvěd, 2011), než o explicitnú vedeckú teóriu. Chcem však zdôrazniť, že všetky vedecké teórie vznikajú explifikáciou implicitných teórií odborníkov. Samozrejme, že pri tom okrem osobného presvedčenia zohrávajú dôležitú úlohu informácie získané štúdiom literatúry a poznatky získané výskumom, vrátane jeho kvantitatívneho spracovania.

Nové testy sa zväčša opierajú o explicitné teórie. Teoretické ukotvenie testu sa pokladá za jeden zo znakov jeho kvality (Urbánek, Širůček, 2010). Vplyvná skupina autorov je presvedčená, že teóriu treba tvoriť na základe matematických výpočtov, predovšetkým pomocou faktorovej analýzy. Tak sa napríklad Woodcockov-Johnsonovej kognitívna batéria (Ruef et al., 2003) opiera o faktorovoanalytickú teóriu inteligencie známu ako CHC (Cattell-Horn-Carroll – pozri napr. Furman, 2005). Autori poslednej revízie známeho inteligenčného testu R. Amthauera (2005) sa zasa pokúsili prispôbiť pôvodnú štruktúru testu teórii, ktorú vytvorili ako kompiláciu faktorovoanalytických teórií vrátane CHC.

Som presvedčený, že aplikácia faktorovej analýzy môže pôvodnú teóriu spresniť, žiadna teória sa však nedá vypočítať. O údajoch, ktoré do výpočtu vložíme, nerozhoduje totiž matematika, ale úvaha (implicitná teória) autora. Napríklad: CHC obsahuje faktor sluchovej diskriminácie – jednoducho preto, že J. B. Carroll na základe svojej implicitnej teórie zaradil do výpočtov aj úlohy postihujúce túto schopnosť. Podobne ako väčšina teórií inteligencie však neobsahuje faktor tvorivosti. O tom, či tvorivosť možno považovať za súčasť inteligencie, však nerozhoduje výsledok faktorovej analýzy, ale autor teórie. J. P. Guilford (1967) na základe svojich teoretických úvah zaradil do štruktúry inteligencie aj faktory tvorivosti a iným typom faktorovej analýzy ich potvrdil. J. R. Sternberg (2001) hovorí o tvorivej inteligencii len na základe teoretickej analýzy problematiky, faktorovú analýzu k tomu nepotrebuje. Je preto jeho teória menejcenná?

I. Ruisel (2004), ktorý študuje inteligenciu mimo rámca psychodiagnostiky, konštatuje, že faktorovoanalytické teórie inteligencie sú dnes na ústupe.

Stanovenie faktorovej štruktúry testov S-B (Thorndike et al., 1995) či WISC (Wechsler, 2002) možno uviesť ako príklady adekvátneho použitia faktorovej analýzy na spresnenie teórie konkrétneho testu. Aplikácia faktorovej analýzy za účelom obsahového vymedzenia meraného konštruktu však nemusí byť adekvátne. Konkrétne WJ-IE (Ruef et al., 2003): je pamäť a sluchová diskriminácia naozaj súčasťou „schopnosti myslieť“, kam ich zaraďuje príručka?

Aké teda vidím problémy pri aplikácii faktorovoanalytických koncepcií v psychometrii? Zamyslieť sa treba predovšetkým nad dvomi otázkami:

- Je obsah každého faktoru či „trsu subtestov“ dostatočne logický (a psychologický)?
- Je vždy adekvátne interpretovať testové výsledky podľa faktorov, ako to odporúčajú príručky (napr. Amthauer et al. 2005)? Faktory boli odvodené z korelácií medzi subtestami vo veľkom súbore, v individuálnom prípade však takéto vzťahy nemusia platiť. Faktorové (každé sumárne) skóre možno zmysluplne interpretovať iba v prípade, že medzi skóre subtestov vo vnútri faktora nie sú veľké rozdiely.

## Výber testových položiek

Výber konkrétnych položiek sa robí z väčšieho množstva navrhnutých položiek na základe položkovej analýzy tak, aby položky neboli príliš ľahké ani príliš ťažké a aby dobre diferencovali medzi skúšanými (Halama, 2011). Údaje o obtiažnosti a diskriminačnej sile položiek sú pre zostavenie kvalitného testu nevyhnutné. Nemožno sa riadiť „pocitmi“ autora, tieto parametre treba skutočne vypočítať. Aby mal výpočet zmysel, musí sa samozrejme realizovať s údajmi získanými v adekvátnej referenčnej skupine (reprezentatívnom súbore, pozri nižšie).

Teória odpovede na položku (IRT) pracuje s komplexnejšími charakteristikami (pozri napr. Žitný, 2014). Je však otázne, či testy založené na IRT poskytujú skutočne presnejšie výsledky, a to vzhľadom na dve skutočnosti:

- vstupné údaje (odpovede testovaného) sú vždy „soft data“;
- aj v prípade IRT sa na individuum aplikujú pravdepodobnostné poznatky získané výpočtom z údajov veľkého súboru.

Napríklad obtiažnosť: V klasickej teórii testov vypočítame jediný koeficient, ktorý nám hovorí, aké percento všetkých testovaných konkrétnu položku správne vyrieši. Aj keď vieme, že obtiažnosť položky je dajme tomu 0,8 (80 % správnych riešení), nič nám to nepovie o tom, ako bude túto položku riešiť konkrétna osoba. Dá sa predpokladať, že pre jedinca s vyššou úrovňou meranej schopnosti bude úloha ľahká, no pre testovaného

s podpriemernými schopnosťami ľahká byť nemusí. Teória odpovede na položku nám ponúka sofistikovaný výpočet krivky, z ktorej odčítame, nakoľko je položka obtiažná pre jedinca s priemernou, podpriemernou i nadpriemernou úrovňou meranej schopnosti. No aj keď viem, aká je odhadovaná úroveň testovaného, a na základe toho konštatujem, že má pri odpovedaní na konkrétnu položku 80-percentnú pravdepodobnosť úspechu, stále to neznamená, že práve on nepatrí medzi tých 20 % ľudí, ktorí položku nevyriešia.

V tomto vidím zásadný problém tzv. adaptívneho testovania, t. j. prispôsobenia rozsahu testu konkrétnej osobe, a to tak pri klasickom testovaní, ako aj pri adaptívnom testovaní pomocou počítača, ktoré využíva teóriu odpovede na položku. Hoci podľa predbežných výskumov majú adaptívne počítačové testy založené na IRT aj pri podstatnom znížení počtu zadávaných položiek stále dobré psychometrické charakteristiky (Žitný, 2014), riziko, že skrátenou verziou testu neodhadneme úroveň schopností konkrétneho testovaného správne, nie je zanedbateľné. Je napríklad známe, že rozumovo nadané deti niekedy nevyriešia relatívne ľahké úlohy a zabodujú až pri úlohách ťažkých (Portešová, 2011). Adaptívne testovanie takéhoto dieťaťa by sa mohlo skončiť skôr, než na tie ťažké úlohy dôjde. Počítač mu ich, na základe slabého výkonu v ľahkých úlohách, vôbec neponúkne a potenciál dieťaťa ostane neodhalený.

## Overovanie psychometrických charakteristík

Všetky druhy reliability (Halama, 2011; Urbánek a kol., 2011) je vhodné vyjadriť pomocou vypočítaných koeficientov. Úlohou tvorca či upravovateľa testu je stanoviť adekvátne zdroje informácií o reliabilite (Standardy..., 2001) a im primerané štatistické ukazovatele. V testoch zameraných na osobnostné vlastnosti je namieste zaoberať sa stabilitou výsledkov v čase, v testoch sledujúcich aktuálne psychické stavy by to bolo kontraproduktívne.

Vo väčšine prípadov je vhodné posudzovať reliabilitu okrem iného podľa konzistencie odpovedí na položky zamerané na rovnakú vlastnosť. Zdá sa, že v súčasnosti sa pri tom nadožíva Cronbachov koeficient alfa (Urbánek, Širůček, 2010), ktorého výpočet ponúkajú bežne dostupné štatistické počítačové programy (napr. SPSS). Údaje o vnútornej konzistencii testu prispievajú aj k výberu položiek – položka, ktorá konzistenciu narúša, by mala byť z testu vyradená. Pozor ale na privysoké koeficienty (Halama, 2011)! Môžu znamenať, že položky nezachytávajú dostatočne reprezentatívnu množinu situácií, v ktorých sa má sledovaná vlastnosť prejavovať. Ak napríklad položím desaťkrát tú istú, len rôzne naštylizovanú otázku, dostanem desať zhodných odpovedí a koeficient konzistencie sa bude rovnať jednej. Napriek absolútnej reliabilite bude však validita takto zostaveného „testu“ pochybná. Podľa môjho názoru ekologickú validitu testu spochybňuje vnútorná konzistencia vyjadrená koeficientom vyšším než 0,95. Pri nižších hodnotách samozrejme platí, že vyššia reliabilita je lepším predpokladom toho, aby test mohol byť pokladaný za validný (Halama, 2011; Urbánek a kol., 2011).

Pri hľadaní dôkazov o validite (Standardy..., 2001) sa vychádza z teórie. Pre získanie väčšiny týchto dôkazov je treba realizovať empirický kvantitatívny výskum, v takom prípade ich možno predložiť v podobe čísel (koeficientov validity). D. Boorsboom so spolupracovníkmi (podľa Urbánka a kol., 2011) súčasné prístupy ku kvantifikácii validity ostro kritizujú. Upozorňujú napríklad, že ak výsledok nejakého testu vysoko koreluje s výsledkom iného testu, nemusí to ešte znamenať, že testy merajú totožnú vlastnosť. V každom prípade odporúčam sústrediť sa na to, čo klasická teória testov označuje ako obsahovú validitu a čo sa overuje expertnými posudkami obsahu položiek; ich zhodu samozrejme tiež možno vyjadriť matematicky (Halama, 2011).

Podobne možno s pomocou matematiky vyjadriť aj objektivitu testu chápanú ako zhodu výsledkov, ktoré s jeho pomocou získajú viacerí posudzovatelia (tamtiež; Urbánek et al., 2011). To je dôležité najmä pri takých psychodiagnostických nástrojoch, ktoré nemajú jednoznačne formalizované kritériá vyhodnocovania (napríklad pri projektívnych testoch či pri testoch tvorivosti).

## Tvorba testových noriem

Pre možnosť využívať test nielen vo výskume, ale ako nástroj individuálnej psychodiagnostiky, je nevyhnutné porovnať odpovede testovaného s odpoveďami referenčnej skupiny (teda s odpoveďami typickými pre populáciu, do ktorej skúšaný patrí). To nám umožňuje aplikácia vhodných metód matematickej štatistiky.

Kruciálnou otázkou tvorby noriem je výber adekvátnej referenčnej skupiny (reprezentatívneho štandardizačného súboru). Náhodný výber z populácie je nerealizovateľný, ďalšie možné typy výberov uvádzajú všetky príručky kvantitatívnej metodológie (Ritomský, 2002; Hendl, 2004 a d.). Najčastejšie sa používa tzv. kvótny výber; reprezentatívnosť súboru sa dosahuje stratifikáciou podľa adekvátnych demografických kritérií. Je na odborníkoch realizujúcich štandardizáciu, aby zvolili kritériá, ktoré majú zásadný význam vo vzťahu k sledovanej vlastnosti. Súbor by mal byť dostatočne veľký na to, aby sa zvolené kritériá dali uplatniť (Dočkal, 2010). Pri „doladovaní“ (vyvažovaní) súboru podľa stratifikačných kritérií môže pomôcť matematika, nevyrieši však situáciu, ak je v niektorom podsúbore nedostatok osôb.

Príkladom takto nesprávne použitej matematiky môže byť tvorba slovenského štandardizačného súboru testu WJ-IE. Podľa príručky (Ruef et al., 2003) mal slovenský súbor 962 členov, z toho sedem (0,7 %) malo maďarskú národnosť. Keďže upravovatelia chceli zachovať stratifikáciu podľa etnickej príslušnosti a na Slovensku žije podľa demografických údajov cca 11 % Maďarov, prisúdili výsledkom maďarských probandov váhu 16. Je otázne, či je tvorba spoločných noriem pre rôzne etnické skupiny adekvátna. Ak sa ale mala dodržať národnostná skladba štandardizačného súboru, bolo nevyhnutné zaradiť do štandardizačného súboru väčší počet maďarských probandov. Veľkosť súboru síce nie je zárukou jeho reprezentatívnosti (Dočkal, 2010), podstatne však ovplyvňuje presnosť merania: čím početnejší je súbor, tým užší je interval

spoľahlivosti sledovaného parametra (Hendl, 2004). Priemerný výsledok siedmich osôb môže byť zaťažený veľkou chybou; prisúdením uvedenej váhy sa táto (neznáma) chyba šestnásťkrát vniesla do výsledných noriem testu.

Zásadnou myšlienkou, ktorú chcem na tomto mieste vysloviť, je, že testové normy sa nepočítajú, ale tvoria; výpočty sú iba nevyhnutným východiskom pre ich vypracovanie. Psychologickú realitu totiž nemožno vždy vyjadriť pomocou matematického modelu. Podrobnejšie som sa tejto problematike venoval v iných štúdiách (Dočkal, 2009, 2010); tu uvediem niekoľko príkladov neadekvátne vytvorených noriem:

Nalistujme si prevodové tabuľky v českej príručke SON-R 2½-7 (Tellegen et al., 2008). Až do veku temer 4 rokov je možné za nulový výkon v niektorom zo šiestich subtestov získať štandardné skóre (ŠS) vyššie ako 1. Napríklad v subteste Mozaiky získa 2;5 ročné dieťa, ktoré nevyriešilo ani jednu úlohu, štandardné skóre 7 bodov (čo zodpovedá IQ=85, teda nižšiemu priemeru). Transformácia hrubého skóre na štandardné prostredníctvom z-skóre sa realizuje na celej číselnej osi. Pri nízkej priemernej hodnote a sprava zošikmenej distribúcii, či veľkej smerodajnej odchýlke je možné vypočítať nenulové ŠS aj pre záporné hrubé skóre (bližšie pozri Dočkal, 2009), hoci záporné HS reálne neexistuje. Túto anomáliu autori noriem korigujú tým, že hodnoty ŠS pre záporné HS v tabuľkách neuvádzajú. Štandardné skóre pre nulový výkon však akceptujú tak, ako bolo vypočítané podľa príslušného vzorca.

Hrubé skóre rovné nule znamená, že výkonový potenciál dieťaťa je maximálne taký, ako ukazuje príslušné ŠS, ale môže byť aj ľubovoľne nižší. Test nie je v oblasti nízkeho výkonu v uvedenom veku dostatočne citlivý, preto neumožňuje vyjadriť sa o reálnej výkonovej úrovni testovaného. Napriek tomu sa ŠS uvedené v tabuľke započítava do sumárnych výsledkov, čím dochádza k nekontrolovateľnému nadhodnoteniu výkonov dieťaťa. Rovnaký postup je použitý v holandskej, ale aj nemeckej a slovenskej príručke SON-R 2½-7, lebo P. J. Tellegen, ktorý normy vypracoval, sa plne spolieha na presnosť výpočtov. Ten istý princíp (započítanie nenulového ŠS za nulový výkon v subteste) využívajú aj britské a české vydania testu WISC-III či nemecké a české vydania I-S-T 2000 R (Dočkal, 2010), čo pokladám za nenáležité.

Som presvedčený, že ak nechceme odhad schopností testovaného jednotlivca nekontrolovateľne nadhodnotiť, nemali by sme subtesty, v ktorých nezískal ani bod, brať pri výpočte sumárnych skóre do úvahy. Preto sme v slovenskej verzii WISC-III (Wechsler, 2006) nulovému výkonu neprisudzovali vážené skóre a výpočet celkového IQ umožňujeme robiť iba na základe subtestov, v ktorých dieťa vyriešilo aspoň jednu položku. Interpretácia nulového výkonu v tom-ktorom subteste, ako aj interpretácia IQ stanoveného na základe menšieho počtu subtestov je potom vecou klinickej úvahy psychológa.

Posledným problémom, na ktorý chcem upozorniť, je tvorba noriem vývinového testu, ktorého štandardizácia sa bežne robí prierezovým výskumom s rôznymi vekovými skupinami. Keďže výkon dieťaťa s vekom narastá, malo by platiť, že na dosiahnutie

rovnakého štandardného skóre potrebuje starší proband získať viac bodov hrubého skóre. To sa darí dodržať pri výkonoch v pásme priemeru – prípadné nerovnomernosti v priemeroch získaných z výskumných údajov sa vyrovnávajú graficky alebo pomocou regresných rovníc (Říčan, 1977). V extrémnych polohách sa však (vd'aka nerovnakým rozptylom v rôznych vekových skupinách štandardizačného súboru) môžeme stretnúť s anomáliami, ktoré treba pri tvorbe prevodovej tabuľky odstrániť manuálne, ako to urobili napríklad autori britskej príručky k testu WISC-III (Wechsler, 1992) a ako sme to realizovali aj v slovenskej príručke (Wechsler, 2006). Autor českých noriem (Boschek, 2002) sa však pýši tým, že k žiadnym úpravám nepristúpil a opiera sa o presne vypočítané hodnoty.

Ako sa takáto matematická rigoróznosť prejavila? Nalistujme si v českej príručke (Wechsler, 2002) napríklad prevodové tabuľky pre 10;8-10;11 ročných a 11;0-11;3 ročných. Ak proband, ktorého sme testovali vo veku 10;8 rokov získal v subteste Kocky HS=15, hodnotíme ho štandardným skóre 3 (čo zodpovedá IQ=65, teda úrovni mentálnej retardácie). Predpokladajme, že toho istého probanda budeme retestovať po pol roku a jeho absolútny výkon sa nezmení. Napriek tomu mu podľa prevodovej tabuľky prisúdime ŠS = 4 (čo zodpovedá IQ=70 a posúva ho do tzv. „hraničného pásma“). Podobne je tomu aj pri iných subtestoch; v prípade Kódovania je dokonca možné získať vo vyššom veku vyššie ŠS aj pri niekoľkobodovom znížení reálneho výkonu oproti nižšiemu veku.

Bolo by samozrejme nezmyslom interpretovať stagnujúci (či dokonca zhoršený) výkon dieťaťa ako jeho zlepšenie, ako k tomu nabáda tabuľkové štandardné skóre. Príčinu uvedenej anomálie som podrobne vysvetlil inde (Dočkal, 2009). Riešiť ju treba úpravou rozptylov v susedných vekových skupinách pred výpočtom štandardných skóre a manuálnou úpravou prevodovej tabuľky, ak aj napriek tomu anomália pretrváva.

## Záver

Matematika je nepostrádateľnou pomocníčkou pri konštrukcii, štandardizácii aj interpretácii výsledkov psychodiagnostických testov. Bez uplatnenia matematicko-štatistických postupov nemožno vytvoriť kvalitný psychometrický nástroj. Avšak matematika je pritom naozaj iba pomocníčkou:

- Teóriu, na ktorej je test založený, rozhodne nie je možné „vypočítať“. Exploračná faktorová analýza je vhodným nástrojom istého typu výskumov a jej výsledky môžu prinášať zaujímavé hypotézy. Nemala by sa však používať na generovanie psychologickéj teórie. Vypočítané faktory sú totiž matematickým artefaktom a nemusia zodpovedať žiadnej psychickej realite (Gould, 1998). Niektoré aspekty teórie možno pomocou faktorovoanalytických postupov overovať, ale až ex post – na to slúži tzv. konfirmačná faktorová analýza (Halama, 2011).

- Tvorba testových noriem sa vždy opiera o presné výpočty. Konečné slovo pri nej však musí mať psychologický úsudok, ktorý môže výsledky uskutočnených výpočtov korigovať.
- Najpriamejšie je použitie matematiky pri zisťovaní psychometrických charakteristík jednotlivých položiek i celého testu. Ktoré charakteristiky a akým spôsobom zisťovať, o tom však opäť rozhoduje kvalifikovaný psychológ a interpretácia zistených koeficientov vrátane rozhodovania sa na ich základe o ďalšom postupe pri vývoji testu je tiež vecou jeho úvahy.
- Správna interpretácia testových výsledkov je nemysliteľná bez poznania ich štatistického pozadia (čo žiaľ nebyva našou silnou stránkou), musí však vždy ísť o psychologickú, nie štatistickú interpretáciu.

Odmietanie matematických aplikácií v psychológii je zaiste nerozumným extrémom, napokon aj kvalitatívny prístup s istým využitím matematiky ráta (pozri napr. Silverman, 2005). Rovnako nerozumné by však bolo spoliehať sa na to, že matematika vyrieši psychologické problémy namiesto psychológie, a podľahnúť zdanlivej elegancii matematických riešení. Konečným arbitrom podoby testu, testových noriem aj interpretácie získaných výsledkov musí byť odborná psychologická úvaha. O využití konkrétnych matematických postupov v psychodiagnostike vždy rozhoduje autor (upravovateľ, používateľ) testu.

## Literatúra

- Amthauer, R., Brocke, B., Liepman, D., & Beauducel, A. (2005). *Test struktury inteligence I-S-T 2000 R*. První české vydání. Upravila: A. Plháková. Praha: Testcentrum.
- Bačová, V. (2009). *Súčasný smery v psychológii*. Bratislava: Veda.
- Boschek, P. (2000). *Tabulky norem a psychometrické vlastnosti české verze testu WISC-III<sup>UK</sup>*. Praha: IPP ČR.
- Dočkal, V. (2009). Matematika a zdravý rozum pri tvorbe testových noriem [CD]. In L. Golecká, J. Gurňáková, & I. Ruisel (eds), *Sociálne procesy a osobnosť 2008* (pp. 45-56). Bratislava: ÚEPsSAV.
- Dočkal, V. (2010). Metodologické problémy štandardizácie testov schopností určených pre deti a mládež [CD]. In M. Gregussová, & E. Farkašová (eds), *Psychológia dieťaťa: Výskum a prax*. Bratislava: VÚDPaP.
- Furman, A. (2005). Teória inteligencie Gf-Gc ako východisko testovej batérie Woodcock-Johnson international editions. *Psychológia a patopsychológia dieťaťa*, 40, 347-361.
- Gould, S. J. (1998). *Jak neměřit člověka*. Praha: NLN.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Halama, P. (2011). *Princípy psychologické diagnostiky*. 2. dopln. vyd. Trnava: FFTU.
- Hendl, J. (2004). *Přehled statistických metod zpracování dat*. Praha: Portál.



- Nedvěd, J. (2011). Střípky z historie Wechslerových škál [CD]. In: D. Heller & P. Michálek (eds), *28. Psychologické dny. Cesty psychologie a psychologie cest* (pp. 179-187). Praha: PEF ČZU; ČMPS.
- Portešová, Š. (2011). *Rozumově nadané děti s dyslexií*. Praha: Portál.
- Ritomský, A. (2002). *Metódy psychologického výskumu: kvantitatívna analýza dát*. Bratislava: Medzinárodné stredisko pre štúdium rodiny.
- Ruef, M., Furman, A., & Muñoz-Sandoval, A. (2003). *Woodcock-Johnson. Medzinárodná edícia*. Nashville: The Woodcock-Muñoz Foundation.
- Ruisel, I. (2004). *Inteligencia a myslenie*. Bratislava: Ikar.
- Říčan, P. (1977). *Úvod do psychometrie*. Bratislava: Psychodiagnostické a didaktické testy.
- Silverman, D. 2005. *Ako robiť kvalitatívny výskum*. Bratislava: Ikar.
- Standarty pro pedagogické a psychologické testování*. (2001). Praha: Testcentrum.
- Sternberg, R. J. (2001). *Úspěšná inteligence*. Praha: Grada.
- Tellegen, P. J., Laros, J. A., & Heider, D. (2008). *SON-R 2½-7. Neverbální inteligenční test*. Praha: Hogrefe – Testcentrum.
- Urbánek, T., Denglerová, D., & Širůček, J. (2011). *Psychometrika*. Praha: Portál.
- Urbánek, T., & Širůček, J. (2010). Recenze psychodiagnostické metody, její průběh, výsledky a důsledky. Workshop. *28. Psychologické dny. Cesty psychologie a psychologie cest*. ČMPS, Olomouc 9.9.2010.
- Wechsler, D. (1992). *WISC-III<sup>UK</sup>. Manual*. London: The Psychological Corporation.
- Wechsler, D. (2002). *WISC-III. Wechslerova inteligenční škála pro děti*. Autoři české verze D. Krejčířová, P. Boschek, J. Dan. Praha: Testcentrum.
- Wechsler, D. (2006). *WISC-III<sup>SK</sup>. Wechslerova inteligenčná škála pre deti*. Autori slovenskej verzie: V. Dočkal, E. Kretová, B. Kundrátová, B. Sedlačová, M. Tesař. Praha: Hogrefe – Testcentrum.
- Žitný, P. (2014). *Aplikácia teórie odpovede na položku v psychometrickej analýze a počítačovom adaptívnom testovaní*. Kraków: Towarzystwo Słowaków w Polsce; FF TU.

## **Vladimír Dočkal (2016): The splendor and miseries of mathematical solutions in psychodiagnosics**

**Abstract:** When a hundred years ago psychology started to quantify its findings, it won its respect and an opportunity to integrate itself into the world of science – that is, as science was understood in those days. The quantitative approach is characteristic mostly for the psychometrics, which has become almost a complete synonym of psychodiagnosics in the light of positivism; if that is not so, then it has become at least a prevalent direction in it. Even today, the positivist paradigm in psychology has a strong position; however, rather new and diverse (qualitative) views receive an increasing support. It is therefore appropriate to reflect on the role of mathematics in psychological knowledge, particularly in psychodiagnosics. The contribution critically evaluates the possibilities of mathematics in framing the theory on which the psychological test is based, in selecting the items, and in verification of psychometric characteristics of the test and in development of test standards. It demonstrates the usefulness of mathematical approaches, but also draws attention to their limits. Mathematics is unsubstitutable in psychodiagnosics; however, it has only an assistant role that cannot replace qualified psychological reasoning.

**Key words:** psychometrics, mathematical statistics, test theory, items, norms