

Moderné vzdelávanie pre vedomostnú spoločnosť/Projekt je spolufinancovaný zo zdrojov EÚ

METODIKA TVORBY TESTOVÝCH ÚLOH A TESTOV

14. Na divadelnom predstavení sa zúčastnili žiaci z troch základných škôl. Počet žiakov je uvedený v tabuľke. Jedna vstupenka stála 1 euro. Ak mála skupina viac ako 20 žiakov, vstupné

15. Martin ktorá že tie

6

Úloha č. 1

Vo ve na stí a vo národí

Ako m Uvedt

1. 2.

Úloha

Prečo

A. B. C. D.

Úloha

Martin Ako n

A. B. C. D.

Úloha

Čo je t

A. B. C. D.

Ukážka 2 *Listy od Félixa*

2. Prehľad zistení štúdie OECD TALIS 2008 zverejnených v medzinárodných publikáciách

úroveň 3 CIL

MATEMATIKA – 1203

PREHLAD VZŤAHOV

Mocniny: $a^x \cdot a^y = a^{x+y}$, $\frac{a^x}{a^y} = a^{x-y}$, $(a^x)^y = a^{x \cdot y}$, $(a \cdot b)^x = a^x \cdot b^x$, $\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$, $a^{-x} = \frac{1}{a^x}$, $a^{\frac{1}{x}} = \sqrt[x]{a}$

Goniometrické funkcie:

$\sin^2 x + \cos^2 x = 1$	$\operatorname{tg} x = \frac{\sin x}{\cos x}$	x	0°	30°	45°	60°	90°
$\sin 2x = 2 \sin x \cos x$	$\cos 2x = \cos^2 x - \sin^2 x$	$\sin x$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\sin\left(\frac{\pi}{2} - x\right) = \cos x$	$\cos\left(\frac{\pi}{2} - x\right) = \sin x$	$\cos x$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0

Trigonometria:

Sinusová veta: $\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma} = 2r$

Kosínusová veta: $a^2 = b^2 + c^2 - 2bc \cos \alpha$

Logaritmus: $\log_a(x \cdot y) = \log_a x + \log_a y$, $\log_a \frac{x}{y} = \log_a x - \log_a y$, $\log_a x^k = k \log_a x$

Aritmetická postupnosť: $a_n = a_1 + (n-1)d$, $s_n = \frac{n}{2}(2a_1 + (n-1)d)$

Geometrická postupnosť: $a_n = a_1 \cdot q^{n-1}$, $s_n = \frac{a_1(q^n - 1)}{q - 1}$

Kombinatorika:

$P(n) = n!$, $C(k, n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$, $C'(k, n) = \binom{n+k-1}{k}$

Analytická geometria

Parametrické vyjadrenie priamky: $Ax + By + C = 0$

Všeobecná rovnica priamky: $Ax + By + C = 0$

Uhloví vektorov: $\cos \varphi = \frac{a \cdot b}{|a| |b|}$

Vzdialenosť bodu $M[m; n; p]$ od roviny: $|Mp| = \frac{|am_1 + bm_2 + cm_3 + c|}{\sqrt{a^2 + b^2 + c^2}}$

Stredový tvar rovnice kružnice

Objemy a povrchy telies:

	kváder	valec	ihlan	kužeľ	gufa
objem	abc	$\pi r^2 v$	$\frac{1}{3} S_p v$	$\frac{1}{3} \pi r^2 v$	$\frac{4}{3} \pi r^3$
povrch	$2(ab + bc + ca)$	$2\pi r^2 + 2\pi r v$	$S_p + S_p$	$\pi r^2 + \pi r s$	$4\pi r^2$

19. marec 2015

Krajina	% správnych odpovedí (S.E)
Slovenská republika	23 (1.5)
Česká republika	21 (1.2)
Čie	19 (1.2)
Slovinsko	16 (1.0)
Chorvátsko	14 (1.2)
Thajsko	7 (0.9)
Turecko	4 (0.7)



Európska únia
Európsky sociálny fond



MINISTERSTVO ŠKOLSTVA,
VEDY, VÝSKUMU A ŠPORTU
SLOVENSKEJ REPUBLIKY

NÚCEM
NÁRODNÝ ÚSTAV CERTIFIKOVANÝCH
MERANÍ VZDELÁVANIA



ZYŠOVANIE KVALITY VZDELÁVANIA
NA ZÁKLADNÝCH A STREDNÝCH
ŠKOLÁCH S VYUŽITÍM ELEKTRONICKEHO TESTOVANIA

Moderné vzdelávanie pre vedomostnú spoločnosť/Projekt je spolufinancovaný zo zdrojov EÚ

METODIKA TVORBY TESTOVÝCH ÚLOH A TESTOV



METODIKA TVORBY TESTOVÝCH ÚLOH A TESTOV

Publikácia vznikla ako výstup z projektovej aktivity 1.1 projektu Zvyšovanie kvality vzdelávania na základných a stredných školách s využitím elektronického testovania, spolufinancovaného z prostriedkov EÚ, ktorého riešiteľom bol Národný ústav certifikovaných meraní vzdelávania.

Zostavil: Mgr. Timotej Kubiš

Kolektív autorov: doc. RNDr. Peter Demkanin, PhD.
PhDr. Helena Hanuljaková
Mgr. Timotej Kubiš
PhDr. Marián Lapitka, CSc.

Autori: doc. RNDr. Peter Demkanin, PhD. (PD)
Mgr. Michal Hajdúk (MH)
PhDr. Helena Hanuljaková (HH)
Mgr. Timotej Kubiš (TK)
PhDr. Marián Lapitka, CSc. (ML)
doc. RNDr. Martin Malčík, PhD. (MM)

Recenzenti: doc. RNDr. Renáta Bernátová, PhD.

Štylistická a jazyková úprava: Mgr. Timotej Kubiš

Grafická úprava: Ing. Jana Madarászová, Jaroslava Gašparíková
Obálka: Ing. Jana Madarászová

Ilustrácia: akad. mal. Jindřich Prášil

OBSAH

Predslov.....	7
1. Úvod do problematiky testovania	9
1.1. K histórii testovania na Slovensku (TK).....	9
1.2. Kontrola výsledkov vzdelávania (ML a kol.).....	13
1.3. Problematika merania vo vzdelávaní (ML a kol.).....	15
1.5. Etika hodnotenia žiakov (ML a kol.).....	21
1.6. Vplyv testovania na školské vzdelávanie (ML a kol.).....	23
1.7. Motivačná funkcia didaktických testov (TK).....	24
2. Teória tvorby testov	27
2.1. Predmet didaktického testu (ML, TK).....	29
2.2. Bloomova taxonómia (TK).....	32
2.3. Druhy testov (TK).....	49
2.3.1. Podľa meranej charakteristiky výkonu.....	49
2.3.2. Podľa typu učenia.....	50
2.3.3. Podľa miery špecifickosti učenia zisťovaného testom.....	51
2.3.4. Podľa fázy učebného procesu.....	52
2.3.5. Podľa prípravy testu.....	53
2.3.6. Podľa spôsobu interpretácie výsledkov.....	54
2.3.7. Podľa typu úloh.....	59
2.3.8. Podľa rozsahu testovaného učiva.....	59
2.3.9. Podľa úlohy učiteľa v testovaní.....	60
2.4. Vlastnosti didaktického testu.....	61
2.4.1. Validita (HH a kol.).....	61
2.4.2. Reliabilita (HH a kol.).....	63
2.4.3. Obťažnosť (TK).....	65
2.4.4. Citlivosť (HH a kol.).....	66
2.4.5. Praktickosť testu (HH a kol.).....	67
2.4.6. Objektívnosť testu (HH a kol.).....	68
3. Plánovanie a tvorba testu	71
3.1. Vymedzenie účelu testu (TK).....	71
3.2. Výber cieľovej skupiny (TK).....	71
3.3. Formulácia výskumnej otázky alebo hypotézy (TK).....	73
3.4. Stanovenie premenných (TK).....	74
3.5. Návrh testovej špecifikácie (TK).....	76
3.6. Operacionalizácia premenných.....	80
3.7. Výber alebo vytvorenie úloh testu a prvotné posúdenie ich validity (TK).....	80
3.8. Zostavenie kľúča správnych odpovedí k testu.....	81

3.9.	Určenie obťažnosti testových úloh (TK)	81
3.10.	Stanovenie validity a reliability testu (TK)	82
3.11.	Overenie prvotného návrhu testu a úprava testu (TK)	82
3.12.	Analýza výsledkov testu a nesprávnych odpovedí (TK)	83
3.13.	Kontrola a úpravy testu (PD a kol.)	83
3.14.	Tvorba variantov testu (PD a kol.)	84
3.15.	Vecná a jazyková kontrola testu (PD a kol.)	84
3.16.	Grafická úprava testu (PD a kol.)	84
3.17.	Overovanie testu (PD a kol.)	85
3.18.	Vypracovanie manuálu pre zadávateľov testu a hodnotiteľov (PD a kol.)	85
4.	Tvorba testových úloh	87
4.1.	Princípy tvorby testových úloh (PD a kol.)	88
4.2.	Druhy a formy testových úloh	91
4.2.1.	Uzavreté úlohy (PD a kol.)	92
4.2.2.	Otvorené úlohy (PD a kol.)	94
4.2.3.	Ďalšie formy testových úloh v systéme eTest (TK)	96
4.3.	Východiskový text (ML, TK)	97
4.3.1.	Druhy textov podľa pôvodu	98
4.3.2.	Druhy textov podľa typu	98
4.3.3.	Implicitné a explicitné informácie v texte	99
4.3.4.	Rozsah a obsahová náročnosť textu	99
4.3.5.	Formatívne aspekty východiskového textu	100
4.3.6.	Intaktnosť prevzatého textu a autorské práva	101
4.3.7.	Spôsob citovania pri prevzatých východiskových textoch	101
4.4.	Tvorba, výber a počet distraktorov (PD a kol.)	102
5.	Štatistická analýza testových úloh (MH)	107
6.	Banka úloh pro adaptivní testování (MM)	129
6.1.	Obecný popis adaptivního testování	130
6.2.	Banka úloh	132
6.2.1.	Tvorba banky úloh	133
6.2.2.	Metody pro vytvoření banky úloh	134
6.2.3.	Podmínky nutné ke srovnávání úloh	134
6.2.4.	Metody srovnávání	134
6.3.	Adaptivní testování – Computer Adaptive Testing	136
	Literatúra	147

PREDSLOV

Predkladaná publikácia vznikala počas realizácie národného projektu *Zvyšovanie kvality vzdelávania na základných a stredných školách s využitím elektronického testovania* (E-test), ktorého realizátorom bol Národný ústav certifikovaných meraní vzdelávania (NÚCEM).

Je určená predovšetkým autorom testových úloh a testov, učiteľom škôl zapojených do projektu E-test, ktorí budú pracovať s elektronickým testovacím systémom e-Test a s jeho bankou úloh, ale aj širšej odbornej verejnosti, ktorá sa zaoberá testovaním.

V publikácii prinášame základné informácie o rôznych druhoch testových úloh a testov, zameriavame sa aj na tvorbu úloh a tvorbu testov. Pre zainteresovanejších čitateľov sme do publikácie zaradili aj kapitolu o štatistickej analýze testových úloh, kapitolu o Bloomovej taxonómii a aj kapitolu o adaptívnom testovaní. Pri tom všetkom sa snažíme zohľadňovať naše skúsenosti z praxe v oblasti testovania na Slovensku. Práve v snahe vyvarovať sa starým chybám sme v tejto publikácii zhromaždili aj rôzne zásady, pravidlá a tiež kroky, ktoré je potrebné dodržať pri tvorbe testu. Postup vytvárania premysleného výskumného nástroja pomôže porozumieť tomu, čo test vlastne je (resp. čím môže byť), a zároveň pomôže urobiť testovanie rigoróznejším a prínosnejším pre tých, ktorí sa testovania zúčastňujú, ako aj pre tých, ktorí testy zadávajú.

Na príprave textu tejto publikácie sa podieľal užší kolektív autorov. Preto v obsahu vyznačujeme, ktoré časti vznikli v procese kolektívnej tvorby, a ktoré boli vytvorené individuálne. Neskôr sme prizvali aj autorov – špecialistov, ktorí našu publikáciu obohatili o dve rozsiahle kapitoly: *Štatistická analýza testových úloh* (Michal Hajdúk) a *Banka úloh pro adaptivní testování* (Martin Malčík).

Veríme, že táto publikácia pomôže učiteľom a tvorcom testových úloh a testov získať vhľad do problematiky testovania a bude pre nich podnetom pre hľadanie ďalších dôležitých stimulov z oblasti testovej teórie pre školskú prax.

1. ÚVOD DO PROBLEMATIKY TESTOVANIA

1.1. K histórii testovania na Slovensku

Historickou príčinou zavádzania didaktických testov pri hodnotení výsledkov vzdelávania bol postupný nárast kritiky príliš subjektívneho hodnotenia vedomostí žiakov. Táto kritika sprevádzala školstvo už od čias zapojenia širších vrstiev obyvateľstva do vzdelávania (povinná školská dochádzka fungovala za čias Uhorska iba obmedzene po r. 1868, potom až za 1. ČSR po r. 1922) a trvá dodnes. V školách sa vtedy používal klasický výkladový a dogmatický systém, ktorý miestami pretrval dodnes.¹

„Pri ústnom skúšaní (s cieľom klasifikačným) je nemožné, aby sme dali každému žiakovi tú samú otázku... Jeden žiak dostane „výbornú“, že odpovedal na otázku z miestopisu Slovenska. Druhý dostane „nedostatočnú“, že napr. nevedel vodopis Moravy, hoci by na otázku z miestopisu Slovenska bol vedel odpovedať práve tak dobre, ako ten prvý, kdežto ten zas nebol by vedel vodopis Moravy ako tento druhý. Kde je tu spravodlivosť?“

Sú tu rozdiely v požiadavkách ohľadne povahy odpovedi samotnej. Jeden učiteľ žiada voľný spôsob vypravovania, ktorý má odzrkadliť osobnosť, spôsob pojatia, spracovania látky u žiaka. Naopak, najdú sa aj takí učitelia, ktorí sami s už v mysli formulujú odpoveď na danú otázku a klasifikujú podľa toho, ako sa odpoveď žiaka priblížila k tejto nimi formulovanej odpovedi.

Keď sa pozeráme na všetky tieto veci, nemožno sa diviť tomu, ako ďaleko je takáto bežná skúška od požadovanej objektivity...“²

Snahy zlepšovať školstvo sa objavili od jeho počiatkov. V období humanizmu to bol J. A. Komenský, ktorý vniesol do vyučovania prvok kontroly tým, že rozčlenil vyučovaciu hodinu na fázy: „*Druhé čtvrt hodiny at' je* (učiteľ – žiakov) *examinuje, vyrozuměli-li.*“³ V druhej polovici 19. storočia ovplyvňoval naše (vtedy rakúsko-uhorské) školstvo už aj nový pedagogický smer – herbartizmus, vychádzajúci z myšlienok nemeckého pedagóga J. F. Herbarta (1776 – 1841).

Herbartizmus bol v strednej Európe novým fenoménom a vlády sa ho pokúšali zaviesť do vzdelávania, vštepiť jeho zásady učiteľom. Aj v Uhorsku prechádzali učitelia dôkladnými prázdninovými vzdelávaniami, tzv. náukobehmi. Nemalo to však želaný efekt. Môžeme si ľahko predstaviť, koľko mal tento nový smer medzi učiteľmi priaznivcov. Učitelia predstierali, že sa riadia stupňami, ale v triedach sa naďalej držali strnulého výkladu a učenia poučiek naspamäť.⁴

Herbartovými stupňami sa inšpiroval B. S. Bloom, ktorý potom skoncipoval tzv. taxonómiu cieľov vzdelávania. Tá je už dlhšiu dobu v pedagogike hodnotená ako veľmi prínosná (pozri kapitolu 2.2. Bloomova taxonómia).

¹ Tomášková, J.: *Inovatívne spôsoby hodnotenia žiakov v predmete slovenský jazyk a literatúra*. Bratislava : Metodicko-pedagogické centrum, 2015, s. 13 – 18. ISBN 978-80-565-0379-9.

http://www.mpc-edu.sk/library/files/tomaskova_inovacia_hodnotenie_sjl.pdf (1. 4. 2015)

² Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 7 – 10.

³ Komenský, J. A.: *Didaktika, to jest umění umělého vyučování*. Praha : Nákladem knihkupectví Theodora Mourka, 1871. s. 146.

⁴ Lapitka, M. a kol.: *Didaktické základy novej koncepcie vyučovania literatúry a štruktúry učebníc na strednej škole*. Bratislava : Metodicko-pedagogické centrum, 2009, s. 8. ISBN 978-80-8052-340-4.

Herbartove myšlienky rôznym spôsobom ovplyvnili súčasnú pedagogiku. Napríklad vo vyspelejších školských systémoch už prekonaný *dualizmus* v odčlenení výchovy od vzdelávania sa u nás ešte traďuje. Jeho pôvodným zámerom bolo cielavedomejšie vychovávať a plniť afektívne ciele, no, žiaľ, viedlo to skôr k marginalizácii výchovy.⁵

U nás sa už po roku 1918 postupne začali dostávať k slovu súčasnejšie prúdy v pedagogike. Za priekopníka slovenskej pedagogiky v oblasti testovania považujeme Juraja Čečetku (1907–1983). Zaujímal sa najmä o adlerovskú psychológiu a jej využitie v pedagogike a prinášal do slovenského myslenia o vzdelávaní mnohé nové myšlienky.

V roku 1934 publikoval popularizačnú knižku o testovaní s názvom *Testovanie na školách a jeho štatistické praktikum*. Táto publikácia je po 80 rokoch stále aktuálna. Spôsobila to dlhá odmlka v testovaní u nás. Publikácia obsahuje veľmi názorné vysvetlenia základných javov súvisiacich s testovaním, štatistické parametre aj klasické štatistické spracovanie, ktoré sa u nás popri novších postupoch stále najviac používa.

V úvode knihy sa J. Čečetka zamýšľa nad tým, prečo je testovanie učiteľskou verejnosťou nedostatočne chápané: „*Problém testovania samotného nie je ešte ani v širších vrstvách nášho učiteľstva taký vyjasnený, ako by bolo si želať. Nemôžeme sa tomu ani diviť. Inštrukcie, ktoré naše učiteľstvo v tomto smere dostalo, nie sú dostatočné, aby mohlo aspoň prakticky, po technickej stránke pojať testovanie tak, aby toto v každom smere bolo objektívnou skúšobnou metódou a dovoľovalo učiteľstvu veľkorysejšiu porovnávaciu prácu vedeckú.*“

Čečetkove hodnotenie z roku 1934 potvrdzuje aj česká publikácia venovaná tvorbe didaktických testov z roku 2010: „*V súčasnej didaktike není hodnocení výsledků vyučovacího procesu stále věnován dostatečný prostor a celkově se dá říci, že výuka budoucích učitelů je zaměřena spíše na schopnost vytvářet vědomosti. Nedostatečná příprava budoucích učitelů v oblasti hodnocení vyučovacího procesu vede k tomu, že učitel má sklon hodnotit vyučovací proces a vědomosti studentů na základě pouze svých osobních zkušeností a intuice... Potom si každý učitel vytváří své vlastní normy, které většinou vedou k neobjektivnímu hodnocení žáků a v konečném důsledku i ke snižování úrovně jejich vědomostí. ...rozmanitost školních známek... spočívá v omylech a v povrchnosti při posuzování žáků. ...známka není přesným měřítkem výkonu žáka, naopak mnohdy zobrazuje přístup učitele k předmětu, hodnocení chování žáka a jiná druhotná kritéria.*“⁶

O tom, že aj štvorkári môžu dobre obstáť v testoch študijných predpokladov pre štúdium na vysokej škole, a naopak, jednotkári môžu aj neuspieť, nás presvedčil nedávny výskum českej spoločnosti www.scio.cz, s. r. o., realizovaný v spolupráci s Ministerstvom školstva vedy, výskumu a športu SR na vzorke vyše 14 000 maturantov v októbri 2014. „*Aj to je dôkaz, že známky nemusia vždy zodpovedať reálnym schopnostiam žiaka, pretože ich ovplyvňuje veľa subjektívnych faktorov...*“^{7,8} Medzi najčastejšie subjektívne faktory pri hodnotení žiakov učiteľmi sa považuje **haló efekt** (podľahnutie prvému dojmu) a tzv. **labeling** (nálepkovanie, prípadne až stigmatizácia žiakov), t. j. negatívne očakávania u opakovane neúspešných žiakov.⁹

Spoločnosť SCIO sa podobným meraniam v Českej republike venuje systematicky, a ak pripustíme, že situácia v českom školstve nie je úplne odlišná od našej, mali by sme venovať pozornosť aj zisteniam

⁵ Napríklad, za kurióznou už dnes môžeme považovať potrebu premenovať občiansku výchovu na občiansku náuku po roku 1989. Pritom už aj medzinárodný výskum CIVED 1999 zistil u nás nedostatky v oblasti výchovy k občianstvu, výskum ICCS 2009 ich znovu potvrdil.

⁶ Jeřábek, O. – Bílek, M.: *Teorie a praxe tvorby didaktických testů*. Olomouc : Univerzita Palackého v Olomouci, 2010, s. 6. ISBN 978-80-244-2494-1.

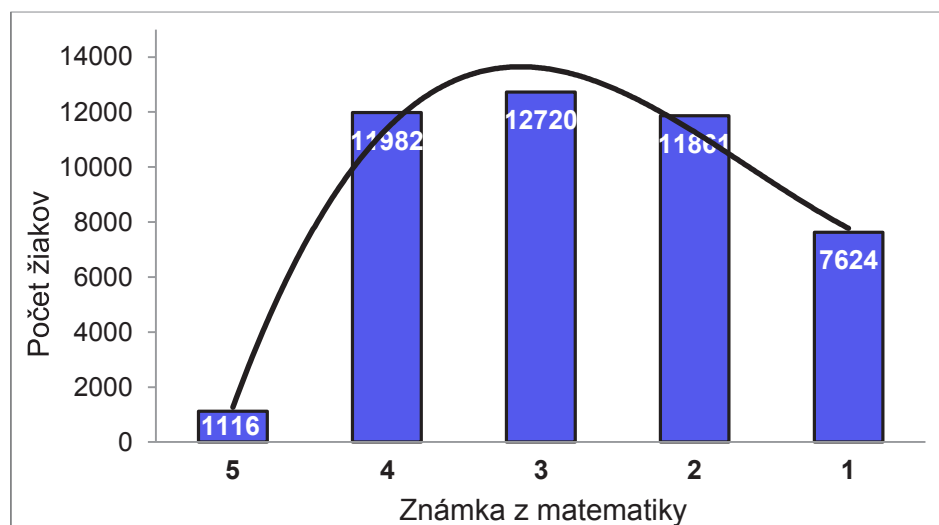
⁷ Vyvialová, L.: *Aj štvorkár sa môže uplatniť na vysokej škole, dôkazom sú výsledky testu študijných predpokladov*. <https://www.scio.cz/o-vzdelavani/analyzy-a-studie-spolecnosti-scio/aj-stvorkar.asp> (1. 4. 2015)

⁸ SCIO: *Analýza výsledkov testovania VŠP*. Súhrnná správa. 2014. <https://www.scio.sk/download/zprava-skoly-VSP-2014.pdf> (1. 4. 2015)

⁹ Labeling theory. http://en.wikipedia.org/wiki/Labeling_theory (1. 4. 2015)

spoločnosti SCIO z výskumov v prostredí českého školstva: „Přísnost známkování na ZŠ se velmi liší a ze známek nelze zjistit téměř nic o tom, jak dopadnou žáci u přijímacích zkoušek v předmětovém testu připraveném nezávislým subjektem.“¹⁰

Ako vidno na nasledujúcom grafe (obr. 1) rozloženie žiakov 9. ročníka ZŠ v SR v školskom roku 2010/2011 podľa známok z matematiky na polročnom vysvedčení nezodpovedá normálnemu rozloženiu. Z grafu vidno, že v praxi sa piaty klasifikačný stupeň na hodnotenie takmer nepoužíva. Ak by sme chceli v grafe vidieť normálne rozloženie, museli by sme toto nepravdepodobné rozdelenie pretransformovať na 3 stupne: štvorkári by predstavovali podpriemer, dvojkári a trojkári (obvykle široký) priemer a jednotkári nadpriemer. V tejto súvislosti spomína aj J. Čečetka snahu niektorých našich pedagógov o zavedenie jednoduchšieho, trojznámkového systému v časoch 1. ČSR.¹¹



Obrázok 1. Rozdelenie žiakov 9. ročníka ZŠ v SR podľa známky z matematiky na polročnom vysvedčení v školskom roku 2010/2011¹²

Čečetka rozdeľuje učiteľov podľa vzťahu k testovaniu do dvoch kategórií: „Jedni sa o vec skutočne zaujímajú, iní ju konajú len akosi v túžbe po „novote“. Podľa toho je potom aj postoj učiteľstva k testovaniu: jedni kriticky posudzujú jeho dobré a zlé stránky, usilujú sa zaujať k nemu objektívne stanovisko, iní a priori alebo bez rozmyslu odmietajú túto metódu alebo voslep ju chvália ako vec, ktorá za deň vykoná v škole svetoborný prevrat. Práve takéto nekritické prijímanie testovania narobí najviac zla, lebo nám prináša diletantov, blúznivcov, ktorí svojím povrchným ponímaním a vykonávaním vecí zatemňujú dobré stránky testovania a časom sami sa stanú, sklamaní, jeho najväčšími odporcami.“¹³

Hneď po Druhej svetovej vojne sa naplno prejavil istý negatívny vplyv sovietskej pedagogiky na naše školstvo a odpor k testovaniu bol zavedený celoplošne a zhora. Doktrína Komunistickej strany testovanie už dávnejšie odsudzovala,¹⁴ a keď na to po roku 1948 bola príležitosť, takmer okamžite bol prijatý

¹⁰ Hučín, J.: *Aktuální obraz vzdělávání při přechodu na střední školu*. Vybraná zjištění z přijímacích zkoušek. Tisková zpráva. Scio, 2013.

https://www.scio.cz/download/Co_rikaji_data_z_prijimacek_o_vzdelavani_13-10-15.pdf (1. 4. 2015)

¹¹ Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 9.

¹² Kubiš, T. – Polgáryová, E. – Stopková Kurajová, J.: *Testovanie 9-2011 – priebeh, výsledky a analýzy*. Bratislava : NÚCEM, 2011, s. 20.

http://www.nucem.sk/documents//26/testovanie_9_2011/Final_Sprava_T9-2011.pdf (1. 4. 2015)

¹³ Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 5.

¹⁴ „Praxe pedagogů, naprosto odtržená od pedagoga a od vyučovacího procesu, vedla v podstatě k pavědeckým experimentům a k nesčíslnému množství výzkumů mezi žáky a jejich rodiči, jako jsou **nesmyslné a škodlivé dotazníky, testy apod., které strana již dávno odsoudila**. Ty zdánlivě vědecké výzkumy, konané s velkým počtem žáků a jejich rodičů, směřovaly

zákon (v r. 1950) a náležité vykonávacie predpisy (školský a skúšobný poriadok pre 1. až 3. stupeň škôl a pod.), ktoré okrem toho, že zaviedli používanie žiackych knižiek, zakazovali používanie didaktických testov, bodovacích systémov a podobných aktivít.

Progresívne vedecké smery v psychológii (napr. adlerovská psychológia, ktorej sa venoval J. Čečetka, a mnohé iné) boli odsudzované ako buržoázne pavedy. Naši poprední pedagógovia (medzi nimi aj prof. Juraj Čečetka) v 50-tych rokoch neprešli previerkami. Pedagogický ústav, ktorý Čečetka viedol, bol zrušený a z Filozofickej fakulty musel po roku 1958 úplne odísť. Rehabilitovaný bol až in memoriam, v roku 1990.

Hoci predpisy zakazujúce testovanie boli zrušené predpisom č. 129 už v roku 1967, ani potom nastal v oblasti testovania u nás žiadny vážny zvrät. Zbierky úloh sa síce vydávali, ale testy zostavené zo zbierkových úloh nemali vlastnosti dnešných didaktických testov. Viac sa začali rozvíjať testy najmä v oblasti psychológie a v školstve neskôr najmä prijímačkové testy.

Aj po roku 1989 nejaký čas trvalo, kým sa testovanie znovu dostalo do pedagogického diskurzu. Medzinárodné testovania až do roku 1990 prebiehali bez našej účasti. Po politickej zmene v našom štáte sa do nich postupne, i keď veľmi váhavo, zapojilo aj Slovensko. V našich školách sa administrovali napríklad medzinárodné výskumy TIMSS, PISA, PIRLS, niektorých sme sa zúčastnili už aj viackrát. Hlavným zmyslom týchto výskumov je porovnať úspešnosť žiakov podľa zvolených kritérií na svetovej alebo európskej úrovni. Bolo to po prvý raz, čo sa náš vzdelávací systém ocitol pod takýmto drobnohľadom. Autori týchto projektov sa zamerali na kľúčové kompetencie a všeobecné intelektuálne zručnosti, ktoré nie sú národným špecifikom. A práve takáto orientácia testovania ukázala slabé stránky nášho vzdelávacieho systému. Na našich školách, a teraz to už nepredpokladáme, ale vieme, sa práve tieto vzdelávacie ciele orientované na kľúčové kompetencie a všeobecné intelektuálne zručnosti zanedbávajú, dokonca neraz až ignorujú, v prospech zaužívaného hromadenia krátkodobých faktických vedomostí („aby žiaci veľa vedeli...“). Pozitívny spätný vplyv týchto medzinárodných výsledkov sa do nášho školstva zatiaľ nepremietol.

Okolo roku 1998 začal projekt Milénium, ktorý vypracovala komisia vymenovaná Ministerstvom školstva, vedy, výskumu a športu Slovenskej republiky (Miron Zelina, Ján Turek, Vladislav Rosa) dostávať reálnu podobu. Projekt si dával za cieľ zmenu obsahu vzdelávania: „znížiť počet nepotrebných, nefunkčných informácií a nahradiť ich obsahom, poznatkami potrebnými pre život, kľúčovými kompetenciami.“¹⁵ V tom čase začal ŠPÚ pilotovať vedomostné testy (tzv. Monitor) a po rozsiahlej reforme školstva v roku 2008 začal aj novovzniknutý NÚCEM vytvárať testy novšieho typu, a od tradičného testovania vedomostí a školských zručností sa začal uberať smerom k testovaniu predmetových a kľúčových kompetencií.

Reforma školstva, na ktorej trvala Európska komisia, totiž priniesla hlavne zmeny v oblasti cieľov výchovy a vzdelávania, začala školstvo orientovať smerom od vedomostí ku kompetenciám. Zmeny sú zakotvené v kľúčových dokumentoch, akým je nový Školský zákon (zákon 245/2008 Z. z. o výchove a vzdelávaní o zmene a doplnení niektorých zákonov)¹⁶ a reformný Štátny vzdelávací program (ŠVP).¹⁷

prevažne proti žákům se špatným prospěchem nebo proti těm, kteří plně do rámce školního režimu nezapadali, a jejich účelem bylo dokázat na základě zdánlivého „vědeckého“ a „biosociálního“ hlediska moderní pedagogie **dědičnou a sociální podmíněnost špatného prospěchu** žáka nebo jednotlivých závad v jeho chování, dále **najít maximum záporných vlivů a patologických nedostatků samotného žáka, jeho rodiny, příbuzných, předků a společenského prostředí** a tím právě najít důvod pro odstraňování žáků z normálního školního kolektivu.“ (Úryvek zo správy Ústredného výboru Vsesvazovej komunistickéj strany zo 4. 7. 1936. In: Jeřábek, O. – Bílek, M.: Teorie a praxe tvorby didaktických testů. Olomouc: Univerzita Palackého v Olomouci, 2010, s. 9 – 10. ISBN 978-80-244-2494-1.)

¹⁵ Konceptia rozvoja výchovy a vzdelávania v Slovenskej republike na najbližších 15 – 20 rokov (projekt Milénium). S. 16. <http://www.cpk.sk/web/dokumenty/krvv.pdf> (1. 4. 2015)

¹⁶ Zákon 245/2008 Z. z. o výchove a vzdelávaní (školský zákon) a o zmene a doplnení niektorých zákonov. http://www.uips.sk/sub/uips.sk/images/PKvs/z245_2008.pdf (1. 4. 2015)

¹⁷ Štátny pedagogický ústav: Štátny vzdelávací program. <http://www.statpedu.sk/sk/Statny-vzdelavaci-program.alej> (1. 4. 2015)

Reforma priniesla ešte jeden výrazný posun v myslení o vzdelávaní – zaviedla chápanie obsahu (učiva) ako prostriedku pre dosahovanie cieľov. Týmto zmenám sa musí prispôsobovať aj koncepcia testovania.

Odborníkov na testovanie, najmä v oblasti teórie, máme však na Slovensku stále nemnoho, podobne ani študijný odbor alebo predmet venovaný edukometrii u nás stále na pedagogických fakultách chýba. Nedávny Medzinárodný výskum vyučovania a vzdelávania TALIS 2013 realizovaný aj na Slovensku zistil, že iba približne 10 % slovenských učiteľov realizovalo individuálny alebo skupinový výskum v oblasti svojho profesionálneho záujmu,¹⁸ t. j. v oblasti vzdelávania. Chýbajúca teoretická príprava a malá prax našich pedagógov s vedecko-výskumnou prácou v oblasti vzdelávania a s prácou s informáciami o žiakoch a vzdelávaní znižuje aj možnosti uplatnenia spätnej väzby z výskumov vzdelávania realizovaných externými inštitúciami.

1.2. Kontrola výsledkov vzdelávania

Pojem **pedagogická kontrola** výstupov z učebného procesu sa sformoval po aplikácii kybernetickej teórie do pedagogiky v 60. rokoch minulého storočia. E. Stračár vo svojej priekopníckej knihe *Systém a metódy riadenia učebného procesu* zaviedol ako rovnocenný výraz pre vyučovaciu činnosť učiteľa termín „riadenie učebného procesu“.¹⁹ A tam, kde sa uplatní model riadenia štruktúrovaných celkov, celkom logicky nastupuje etapa „spätnej informácie“, etapa *kontroly* činnosti systému, ktorá ukazuje – v prvom rade účastníkom – mieru účinnosti ich práce.

V autoregulatívnom systéme, akým má byť každá škola, sa dáta získané kontrolou stávajú kritériom na *korekcie* postupov. Bez týchto úprav by učebný proces stratil mobilitu, teda prestal by smerovať dopredu, ku kontinuálnemu plneniu výchovných cieľov, alebo – a to je v školskej praxi pravdepodobnejšie – zmenil by sa na mechanizmus, ktorý postupuje podľa učebného programu formálne, bez ohľadu na to, čo si žiaci z vyučovania odniesli.

Spätnú väzbu, ktorá sa realizuje skúšaním, treba chápať ako organickú súčasť vyučovacieho procesu. Okrem významu pre efektívne riadenie má skúšanie význam aj ako motivačný prostriedok.²⁰

Skúšanie preto nemožno zrušiť, nie je prejavom niečej administratívno-právnej horlivosti, ale motorom, ktorý poháňa vzdelávanie dopredu a spoluurčuje jeho priebeh. Zmyslom spätnej väzby je poskytnúť žiakovi pomoc, aby zlepšil to, čo robí. Z toho vyplýva, že spätá väzba má byť pre učiaci sa subjekt užitočná a zmysluplná.²¹

Učiteľ zisťovaním výsledkov učebnej činnosti žiakov, ktorú riadi, dostáva podklady pre ďalší postup. Koriguje pracovné plány z obsahového hľadiska i s ohľadom na časový rozvrh. Táto úprava má v tradičnej škole veľmi obmedzený priestor, ale predsa len existuje a má svoju dôležitosť. Len mimochodom poznamenávame, že autor vzdelávacieho programu „mastery learning“, B. Bloom, bol v tomto smere zásadový: kým si v jeho newyorskej škole aj najpomalší žiak neosvojil určené učivo (t. j. kým nedosiahol „mastery“ úroveň), trieda ďalej nepostupovala. V tradičnej škole, akou je aj naša škola, učiteľ odučí hodinu, resp. tematický celok a učebnú úspešnosť **väčšiny** žiakov iba odhaduje „od oka“, na základe odpovedí **niekoľkých** vyvolaných žiakov.

¹⁸ Základná informácia o výsledkoch štúdie OECD TALIS 2013. Prvotné zistenia. s. 7.

http://www.nucem.sk/documents/27/medzinarodne_merania/talis/publikacie/ine/TALIS_2013_-_prvotne_zistenia.pdf (1. 4. 2015)

¹⁹ Stračár, E.: *Systém a metódy riadenia učebného procesu*. Bratislava : SPN, 1967, 355 s.

²⁰ Chráška, M.: *Didaktické testy*. Příručka pro učitele a studenty. Brno : Paido – edice pedagogické literatury, 1999, s. 10. ISBN 80-85931-68-0

²¹ Brown, G. – Bull, J. – Pendlebury, M.: *Assessing Student Learning in Higher Education*. London: Routledge, 1997, s. 8. ISBN 0-415-16226-2

Ihneď po r. 1990 sa začalo na pokyn MŠ SR experimentovať s tzv. **modulovým usporiadaním** vyučovania, ktoré sa v mnohých zahraničných školách úspešne uplatňuje.

Modul je väčší tematický celok (3 – 5 týždňov), ktorý má vlastné obsahové a výkonové štandardy. Počas jeho realizácie sa (zväčša) neskúša, ale po jeho ukončení sa všetci žiaci podrobujú výstupnému testu či inej forme kontroly. Tento projekt sa usiluje eliminovať vyššie spomenutý nedostatok tradičnej školy. Učiteľ po odučení modulu získa veľmi slušnú súhrnnú informáciu o tom, ako sa mu podarilo so žiakmi naplniť vzdelávacie ciele modulu. Ani tento pokus o inováciu v našom školstve sa nepodaril (hoci nevylučujeme, že sa na školách predsa len používa).

Pedagogická diagnostika. Každý (erudovaný) učiteľ vie, že skúšanie nie je ani zďaleka jedinou formou poznávania žiaka a vytvárania predstavy o jeho osobnosti. Učiteľ zhromažďuje poznatky o žiakovi prostredníctvom každého rozhovoru, na základe pozorovania jeho správania počas vyučovania i mimo neho, ako aj skúmaním rôznych produktov, ktoré žiak vytvoril.

Všetky pokusy o systematické uplatnenie pedagogickej diagnostiky žiakov na Slovensku zlyhali. Najďalej sa postúpilo na začiatku 80. rokov, keď bol do základných škôl zavedený tzv. *evidenčný list žiaka*, ale po búrlivých protestoch učiteľov, že ide o prácu navyše a o zbytočnosť, bol experiment po niekoľkých rokoch aplikácie zrušený. Prišlo to v čase, keď v mnohých vyspelých vzdelávacích systémoch sa diagnostikovanie zaviedlo do škôl celoplošne a uplatňuje sa tam dodnes.

Objasníme si rozdiel medzi **bežnou kontrolou** a **kontrolou spojenou s pedagogickou diagnostikou** žiaka. Zmyslom skúšania, ako ho poznáme z našich škôl, je zistiť vedomosti a zručnosti, ktorými žiak disponuje.

Učiteľ-diagnostik však skúma pozadie a okolnosti žiakovho výkonu, chce vedieť, v čom spočívajú jeho školské úspechy či neúspechy, čo ich zapríčiňuje a podmieňuje (**schopnosti, postoje k vzdelaniu, motivácia, vplyv rodičov**). Okrem toho učiteľ analyzuje žiakov výkon (odpoveď, prezentácia, písomné alebo materiálne dielo), identifikuje v ňom to, čo je podľa spoločensky relevantných kritérií (vedecká či vecná správnosť, originalita, estetika atď.) prijateľné alebo dokonca vynikajúce a čo objektivizovaným požiadavkám nevyhovuje. Výkon žiaka tak prestáva byť uzavretým javom, pretože sa zviditeľňuje jeho vnútorná štruktúra i vonkajší kontext.

Klasifikácia tvorí súčasť vyučovania približne od 17. storočia, u nás sa uplatnila ako oficiálna informácia o úspešnosti žiaka na základe tereziánskeho dokumentu *Ratio educationis...* Etymológia ukazuje na sémantické spojenie výrazu klasifikácia so slovom „skupina“, „trieda“. Je to teda zaradovanie objektov, v škole sú to žiaci, do stratifikovaných, výkonovo odstupňovaných skupín. V tzv. alternatívnych školách to tak nevyhnutne nemusí byť, pre ich pedagogickú teóriu je výkonová hierarchizácia žiakov neprijateľná, a preto sa tam ani obvyklým spôsobom neskúša.

Naopak, v tradičných školách, akými je aj väčšina našich škôl, je triedenie žiakov podľa školskej výkonnosti inherentnou súčasťou uplatňovanej pedagogiky. U nás pri bežnej klasifikácii zaradujeme žiakov do piatich výkonnostných stupňov, ale to je len vec tradície. Ak vytvoríme poradie žiakov podľa počtu bodov z testu (testové skóre) alebo kumulovaných percent, môže byť výkonových tried aj viac, napr. 64 (maturitný test zo slovenského jazyka a literatúry) alebo 100, ak sa výkony vyjadria v percentách.

Vnútri danej výkonnostnej skupiny sú oficiálne všetci žiaci rovnocenní, napr. všetci žiaci v skupine s poradovým číslom 2 (chváľitebný); to, že si učelia vypomáhajú plusom a mínusom či hviezdikami, je ich súkromná vec, oficiálne to neplatí. (V skutočnosti potom učiteľ používa viacstupňovú klasifikačnú škálu.)

Slovné hodnotenie. Po prevrate v roku 1989 – často na základe osobných návštev – veľa pedagogických pracovníkov zistilo, že v mnohých zahraničných školách žiak dostáva aj certifikáty s podrobným vyhodnotením jeho školskej úspešnosti; niekedy popri vysvedčení či protokole o jeho testovom výkone,

inokedy namiesto nich. Pre odbornú nepripravenosť si však neuvedomili, že si všimli iba výsledok, ktorý vyrastá z veľmi hlbokého a rozvetveného koreňového systému – pravidelnej a mnohostrannej diagnózy každého žiaka, teda z podobného systému, aký sme u nás zopár rokov predtým zlikvidovali. Pod dojmom povrchných poznatkov sa upli na sekundárny znak, totiž, že výsledky diagnostikovania sa uvádzajú v prirodzenom jazyku. A tak sa aj u nás rozbehli iniciatívy na nahradenie „zastaraného“ známkovania „modernými“ **slovnými hodnoteniami**. Dokonca MŠ SR iniciovalo niekoľko experimentov, ale všetky skončili neúspechom, pretože kváziexperti nedokázali navrhnúť nič, okrem trocha rozkošateného pomenovania známky. Pokusy a experimenty nemohli uspieť, pretože iniciátori a realizátori nepochopili, že v príslušných dokumentoch zo zahraničia nejde o slovné vyjadrenie známky, ale o demonštráciu výsledkov sústavného a systematického diagnostikovania osobnosti žiaka a analýzy výkonov žiaka.

1.3. Problematika merania vo vzdelávaní

V predchádzajúcej kapitole sa hovorilo o zisťovaní výsledkov vzdelávania pre potreby školy, učiteľov, žiaka a jeho rodičov. Lenže vzdelávanie a jeho výsledok – vzdelanosť – je aj významný spoločenský a politický jav, takže jeho výsledky zaujímajú celú spoločnosť a najmä jej vedúce osobnosti. Tie do vzdelávania detí a mládeže uvoľňujú zo štátnych prostriedkov veľké sumy peňazí a chcú sa presvedčiť o tom, aká bola ich návratnosť v náraste sociálneho blaha.

Požadovanú odpoveď môžu získať prostredníctvom *merania výstupov* zo vzdelávacieho systému. Veličín, ktoré sa merajú, je pomerne dosť (napr. počet žiakov a učiteľov, finančné náklady na vzdelávanie, úspešnosť a efektívnosť vyučovania atď.). Kontrola a klasifikácia žiackej výkonnosti, o ktorých sme doteraz hovorili, sú súčasťou tohto rozsiahleho spôsobu overovania funkčnosti vzdelávacieho systému a sú jeho významným, i keď málokedy hodnoverným prvkom.

Meranie je relatívne presná procedúra na poznávanie skutočnosti. Najskôr sa rozvinulo pri zisťovaní vlastností materiálnych vecí (dĺžka, plocha) a až oveľa neskôr (v 19. storočí) sa vedci začali zaoberať aj myšlienkou merania sociálnych a psychologických javov. Tento spôsob zmocňovania sa reality sa vždy prijímal s istou nedôverou a pochybnosťami, ktoré v rôznej podobe prežívajú až dodnes.

Podstatou merania fyzikálnych veličín je prikladanie etalónu (meracej jednotky) k zvolenému javu, v našom prípade *k edukačnému javu*. Jeden z takýchto javov je stav vedomostí a zručností, ktoré sú základom pre ďalšie vzdelávanie. Po absolvovaní školy budú žiaci nadobudnuté vedomosti a zručnosti využívať v spoločenskom živote a v práci.

Meranie edukačných javov sa realizuje ako kvantifikácia meraných javov alebo pomocou kvalitatívnych výskumných procedúr. V rámci **kvalitatívnej** analýzy sa najčastejšie používa triedenie (tabelovanie), v ktorom vystupuje jednak pomenovanie kvalitatívnych znakov (vek a pohlavie žiakov, chybovosť, prospech/neprospech na konci klasifikačného obdobia atď.). Pravda, aj v práci s takýmito znakmi sa používajú numerické operácie (početnosť jednotiek v tabuľke, priemerná početnosť atď.), ale ich význam je iba pomocný, sprostredkujúci.

Meracie procedúry uplatňované pri **kvantitatívnych** znakoch edukačných javov sú oveľa častejšie a rezultatívne efektívnejšie. Na druhej strane vo vzdelávacom systéme nie je príliš mnoho edukačných javov vhodných na kvantifikáciu a niekedy vznikajú ťažkosti pri interpretácii výsledkov ich merania. (Pozri kapitolu 3.4. Stanovenie premenných.)

Väčšina edukačných javov sú tzv. **konštrukty**, teda abstraktné pojmy, ktoré nie sú zmyslami uchopiteľné a ktorých predstavu sme si vytvorili v hlave dodatočne. Ide o výrazy ako **motivácia, charakter, pozornosť, vkus, ohľaduplnosť** atď. Aby sme si tieto pojmy dokázali predstaviť, musíme ich obsahovo

vymedziť. Spoločenské vedy, medzi nimi aj pedagogika, sa vydali do sféry merania svojich javov až potom, keď sa plánovane naučili vytvárať tieto konštrukty tak, aby sa niektoré z ich vlastností dali kvantifikovať (napr. kvocient inteligencie). *Konštruktom sa myslí celok znakov, o ktorých sa nedá uvažovať ako o izolovane pozorovateľných dimenziách správania. Jeho pochopenie vychádza z presnej a premyslenej konceptuálnej definície, ktorá vyžaduje operacionalizáciu, prevedenie na súbor úloh, ktoré sa majú riešiť, alebo otázok, ktoré majú byť zodpovedané.*²²

Aj vedomosti žiakov, o ktoré ide pri školskom testovaní, sú konštrukty. Merať ich dokážeme iba od vtedy, ako sa podarilo – napriek neustávajúcej kritike – nájsť spôsob ich *operacionalizácie*. Merať možno iba taký konštrukt, ktorému vieme priradiť číselnú hodnotu vyjadrujúcu jeho mieru. (Pozri kapitolu 3.6. Operacionalizácia premenných.)

Meranie edukačných javov si vyžaduje:

- a) **jednotky merania**,
- b) **nástroje na meranie** (testovacie nástroje).

Usporiadaná množina jednotiek merania sa nazýva **stupnica**, v niektorých starších publikáciách aj škála. Stupnic, ktoré sa používajú pri štatistickom spracovaní edukačných javov, je viacero druhov, ale najznámejšie sú dve: *poradová stupnica* a *podielová stupnica*.

Poradová stupnicu vytvára postupnosť prirodzených čísel od hodnoty 1 (prvý v poradí) až po n (n-tý v poradí). V zásade (ale neplatí to vždy) zaujíma každý meraný objekt jednu, svoju vlastnú pozíciu v poradí. Usporiadanie objektov do poradia znamená, že podľa meraného znaku sa každý nasledujúci člen vyznačuje o niečo menšou mierou posudzovanej vlastnosti ako ten, ktorý zaujíma miesto pred ním. Miery týchto rozdielov sa neskúmajú, iba sa navzájom porovnávajú: $1. > 2. > 3. > 4. > \dots > n$. Miera vlastnosti každého člena sa pri alokácii môže merať, alebo iba odhadnúť; používa sa oboje.

Osobitným prípadom aplikácie poradovej stupnice je vytvorenie poradových tried – napr. 1. – skupina najlepších; 2. – skupina stredných; 3. – skupina najslabších. V každej z takýchto skupín sa môže nachádzať akýkoľvek počet prvkov (osôb). Jednotlivci zaradení do danej skupiny sa považujú podľa meraného znaku za navzájom rovnocenných. V školstve je to triedenie (klasifikácia) žiakov podľa množstva vedomostí zistených skúšaním. Tento typ vyhodnocovania sa používa aj pri kritériových testoch. Typom poradovej stupnice je aj percentilová stupnica.

Podielová stupnica sa považuje za najlepšie vyjadrenie kvantitatívnych vlastností edukačných javov, pretože umožňuje nielen zaradenie, ale aj porovnanie veľkosti rozdielov medzi objektmi. Podielovú stupnicu tvorí postupnosť všetkých prirodzených čísel od 0 do nekonečna, pričom všetky jednotky na stupnici i v realite majú rovnaký rozmer. Sú vždy nejakým násobkom etalónu (meracej jednotky). Zo základných číselných údajov možno vypočítať napríklad priemery a iné reprezentatívne štatistické hodnoty. Používa sa napríklad na percentuálne vyjadrenie skóre.

Pravda, na to, aby bolo možné použiť jednotky podielovej stupnice ako číselné zobrazenie vlastností edukačných javov, museli byť vyvinuté nástroje na ich meranie. Takýmto nástrojom, keď zvolíme za predmet merania pedagogicky relevantné konštrukty (vedomosti, zručnosti, kognitívne operácie), sú napríklad didaktické testy – tzv. testovacie nástroje.

Aplikácia pomerovej škály a konzekventných matematických operácií nastoľuje vždy znova základnú otázku validity merania, teda: *aký je vzťah výsledných metrických údajov k realite?* Napríklad v čase vytvárania tejto publikácie televízne stanice oznamovali výsledky merania testovania deviatakov. Poslucháči sa zo správ dozvedeli, že v tomto roku slovenský jazyk dopadol lepšie ako matematika, pretože tak to ukazovali percentuálne hodnoty. Toto každoročne sa opakujúce tvrdenie spravodajských médií

²² Papica, J.: *Psychometrické metody ke konstrukci testu a sestavování testových systémů*. Olomouc : Universita Palackého, 1987, s. 4.

je pedagogicky chybné. Porovnatelnosť číselných údajov získaných meraním závisí nielen od toho, či boli podrobené správnym aritmetickým spracovaniam, ale aj od porovnateľnosti konštruktov. Pokiaľ sa konštrukty rôznej triedy neprevedú na spoločného menovateľa (príkladom je operacionalizácia čitateľskej gramotnosti, ako ju poznáme z medzinárodných meraní), tak ich číselné zobrazenia nie sú porovnateľné. Ak máme dva testy, ktoré merajú celkom odlišné konštrukty (vedomosti z matematiky a vedomosti z materinského jazyka) a ak tieto testy neboli zrovnocnené podľa špecifických kritérií, napr. podľa obťažnosti testov, tak výsledky porovnávania nemôžu byť predmetom výkonnostnej interpretácie. To isté platí napríklad aj pri medziročnom porovnávaní výsledkov testovania vedomostí z matematiky, keď každý rok je použitý iný test.

1.4. Testovanie ako forma písomnej kontroly

Ústne a **písomné skúšanie** – tieto dva spôsoby kontroly sa v škole používajú najčastejšie. Okrem nich existuje napr. meranie telesných výkonov (športové disciplíny, tanec), vyhodnocovanie materiálnych produktov (technické modely, umelecké diela, projekty).

Pri **ústnom skúšaní** je pre súčasnú didaktiku kľúčovou otázkou využitie monologického alebo dialogického spôsobu examinácie. V minulosti dominovala v škole *súvislá monologická odpoveď*, teda reprodukcia vopred zadaného textu, ktorý sa žiak naučil doma naspamäť. Pedagogické príručky sa vždy zmieňovali o tom, že učiteľ by mal nejako zasahovať do tejto predpripravenej verbálnej aktivity žiaka, a určite sa to občas aj stávalo, ale vždy išlo o zásah nepodstatný a často nevhodný. Ak žiak látku ovládal, t. j. vedel text odriekať plynulo, učiteľov zásah ho často iba zmiatol a vyviedol z tempa. A čo bolo najpodstatnejšie, tento vstup do žiakovej odpovede nemal hlbší zmysel, veď žiak opakoval kanonizovaný text. Iná vec je *dialogická forma* overovania žiakových vedomostí a zručností, ktorá tkvie v progresívnej didaktickej koncepcii učebného procesu. Takýto dialóg nie je trápny komunikatívnym hľadáním dvoch nesúladných jednotlivcov, ale partnerský rozhovor o zadanom probléme. Prvým predpokladom úspešného použitia dialógu na skúšanie je to, že takýto dialóg bol neoddeliteľnou súčasťou predchádzajúceho vyučovania. Aktivizujúce vyučovanie vedie žiakov k riadenému samostatnému poznávaniu sveta a hľadaniu odpovedí na otázky, k hodnoteniu javov, k zaujímaniu uvedomelých postojov a k spontánnemu citovému prežívaniu udalostí, ktoré sú predmetom vyučovania. V takomto prípade sa hranica medzi učením a kontrolou stiera, nejde už o dve oddelené aktivity, ale o to, že kontrola sa stáva integrálnou súčasťou učebného procesu.

Písomná forma kontroly má dve základné podoby:

1. tvorbu textu,
2. písomné riešenie testových zadaní.

Žiakova tvorba vlastného textu je náročný a pre svoju komplexnosť a originalitu aj veľmi oceňovaný spôsob demonštrácie výsledkov vzdelávania. Slohové práce, písanie textu verejnej prezentácie, vlastná žiakova umelecká tvorba či návrhy dokumentov sú len stručným výpočtom viacerých možností. Vytvorenie textu je kreatívna činnosť, ktorá aktivizuje celú osobnosť žiaka a je prejavom jeho individuality. Písomný produkt umožňuje učiteľovi dôkladne preskúmať jeho obsah a formálnu stránku a zhodnotiť ich vzájomnú prepojenosť.

Charakteristika didaktického testu. Didaktický test (ďalej len „test“) je druh skúšky. Žiak zväčša odpovedá presne určeným spôsobom (uzavreté úlohy so štruktúrovanou odpoveďou), len v menšom počte prípadov odpoveď aj tvorí (otvorené úlohy). Toto konštatovanie sa spája predovšetkým s testami používanými na celoštátnej úrovni. V testoch, ktoré si tvoria sami učitelia, by mohli otvorené úlohy zohrať významnejšiu úlohu, pretože každú odpoveď preskúma a vyhodnotí tvorca testu individuálne.

Za test nemôžeme považovať akýkoľvek list papiera, na ktorom sú napísané nejaké otázky a úlohy. Takáto chybná predstava o teste je hádam najzávažnejší a najrozšírenejší predsudok učiteľov

o testovaní. Z nej pramení ilúzia, že zostaviť test nie je príliš veľké umenie. A čo je ešte horšie, takáto predstava pre mnohých učiteľov, najmä učiteľov slovenského jazyka a matematiky, spĺňa funkciu modelu, podľa ktorého sa posudzujú štandardizované testy pre testovanie deviatakov či externú maturitu. Tieto testy sa totiž nastavujú na *optimálne cieľové požiadavky* z ročníkových a maturitných štandardov a ich výkonová stránka sa odvodzuje od najproduktívnejšej didaktickej teórie predmetu – konštruktivistickej didaktiky. Ak ju učiteľ neovláda a jeho realizácia učebného procesu za ňou veľmi zaostáva, ocitne sa on – ale predovšetkým jeho žiaci – v nepriaznivej situácii. Externý test v takomto prípade nie je kompatibilný s cieľmi a realizáciou učebného procesu v konkrétnej triede a doplácajú na to žiaci.

Didaktický test – ako naostatok každá skúška – má byť obrazom učebného procesu. Súbor viac-menej náhodne vybraných úloh by bol nepravým obrazom aj toho najhoršieho učebného procesu.

Test v zásade nie je ani horší, ani lepší nástroj kontroly ako iné formy písomnej examinácie alebo ako ústna skúška, ale je *výhodnejší* a za túto vlastnosť platí tým, že poskytuje obmedzené možnosti pokiaľ ide o predmet skúšania: v cieľoch a v obsahu vzdelávania je mnoho prvkov, ktoré sa nedajú merať testom. Použitie testu je vždy *veľký kompromis*. Ak chceme s jeho pomocou niečo získať, musíme aj niečo obetovať.

Nesporne existuje niekoľko predností didaktického testu v porovnaní so všetkými ostatnými. Hlavné prednosti testu:

1. **Vyskúšanie celej skupiny naraz** (napr. triedy, školy, ročníka). Ústnou skúškou to v porovnateľnom čase nedosiahneme nikdy. Možno sa k tomu priblížiť inými písomnými skúškami (diktát, úvaha, projekt), ale pri nich zväčša vystupuje kritická otázka vyhodnocovania, takzvaného *opravovania*, náročného na čas.
2. **Rýchla kontrola a spoľahlivosť.** Čím viac sa test konštrukčne otvára, teda čím viac je v ňom otvorených úloh, ktoré vyžadujú od žiakov tvorbu odpovede, tým viac sa prednosť rýchleho a spoľahlivého (reliabilného) overenia výsledku stráca.
3. **Možnosť dôkladnej prípravy.**

Kvalitatívnu úroveň testu zvyšuje možnosť viacnásobnej revízie, overovacieho použitia, resp. vyskúšania jeho jednotlivých úloh. Pripravený návrh testu môže posúdiť kolega či predmetová komisia a ich pripomienky môžu veľmi prospieť kvalite testu. Takmer nič z toho nie je možné urobiť pri iných písomných skúškach a len máločo z toho je možné urobiť pri ústnych skúškach.

Ak tvorca testu dobre rozumie testom, ich tvorbe a správne použitiu, vystupujú prednosti testu do popredia, pretože tvoriť testy sa *možno naučiť!*

Každá z predností testu má aj svoj protiklad. Test je hromadná skúška, ktorá umožňuje poznať aktuálnu úroveň vedomostí triedy, ale potláča individuálny rozmer žiaka. Test síce dobre diferencuje množstvo vedomostí jednotlivých žiakov, ale je oveľa menej citlivý k ich kvalitatívnym osobitostiam. Všetci žiaci, ktorí niečo vedia (sú priemerní, resp. vedľa veľmi málo), vyznievajú podobne, pretože obraz ich školskej úrovne je štandardizovaný a vyjadrený iba nejakým číslom (percentom, percentilom, individuálnym skóre).

Sumatívna a formatívna kontrola a hodnotenie výsledkov. Jedným z druhov kontroly je skúška *súhrnného typu*, ktorá smeruje k výstupnému hodnoteniu či klasifikácii, najčastejšie na konci učebného, resp. klasifikačného obdobia. Potreba overiť si a zdokumentovať, čo sa žiak za dané obdobie naučil, a podať o tom správu, je v tomto prípade primárna, ba ponajviac aj jediná úloha, ktorá sa zohľadňuje. V teórii testovania sa v takomto prípade hovorí o *sumatívnej skúške*, pretože dôraz sa kladie práve na celkový výstupný obraz (sumu) žiakových vedomostí. Takéto testovanie sa najčastejšie používa aj u nás.

Test však umožňuje aj dodatočne analyzovať žiakov výkon, preskúmať postupy, akými žiak dospel k správne alebo nesprávne výkonu, a v prípade negatívneho výsledku umožňuje začať s nápravou, teda špecifikovať a individualizovať ďalšie učenie žiaka. Pretože v tomto prípade je primárnou úlohou diagnostikovanie skrytých aspektov žiakovho výkonu a eventuálne aj jeho náprava, hovorí sa tomuto testovaniu *formatívne*.

Rozdiely medzi formatívnym a sumatívnym hodnotením nespočívajú v spôsoboch získavania informácií o školskej úspešnosti žiakov, ale v spracovaní a využití dát.

Diagnostikovanie žiakovho výkonu testom sa uskutočňuje niekoľkými spôsobmi. Najčastejšie ide o žiakovú *autokomentáre*. Učiteľ najskôr vyhodnotí test, vyznačí správne a chybné odpovede a vráti ho žiakovi. Ten rekonštruje a vysvetľuje myšlienkový postup, ktorým dospel k (správne/nesprávne) výsledku. Odhaľuje sa tak zlé prečítanie zadania, technické chyby napríklad v matematických operáciách, nesprávny úsudok alebo absencia znalosti vecného poznatku či predmetovej zručnosti. Pri komentovaní sa žiak môže radiť s učiteľom, resp. so spolužiakmi. Dôležitá je presná identifikácia nesprávneho výkonu chýbajúceho žiaka a porovnanie jeho spôsobu uvažovania a práce s postupom úspešného spolužiaka.

Diagnostická informácia o pozadí testového výkonu je nenahraditeľnou informáciou a poskytuje o žiakovi iný poznatok, ako iba výsledné testové skóre či umiestnenie žiaka v poradovníku triedy.

Ako sme už uviedli predtým, všetky tieto možnosti použitia testov a využitia ich potenciálu sú veľmi významne podmienené didaktickým modelom usporiadania učebného procesu. **V takzvanom tradičnom spôsobe vyučovania, ktorého najvýraznejšou vlastnosťou je jednostranná aktivita učiteľa a pasivita žiakov počas vyučovacej hodiny, sa spomenuté výhody testovej skúšky celkom strácajú, a preto sa testovanie vlastne ani veľmi nevyužíva.** Naopak, aktivizujúce formy učebného procesu (zaujímavý je anglický termín „learning with understanding“ – učenie sa s porozumením²³) poskytujú všetky možnosti na formatívne využitie spätnej informácie získanej kontrolou výsledkov. Učebný proces sa realizuje v relatívne uzavretých špirálach: vychádza z naplánovaných cieľov, smeruje k zisteniu, do akej miery sa podarilo počas učenia tieto ciele naplniť, a pokračuje korekčnou fázou, v ktorej sa dopĺňajú, upevňujú a systematizujú vedomosti, získané žiakmi v predchádzajúcej etape učenia. Až potom sa učebný proces presunie k novej téme (pozri obr. 2).



Obrázok 2. Schéma učebného procesu

²³ Harlen, W.: *The Role of Assessment in the Implementation of Science into Primary School*. In: Science and Primary School. Amsterdam : University of Amsterdam, 2005, s. 28.

Táto pedagogická koncepcia sa opiera o zásadne odlišné chápanie chyby v školských výkonoch, ako je to, žiaľ, dosť bežné v škole či u našej verejnosti. **Iné je nazerať na chybu ako na zlyhanie, ktoré sa hodnotí negatívne (čo je celkom samozrejme pri sumatívnej kontrole), a iné vnímať ju ako indikátor, zviditeľnenie vnútorných faktorov, ktoré ju vyvolali.** Poznávame, že nejde o žiadnu pedagogickú inováciu. Nemeckí asociacionistickí psychológovia sa zaoberali teóriou chyby už v 20. rokoch minulého storočia. Naopak, americký psychológ B. F. Skinner v 50. rokoch rozpracoval teoretickú predstavu učenia, v ktorej je chyba celkom vylúčená (error-free-learning): „Tvorca programu musí chybu predvídať a odstrániť ju...“ Český pedagóg V. Kulič (1971) vypracoval koncepciu riadeného učenia so spätnou väzbou, v ktorej je chyba v plnej miere akceptovaná a považuje sa za veľmi cenný spätnoväzbový poznatok: *Včas identifikovaná a korigovaná chyba výsledok učenia nijako neznižuje a naopak, neodhalená chyba učenie naozaj ohrozuje... V mnohých prípadoch je to práve dočasný rozpor, neúspech, chybný výkon, ktorý sa môže stať výzvou k vyššej aktivizácii činnosti a učenia, ktorý vyvoláva kognitívnu a intelektuálnu aktivitu, mobilizuje psychické sily v učení a v riešení problému.*²⁴

Testovanie v škole je súčasťou kontroly výsledkov učebného procesu, preto napĺňa všetky ciele, ktoré sa vzťahujú na školskú kontrolu. Česká pedagogička Jana Rymešová v príspevku o evalvácii – hodnotení škôl ako vzdelávacích inštitúcií – napísala: *„Škola nás zaujíma ako miesto, kde sa vzdelávací proces realizuje. Jej evalváciu však považujeme hlavne za prostriedok na zvyšovanie úrovne samotnej školy a na zlepšovanie podmienok jej činnosti. Na hodnotenie žiakov sa pozeráme ako na **súčasť** kurikulárneho procesu vnútri školy.*²⁵

Z citátu vyplýva, že testovanie môže v škole a v celom vzdelávacom systéme plniť viacero funkcií. Slúži ako nástroj na získavanie informácií dôležitých:

- a) na diagnostikovanie a klasifikáciu žiakov,
- b) pre učiteľa pri úprave plánov učebného procesu,
- c) pre riadenie školy,
- d) na externú kontrolu práce školy – inšpekciu, poradenstvo,
- e) pre riadenie rezortu školstva a celého vzdelávacieho systému (evalvácia),
- f) pre nadnárodné porovnávanie výsledkov vzdelávania.

Pre učiteľa má zmysel predovšetkým intenzívne testovanie z pomerne úzkeho rozsahu učiva. Testy pre riaditeľov škôl a školských inšpektorov sa zameriavajú na oveľa širšiu časť učebnej látky, pretože tá poskytuje príjemcom určitý nadhľad, ktorý je pre ich prácu potrebný.

Testy s celoštátnym záberom sú najčastejšie sumatívne (u nás testovanie deviatakov a externá maturita) a v mnohých prípadoch poskytujú informáciu o výsledkoch na konci príslušného stupňa školského vzdelávania.

Nadnárodné testy sa obyčajne neviažu na konkrétne učivo, ale zameriavajú sa na niektoré kľúčové kompetencie alebo zručnosti so širokým dosahom (čitateľská, matematická, štatistická, finančná, prírodovedná, počítačová gramotnosť a i.).

Častou témou na diskusiu býva vzťah testovania k iným druhom skúšok – predovšetkým k ústnym skúškam. O špecifikách oboch spôsobov examinácie sme už hovorili a práve z týchto osobitností vyvieira aj potreba kombinovania oboch foriem skúšania. V mnohých západných krajinách sa iba testuje (ale aj tam sa situácia mení, podľa toho, ktorá pedagogická tendencia má momentálnu prevahu). Presnejšie povedané, výstupné skúšky z ročníka alebo zo školského stupňa sa realizujú výlučne testovou formou (anonymne na cudzích školách a pod dozorom neznámych učiteľov; výsledky sa doručujú poštou).

²⁴ Kulič, V.: *Chyba a učení*. Praha: Státní pedagogické nakladatelství, 1971, s. 211, 212.

²⁵ Rymešová, J.: *Sledování evaluačních systémů v různých zemích*. In: Metodolgia hodnotenia výsledkov vzdelávania. Zborník príspevkov zo seminára v Budmericiach, 6. – 8. mája 1996. Bratislava : Metodické centrum v Bratislave, 1996, s. 17. ISBN 80-8052-008-9

Formatívne skúšanie je naopak veľmi uvoľnené, zameriava sa prednostne na informovanie žiaka, aby vedel, aká je jeho doterajšia úspešnosť v preberanej látke a na diagnostikovanie účinnosti vyučovacích postupov učiteľa. Výsledok sa neklasifikuje. Je vecou žiaka a učiteľa, ako naložia so získanými údajmi o svojej výkonnosti.

U nás – ako všetci vieme – sa skúša pravidelne a prevažne ústne. Len zriedka sa v priebehu učenia administruje vlastný alebo komerčný test na parciálne učivo, ktorý podá informáciu učiteľovi a zároveň upozorní žiaka, aké má vyhliadky uspieť na konci klasifikačného obdobia. Z tohto dôvodu sa trochu chýlostivá otázka vzťahu medzi výsledkami pravidelnej kontroly výsledkov vzdelávania a medzi výsledkami súhrnných testovaní na konci školského stupňa v podstate odsúva z centra našej pozornosti. Pritom je konfrontácia známok od učiteľa a výkonov v teste, keď sú aspoň čiastočne obsahovo porovnateľné, veľmi lákavá. V prvých rokoch 21. storočia V. Burjan realizoval napríklad sumatívne testovanie z matematiky, pričom porovnával poslednú klasifikáciu žiaka z daného predmetu s výsledkami dosiahnutými v teste. Porovnanie ukázalo, že vo väčšine prípadov žiaci dosiahli v teste horšie výsledky, ako naznačovala výročná známka. Pravda, toto porovnávanie nie je pedagogicky ani sociálne jednoduchá záležitosť, pretože klasifikačný stupeň sa vzťahuje na dosť odlišné okolnosti žiakovho výkonu, než za akých vzniká výkon v teste. Lenže – ako sme už povedali – táto otázka sa u nás nerieši, naopak, vedome sa obchádza.

1.5. Etika hodnotenia žiakov

Základom moderného chápania učebného procesu je interakcia medzi učiteľom a žiakmi a medzi žiakmi navzájom, a to je spoločenský vzťah. Ako každá sociálna interakcia aj táto školská sa riadi spoločenskými normami. Spoločenské pravidlá a sociálne aprobované ľudské hodnoty určujú pozíciu učiteľa, pozíciu žiaka i pozíciu rodičov, ktorí tvoria vonkajší spoločenský rámec školského vzdelávania.

Klasifikácia a hodnotenie sú zo všetkých aspektov sociálnych väzieb medzi učiteľmi a žiakmi psychologicky najvyhranenejšie, pretože sa hlboko dotýkajú predstavy žiaka o sebe samom, ovplyvňujú jeho sebavedomie a sebahodnotenie a spoluurčujú jeho miesto v sociálnom prostredí triedy a školy. Navyše, výsledky týchto pedagogických procesov prekračujú hranice školy a vyvolávajú ohlasy v širšom spoločenskom kontexte, pretože majú príznak oficiálnosti.

Z týchto dôvodov sa so školskou kontrolou spájajú všeobecne akceptované očakávania, z ktorých niektoré dostali aj podobu legislatívnych noriem a školských predpisov, ktoré zaväzujú tak žiaka, ako aj učiteľa.

Prvým pravidlom je, že učiteľ *musí skúšať a známkovať*, pretože ho k tomu zaväzujú aj školské predpisy. O pedagogických dôvodoch tejto potreby sme už hovorili. Rovnako je stanovené, že žiak sa *musí podrobiť skúške*, vyplýva to z jeho štatútu žiaka, účastníka formálneho vzdelávania, ktorý je vymedzený zákonmi. Táto záväznosť examinačnej zložky učebného procesu vytvorila aj profesionálno-etické požiadavky, ktoré sa kladú na akt skúšania, klasifikácie a hodnotenia.

Dôležitou požiadavkou kladenou na učiteľa je povinnosť vytvoriť pre kontrolovaného žiaka *optimálne sociálne prostredie*, aby spoločensky a citovo náročnou situáciou skúšaného človeka prechádzal pod čo najmenším vonkajším tlakom. V triede pri skúšaní má panovať uvoľnená, priateľská atmosféra, bez nervozity a vonkajších zásahov. Pri ústnom skúšaní v triede je to pre učiteľa skoro nezvládnuteľná situácia, pri písomných prácach a testovaní sa naopak darí dosiahnuť takýto stav pomerne ľahko.

V prípade testovej alebo inej písomnej skúšky je dôležité, aby všetci examinovaní žiaci mali *rovnaké podmienky*. Pri individuálnej skúške je to irelevantné, veď medzi odpoveďami dvoch žiakov môže vzniknúť aj niekoľkomesačná pauza, pri testovaní je to však podmienka s absolútnou platnosťou.

Puntičkárski kritici testov radi vyťahujú proti tomuto tvrdeniu námietky a dokazujú, že v triede, kde žiaci pracujú, existujú rôzne lokality vytvárajúce drobné rozdiely (okno, dopad slnečných lúčov, prehriaty radiátor atď.), ale tieto drobné odchýlky sa nepovažujú za priestupky voči zásade rovnosti podmienok.

Ďalším aspektom, s ktorým sa skúšajúci učitelia pri ústnej odpovedi vyrovnávajú veľmi ťažko, je zachovanie *dôstojnosti skúšaného žiaka*. Pokiaľ ide o výborného žiaka, ktorý si pri odpovedi počína suverénne a je si svojej pozície vedomý, nevzniká problém. Ťažkosti sa rodia v situácii, ktorú vytvorí nepripravený a celkovo neúspešný žiak. Pedagogika pripomína učiteľom potrebu *pozitívneho postoja* aj k takémuto žiakovi, zdôrazňuje potrebu vyzdvihovať akýkoľvek náznak úspešného prejavu žiaka, potláčať hodnotiace reakcie na žiakove chyby a iné nedostatky v jeho prezentácii. Je to požiadavka-ideál, ktorý sa v praxi takmer nedá dosiahnuť. Žiak zlyhávajúci pri odpovedi neposkytuje zväčša žiadne podnety, ktoré učiteľ môže zachytiť a využiť na jeho progresívnu motiváciu.

Práve tu sa prejavuje najväčšia sila konštruktivistickej didaktiky. Pri aktivizujúcom vyučovaní sa verejne prezentuje predovšetkým ten žiak, ktorý má čo povedať a čím argumentovať, a aj keď jeho príspevok učiteľ z vecného hľadiska neprijme alebo posúdi iba ako čiastočne správny, vždy je to zo strany žiaka *aktívny* čin, ktorý už len pre túto svoju stránku môže byť prijatý s porozumením, a to nielen zo strany učiteľa, ale aj zo strany žiakov, ktorí môžu na podnet spolužiaka zareagovať súhlasom, kriticky alebo rozvinutím a doplnením jeho nápadu.

Pri testovaní sa aspekt *pozitívneho postoja* k žiakovi nerealizuje vôbec. Žiak je pri riešení testu odkázaný sám na seba, nikto od neho bezprostredne „nič nepožaduje“, nikto nereaguje na jeho pokusy o riešenie ani podporou, ani kritikou. V mnohých smeroch je táto pozícia žiaka oveľa ťažšia ako pri ústnej odpovedi, ale zasa na druhej strane neexistuje ani možnosť negatívneho zásahu učiteľa či spolužiakov.

Eticky zložitá je aj *klasifikácia žiakovho výkonu/výtvoru*. Najčastejšie sa v tejto súvislosti uvádza požiadavka **objektívnosti klasifikácie**, občas dokonca aj požiadavka **objektívnosti hodnotenia**. Objektívne hodnotenie prirodzene nejestvuje, pretože hodnotenie je vždy subjektívne, ale možno uvažovať **o objektívnosti kontroly**. Pod objektívnosťou sa však myslí skôr nestrannosť, nepredpojatosť, teda to, že učiteľ pristupuje pri vyučovaní a najmä pri skúšaní k všetkým žiakom rovnako, bez predsudkov a halóefektov, nesignalizuje obľúbenosť jedných či neobľúbenosť iných.

Pri testovaní sa potlačenie možnosti segregácie niektorých skupín žiakov zdôrazňuje ako jedna z jeho najdôležitejších predností. Najmä pri externých testoch je celkom vylúčený prvok zámernej diskriminácie niektorých žiakov. Autori testu ich nepoznajú a nikdy s nimi neprídu do styku.

V teórii testovania sa však uvádzajú aj okolnosti, ktoré môžu porušovať objektívnosť a nezáujatosť testu alebo vytvárať na žiaka negatívny citový nátlak (v USA zvädzali černošskí aktivisti veľké boje so zástancami inteligenčných testov, v ktorých černošské obyvateľstvo skórovalo veľmi nízko, a označovali ich ako obsahovo diskriminačné voči afroamerickému obyvateľstvu).

Pri testovaní dochádza k diskriminácii u žiakov zo škôl, v ktorých sa ignorujú spoločensky určené ciele (žiaci sa neučia v duchu štátnych vzdelávacích programov), v ktorých sa žiaci nenaučia čítať náročné testové zadania, v ktorých sa preferuje pamäťové učenie a mechanické reprodukovanie vedomostí atď. **Všetci títo žiaci sú diskriminovaní oproti svojim rovesníkom z iných škôl, ale príčina diskriminácie nespočíva v testovaní, ale v nekvalitnom vyučovaní a nezodpovednosti pedagogických pracovníkov školy.**

Diskriminačné prvky sa môžu prejavovať v testoch aj vtedy, ak sa v ich obsahu objavujú vecné alebo jazykové prvky, ktoré v niektorých regiónoch – na rozdiel od iných – *netvoria súčasť každodennej skúsenosti* žiakov (velkomestské deti sa málo stretávajú s prírodnými javmi, ako sú stromy, poľnohospodárske rastliny alebo zvieratá; vidiecke deti majú oveľa menej skúseností s civilizačnými a kultúrnymi vymoženosťami). Tieto rozdiely do istej miery vyrovnávajú informácie masovokomunikačných prostriedkov, čo je jeden z mála ich pozitívnych príspevkov k formovaniu detí a mládeže.

Nežiaduci vplyv môžu vyvolať aj zadania úloh, ktoré sú poplatné rodovým (gendrovým) sociálnym tradíciám: v úlohách preferujúcich maskulínne vzory – silovosť, šport, motorizmus, technika – významne vyššie skórujú chlapci, v úlohách zameraných na rodinný život, domáce práce, zdravie, výživu sú zvýhodnené dievčatá.

1.6. Vplyv testovania na školské vzdelávanie

Vplyvy testovania na školské vzdelávanie môžu byť pozitívne i negatívne. O tom, ako môže testovanie pomôcť učiteľovi spoznať stav vedomostí žiakov v danom čase a upraviť vzdelávací program vzhľadom na tieto informácie, sme už hovorili. Ústnym skúšaním sa takáto hromadná informácia o žiakoch nedá získať. Pozitívny vplyv má test aj vtedy, keď povzbudí učiteľov na využívanie lepších učebných metód a foriem. Napríklad použitie interview na záverečných skúškach môže povzbudiť učiteľov k precvičovaniu konverzácie so žiakmi.²⁶

Téma neželaných dôsledkov testovania sa v teórii 60. a 70. rokov minulého storočia tešila veľkej pozornosti. Skúmal sa „spätný“ vplyv obsahu testov na obsahové zameranie učebného procesu v školách a na preferenciu jednotlivých tém v ich vzdelávacích programoch. V spomínanom období vrcholil v mnohých vyspelých krajinách vplyv pedagogického liberalizmu (Veľká Británia, USA, škandinávske krajiny). Jedným z jeho prejavov bola neexistencia celoštátnych dokumentov, ktoré by akýmkoľvek spôsobom určovali školám obsah ich kurikula. V niektorých krajinách dokonca nejstvovali ani definície najvšeobecnejších – hlavných cieľov vzdelávania. Tvorba vzdelávacieho obsahu bola výlučne v rukách školy a učiteľov. Napriek tomu sa v týchto štátoch uplatňovalo celoštátne testovanie, ktorému v rôznych krajinách pripisovali rôzny význam či dôležitosť. V dôsledku tohto sa školy dostávali pod tlak, pretože žiaci protestovali proti tomu, aby v celoštátnych testoch boli podrobovaní skúške aj z takých tém, ktoré vôbec neboli zaradené do kurikula ich školy alebo sa preberali iba okrajovo.

Školy sa usilovali tomuto problému čeliť tým, že učitelia analyzovali testy z predchádzajúcich rokov a podľa ich obsahu upravovali svoje vzdelávacie programy. To, čo sa netestovalo, odsúvali do úzadia alebo celkom vynechávali a to, čo bolo v testoch zdôraznené, posilnili aj v kurikulu. Túto prax však kritizovali pedagógovia a didaktici, ktorí považovali pragmatickú transformáciu obsahu vzdelávania za deformáciu a poukazovali na jej negatívny účinok na absolventov školy. Vedľajší účinok testovania (v angl. tzv. *backwash*) je negatívny vplyv testu na obsah vyučovania, ktoré sa prednostne zameriava na to, aby žiaci boli pripravení na test. Ak sa napríklad zručnosť písania testuje iba na uzavretých úlohách s výberom odpovede (*single choice*), potom možno očakávať, že učitelia budú precvičovať najmä takéto typy úloh, namiesto toho, aby rozvíjali samotnú zručnosť písania.

V sedemdesiatych rokoch však vo Veľkej Británii vtedajšia premiérka presadila historickú reformu školstva a nechala vypracovať prvé obsahové a výkonové štandardy pre jednotlivé predmety (zároveň sa rozhodlo, v ktorých ročníkoch britského školstva sa bude povinne testovať všetka populácia celoštátnymi testami). Britský krok rýchlo nasledovali Američania, severania, Holanďania a ďalšie krajiny. Tým, že vznikla jednotná obsahová základňa pre vyučovanie a testovanie, sa efekt spätného účinku testovania na obsah vzdelávania logicky začal vytrácať. Zároveň s tým poklesol aj záujem o teoretické štúdium spätných účinkov.

U nás pre takýto jav nejstvovali porovnateľné podmienky a navyše sa v našich školách testy používajú stále v dosť malej miere. Ale jeden príklad zmeny dôrazu v cieľových kategóriách pod vplyvom testov sme zaznamenali aj my. Celé desaťročia špecialisti na čitateľskú gramotnosť upozorňovali zodpovedné kruhy na pretrvávajúce podceňovanie problematiky čítania najmä v základných školách. Až naša

²⁶ Porovnaj: Hughes, A.: *Testing for Language Teachers*. Cambridge University Press. 1989, s. 19. ISBN 0-521-27260 a *Dictionary of language testing*. (red. Alan Davies) Cambridge : Cambridge University Press, 1999, s. 225. ISBN 0521658764.

účasť v dvoch významných medzinárodných testovacích projektoch (TIMSS a PISA) ukázala v plnom svetle tento vleký nedostatok nášho vzdelávania. Pod vplyvom nevelmi úspešného vystúpenia našich žiakov na medzinárodnej scéne sa začalo búrlivo diskutovať o probléme „čítania s porozumením“, na školy sa začal vytvárať nátlak, aby venovali tomuto cieľu vzdelávania potrebnú pozornosť a aby hľadali spôsoby na zlepšenie situácie. Či sa tento spätný vplyv medzinárodného testovania obmedzil iba na publicistiku, alebo sa preniesol aj do nášho vyučovania, nateraz nevieme.

1.7. Motivačná funkcia didaktických testov

Kedže testovanie má stále ešte svojich odporcov, spomeňme niektoré jeho výhody. Testovanie vzbudzuje v žiakoch prirodzenú súťaživosť a zvedavosť na svoj výkon v porovnaní s ostatnými. Žiakov môžeme pomocou testovania motivovať k učeniu, a to prostredníctvom netradičnej konštrukcie úloh, pomocou východiskových textov orientovaných na prax a na problémové situácie z reálneho života, používaním nesúvislých textov, najmä obrázkov, grafov, tabuliek, máp a schém, pomocou úloh na dopĺňanie a priradovanie či používaním obrázkov a symbolov v odpovediach. (V banke úloh systému eTest, ktorú pripravovali učitelia v spolupráci s posudzovateľmi, garantmi a supervízormi sa nachádza množstvo takýchto úloh.)

Testovanie má oproti ústnemu skúšaniu celý rad ďalších výhod. Oproti ústnemu skúšaniu, pri ktorom sa najviac preskúšavajú zapamätané vedomosti, sa testovanie viac orientuje na riešenie problémov. V našom školstve, ktoré je ešte stále zaťažené hodnotením vedomostí, môžu pri bežnom skúšaní teda len ťažko uspieť žiaci, ktorí sú obdarení invenciou alebo schopnosťou analyzovať. Testovanie je aj menej časovo náročné, učiteľ efektívne odskúša naraz celú triedu. Vo vzťahu k žiakom je testovanie viac objektívne (v pozitívnom aj negatívnom zmysle, teda, že žiakovi ani nezhoršuje známku, ani nenaďžda) a má pevne stanovené normy hodnotenia. Testovanie nezahŕňa do hodnotenia minulé neúspechy žiaka ani jeho správanie.

Pozoruhodné je aj nedávne zistenie z krajín Spojeného kráľovstva Veľkej Británie a Severného Írska, že zrušenie celoplošného testovania vedie k poklesu učebných výkonov žiakov. Teda aj také „testovanie pre testovanie“ má aspoň motivačný význam.

„Testovanie nie je takou ľahkou a prostou vecou, ako by sa na prvý pohľad zdalo. Nejde tu len o technickú prácu pri zostrojení testu, ale najmä o ideu, zmysel, účel testu samotného, o vyhovenie nie len psychologickým, ale aj technickým požiadavkám... Zostrojiť dobrý test znamená vyznať sa v otázkach psychológie skúšky, aj v niektorých otázkach štatistickej praxe. Testovanie nemá slúžiť len na diletantskú zábavu, ale na zobjektívnenie skúšania v škole, v neposlednom rade na poľahčenie práce učiteľa. Umožňuje ale najmä zdokonalenie vyučovania samotného, dovoľuje hlbšie poznávať, ako sa dieťa učí a tiež ako učiteľ vyučuje.“²⁷

²⁷ Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 6.

2. TEÓRIA TVORBY TESTOV

Podstatou tvorby didaktického testu je **preformulovanie špecifických cieľov do podoby úloh**.²⁸ Výber cieľov merateľných didaktickým testom z cieľov zakotvených v Štátnom vzdelávacom programe (ďalej ŠVP), identifikácia výkonov viažucich sa k vybraným cieľom, identifikácia ukazovateľov jednotlivých výkonov a operacionalizácia ukazovateľov výkonu do podoby indikátorov – testových úloh – to sú nevyhnutné kroky potrebné na zostavenie výskumného nástroja, ktorým je každý didaktický test.

Náhodný alebo reprezentatívny výber osamotených jednotlivých výkonov zo ŠVP bez identifikácie nadväznosti na konkrétne hlavné ciele vzdelávania nemôže viesť k zostaveniu kvalitného testu. Každý hierarchicky nadradený cieľ musí byť v súlade s čiastkovými cieľmi a každý cieľ musí byť v teste „ilustrovaný“ – teda reprezentovaný – dostatočným počtom úloh. Pri reprezentatívnom alebo náhodnom výbere zložiek ŠVP do špecifikácie testu sa môže stať, že niektoré osamotené úlohy budú nedostatočne reprezentovať napĺňanie cieľa, na ktorý sa vzťahujú. Pritom pri sumatívnom didaktickom teste chceme úlohami reprezentovať všetky relevantne merateľné ciele, kým pri diagnostickom (formatívnom) teste sa sústredíme na tie, ktoré sa týkajú stanovenej výskumnej otázky alebo hypotézy. V oboch prípadoch však musíme presne vedieť, na ktorý cieľ sa každá jedna úloha viaže. Ak zložíme test bez tejto nadväznosti, nemôžeme na základe výsledkov takéhoto testu seriózne povedať o žiakovi nič viac, ako dokážeme *odhadnúť* rozborom jeho výkonov v jednotlivých úlohách. Ale test má byť predsa praktickým (efektívnym) nástrojom na *meranie*. No skutočnú efektivitu (nie ekonomickosť testovania) pri hodnotení žiakov zabezpečuje práve možnosť presnej interpretácie výsledkov testu, ktorá bez premyslenej koncepcie a bez interpretačnej škály nie je možná.

Didaktický test je obvykle *batériou testových úloh*. V učiteľskom teste by malo byť minimálne 12 úloh (výnimkou môžu byť testy pre žiakov na stupni vzdelávania ISCED 1). Veľmi zriedka tvoria test 2 – 3 úlohy, testy s malým počtom úloh sa používajú vtedy, keď jednotlivé úlohy vyžadujú časovo náročné riešenie (napríklad úlohy s tvorbou dlhej odpovede – slohové úlohy a pod.). Naopak, štandardizované didaktické testy sa skladajú aj z niekoľkých desiatok úloh. Zaužívaný výraz *batéria úloh* znamená, že test nevznikne tak, že sa zhromaždí dohodnutý počet akýchkoľvek úloh, ale že sa vytvára podľa premysleného plánu, úlohy doň sa účelovo tvoria alebo starostlivo vyberajú z banky a tento výber podlieha viacerým odborným kritériám. Obsahovú a funkčnú vyváženosť testu zabezpečuje kompatibilita úloh rôzneho charakteru. Tvorca testu i testových úloh preto musí vedieť, aké druhy úloh v teste možno použiť, čím sa medzi sebou odlišujú, resp. čím sa navzájom dopĺňajú a ako sa tieto ich vlastnosti prejavujú v „zdravom“ teste, ktorý má byť jednotou mnohorakosti.

Výraz *testová úloha* je základný a pomenúva samostatnú časť testu, ktorá je graficky jasne odčlenená, viaže sa na určený prvok učebného obsahu, zahŕňa konkrétny podnet na žiakovú činnosť, má priradené poradové číslo a je vyhodnocovaná (skórovaná) nezávisle na ostatných úlohách a čiastkových úlohách.

Úloha môže byť *zložená* z viacerých čiastkových úloh, ktoré sú od seba neoddeliteľné (napr. sa viažu na ten istý východiskový text alebo na spoločnú úlohovú situáciu), sú s celým úlohovým komplexom obsahovo spojené, každá z nich má priradené svoje poradové číslo a každá z nich je vyhodnocovaná (skórovaná) nezávisle na ostatných čiastkových úlohách.

²⁸ Turek, I.: *Kapitoly z didaktiky*. Didaktické testy. Bratislava : Metodicko-pedagogické centrum, 1995, s. 34 – 35. ISBN 80-85185-96-2

Testová položka je užší pojem, rozumieť sa ňou môžu aj voľne stojace úlohy, aj jednotlivé čiastkové úlohy. Pojem položka sa najčastejšie používa pri vyhodnocovaní testu a v štatistike. O položke teda platí to, čo o úlohách a čiastkových úlohách: každá položka má v teste vlastné poradové číslo a vyhodnocuje (skóruje) sa samostatne, nezávisle na ostatných položkách. Položkou môže byť jedna voľne stojaca úloha, ale úloha sa môže skladať aj z viacerých položiek (čiastkových úloh).

Ako sme už povedali, test sa zostavuje podľa premysleného plánu z úloh, ktoré majú byť obsahovo, procesuálne i vzhľadom na svoju formu rozmanité. Výraz *batéria úloh* hovorí okrem iného aj to, že každá úloha v didaktickom teste je vlastne akýsi *samostatný test*! Úlohy spolu tvoria vyšší celok, ktorý má určité synergické vlastnosti, ale zároveň **každá úloha (aj čiastková) je zmysluplná sama o sebe a možno ju analyzovať i hodnotiť ako takú**. Tým sa didaktické testy líšia od iných typov testov, napríklad od psychologických, kde jednotlivá úloha je nahraditeľnou zložkou homogénneho testu.

S klastrami testových úloh sa však môžeme stretnúť aj pri školskom testovaní. Býva to častejšie až pri sekundárnom spracovaní testových výsledkov. Pravda, dodatočné zoskupovanie príbuzných úloh by nebolo možné, keby ich autor netvoril, resp. nevyberal do testu už s výhľadom na toto preskupenie. Na vytvorenie klastra úloh stačí väčšinou jedno kritérium, či už je to spoločná zložka predmetu (napr. tri klastre úloh: 1. z literatúry, 2. zo slohu, 3. z gramatiky), spoločná kognitívna úroveň (napr. úlohy na analýzu), alebo typ úlohy (otvorené a uzavreté). Niekedy sa vytvárajú aj skupiny úloh, ktoré spája spôsob, akým môžu byť riešené: napríklad, keď sa v teste z matematiky polovica úloh rieši bez kalkulačky a polovica s kalkulačkou.

Sekundárne transformácie testu. Uvedené poznatky o predmete testovania nie sú iba nejakým formálnym výpočtom teoretických propozícií, ale ukazujú, ako možno hodnotiť výsledky testov, ako zistiť, čo všetko vypovedajú o testovanom žiakovi. Pretože test je štruktúrovaný podľa plánu vyjadreného napr. v špecifikačnej tabuľke, možno aj jeho výsledky sekundárne triediť podľa týchto hľadísk. Obvyklý výsledok – individuálne skóre žiaka – je iba jeden z údajov, ktorý sa dá z výsledku testovania získať. Pritom nie je podstatné, či sa analyzuje jeden test, teda jeden test určitého žiaka, alebo všetky testy danej triedy či inej skupiny žiakov. Oboje má pre učiteľa a pri korigovaní plánov ďalšieho vyučovania veľkú výpovednú hodnotu.

Test možno sekundárne členiť na sektory (tzv. klastre) napríklad preskupením úloh podľa ich obsahu, takže sa vedľa seba dostanú položky vzťahujúce sa na *rovnaký tematický celok* (napr. *voľný pád; vznik, priebeh a následky tridsaťročnej vojny; domáce cicavce* atď.). Parciálne súhrnné skóre z takýchto sektorov (klastrov) podáva informácie o tom, ako žiak (trieda, skupina) ovláda príslušnú časť učiva.

Sektorovanie testu podľa obsahových položiek nie je jedinou možnosťou jeho sekundárneho čítania. Druhú možnosť poznania žiaka (triedy, skupiny) ponúka procesuálna stránka úloh. Ak si test dodatočne preorganizujeme tak, aby sa do jedného sektora (klastra) dostali úlohy vyžadujúce *rovnakú poznávaciu operáciu* (t. j. s približne rovnakou kognitívnou obťažnosťou), dozvieme sa o jednotlivcovi i o kolektíve triedy, ako vysoko sa dostali žiaci počas vyučovania na pomyselnom taxonomickom rebríčku poznávacích operácií.

Čiastkové skóre z týchto homogénnych zoskupení, sektorov či klastrov, poskytuje učiteľovi, ale i žiakom, pomerne spoľahlivú informáciu o zvládnutí príslušnej časti vecného učiva, resp. zohľadnenej kognitívnej operácie. **To znamená, že z ovládania učiva patriaceho do jedného sektoru testu nevyplýva žiadna zmysluplná pravdepodobnosť, že žiak ovláda aj učivo iných tematických celkov.**²⁹

²⁹ Systém E-test umožňuje vytváranie tzv. sekcií, do ktorých môže učiteľ začleniť úlohy podľa ľubovoľného kritéria.

Tabuľka 1. Rozvrh klastra úloh v teste

Čísla úloh testu (zvýraznené sú homogénne úlohy)																									
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
Klaster																									
Počet bodov	1 bod						1 bod		1 bod							1 bod									6 možných bodov

Didaktický test, resp. jeden jeho sektor, sa neskladá zo súboru homogénnych, navzájom nahraditeľných úloh; každá úloha je z hľadiska svojej examinačnej funkcie samostatná a jedinečná (výnimkou sú zdvojené testové úlohy pri overovaní reliability testu).

2.1. Predmet didaktického testu

V predchádzajúcom výklade sme sa už viackrát dotkli otázky, na čo sa vlastne didaktické testy zameriavajú, čo je predmetom ich merania. Didaktické testy (staršie názvy „školské testy“, „testy školskej úspešnosti“, „vedomostné testy“) zisťujú, čo sa žiaci v danom období naučili, čo si osvojili, prípadne, čo vo svojej intelektuálnej vybavenosti zmenili v súlade s externými vzdelávacími cieľmi.

Už sme povedali, že predmetom merania didaktickými testami tohto druhu sú merateľné konštrukty vychádzajúce zo vzdelávacích programov. Štátny vzdelávací program je ukotvený v Školskom zákone, ktorý stanovuje aj cieľ výchovy a vzdelávania. Z nášho hľadiska sú dôležité tie formulácie, v ktorých sa hovorí o kognitívnych cieľoch, teda o cieľoch, ktoré môžu byť predmetom didaktického testu. Cieľom teda (okrem iného) je „umožniť dieťaťu alebo žiakovi:

- a) získať **vzdelanie** podľa tohto zákona,
- b) získať **kompetencie**, a to najmä v oblasti komunikačných schopností, ústnych spôsobilostí a písomných spôsobilostí, využívania informačno-komunikačných technológií, komunikácie v štátnom jazyku, materinskom jazyku a cudzom jazyku, matematickej gramotnosti, a kompetencie v oblasti technických prírodných vied a technológií, k celoživotnému učeniu, sociálne kompetencie a občianske kompetencie, podnikateľské schopnosti a kultúrne kompetencie,
- c) ovládať anglický jazyk a aspoň jeden ďalší **cudzí jazyk** a vedieť ich používať,
- d) naučiť sa správne **identifikovať a analyzovať problémy a navrhovať ich riešenia a vedieť ich riešiť...**“³⁰

Ako východisko pre teoretické určenie pojmu „vzdelanie“ je najlepšie použiť niektorý stručný, pragmatický opis vzdelávania. Juraj Čečetka v staršom výkladovom slovníku vysvetľuje podstatu vzdelávania ako „výchovu rozumovú... ako usposobovanie človeka zo stránky rozumovej aktívne sa začleniť do ľudskej kultúry (duchovnej i hmotnej)“.³¹ Pre českého didaktika J. Maňáka (1988) je vzdelanie výsledkom komplexného edukačného pôsobenia: „...najčastejšie sa za výsledky (vzdelávania) považujú osvojené vedomosti, zručnosti a návyky, prípadne rozvoj myšlienkových operácií; úplne zriedka sa hodnotia aj vytvorené postoje a vlastnosti osobnosti.“³² Súčasný pedagogický odborník Š. Švec vníma ciele vzdelávania na najvšeobecnejšej možnej rovine: učiť sa byť, poznávať a konať.³³

³⁰ § 4 zákona 245/2008 Z. z. o výchove a vzdelávaní o zmene a doplnení niektorých zákonov.

³¹ Čečetka, J.: *Príručný pedagogický lexikon*. Martin : Kompas, 1943, s. 326 – 327.

³² Maňák, J.: *Náčrt didaktiky pre ďalšie vzdelávanie pedagogických pracovníkov*. Bratislava : Slovenské pedagogické nakladateľstvo, 1989, s. 68.

³³ Švec, Š.: *Učiť sa byť a uspieť*. In: *Pedagogická revue*. R. 54, 2002, č. 1, s. 9. ISSN 1335-1982

V odbornej literatúre nájdeme aj takúto typológiu indikátorov vzdelávania:

- emocionálne – postoje, ašpirácie, kvalita života,
- behaviorálne (týkajúce sa správania) – **zručnosti**, kooperácia, iniciatíva,
- kognitívne – **vedomosti**, presvedčenie,
- demografické – pohlavie, etnicita, sociálno-ekonomický status;
- náklady – zdroje, čas a peniaze;
- priebeh – kto ako dlho bol vzdelávaný a čo sa učil, vyváženosť kurikula, kontinuita a trvácnosť, dochádzka.³⁴

Operacionalizovaným predmetom didaktických testov sa môžu stať iba vyňaté zložky týchto široko načrtnutých cieľových kategórií – najčastejšie teda *vedomosti*, *intelektuálne* a *manuálne zručnosti*, a tiež vyššie *kognitívne* operácie, ktoré sú prejavom kompetencií.

Do predmetu didaktického testovania sa nezahŕňajú psychomotorické a afektívne ciele, teda *postoje* a ich *hodnotové východiská*, ani *presvedčenia*, *spoločenské názory*, *estetické prežitky*, ani *schopnosti*, *sklony*, *nadanie* či *motivácia*.³⁵ So zámerom zjednodušovať a spresňovať môžeme skonštatovať, že meranie didaktickými testami sa zameriava iba na tie edukačné javy, ktoré dokážeme presne zmerať. V praxi to znamená, že ide o také duchovné a materiálne vlastnosti žiakov, ktoré títo dokážu vyjadriť navonok a **prezentácia ktorých sa im javí ako prirodzený prejav ich sebarealizácie**.

Ako negatívna súvislosť s týmto sa prejavuje snaha žiakov predstierať lepšie vlastnosti, než akými reálne disponujú (odpisovanie a iné podvody pri testovaní), alebo zriedkavý, no existujúci prejav sebapoškodzovania, pri ktorom žiak vedome pracuje pod svoje možnosti (napr. deprivovaný žiak, ktorý sa chce takto pomstiť rodičom, zriedkavejšie učiteľovi).

Naše Štátne vzdelávacie programy³⁶ po roku 2008 na základe nového Školského zákona stanovili nové ciele výchovy a vzdelávania – opísané ako kompetencie, ktorých definícia v zákone sa prekrýva s *všeobecne rozšírenou* definíciou funkčnej gramotnosti.³⁷ Aj keď sa pojem kompetencia niekedy používa ako variant k pojmu gramotnosť (najmä tzv. kľúčové kompetencie), funkčná gramotnosť sa v odbornej literatúre považuje voči kompetenciám za nadradený pojem. Kým funkčná gramotnosť je chápaná ako schopnosť človeka plnohodnotne sa integrovať do spoločnosti a uplatniť sa v nej, kompetencia sa vníma len ako komplex vedomostí, zručností, postojov a hodnôt potrebných na osvojovanie si kultúry.³⁸ Zároveň sa pojem kompetencie niekedy prekrýva s pojmom zručnosti (pozri kapitolu 2.1. Predmet didaktického testu). Podľa Štátneho vzdelávacieho programu pre stupeň ISCED 2, ktorý stanovuje ciele pre základ vzdelania, na ktorý je potom na stupni ISCED 3 ďalej rozvíjaný, je hlavným cieľom vzdelávania, aby *žiaci* nadobudli primerane veku rozvinuté **klúčové spôsobilosti** (kompetencie) a **zmysluplné základné vedomosti**.³⁹

³⁴ Fitz-Gibbon, C. T. – Tymms, P.: *Technical and ethical issues in indicator systems*. Doing things right and doing wrong things. Education Policy Analysis Archives, Vol. 10, Nr. 6, 16. January 2002, s. 6. ISSN 1068-2341. <http://epaa.asu.edu/ojs/article/viewFile/285/411> (8. 7. 2015)

³⁵ Tieto vnútorné kvality človeka sú predmetom vedeckého skúmania, ale výsledky týchto výskumov netvorí súčasť procesu vnútroškolskej kontroly.

³⁶ Štátny (celonárodný) vzdelávací program (ŠVP) škôl je podľa Školského zákona hierarchicky najvyšší cieľovo-programový projekt vzdelávania.

³⁷ „Na účely tohto zákona **sa rozumie: t) kompetenciou preukázaná schopnosť využívať vedomosti, zručnosti, postoje, hodnotovú orientáciu a iné spôsobilosti na predvedenie a vykonávanie funkcií podľa daných štandardov v práci, pri štúdiu, v osobnom a odbornom rozvoji jedinca a pri jeho aktívnom zapojení sa do spoločnosti, v budúcom uplatnení sa v pracovnom a mimo-pracovnom živote a pre jeho ďalšie vzdelávanie...**“ § 2. Základné pojmy. In: Zákon 245/2008 Z. z. o výchove a vzdelávaní (školský zákon) a o zmene a doplnení niektorých zákonov.

³⁸ Švrčková, M.: *Kvalita počáteční čtenářské gramotnosti*. Výzkumná analýza a popis soudobého stavu. Ostrava : Ostravská univerzita v Ostrave, 2011, 242 s. ISBN 978-80-7 464-020-9.

³⁹ Štátny vzdelávací program pre 2. stupeň základnej školy v Slovenskej republike. ISCED 2 – nižšie sekundárne vzdelávanie. Bratislava : ŠPÚ. s. 6. http://www.statpedu.sk/files/documents/svp/2stzs/isced2/isced2_spu_uprava.pdf (1. 7. 2015)

Vedomosti sa tradične považujú za najsamozrejmejšiu cieľovú kategóriu školského testovania. Do tejto triedy edukačných javov zahrňame *pojmy, fakty, udalosti a vzťahy medzi nimi*, (napr. zákony, poučky, definície), *významné texty* (napr. ústavné zákony, texty umeleckých diel), *matematické vzorce, jednotky fyzikálnych veličín, dôležité číselné a číslkové hodnoty* (napr. významné historické dátumy – 28. 10. 1918; hodnotu čísla π atď.). Podľa nového školského zákona má škola žiakom umožniť získať **zmysluplné** základné vedomosti.

Školstvo v celom modernom svete konečne nabralo smer od vedomostí ku kompetenciám, a taxonómie cieľov nám pomáhajú vidieť, kam sa uberať (pozri tiež tab. 2 v kapitole 2.2. Bloomova taxonómia): vedomosti zaujímajú v taxonómiách kognitívnych cieľov dve najspodnejšie priečky hierarchie: v Bloomovej aj v Niemiarkovej ich nájdeme na 1. a 2. úrovni (1. zapamätať si, 2. rozumieť).

Zručnosti tvoria druhú skupinu edukačných javov a mnohé z nich, hoci nie všetky, sa stávajú predmetom overovania testom. Väčšinou ide o tzv. predmetové zručnosti, menej často o **zručnosti nadpredmetové, kľúčové**. Medzi tieto patria aj rečové zručnosti – *schopnosť viesť dialóg (rečová interakcia), počúvanie a čítanie s porozumením (rečová recepcia) aj písanie a monologický ústny prejav (rečová produkcia)*, ktoré sú súčasťou komunikačnej kompetencie a tvoria základ pre testovanie školských výsledkov žiakov z materinského jazyka alebo z cudzích jazykov. Do predmetu testovania sa však dostávajú napríklad aj konštrukčné geometrické a operačné aritmetické zručnosti, ovládanie prístrojov, výpočtovej techniky a pod.

Zručnosti spadajú u Niemiarka do 3. úrovne taxonómie (použitie v typických situáciách – špecifický transfer) a u Blooma (v revidovanej taxonómii) do podkategórie 3.1. vykonávať (spracovať) – v známych typoch úloh.

Kompetencie. Názvoslovie, ktoré by rozlišovalo kompetencie a zručnosti ešte nie je ustálené. Niekedy sú používané v horizontálnom postavení, inokedy sú zručnosti súčasťou kompetencií. (Európsky kvalifikačný rámec, ktorý bol vytvorený práve pre potrebu spoločného posudzovania kvalifikácií a pre potrebu reflexie výsledkov vzdelávania, vymedzuje kompetencie pomocou pojmov samostatnosti, kooperácie a zodpovednosti.)⁴⁰ Najčastejšie sa však kompetencie chápu ako schopnosť použiť osvojené vedomosti v zložitých, rozmanitých a nepredvídateľných situáciách – to znamená použiť vedomosti a zručnosti pri štúdiu, v práci, v bežnom živote a pri riešení problémov.

V Niemiarkovej taxonómii sa úlohy na takéto použitie vedomostí a zručností nachádzajú na úrovni 4 (použitie v problémových situáciách – nešpecifický transfer) a v revidovanej Bloomovej taxonómii spadajú takéto úlohy do podkategórie 3.2 (zavádzať (implementovať) – v neznámych typoch úloh), a do úrovne 4 (rozobrať), 5 (hodnotiť) a 6 (tvoriť).

Riešenie problémov. Škola musí žiakovi umožniť naučiť sa aj identifikovať a analyzovať problémy, navrhnúť riešenie a vedieť problém vyriešiť (pozri pozn. 29). Riešenie problémov sa chápe ako schopnosť reflektovať vedomosti a zručnosti a použiť ich v reálnych životných situáciách. V medzinárodnom výskume OECD PISA sa testuje riešenie problémov v samostatnom testovom module. U nás prebehlo takéto testovanie už v roku 2003 – 5 rokov pred školskou reformou, ktorá vniesla riešenie problémov do školského zákona. Výskum OECD PISA rozlišuje 3 úrovne schopnosti riešiť problémy: 1. riešiteľ základných problémov, 2. zdôvodňujúci a rozhodujúci sa riešiteľ problémov, 3. premýšľavý, komunikatívny riešiteľ problémov. Naši žiaci zlyhávali v triviálnych úlohách, v ktorých stačilo použiť analógiu – priemer výkonov žiakov SR bol v tejto oblasti v roku 2003 na najnižšej, t. j. na 1. úrovni.⁴¹

⁴⁰ Malčík, M.: *Role informačných a komunikačných technológií v diagnostice edukace*. Ostrava : Ostravská univerzita v Ostravě, 2013, s. 31. ISBN 978-80-7464-217-3.

⁴¹ Koršňáková, P. – Tomengová, A.: *PISA SK 2003*. Príloha. Bratislava : ŠPÚ, 2005. s. 9. http://www.nucem.sk/documents//27/medzinarodne_merania/pisa/publikacie_a_diseminacia/4_ine/PISA_2003_priloha.pdf (1. 7. 2015)

Nie všetky úlohy, ktoré sú vhodné na použitie na vyučovaní, sú vhodné aj na použitie v didaktickom teste. Na vyučovaní je niekedy vhodné ukázať žiakom mnohorakosť riešení a ukázať im viaceré hypotézy a na nich postupy vedeckého bádania a dokazovania. Pri tvorbe úloh na širšie použitie v didaktických testoch sa však vyhýbame témam, ktoré sú z vedeckého hľadiska sporné, nedoriešené, panuje okolo nich polemika, alebo sa v danej oblasti v nedávnej dobe objavili výskumy, na základe ktorých sa poznatky o danej téme prehodnocujú alebo budú prehodnocovať. Takých tém je dosť v humánnych a rovnako aj v prírodných vedách. Patria sem aj aktuálne a nedávne udalosti a nedávno vzniknuté problémy a otázky, ktoré ešte nie sú vyriešené alebo na ne existuje viacero vysvetlení a sú chápané rozlične alebo sú kontroverzné. V slovenskom jazyku sa napríklad neodporúča používať na rozbor príklady väzieb, ktoré boli v minulosti považované za primkýnanie ale dnes už sa považujú za tzv. nulovú realizáciu syntagmy, ktorú opísal v slovenskej syntaxi až nedávno profesor Ján Kačala. Neodporúča sa používať ani témy, ktoré síce už sú vyriešené v odborných diskusiách, ale výsledky ktorých sa ešte nedostali do školských učebníc. Takou je v slovenskom jazyku a literatúre napríklad téma vedľajšieho prízvuku, ktorého neexistenciu laboratórne už dávno potvrdil dnes už neexistujúci kabinet komunikácie v Nitre pod vedením profesora F. Mika, alebo napríklad otázka viacnásobne definitívne vyriešenej Sládkovičovej strofy, ktorú nakoniec vyriešil až prof. František Štraus (ide o tzv. bezhlavý sonet) a pod.

2.2. Bloomova taxonómia

Spomedzi prejavov duševnej vyspelosti žiaka spomenutých v predchádzajúcej kapitole sa pre svoju mimoriadnu dôležitosť v živote i v priebehu vzdelávania osamostatnili **logicko-poznávacie (kognitívne) operácie**. Dlhो zostávali skryté, bez explikácie v obsahu vzdelávania, a teda aj mimo predmetu examinácie.

Generálny vpád do problematiky intelektuálneho rozvoja žiaka podnietili revolučné výskumy amerického pedagóga B. Blooma, ktorý svoje zistenia publikoval na prelome 50. a 60. rokov minulého storočia. Asi o štyridsať rokov neskôr (už po Bloomovej smrti) vydal na účely vzdelávania upravenú verziu pôvodnej Bloomovej taxonómie kognitívnych operácií jeho dlhoročný spolupracovník L. Anderson.⁴²

Bloomova taxonómia kognitívnych cieľov

Pôvodná Bloomova *Taxonómia cieľov vzdelávania* (Taxonomy of educational objectives: the classification of educational goals) bola publikovaná v roku 1956 a bola vytvorená ako pomôcka pre klasifikáciu testových úloh.⁴³ Kým pôvodná Bloomova taxonómia (ďalej BT) ovplyvňovala donedávna tvorbu vzdelávacích programov najmä v západnom svete, revidovaná Bloomova taxonómia (ďalej RBT) upravená pre vyučovanie z ruky L. W. Andersona a D. R. Krathwohla (spolupracovníka B. S. Blooma) publikovaná vlastne len nedávno (v roku 2001) začala v súčasnosti progresívne ovplyvňovať tvorbu vzdelávacích programov v celom modernom svete, aj u nás.

Každá taxonómia je vlastne len hierarchicky usporiadaná klasifikácia. Jednotlivé klasifikačné kategórie sa v taxonómii nazývajú úrovne. Užitočnosť používania taxonómie pri takom všestrannom procese, ako je vzdelávanie, je viac ako zrejmá. Rovnako by sme si nevedeli predstaviť prípravu jedla pre žiakov v školskej jedálni, keby sme nemali základnú predstavu o druhoch potravín (rastlinné a živočíšne produkty) a o zložení potravín (cukry, tuky, bielkoviny a vláknina atď.). Tak, ako strava môže byť vyvážená

⁴² Anderson, L. W. – Krathwohl, D. R.: *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York : Longman, 2001, 305 s. ISBN 80-200-0524-2

⁴³ Bloom, S. B. (ed.): *Taxonomy of Educational Objectives*. Handbook I: Cognitive Domain. New York, London, Toronto: Longmans, Green & Co., 1956, s. 16.

a pestrá, aj vzdelávanie, (ktoré je určite zložitejšie, ako príprava obeda pre školáka) môže byť už pri uvedomovaní si úrovni akejkoľvek taxonómie kognitívnych cieľov pestrejším, vyváženejším, cieľavedomejším procesom, vedúcim k stanoveným cieľom – ku kompetenciám. Bez uvedomovania si pojmov ako zapamätanie, porozumenie, implementácia, rozbor, hodnotenie alebo produkovanie môže aj najlepší učiteľ skĺzať k používaniu vyučovacích metód a postupov, ktoré sú mu pohodlné a vyučovanie (a tiež skúšanie) sa tak môže stať nevdojak jednostranným, zameraným len na memorovanie zapamätaných vedomostí a pod.

Taxonómii cieľov vzdelávania je viacero, z histórie poznáme aj Herbartove štyri formálne stupne vzdelávania, ktorými sa inšpiroval B. S. Bloom:

- **fáza prehlbovania:**
 1. *vyjasniť* si predchádzajúce vedomosti
 2. prijímať nové vedomosti – *zlučovať*
- **fáza reflexie:**
 3. začleniť nové poznatky do *systému* existujúcich vedomostí
 4. použiť nové poznatky – *metóda precvičovania*.⁴⁴

M. D. Gall uvádza v tabuľkovom prehľade taxonómii teoretikov ako Adams, Aschner, Bloom, Bremer, Carner, Clements, Schreiber, Guszak, Pate atď., ktorých taxonómie sa rôzne prekrývajú a rôzne líšia.⁴⁵ U nás je najznámejšou Niemierkova a Bloomova taxonómia, a rokmi preverená (už aj revidovaná) Bloomova taxonómia sa javí ako najperspektívnejšia. Oproti Niemierkovej taxonómii má Bloomova členenie aj pre vyššie kognitívne procesy (4., 5. a 6. úroveň) a navyše je podrobnejšie rozčlenená aj v podkategóriách. Niemierkova taxonómia všetky vyššie logicko-poznávacie procesy zhrňa do jednej kategórie (nešpecifický transfer).

Tabuľka 2. Porovnanie revidovanej Bloomovej a Niemierkovej taxonómie

Revidovaná Bloomova taxonómia			Niemierkova taxonómia
1. Zapamätať si		Vedomosti	1. Zapamätanie
2. Rozumieť			2. Porozumenie
3. Použiť (v danej situácii)	3.1. Vykonávať (spracovať) – v známych typoch úloh	Zručnosti	3. Použitie v typických situáciách (špecifický transfer)
	3.2. Zavádzať (implementovať) – v neznámych typoch úloh		4. Použitie v problémových situáciách (nešpecifický transfer)
4. Rozobrať		Kompetencie	
5. Hodnotiť			
6. Tvoriť			

⁴⁴ Formalstufe. <https://de.wikipedia.org/wiki/Formalstufe> (1. 4. 2015)

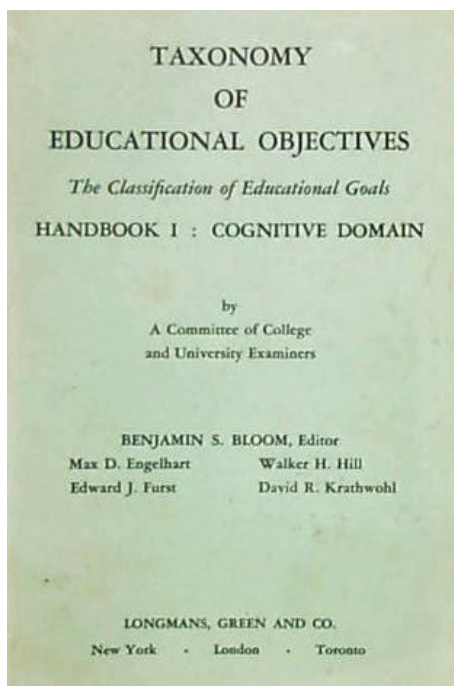
⁴⁵ Gall, M. D.: *The Use of Questions in Teaching*. In: Review of Educational Research. Vol. 40, No. 5, 1970, s. 707 – 721. ISSN 0034-6543.

Tabulka 3. Podkategórie pôvodnej Bloomovej taxonómie⁴⁶

Pôvodná Bloomova taxonómia (1956)	
Úroveň (kategória)	Podkategória
1. Vedomosti	1.10. Vedomosti o prvkoch
	1.11. Terminologické vedomosti
	1.12. Vedomosti o špecifických faktoch
	1.20. Vedomosti o spôsoboch a prostriedkoch narábania so špecifickými faktami
	1.21. Vedomosti o konvenciách (o zaužívaných spôsoboch)
	1.22. Vedomosti o trendoch a postupoch
	1.23. Vedomosti o klasifikáciách a kategóriách
	1.24. Vedomosti o kritériách
	1.25. Vedomosti o metódach
	1.30. Vedomosti o univerzáliách a abstrakcii
	1.31. Vedomosti o princípoch a generalizácii
	1.32. Vedomosti o teóriách a štruktúrach
2. Chápanie	2.10. Translácia (podať informáciu)
	2.20. Interpretácia (podať informáciu zrozumiteľnejšie alebo inak)
	2.30. Extrapolácia (rozšíriť informáciu na základe implikácií)
3. Používanie	
4. Rozbor	4.10. Rozbor prvkov
	4.20. Rozbor vzťahov
	4.30. Rozbor organizačných princípov
5. Spájanie	5.10. Vytváranie komunikácie
	5.20. Vytváranie plánu práce alebo plánovaných úkonov
	5.30. Odvodzovanie abstraktných vzťahov
6. Hodnotenie	6.10. Posudzovanie interných prvkov
	6.20. Posudzovanie na základe vonkajších kritérií

⁴⁶ Krathwohl, R. D.: *A Revision of Bloom's Taxonomy: An Overview*.
http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf (1. 3. 2015)

Revízia Bloomovej taxonómie



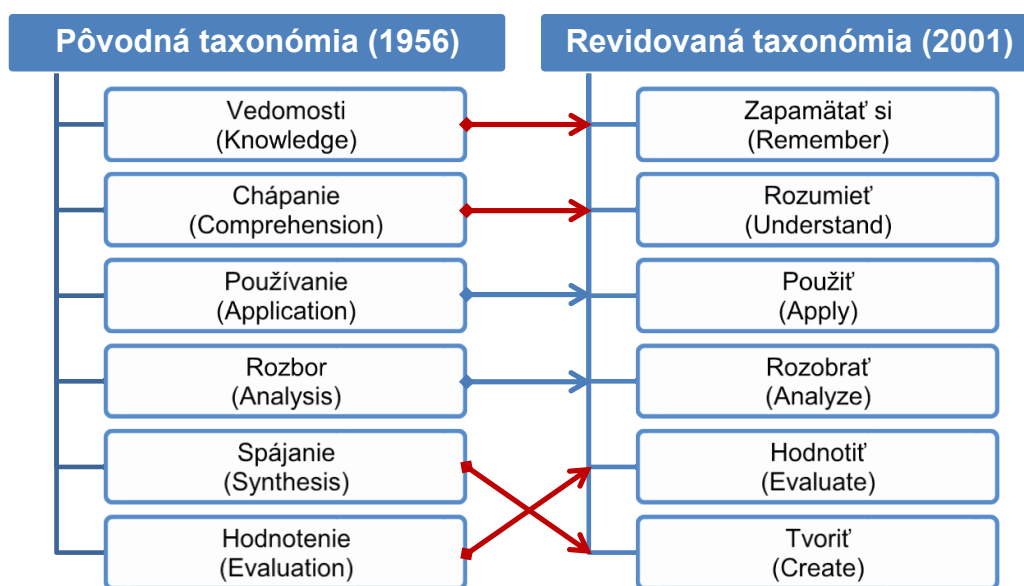
Obrázok 3. Prebal prvého vydania Bloomovej taxonómie z roku 1956 – I. Kognitívna doména⁴⁸

Publikovaná revízia Bloomovej taxonómie z roku 2001, zaviedla do pôvodnej taxonómie iba malé zmeny. Cieľom revízie bolo pôvodnú taxonómiu zostavenú pre testovanie prispôbiť **vzdelávaniu**. Revízia BT nevychádzala z potreby ju opraviť, pôvodnú BT nevyvrátili ani nové poznatky v oblasti kognitívnej psychológie, v podstate sa stále potvrdzuje správnosť pôvodného konceptu.⁴⁷

Známou črtou RBT je aj zmena názvov šiestich úrovní dimenzie kognitívnych procesov z podstatných mien na slovesá (analýza à analyzovať). K zmene na slovesá došlo preto, aby bolo možné priradiť sloveso pred podstatné meno z druhej, vedomostnej dimenzie: napr. **spracovať** (3. úroveň) podľa **postupu** (C. procedurálne vedomosti) atď. a tým pomôcť identifikovať cieľ vzdelávania, ktorý býva takto formulovaný.

Pôvodná BT zahŕňa tri domény (kognitívnu, afektívnu a psychomotorickú) a každá je venovaná celá kniha. Pre nás je z týchto troch kníh zaujímavá iba prvá, venovaná kognitívnej doméne (obr. 3). Najznámejšou črtou RBT je pridaná dimenzia vedomostí (tab. 4), ktorá vlastne vychádza z príliš veľkého počtu podkategórií pre 1. úroveň (knowledge) vyčlenených už v pôvodnej BT (tab. 3).

Jednoduchosť BT však neznamená, že jej používanie je jednoduché. V pripomienkach k revízii pôvodnej BT zaznela okrem iného aj potreba častejšie sa odvolávať na pôvodnú BT, pretože hoci sa v mnohých krajinách podľa nej pracuje, málokto pozná jej skutočné znenie).



Obrázok 4. Porovnanie pôvodnej a revidovanej Bloomovej taxonómie

⁴⁷ Zaujímavosťou je, že posledné tri vyššie úrovne pôvodnej BT (analýza, syntéza a hodnotenie) sa prekrývajú s Komenského tromi metódami pansofie opísanými v jeho Analytickej didaktike z roku 1646, pomocou ktorých je podľa Komenského možné spoznať podstatu všetkých javov: 1. analýzou (rozborom), 2. syntézou (skladaním) a 3. synkrikou (usúvzťažnením). Komenský, J. A.: Didaktika analytická. <http://muj.optol.cz/richterek/lib/exe/fetch.php?media=didaktika.pdf> (1. 3. 2015)

⁴⁸ Taxonomy of Educational Objectives: *The Classification of Educational Goals*.

<http://www.amazon.com/Taxonomy-Educational-Objectives-Classification-Goals/dp/0679302093> (1. 4. 2015)

Tabuľka 4. Dimenzia vedomostí Revidovanej Bloomovej taxonómie

Dimenzia vedomostí revidovanej Bloomovej taxonómie (2001)	
A. Faktické vedomosti	(Základné vedomosti, s ktorými sa žiak musí oboznámiť)
a)	Vedomosti a poznatky o jednotlivostiach a zložkách a ich špecifických detailoch
b)	Terminologické vedomosti (pojmy, definície)
B. Konceptuálne vedomosti	(Vedomosti o vzťahoch medzi prvkami v rámci väčších štruktúr)
a)	Vedomosti o klasifikáciách a kategóriách (triedenie faktov, pojmov, myšlienok, hodnôt na skupiny, triedy, druhy atď.)
b)	Vedomosti o princípoch a zovšeobecneniach (vzťahy, zákonitosti, súvislosti, oblasti a kontexty)
c)	Vedomosti o teóriách, modeloch a štruktúrach (princípy, hypotézy, schémy, stromy)
C. Procedurálne vedomosti	(Vedomosti o špecifických postupoch, metódach, algoritmoch, technikách, o použití zručností pri riešení otázky/úlohy/problému)
a)	Vedomosti o subjektívne špecifických zručnostiach a algoritmoch
b)	Vedomosti o subjektívne špecifických technikách a metódach
c)	Vedomosti o kritériách určujúcich, kedy použiť vhodnú procedúru
D. Metakognitívne vedomosti	(Vedomosti o poznaní, ale tiež o uvedomení si vlastného poznania)
a)	Vedomosti o všeobecných stratégiách poznávacích procesov (všeobecné metódy učenia sa, metódy rýchleho čítania a získavania informácií, práca s databázami a registrami, indexmi a katalógmi, heuristika, metódy a stratégie na riešenie problémov)
b)	Vedomosti o kognitívnych úlohách (vedomosti o kognitívnej náročnosti riešenia rôznych učebných úloh alebo problémov, o východiskách a spôsobe ich riešenia vrátane kontextu a podmienok)
c)	Vedomosti o sebe (sebapoznanie, poznanie svojich silných a slabých stránok, vedomosti o vlastnej úrovni vedomostí a zručností, o svojej vlastnej stratégii učenia sa, o vhodných podmienkach pre seba na učenie)

Pre používanie BT v prostredí tvorby testových úloh a testov sa musíme znovu vrátiť aj k pôvodnej Bloomovej taxonómii, ktorá bola vytvorená na účely klasifikácie testových úloh. Dôvodom je pomenovanie niektorých úrovní, ktoré je iné pri vyučovaní, a iné pri testovaní. Napríklad, kým pri vyučovaní vedomostí je dôležité, aby si žiak vedomosť zapamätal, pri testovaní je dôležité, aby si žiak vedomosti vybavil z pamäti. Testových úloh, pri ktorých žiak používa zapamätávanie, je málo a väčšinou sa týkajú testov všeobecných schopností alebo testov z cudzích jazykov. V takých úlohách si má žiak pozorne vypočítať alebo prečítať ukážku v stanovenom čase, a v ďalšej fáze riešenia úlohy, keď si ukážku už nemôže znovu vypočítať alebo prečítať, má na základe zapamätaných informácií (z krátkodobej pamäti) vyriešiť úlohu a odpovedať. Naopak, väčšina úloh na 1. úroveň (BT: Vedomosti, RBT: Zapamätať) využíva proces vybavenia si vedomostí z dlhodobej pamäti alebo rozpoznanie určitých prvkov v ukážke úlohy na základe predchádzajúcich vedomostí.

Pri práci s BT je veľmi potrebné naštudovať si aj jej **podkategórie**, inak sa budeme pohybovať vo veľmi všeobecných pojmoch. V našom prostredí je to dôležité o to viac, že mnohé cudzie slová (vrátane anglických pomenovaní úrovní BT) sa v slovenčine používajú pomerne voľne, a nie vždy správne. Problémy vznikajúce s používaním už zaužívaných cudzích slov spôsobujú celý rad chýb pri používaní Bloomovej taxonómie na klasifikáciu testových úloh. Najtypickejším príkladom sú názvy úrovní použitých (aplikovať) a rozobrať (analyzovať), do ktorých majú mnohí tvorcovia úloh tendenciu zaraďovať úlohy, pri ktorých žiak vykonáva akúkoľvek logickú operáciu („aplikuje“) alebo pri ktorých akokoľvek rozmýšľa nad úlohou a možnosťami odpovede („analyzuje“). Problematickým sa v našom prostredí javí aj názov 6. úrovne (*tvoriť*), ktorý sa u nás prekrýva s pojmom *tvorivosť* (v školstve v súčasnosti frekventovaným), a to spôsobuje, že sa niektorí autori nazdávajú, že nie je možné tvoriť úlohy pre túto úroveň. Pritom do tejto úrovne patria napríklad úlohy na plánovanie, zjednodušenie, zovšeobecňovanie (abstrakciu) a pod. Spomenutý problém s používaním cudzích slov sa týka aj samotných činnostných (aktívnych) slovies.

Tabuľka 5. Podkategórie revidovanej Bloomovej taxonómie

Revidovaná Bloomova taxonómia (2001)	
Úroveň (kategória)	Podkategória
1. Zapamätať si (Remember) (vybavovať si relevantné vedomosti z dlhodobej pamäti)	1.1. Rozpoznávať
	1.2. Vybavovať si
2. Rozumieť (Understand) (určiť význam na základe predložených informácií a ústne, písomne alebo graficky ho vyjadriť)	2.1. Interpretovať
	2.2. Uviesť príklad
	2.3. Klasifikovať
	2.4. Sumarizovať
	2.5. Usudzovať
	2.6. Porovnávať
	2.7. Vysvetľovať
3. Použiť (vykonať podľa návodu alebo použiť postup v danej situácii (in a given situation))	3.1. Vykonávať (spracovať) – v známych typoch úloh (familiar task)
	3.2. Zavádzať (implementovať) – v neznámych typoch úloh (unfamiliar task)
4. Rozobrať (rozložiť materiál na jednotlivé súčasti a zistiť, v akom vzťahu sú tieto časti, a aký je ich vzťah k celku alebo k hlavnej funkcii materiálu)	4.1. Rozlišovať
	4.2. Usporiadať (organizing)
	4.3. Prisudzovať (attributing)
5. Hodnotiť (posudzovať na základe kritérií a štandardov)	5.1. Kontrolovať
	5.2. Kritizovať
6. Tvoriť (spájaním častí vytvárať nové, zaujímavé vnútorne súdržné celky alebo vytvárať originálne produkty)	6.1. Vytvárať (generating)
	6.2. Plánovať
	6.3. Produkovať (Producing)

Niektoré publikácie a voľne šírené „pomôcky“ ponúkajú stručné, prehľadné tabuľky dvojdimenzionálnej RBT, v ktorých sú v jednotlivých bunkách uvedené niektoré „typické“ činnostné (aktívne) slovesá (viac či menej správne). Pri práci s BT však nestačí pracovať so zjednodušenými prehľadmi alebo iba s názvami šiestich úrovní. Ani obmedziť sa na známu dvojdimenzionálnu tabuľku s prehľadom šiestich základných úrovní kognitívnych procesov v riadku a so štyrmi druhmi vedomostí vo vedomostnej dimenzii v stĺpci vonkoncom nemôže byť postačujúce (tab. 6). Pritom v mnohých publikáciách sa pri opise RBT autori obmedzujú práve na túto tabuľku. Táto tabuľka však má slúžiť ako sumárny plán rozloženia úloh v taxonómii pre tvorbu testu alebo ako prehľadný sumár úloh podľa úrovní po vytvorení testu, nie ako pracovný nástroj pri zaraďovaní úloh.

Tabuľka 6. Sumárna tabuľka revidovanej Bloomovej taxonómie

			Dimenzia kognitívnych procesov					
			1	2	3	4	5	6
			Pamätať si	Rozumieť	Použiť	Rozobrať	Hodnotiť	Tvoriť
Dimenzia vedomostí	A	Faktické vedomosti						
	B	Konceptuálne vedomosti						
	C	Procedurálne vedomosti						
	D	Metakognitívne vedomosti						

Kým u nás sa teória testovania v školskom prostredí nevyučuje ani ako voliteľný predmet, na Univerzite v Pensacole (USA) existuje centrum CUTLA (Center for University Teaching, Learning, and Assessment) venujúce sa meraniu vzdelávania v univerzitnom prostredí. Na ich internetovej stránke je publikovaný zoznam slovíes pre Bloomovu taxonómiu,⁴⁹ ktorá je fúziou pôvodnej a revidovanej BT. Poradie posledných dvoch úrovní a tvar názvov prevzali z RBT, ale názov prvej úrovne opravili podľa pôvodnej BT. V tejto taxonómii sa nachádza pod jednotlivými úrovňami spolu až 160 rozličných *činnostných* (aktívnych) slovíes (pre všetkých 6. úrovní ich v preklade do slovenčiny uvádzame nižšie).

V našom prostredí vznikajú problémy nielen pri preklade názvov úrovní ale aj s prekladom jednotlivých slovíes. Tie majú neraz veľa významov, z ktorých nie všetky sa vzťahujú na činnosti, ktoré je možné použiť v oblasti vzdelávania. Zároveň nie je možné tieto slovesá vhodne preložiť vždy iba jedným slovom – slovesom. Špecifikom slovenčiny je aj to, že dokážeme formulovať úlohy jednoducho aj bez toho, aby sme použili náležité *činnostné* (aktívne) sloveso. Zároveň však slovenčina umožňuje použiť v zadaní aj slovesá, ktoré sú z hľadiska kognitívnych procesov, potrebných na vyriešenie úlohy, nenáležité alebo nesprávne, resp. na základe ktorých nie je možné mechanicky zaradiť úlohu do príslušnej úrovne, ako to zjednodušene podávajú niektoré príručky. Napríklad: posúď, na ktorom obrázku je hmyz (správne: 1. rozoznať, poznať); napíš názov najvyššieho zákonodarného orgánu SR (správne: 1. vedieť si vybaviť z pamäti – úloha netestuje písanie); urči, ktorý postup je efektívnejší (správne: 5. porovnaj postupy na základe daných kritérií a vyber efektívnejší atď.). Preto je potrebné venovať zvláštnu pozornosť zadaniu úloh, a nenáležité slová vypustiť, nahradiť adekvátnymi, alebo aspoň zaradiť úlohu do príslušnej úrovne na základe postupu riešenia, nie na základe použitého slovesa.

Mechanické priradovanie na základe *činnostného* (aktívneho) slovesa nie je možné vykonať ani v anglickom znení úloh, a to ani za predpokladu, že sa použilo to správne sloveso – niektoré slovesá sú totiž charakteristické pre viaceré úrovne (napríklad sloveso *vybrať* (select) sa uvádza až pri štyroch úrovniach, 1., 2., 3. a 5.). V zozname slovíes, ktorý vypracovalo centrum CUTLA sa ešte ďalších 51 slovíes zo 160 nachádza v dvoch alebo troch úrovniach RBT naraz.

Pre utvorenie lepšej predstavy o úrovniach RBT sme tieto slovesá z tabuľky vybrali a uvádzame ich v nasledovnej samostatnej tabuľke. Navrhli sme preklad jednotlivých významov týchto slovíes tak, aby sa dali identifikovať s jednotlivými úrovňami RBT. Významy sú radené za sebou a hierarchicky usporiadané od najjednoduchšieho po najzložitejšie (zľava doprava). Je zaujímavé, ako rozličné významy toho istého slovesa pri rôznych úrovniach ilustrujú hierarchiu úrovní RBT – hierarchia je takto zreteľnejšia, ako keď sú úrovne reprezentované iba skupinami rozličných slovíes.

⁴⁹ Action Words for Bloom's Taxonomy. Center for University Teaching, Learning, and Assessment. University of West Florida, 11 University Pkwy, Pensacola, FL 32514.
<http://www.cte.cornell.edu/documents/Assessment%20-%20Blooms%20Taxonomy%20Action%20Verbs.pdf>
 (15. 07. 2015)

Tabuľka 7. Preklady činnostných (aktívnych) sloviess, ktoré sa vzťahujú na viaceré úrovne RBT

Činnostné slovesá ktoré sa vzťahujú k viacerým úrovňam RBT	Preklady činnostných sloviess pre jednotlivé úrovne RBT
appraise (4, 5)	stanoviť hodnotu (4), stanoviť kvalitu (5)
calculate (3, 4)	vypočítať (3), spočítať niektoré prvky (4)
classify (2, 4)	zaradiť do širšej skupiny alebo triedy (2), roztriediť podľa podrobnej klasifikácie (4)
collect (3, 6)	zhromaždiť informácie (3), vytvoriť z častí zmysluplný celok, súbor (6)
compare (2, 4, 5)	porovnať s odlišným (2), porovnať s podobným (4), porovnať s predpisom, s normou (5)
conclude (4, 5)	dospieť k záveru (4), urobiť záver na základe kritérií (5)
construct (3, 6)	zostrojiť podľa predlohy (3), vytvoriť niečo (6)
contrast (2, 4)	nájsť jednoduchý protiklad (2), odhaliť negatívne stránky niečoho (4)
criticize (4, 5)	pripomienkovať (4), kriticky komentovať (5)
demonstrate (2, 3)	ukázať na príklade (2), dokázať vykonať nejakú operáciu, činnosť alebo postup podľa inštrukcií (3)
describe (1, 2)	vyznačiť, popísať (1), podrobne opísať (2)
develop (3, 6)	dokončiť podľa návodu (3), rozvinúť, zmysluplne dokončiť (6)
devise (4, 6)	navrhnuť (4), upraviť plán, zjednodušiť, vylepšiť niečo (6)
differentiate (2, 4)	vidieť rozdielnosť (2), nájsť rozdiely (4)
discover (1, 2, 3)	rozpoznať (1), uvedomiť si (2), odhaliť, nájsť (3)
discriminate (4, 5)	nájsť rozlišujúci prvok (4), vnímať podstatu rozdielu (5)
distinguish (2, 4, 5)	rozlíšiť zjavné rozdiely (2), rozoznať ťažko rozoznateľné (4), identifikovať charakteristický znak (5)
estimate (2, 4, 5)	zhruba odhadnúť (2), stanoviť pravdepodobnú hodnotu (4), oceniť hodnotu niečoho (5)
evaluate (4, 5)	číselne vyjadriť počet alebo hodnotu, vytvoriť si predstavu o počte alebo o hodnote (4), zhodnotiť (5)
examine (1, 3)	odpovedať na otázku (1), položiť otázku, preveriť niečo podľa inštrukcií (3)
experiment (3, 4)	vykonať pokus podľa inštrukcií (3), pokusom overiť (4)
explain (2, 3, 4)	objasniť niečo (2), vysvetliť postup (3), popísať prvky a ich vzťahy (4)
express (2, 6)	nájsť zástupný znak, symbol prvku alebo vzťahu, vyjadriť inak (2), sformulovať myšlienku (6)
generalize (2, 6)	vztiahnuť na väčší súbor (2), zovšeobecniť (6)
choose (3, 5, 6)	vybrať niečo podľa parametrov (3), vybrať niečo podstatné (5), zvoliť niečo najvhodnejšie (6)
identify (1, 2)	spoznať niečo (1), priradiť časť alebo prvok k správnejmu celku alebo k štruktúre (2)
illustrate (2, 3, 4)	uviesť príklad na niečo (2), vytvoriť ukážku niečoho podľa inštrukcií (3), nájsť príklad niečoho v danom materiáli (4)
infer (2, 4, 6)	pochopiť zo slov a údajov (2), dedukovať z dôkazov (4), urobiť záver z dôkazov a úsudkov (6)
interpret (2, 3)	chápať význam celku (2), vysvetliť, podať niečo v zrozumiteľnejšej forme (3)
judge (2, 3, 5)	rozumieť názoru alebo rozhodnutiu (2), dospieť k rozhodnutiu podľa inštrukcií (3), rozhodnúť na základe kritérií (5)

justify (5, 6)	dokázať, deklarováť správnosť niečoho (5), dať niečo do poriadku (6)
list (1, 3)	vymenovať, nájsť v zozname, zaradiť do zoznamu (1), pracovať so zoznamom (3)
modify (3, 6)	upraviť podľa inštrukcií (3), invenčne upraviť na niečo iné (6)
observe (1, 2)	pozorovať a zaregistrovať niečo (1), všímať si s prihliadnutím na kontext (2)
order (2, 4, 5)	rozumieť pokynom (2), nájsť pravidlo usporiadania (4), usporiadať niečo podľa kritérií (5)
organize (4, 6)	systematicky usporiadať prvky (4), skoordinať aktivity (6)
plan (4, 6)	plánovite postupovať (4), napláňovať (6)
predict (2, 3, 5)	odhadnúť význam (2), predvídať pokračovanie procesu, činnosti (3), predpovedať dopad, efekt (5)
prepare (3, 6)	predpripraviť (3), spraviť niečo ľahšie použiteľným (6)
produce (3, 6)	vyrobiť podľa návodu (3), vyprodukovať niečo nové (6)
record (1, 3)	presne zapísať, vypísať (1), zaznamenať niečo konvenčným alebo predpísaným spôsobom (3)
relate (2, 3)	vidieť súvislosť (2), usúvzťažniť na základe pravidiel (3)
report (2, 3, 6)	informovať (2), zostaviť informáciu podľa inštrukcií (3), sformulovať správu (6)
rewrite (2, 6)	nanovo napísať (2), vylepšiť prepísaním (6)
select (1, 2, 4, 5)	vybrať niečo explicitné (1), vybrať niečo implicitné (2), vybrať časti z celku na základe viacerých kritérií (4), vybrať niečo najvhodnejšie (5)
show (2, 3)	ukázať na niečo na základe porozumenia (2), ukázať na niečo konkrétne na základe inštrukcií (3)
simulate (3, 6)	napodobniť niečo na základe inštrukcií (3), tvorivo niečo napodobniť (6)
solve (3, 6)	vyriešiť problém (3), vysvetliť záhadu (6)
summarize (2, 5)	stručne vyjadriť podstatu (2), zhrnúť (5)
support (5, 6)	podoprieť tvrdenie, odporučiť niečo (5), sformulovať odporúčenie (6)
test (4, 5, 6)	urobiť skúšku správnosti (4), preveriť súlad s kritériami (5), odhaliť skryté kvality niečoho (6)
write (3, 6)	napísať na základe inštrukcií (3), tvorivo niečo napísať (6)

Nielen pri takýchto slovesách preto musíme úlohu zaraďovať do príslušnej úrovne nie mechanicky, podľa použitého slovesa, ale na základe posúdenia logicko-poznávacích **procesov, ktoré sú potrebné na vyriešenie úlohy**. Posúdenie vykonáme tak, že úlohu vyriešime a celý postup si zaznamenáme. Jednotlivé operácie, ktoré vykonávame, analyzujeme a snažíme sa identifikovať rozhodujúci kognitívny proces. Obzvlášť si treba dať pozor, aby úloha nebola nadhodnotená – zaradená do vyššej kategórie, čo je pomerne *častá chyba*. Aj keď je úloha už klasifikovaná podľa RBT, odporúčame pred použitím úlohy v teste vykonať kontrolu jej klasifikácie podľa tohto postupu:

1. vyriešiť úlohu tak, že si celý postup od prvého kroku po posledný (každú maličkosť) zaznačíme (prečítame si ukážku, prečítame si zadanie úlohy, vyhladáme explicitne uvedené informácie v grafe, porovnáme vyhladané hodnoty, zvýrazníme si vyššiu hodnotu, prezrieme si legendu ku grafu, priradíme hodnoty k názvom skupín, zvýrazníme si vybranú skupinu, znovu si prečítame zadanie úlohy a skontrolujeme si výsledok (ak na to upozorňuje úloha), označíme správnou odpoveď a pod.)
2. každý jednotlivý krok analyzujeme a klasifikujeme pomocou BT,
3. určíme rozhodujúci logicko-poznávací proces (**najvyšší** z tých, ktoré sa pri celom postupe riešenia úlohy **musia** použiť).

Pre lepšiu orientáciu v úrovniach RBT ale aj pre lepšie rozlišovanie medzi nimi je potrebné neustále sa vracieť k podkategóriám a k celému zoznamu slovies jednotlivých úrovní, a dlhšie uvažovať nad tým, ktorý logicko-poznávací proces pri riešení naozaj prebieha, ktorý nie, ktorý postačuje atď. Ak pri kontrole zaradenia úlohy do BT odhalíme v postupe kognitívny proces, ktorý je nevyhnutný na vyriešenie úlohy ale nie je našim cieľom ho v úlohe merať (napríklad nejaká vedomosť, poznanie špecifického detailu a pod. pri úlohách na vyššie kogn. procesy), môžeme zadanie úlohy doplniť o definície, vysvetlenia alebo potrebné údaje, postupy či vzorce.

Potreba správnej klasifikácie úloh bola pre B. Blooma podnetom pre vytvorenie pôvodnej taxonómie v roku 1956. Ani zďaleka to nebola potreba niečo škatulkovať alebo klasifikovať pre potešenie nejakého štátneho úradníka. Bez správne zaradených úloh by nebolo možné zostavovať testy, ako sú PISA a pod., ktoré žiakov rozdeľujú podľa schopností do skupín. Bez správne kalibrovaných a zaradených úloh nemožno realizovať ani adaptívne testovanie (pozri kapitolu o adaptívnom testovaní). V takom prípade by sme žiakovi, ktorý nezvládol nejakú úlohu, zadali ako ďalšiu nie jednoduchšiu, ale možno práve kognitívne náročnejšiu úlohu, a pod. Bez správne zaradených úloh nie je možné zostrojiť ani dobrý formatívny test na diagnostické účely a úpravu vyučovacích postupov. Ak by boli úlohy zaradené nesprávne, mohli by sme mať úplne pomýlený obraz o schopnostiach našich žiakov. Nehovoriac o tom, že progres v rozvoji gramotnosti (matematickej, finančnej, štatistickej, čitateľskej alebo v občianskej) by sme mohli naďalej usmerňovať iba na základe odhadov a sond – pomocou analýzy vybraných osamotených úloh. Alebo by sme mohli dúfať, že sa nám podarí nájsť finančné príspevky na zapojenie sa do medzinárodných meraní aj v oblastiach, ktoré sú u nás síce považované za kľúčové, ale monitorované sú iba okrajovo, napr. občianska gramotnosť (IEA: ICCS – Civic and Citizenship Study) a pod. Bez správneho zaraďovania úloh do taxonómie sa stáva aj pestrosť testovania a samotného vyučovania iba iluzórnou a formálnou, veď často úlohy zaradené do vyšších úrovní po podrobnejšom preskúmaní spadajú nanajvýš do prvých dvoch úrovní. Preto správne a uvážlivé používanie BT jednak zaručuje všestrannosť, pestrosť a vyváženosť vyučovania, pri ktorom sa aj naplňujú vznešené ciele v Štátnych vzdelávacích programoch, a zároveň zaručuje aj cieľavedomosť, zmysluplnosť, konzistentnosť a rigoróznosť testovania.

Činnostné (aktívne) slovesá v úrovniach Bloomovej taxonómie

1. úroveň: Pamätať si (RBT); Vedomosti (BT)

• vytvoriť kópiu, napodobniť • presne zopakovať, duplikovať • **vymenovať jedno po druhom** • **označiť**, pooznačovať • **nájsť zhodu**, spárovať • zapamätať si • **pomenovať**, nazvať, **dať prívlastok** • vynechať, **vyľúčiť** • zopakovať, citovať • čítať (s porozumením) • **vybaviť si** • recitovať, opakovať z pamäti • zopakovať, znovu vykonať • zopakovať, vytvoriť kópiu • prerozprávať, povedať inak • stanoviť • vystihnúť povahu, význam alebo rozsah niečoho, **definovať** • usporiadať • podať (informáciu) • znázorniť • venovať niečomu pozornosť, vypočítať si • umiestniť, **nájsť miesto alebo pozíciu niečoho** • znovu spoznať na základe predchádzajúcich vedomostí • vyznačiť, popísať (1) • **rozpoznať** (1) • odpovedať na otázku (1) • spoznať niečo (1) • **presne zapísať**, vypísať (1) • **vybrať niečo explicitné** (1) • vymenovať, **nájsť v zozname**, zaradiť do zoznamu (1) • pozorovať a zaregistrovať niečo (1)⁵⁰

Každá úloha, ktorú je možné vyriešiť **iba** na základe zapamätaných vedomostí, môže byť zaradená iba do 1. úrovne. Do tejto úrovne patria úlohy, ktoré prevažne overujú schopnosť vybaviť si zapamätané **zmysluplné základné** vedomosti z dlhodobej pamäti pri riešení úlohy. Žiaci si dokážu spomenúť na definície, dôležité dátumy, osoby, udalosti, javy, procesy a pod. Tento typ úloh býva zväčša formulovaný ako zisťovacia otázka (Ako sa nazýva to a to? Ako prebieha/prebiehalo to a to? Kde sa nachádza to a to? Kto sa zaslúžil o to a to?), ale aj ako výzva k činnosti, a môže byť spojená s úlohovou situáciou aj s východiskovým textom.

V úlohách s východiskovým textom na tejto úrovni dokážu žiaci vypísať informácie explicitne uvedené v texte, vyhľadať informácie v zozname či v inom nesúvislom texte, dokážu odčítať hodnoty na stupnici, nájsť konkrétne miesto na mape, vybrať z ponuky obrázkov, na ktorom je zobrazená požadovaná vec, osoba alebo požadovaný jav, proces a pod. Úlohy na vyhľadanie informácií explicitne uvedených v texte bývajú niekedy nesprávne zaradované do 2. úrovne s nesprávne položenou otázkou: „čo vyplýva z textu ukážky?“ – explicitné informácie však z textu nevyplývajú, naopak, sú v ňom uvedené. O explicitne uvedenej informácii hovoríme aj vtedy, keď je časové a miestne rozpätie alebo rozpätie v počte naznačené pomlčkou alebo predložkami (strana 5 – 12, v rokoch 1989 – 1993, január – február, Bratislava – Budapešť (na vlakovej trati), od 8.00 do 18.00 h) alebo ak je opísané slovami (v prvej polovici roka, v poslednom desaťročí a pod.).

42

Negatívne formulovaná úloha testujúca zapamätané vedomosti môže nechtiac merať schopnosť kontroly, ktorá v tomto prípade síce spadá do 1. úrovne BT, ale úloha testuje napĺňanie iného cieľa, ako bolo zamýšľané. Ak žiak za takúto úlohu nezíska bod, nebude sa dať určiť, ktorý cieľ sa tu nepodarilo dosiahnuť, či si žiak nevšimol negatívnu formuláciu a potom nevenoval pozornosť všetkým možnostiam odpovede a vybral prvú správnu odpoveď na pozitívnu formuláciu, alebo si požadovanú vedomosť nevedel vybaviť z pamäti. Takáto úloha by teda merala inú schopnosť, ako sme zamýšľali, a bola by teda nevalidná. Ak sa nemôžeme takýmto formuláciám vyhnúť, je potrebné v inštrukciách k úlohe žiaka upozorniť, aby sa k zadaniu vrátil alebo aby si ho pozorne prečítal, resp. aby si skontroloval výsledok.

Pozorne posúdiť je potrebné aj úlohy, v ktorých žiakom podávame inštrukcie k spôsobu zaznamenania odpovede (najmä pri otvorených úlohách s tvorbou krátkej odpovede). Tu je potrebné zvážiť najvyššiu vyžadovanú kognitívnu úroveň a dať ju do súladu s cieľom testovej úlohy. Ak by sme totiž v úlohe na 1. úroveň vyžadovali od žiakov, aby svoju odpoveď sformulovali do jedného slova alebo slovného spojenia, mohlo by sa stať, že presiahneme zamýšľaný cieľ, a vytvoríme úlohu vyžadujúcu schopnosť

⁵⁰ **Knowledge:** copy • duplicate • enumerate • label • match • memorize • name • omit • quote • read • recall • recite • repeat • re-produce • retell • state • define • tabulate • tell • visualize • listen • locate • recognize • describe (1, 2) • discover (1, 2, 3) • examine (1, 3) • identify (1, 2) • record (1, 3) • select (1, 2, 4, 5) • list (1, 3) • observe (1, 2)

zovšeobecniť, zhrnúť, zostručniť myšlienku, čo by už mohla byť 6. úroveň Bloomovej taxonómie. Podobne v matematike: požiadavka na formu odpovede, ktorú aj zohľadňujeme pri vyhodnocovaní úlohy (napríklad: uviesť výsledok zaokrúhlený na dve desatinné miesta) môže znamenať posun do úrovne 3.1 v RBT (vykonávať), keďže sa tu očakáva zručnosť pri vykonávaní rutinej operácie.

Vytvoriť úlohu na túto úroveň je asi najjednoduchšie. V procese tvorby úloh sme si dokonca všimli, že niektorí autori úloh napriek snahe vytvoriť úlohy na vyššie úrovne kognitívnych procesov nechtiac vytvárali stále úlohy na 1. úroveň, akurát, že ich zaraďovali do nesprávnej úrovne BT.

Napríklad:

Označ drogu, ktorá patrí v Slovenskej republike medzi nelegálne.⁵¹

- A. tabak
- B. marihuana
- C. káva
- D. alkohol

Úloha bola **nesprávne zaradená** do úrovne 4. (rozbor/analýza), a teda merala niečo iné, ako chcela, **bola teda nevalidná** (pozri obrázok 9/b v kapitole 2.4.2. Reliabilita). Ako dôvod zaradenia autorka uviedla, že žiak musí porovnať účinky drog a určiť ich legálnosť. Po diskusii s posudzovateľom sa zhodli, že úloha patrí do 1. úrovne (zapamätať si), pretože žiaci získali informácie o legálnosti drog priamo na vyučovaní.⁵²

V takomto type úloh nie je čo analyzovať – nepredložili sme žiakovi žiadny materiál na rozbor. Účinky drog, ak by ich aj mal žiak porovnávať, hoci úloha to od neho nevyžaduje, by musel poznať tiež z pamäti. Ale ak by sme žiakovi predložili text s popisom účinkov viacerých drog bez uvedenia ich názvov, a spýtali by sme sa, či sa v článku opisovali aj účinky niektorej nelegálnej drogy, (poprípade aj ktorej) mohlo by už ísť o analýzu.

Naopak, ak zaradíme do 1. úrovne úlohu, v ktorej si žiaci dokážu pomocou rozboru (4. úroveň BT) predloženého východiskového textu (obrázka a pod.) odvodiť správnu odpoveď, úloha nemeria to, čo si myslíme, je zaradená nesprávne, a je teda nevalidná.

Napríklad, ak sa od žiakov gymnázia v ŠVP pre biológiu vyžaduje, aby vedeli *prostredníctvom obrazového materiálu opísať stavbu tráviacej sústavy cicavcov*, neznamená to, že musia tráviace orgány iba poznať z pamäti a vedieť ich označiť na obrázku. Poloha a tvar jednotlivých orgánov žiakovi slúžia nielen ako predmet, na ktorom má predviesť svoje vedomosti, ale aj ako materiál na rozbor. Ak by sme teda pre 1. úroveň BT vytvorili takúto úlohu, merali by sme u niektorých žiakov aj schopnosť dedukcie a rozboru, a úloha by tak nebola validná.

Tu treba poznamenať, že v slovenských vzdelávacích programoch sa vo výkonových požiadavkách často používa **pomocné (modálne) sloveso vedieť**, ktoré ale nie vždy predznamenáva 1. úroveň BT – vedieť, poznať, zapamätať si, a netreba si ho zamieňať s činnostným (aktívnym) slovesom. Toto pomocné sloveso označuje v slovenčine všeobecne schopnosť, vie = je niečoho schopný (vie trafiť domov = dokáže trafiť domov). K prvej úrovni sa viažu požiadavky formulované pomocou slovesa vedieť napríklad takto: vie vymenovať, vie vyhľadať, vie zaznamenať, vie popísať, vie lokalizovať a pod. K vyšším

⁵¹ Úlohu sme pre potreby tejto publikácie upravili, nebola vecne správna: v zadaní sa nepýtala na drogu, ktorá je nelegálna na Slovensku, ale všeobecne. Je pritom známe, že marihuana je legalizovaná vo viacerých krajinách sveta. Menej známa je tiež skutočnosť, že alkohol je tiež nelegálna droga, a to až v 16-tich krajinách sveta. Podobne distraktor A sme zmenili z cigariet na tabak, pretože neuvádzal žiadnu konkrétnu drogu – vo forme cigariet je možné užívať rôzne druhy drog. Tabak je zas úplne ilegálny v Bhutáne.

⁵² e-Testovanie. Informačný bulletin. Bratislava : NÚCEM, 2014, s. 5.
<http://www.etest.sk/data/att/132.pdf> (1. 7. 2015)

úrovniam sa zas viažu požiadavky formulované pomocou slovesa vedieť takto: vie uviesť príklady = *illustrate* (najmenej 2. úroveň); vie vysvetliť, vie použiť, vie spracovať, vie analyzovať, vie posudzovať, vie naplánovať atď.

Testy zložené prevažne z úloh na overovanie vedomostí môžu mať *nízku reliabilitu*. Spôsobuje to známy fakt, ktorý spomína vo svojej knihe o testovaní už J. Čečetka, že žiaci nevedia v množstve vedomostí rozlíšiť dôležité od menej dôležitých vedomostí, a pod vplyvom rôznych faktorov si každý žiak zapamätá iné vedomosti viac, a iné menej.

2. úroveň: Rozumieť (RBT); Chápanie (BT)

• klásť otázky • citovať, odvolať sa na zdroj informácie • prepojiť, **spojiť (na základe spoločných znakov alebo súvislosti)** • previesť (do inej sústavy), zmeniť formu alebo funkciu • formálne prezrieť, posúdiť s úmyslom zmeniť • nápadito rozviesť tému, diskutovať • rozšíriť (napr. opis) • **uviesť príklady** • **zaradiť do skupín, zoskupiť** • ukázať na niečo • prerozprávať, vyjadriť inými slovami • **vybrať zástupný prvok** • skúmať • znovu nastoliť • obkresliť • zmeniť • preložiť • **nájsť jednoduchý protiklad (2)** • vidieť súvislosť (2) • zhruba odhadnúť (2) • odhadnúť význam (2) • **pochopiť zo slov a údajov (2)** • porovnať s odlišným (2) • ukázať na príklade (2) • všímať si s prihliadnutím na kontext (2) • nanovo napísať (2) • **rozlíšiť zjavné rozdiely (2)** • priradiť časť alebo prvok k správne celku alebo k štruktúre (2) • uvedomiť si (2) • **zaradiť do širšej skupiny alebo triedy (2)** • informovať (2) • ukázať na niečo na základe porozumenia (2) • **rozumieť pokynom (2)** • vidieť rozdielnosť (2) • **vybrať niečo implicitné (2)** • nájsť zástupný znak, symbol prvku alebo vzťahu, vyjadriť inak (2) • **uviesť príklad na niečo (2)** • **chápať význam celku (2)** • **objasniť niečo (2)** • rozumieť názoru alebo rozhodnutiu (2) • podrobne opísať (2) • vzťahnuť na väčší súbor (2) • **stručne vyjadriť podstatu (2)**⁵³

Do tejto úrovne patria prevažne úlohy, pri ktorých žiak vo svojej odpovedi uvedie paralelu na význam, ktorý pochopil, a tým dokáže svoju schopnosť. Môže však uviesť aj protiklad, konkrétny príklad, uviesť dominantnú črtu bez toho, aby zovšeobecňoval, objasnením dokáže, že rozumie danej veci, alebo zaradí vec do skupiny, čím preukáže, že rozumie povahe veci a vidí spoločné znaky danej veci s ostatnými podobnými vecami. Na testovanie porozumenia sú veľmi vhodné východiskové texty, ktorých význam je utajený pomocou alegórie, texty, ktorých význam je citovo zafarbený (expresívny), osobne motivovaný (subjektívny), texty, v ktorých sa používajú hyperboly (zveličenie, nadsázka) či iné persuzívne postupy atď. (fejtóny, satiry, karikatúry, aforizmy, vtipy, príslovia, bájky, recenzie, komentáre, glosy, pamflety atď.).

Naopak, nepatria sem úlohy na vyhľadanie informácií, ktoré sú vo východiskovom texte uvedené explicitne – to je tzv. bazálna alebo základná gramotnosť (gramotnosť v pôvodnom význame, nie kľúčová kompetencia). Ak testujeme u žiakov porozumenie, musia nám, takpovediac, predložiť dôkaz toho, že porozumeli – najlepšie uvedením odpovede v otvorenej úlohe alebo pri zložitejších úlohách aj výberom odpovede z uvedených možností (v takom prípade sa musia spomenuté príklady, objasnenia, zástupné prvky a pod. nachádzať v možnostiach odpovede).

Nepatria sem ani úlohy, pri ktorých v teste od žiakov vyžadujeme, aby odpovedali na základe porozumenia nejakému problému alebo východiskovému textu, s ktorým sa stretli v minulosti (mimo testu, ktorý riešia) a úlohu postavíme na zisťovacej otázke bez východiskového textu. Napríklad, keď sa spýtame na dôvody vypuknutia Veľkej francúzskej revolúcie. V takom prípade totiž znova meriame

⁵³ **Understand:** ask • cite • associate • convert • review • discuss • extend • give examples • group • indicate • paraphrase • re-present • research • restate • trace • transform • translate • contrast (2, 4) • relate (2, 3) • estimate (2, 4, 5) • predict (2, 3, 5) • infer (2, 4, 6) • compare (2, 4, 5) • demonstrate (2, 3) • observe (1, 2) • rewrite (2, 6) • distinguish (2, 4, 5) • identify (1, 2) • discover (1, 2, 3) • classify (2, 4) • report (2, 3, 6) • show (2, 3) • order (2, 4, 5) • differentiate (2, 4) • select (1, 2, 4, 5) • express (2, 6) • illustrate (2, 3, 4) • interpret (2, 3) • explain (2, 3, 4) • judge (2, 3, 5) • describe (1, 2) • generalize (2, 6) • summarize (2, 5)

zapamätanú vedomosť, hoci možno aj získanú vyšším kognitívnym procesom. Ak má úloha merať porozumenie, musí u žiaka počas riešenia úlohy prebehnúť tento proces, nie vybavenie si vedomosti (získanej a upevnenej akokoľvek) z pamäti. Inak riskujeme nevalidnosť úlohy – keďže pri externom meraní nevieme, ako žiak danú vedomosť nadobudol, či u neho proces porozumenia prebehol alebo učiteľ použil výklad, a žiak vedomosť nadobudol procesom zapamätania. Samozrejme, ak učiteľ zostavuje učiteľský test pre svoju triedu a vie, že môže danú vedomosť od žiakov vyžadovať iba na základe ich porozumenia danej veci, môže takúto úlohu postaviť aj bez východiskového textu.

3. úroveň: Použití (RBT); Používanie (BT)

- vykonať • použiť • artikulovať • **vykonať podľa zvyčajného postupu** • skompletizovať • **vypočítať**
- rozhodnúť • dramatizovať • **upotrebiť, využiť (metódu, pravidlo)** • ustanoviť • zameniť, zmeniť
- vytvoriť graf, mapu, plán • klásť otázky • obratne zmeniť, upraviť, presunúť • pracovať (podľa niečoho)
- vyjadriť farbou, zafarbiť • naplánovať • naskicovať • naučiť sa postupovať • **urobiť prevod • uplatniť**
- spočítať (3) • zostrojiť podľa predlohy (3) • **dokončiť podľa návodu** (3) • odhaliť, nájsť (3) • položiť otázku, **preveriť niečo podľa inštrukcií** (3) • vykonať pokus podľa inštrukcií (3) • dokázať **vykonať nejakú operáciu, činnosť alebo postup podľa inštrukcií** (3) • zhromaždiť informácie (3) • vysvetliť postup (3) • **vybrať niečo podľa parametrov** (3) • vytvoriť ukážku niečoho podľa inštrukcií (3) • vysvetliť, podať niečo v zrozumiteľnejšej forme (3) • dospieť k rozhodnutiu podľa inštrukcií (3) • upraviť podľa inštrukcií (3) • predvídať pokračovanie procesu, činnosti (3) • predpripraviť (3) • vyrobiť podľa návodu (3) • **zaznamenať niečo konvenčným alebo predpísaným spôsobom** (3) • **usúvzťažniť na základe pravidiel** (3) • zostaviť informáciu podľa inštrukcií (3) • ukázať na niečo konkrétne na základe inštrukcií (3) • napodobniť niečo na základe inštrukcií (3) • vyriešiť problém (3) • napísať na základe inštrukcií (3) • pracovať so zoznamom (3)⁵⁴

Do úrovne 3.1 patria najmä úlohy, v ktorých žiak skôr ako odpovie, musí vykonať nejakú operáciu v známom type úloh (*familiar task*), v tzv. typickom školskom kontexte. V tomto type úloh už iba s otázkou nevystačíme. Pre tento typ úloh musíme naformulovať úlohovú situáciu, prípadne dodať východiskový text a potrebné informácie, a až potom môžeme formulovať výzvu k činnosti alebo položiť otázku. Testujú sa tu zručnosti, nie vedomosti. Nesprávne sa sem zaraďujú úlohy, v ktorých žiak takpovediac „použije“ svoje zapamätané vedomosti na to, aby vybral správnu odpoveď – tu ide iba o vybavovanie si z pamäti.

V úrovni 3.2 (implementovať, zavádzať) ide o to, aby žiak zručnosť, ktorou disponuje, vedel upotrebiť v situácii, ktorá je pre žiakov neobvyklá (*unfamiliar task*), ktorá nie je typicky školská. Napríklad, kým v úrovni 3.1 (vykonávať, spracovať) žiak dokáže vypočítať obsah rovinných útvarov (štvorec, obdĺžnik a pod.), v úrovni 3.2 už musí byť schopný vypočítať, koľko metrov štvorcových krytiny potrebujeme na pokrytie podlahy v izbe, ktorej pôdorys má napríklad aj výbežky, a zaokrúhliť výsledok nahor na celé balenia krytiny. **Neobvyklou situáciou** tu môžeme rozumieť každú situáciu, ktorá pravdepodobne **nebola použitá počas vyučovania**, alebo s takou situáciou aspoň nie je úplne identická.⁵⁵

Aby sme v týchto úlohách vylúčili faktor nepamätania si čiastkových vedomostí, môžeme žiakovi poskytnúť vzorce, ale aj iné informácie, ktoré sú potrebné na riešenie úlohy. Môže ísť o definície, vzorce, číselné hodnoty, heslá z encyklopédie, poučky, návody, inštrukcie, fakty, ukážky a pod. Ak žiakom tieto

⁵⁴ **Apply:** administer • apply • articulate • practice • complete • compute • determine • dramatize • employ • establish • change • chart • interview • manipulate • operate • paint • schedule • sketch • teach act • transfer • use • calculate (3, 4) • construct (3, 6) • develop (3, 6) • discover (1, 2, 3) • examine (1, 3) • experiment (3, 4) • demonstrate (2, 3) • collect (3, 6) • explain (2, 3, 4) • choose (3, 5, 6) • illustrate (2, 3, 4) • interpret (2, 3) • judge (2, 3, 5) • modify (3, 6) • predict (2, 3, 5) • prepare (3, 6) • produce (3, 6) • record (1, 3) • relate (2, 3) • report (2, 3, 6) • show (2, 3) • simulate (3, 6) • solve (3, 6) • write (3, 6) • list (1, 3)

⁵⁵ Arreola, A. Raoul: *Writing Learning Objectives*. Memphis : The University of Tennessee, 1998, s. 6.
http://www.uwo.ca/tsc/graduate_student_programs/pdf/LearningObjectivesArreola.pdf (1. 7. 2015)

vedomosti v úlohách na vyššie úrovne neposkytneme, nebudeme vedieť, či žiaci zlyhali v týchto úlohách preto, že nedokázali vykonať požadovanú operáciu, alebo preto, že si iba nepamätali čiastkovú vedomosť – otáznu potom ostane validita takejto úlohy, pretože nebude merať presne to, čo chceme. Ak by takáto úloha aj nebola osamotená a nadväzovala by na úlohy, v ktorých by sme separátne overili, či žiak tieto čiastkové vedomosti má, bola by takáto úloha pre žiaka, ktorý nemá potrebné vedomosti, neriešiteľná. Pritom žiak môže mať požadované zručnosti, ale my ich nenameriame...

4. úroveň: Rozobrať (RBT); Rozbor (BT)

• vyzdvihnúť (vlastnosť, rys) • analyzovať • určiť poradie pre riešenie problémov podľa ich relatívneho významu • **rozložiť** • zaradiť, triediť • prepojiť, spojiť • usúvzťažniť • odvodiť, vyvodiť • zostaviť diagram • rozčleniť • oddeliť, deliť • zamerať sa • zachytiť hlavné rysy • poukázať na niečo, naznačiť • položiť otázku • odčleniť • rozdeliť • **preskúmať** • stanoviť hodnotu (4) • **spočítať niektoré prvky** (4) • **roztriediť podľa podrobnej klasifikácie** (4) • porovnať s podobným (4) • dospieť k záveru (4) • odhaliť negatívne stránky niečoho (4) • **vybrať časti z celku na základe viacerých kritérií** (4) • pripomenkovať (4) • navrhnúť (4) • **nájsť rozdiely** (4) • **nájsť rozlišujúci prvok** (4) • **rozoznať ťažko rozoznatelné** (4) • stanoviť pravdepodobnú hodnotu (4) • číselne vyjadriť počet alebo hodnotu, vytvoriť si predstavu o počte alebo o hodnote (4) • pokusom overiť (4) • popísať prvky a ich vzťahy (4) • **nájsť príklad niečoho v danom materiáli** (4) • **dedukovať z dôkazov** (4) • **nájsť pravidlo usporiadania** (4) • systematicky usporiadať prvky (4) • plánovite postupovať (4) • urobiť skúšku správnosti (4)⁵⁶

Do tejto úrovne patria úlohy, ktoré nie je možné postaviť bez východiskového textu. Žiak musí dostať materiál na rozbor, na preskúmanie.⁵⁷ Môže ísť aj o slovo, vetu, či obrázok atď. Podobne, ako to je pri porozumení a používaní, ani tu nemôžeme od žiaka chcieť, aby odpovedal na základe zapamätaných vedomostí, ktoré získal v minulosti (hoci aj analýzou) – išlo by o proces vybavovania si z pamäte. Nemôžeme ani chcieť, aby analyzoval vedomosti, o ktorých predpokladáme, že si ich pamätá. Tak by sme totiž zase merali aj schopnosť vybaviť si vedomosti z pamäte, a ak by si žiak nepamätal okolnosti, prvky a vzťahy, ktoré má analyzovať, alebo by si ich nepamätal dokonale, nemohol by preukázať, že proces rozboru ovláda – takáto úloha by nebola validná.

Napríklad, predložíme žiakovi ukážku statického opisu nejakej veci a v zadaní úlohy budeme vyžadovať, aby žiak zistil, či ide o statický alebo dynamický opis a vybral odpoveď aj so správnym zdôvodnením. Správna odpoveď by bola *statický opis* so zdôvodnením, že v texte sa vyskytujú prevažne podstatné a prídavné mená. Vybrať túto odpoveď bez analýzy textu (alebo hádania) nie je možné. Ak by sme úlohu otočili, postup analýzy by sme vložili do správnej odpovede obklopenej distraktormi a spýtali by sme sa, ktorý z nasledovných postupov je vhodný na zistenie statickosti alebo dynamickosti opisu v ukážke, merali by sme iba schopnosť vybaviť si alebo znovu spoznať vedomosti o postupoch (1. úroveň).

⁵⁶ **Analyze:** advertise • analyze • prioritize • break down • categorize • connect • correlate • deduce • diagram • dissect • divide • focus • outline • point out • question • separate • subdivide • survey • appraise (4, 5) • calculate (3, 4) • classify (2, 4) • compare (2, 4, 5) • conclude (4, 5) • contrast (2, 4) • select (1, 2, 4, 5) • criticize (4, 5) • devise (4, 6) • differentiate (2, 4) • discriminate (4, 5) • distinguish (2, 4, 5) • estimate (2, 4, 5) • evaluate (4, 5) • experiment (3, 4) • explain (2, 3, 4) • illustrate (2, 3, 4) • infer (2, 4, 6) • order (2, 4, 5) • organize (4, 6) • plan (4, 6) • test (4, 5, 6)

⁵⁷ Pozri napríklad úlohu na rozpoznávanie manipulatívnej komunikácie na s. 30. (správna odpoveď je možnosť E.): *In: Opatření ministra školství, mládeže a tělovýchovy, kterým se mění Rámcový vzdělávací program pro základní vzdělávání*. Příloha 2: Standardy pro základní vzdělávání. Český jazyk a literatura. 2013. http://www.msmt.cz/file/30521_1_1 (1. 7. 2015)

5. úroveň: Hodnotiť (RBT); Hodnotenie (BT)

• **dokazovať, argumentovať** • určiť, stanoviť • **zohľadniť, uvážiť** • presvedčiť • napísať kritiku, recenzovať • debatovať • rozhodnúť • presvedčiť (pomocou dôkazov, dôvodov) • vyjadriť svoje názory • **nájsť chyby** • vyhľadať, nájsť • vyjadriť iným spôsobom, preformulovať • **navrhnuť niečo vhodné pre daný účel** • **posúdiť mieru, význam, vplyv, hodnotu, proporcie** • zaradiť (do poradia) • **určiť hodnotu podľa normy** • oznámkovať • obhajovať (názor) • priradiť body • uvážiť • stanoviť kvalitu (5) • porovnať s predpisom, s normou (5) • **urobiť záver na základe kritérií** (5) • kriticky komentovať (5) • vnímať podstatu rozdielu (5) • identifikovať charakteristický znak (5) • **vybrať niečo najvhodnejšie** (5) • oceniť hodnotu niečoho (5) • zhodnotiť (5) • **vybrať niečo podstatné** (5) • **rozhodnúť na základe kritérií** (5) • **dokázať, deklarováť správnosť niečoho** (5) • **usporiadať niečo podľa kritérií** (5) • predpovedať dopad, efekt (5) • zhrnúť (5) • podoprieť tvrdenie, odporučiť niečo (5) • **preveriť súlad s kritériami** (5)⁵⁸

Úlohy, ktoré patria do tejto úrovne, najčastejšie vyžadujú od žiaka, aby pre nejaké rozhodnutie vybral správny **dôkaz, argument** alebo **zdôvodnenie**. Argument je presvedčivý fakt potvrdzujúci istú mienku, názor alebo tvrdenie. Netreba si tu zamieňať tvrdenie s faktom. **Fakt** je zistená, overená alebo daná jednotlivá skutočnosť, niečo skutočné, objektívne jestvujúce, pravdivý poznatok, udalosť, údaj, kým **tvrdenie** je slovné vyjadrenie myšlienky s cieľom uistiť niekoho o niečom, dokázať niečo alebo niekoho presvedčiť. Tvrdenie, na rozdiel od faktu, nemusí byť vždy pravdivé. Dôležitosť schopnosti rozlišovať medzi faktom, pravdivým tvrdením a obyčajným názorom prezentuje úloha z medzinárodného výskumu občianskeho vzdelávania IEA ICCS 2009:⁵⁹

V ktorej z možností je uvedený názor?

- Každá krajina má svoju vlastnú vlajku a štátnu hymnu.
- Organizácia spojených národov má svoju vlastnú vlajku, hoci to nie je krajina.
- Každý by si mal ctiť vlajku a hymnu svojej krajiny.
- Vlajky často nachádzame na lodiach patriacich nejakému národu.

Táto úloha by patrila do 2. úrovne BT, meria schopnosť pochopiť hlavnú myšlienku výrokov, rozumiť rozdielu vo formulácii týchto výrokov a na základe toho ich priradiť k skupine tvrdení alebo názorov. (Keby šlo o ťažšie rozoznateľné výroky, mohlo by ísť aj o úroveň 4.)

⁵⁸ **Evaluate:** *argue • assess • consider • convince • critique • debate • decide • persuade • editorialize • errors • find • reframe • recommend • measure • rank • rate • grade • defend • score • weigh • appraise (4, 5) • compare (2, 4, 5) • conclude (4, 5) • criticize (4, 5) • discriminate (4, 5) • distinguish (2, 4, 5) • select (1, 2, 4, 5) • estimate (2, 4, 5) • evaluate (4, 5) • choose (3, 5, 6) • judge (2, 3, 5) • justify (5, 6) • order (2, 4, 5) • predict (2, 3, 5) • summarize (2, 5) • support (5, 6) • test (4, 5, 6)*

⁵⁹ ICCS 2009. Uvoľnené úlohy.

http://www.nucem.sk/documents//27/medzinarodne_merania/iccs/ukazky_testovych_uloh/iccs/Ukážky_testovych_uloh.pdf (1. 7. 2015)

Príkladom na 5. úroveň, konkrétne na zdôvodňovanie sú nasledovné úlohy z výskumu ICCS 2009:

V krajine K. je menšinová skupina ľudí, ktorých jazyk je odlišný od oficiálneho (štátneho) jazyka krajiny. Táto skupina má svoje vlastné školy, kde sa deti vyučujú a učia iba vo svojom vlastnom tradičnom jazyku.

Vláda krajiny K. rozhodne, že všetky školy musia vyučovať všetky deti iba v oficiálnom (štátnom) jazyku krajiny. Vláda urobí toto rozhodnutie preto, lebo sa nazdáva, že to pomôže deťom menšinovej skupiny.

Ktorý z nasledujúcich dôvodov najlepšie podporuje rozhodnutie vlády?

- A. Spôsobí to, že deti doma prestanú hovoriť svojím tradičným jazykom.
- B. Škola tak bude pre deti zaujímavejšia.
- C. Deti dostanú väčšiu šancu plne sa začleniť do širšieho spoločenstva.*
- D. Pomôže to deťom ľahšie sa naučiť svoj tradičný jazyk doma.

Čo je najlepší dôvod proti rozhodnutiu vlády?

- A. Vyučovacie predmety by nemali byť ovplyvňované vládami.
- B. Vlády by mali akceptovať, že je potrebný viac ako jeden oficiálny jazyk.
- C. Vlády sú zodpovedné za ochranu kultúr menšinových skupín.*
- D. Deti menšinových skupín sa môžu sťažovať na vyučovanie oficiálneho jazyka.

* Správna odpoveď.

Do tejto úrovne nepatria úlohy, ktoré používajú sloveso *hodnotiť* nenáležite, najčastejšie v súvislosti s informáciami explicitne uvedenými v texte, pri ktorých má žiak „zhodnotiť“ ich pravdivosť alebo nepravdivosť. Napríklad, v ukážke je uvedený dátum a úloha vyzýva žiaka k činnosti nasledovne: „Zhodnot' pravdivosť výroku: „Podujatie sa bude konať vtedy a vtedy“. Ide tu pritom o vyhľadanie a kontrolu explicitne uvedenej informácie (1. úroveň).

6. úroveň: Tvoríť (RBT); Spájanie (BT)

• predvídať, predpokladať • **usporiadať** • zostaviť, zložiť • vytvoriť teóriu alebo domnienku bez dôkazov • spolupracovať • **skombinovať** • zozbierať • robiť alebo plánovať niečo podľa presnej predstavy • prerobiť, preskupiť, inak usporiadať • zložiť, tvoriť • tvoriť, generovať, produkovať • **zjednodušiť**, uľahčiť • **stručne vyjadriť, sformulovať** • navrhnúť hypotézu • utvoriť si predstavu, obraz • združiť, zlúčiť, vytvoriť funkčný celok • zasiahnuť • vymyslieť, vynájsť • vykonať, vyrobiť • **rozumne použiť** • vyjednávať, **nájsť cestu (cez prekážky)** • **určiť vznik, počiatok** • dať návrh • preorganizovať, inak usporiadať • **prehodnotiť a zmeniť na základe dôkazov** • **vžiť sa do roly** • **vytvoriť schému** • zostaviť podľa plánu • **nahradiť** • potvrdiť, schváliť • vytvoriť niečo (6) • **rozvinúť, zmysluplne dokončiť** (6) • upraviť plán, **zjednodušiť, vylepšiť niečo** (6) • sformulovať myšlienku (6) • **urobiť záver z dôkazov a úsudkov** (6) • **zvoliť niečo najvhodnejšie** (6) • **vytvoriť z častí zmysluplný celok**, súbor (6) • **zovšeobecniť** (6) • **dať niečo do poriadku** (6) • invenčne upraviť na niečo iné (6) • skoordinať aktivity (6) • **naplávať** (6) • **spraviť niečo ľahšie použiteľným** (6) • vyprodukovať niečo nové (6) • **sformulovať správu** (6) • vylepšiť prepísaním (6) • tvorivo niečo napodobniť (6) • **vysvetliť záhadu** (6) • sformulovať odporúčenie (6) • **odhaliť skryté kvality niečoho** (6) • tvorivo niečo napísať (6)⁶⁰

⁶⁰ **Create:** anticipate • arrange • assemble • speculate • collaborate • combine • compile • design • rearrange • compose • create • facilitate • formulate • hypothesize • imagine • integrate • intervene • invent • make • manage • negotiate • originate • propose • reorganize • revise • role-play • schematize • structure • substitute • validate • construct (3, 6) • develop (3, 6) • devise (4, 6) • express (2, 6) • infer (2, 4, 6) • choose (3, 5, 6) • collect (3, 6) • generalize (2, 6) • justify (5, 6) • modify (3, 6) • organize (4, 6) • plan (4,

Vôbec nie je správne tvrdiť, že sem patria iba tzv. tvorivé úlohy, ani nie je správne vyhýbať sa tejto úrovni pri tvorbe úloh do didaktických testov. Ako vidíme zo zvýraznených slovies, úlohy, ktoré patria do tejto úrovne, vyžadujú od žiaka činnosti, ktorých výsledkom je niečo jednoduchšie pochopiteľné a pritom stručné, niečo opravené alebo vylepšené zjednodušením. Žiaci dokážu vyriešiť úlohy tak, že sa vžijú do predpísanej roly, dokážu v možnostiach nájsť pre rôzne veci rozumné využitie, dokážu z rôznych príčin a následkov určiť počiatok problému, dokážu vytvoriť schému zložitého procesu, zefektívniť proces zmenou a zlúčením alebo odstránením krokov, zefektívniť plán alebo rozvrh práce, dokážu doplniť do procesu chýbajúce prvky, dokončiť cyklus a pod.

Tento typ úloh sa samozrejme nielenže viaže na zväčša zložené alebo kombinované východiskové texty, ale v didaktickom teste musí byť žiakova odpoveď zväčša štruktúrovaná, čo v tomto prípade vyžaduje väčší priestor pre možnosti odpovede, do ktorých sa musia zmestiť niekedy až celé riešenia. Inou možnosťou sú interaktívne formy úloh v elektronickom testovacom prostredí,⁶¹ ktoré systém e-Test zatiaľ neumožňuje.

2.3. Druhy testov

2.3.1. Podľa meranej charakteristiky výkonu

A. Testy rýchlosti

Testy rýchlosti overujú rýchlosť riešenia určitého (jedného) typu úloh približne rovnakej obťažnosti. Predpokladom správnej aplikácie testu rýchlosti je schopnosť všetkých žiakov riešiť daný typ úloh. Počet úloh musí byť taký, aby nebolo možné vyriešiť všetky úlohy v časovom limite testovania.

Časový limit sa používa ako merateľná premenná napr. aj pri psychologických testoch inteligencie, v ktorých sú úlohy rozdelené do blokov. Každý blok má svoj časový limit a obsahuje iba úlohy jedného druhu a približne rovnakej obťažnosti. Čím je úloh viac, tým je, samozrejme, presnosť takéhoto merania väčšia.

Takýto test nemeria „hĺbku“ vedomostí alebo zručností. Test meria, koľkokrát je žiak schopný bezchybne vykonať určitú operáciu za určitý čas.

„Například, ak dáme žiakom riešiť jedinú úlohu, můžeme ich rozdelit' len do dvou skupin: vyhovel – nevyhovel. Nepomůže, naopak, dať im například pět matematických úloh a klasifikovat... V prvom rade záleží tu na povahe úloh samotných. Ak im dáme všetky úlohy len jedného druhu a dlá odhadu rovnakej (približne) obťažnosti, můžeme z výsledkov vlastne usúdiť len: žiaci, ktorí neurobili ani jednu úlohu, neovládajú tento spôsob rátania, ale dlá počtu vyriešených úloh nemôžeme ešte dobre usudzovať to, ako, do akej miery tento úkon ovládajú, lebo jedine počet riešených úloh môže byť svedectvom o rýchlosti práce jednotlivých žiakov, čo ale v každom prípade nie je ešte znakom, podľa ktorého by sme mohli posudzovať samotné „vedomosti“ žiaka.“⁶²

6) • prepare (3, 6) • produce (3, 6) • report (2, 3, 6) • rewrite(2, 6) • simulate (3, 6) • solve (3, 6) • support (5, 6) • test (4, 5, 6) • write (3, 6)

⁶¹ Pozri napríklad interaktívnu úlohu Climate control z modulu Riešenie problémov PISA: <http://www.oecd.org/pisa/test/testquestions/question3/> (1. 7. 2015)

⁶² Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 10 – 11.

B. Testy úrovne

a) Jednoúrovňové testy

Jednoúrovňový test (na rozdiel od testu rýchlosti) meria mieru jednej schopnosti, pretože väčšinou obsahuje úlohy rôznej náročnosti zamerané na jeden konkrétny poznávací proces (napr. testy zamerané iba na rozbor a pod.). Takýto test môže byť ešte viazaný na užšie vymedzenú oblasť, tematický celok a pod. Takýto test umožňuje hlbšie a detailnejšie analyzovať konkrétne vymedzené výkony žiakov v sledovanej oblasti.

Jednoúrovňový test ale môže byť aj nadpredmetový. Priestor pre nadpredmetovosť vytvára práve obmedzenie sa na jednu úroveň (jeden tem. celok, jeden kogn. proces a pod.). V takomto teste môžeme merať vyššiu kognitívnu úroveň a zisťovať, do akej miery sú žiaci schopní prepájať či zovšeobecňovať vedomosti z rôznych predmetov.

b) Viacúrovňové testy

Vo viacúrovňových testoch bývajú obvykle zastúpené úlohy na meranie všetkých úrovní stanovenej škály. Je však možné zostaviť aj samostatný test na meranie napríklad nižších alebo vyšších úrovní škály.

Na triedenie úloh do rôznych úrovní sa vo vzdelávaní obvykle používajú na to určené taxonómie (Bloomova, Niemiarkova), stupeň obťažnosti alebo pri testovaní v cudzích jazykoch napríklad Spoločný európsky referenčný rámec pre jazyky (SERR).

Počet úloh merajúci jednotlivé úrovne môže byť rôzny, nemal by však byť náhodný. Napríklad, ak v danom predmete nie sú dôležité zapamätané konkrétne fakty, lebo to neurčuje výkonový štandard, môže zostavovateľ testu dať úloh na 1. úroveň Bloomovej taxonómie menej, a naopak, na testovanie cieľov a výkonov, ktoré sú v danom predmete dôležité a kľúčové, zaradiť úloh viac.

Dôležité je tu však zachovať minimálny počet úloh pre každú úroveň. Keďže za minimálny počet úloh v teste sa považuje 10 úloh, pri minimálnom počte úrovní (2) vo viacúrovňovom teste je to 5 úloh pre každú úroveň. Aj tu platí, že čím je úloh pre každú úroveň viac, tým je presnosť takéhoto merania väčšia. Maximálny počet úloh je obmedzený časovým limitom a schopnosťou žiakov sústrediť sa na riešenie úloh.

2.3.2. Podľa typu učenia

Toto rozdelenie navrhol B. S. Bloom. Pre každú oblasť navrhol samostatnú taxonómiu cieľov. Najznámejšia je jeho taxonómia cieľov vzdelávania pre kognitívnu oblasť, ktorá je členená na úrovne podľa logicko-poznávacích procesov.

A. Testy na zisťovanie výsledkov kognitívneho učenia

(vedomosti, zručnosti, kompetencie a pod.)

B. Dotazníky na zisťovanie výsledkov afektívneho učenia

(postoje, hodnoty a pod.)

C. Testy na zisťovanie výsledkov psychomotorického učenia

(napr. test písania na počítači, praktické testy v autoškole (jazda) a pod.)

2.3.3. Podľa miery špecifickosti učenia zisťovaného testom

A. Testy výsledkov vzdelávania

Na meranie výsledkov vzdelávania sa používajú **didaktické testy**. Používajú sa aj v priebehu, aj na konci vzdelávacieho celku alebo stupňa. Za výsledky vzdelávania považujeme merateľné, pedagogicky relevantné konštrukty (vedomosti, zručnosti, kompetencie, kognitívne operácie a pod.).

Podľa Európskeho kvalifikačného rámca (EKR)⁶³ výsledky vzdelávania vyjadrujú to, čo žiak vie, chápe a je schopný urobiť pri ukončení procesu vzdelávania. Výsledky vzdelávania sú definované z hľadiska vedomostí, zručností a kompetencií.

- **Vedomosti** sú výsledkom osvojenia si informácií prostredníctvom učenia. Vedomosti sú súbor faktov, zásad, teórií a postupov, ktoré sa vzťahujú na oblasť práce a štúdia. V kontexte EKR sú vedomosti opísané ako teoretické a/alebo faktické (know-what).
- **Zručnosti** sú spôsobilosti uplatňovať vedomosti a využívať know-how na splnenie úloh a riešenie problémov. V kontexte EKR sú zručnosti opísané ako kognitívne (vrátane využívania logického, intuitívneho a kreatívneho myslenia) alebo praktické (vyžadujúce manuálnu zručnosť a používanie metód, materiálov, prostriedkov a nástrojov).
- **Kompetencie** sú preukázané spôsobilosti⁶⁴ použiť vedomosti, zručnosti a osobné, sociálne alebo metodické schopnosti v pracovných alebo študijných situáciách a v odbornom a osobnom rozvoji.

B. Testy študijných predpokladov

Testy študijných predpokladov (aptitude tests) by sa mali používať najmä pri prijímaní študentov na vyšší typ školy. Bohužiaľ, testy, ktoré sa na školách na tento účel používajú, sa príliš nelíšia od testov výsledkov vzdelávania.⁶⁵ Test všeobecných študijných predpokladov testuje základné zručnosti a schopnosti žiakov, potrebné pre ich ďalšie štúdium. Cieľom testu nie je testovanie vedomostí. Napríklad test študijných predpokladov spoločnosti SCIO je zložený zo štyroch špecifických častí (verbálnej, logickej, argumentačnej a kvantitatívnej), z ktorých každá testuje iný typ myslenia.⁶⁶

Na rozdiel od vedomostného testu sa pri týchto testoch neposudzuje obsahová, ale prediktívna validita (pozri kapitolu 2.4.1. Validita). Výsledky takéhoto testu by mali korelovať s výkonom žiaka v prvom roku štúdia na škole, na ktorú bol na základe testu prijatý (porovnávacím kritériom je buď miera produktivity alebo nameraný výkon). Táto korelácia však u prediktívnych testov nebýva veľmi vysoká. Korelácia okolo 0,35 sa tu hodnotí už ako pomerne významná.⁶⁷ Porovnávať však je možné iba výsledky testov, ktoré merajú rovnaký konštrukt a sú pri vyhodnocovaní na takéto porovnanie prispôsobené – normalizácia skóre. Nízku koreláciu výsledkov študentov vysokých škôl a testov študijných predpokladov, ktoré testujú najmä predmetové a kľúčové kompetencie, môže spôsobovať aj obvyklá prílišná orientácia výučby a skúšania v prvom roku štúdia na vysokej škole na zapamätané vedomosti.

⁶³ Generálne riaditeľstvo pre vzdelávanie a kultúru Európskej komisie: *Európsky kvalifikačný rámec pre celoživotné vzdelávanie*. Luxemburg : Úrad pre vydávanie úradných publikácií Európskych spoločenstiev, 2009, 20 s. ISBN 978-92-79-08489-8. http://ec.europa.eu/ploteus/sites/eac-efq/files/broch_sk.pdf (1. 4. 2015)

⁶⁴ Aj Štátny vzdelávací program uvádza k pojmu kľúčová kompetencia ako *ekvivalent* pojem spôsobilosť. (Nejde o *synonymum*, ako chybné uvádza ŠVP, pretože synonymické vzťahy fungujú iba vnútri jedného jazyka, pozn. T. K.)

⁶⁵ Zásady tvorby didaktického testu. Fakulta informatiky a manažmentu Univerzity v Hradci Králové. http://fim.uhk.cz/oliva/tvorba_vedeni/rekap-www/m01/m01-02.html (1. 4. 2015)

⁶⁶ SCIO: *Ukážka testu všeobecných študijných predpokladov*. <https://www.scio.cz/download/VSP-kratky.pdf> (1. 4. 2015)

⁶⁷ Predictive validity. http://en.wikipedia.org/wiki/Predictive_validity (1. 4. 2015)

2.3.4. Podľa fázy učebného procesu

A. Vstupné testy

Vstupný test je zadávaný žiakom na začiatku štúdia na určitom stupni vzdelávania. Vstupným testom môžeme zisťovať vedomosti, zručnosti a kompetencie **získané pred začiatkom štúdia** na aktuálnom stupni vzdelávania.

Účelom vstupných testov je potreba **rozdeliť** (diferencovať) žiakov na skupiny (tzv. zaradovacie testy) podľa úrovne vedomostí, schopností alebo kompetencií, preto bývajú zároveň **viacúrovňové** (Blom, SERR pri cudzích jazykoch a pod.).

B. Priebežné testy

Priebežným testom môžeme merať výsledky vzdelávania v istom obsahovo vymedzenom celku. Priebežný test poskytuje spätnú väzbu v okamihu, kedy je ešte možné zlepšiť výkon žiakov alebo učiteľa, preto takýmto testom hovoríme aj **formatívne** alebo **evalvačné** testy.

Priebežné testy sa používajú najmä na diagnostiku problémov s učením sa a s cieľom odstrániť ich, zefektívniť učebnú činnosť žiakov.⁶⁸ Používajú sa na korekciu učebných postupov, na spätnoväzbové účely alebo na priebežné hodnotenie práce žiakov a učiteľa. Formatívny test takpovediac nastavuje žiakom aj učiteľovi latku.

Nevyhnutným následkom priebežného testu je, aby žiaci alebo učiteľ dostali spätnú väzbu, teda zrozumiteľnú informáciu o tom, ako majú zmeniť svoje učenie, aby dosiahli to, o čo sa usilujú.

Špeciálnym druhom formatívneho testu sú **inšpekčné** (kontrolné) testy. Takýto test obvykle nevykonáva sám učiteľ, ale osoba alebo externá inštitúcia, poverená kontrolou. Môže ísť o test zadaný riaditeľom, zriaďovateľom, Štátnou školskou inšpekciou (ŠSI). Účelom kontroly je zistenie stavu vzdelávacieho procesu, navrhnutie zmien a ich implementácia do vzdelávacieho procesu.

Pri tvorbe formatívneho testu treba rozlišovať medzi počiatočným, aktuálnym a cieľovým stavom vzdelania žiakov. Formatívny test zisťuje **aktuálny** stav.

Na základe informácií získaných analýzou chýb a výsledkov formatívneho testu je nevyhnutné vykonať zmeny, ktoré nasmerujú proces vyučovania k cieľom stanovených v školskom zákone a v ŠVP.

Pri formatívnom hodnotení je vhodné zvážiť výkony žiaka v priebežnom teste aj s jeho výsledkami v testoch z iných predmetov. Je totiž pravdepodobné, že miera schopnosti žiaka sa prejaví aj v iných predmetoch a zároveň je tak možné postrehnúť aj prítomnosť faktorov pozitívne alebo negatívne ovplyvňujúcich učenie sa žiaka.⁶⁹

C. Výstupné testy

Výstupné testy (sumatívne) sú zadávané žiakom na konci štúdia na určitom stupni vzdelávania. Výstupným testom môžeme zisťovať vedomosti, zručnosti a kompetencie **získané počas štúdia** na aktuálnom stupni vzdelávania.

⁶⁸ Tomášková, J.: *Inovatívne spôsoby hodnotenia žiakov v predmete slovenský jazyk a literatúra*. Bratislava : Metodicko-pedagogické centrum, 2015, s. 20. ISBN 978-80-565-0379-9.

⁶⁹ Hrabal, V. (st.) – Hrabal, V. (ml.): *Diagnostika*. Pedagogickopsychologická diagnostika žaka s úvodom do diagnostické aplikácie statistiky. Praha : Univerzita Karlova v Praze, 2002, s. 69. ISBN 80-246-0319-5.

Výsledky výstupných testov môžu okrem toho, že slúžia na záverečnú klasifikáciu žiakov, ovplyvňovať aj budúcu študijnú a profesijnú kariéru žiakov a teda môžu slúžiť na:

- podmieňovanie postupu do vyššieho ročníka – **vyradovacie** testy (CR-testy),
- výber žiakov na vyšší stupeň vzdelávania – **prijímačkové** testy (NR-testy),
- získanie osvedčenia o odbornej spôsobilosti (**akreditačné** testy – majú stanovenú minimálnu hranicu skóre na získanie osvedčenia).

2.3.5. Podľa prípravy testu

A. Neštandardizované testy

Neštandardizované testy sú testy, pri ktorých neboli zrealizované základné kroky prípravy štandardizovaného didaktického testu a ich tvorcovia nepoznajú ich vlastnosti a parametre. Často nemajú ani dopredu pripravenú hodnotiacu škálu. Používajú sa väčšinou na zisťovanie výsledkov vzdelávania za kratšie časové obdobie. Sú to neformálne testy. V teórii sa pre ne používa pomenovanie školské alebo učiteľské testy.

B. Kvázištandardizované testy

Kvázištandardizované testy sú pripravované dôkladnejšie ako neštandardizované, ale v procese prípravy u nich nebývajú z rôznych dôvodov zrealizované všetky kroky prípravy didaktického testu. Príčinou môže byť nedostatok vedomostí o zostavovaní testu, nedostatok času a podmienok na prípravu testu, ale aj podceňovanie prípravy testu.

Ak je vynechaná prípravná fáza, napríklad stanovenie hypotéz (formatívny test), stanovenie výskumnej otázky (výkonový test) alebo podrobná špecifikácia testu, sú úlohy vybrané viac-menej náhodne. Potom je ťažké na základe výsledkov takéhoto testu o žiakoch niečo povedať. Bez dobre urobenej špecifikácie sa do testu môže dostať napríklad priveľa úloh z jednej úrovne Bloomovej taxonómie, napríklad priveľa úloh na 1. úroveň (zapamätať si, vedieť, poznať). Takýto test môže mať veľmi nízku reliabilitu, ak by sme ho chceli administrovať na viac ako jednej triede. Učiteľ totiž môže dávať rôzny dôraz na jednotlivé vedomosti pri výučbe viacerých tried, a tiež platí, že ani žiaci neprikladajú vždy rovnakú dôležitosť všetkým informáciám.

Ak je vynechaná fáza skúšobného overovania testu, obťažnosť testu sa dá iba odhadovať. Odhadovanie obťažnosti je jedna z metód určovania obťažnosti úloh, ale spoľahlivejšou metódou sa stáva až vtedy, keď sa na nej zúčastňuje viacero odborníkov, ktorí si svoje odhady najprv zapíšu, potom si ich porovnajú, a ak medzi nimi nie sú vážne nezhody, spriemerujú ich. Ak sa v odhadoch veľmi líšia, musia o svojich odhadoch viesť odbornú diskusiu a dohodnúť sa.

C. Štandardizované testy

Príprave štandardizovaného testu sa podrobne venujeme v kapitole 3. Plánovanie a tvorba testu.

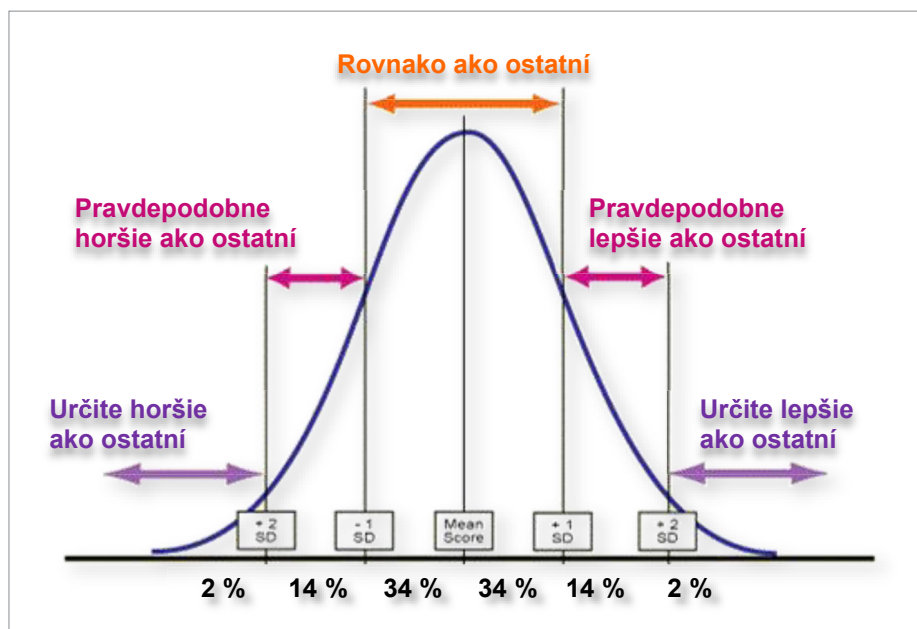
2.3.6. Podľa spôsobu interpretácie výsledkov

V odbornej literatúre sa zvykne hovoriť o dvoch druhoch didaktických testov. V skutočnosti ide skôr o dva spôsoby ich administrácie, spracovania údajov a interpretácie výsledkov.

V pedagogickej praxi sa najčastejšie používajú kritériové testy (CR). Testy relatívneho výkonu (NR) sú oveľa častejšie v psychológii, ale používajú sa aj v školstve – napríklad pri nadnárodných testovaniach, kde neexistujú žiadne nadnárodné cieľové dokumenty, ktoré by mohli byť kritériom na posudzovanie výsledkov žiakov zo všetkých zúčastnených štátov.

A. Rozlišujúce testy (testy relatívneho výkonu, NR-testy, Normatívno-referenčné testy skrátené: normatívne testy)

Sú to testy na meranie **relatívneho výkonu** žiakov na základe dodatočného porovnania výsledkov.⁷⁰ Na základe výsledkov takéhoto testu síce nemôžeme zhodnotiť žiakove vedomosti, ale vieme ho podľa jeho výkonu zaradiť v rebríčku všetkých žiakov na konkrétne miesto. Výsledky NR-testov sa usporadúvajú na podielovej alebo poradovej stupnici (podľa skóre, podľa percentuálneho vyjadrenia podielu skóre alebo podľa percentilov).



Obrázok 5. Normálne rozdelenie pravdepodobnosti (Gaussovo rozdelenie)

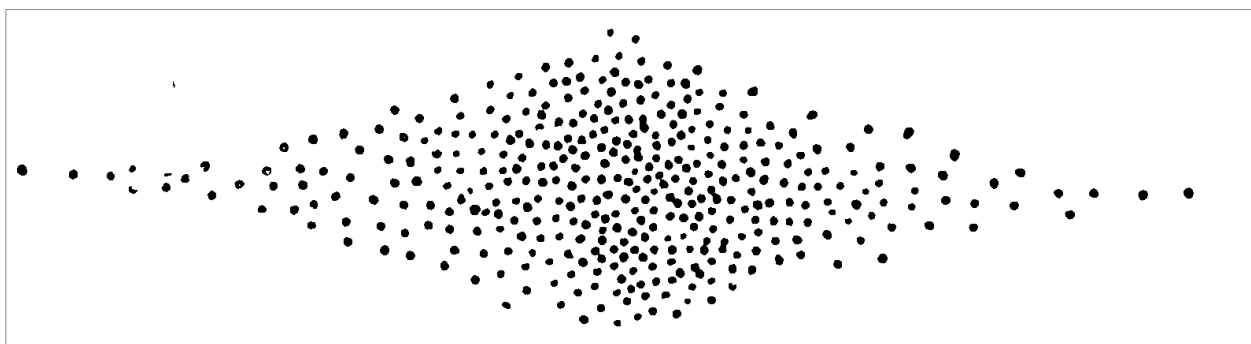
Ak je NR-test správne nastavený, pri väčšej vzorke zodpovedá distribúcia skóre Gaussovemu rozdeleniu. Teda pravdepodobnosť, že maximálny počet bodov v teste získa 100 % žiakov, je blízka nule, rovnako ako pravdepodobnosť, že v teste 100 % žiakov nezíska ani jeden bod. Zato môžeme očakávať, že výsledky väčšiny žiakov (zhruba sedemdesiatich percent) sa budú pohybovať okolo stredných hodnôt. Najlepšie a najhoršie výsledky by malo dosahovať zhruba do 5 % žiakov. Takýto test dobre rozlišuje medzi najlepšími a medzi najhoršími žiakmi.

„...skúmame, či medzi najslabšími a najlepšimi výsledkami je dostatočne veľký rozdiel a taktiež, či medzi týmito krajnými prípadmi sú výsledky ostatných žiakov v primeraných intervaloch rozdelené...“⁷¹

Ak by sme maratónskych bežcov zastavili po pätnástich minútach behu, vyzeralo by ich schematické rozmiestnenie asi takto:

⁷⁰ Turek, I.: *Kapitoly z didaktiky*. Didaktické testy. Bratislava : Metodicko-pedagogické centrum v Bratislave, 1995, s. 22. ISBN 80-85185-96-2

⁷¹ Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 27.



Obrázok 6. Grafické znázornenie rozmiestnenia bežcov v pelotóne⁷²

Vidíme, že vďaka rozdielom v schopnostiach najlepších a najhorších vznikli na oboch koncoch pelotónu akési výbežky – v týchto výbežkoch sú rozostupy medzi bežcami najväčšie – tam maratónsky beh najlepšie rozlišuje medzi bežcami. Čím je beh dlhší, podobne ako test, tým sú rozdiely medzi najlepšími (aj medzi najhoršími) väčšie. Ak by bol beh krátky, povedzme na 100 m, rozdiely medzi najlepšími bežcami by sme museli merať v stotínach sekundy. Pri ľahkom a krátkom teste, v ktorom je počet meracích jednotiek obmedzený počtom úloh, by sme nevedeli rozdiely medzi žiakmi vyjadriť číselne – všetci najlepší žiaci by mali rovnaký počet bodov. S takýmto testom sa bežne môžeme stretnúť pri prijímaní na školy – o prijatí a neprijatí žiakov na obmedzený počet miest potom nerozhoduje test, ale abecedné poradie najlepších doplnené odvolacími konaniami.

Normatívne testy požadujú, aby bol splnený predpoklad normality aj u premenných, ktoré sú nimi analyzované. Veľa premenných má normálne rozdelenie, a preto sa normálne rozdelenie považuje za všeobecnú črtu skúmanej reality. Aby sme však mohli porovnávať výsledky testu s Gaussovou krivkou, musíme použiť test na dostatočne veľkej vzorke. Ak totiž narastá veľkosť vzorky, krivka funkcie sa približuje ku Gaussovej krivke, a to dokonca aj vtedy, ak rozdelenie skúmanej premennej nie je normálne. Tento princíp je dokázaný matematicky (G. Pólya⁷³) aj empiricky, takzvanými Monte-Carlo experimentmi. Preto je používanie testov závislých od normálneho rozdelenia populárne vo všetkých druhoch výskumu.⁷⁴

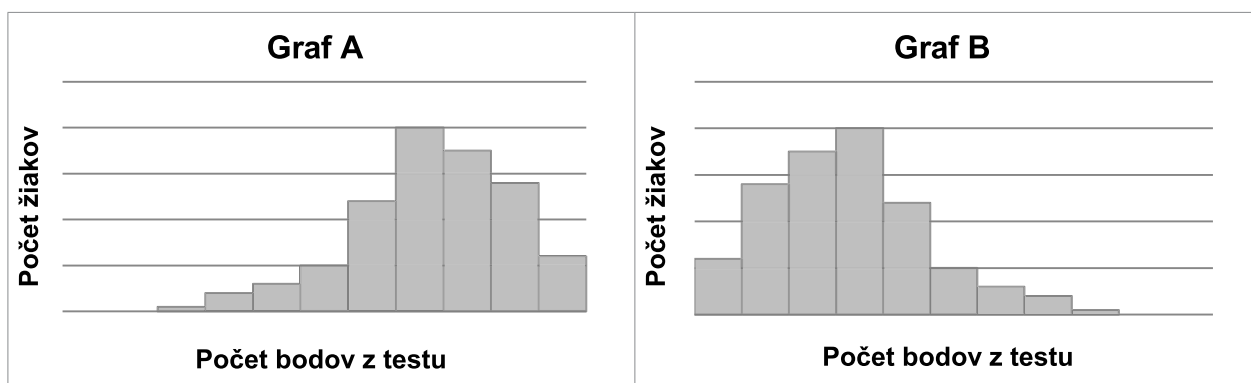
Prijímačkový test však môže byť nastavený aj ako ťažší, vtedy test dobre rozlišuje len najlepších žiakov (graf B). O žiakoch, ktorí sú priemerní a slabší, takýto test nehovorí takmer nič. Vrchol Gaussovej krivky je vtedy posunutý doľava.

Ak je krivka normálneho rozloženia posunutá doprava (graf A), môže to (pri veľkej vzorke žiakov) znamenať: 1. že test je príliš ľahký alebo sa v ňom prejavila zvýšená snaha žiakov uspieť, 2. že testovanie neprebehlo objektívne, a žiaci odpisovali. Ak je test ľahký, môže dobre rozlišovať iba ak najslabších žiakov. Podobne orientovanú krivku majú aj CR testy, ktoré v podstate fungujú ako vyradovacie, teda opačne, ako prijímačkové. Ak však výsledky nezodpovedajú normálnemu rozloženiu pri malej vzorke žiakov (trieda, škola), nemusí to znamenať nič z vyššie spomenutého. Môže jednoducho ísť o triedu s výnimočnými alebo so slabými žiakmi, neprimeraný, príliš ťažký alebo príliš ľahký test a pod.

⁷² Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 32.

⁷³ Central limit theorem. http://en.wikipedia.org/wiki/Central_limit_theorem (1. 4. 2015)

⁷⁴ Litavcová, E.: *Aplikovaná štatistika v príkladoch*. 2002 – 2003. <http://www.fhvp.unipo.sk/cvt/statistika/stuvod.htm> (1. 4. 2015)



Obrázok 7. Grafy rozloženia skóre žiakov (histogramy)

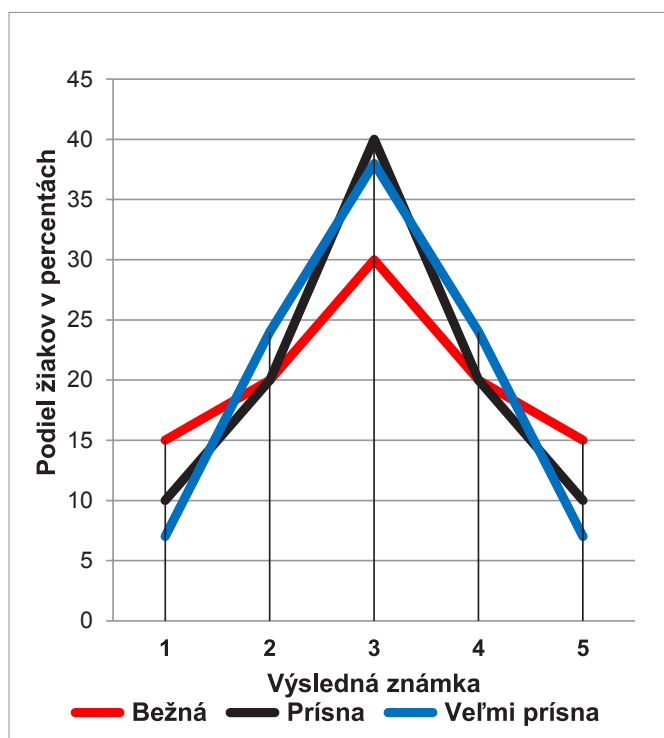
Prostredníctvom štatistických procedúr možno empirické rozloženie výsledkov testovania skupiny žiakov transformovať (normalizovať, resp. štandardizovať) na približné rozloženie s ľubovoľným priemerom a ľubovoľnou smerodajnou odchýlkou. V súčasnosti sa štandardizácia skóre používa napríklad pri vyhodnocovaní medzinárodných výskumov (**IEA**: PIRLS, ICCS, TIMS, ICILS; **OECD**: PISA), pričom skóre sa štandardizuje napríklad na priemer M (mean) = 500 bodov so smerodajnou odchýlkou SD (standard deviation) = 100 bodov. **Len vďaka takejto štandardizácii skóre je potom možné porovnať výkon krajín aj v rámci jednotlivých cyklov výskumov.**

Bodové skóre z didaktického testu je pre potreby klasifikácie možné podľa normálneho rozloženia previesť aj na klasifikačné stupne (tab. 8).

Tabuľka 8. Podiel žiakov v percentách podľa známok pre rôzne typy klasifikácie

Typ klasifikácie			Známka
Bežná	Prísna	Veľmi prísna	
15 %	10 %	7 %	1
20 %	20 %	24 %	2
30 %	40 %	38 %	3
20 %	20 %	24 %	4
15 %	10 %	7 %	5

Pre jednotlivé podiely žiakov sa vypočítajú zodpovedajúce počty bodov v didaktickom teste.



Obrázok 8. Krivky rôznych typov klasifikačných stupníc

Percentuálne vyjadrenie skóre však nemusí presne popisovať schopnosti žiakov oproti skupine. Všetci asi poznáme situáciu, keď žiak príde domov so zlou známku z písomky a argumentuje tým, že vraj nikto nemal jednotku a všetci dostali trojky a horšie známky. Pre presnejší odhad schopností žiakov si na základe skóre preto môžeme dopočítať aj *percentil* podľa nasledovného postupu:

1. Zostavenie tabuľky kumulovanej početnosti výsledkov:

Tabuľka 9. Návod na zostavenie tabuľky kumulovanej početnosti skóre

Počet bodov	Početnosť výsledku (n_i)	Kumulovaná početnosť výsledku (n_k)
n_1	a	$x = a$
n_2	b	$y = x + b$
n_3	c	$z = y + c$ (...atď.)

Tabuľka 10. Tabuľka kumulovanej početnosti skóre v teste, ktorý riešilo 20 žiakov

Počet bodov	Početnosť výsledku (n_i)	Kumulovaná početnosť výsledku (n_k)
2	1	1
7	3	4
8	4	8
9	2	10
10	2	12
12	3	15
13	2	17
15	2	19
18	1	20
Spolu: 20 ž.		

2. Dosadenie do vzorca:

$$PR_i = 100 \cdot \frac{n_k - \frac{n_i}{2}}{n}$$

PR_i – percentilové poradie (percentile rank) žiaka s počtom bodov i

n_i – početnosť výsledku i

n_k – kumulovaná početnosť výsledku i

n – počet testovaných žiakov

Ukážka výpočtu percentilu u žiaka, ktorý získal v teste 13 bodov (max. počet bodov bol 18):⁷⁵

$$PR_{13} = 100 \cdot \frac{17 - \frac{2}{2}}{20} = 80 \%$$

Žiak, ktorý získal 13 bodov, zvládol test iba na 65 %, ale až 80 % žiakov dosiahlo v teste horší výsledok ako on.

B. Overujúce testy

Testy absolútneho výkonu, CR-testy, kritériovo-referenčné testy, skrátene: kritériové testy – názov pre tento spôsob spracovania testov – criterion-referenced test – vymyslel R. Ebel v príspevku s príznačným titulom *Content-Standard Test Scores* (1962).⁷⁶

„Už začiatkom 60. rokov sa prvýkrát objavuje myšlienka konštruovať a vyhodnocovať didaktické testy inak: vopred určiť isté kritériá... podľa ktorých sa budú konkrétne výkony žiakov posudzovať. Výsledkom takéhoto „criterion-referenced“ testu... je potom zhodnotenie výkonu žiaka na pozadí daných kritérií (pričom výkon žiaka sa neporovnáva s výkonmi iných žiakov). Je zrejmé, že **výber kritérií** je jedným z kľúčových problémov takéhoto hodnotenia. Zväčša **vychádza z cieľov vyučovania** a z istej teórie kognitívnych mechanizmov učenia sa.“⁷⁷

Absolútny výkon žiaka sa v našich podmienkach odvodzuje od požiadaviek určených štátom (školský zákon, štátny vzdelávací program, výkonové štandardy) a jeho hodnota sa určuje na základe

⁷⁵ Jeřábek, O. – Bílek, M.: *Teorie a praxe tvorby didaktických testů*. Olomouc : Univerzita Palackého v Olomouci, 2010, s. 20 – 22. ISBN 978-80-244-2494-1.

⁷⁶ Ebel, R. L.: *Content-Standard Test Scores*. In: Educational and Psychological Measurement. R. 13, 1962, č. 1, s. 15 – 20.

⁷⁷ Burjan, V.: *Evalvacia a hodnotenie výsledkov vyučovania matematiky*. Bratislava : Metodické centrum v Bratislave, 1993, 52 s.

porovnania maximálneho možného počtu bodov, ktoré možno v teste získať, s počtom testových položiek, ktoré žiak správne vyriešil. Počet možných výsledkov je pri kritériových testoch obmedzený na niekoľko hodnotiacich stupňov, ktoré sú vopred stanovené, napr. neprospeš – prospel; neprospeš – prospel – prospel veľmi dobre a pod, kým pri normatívnych testoch je výsledkom napríklad percentil, pričom hodnota percentilov závisí od počtu a od výsledkov všetkých testovaných žiakov.

Pri kritériových testoch môže byť medzi riešiteľmi viaceru jednotlivcov s rovnakým výsledkom a nemusí sa vyskytnúť ani jeden žiak s absolútnym výkonom (nulové alebo maximálne skóre). Rozloženie výsledkov v histograme vytvára krivku s vrcholom výrazne posunutým doprava (pozri graf A – obr. 7).

Typickým príkladom kritériového testu je priebežný učiteľský test, zahŕňajúci 1 – 2 tematické celky s väčším počtom úloh z prebratého učiva, pričom väčšina úloh je vyriešiteľná pre väčšinu žiakov (štvrtročné písomky a pod.). Takýto test nemusí byť plánovaný ako formatívny a môže slúžiť aj na klasifikáciu žiakov.

2.3.7. Podľa typu úloh

A. Objektívne skórovateľné

Patria sem testy s uzavretými úlohami a s otvorenými úlohami s tvorbou krátkej odpovede. Na hodnotenie takýchto úloh môže stačiť aj kľúč správnych odpovedí, prípadne hodnotiacia príručka.

B. Subjektívne skórovateľné

Patria sem testy s otvorenými úlohami s tvorbou dlhej odpovede alebo s tvorbou slohových útvarov. Úlohy tohto typu nemôžu byť hodnotené objektívne (nie je možné u nich stanoviť objektívne kritériá hodnotenia) a nie sú vhodné do štandardizovaných didaktických testov, i keď umožňujú skúšať oveľa komplexnejšie vedomosti a zručnosti, ako objektívne skórovateľné úlohy.⁷⁸

V testoch, ktoré si tvoria sami učitelia, však môžu otvorené úlohy zohrať významnejšiu úlohu, pretože každú odpoveď preskúma a vyhodnotí tvorca testu individuálne.

2.3.8. Podľa rozsahu testovaného učiva

A. Monotematické testy

Monotematické didaktické testy sú zväčša zamerané na testovanie jedného tematického celku z jedného vyučovacieho predmetu. Patria sem aj testy zamerané na jednu zručnosť alebo kompetenciu, napríklad testy na čitateľskú gramotnosť a pod.

B. Polytematické testy

Polytematické testy testujú buď viac tematických celkov jedného vyučovacieho predmetu naraz, alebo testujú zložky viacerých vyučovacích predmetov – nadpredmetové testy. Patria sem aj testy mapujúce kľúčové kompetencie vo všeobecnosti alebo testy, zamerané na meranie matematickej, prírodovednej a čitateľskej gramotnosti (PISA) a pod.

⁷⁸ Chráška, M.: *Metody pedagogického výskumu*. Praha : Grada, 2007, s. 188. ISBN 978-80-247-1369-4.

2.3.9. Podľa úlohy učiteľa v testovaní⁷⁹

A. Statické testy

Pri statických testoch učiteľ nemôže zasahovať do priebehu testovania, ani pomáhať žiakom. Statické testy merajú výsledky učenia, ktoré jednotlivec nadobudol pred testovaním. Pri statickom teste je vzdelávacia funkcia testovania vedľajšia, a pri jeho hodnotení sa pri nesprávne zodpovedaných úlohách neprihliada na správny postup riešenia ani na dôvody, ktoré spôsobili chybnú odpoveď žiaka.

B. Dynamické testy

Dynamické testovanie zahŕňa zámerné učenie v testovej situácii ako súčasť diagnostického procesu – učiteľ pri pozorovaní procesu testovania odhaľuje, ako dieťa rieši test, ako postupuje pri riešení úloh a uvažuje nad tým, ako je možné upraviť jeho vyučovanie. Dynamické testy poskytujú informáciu **o učebnom potenciáli** testovaného žiaka: ako sa dokáže učiť v nových situáciách, keď sú prekážky učenia odbúrané (s pomocou učiteľa).

Dynamické testy porovnávajú výkon jednotlivca s ním samým. Používajú sa napríklad na testovanie zdravotne znevýhodnených žiakov, na testovanie žiakov zo sociálne znevýhodňujúceho prostredia, na testovanie nadaných žiakov, na testovanie inteligencie žiakov a pod.

Podľa spôsobu administrácie delíme dynamické testy na:

- a) Adaptívne testy (AT),
- b) Počítačové adaptívne testy (CAT – computer adaptive tests).

Pri CAT testoch nahrádza učiteľa dômyselný počítačový algoritmus, ktorý na základe odpovedí žiaka na zadávané úlohy rozhoduje o tom, akú obťažnú úlohu ďalej bude žiak riešiť. Každý žiak tak môže riešiť test šitý na mieru jeho schopností. Úlohy algoritmus prideluje žiakovi dovtedy, kým sa chyba merania nezníži na stanovené minimum. V literatúre sa tento algoritmus nazýva aj tzv. *virtuálny učiteľ*.

Tento systém testovania výrazne skracuje testovací čas, pretože napríklad slabší žiak nedostáva na riešenie úlohy, s ktorými si neporadí, a nadanejší žiak nemusí riešiť preňho banálne ľahké úlohy. Tento princíp tak zároveň preukázateľne znižuje frustráciu a zvyšuje motiváciu žiakov pri riešení testu. Podrobnejšie sa počítačovým adaptívnym testom venujeme v kapitole 6. Banka úloh pre adaptívne testovanie.

⁷⁹ Mesárošová, M.: *Dynamické testovanie jazykových a verbálnych schopností detí zo sociálne znevýhodňujúceho prostredia*. In: *Pedagogika* : časopis pro vědy o vzdělávání a výchově. Praha : Pedagogická fakulta Univerzity Karlovy v Praze, 2008, r. 58, č. 4, s. 327 – 336. ISSN 2336-2189.
http://pages.pedf.cuni.cz/pedagogika/?attachment_id=1108&edmc=1108 (3. 11. 2014)

2.4. Vlastnosti didaktického testu

V odbornej literatúre zaberajúcej sa testami a testovaním sa môžeme stretnúť s rôznymi charakteristikami či vlastnosťami testov. Mohli by sme ich rozdeliť do dvoch skupín, v rámci ktorých prevláda určitá hierarchia:

- základné alebo hlavné vlastnosti testov,
- ďalšie vlastnosti testov.

Medzi základné vlastnosti didaktických testov patria:

1. validita (obsahová správnosť/platnosť) testu,
2. reliabilita (spoľahlivosť) testu.

Medzi ďalšie vlastnosti didaktického testu patria najmä:

3. obťažnosť,
4. citlivosť,
5. praktickosť,
6. objektívnosť.

Tieto ukazovatele považujú mnohí odborníci v oblasti testovania za štatistické ukazovatele, ktoré ovplyvňujú reliabilitu testu.

2.4.1. Validita

Pri testovaní si musí autor položiť tieto otázky:

- Vzťahuje sa test na ciele výchovy a vzdelávania?
- Odpovedá test na našu výskumnú otázku?
- Meriame týmito úlohami to, čo chceme merať?

Dobry test je **validný**, ak vieme, čo chceme merať, ako to odmerať, a keď test skutočne meria to, čo chceme odmerať. V takom prípade sa hovorí, že test má **obsahovú validitu**. Základným predpokladom validity testu je konkrétna formulácia učebných cieľov a úloh a nadväznosť úloh na tieto ciele.

Jedným zo základných predpokladov validity testu je vypracovanie podrobnej špecifikácie testu, ktorá by mala obsahovať jednoznačné ciele testovania. (Pozri kapitolu 3. Plánovanie a tvorba testu.)

„Akýkoľvek test alebo spôsob hodnotenia sa dajú pokladať za validné, ak je možné dokázať, že to, čo sa reálne hodnotí (konštrukt), je to, čo má byť hodnotené,“⁸⁰ a že získané informácie odzrkadľujú vedomosti a zručnosti každého žiaka. Povedané inými slovami, test je validný, ak skutočne meria cieľ (ciele) testu.

Podľa účelu a metódy stanovenia delíme validitu na tieto druhy:

- **Obsahová validita** (content validity) – obsah testu musí predstavovať reprezentatívnu vzorku učiva, zručností, štruktúr atď. preberaných na vyučovacích hodinách, čiže obsahová validita udáva, do akej miery je obsah testu v súlade s prezentáciou učiva v pedagogických dokumentoch. Niekedy sa označuje táto validita aj ako kurikulárna (niekedy tiež *kurikulárna normativita*).
- **Konštruktová validita** (construct validity) – táto forma validity udáva, do akej miery test meria určitú charakteristiku žiakov (logické myslenie žiakov, tvorivosť a pod.), čiže udáva, do akej miery úlohy v teste odzrkadľujú základné aspekty teórie, na ktorej je test založený.

⁸⁰ Common European Framework of Reference for Languages: *Learning, teaching, assessment* (CEFR). Cambridge : Cambridge University Press, 2007. s. 177. ISBN 0521803136
http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (25. 10. 2008)

- **Súbežná validita** (concurrent validity) – udáva zhodu výsledkov testu s iným testom, ktorý testuje to isté.
- **Prediktívna validita** – schopnosť testu predpovedať výsledok v budúcnosti.
- **Dojmová validita** (face validity) – vypovedá o tom, čo si učitelia a žiaci o teste myslia (test je príliš ľahký/ťažký a pod).

Obsahovú validitu testu majú posudzovať experti, ku ktorým patria v prvom rade odborníci na testovanie, metodici, recenzenti, inšpektori, autori učebníc, skúsení učitelia a pod. Každú úlohu navrhovaného testu posudzujú vo vzťahu k Štátnemu vzdelávaciemu programu.

Pri zostavovaní jednoduchého testu učiteľom sa obvykle posudzuje iba tento druh validity testu.

Faktory, ktoré negatívne ovplyvňujú obsahovú validitu testu:

- Zaraďovanie úloh, ktoré pracujú s okrajovým učivom, tzv. chytáky.
- Zaraďovanie tzv. hlavolamov a úloh s nereálnym (fantastickým) kontextom, ktoré sú vhodné skôr do súťaží a olympiád.
- Chybný výber výkonov.
- Nesúlad úloh s cieľmi testovania a s cieľmi výchovy a vzdelávania.
- Nesprávna konštrukcia úloh, ktorá spočíva v tom, že na vyriešenie úlohy musia žiaci použiť niekoľko myšlienkových operácií, pričom niektoré z nich nemusia ovládať.⁸¹
- Nejednoznačná alebo nesprávna formulácia testovej úlohy (úzko súvisí s reliabilitou).
- Nejednoznačné výrazy, cudzie slová alebo málo používané slová či komplikovaná vetná stavba a pod. môžu spôsobiť, že žiaci pochopia úlohu inak, ako autori pôvodne zamýšľali.
- Výskyt úloh s príliš nízkou alebo vysokou obťažnosťou v teste.

Konštruktová validita. Pri zostavovaní učiteľského testu je expertom aj posudzovateľom sám učiteľ, a ak chce testovať nejaký konštrukt alebo ak chce zostaviť diagnostický formatívny test, musí venovať zvýšenú pozornosť konštruktivej validite. Musí skontrolovať konzistentnosť čiastkových cieľov a výkonov, ktoré sú reprezentované úlohami, s vyššie postavenými cieľmi v Štátnom vzdelávacom programe a zároveň ich zosúladiť s účelom a cieľmi testu.

Súbežná validita je miera zhody výsledkov testu s iným podobným akceptovaným kritériom úspešnosti. Ak sa výsledky testu porovnávajú s výsledkami nezávislého a spoľahlivého hodnotenia kompetencií žiaka, stáva sa toto nezávislé hodnotenie kritériom pre validáciu nového testu. V takom prípade hovoríme o **kritériovej validite**. Je však nevyhnutné, aby test, ktorý sa stáva kritériom, bol nielen skutočne validný, ale aby meral aj rovnaký konštrukt a aby mal rovnakú kognitívnu náročnosť ako posudzovaný test (aby bol porovnateľný), pretože inak by ani pozitívna korelácia obidvoch testov negarantovala ich kritériovú validitu (ak by boli napríklad obidva nevalidné).

Výsledky testu môžu mať aj vysokú **prediktívnu validitu**, ak sa porovnávajú s výkonmi žiaka v ďalšom štúdiu. V tomto prípade na rozdiel od súbežnej validity musí medzi jednotlivými meraniami prebehnúť dlhší čas. Ak ide o vysokú koreláciu medzi výsledkami testu napríklad na konci deviateho ročníka a medzi výsledkami, ktoré žiak dosahuje na strednej škole, je možné na základe výsledkov testu **predpovedať** úspešnosť žiaka aj v ďalšom štúdiu. Prediktívna validita testu je schopnosť testu byť dobrým prediktorom vývoja výkonu v určitej oblasti, je to vlastne predpovedanie (prognózovanie) určitých vlastností a schopností žiakov.⁸²

⁸¹ Autori úloh môžu považovať určitú myšlienkovú operáciu za elementárnu a samozrejmu, a preto ju nepovažujú za cieľ úlohy. Pre žiakov však môže byť práve táto operácia prekážkou úspešného vyriešenia úlohy, t. j. zamýšľaného cieľa autorov testu. Napríklad, ak by žiaci dostali úlohy zamerané na čítanie s porozumením, ktoré by vychádzali z dlhého odborného textu s množstvom pre nich neznámych termínov, s dlhými zloženými súvetiami a pod., tak by sa v ňom nemuseli správne zorientovať a nevedeli by odpovedať na pomerne jednoduché otázky, na ktoré by pri inom východiskovom texte odpovedali správne.

⁸² Rafajlovičová, R. – Štulrajterová, M.: *Skúšanie, testovanie a hodnotenie v edukačnom procese*. Bratislava : ŠPÚ, 2002, s. 19. ISBN 80-85756-67-6

Opäť však platí, že ak chceme stanoviť prediktívnu validitu testu, musíme porovnávať podobné testy a podobné konštrukty. Kým v minulosti vo vzdelávaní dominovali testy vedomostí, dnes sú testy väčšinou konštruované na základe niektorej taxonómie logicko-poznávacích operácií, a je nemožné a nesprávne porovnávať výsledky testu obsahujúceho aplikačné úlohy s výsledkami skúšky, pri ktorej sa overuje schopnosť zapamätať si vedomosti.

2.4.2. Reliabilita

Reliabilita testu je ukazovateľom presnosti a spoľahlivosti merania. Výsledky získané v teste relatívne presne ukazujú skutočné vedomosti žiakov. Reliabilita je štatistická veličina, ktorá vyjadruje spoľahlivosť testu. Dá sa charakterizovať ako relatívna neprítomnosť chyby pri meraní.

Každý test musí byť spoľahlivým meradlom zručnosti, ktorú má hodnotiť, a musí zabezpečovať konzistentne zhodné miery v každom testovaní, t. j. ak sú testovaní tí istí žiaci v totožných situáciách, výsledky testu by mali byť zhodné. Výsledky testu by sa nemali odlišovať ani vtedy, ak je ten istý test vyhodnocovaný dvomi alebo niekoľkými odlišnými hodnotiteľmi. Žiakovo skóre by malo byť rovnaké (alebo približne rovnaké), aj keď má inú verziu testu.

Reliabilita testu tiež vyjadruje, do akej miery je zachované rovnaké poradie žiakov v dvoch samostatných (skutočných alebo simulovaných) administráciách toho istého hodnotenia.⁸³ To znamená, že test by mal mať schopnosť merať konzistentne, a keby ten istý žiak robil rovnaký test niekoľkokrát, mal by vždy dosiahnuť rovnaké výsledky.

Každá hodnota, ktorú žiak v teste dosiahne, sa dá z hľadiska presnosti rozdeliť na dve časti:

- správnu hodnotu – mala by vyjadrovať to, čo žiak naozaj vie,
- chybovú hodnotu – tvorí súčasť výsledku, ale nezodpovedá skutočnému stavu vedomostí žiaka. Na základe chýb môžeme zistiť, čo žiak nevie alebo aký má problém v učení.

Reliabilita sa kvantitatívne vyjadruje koeficientom reliability, ktorý udáva, do akej miery sú zhodné výsledky získané po niekoľkonásobnom opakovaní toho istého testu. Ideálne by mal mať test reliability vyjadrenú číslom 1,00, čo znamená, že ak by sa takýto test opakovane zadal, dosiahli by v ňom všetci žiaci identický počet bodov za každú úlohu ako v prvom meraní. V praxi sa ale reliability s hodnotou 1,00 nevyskytuje. Reliabilita sa pohybuje od čísla 0,00 (žiadna zhoda s predošlými výsledkami v danom teste) až po 1,00 (úplná zhoda s predošlými meraniami).

Koeficient reliability sa spravidla pohybuje medzi hodnotami 0,5 až 0,95:

- nad 0,90 test sa považuje za vynikajúci,
- 0,85 – 0,95 ® na základe testu možno prijať rozhodnutia,
- 0,65 – 0,85 ® test možno použiť ako jeden z predpokladov pre rozhodnutia,
- pod 0,65 ® nespoľahlivý test, nevhodný pre klasifikáciu, prijímaciu skúšku a pod.

Čím je reliability vyššia, tým menší vplyv na výsledok má náhoda. Mieru reliability didaktického testu zlepšuje počet úloh zaradených do testu.

Metódy zisťovania (výpočtu) reliability:

- **Metóda test – retest**

Táto metóda je založená na porovnávaní. Ten istý test zadáme dvakrát po sebe a vypočítame koeficient korelácie medzi výsledkami. Keď žiak rieši ten istý test dvakrát, jeho výsledky by mali byť zhodné, resp. takmer zhodné (výkonnosť človeka kolíše a pod.). Nedostatkom tejto

⁸³ Common European Framework of Reference for Languages: *Learning, teaching, assessment (CEFR)*. Cambridge : Cambridge University Press, 2007. s. 177. ISBN 0521803136
http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (25. 10. 2008)

metódy je, že sa test musí opakovať veľmi rýchlo, pretože výučba v škole pokračuje, t. j. žiak by mohol mať pri opakovanom riešení testu viac vedomostí a zručností, ako mal pri prvom meraní. Na druhej strane, ak zadáme žiakom ten istý test príliš rýchlo po predchádzajúcom, žiaci si môžu niektoré úlohy pamätať.

- **Metóda paralelných/ekvivalentných foriem**

Pri tejto metóde sa obe merania vykonajú počas jedného testovania. V teste sú umiestnené dvojice otázok, ktoré vecne patria k sebe, pýtajú sa na tú istú vec, ale majú odlišne znejúce zadania. Dvojice paralelných úloh musia byť operačne na rovnakej úrovni. Skóre každého žiaka je súborom dvojíc výsledkov (0/0, čiže nesprávna odpoveď/nesprávna odpoveď a 1/1, čiže správna odpoveď/správna odpoveď – to je želaný výsledok, ale 0/1, resp. 1/0, čiže kombinácia nesprávnej a správnej odpovede – to je výsledok, ktorý naznačuje chyby merania).

Reliabilita testu je ovplyvňovaná viacerými faktormi:

- **Počet úloh:** čím je úloh viac a čím viac aspektov učiva testujú, tým sú výsledky testu reliabilnejšie. To znamená, že ak žiaci majú v teste málo úloh, nemusia podať najlepší výkon, pretože v teste sa náhodou testujú tie časti učiva, ktoré sú pre nich ťažké. Platí to však aj opačne – ak majú šťastie a dostanú „lahkú úlohu“, ich výkon môže byť oveľa lepší ako by bol ich priemerný výkon. Pretože riešia len pár úloh a nikto nedostane inú šancu, výsledky testu nie sú reliabilné. Preto je lepšie mať niekoľko nezávislých úloh a každá z nich by mala byť pre žiaka novým začiatkom. Výsledok testu, najmä ak ide o výstupný test, by nemal byť záležitosťou šťastia.
- **Úroveň obťažnosti jednotlivých úloh alebo položiek v rôznych variantoch testu má byť rovnaká.**
- **Výber úloh** v teste má predstavovať reprezentatívnu vzorku učiva.
- **Volba distraktorov:** úlohy so zle zvolenými distraktormi (ponúknutými nesprávnymi odpoveďami) znižujú reliabilitu.
- **Pokyny a grafické spracovanie úloh:** nejasné, nezrozumiteľné pokyny a grafické objekty znižujú reliabilitu,
- **Prediktabilita** (predpovedateľnosť, predvídateľnosť, uhádnuteľnosť správnej odpovede). Ak žiak dokáže vyriešiť napríklad aplikačnú úlohu na základe svojich predchádzajúcich vedomostí bez toho, aby skutočne úlohu riešil, úloha je chybná.

Reliabilita samotného testovania je pri administrácii a vyhodnocovaní ovplyvňovaná ďalšími faktormi:

- Testovaní žiaci, ich osobnosť a momentálna dispozícia (strach, nepresnosť, unáhlenosť, únava a pod.).
- Rozptyl v úrovni schopností testovaných žiakov (väčší rozptyl ⇒ vyššia reliabilita).
- Počet žiakov (viac žiakov ⇒ vyššia reliabilita).
- Nesprávna administrácia testu.
- Zlé podmienky testovania (hluk, nekvalitná nahrávka, nedostatok času a pod.).
- Odpisovanie alebo napovedanie správnych odpovedí.
- Nesprávne vyhodnocovanie (subjektívne hodnotenie, neobjektivita a nejednoznačnosť pri hodnotení výsledkov).

Vzťah medzi reliabilitou a validitou

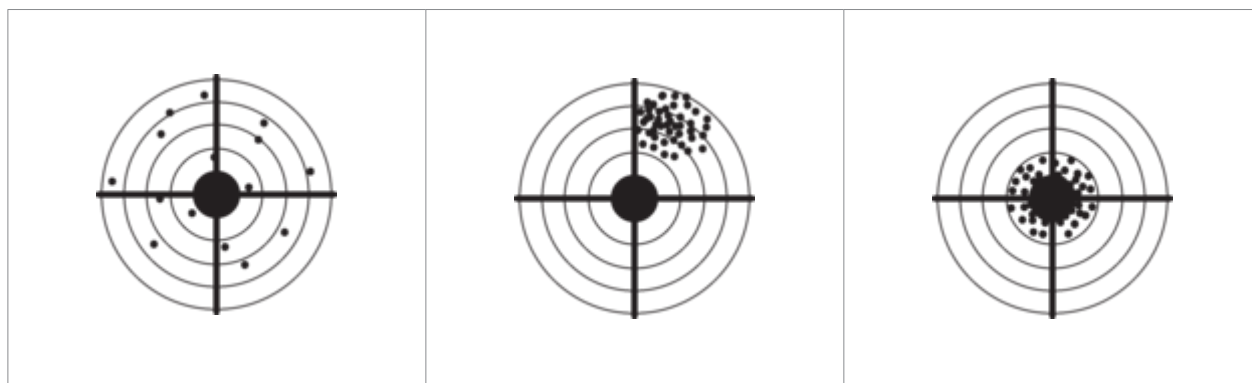
Reliabilitou sa zisťuje, do akej miery sa v teste podarilo potlačiť vplyv náhody. Udáva, nakoľko by žiaci dosiahli rovnaké výsledky v opakovanom testovaní. Reliabilita je nevyhnutnou, ale nie postačujúcou podmienkou validity.

Vysvetlime si vzťah medzi reliabilitou a validitou na príkladoch strieľania do terča. Reliabilita vypovedá o tom, ako veľmi sú strely rozptýlené po terči. Validita vypovedá o tom, ako často strely trafili stred (cieľ).

Ak sú strely rozptýlené, presnosť a spoľahlivosť (reliabilita) strelca je nízka a validita tiež, pretože ak sú strely rozptýlené, málokedy trafili cieľ.

Ak sú strely v terči nahusto blízko seba, spoľahlivosť a presnosť strelca je vysoká. To však nezaručuje, že sa strelec trafil do stredu terča (podobne didaktický test môže pomerne presne merať úplne iný konštrukt, ako sme zamýšľali).

Spoľahlivosť a presnosť je teda nevyhnutným predpokladom validity, ale nie jediným. Presný strelec sa musí aj presne trafiť do stredu terča. Aj validný didaktický test musí presne merať práve to, čo chceme merať.⁸⁴



Obrázok 9/a. Nízka reliabilita aj validita

Obrázok 9/b. Vysoká reliabilita a nízka validita

Obrázok 9/c. Vysoká reliabilita aj validita

Test môže byť **reliabilný** bez toho, aby bol **validný**. Test nemôže byť **validný** bez toho, aby bol reliabilný.

2.4.3. Obťažnosť

Obťažnosť jednotlivých testových úloh a následne celého testu vytvára autor testu už pri konštrukcii testových úloh a ich zaradení do testu.

Obťažnosť je z didaktického hľadiska komplexný údaj, ktorého štruktúru podmieňuje štruktúra učebnej látky a požiadavky na logicko-poznávacie operácie v zadaní úloh. Zamýšľaná obťažnosť testu musí vychádzať z vlastností zložiek, ktoré tvoria učebnú látku i samotnú úlohu:

1. stupeň dôležitosti učebného prvku v štruktúre učiva,
2. náročnosť a vekovú primeranosť logicko-poznávacjej operácie,
3. zložitnosť požadovaných intelektuálnych a manipulačných zručností.

⁸⁴ Analýza výsledkov a hodnotení kvality testu. http://www.wikiskripta.eu/index.php/Fórum:Testy#Standardizace_a_normov.C3.A1n.C3.AD_testu (1. 7. 2015)

Obťažnosť ako číselnú hodnotu vypočítanú štatistickými metódami (po spracovaní výsledkov testovania) vyjadrujeme v klasickej testovej teórii buď indexom obťažnosti (P) alebo hodnotou obťažnosti (Q).⁸⁵

Ak odhadujeme percentuálny podiel žiakov, ktorí na danú úlohu odpovedali **správne**, vypočítame index obťažnosti (P).

$$P = \frac{n_S}{n} \cdot 100$$

n_S – počet žiakov, ktorí odpovedali správne
 n – počet žiakov, ktorí riešili úlohu

Index obťažnosti môže mať hodnotu od 0 do 100 %. Percento je tým väčšie, čím viac žiakov úlohu vyriešilo správne, teda **čím väčšie percento, tým ľahšia úloha**. V zahraničných publikáciách sa pri obťažnosti P môžeme stretnúť aj s pomenovaním *easybility* (jednoduchosť) testovej úlohy.

NÚCEM pri položkovej analýze používa v posledných rokoch index obťažnosti (P). Hodnoty indexu obťažnosti (P) môžeme nájsť v správach z testovaní aj pod označením „úspešnosť“.

Mohlo by sa nám zdať, že index obťažnosti vyjadruje skôr jednoduchosť úlohy, a preto niektorí autori používajú tzv. hodnotu obťažnosti (Q). Hodnota obťažnosti sa odhaduje ako percentuálny podiel žiakov, ktorí na danú úlohu odpovedali **nesprávne**.

$$Q = \frac{n_N}{n} \cdot 100$$

n_N – počet žiakov, ktorí odpovedali nesprávne
 n – počet žiakov, ktorí riešili úlohu

Hodnota obťažnosti môže mať tiež hodnotu od 0 do 100 %. Percento je tým väčšie, čím viac žiakov úlohu vyriešilo nesprávne, teda **čím väčšie percento, tým ťažšia úloha**.

V starších výskumných správach NÚCEM-u sa môžeme stretnúť aj s hodnotami obťažnosti (Q), napr. ešte v národnej správe z Testovania 9-2012.⁸⁶

Vzťah medzi indexom obťažnosti (P) a hodnotou obťažnosti (Q) vyjadruje tento vzorec:

$$Q + P = 100 \%$$

Ak teda vypočítame hodnotu P a odčítame ju od čísla 100, získame hodnotu Q (a naopak).

2.4.4. Citlivosť

Citlivosť testu vyjadruje mieru, do akej test dokáže rozlíšiť žiakov s dobrými a slabými vedomosťami (diskriminačná sila testu): dôležitá je najmä u NR testov (rozlišujúcich). Nevyhnutná je tu pilotáž úloh, po ktorej sa analyzuje každá úloha zvlášť. Treba určiť diskriminačnú hodnotu úloh, čiže citlivosť.

Test s vysokou citlivosťou obsahuje úlohy rôznej obťažnosti (ľahké, stredne náročné, ťažké).

Citlivosť vypovedá o schopnosti testu rozlišovať medzi žiakmi s rôznou úrovňou skutočných vedomostí a zručností. V citlivom teste by výsledky žiakov mali byť primerane rozmiestnené po celej bodovej škále. Test nie je citlivý, ak všetci žiaci dosiahnu v teste výborný alebo zlý výsledok – nerozlišuje žiakov.

Optimálna miera citlivosti sa líši v závislosti od účelu testu. Vysokú citlivosť musia mať rozmiestňovacie testy, napríklad testy na zaradenie žiakov do skupín podľa počiatocnej úrovne jazykovej kompetencie vo vyučovaní CJ. Test môže dobre rozlišovať medzi najlepšimi žiakmi (výberové – prijímačkové testy) alebo medzi najslabšími žiakmi (vyraďovacie – postupové testy). Výberové testy majú obsahovať úlohy s vysokou citlivosťou a obťažnejšie úlohy. Ak test málo diskriminuje, vyradia sa úlohy, ktoré takmer

⁸⁵ Chráska, M.: *Didaktické testy*. Brno : Paido, 1999, 91 s. ISBN 80-85931-68-0.

⁸⁶ Polgáryová, E. – Alföldyová, I. (zost.): *Testovanie 9-2012. Priebeh, výsledky a analýzy*. Bratislava : NÚCEM, 2012, s. 7. http://www.nucem.sk/documents//26/testovanie_9_2012/Spr%C3%A1va_T9-2012_final.pdf (1. 4. 2015)

všetci žiaci vyriešili správne. Vyradovacie testy, naopak, obsahujú viac ľahkých úloh, ktoré majú dobrú rozlišovaciu schopnosť pri menej výkonných žiakoch. Test, prostredníctvom ktorého chceme overiť, či si každý jednotlivý žiak osvojil určité učivo, nemusí mať vysokú citlivosť.

2.4.5. Praktickosť testu

Praktickosť testu (*practicality*) je aspoň taká dôležitá, ako vlastnosti uvedené predtým. Pri tvorbe testu je potrebné uvažovať aj o tom, koľko údajov o žiakovi z testu získame v porovnaní s časom, ktorý je potrebný na jeho tvorbu, administrovanie, vyhodnocovanie a spracovávanie výsledkov. Povedané inými slovami, je potrebné uvažovať o tom, či informácie, ktoré z testu získame, stoja za prácu a peniaze, ktoré do testu vložíme.⁸⁷

Praktickosť testu spočíva v tom, že test musí učiteľom i žiakom ich prácu uľahčovať a zefektívňovať. Pomocou testu sa majú dať za krátky čas získať od väčšieho počtu žiakov relatívne spoľahlivé informácie, ktoré umožňujú porovnateľnosť výsledkov. Praktickosť testu je v jeho **jednoduchom administrovaní**, v skórovaní, ale aj vo vyhodnocovaní testových výsledkov – preto má mať test premyslenú štruktúru a pripravenú aj interpretačnú škálu.

Praktickosť testu je ovplyvňovaná zdrojmi, ktoré sú často limitované. Predovšetkým je potrebné brať do úvahy ľudské zdroje: koľko autorov úloh potrebujeme, koľko času bude treba na zostavenie testu, jeho validáciu a pilotovanie. Okrem toho treba brať do úvahy aj čas, ktorý budú potrebovať žiaci na jeho vypracovanie, a čas potrebný na jeho vyhodnocovanie.⁸⁸ Treba brať do úvahy aj množstvo papiera a náklady na rozmnoženiny testu a pod.

Praktickosť testovania docielime dôsledným naplánovaním. Do úvahy musíme zobrať: trvanie testu, počet testov, zadanie/inštrukcie, úpravu testu (formu odpovedových hárkov, grafickú kvalitu obrázkov, vzorcov, grafov, kvalitu audiovizuálnych nahrávok a pod.), náklady na jazykové mutácie testu, podmienky pre testovanie žiakov.

Žiaci by mali mať dostatok času na vypracovanie všetkých úloh v teste. Pritom ide o pomerne zložitý problém, pretože žiaci nepracujú rovnakým tempom. Ak žiaci nemajú k dispozícii dostatok času, cítia sa byť stresovaní. Ak však majú žiaci veľa času, môže dôjsť k pokusom o podvod a k nesústredeniu sa na svoju vlastnú prácu, čo samozrejme skresľuje výsledky.

Test je **ekonomický** preto, lebo sa dá rýchlo vypracovať. Jeho úlohy sú skonštruované tak, aby žiak nemusel písať dlhé texty. Preto za daný čas možno testom zistiť oveľa vedomostí a zručností, ako pri voľnej písomnej práci.

Test je ekonomický aj preto, lebo sa dá rýchlo vyhodnocovať. Osoba, ktorá test vyhodnocuje, môže pomocou kľúča správnych odpovedí za daný čas vyhodnotiť oveľa viac testov ako voľných písomných prác.

Použitelnosť testu je miera vhodnosti testu pre daný účel. Je to miera, do akej test zabezpečuje zadávateľovi ľahkú opravu, vyhodnotenie, formuláciu záverov, administráciu (typ testu, spracovanie testových a odpovedových hárkov, technické nároky na zadávanie testu a vyhodnocovanie a pod.). Súvisí s plánovaním a konštrukciou testu, s jeho vzhľadom a s použitím techniky na jeho vyhodnotenie a spracovanie.

⁸⁷ Weir, C. J.: *Understanding and Developing Language Tests*. New York : Phoenix ELT, 1993, s. 21. ISBN 978-0-13-947532-0

⁸⁸ Gondová, D.: *Taking first steps in teaching English: assessing learners*. Žilina : University of Žilina, 2010, s. 41. ISBN 978-80-554-0236-9

2.4.6. Objektívnosť testu

Objektívnosť patrí k vlastnostiam testu. Niektorí autori ju považujú za základnú vlastnosť (faktor) ovplyvňujúci reliabilitu. Znamená aj vylúčenie (alebo silné zredukovanie) náhodných alebo subjektívnych činiteľov pri testovaní. Objektívnosť je opakom subjektívnosti.

Objektívnosť testovania sa zabezpečuje:

- a) objektívnosťou testu,
- b) objektívnosťou vyhodnocovania,
- c) objektívnosťou priebehu testovania.

Objektívnosť testu sa dosahuje správnym výberom úloh, jednoznačnou formuláciou otázky a možností odpovede, dobrým pokrytím učiva úlohami a zároveň vylúčením všetkých pochybností o správnych odpovediach na jednotlivé úlohy. Obsahovú objektivitu testu zabezpečuje teda dôkladná kontrola každej úlohy. Úloha musí byť vecne úplne správna a správna odpoveď nesmie zodpovedať iba zjednodušenému výkladu v škole, musí byť pravdivá. Úloha nesmie diskriminovať ani najlepšieho žiaka, ktorý má vedomosti nad rámec školského vzdelávania. Zároveň nesmie test ani zvýhodňovať určitú skupinu žiakov (tematické rozvrhnutie položiek).

Objektívnosť vyhodnocovania zabezpečíme presnými a rovnakými pravidlami vyhodnocovania, rovnakým spôsobom skórovania (hodnotenia) tej istej odpovede (t. j. tá istá odpoveď sa hodnotí rovnako, nezávisle od toho, kto test vyhodnocuje). Nejasnosti môžu viesť k odlišnostiam u rôznych hodnotiteľov.

Spôsob vyhodnocovania úloh ovplyvňuje objektívnosť testovania, preto ho treba premyslieť dopredu. Treba sa rozhodnúť, či všetky položky budú hodnotené rovnako (t. j. jeden bod za každú správnu odpoveď), alebo či niektoré z nich budú vážené vyššie (t. j. dva body za každú správnu odpoveď). Ak sú niektoré položky hodnotené vyššie, malo by to byť v teste vyznačené. Pritom však treba dodržať pravidlo, že ak je odpoveď správna, získa žiak dva body, a ak je nesprávna, nezíska žiadny bod. Ak chceme pridelovať jeden bod za čiastočne správne odpovede, je potrebné rozhodnúť dopredu, za aké odpovede ho budeme pridelovať.

Skúsenosti s viacbodovým hodnotením úloh však ukázali, že vhodnejšie je bodovať všetky úlohy jedným bodom. Obvykle totiž získa čiastkové skóre (neúplný počet bodov) viac žiakov, ako plný počet – a to sa pri vyhodnocovaní prejavuje tak, akoby sme do testu umiestnili jednu príliš ľahkú úlohu a jednu príliš ťažkú. Vhodnejším spôsobom váženia je vyšší počet úloh testujúcich tú časť učiva, ktorú považujeme za mimoriadne dôležitú.

Aj napriek tomu, že vyhodnocovanie objektívne hodnotiteľných testov je pomerne jednoduché a obvykle si vyžaduje len označovanie správnych a nesprávnych odpovedí, je potrebné zabezpečiť, **aby bol každý test hodnotený dvomi hodnotiteľmi**, pretože každý z nich sa môže dopustiť chyby. Pretože takéto vyhodnocovanie je náročné na čas, pri hodnotení priebežných testov nie je vždy praktické požiadať o pomoc iného učiteľa. Druhé hodnotenie môžu v rámci spätnej väzby urobiť aj samotní žiaci. Test najprv vyhodnotí učiteľ, potom ho môže použiť na diagnostické účely, a nakoniec žiaci dostanú testy a kontrolujú si pridelené body s kľúčom, ktorý im dá učiteľ. Žiakom je potrebné vysvetliť, že skutočne kontrolujú, či učiteľ opravil test správne. Musia sa naučiť chápať, že aj učiteľ sa môže dopustiť chyby, a preto, aby sa možnosť nesprávneho hodnotenia zredukovala na minimum, sami urobia druhé hodnotenie.

Subjektívne hodnotiteľné testy sa hodnotia na základe dopredu vypracovaných kritérií. Okrem toho je však potrebné, aby všetci hodnotitelia absolvovali tréning hodnotenia. Pretože subjektívne hodnotiteľné testy nie sú súčasťou externého hodnotenia, nebudeme sa nimi podrobnejšie zaoberať.

Objektívnosť priebehu testovania dosiahneme zabezpečením rovnakých podmienok testovania pre všetkých testovaných žiakov (rovnaký čas, rovnaká miestnosť, inštrukcie, rovnaké úlohy a rovnaká možnosť odpovedania pre všetkých žiakov).

Výsledok testu nemajú ovplyvňovať žiadne negatívne vonkajšie činitele (*hluk, nedostatočné osvetlenie, nevhodný čas testovania, odpisovanie, nerovnaké pravidlá administrácie a pod.*).

3. PLÁNOVANIE A TVORBA TESTU

Ako sme uviedli už vyššie, kvalitný didaktický test nevzniká náhodným skladaním nejakých vopred pripravených úloh, ale dôkladne premysleným postupom. Tento zámer môže mať dve podoby:

- a) plán testu,
- b) špecifikácia testu.

Ak si test zostavuje učiteľ sám pre vlastné potreby, prvým krokom má byť vytvorenie **plánu testu**. Takyto plán nemusí mať podobu veľkolepého administratívneho dokumentu. Špecifikácia má zabezpečiť obsahovú a konštruktívnu validitu testu. Špecifikáciu testu zvyčajne pripravuje expert na testovanie v spolupráci s expertom na daný vyučovací predmet.

Ak zadávateľ testu a jeho tvorca nie je tá istá osoba, zadávateľ vytvorí na základe svojho plánu písomnú **špecifikáciu**, prípadne aj prehľadný a štruktúrovaný rozpis úloh a ich vlastností, ktorý sa nazýva **špecifikačná tabuľka**. V týchto dokumentoch zadávateľ určí požiadavky na test, ktoré má autor dodržať. V mnohých smeroch sa špecifikácia obsahovo podobá na plán testu, len s tým rozdielom, že zadávateľ presne uvedie obsah viacerých bodov. Vytvorenie, resp. dopracovanie špecifikačnej tabuľky môže zadávateľ zveriť aj autorovi testu. V takom prípade sa výsledok tejto práce aj recenzuje, podobne ako celý test.

Pri príprave štandardizovaného didaktického testu je potrebné vykonať nasledovné kroky:⁸⁹

3.1. Vymedzenie účelu testu

Na čo má test slúžiť? Aby sme mohli prikročiť k stanoveniu výskumnej otázky alebo k formulácii hypotézy, musíme si ujasniť, na čo chceme výsledky testu použiť.

Je rozdiel, či chceme pomocou testu vybrať najlepších žiakov, napríklad na olympiádu z daného predmetu, na strednú školu, či chceme vykonať inšpekčný test, či chceme zistiť výkony žiakov z daného predmetu na konci niektorého tematického celku, alebo či chceme zistiť, ako sa žiaci učia, a či sú naše vyučovacie postupy účinné.

Od účelu testu závisí celá špecifikácia testu, stupeň zvolenej obťažnosti úloh, dĺžka testu, počet úrovní testu aj obsahové vymedzenie testu.

3.2. Výber cieľovej skupiny

Súčasťou plánovania testu je aj vymedzenie všetkých účelových a okolnostných parametrov. Hneď na začiatku sa musí jasne určiť, aká skupina žiakov bude podrobená skúške.

⁸⁹ Zásady tvorby didaktického testu. Fakulta informatiky a manažmentu Univerzity v Hradci Králové. http://fim.uhk.cz/oliva/tvorba_vedeni/rekap-www/m01/m01-03.html (1. 4. 2015)

Asi najdôležitejšie hľadisko pri formovaní zamerania testu je určenie cieľovej skupiny žiakov podľa úrovne vzdelávania (stupeň školského vzdelávania) a podľa veku a zohľadnenie okolností, ktoré z toho vyplývajú. Inak sa tvoria testy pre deti zo základného stupňa, inak testy pre starších žiakov. Testy pre nižšie ročníky vzdelávania musia mať jednoduchšie texty, častejšie sa používajú aj obrázky, orientujú sa na životné skúsenosti z každodenného života danej vekovej skupiny. Najviac chýb sa zvykne urobiť ani nie tak pri obsahovom zameraní testu, veď tam je rozhodujúcim východiskovým materiálom platný pedagogický dokument (Školský zákon, Štátny vzdelávací program), ale pri informáciách, ktoré vyberá sám tvorca testu (napr. východiskové texty). **Autor nesmie do zadania testových úloh zahŕňať javy, s ktorými väčšina žiakov nemá ešte potrebné skúsenosti.** (Pozri kapitolu 4.3. Východiskový text.)

Základnou cieľovou skupinou pri administrácii testu je určenie triedy alebo školy. Ak sa testovanie týka len konkrétnej školy, je otázka cieľovej skupiny určená automaticky, alebo sa obmedzí na určenie triedy v rámci danej školy. Takýto cieľ sleduje napríklad testovanie na pokyn zriaďovateľa školy, najmä ak ide o nešťátne školy.

Inak je to so širokospektrálnym testovaním. Hromadné testy zadávajú napríklad školskí inšpektori, ktorí do svojho výberu zahrnú obyčajne väčší počet škôl, od jednej až povedzme po sto. Jednu školu testujú iba vtedy, ak má byť ich inšpekcia zacielená práve len na ňu. Ak však chcú získať prehľad o stave vyučovania, otestujú veľké počty škôl alebo tried.

Najväčšiu pozornosť verejnosti priťahujú *celoštátne testy*, u nás je to zatiaľ testovanie deviatakov a externé maturitné skúšky. V tomto prípade sa testovanie vzťahuje na všetkých žiakov príslušného populačného ročníka. Ak sa niektorý žiak pre chorobu alebo iné príčiny testovania nezúčastní, podrobí sa skúške neskôr, v náhradnom termíne, aby sa zabezpečila jednotnosť podmienok a objektivita skúšky.

Osobitným prípadom testovania sú *výberové testovania*. Na Slovensku sa s nimi stretáme napríklad pri pedagogických experimentoch, keď odborníci z výskumných ústavov a vysokých škôl overujú nové pedagogické nástroje (učebné osnovy, vzdelávacie štandardy, vyučovacie metódy). V takomto prípade sa experimentálne nástroje aplikujú na výberoch. Tie bývajú *reprezentatívne, stratifikované* alebo *typologické*.

Vytváraním štatisticky reprezentatívnych výberov sa v tejto publikácii nebudeme podrobnejšie zaoberať. Platí tu však pravidlo optimálneho kompromisu. Čím je (správne urobený) výber väčší (početnejší), tým presnejšie výsledky sa dosiahnu, ale vzrastú aj náklady na testovanie. Zároveň však platí, že čím je vnútorná rozmanitosť žiakov (z hľadiska znaku, ktorý sa bude zisťovať) väčšia, tým, väčší výber potrebujeme na zabezpečenie spoľahlivých výsledkov. Preto výber z populácie gymnazistov môže byť (percentuálne) skromnejší ako pri testovaní na základnej škole, kde je výkonový (vedomostný) rozptyl žiakov v ročníku oveľa väčší.

Dotkli sme sa problematiky miery *spoľahlivosti* výberového testovania. Rozumie sa tým miera zhody medzi skutočnou distribúciou meraného znaku v celej populácii žiakov a medzi výsledkom, ktorý získame pomocou pomerne malého výberu. Zhoda týchto dvoch hodnôt – výberového a populačného – prakticky nemôže byť úplná; oba údaje sa vždy mierne líšia (hovoríme o chybe merania). Ide však o to, aká veľká bude táto chyba. Ak sa pri testovaní priemerné empirické skóre (teoreticky sa pohybujúce medzi hodnotami nula a povedzme 80) odlišuje od skutočného o 1 % alebo o 2 body, považuje sa výsledok za vynikajúci. Ak by sa však rozdiel vyšplhal až na 20 bodov, tak celé výberové testovanie zlyhalo a získané poznatky o žiakoch nemajú žiadnu cenu.

Reprezentatívne výbery sa robia väčšinou náhodnou voľbou respondentov alebo skupín respondentov (tried, škôl). V škole platí skôr to druhé, nevyberáme jednotlivých žiakov, ale jednotlivé triedy. V minulosti sa používalo najmä losovanie, teraz náhodný výber z úplného zoznamu tried (škôl) na Slovensku spoľahlivo urobí počítač (napríklad pri medzinárodných meraniach). Teoretickým základom pre všetky tieto operácie je matematický model distribúcie náhodných javov.

Stratifikované výbery zohľadňujú zloženie základného súboru podľa kraja, veľkosti školy a ďalších charakteristík. Stratifikované a reprezentatívne výbery sa najčastejšie používajú pri pilotážach (skúšobných testoch, „predtestoch“).

Na získavanie parciálnych poznatkov o výsledkoch vzdelávacieho procesu sa veľmi často používajú tzv. *typologické vzorky*. Ich isté nadužívanie vyplýva z toho, že sa dajú získať oveľa ľahšie ako štatistické výbery. Teoretickým odôvodnením pre ich použitie je poznatok, že z hľadiska kritérií, ktoré sú školám určené zvonka (Štátny vzdelávací program, platné učebnice), sa triedy aj vo vzdialených školách na seba dosť podobajú. Najmä ak sa nachádzajú v podobnom sociálnom prostredí. To vedie pracovníkov pedagogických inštitúcií k voľbe jednej-dvoch „typických“ škôl, teda škôl, ktoré sa veľmi nelíšia od mnohých iných škôl s podobným sociálnym zaradením, a výskum urobia tam. Výsledky sa zväčša ani neprezentujú ako obraz populačných parametrov (ak sa to urobí, ide o zavádzanie verejnosti), ale len ako informácia, ktorá má predbežne, orientačne potvrdiť či spochybníť užitočnosť vyvíjaných didaktických prostriedkov.

Výber škôl pre nadnárodné testovania vykonáva zadávateľ testu podľa kritérií, ktoré sú pre všetky zúčastnené krajiny rovnaké. Výber vzorky preto nemôže ovplyvniť žiadny administratívny orgán danej krajiny.

3.3. Formulácia výskumnej otázky alebo hypotézy

Každý test je výskumný nástroj. Aby bol skutočným výskumným nástrojom, musí byť skonštruovaný tak, aby odpovedal na výskumnú otázku alebo aby potvrdil či vyvrátil stanovenú hypotézu.

Ak sa vo výsledkoch národných testov, ktoré škola absolvuje, vyskytnú javy, ktoré si učiteľ nevie vysvetliť, napríklad, že žiaci s lepším prospechom v jeho triede dosiahli horšie výsledky ako žiaci s priemerným prospechom, môže sformulovať hypotézy a zostaviť si test na formatívne účely.

Na základe analýzy výsledkov formatívneho testu môže učiteľ zmeniť svoje vyučovacie metódy alebo zdokonaľiť svoj spôsob hodnotenia žiakov.

Hypotéza je jednoznačný, výstižný a stručne formulovaný hypotetický výrok obsahujúci najmenej dve premenné. Je to jeden z prvkov vedeckého bádania, ktorý dáva testovaniu presnosť a rigoróznosť. Hypotéza a výskumná otázka obsahuje odpoveď na otázku, čo chceme testom zistiť.

Pravidlá, podľa ktorých sa hypotézy formulujú, sú veľmi jednoduché. Zlá, nesprávne naformulovaná hypotéza alebo výskumná otázka (v horšom prípade žiadna) je najčastejším nedostatkom väčšiny výskumných prác v pedagogike.⁹⁰

Tvar hypotéz je presne predpísaný:

- hypotéza je oznamovacia veta, ktorá obsahuje dve premenné,
- premenné sa dajú presne zisťovať (merať, kategorizovať),
- jedna premenná má dve roviny (nezávislá), druhá premenná nemá roviny (závislá),
- medzi rovinami jednej z premenných je vzťah, ktorý sa vyjadruje 2. stupňom prídavného mena alebo príslovky (menšie ako, slabšie ako, intenzívnejšie ako, viac ako, menej ako, častejšie ako, vyšší ako, efektívnejší ako a pod.).

⁹⁰ Gavora, P. – Koldeová, L. – Dvorská, D.: *Elektronická učebnica pedagogického výskumu*.
<http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/hypotezy.php?id=i7> (1. 4. 2015)

Príklady hypotéz:

- **Dievčatá** viac využívajú stratégiu **kontroly** ako **chlapci**. (PISA) – *Závislou premennou je tu sklon kontrolovať, ktorú môže ovplyvňovať nezávislá premenná (pohlavie žiakov).*
- Chlapci viac využívajú stratégiu hľadania súvislostí ako dievčatá. (PISA)
- Žiaci si viac pamätajú faktické vedomosti, ktoré využívajú v každodennom živote.
- Žiaci so známku 3 a horšou majú lepšie schopnosti analyzovať, ako žiaci so známku 2 a lepšou.
- Žiaci so známku 2 a lepšou potrebujú na úspešné vyriešenie úloh so súvislým textom viac času ako žiaci so známku 3 a horšou.

Hypotézy formulované pomocou výrazov vyjadrujúcich príčinu a následok sa dajú jednoducho pretransformovať:

III. **Ak** učiteľ zvýši počet pochvál, **tak** sa žiaci viac naučia.

IV. Učitelia, ktorí žiakov viac chvália, naučia žiakov viac ako učitelia, ktorí žiakov chvália menej.

Vo výskume, ktorý zisťuje smer a silu vzťahov medzi premennými (korelácia, súvzťažnosť), sa hypotézy vyjadrujú odlišne:

- Medzi schopnosťou analyzovať a schopnosťou kontrolovať je pozitívny vzťah.

Nie každý výskum je založený na hypotézach. Výskum, ktorý je deskriptívny, a ktorý má len jednu premennú, nemôže mať hypotézy. V takom prípade výskumníkovi stačí formulovať výskumnú otázku:

- Aký je výkon žiakov 9. ročníka v teste z matematiky?
- Aký je výkon žiakov 9. ročníka v tematickom celku *riešenie aplikačných úloh a úloh rozvíjajúcich špecifické matematické myslenie?*
- Ktorí žiaci majú najlepšie výkony v matematike?
- Ako sa naplnili ciele stanovené v ŠVP pre matematiku?

Zdroje pre tvorbu hypotéz a výskumných otázok:

- Odborná literatúra, učebnice, odborné časopisy, zahraničná literatúra, e-zdroje.
- Praktická skúsenosť učiteľa so školou a vedomosti o rodinnom prostredí žiakov.
- Diskusia s odborníkmi, so špecialistami na vzdelávanie a pod.⁹¹

3.4. Stanovenie premenných

Premenná (štatistický znak) je jeden z najdôležitejších prvkov výskumu. Jej náprotivkom v iných vedách je **veličina** (dĺžka, čas, hmotnosť, objem, teplota, el. prúd, cena, natalita atď.).⁹² Je to vlastne predmet výskumu „rozmenený na drobné“. Je to to, o čom chceme zhromažďovať údaje, čo chceme merať.⁹³ Ak nemáme stanovené premenné výskumu, nemôžeme mať ani stanovené vhodné (skúmateľné) výskumné otázky, ani správne hypotézy výskumu.⁹⁴ Pomocou štatistických metód môžeme premenné popisovať, skúmať ich vzájomné vzťahy alebo ich ďalší vývoj.

⁹¹ Gavora, P. – Koldeová, L. – Dvorská, D.: *Elektronická učebnica pedagogického výskumu*. <http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/volba-vyskumnej-temy.php?id=i5> (1. 4. 2015)

⁹² Veličina. <http://cs.wikipedia.org/wiki/Veličina> (1. 4. 2015)

⁹³ Vladimír Dočkal, docent v odbore pedagogická psychológia, navrhuje pojem „meranie“ v humánnych vedách úplne opustiť a nahradiť ho pojmom „odhadovanie“. Pozri: *Dočkal, V.: Meranie v psychológii*. In: Psychologie. Elektronický časopis ČMPS. 2012, roč. 6, č. 4, s. 37. <http://e-psycholog.eu/pdf/dockal.pdf> (19. 1. 2015)

⁹⁴ Gavora, P. – Koldeová, L. – Dvorská, D.: *Elektronická učebnica pedagogického výskumu*. <http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/premenne.php?id=i6> (1. 4. 2015)

Premenné rozlišujeme podľa toho, ako dobre sú merateľné, teda koľko informácií o premennej získame jej meraním.⁹⁵

Druhy premenných podľa charakteru:

1. **Kvantitatívna** (vyčísliteľná) – charakterizujú vlastnosti, ktoré sa vyjadrujú číselne.
2. **Kvalitatívna** (kategorická) – charakterizujú vlastnosti, ktoré sa opisujú slovom alebo definíciou.

Druhy premenných podľa hodnôt:

1. **Diskrétna** (počítateľná) – dosahuje konečný počet presne vyjadriteľných hodnôt.
2. **Kontinuálna** (spojitá) – dosahuje nekonečný počet hodnôt (napr. teplota) alebo sa nedá presne vyjadriť (napr. nálada).

Druhy premenných podľa škály merania:

1. **Neusporiadaná** (nominálna) – meraný jav je zaradený (klasifikovaný) vždy do jednej z konečného počtu skupín či kategórií (pohlavie, národnosť, mesto, škola, trieda a pod.). Poradie skupín sa tu neurčuje, nie je rozhodujúce.
 - a) **Binárna** – špeciálny druh neusporiadanej premennej (s hodnotami 0 a 1, áno – nie a pod.).
2. **Usporiadaná** (ordinálna) – umožňuje hierarchicky usporiadať objekty na poradovej stupnici podľa toho, ktoré majú viac alebo menej kvality reprezentovanej premennou. Ak vezmeme dvoch žiakov dosahujúcich rôzne hodnoty na ordinálnej škále, vieme určiť, ktorý z nich je lepší, ale medzi výsledkami nie sú rovnako veľké intervaly a nevieme odmerať, aký je medzi nimi rozdiel. (Napr. sociálno-ekonomický index rodiny žiaka, poradie šachistov na turnaji a pod.)
3. **Intervalová** (kardinálna) – umožňuje zoradenie objektov, ale aj kvantifikáciu a porovnanie veľkosti rozdielov medzi nimi. Medzi mernými jednotkami sú rovnako veľké intervaly. Intervalové premenné umožňujú nielen určiť poradie, ale aj vyjadriť rozdiel medzi žiakmi.
 - a) **Pomerová** – umožňuje určiť nielen rozdiel, ale aj pomer medzi hodnotami premenných dvoch žiakov, pretože škála pri pomerovej premennej má aj nulovú hodnotu.

Druhy premenných podľa metódy merania:

1. **Latentná** (skrytá, nepozorovateľná, hypotetická) – nepriamo merateľná (napríklad IQ, vzdelanie a pod.).
2. **Manifestná** (zrejmá, zjavná, pozorovateľná) – priamo merateľná.

Druhy premenných podľa funkcie:

1. **Závislá** – (odpovedová, cieľová, kritériová) mení sa pod vplyvom nezávislej premennej (miera závislosti = významnosť vzťahu).
2. **Nezávislá** (explanačná, vysvetľujúca) – spôsobuje zmenu závislej premennej.
3. **Intervenčná** (rušivá) – hypotetická premenná, ktorá ovplyvňuje vzťah medzi nezávislou a závislou premennou, napr. podmienky testovania a pod. Nie je súčasťou daného výskumu.⁹⁶

Premenná nadobúda rôzne hodnoty. Ak je premennou napríklad výkon vo vedomostnom teste, môže táto premenná nadobúdať hodnoty od 0 do povedzme 40 bodov. Ak by bolo premennou napríklad pohlavie žiakov, mohla by táto premenná na Slovensku dosahovať iba dve hodnoty (chlapec a dievča).⁹⁷

Pri stanovení hypotézy z nej vyberieme premenné a stanovíme rozsah očakávaných hodnôt premenných. Pri formulácii výskumnej otázky je meranou premennou napr. výkon žiaka, ktorý by mal

⁹⁵ Litavcová, E.: *Aplikovaná štatistika v príkladoch*. 2002 – 2003.

<http://www.fhvp.unipo.sk/cvt/statistika/stuvod.htm> (1. 4. 2015)

⁹⁶ Proměnná. <https://wikisofia.cz/index.php/Proměnná> (1. 4. 2015)

⁹⁷ V niektorých štátoch EÚ, napríklad v Nemecku, uznávajú aj tretie, tzv. neurčené pohlavie.

zodpovedať počtu získaných bodov v teste, pričom maximálna hodnota takejto premennej je maximálny počet bodov, ktoré sa dajú získať v teste.

3.5. Návrh testovej špecifikácie

Pred začiatkom konštrukcie testu je potrebné stanoviť si podrobnú špecifikáciu testu. Mala by obsahovať najmä informácie o týchto charakteristikách testu:

a) Presnejšie vymedzenie obsahu testu

Keďže obsah, ako ho chápeme vo vyučovaní, nie je totožný s obsahom testu, musíme ich rozlišovať.

Obsah testu stanovujeme obvyčajne napríklad vymenovaním tematických celkov príslušného ŠVP pre konkrétny predmet. Neznamená to však, že všetok obsah uvedený v ŠVP je predmetom didaktického testu. **Obsah** vyučovania stanovený v ŠVP je **prostriedkom**, pomocou ktorého plníme ciele ŠVP, teda je materiálom, na ktorom sa žiaci učia dosahovať výkonové požiadavky stanovené v ŠVP, a **len niektoré zložky obsahu sú určené na zapamätanie**.

Výkon žiakov pozostáva z radu predmetových a nadpredmetových kompetencií, z čiastkových výkonov, zručností, a aj zo zapamätaných vedomostí. Pod obsahom testu teda rozumieme to, čo môžeme merať, nie to, pomocou čoho to môžeme učiť a dosahovať.

Predmetom testu teda môže byť všetko, čo je stanovené v cieľoch ŠVP pre daný stupeň, a to, čo je stanovené v cieľoch, vo výkonových štandardoch a v kompetenčných štandardoch ŠVP pre daný predmet – v závislosti od účelu testu a cieľov testovania (reprezentatívny výber – sumatívny test/čiastkový výber – priebežný test/špeciálny výber – formatívny test atď.).

Test teda môže obsahovať aj dôležité časti **obsahu** výučby, ak je to v cieľoch predmetu alebo ak sa to vyžaduje od žiakov vo výkonovom štandarde ŠVP, alebo aj ak sa vo formatívnom teste chceme zamerať na obsah vyučovania, pre nejaký konkrétny, diagnostický cieľ. Treba však poznamenať, že testy zamerané iba na vedomosti (obsah) majú veľmi nízku reliabilitu, každý žiak si totiž pamätá iné prvky obsahu dobre, a zároveň každý učiteľ dáva dôraz na iné prvky obsahu.

Postoje žiakov a psychomotorické zručnosti nie sú predmetom didaktických testov, i keď niektoré postoje sa dajú merať pomocou špecificky formulovaných testových úloh.

Súčasťou testov určených pre väčšiu vzorku žiakov nebývajú ani tie výkonové požiadavky ŠVP, ktoré sú obmedzené na lokálne alebo regionálne vedomosti žiakov. Takou je napríklad požiadavka v ŠVP pre občiansku náuku ISCED 2: žiak pozná starostu svojej obce. Takéto výkonové požiadavky je však možné zahrnúť do učiteľského testu, ktorý nezasiahne žiakov z inej obce a pod.

Didaktický test má byť obrazom učebného procesu. Neznamená to však, že si zostavíme test, ktorý nebude v súlade so ŠVP. Znamená to, že to, čo sa rozhodneme testovať, by malo byť aspoň podmnožinou toho, čo máme podľa ŠVP spolu so žiakmi dosiahnuť (pozri nižšie o konzistentnosti cieľa).

„Vzdelávanie je vtedy úspešné, ak vieme, k akým cieľom podľa požiadaviek majú žiaci dospieť.“⁹⁸

Výsledky vzdelávania je možné merať len na základe plánovaných výsledkov vzdelávania – t. j. cieľov vzdelávania.

⁹⁸ „Štátny vzdelávací program pre 2. stupeň základnej školy v Slovenskej republike. ISCED 2 – nižšie sekundárne vzdelávanie. Bratislava : ŠPÚ, 2014, s. 24.

Cieľ vzdelávania podľa Pedagogického slovníka vymedzuje:

- účel a zámer vzdelávania,
- výstup, výsledok vzdelávania.⁹⁹

Trendom, zaužívaným aj v ŠVP, je stanoviť **ciele** vo forme dosahovaných **kompetencií** žiakov. Takto chápaný cieľ vzdelávania je plán na relatívne trvalú zmenu osobnosti žiaka, ktorá má byť dosiahnutá vzdelávaním.

Plnenie všeobecných cieľov sa realizuje dosahovaním čiastkových cieľov. Správne stanovené ciele by mali spĺňať požiadavky: konzistentnosti, primeranosti, kontrolovateľnosti, a jednoznačnosti.

Konzistentným nazývame taký cieľ, ktorý má vnútornú väzbu na vyššie, všeobecné ciele vzdelávania. **Cieľ, ktorý nepomáha naplniť hierarchicky vyšší cieľ, alebo dokonca jeho naplneniu bráni, nie je konzistentný.** Do špecifikácie testu vyberáme také výkony, ktoré pomáhajú naplňovať hlavné ciele vzdelávania.

Primeranosťou cieľov rozumieme optimálny vzťah medzi cieľom a reálnymi podmienkami vzdelávania. Znamená to, že merané ciele majú byť **splniteľné**.

Kontrolovateľnosť je vlastnosť cieľa, ktorá umožňuje zisťovať, či bol cieľ splnený alebo nie.

Jednoznačnosť cieľa je daná jeho formuláciou. Jednoznačne stanovený cieľ nemôže byť rôzne chápaný. Ak opis (vymedzenie) cieľa nie je viazaný na obsah vzdelávania, ide o formálne vymedzený cieľ. Hierarchicky usporiadaný klasifikačný model formálne vymedzených cieľov vzdelávania označujeme pojmom *taxonómia cieľov vzdelávania*. (Pozri kapitolu 2.2. Bloomova taxonómia.)

Pre spomenuté vlastnosti pri špecifikácii testu vyberáme hlavne z čiastkových cieľov, ktoré sú jednoznačnejšie a kontrolovateľnejšie, ako všeobecné ciele. Dbáme však na ich konzistentnosť s vyššími cieľmi a na ich primeranosť reálnym podmienkam vyučovania (splniteľnosť).¹⁰⁰

b) Testovací čas

Čas, ktorý si na test vymedzíme, ovplyvňuje druh testu (test rýchlosti/úrovne), počet a obťažnosť úloh, počet, dĺžka a zložitost' východiskových textov, zloženie testovanej skupiny žiakov, typ úloh podľa nárokov na tvorbu odpovede (otvorené/uzavreté, úlohy vyžadujúce rýsovanie a pod.).

Čím viac obťažných úloh, tým viac času musíme na test vymedziť. Pri stanovení času počítame s istou rezervou (okrem testu rýchlosti). Predchádzame tým zvýšeniu tzv. **vynechanosti** (neriešenosti, t. j. vynechávaniu zložitých úloh pre nedostatok času) a tzv. **nedosiahnutosti** (neriešenosti úloh na konci testu pre nedostatok času).

c) Počet a druh úloh

Test sa skladá v viacerých testových úloh, minimálny počet úloh je 10. Na rýchle školské skúšanie postačujú aj krátke testy (10 – 15 úloh). Testy určené na výskum edukačných javov majú mať niekoľko desiatok úloh (40 a viac). S rastúcim počtom úloh totiž stúpa reliabilita testu, t. j. presnosť merania testu. Kým krátke testy s počtom úloh 10 – 12 dosahujú hodnotu reliability cca 0,5, kým profesionálne didaktické testy dosahujú hodnoty reliability aj nad 0,8.

Napríklad pri zostavovaní kritériového testu musíme počítať na každý prah obťažnosti v každom tematickom celku najmenej s piatimi úlohami. Ak si stanovíme 3 stupne obťažnosti a zvolíme si 2 tematické celky, musí náš test obsahovať minimálne 30 úloh.

⁹⁹ Průcha, J. – Walterová, E. – Mareš, J.: *Pedagogický slovník*. Praha : Portál, 2009, 395 s. ISBN 978-80-7367-647-6.

¹⁰⁰ Zásady tvorby didaktického testu. Fakulta informatiky a manažmentu Univerzity v Hradci Králové.

http://fim.uhk.cz/oliva/tvorba_vedeni/rekap-www/m01/m01-03.html (1. 4. 2015)

Určenie počtu úloh testu závisí od viacerých faktorov:

- účel testu (napr. pri prijímačkovom teste musí byť veľa stredne obťažných a náročnejších úloh na dostatočné rozlíšenie najlepších žiakov),
- druh testu (test rýchlosti/úrovne),
- celkový testovací čas,
- počet a dĺžka východiskových textov,
- počet obťažných, časovo náročnejších úloh,
- počet prahov obťažnosti,
- počet testovaných tematických celkov,
- počet kognitívnych úrovní,
- počet uzavretých a otvorených úloh (otvorené úlohy sú časovo náročnejšie na zápis odpovede),
- spôsob interpretácie výsledkov (pozri nižšie).

d) Plánovaný spôsob interpretácie výsledkov

Špecifikácia testu by mala obsahovať aj informáciu o plánovanom spôsobe interpretácie výsledkov. Teda, či to bude test absolútneho (CR) alebo relatívneho výkonu (NR). Zjednodušene sa dá povedať, že NR-testy môžu byť kratšie, teda že môžu obsahovať menej testových úloh, ako CR testy, ktorých výsledky chceme použiť na meranie absolútneho výkonu. Pri meraní absolútneho výkonu totiž musíme do testu zaradiť úlohy, ktoré budú merať výkony vo všetkých cieľových a výkonových požiadavkách stanovených v ŠVP. Pri NR-testoch môžu úlohy celý požadovaný výkon odzrkadľovať „reprezentatívne“, teda výberovo. CR testy preto bývajú väčšinou obmedzené na meranie 1 – 2, nanajvýš tak troch tematických celkov jedného predmetu. (O CR a NR-testoch pozri kap. 2.3.6.)

e) Spôsob skórovania

Skórovanie je vyhodnocovanie testu pomocou pridelovania bodov. Špecifikačná tabuľka má obsahovať, koľko bodov za jednotlivé úlohy žiak môže získať.

f) Príprava interpretačnej škály

Pri formatívnych testoch sa učiteľ musí zaoberať aj chybami žiakov. Aby táto práca bola efektívna a jednoduchá, je potrebné pripraviť si pomocnú interpretačnú škálu. Tá sa obvykle stanovuje už po skúšobnom overení testu. Škála má obsahovať, čo žiak vie, a čo asi nevie, podľa toho, ktoré úlohy alebo skupiny úloh (klastre) vyriešil, podľa toho, ktoré distraktory volil v uzavretých úlohách s výberom odpovede, a podľa toho, aké nesprávne odpovede uviedol v otvorených úlohách s tvorbou odpovede. O žiakovi teda musíme na základe jeho odpovedí vedieť povedať aj čo nevie a v čom robí chybu.¹⁰¹ Napríklad, ak si žiak vybral v konkrétnych úlohách príslušné správne odpovede (4/A, 7/B, 11/A, 18/D), je dobrý v kontrole, a ak si vybral v tých istých úlohách konkrétne distraktory (4/B, 7/A, 11/D, 18/C), má v kontrole veľké nedostatky atď.

g) Určenie poradia testových úloh

Štatistické spracovanie testov ukazuje, že poradie úloh v teste má vplyv na ich riešenosť aj na úspešnosť ich riešenia. Úlohy zaradené na koniec testu môžu ostať nevyriešené, aj keď nie sú obťažné (tzv. nedosiahnuté úlohy). Na počet vyriešených úloh má zas vplyv to, v akom poradí sa v teste nachádzajú. Vyriešenie niektorej úlohy môže žiakovi pomôcť pri riešení ďalšej úlohy, a, naopak, riešenie úloh v inom poradí môže spôsobiť, že žiak niektorú z úloh nevyrieši.

¹⁰¹ Hejného metoda. 12 klíčových principů: 10. Práce s chybou: předcházíme u dětí zbytečnému strachu. <http://www.h-mat.cz/principy/prace-s-chybou> (1. 4. 2015)

Preto treba zvážiť, či existuje medzi úlohami nejaká interferencia, a následne zvážiť, či je vhodné v systéme e-Test zvoliť miešanie poradia úloh, alebo nie. Tiež je potrebné sa zamyslieť nad tým, ktoré úlohy budú na začiatku, a ktoré na konci testu.

Z hľadiska obsahu by úlohy mali nasledovať v premyslenom, logickom poradí. Náhodné usporiadanie úloh obyčajne sťažuje riešenie testu.¹⁰²

Niekoľko ľahších úloh je vhodné umiestniť hneď na začiatku testu – zvyšuje to motiváciu žiakov riešiť test. Ťažšie úlohy zas nie je vhodné umiestňovať až na koniec testu – môžu ostať nevyriešené.¹⁰³

h) Popis testovanej skupiny žiakov

Súčasťou špecifikácie je aj opis skupiny žiakov, ktorú chceme testovať. Dôležité je uviesť aspoň typ školy, vzdelávací stupeň, ročník, a priemerný vek žiakov a približný počet žiakov, ktorých budeme testovať. Najmä pri formatívnych testoch je dôležité uviesť a zvážiť pri zostavovaní testu aj zloženie vzorky z hľadiska pohlavia, zdravotného alebo sociálneho znevýhodnenia, nedostatku výchovných podnetov a pod. Tieto faktory podstatne ovplyvňujú zloženie testu, ak má byť primeraný schopnostiam a veku žiakov, a ak má poskytovať učiteľovi vhodnú spätnú väzbu.

i) Špecifikačná tabuľka

Vytvorenie špecifikačnej tabuľky nie je totožné so špecifikáciou testu. Tabuľka slúži na prehľadné zosumarizovanie stanovených požiadaviek testu a naplánovanie tvorby úloh. Na rozdiel od špecifikácie testu tabuľka obsahuje iba informácie, ktoré sa týkajú samotných úloh.

Špecifikačná tabuľka sa vytvára pre každý test. V záhlaví tabuľky uvádzame typ školy, ročník a predmet, prípadne aj názov testu, ak ide o nadpredmetový test alebo test gramotnosti. Tabuľku vytvárame uvážlivo a tak, aby sme naplánovali test s takými parametrami a na taký účel, ako sme si stanovili v špecifikácii.

Špecifikačná tabuľka býva organizovaná tak, že v riadkoch sa nachádzajú poradové čísla jednotlivých zamýšľaných úloh a v stĺpcoch sa nachádzajú jednotlivé zložky špecifikácie testu. V stĺpcoch pri každej úlohe v tabuľke vyplníme alebo vyznačíme:

- ciele výchovy a vzdelávania, ku ktorým sa vzťahuje úloha (čo chceme zistiť),
- tematický celok a tematická klasifikácia,
- ukazovateľ (čo úlohou meriame),
- výkonová požiadavka (čo žiak má vedieť),
- požadovaná logicko-poznávací operácia (Bloomova taxonómia),
- kontext (druh východiskového textu),
- druh úlohy (otvorená, uzavretá a pod.),
- bodové hodnotenie úlohy atď.

Ak je tvorca úloh nezávislý od zostavovateľa špecifikácie a špecifikačnej tabuľky, vytvára sa pri posudzovaní vytvoreného testu ešte druhá, kontrolná špecifikačná tabuľka, do ktorej sa všetky údaje vyplnia na základe posúdenia úloh. Nakoniec sa výsledná tabuľka porovná s plánovacou špecifikačnou tabuľkou a v prípade zistených rozdielov sa vykonajú korektúry (buď v úlohách alebo v špecifikačnej tabuľke k testu).

¹⁰²Gavora, P. – Koldeová, L. – Dvorská, D.: *Elektronická učebnica pedagogického výskumu*.

<http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/testy-vedomosti/usporiadanie-uloh.php> (1. 4. 2015)

¹⁰³Byčkovský, P. – Zvára, K.: *Konstrukce a analýza testů pro přijímací řízení*. Praha : Univerzita Karlova v Praze, Pedagogická fakulta, 2007, 79 s. ISBN 978-80-7290-331-3.

3.6. Operacionalizácia premenných

Pri meraní vzdelávania sa nedajú všetky premenné zmerať tak jednoducho, ako napríklad vo fyzike. Operacionalizácia je postup, pri ktorom zložité koncepty diagnostikujeme pomocou merateľných príznakov.

Vezmime si napríklad výskumnú hypotézu: Študenti prírodovedných odborov sú pri vyhľadávaní zdrojov v univerzitnej knižnici úspešnejší ako študenti humanitných odborov.

V tejto hypotéze je nezávislou premennou študijný odbor a závislou premennou je úspešnosť vo vyhľadávaní zdrojov.

Na zmeranie prvej premennej nám v dotazníku bude stačiť jediná otázka pýtajúca sa na študijný odbor.

Ako však zmerať „úspešnosť vo vyhľadávaní zdrojov“? Musíme túto premennú operacionalizovať nasledovným postupom: najprv urobíme dekompozíciu premennej –rozložíme ju na čiastkové ukazovatele, potom ku každému ukazovateľu priradíme indikátor (otázku alebo úlohu).

Jedným z ukazovateľov úspešnosti môže byť napríklad čas, za ktorý študent konkrétny zdroj vyhľadá alebo ako priemer časov, za ktoré vyhľadá viacero zadaných zdrojov, počet vyhľadaných zdrojov alebo môžeme za ukazovateľ úspešnosti považovať fakt, že v zozname vyhľadaných zdrojov žiak uvedie konkrétny zdroj (napr. kľúčovú odbornú literatúru).¹⁰⁴

Každú premennú by mala vo výskume reprezentovať celá batéria úloh.

Alebo si vezmime inú hypotézu, napríklad: Žiaci s výborným prospechom z matematiky dokážu riešiť úlohy s kontextom z reálneho života lepšie ako žiaci so slabým prospechom.

Čo je výborný, a čo slabý prospech? Ak neexistujú takéto kritériá kategorizácie prospechu, výskumník si ich musí stanoviť. Takýto prístup nazývame arbitrárny (na základe zväženého rozhodnutia). Napríklad výborný prospech si vymedzíme ako priemer známok menší ako 1,5, slabý prospech ako priemer známok väčší ako 3,5 a pod. Spôsob vymedzenia výskumník vždy uvedie pri špecifikácii testu, aby bola jeho hypotéza jasná a zrejmalá.

3.7. Výber alebo vytvorenie úloh testu a prvé posúdenie ich validity

Ak nedisponujeme bankou úloh, v tejto fáze nasleduje tvorba testových úloh podľa zostavenej špecifikácie (pozri kapitolu 4. Tvorba testových úloh). V prípade, že máme k dispozícii banku úloh, nasleduje výber vhodných úloh do testu z banky úloh podľa špecifikácie.

Pri tvorbe a aj pri výbere úloh do testu sa zaoberáme aj validitou úloh (kurikulárnou normatívnosťou). Pri tvorbe posudzujeme validitu úloh priebežne, pri výbere úloh z banky úloh je potrebné sa validitou úloh zaoberať preto, lebo úloha už nemusí byť v súlade s pedagogickými dokumentmi alebo nie je validná z iných dôvodov.

¹⁰⁴Kabinet informačných štúdií a knihovníctví v Brně: *Učebnice*. Operacionalizace. <http://vyzkumy.knihovna.cz/ucebnice/operacionalizace-typy-promennych> (1. 4. 2015)

3.8. Zostavenie kľúča správnych odpovedí k testu

Dôležitou súčasťou testu je **klúč správnych odpovedí**. Bez kľúča správnych odpovedí je test z hľadiska reliability, spoľahlivosti merania, nepoužiteľný. Nepríjemné je to, že autor kľúča správnych odpovedí prakticky nikdy nedokáže sformulovať zadania otvorených úloh tak, aby existovala iba jediná správna odpoveď, a takmer nikdy nedokáže identifikovať absolútne všetky varianty, ktoré sa vyhodnocovateľom, žiakom, rodičom, príčinným kritikom môžu javiť tiež ako prijateľné.

Pri automatickom vyhodnocovaní otvorených úloh s tvorbou krátkej odpovede, pri ktorej nie je predmetom testovania pravopis, musí autor úlohy uviesť do kľúča správnych odpovedí všetky možné aj nemožné tvary správnej odpovede, aj gramaticky nesprávne, s preklepmi, s diakritikou aj bez nej, s malými aj veľkými písmenami a podobne. Napriek tomu je takéto úlohy často potrebné prehodnotiť aj manuálne. V odpovediach zadaných do elektronického systému zaväzujú každý znak, aj medzera, čiarka, či spojka medzi slovami.

Predbežné použitie testu (pilotáž, predtest) býva pri tvorbe kľúča veľmi prospešné, ale ani ono nedokáže identifikovať všetky možnosti odpovedí, o ktorých sa dá viesť diskusia. V takomto prípade sa vo vyspelých krajinách, v ktorých celoštátne testovanie neslúži ako terč na ventilovanie komplexov ambiciózných laikov, sa takéto prípady riešia celkom pragmaticky: *správne odpovede sú tie, ktoré sú v kľúči*. Ak však pochybnosti napriek tomu pretrvávajú, a nájde sa žalobca, nasleduje súdna dohra.

Jednoduchším problémom sú uzavreté úlohy s výberom odpovede z viacerých možností. Tu ide skôr o to, aby sa varianty mnohokrát kriticky preskúmali a aby sa každý čo len trochu spochybniteľný variant vylúčil, aby do kľúča neprenikla žiadna chyba.

S kľúčom správnych odpovedí je však spojená aj veľká spoločenská zodpovednosť. Kľúč sa musí udržať v tajnosti a uchovávať oddelene od vlastného testu. Aj Slovensko si pamätá verejné pobúrenie, ktoré vzniklo pre neopatrné manipulovanie s kľúčom správnych odpovedí pri maturitnom teste s matematiky.

3.9. Určenie obťažnosti testových úloh

Zastúpenie úloh podľa obťažnosti v teste závisí od typu testu a cieľov testovania. Test musí byť primeraný testovanej skupine žiakov, ich schopnostiam, ktoré vie učiteľ odhadnúť najlepšie (pri celo-slovenských testoch je to ťažšie). Všeobecne platí, že najlepšiu rozlišovaciu schopnosť majú stredne obťažné úlohy, a tých môže byť v teste najviac (okrem výberových – prijímačkových a vyradovacích testov).

Zostavovateľ testu musí najprv uvážiť obťažnosť jednotlivých testových úloh, ktoré by chcel do testu zaradiť. Pri úlohách, ktoré už boli použité v skúšobnom alebo riadnom testovaní, môžeme obťažnosť poznať zo štatistickej analýzy výsledkov testovania. Pritom pri hodnotách obťažnosti získaných v skúšobných testovaniach treba počítať s tým, že sa hodnota môže výrazne zmeniť podľa toho, akú dôležitosť bude hrať výsledok testu v živote žiaka, teda či rozhodne o jeho známke, zaradení do skupiny, alebo či pôjde dokonca o postupový alebo prijímačkový test. Pri testoch s veľkým dopadom na budúcnosť žiaka sa výkon žiakov významne zvyšuje a úlohy, ktoré vykazovali vlastnosti obťažných úloh dopadnú niekedy aj ako ľahké. Ak úlohy nemajú stanovenú obťažnosť, musí učiteľ ako odborník na testovaný predmet úlohy sám vyriešiť a obťažnosť úlohy **odhadnúť**.

V čase, keď J. Čečetka publikoval svoju knihu o testovaní, ešte nejestvovala Bloomova taxonómia poznávacích procesov a preto určovanie obťažnosti hodnotil Čečetka ako veľmi subjektívnu a neobjektívnu metódu. Rovnako hodnotil aj klasifikáciu žiakov v školách: „...učiteľ necíti sa vždy schopným

klasifikovať jemnejším, päťnámkovým systémom, ako o tom svedčia niektoré hlasy učiteľstva, volajúce po zavedení systému len troch známok.“¹⁰⁵

Práve Bloomova taxonómia vniesla do tejto problematiky viac svetla. Určovanie obťažnosti úloh pomocou ich zaradovania do taxonómie podľa najvyššieho poznávacieho procesu, ktorý žiak na vyriešenie úlohy musí použiť, pomáha objektívnejšie určovať obťažnosť testových úloh. Mechanické zaradovanie úloh do úrovni Bloomovej taxonómie podľa slovies bez zamyslenia sa nad postupom riešenia úlohy však môže viesť k ešte väčšej neobjektivite, ako prosté posudzovanie obťažnosti. (Pozri kapitolu 2.2. Bloomova taxonómia.)

Do testu by nemali byť zaradené úlohy, ktoré sú príliš ľahké alebo príliš ťažké (s hodnotou obťažnosti nad 80 a pod 20 %, optimálne ani pod 30 %, ak počítame s nárastom úspešnosti (s poklesom hodnoty obťažnosti) pri teste s veľkým vplyvom na študijnú kariéru žiaka). Ak prevažujú úlohy ľahké alebo ťažké, musí tieto úlohy nahradiť inými a proces stanovenia obťažnosti testu zopakovať. (Pozri kapitolu 2.4.3. Obťažnosť.)

3.10. Stanovenie validity a reliability testu

Pri zostavovaní štandardizovaného testu sa stanovuje štatistický ukazovateľ reliability jednou z možných metód. Pri zostavovaní školských testov sa tento krok pravdepodobne nepodarí zrealizovať. (Pozri kapitolu 2.4.2. Reliabilita.)

Validita testu, hoci ju je možné aj vypočítať (napr. metódou konfirmačnej faktorovej analýzy), sa väčšinou len odhaduje. Odhaduje sa tak, že odborník na daný predmet posúdi, či špecifikácia testu, jednotlivé úlohy, aj test ako celok sú v súlade s platnými, záväznými pedagogickými dokumentmi (Školský zákon, ŠVP).

3.11. Overenie prvotného návrhu testu a úprava testu¹⁰⁶

Skúšobné (pilotné) testovanie testu sa robí na vzorke žiakov, ktorá je ekvivalentná so skupinou žiakov, ktorej je test určený. (O výbere vzorky pozri kapitolu 3.2.)

Na základe výsledkov skúšobného testu môžeme upraviť jednotlivé parametre testu. Sleduje sa aj neriešenosť úloh ale aj tzv. vynechanosť (súvisí s časovou náročnosťou, s obťažnosťou aj s formuláciou úloh). Na základe analýzy výsledkov skúšobného overovania testu a na základe analýzy testu a odpovedí žiakov môžeme zmeniť počet úloh, vymeniť príliš ľahké a príliš ťažké úlohy za iné, môžeme upraviť testovací čas, znovu posúdiť validitu testu a pod. Pri príprave štandardizovaného testu sa v tejto fáze môžu aj upravovať úlohy a kľúče odpovedí, po každej úprave sa však musí zopakovať skúšobné overenie testu.

¹⁰⁵ Čečetka, J.: *Testovanie na školách a jeho štatistické praktikum*. Bratislava : Učiteľské nakladateľstvo slovenské, Družstvo „U nás“, 1934, s. 9.

¹⁰⁶ Gavora, P. – Koldeová, L. – Dvorská, D.: *Elektronická učebnica pedagogického výskumu*.

<http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/testy-vedomosti/kroky-pri-tvorbe-testu.php?id=i15p11> (1. 4. 2015)

3.12. Analýza výsledkov testu a nesprávnych odpovedí

Keď je test hotový, môžeme pristúpiť k jeho administrácii. Po vyhodnotení testu nasleduje analýza výsledkov, stanovenie validity, reliability a ďalších parametrov testu.

Najjednoduchším spôsobom hromadnej analýzy nazbieraných údajov (odpovedí žiakov) je vytvorenie tabuľky početnosti odpovedí.

Odpovede žiakov roztriedime na správne a nesprávne. Pri otvorených úlohách skúmame aj odpovede, ktoré nie sú uvedené v kľúči odpovedí ako správne, pretože niektorí žiaci môžu objaviť originálne správne riešenie úlohy.

V tejto fáze treba dbať na to, aby sme za nesprávne odpovede nepokladali napríklad odpovede s gramatickými chybami a pod., ak úlohou netestujeme pravopis.

Pri analýze výsledkov formatívneho testu **nesmie chýbať ani dôkladná analýza nesprávnych odpovedí** jednotlivých žiakov, ktorú môžeme robiť aj efektívne, ak sa na ňu vopred pripravíme (pozri bod *f* v kapitole 3.5). Zaoberáme sa nielen tým, ktorá úloha má akú obťažnosť, teda koľko percent žiakov si vybralo správnu odpoveď, ale venujeme sa aj nesprávnym odpovediam v otvorených úlohách aj zvoleným distraktorom (pozri kapitolu 4.4. Tvorba, výber a počet distraktorov).

Podrobnejšie sa analýzou analýzy výsledkov testovania zaoberáme v kapitole 5. Štatistická analýza testových úloh.

3.13. Kontrola a úpravy testu

Ku konečnému návrhu testu sa spravidla dostaneme tak, že na základe špecifikácie a špecifikačnej tabuľky (pozri kapitolu 3.5. Návrh testovej špecifikácie) test zostavíme z testových úloh, ktorých vlastnosti už poznáme. Aj v takomto prípade potrebujeme test ako celok skontrolovať. Robíme to viacerými spôsobmi, pričom komplexnosť závisí najmä od účelu testu. Znovu si treba uvedomiť, či ide o test s veľkým vplyvom na žiaka alebo ide o iný druh testu. Pri testoch s veľkým vplyvom na žiaka musíme dbať najmä na objektivitu a reliabilitu. Pri testoch určených na hodnotenie vzdelávania je zvyčajne najdôležitejšia validita testu.

Kontrola testu by mala pozostávať z nasledovných úkonov:

- jazyková korektúra,
- obsahová a formálna oponentúra,
- kontrola jednoznačnosti kľúča správnych odpovedí a návodu pre vyhodnocovanie otvorených úloh,
- kontrola grafickej úpravy (čitateľnosť informácií zadaných obrázkom, grafom a pod.).

Potom nasleduje pilotáž testu, tá je však často z dôvodov utajenia testu nahrádzaná pilotážou jednotlivých úloh. Tu sa dostávame k veľmi dôležitej stránke tvorby testu – k jeho utajeniu. Miera dôležitosti utajenia má priamy vplyv na iné stránky procesu tvorby testu – ovplyvňuje napríklad množstvo expertov zúčastnených na tvorbe, oponentúre a pilotáži testu. Drobná chyba v návode na vyhodnocovanie otvorenej úlohy (opravená až v počiatkovej fáze vyhodnocovania) nie je v porovnaní s dôležitosťou utajenia testu až taká podstatná.

3.14. Tvorba variantov testu

Testy niekedy vytvárame vo viacerých variantoch. Vytvárame ich zmenou poradia úloh a správnych odpovedí a distraktorov. Aby sme mohli dva testy považovať za dva varianty jedného testu, musia mať rovnaké vlastnosti.

Varianty jednotlivých úloh spravidla vytvára autor testu – varianty nezvyknú byť v banke úloh.

Niekedy musíme vytvárať aj varianty testu pre rôzne skupiny žiakov so zdravotným znevýhodnením. Robieva sa to v spolupráci so špeciálnymi pedagógmi a ďalšími odborníkmi, napríklad zmenou veľkosti písma, zvýrazňovaním niektorých slov v zadaniach pre žiakov s poruchami učenia, tvorbou variantu testu v Brailovom písme a pod.

Samostatným problémom je tvorba jazykových mutácií testu. Preklad jednotlivých úloh do iných jazykov je špecifická činnosť, ktorá prekračuje rámec tejto publikácie.

3.15. Vecná a jazyková kontrola testu

Pred použitím testu je potrebné skontrolovať súlad testu ako celku s jeho špecifikačnou tabuľkou a špecifikáciou. Vlastnosti niektorých úloh sa pri použití s inými úlohami môžu zmeniť – tento jav nazývame **interferencia** (vzájomné ovplyvňovanie) úloh. Spravidla sa v rámci kontroly testu vraciame na samý začiatok tvorby testu a kontrolujeme aj súlad testu s cieľmi testovania (kontrolujeme aj samotnú špecifikačnú tabuľku). Hovoríme o posudzovaní vhodnosti a použiteľnosti testu.

Jazyková kontrola testu sa zameriava na jednoznačnosť a jazykovú správnosť formulácií, súlad zadania a distraktorov (alebo návodu na opravovanie otvorených úloh). Samostatne sa kontroluje, či niektoré formulácie nie sú nápovedné (ak nápoveda nebola použitá zámerne). Jazykovej kontrole pomáha, ak testovaný vyučovací predmet má vypracovaný školský slovník používaných pojmov.

3.16. Grafická úprava testu

84

Konečnou fázou prípravy testu je spracovanie jeho grafickej podoby. Zvyčajne si všímame veľkosť a typ písma, formátovanie distraktorov, viditeľnosť inštrukcií, spôsob a formu písania inštrukcií pre žiaka (musí byť jasné, ku ktorým úlohám a ukážkam sa inštrukcie vzťahujú), formátovanie miest pre prácu žiakov a pre odpovede. Samostatne sa kladie dôraz na formátovanie odpovedových hárkov (ak sú použité). Napríklad chýbajúca inštrukcia k spôsobu označenia správnej odpovede môže mať za následok odpútanie žiakovej pozornosti od testovaného obsahu k formálnej stránke testu. Podobne, chýbajúca zmienka o tom, že testová úloha pokračuje aj na ďalšej strane, môže mať za následok vynechanie časti odpovede.

Testy, ktoré majú svoje pevné miesto vo vzdelávacom procese (napr. maturitné), majú zaužívanú grafickú podobu a táto by sa nemala často a radikálne meniť – jej zmena môže mať vplyv na pozornosť žiakov pri práci s testom.

3.17. Overovanie testu

Ak vlastnosti testu overujeme pilotážou na starostlivo vybratej (reprezentatívnej) vzorke žiakov, môžeme pristúpiť k posudzovaniu jednotlivých úloh (testových položiek). Samozrejme, ak chceme, aby takto pilotovaný test bol kvalitnejší ako test bez pilotáže, musíme zabezpečiť utajenie testu. Žiaci (spravidla ani učitelia) nesmú vedieť, aký zámer v budúcnosti s týmto testom máme. V záujme utajenia väčšinou pilotujeme iba jednotlivé úlohy, nie celý test – avšak pilotovaním celého testu vieme získať presnejšie informácie.

Na základe pilotovania testu posudzujeme test ako celok i jeho jednotlivé úlohy. Dôležitou vlastnosťou úloh v teste je ich rozlišovacia schopnosť, teda citlivosť.

Po pilotovaní testu a jeho štatistickom vyhodnotení nasledujú korekcie testu. Korekcie sa realizujú viacerými spôsobmi.

- **Korekcie spôsobu vyhodnocovania** testu robievame po každej pilotáži testu a niekedy aj po ostrom použití testu. Na základe žiackych riešení kontrolujeme jednoznačnosť správnej odpovede v uzavretých úlohách. Pri otvorených úlohách na základe žiackych odpovedí upravujeme formulácie pre hodnotiteľov a starostlivo zvažujeme skórovanie – teda pridelovanie bodov za jednotlivé typy odpovedí.
- **Korekcia samotného testu** na základe pilotovania je veľmi citlivá záležitosť, pri ktorej sa musí dbať na to, aby každá prípadná zmena prešla celým procesom opísaným v tvorbe testu. Zanedbanie niektorého z krokov môže viesť do testu vážnu chybu. Pri väčších korekciách testu je nevyhnutné nové pilotovanie testu pre overenie jeho vlastností.

(Pozri tiež kapitolu 5. Štatistická analýza testových úloh.)

3.18. Vypracovanie manuálu pre zadávateľov testu a hodnotiteľov

Pri tvorbe testových úloh do banky úloh je dobré poznať a mať na pamäti spôsob zadávania testu a všeobecné inštrukcie pre hodnotiteľov. Prinášame iba niekoľko základných otázok:

- Aké informácie a inštrukcie dostanú žiaci a učitelia pripravujúci žiakov v rámci prípravy na testovanie?
- Aké inštrukcie dostanú žiaci pri administrovaní testu?
- Aké iné materiály majú žiaci počas testovania k dispozícii (dáta, súbory vzorcov, kalkulačky, slovníky, prístup na internet a pod.)?
- Bude žiakom navrhnutá stratégia riešenia testu – ak áno, tak aká?

Napríklad:

- Má žiak preskakovať úlohy týkajúce sa istého obsahu, istého formátu?
- Má si žiak pred začatím riešenia celý test prečítať?
- Má si žiak po skončení riešenia všetky odpovede skontrolovať?

Vo fáze prípravy testových úloh je dobré, ak sú známe spôsoby zapisovania správnych odpovedí pre otvorené úlohy. Tieto spôsoby by mali byť v súlade s manuálom pre hodnotiteľov – s časťou manuálu pre hodnotiteľov venovanú jednotlivým úlohám ako aj so všeobecnou časťou manuálu venovanou všeobecným pravidlám posudzovania správnosti odpovede.

4. TVORBA TESTOVÝCH ÚLOH

Zatiaľ čo my, autori testových úloh, sa venujeme tvorbe úloh pre testovanie vo vlastnom vyučovaní predmete, žiaci musia riešiť testy zo všetkých vzdelávacích oblastí. Uvedomujeme si špecifickosť testovania v jednotlivých predmetoch, avšak za dôležité pokladáme zjednotenie niektorých vlastností úloh (parametrov) tak, aby sa pozornosť žiaka mohla upriamovať viac na testované vedomosti, zručnosti a kompetencie a čo najmenej na formu testu. Pokladáme teda za zmysluplné venovať sa niektorým spoločným vlastnostiam procesu tvorby testových úloh.

Pri tvorbe úlohy spravidla začíname **formulovaním úlohovej situácie** alebo s tvorbou či výberom **východiskového textu**. Úlohová situácia zabezpečuje zámerné nasmerovanie pozornosti alebo predstavivosti žiaka na situácie, javy, objekty a pojmy, ku ktorým sa viaže zadanie úlohy. Úlohová situácia, podobne ako východiskový text, dodáva úlohe kontext (niektoré úlohové situácie tento kontext zámerné vynechávajú). Úlohové situácie by nemali žiaka uvádzať k riešeniu nereálnych problémov, nákup sedemdesiatich ananásov, ani k riešeniu problémov neracionálnymi postupmi, napr. delenie koláča pomocou trigonometrických funkcií. Východiskový text alebo úlohová situácia s naformulovanou otázkou alebo zadáním je základom tvorby testovej položky alebo celej testovej úlohy.

Testová úloha má spravidla určitú štruktúru:

1. úvodná informácia a inštrukcie,
2. východiskový text alebo úlohová situácia,¹⁰⁷
3. zadanie úlohy alebo otázka,
4. doplnkové informácie alebo pomôcky,
5. inštrukcie k riešeniu,
6. inštrukcie k odpovedi,
7. štrukturovaná odpoveď alebo priestor na tvorbu odpovede.

Naformulujte úlohovú situáciu súvisiacu s udalosťou, ktorú ste zažili nedávny víkend. Napíšte krátky príbeh, ktorý povedie k formulovaniu úlohy (nastolenie problému). Následne vytvorte jednu úlohu viažucu sa k tejto situácii. Vráťte sa k úlohovej situácii, a ak je to potrebné, tak ju preformulujte tak, aby odpoveď bola jednoznačná. Následne naformulujte správnu odpoveď a tri distraktory (nesprávne odpovede). Ak sa to nedá, vráťte sa k úlohovej situácii a znovu ju upravte.

Pri tvorbe testovej úlohy sme sa zatiaľ ničím neobmedzovali. Avšak za bežných okolností pripravujeme testové úlohy pre použitie v teste s vopred určenými parametrami (špecifikácia testu).

Proces tvorby úlohy si ukážeme na príkladoch.

Úloha 1. Jozef šiel v pondelok ráno do školy. Z domu odišiel o 7.23 h a do školy vošiel o 7.38 h. Určte, ako dlho Jozefovi trvá cesta do školy.

(A) 10 min **(B)** 15 min **(C)** 20 min **(D)** zo zadania sa nedá určiť

Pozrime sa na úlohu 1. Je to uzavretá úloha. Uvedomme si, že Jozefovi v pondelok trvala cesta do školy 15 min, ale napriek tomu odpoveď B nemusí byť správna. Upravte úlohu tak, aby správnou odpoveďou bola možnosť B.

¹⁰⁷V systéme e-Test sa východiskový text nazýva „spoločný obsah“.

Ukážka 1: Galileo Galilei sa narodil v meste Pisa, v Taliansku, v roku 1564. Istý čas študoval medicínu, avšak nadovšetko mal rád matematiku a zákonitosti prírody. Neskôr, keď bol učiteľom na Univerzite v Pise, učil študentov že Aristoteles, starogrécky polyhistor, nemal vždy pravdu. Niektorí kolegovia preto Galileiho kritizovali a robili mu problémy. Galilei dokázal, že telesá s rôznymi hmotnosťami a padajúce z rovnakej výšky dopadajú na povrch zeme rovnakou rýchlosťou. Známy je jeho pokus, pri ktorom pustil z vrcholu šikmej veže v Pise súčasne dve delostrelecké gule, jednu s hmotnosťou 4,5 kg a druhú s hmotnosťou 0,5 kg. Obe dopadli v tom istom okamihu.

Úloha 2.

Na základe ukážky môžeme povedať, že Galileo Galilei žil v druhej polovici _____.

(A) 15. storočia

(B) 16. storočia

(C) 17. storočia

(D) 18. storočia

Úloha 3.

Galileo Galilei a Aristoteles sa _____.

(A) stretávali a diskutovali o zákonitostiach prírody

(B) nestretávali, ale aj tak robil Aristoteles Galileiovi problémy

(C) nemohli stretnúť, avšak Galileo Galilei poznal mnohé z diel Aristotela

(D) nemohli stretnúť, ale aj tak Aristoteles komentoval Galileiho učenie

Pozrime sa na úlohy 2 a 3. Sú to uzavreté úlohy viažuce sa na jeden východiskový text. Úlohy sú formulované nedostatočne – úlohová situácia umožňuje žiakovi považovať za správnu viac ako jednu z ponúknutých možností. Identifikujte chyby v úlohách. Diskutujte o možnosti vytvorenia ďalších úloh viažucich sa k druhej časti textu.

4.1. Princípy tvorby testových úloh

V tejto časti sa venujeme iba testovaniu výsledkov takzvaného konvergentného učenia, teda takého, pri ktorom sa žiaci majú naučiť niečo, o čom žiak dokáže jednoznačne vyhlásiť, že to je alebo nie je pravda.

Porovnajme úlohu 4 s úlohou 5 a všimnime si, či možno jednoznačne rozhodnúť o pravdivosti uvedených výrokov.

88

Úloha 4. Bodka sa v slovenčine píše na konci oznamovacej vety.

Pravda/Nepravda

(Žiaci nadobudli takúto vedomosť a jednoznačne môžeme povedať, že to je pravda.)

Úloha 5. Štebot vrabcov ľudí uspáva.

Pravda/Nepravda

(O tomto výroku sa dá polemizovať, nedá sa jednoznačne povedať, že to je pravda.)

V nasledujúcom príklade sme výrok nevhodný na testovanie upravili na testovú úlohu.

Úloha 6. Rozhodnite, ktorý z nasledujúcich typov elektrární najmenej znečisťuje životné prostredie.

(A) jadrová

(B) tepelná

(C) solárna

(D) veterná

(Ak sa žiak stretol s odpoveďou na túto otázku len nedávno, tak bez váhania vyberie odpoveď prezentovanú učiteľom v procese vyučovania ako správnu. Ak však žiak zapojí vyššie kognitívne zručnosti a začne porovnávať vplyv jednotlivých elektrární na životné prostredie, môže (mal by) prísť k záveru, že jednoznačná odpoveď neexistuje. Táto úloha je príkladom nevhodnej formulácie.)

Testy zvyčajne obsahujú aj všeobecnejšie úlohy, ktorými testujeme úroveň dosiahnutú v nadpredmetových zručnostiach, v kompetenciách. Príkladom takýchto meraní je meranie schopnosti žiaka čítať informácie (zadané textom, tabuľkou, grafom, obrázkom a pod.). Čítanie s porozumením rozvíjame vo všetkých vyučovacích predmetoch a merať mieru schopnosti čítať s porozumením tiež môžeme (aj musíme) vo všetkých vyučovacích predmetoch. Aj pri takto zameraných úlohách musia existovať jednoznačne správne odpovede (to však neznamená, že správna odpoveď musí byť vždy iba jedna – napríklad pri otvorených úlohách alebo pri uzavretých úlohách s viacerými správnymi odpoveďami).

Jednou zo základných úloh tvorcu testových úloh je položiť žiakovi otázku – osloviť ho otázkou, vyzvať ho k činnosti – dať žiakovi úlohu: vyvolať v žiakovi reakciu vedúcu k zodpovedaniu otázky alebo k vyriešeniu zadanej úlohy. Musíme si uvedomiť, že je veľký rozdiel medzi oslovovaním žiakov, ktorých vyučujeme, a medzi oslovovaním žiakov, s ktorými sme sa nikdy nestretli.

Pri písomnom testovaní môžeme formulovať buď **výzvu k činnosti** alebo **priamu otázku**. Používame oba prístupy. Vráťme sa k úlohe 1 a všimnime si možnosť, ako sa dá otázka preformulovať na výzvu k činnosti a naopak. Ktorá z možností je vám prirodzenejšia?

Určte, ako dlho Jozefovi trvala cesta do školy. *(Výzva k činnosti – žiak má niečo určiť.)*

Ako dlho trvala Jozefovi cesta do školy? *(Priamo položená otázka – žiak má odpovedať.)*

Pri formulovaní výziev spravidla používame zaužívané činnostné slovesá a zaužívané formulácie. Tieto musia byť jednoznačné, žiaka nechceme príliš zamestnávať identifikovaním toho, čo má robiť. (Pozri kapitolu 2.2. Bloomova taxonómia.)

- Kategorizujte; navrhňte; upravte; vytvorte; predpovedzte; posúďte; zdôvodnite; obháňte alebo vyvráťte; povedzte; vymenujte; doplňte; opíšte (kto; kedy; kde; ako); definujte; vyberte; označte; zoradte; vysvetlite vlastnými slovami; porovnajte; dajte do vzájomného vzťahu; odhadnite; načrtnite; skontrolujte; pomenujte; vyriešte; vypočítajte; zaradte; navrhňte; naplánujte; zrealizujte; použite; dokážte; vysvetlite; prečo sa deje *to a to!*
- Urobte rozbor; uveďte príklad; nakreslite schému; špecifikujte podmienky; určte obmedzenia; vymenujte problémy; uveďte klady a zápory; vysvetlite príčiny a dôsledky *toho a toho!*
- Rozhodnite; napíšte správu; diskutujte *o tom a tom!*
- Zaujmite stanovisko *k tomu a tomu!*
- Vывodte závery *z toho a toho!*

Testy s veľkým vplyvom na študijnú kariéru žiaka nevyhnutne musia byť pripravované tak objektívne ako sa len dá. Dôraz je kladený na hodnoverné, spoľahlivé bodovanie a spravodlivosť. Toto hľadisko má priamy vplyv na zužovanie testovaných výstupov na tie, ktoré sa dajú jednoduchým spôsobom a hodnoverné, spoľahlivo vyhodnotiť. Vo všeobecnosti tento prístup obmedzuje možnosť hodnotiť

viaceré hodnotné výstupy vzdelávania ako napríklad riešenie problémov a kritické myslenie. Ako autori testových úloh sa snažíme hľadať možnosti na hodnotenie aj takýchto výstupov vzdelávania.

Vo všeobecnosti hovoríme, že testovať môžeme iba to, čo je určené v dokumentoch upravujúcich vzdelávanie. V našom prípade testujeme naplnenie Štátnych vzdelávacích programov (ŠVP) a cieľových požiadaviek. Je však potrebné uvedomiť si, že tieto dokumenty určujú aj vyššie kognitívne úrovne, než je iba zapamätanie si vedomostí. Toto nás oprávňuje tvoriť testové úlohy s kontextom (s obsahom) aj mimo tém určených v ŠVP. Napríklad schopnosti žiakov pracovať s pojmom *percentá* môžeme hodnotiť v kontexte z oblasti účtovníctva alebo chémie, aj keď v ŠVP z matematiky nie je použitá téma z účtovníctva ani z chémie. Najzákladnejším cieľom vzdelávania je žiaka vzdelávať, a vzdelávame ho aj pri hodnotení. Niektoré z úloh testu teda môžu (mali by) obsahovať pre žiaka nové informácie, pre väčšinu žiakov aj zaujímavé, s ktorými majú v rámci riešenia testu pracovať a pri práci s nimi preukázať svoje vedomosti, schopnosti a zručnosti.

Samostatnou otázkou súvisiacou s tvorbou testových úloh sú zdroje informácií a iné pomôcky, ktoré žiakovi počas testovania poskytneme. Značne ovplyvňujú formulovanie testovej úlohy a je potrebné ich brať do úvahy. Medzi takéto pomôcky patria napríklad kalkulačky (s vyresetovanou pamäťou), počítače (s presne vymedzeným obsahom pamäte a s presne vymedzeným spôsobom prístupu k internetu), slovníky, prehľady vzťahov a konštant.

Už na tomto mieste je vhodné sa zamyslieť nad kvalitou testovej úlohy. Jednou z možností je pozrieť sa na vlastnú testovú úlohu očami recenzenta a zodpovedať si nasledovné otázky:

(Zvážte, na ktoré otázky by pri vhodne sformulovanej úlohe mala byť odpoveď nie.)

- Je jasné, čo úloha testuje? Vieme jednoznačne povedať, akú vedomosť žiak nevie alebo ktorú zručnosť nemá žiak, ktorý v danej úlohe zlyhá?
- Je úloha v súlade s cieľmi vzdelávania v príslušnom predmete a na príslušnom stupni/v príslušnom ročníku?
- Vieme sformulovať nosnú myšlienku použitej úlohovej situácie?
- Je úloha sformulovaná v súlade s odporúčaniami pre daný druh testovej úlohy?
- Ak je v úlohe naznačené riešenie, je to urobené úmyselne? (Spravidla sa snažíme vyhnúť naznačeniu správnej odpovede.)
- Je odpoveď, ktorá je správna v použitom modeli, správna aj v realite? (Mnohé školské formulácie predpokladajú použitie v istých zjednodušených modeloch.)
- Nabáda úloha žiaka k uvažovaniu smerom k neexistujúcim problémom?
- Je úloha primerane obťažná?
- Je použitie obrázkov, grafov, tabuliek atď. opodstatnené?
- Sú obrázky, grafy a pod. jasné?
- Sú poznámky na obrázkoch, atď. správne, jednoznačné, stručné a jasné?
- Obsahuje úlohová situácia informácie nepodstatné vo vzťahu k testovaným vedomostiam, schopnostiam alebo zručnostiam?
- Je úloha sformulovaná gramaticky správne?
- Môžu alternatívne interpretácie použitých formulácií viesť k zmene významu?
- Obsahuje úloha dvojitú negáciu alebo nepotrebné vsuvky?
- Je úloha zbytočne sformulovaná pomocou negatívnych pojmov?
- Môže nastať výrazná zmena významu presunutím dôrazu?
- Je úloha jasne sformulovaná a oddelená od úlohovej situácie?
- Obsahuje úloha dostatok informácií potrebných na vyriešenie?
- Poskytuje úloha dostatok informácií o požadovanej forme a hĺbke odpovede?
- Vie žiak, či má byť odpoveď zdôvodnená?
- Sú informácie a úloha zreteľne odlíšené?

- Nezvýhodňuje alebo nediskriminuje úloha alebo text niektoré skupiny žiakov? (Např. pohlavie, región, vierovyznanie a pod.)
- Je grafická stránka úlohy v súlade s odporúčaniami pre daný druh testovej úlohy?

Pri tvorbe testových úloh do banky úloh sa vynára otázka originality testovej úlohy. Tomuto aspektu sa v tejto publikácii venujeme iba okrajovo, a tak si to naznačíme na príklade **chybnej** úlohy. Všímnite si alternatívy tejto úlohy a diskutujte o tom, ktoré alternatívy sú variantom tej istej úlohy a ktoré sú novou úlohou. Závěry sa pri autoroch úloh v jednotlivých predmetoch môžu mierne líšiť.

Úloha 7. Marienka šla v pondelok ráno do školy. Z domu odišla o 7.23 a do školy vošla o 7.38. Určte, ako dlho Marienke trvá cesta do školy.

(A) 10 min (B) 15 min (C) 20 min (D) zo zadania sa nedá určiť

Úloha 8. Jozef šiel v pondelok ráno do školy. Z domu odišiel o 7.15 h a do školy vošiel o 7.30 h. Určte, ako dlho Jozefovi v to ráno trvá cesta do školy.

(A) 10 min (B) 15 min (C) 20 min (D) zo zadania sa nedá určiť

Úloha 9. Marienka šla v pondelok ráno do školy. Z domu odišla o 7.23 h a do školy vošla o 7.38 h. Vypočítajte, ako dlho Marienke trvá cesta do školy. Výsledok uveďte v minútach.

Výpočet:

Výsledok: _____ min

4.2. Druhy a formy testových úloh

V predchádzajúcich častiach sme sa pozorného čitateľa snažili presvedčiť o tom, že základom dobrej testovej úlohy je jasný zámer vhodne vložený do úlohovej situácie (pozri tiež kapitolu 3.6. Operacionalizácia premenných), sformulovaný v podobe zadania a jasnej otázky, prípadne výzvy k činnosti. K vytvoreniu dobrej testovej úlohy nám ešte ostáva upraviť celok do podoby, v akej ho použijeme v teste. Aj pri tomto kroku je niekedy potrebné úlohu mierne upraviť. Možno by sa zdalo, že predchádzajúce časti predstavovali zbytočnú príťaž a že najlepšou cestou pri tvorbe testovej úlohy je začať tu – komplexnejším procesom tvorby, ktorý opisujeme v tejto publikácii, však spravidla pripravujeme kvalitnejšie a náročnejšie úlohy a tiež úlohy, ktoré majú ambíciu nasmerovať vyučovanie želaným smerom (teda menej obvyklé úlohy).

Základné rozdelenie testových úloh:

1. otvorené,
2. uzavreté.

Na nasledujúcich príkladoch si všimnime otvorenosť a uzavretosť úloh (použiteľných aj ako testové úlohy).

Úloha 10. Vysvetlite, kedy sa vo vete pred slovom „alebo“ píše čiarka.

Odpoveď: _____

Podobný príklad vo forme uzavretej úlohy:

Úloha 11. Vo vete sa pred slovom „alebo“ píše čiarka vtedy, keď _____.

- (A) spája viacnásobný vetný člen
- (B) vyčleňuje vloženú vedľajšiu vetu
- (C) ...
- (D) ...
- (E) ...

Príklad preformulovania iným spôsobom do uzavretej úlohy:

Úloha 12. Určte, v ktorej vete chýba čiarka.

- (A) Zbierali ste maliny alebo jahody?
- (B) Pomôžete nám alebo tu budete iba postávať?
- (C) Buď ma poslušneš alebo si to urob sám.
- (D) ...
- (E) ...

(Koľko správnych odpovedí je v tejto úlohe?)

4.2.1. Uzavreté úlohy

V uzavretých úlohách žiak vyberá alebo iným spôsobom určuje správnu odpoveď z ponúknutých možností (štruktúrovaná odpoveď).

A. Binárne (dichotomické) úlohy

Správna odpoveď sa vyberá z dvoch možností:

- áno/nie,
- pravda/nepravda,
- správne/nesprávne,
- súhlasím/nesúhlasím.

Používať takéto úlohy samostatne sa neodporúča. Spravidla ich používame v skupinách. Po vypočutí alebo prečítaní textu môžeme žiakom položiť viac otázok zameraných na meranie toho istého javu. Napríklad päť úloh typu: Peter pracuje ako informatik. *Áno/Nie*

Niekedy pridávame aj tretiu možnosť: *Na základe vypočutého (prečítaného) sa nedá rozhodnúť.*

V systéme eTest nájdeme tento druh úlohy pod označením *True/False*.

B. Úlohy s výberom odpovede z viacerých možností

Úlohy s výberom odpovede môžu byť s jednou správnou a s viacerými nesprávnymi odpoveďami, alebo s viacerými správnymi a viacerými nesprávnymi odpoveďami. Ponúknuté nesprávne odpovede nazývame distraktory. Spravidla používame úlohy s tromi distraktormi a s jednou správnou odpoveďou. Samostatne stojí za zmienku úloha s negatívnym výberom, pri ktorej je úloha formulovaná tak,

že testovaný má z ponúknutých odpovedí vybrať nesprávnu odpoveď. V literatúre sa objavuje aj druh úlohy s výberom najsprávnejšej odpovede.

V systéme eTest nájdeme tento druh úlohy pod označením *Single choice* (jedna správna odpoveď) a *Multiple choice* (viac správnych odpovedí).

C. Priradovacie úlohy

V tomto druhu úloh sa predkladajú žiakom dva stĺpce (dve zoskupenia) výrazov/častí textu/obrázok. Žiak má priradiť správne výrazy/časti textu/obrázky z jedného stĺpca k výrazom/častiam textu/obrázkom v druhom stĺpci.

Úloha 13. Priradte jednotlivé živočíchy k triede, do ktorej patria:

mačka domáca	ryby
lienka sedembodková	cicavce
kapor obyčajný	plazy
krokodíl nílsky	hmyz

V systéme eTest nájdeme tento druh úlohy pod označením *Drag and Drop* alebo *umiestňovacia úloha*.

D. Zoradovacie úlohy

Žiak musí zoradiť príp. usporiadať predložené slová alebo časti textu podľa určitého princípu, podľa nejakej štruktúry, podľa chronológie, súvislosti, logiky a pod. Môžeme ich nazvať aj úlohami s násobným výberom alebo doplňovaním. Prvý člen býva niekedy už určený, aby nedošlo k zámene inak správneho poradia (zhora – nadol/zdola – nahor). V literatúre býva tento druh aj poddruhom priradovacej úlohy, pretože sa tu k jednotlivým prvkom vlastne priraduje poradie. Pri tomto druhu úloh je potrebné precízne naformulovať inštrukcie a zadanie úlohy, aby žiak vedel, ako presne má jednotlivé prvky zoradiť.

Úloha 14. Zoradte obdobia podľa chronológie od najstaršieho po najmladšie:

neolit, mezolit, doba železná, doba bronzová

Najstaršie:	1. paleolit
	2.
	3.
	4.
Najmladšie:	5.

V systéme eTest nájdeme tento druh úlohy pod označením *Ordering*.

E. Úlohy na dopĺňanie textu, doplňovacie úlohy:

Časť vety (na začiatku, vo vnútri alebo na konci) je vynechaná, žiaci na základe kontextu text dopĺňajú ponúkaným slovom, vytvoreným tvarom slova, alebo slovným spojením.

Úloha 15. V nasledujúcom texte doplňte vynechané slová:

Nevyhnutnou východiskovou podmienkou ____ vývoj testu ____ jeho modifikácie ____ teoretické odôvodnenie konštruktu, ktorý chceme merať, a rozhodnutie, na ____ účel – predikačný alebo explanačný – bude slúžiť. Konštruktom ____ myslí celok znakov, o ktorých sa ____ uvažovať ako o izolovane pozorovateľných dimenziách správania. Jeho pochopenie vychádza ____ presnej a premyslenej konceptuálnej definície, ktorá vyžaduje operacionalizáciu, prevedenie ____ súbor úloh, ktoré sa majú riešiť, alebo otázok, ktoré majú byť zodpovedané.¹⁰⁸

Ak v tomto príklade ponúkame slová na doplnenie, je to úloha uzavretá. Môžeme ponúknuť aj viac slov, ako je vynechaných, napríklad vynecháme 10 slov a na doplnenie ponúkame 15 slov. Ak tieto slová neponúkame, môžeme ju pokladať za otvorenú (niekedy takéto úlohy nazývame aj polootvorené).

V systéme eTest nájdeme tento druh úlohy pod označením *Custom fill*.

4.2.2. Otvorené úlohy

Otvorené úlohy sú produktívne úlohy, ktoré si vyžadujú od žiakov krátke alebo dlhšie odpovede (žiak má niečo napísať, nakresliť, načrtnúť). Otvorené úlohy zahŕňajú doplňovacie úlohy, úlohy s tvorbou krátkej odpovede, úlohy s tvorbou dlhšej odpovede a úlohy s tvorbou slohových útvarov.¹⁰⁹

A. Otvorené úlohy s tvorbou krátkej odpovede¹¹⁰

Príklady na úlohy s tvorbou krátkej odpovede:

Úloha 16. V ktorom roku bola založená Univerzita Komenského?

Odpoveď: _____

Úloha 17. Prečo sa dievča bálo?¹¹¹

Odpoveď: _____

Úloha 18. Vypočítajte dĺžku prepony pravouhlého trojuholníka, ak poznáte veľkosti odvesien 12,4 cm a 23,6 cm. Výsledok vyjadrite v centimetroch a zaokrúhlite na desatiny.

Výpočet:

Výsledok: _____ cm

V systéme eTest nájdeme tento druh úlohy pod označením *Fill*.

B. Otvorené úlohy s tvorbou dlhšej odpovede

Príklad úloh s tvorbou dlhšej odpovede (úlohy 20 – 22).

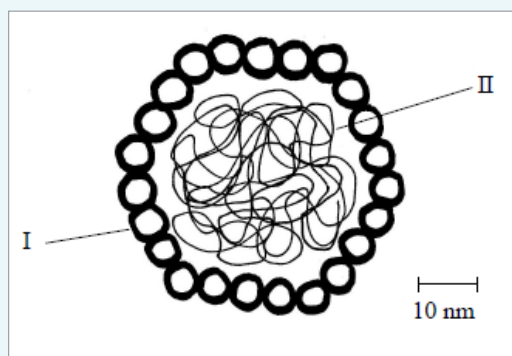
¹⁰⁸Papica, J.: *Psychometrické metody ke konstrukci testu a sestavování testových sistemů*. Olomouc : Univerzita Palackého, 1987, s. 4.

¹⁰⁹Úlohám s tvorbou slohových útvarov sa v tejto časti bližšie nevenujeme.

¹¹⁰Na ich označenie sa používa aj zaužívaná skratka ÚKO.

¹¹¹Úloha sa viaže na textovú ukážku.

Ukážka 2: Náčrt vírusu



Úloha 19. Identifikujte štruktúry označené I a II.

Štruktúra I:

Štruktúra II:

Úloha 20. Použitím mierky na obrázku určte priemer vírusu. Svoj postup vysvetlite.

.....

Úloha 21. V krátkosti vysvetlite, prečo sú antibiotiká účinné proti baktériám a nie proti vírusom.

.....

Úloha 22. Vysvetlite proces, ktorým sa baktérie stávajú odolnými proti antibiotikám.

.....

Systém eTest umožňuje aj využitie druhu úlohy s tvorbou slohových útvarov. V systéme eTest nájdeme tento druh úlohy pod označením *File* (Úloha s odovzdaním riešenia v súbore). Žiak úlohu vypracuje v textovom alebo inom editore a nahrá riešenie v podobe súboru do systému eTest. Tento druh úlohy sa však dá využiť aj na zložitejšie výpočty, pri ktorých hodnotíme aj postup, alebo na iné tvorivé úlohy.

Rozdelenie testových úloh na otvorené a uzavreté úzko súvisí aj s procesom vyhodnocovania žiacich odpovedí. Pri **uzavretých** úlohách autor úlohy určuje, ktorá z ponúknutých odpovedí je správna. Odpovede na uzavreté úlohy môže vyhodnocovať stroj (napr. počítač) alebo laik. Pri **otvorených** úlohách autor úlohy tvorí inštrukcie na vyhodnocovanie odpovedí. Úlohy s tvorbou krátkej odpovede spravidla môže hodnotiť laik, potrebuje však úplnú, presnú a jednoznačnú inštrukciu na vyhodnocovanie. Úlohy s tvorbou dlhšej odpovede musí hodnotiť expert, ale aj v takomto prípade, v záujme objektívnosti, musí dodržiavať pomerne presne vypracované inštrukcie na vyhodnocovanie.

Inštrukcie pre hodnotiteľov na vyhodnocovanie žiacich odpovedí sú súčasťou tvorby testovej úlohy, ale neskôr, v procese tvorby testu, sa môžu upravovať. Upravujú sa tiež v procese pilotného overovania testu a tiež po administrovaní testu na základe vybranej vzorky žiacich odpovedí.

Ďalším prístupom ku klasifikácii testov je jeho materiálne spracovanie. Testy poznáme:

- ústne (štruktúrovaný, alebo pološtruktúrovaný rozhovor),
- písomné (v tlačenej forme),
- písomné (v elektronickej forme, napr. najnovšie už aj testové moduly na riešenie problémov IEA PISA),
- praktické (testovaný realizuje praktickú činnosť),
- praktické (v elektronickej forme; najmä, ale nie len v informatike).

4.2.3. Ďalšie formy testových úloh v systéme eTest

Okrem spomenutých druhov úloh umožňuje elektronický testovací systém eTest, vyvinutý v projekte *Zvyšovanie kvality vzdelávania na základných a stredných školách s využitím elektronického testovania*, vytváranie aj ďalších foriem úloh.

Špeciálnou formou, ktorú umožňuje systém eTest, je tzv. matica. V systéme eTest nájdeme túto formu úlohy pod označením *Single matrix* (jedna správna odpoveď v riadku) a *Multiple matrix* (viac správnych odpovedí v riadku).

Do tejto formy je možné transformovať viaceré druhy úloh:

- zloženú úlohu s výberom odpovede z viacerých možností,
- priraďovaciu úlohu,
- zoradovaciu úlohu
- zloženú dichotomickú (binárnu) úlohu.

V prvom prípade je to zložená úloha pozostávajúca z viacerých čiastkových úloh s výberom odpovede z viacerých možností, pričom čiastkové úlohy sú zoradené pod sebou v riadkoch a možnosti odpovede sú v stĺpcoch. Pri tomto druhu úlohy musia byť možnosti v stĺpcoch vybrané tak, aby žiak uvažoval so všetkými možnosťami pri každej čiastkovej úlohe (jedna možnosť môže byť správnou odpoveďou vo viacerých čiastkových úlohách).

Úloha 23. Priradiť k členským a partnerským krajinám OECD kontinent/kontinenty, na ktorom/na ktorých sa rozprestierajú. (Niektoré krajiny sa rozprestierajú na viacerých kontinentoch – tzv. transkontinentálne krajiny.)

	Európa	Ázia	Afrika	Amerika	Austrália	Antarktída
Austrália	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Gruzínsko	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mexiko	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Turecko	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Panama	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kazachstan	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ruská federácia	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

V prípade, že chceme do formy matice pretransformovať priraďovaciu úlohu, je postup nasledovný: do riadkov umiestnime vždy prvú časť rozdelenej dvojice a do stĺpcov druhú časť.

Úloha 24. Priradte jednotlivé živočíchy k triede, do ktorej patria:

	Ryby	Cicavce	Hmyz	Plazy
Mačka domáca	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komár dotieravý	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Kapor obyčajný	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Krokodíl nílsky	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

V prípade zoradovacej úlohy ponúkame v stĺpcoch poradové čísla. Pri dichotomickej (binárnej) úlohe ponúkame v stĺpcoch možnosti: Pravda/Nepravda a pod.

Systém eTest umožňuje aj vytváranie úloh, pri ktorých žiak jednoduchým kliknutím označí správnu odpoveď buď v texte alebo na obrázku. V prípade kliknutia do textu sa vybrané slovo/slová podfarbia a odpoveď sa tak zaznamená a vyhodnotí automaticky. V prípade označenia miesta alebo časti na obrázku žiak presunie označovací symbol na vybrané miesto a odpoveď sa zaznamená a vyhodnotí automaticky. V systéme eTest nájdeme tento druh úlohy pod označením *Marking text (Úloha s označením odpovede v texte)*, v prípade obrázkov je to označenie *Hotspot (Úloha s označením odpovede v objekte)*.

4.3. Východiskový text

V jednotlivých vyučovacích predmetoch rozvíjame okrem všeobecných aj špecifické schopnosti, a to aj v oblasti komunikácie. Pri tvorbe testových úloh využívame v zadaniach najrôznejšie texty, súvislé aj nesúvislé, napríklad v matematike a fyzike sú to vo veľkej miere rôzne vzťahy, náčrty, obrazce, tabuľky, grafy, v biológii sú to najmä popísané schematické obrázky a v jazykoch najviac využívame ukážky súvislých textov v písanej i zvukovej podobe, dávame dôraz aj na gramatiku a štylistiku. Tieto špecifiká sa prejavujú aj v testových úlohách.

Úlohy v teste nebývajú osamotené, ale obyčajne tvoria tematický sektor, súbor úloh spojených určitým tematickým okruhom. Tieto sektory môžu byť nepravidelné, s nerovnakým počtom úloh, alebo sa môžu vyznačovať systematickosťou. Napríklad didaktický test pre externú maturitu zo slovenského jazyka a literatúry sa vždy skladá zo 64 úloh rozdelených do 6 sektorov, každý s ôsmimi úlohami. Takéto usporiadanie úloh je jedným z aspektov štandardizácie didaktického testu.

Zjednocujúcim prostriedkom pre jeden tematický sektor je:

- a) východisková/úlohová situácia,
- b) východiskový text.

V skutočnosti v oboch prípadoch ide o text, no medzi nimi je určitý rozdiel. Text úlohovej situácie vytvára tvorca testu (resp. tvorca príslušného sektora) ako špecifický nástroj obsahujúci všetky potrebné informácie, ktoré sa premietnu do testových úloh a bez ktorých úlohy nie sú riešiteľné (pozri kapitolu 4. Tvorba testových úloh).

Je dôležité si uvedomiť, že východiskový text má v teste funkciu spúšťačieho stimulu pre žiakovu odpoveď. Obyčajne sa nevzťahuje na jednotlivú úlohu, ale na celý test či sektor testu. Z toho vyplýva pre autora veľká zodpovednosť pri jeho hľadaní a spracovaní. V systéme e-Test sa takýto východiskový text nazýva „spoločný obsah“.

4.3.1. Druhy textov podľa pôvodu

Východiskové texty posudzujeme z viacerých hľadísk. Podľa pôvodu rozlišujeme východiskové texty:

- a) prevzaté,
- b) vytvorené,
- c) zmiešané.

Prevzatý východiskový text je originálny, autorizovaný text, ktorý tvorca testu prevezme z nejakého *verejného zdroja*. Zdrojom nemusia byť len umelecké, populárne či odborné diela, ale rozličné druhy písaných materiálov rôznych žánrov pochádzajúcich z rôznych oblastí ľudského života. Prevzatý východiskový text nie je zámerne vytvorený pre potreby testových úloh, takže musí byť prácne vyhľadaný a starostlivo vybraný. Z tejto skutočnosti vyplývajú takmer všetky ťažkosti a problémy, ktoré sa s voľbou prevzatého východiskového textu spájajú.

Pre výber takéhoto textu platí jedno zásadné pravidlo: východiskový text sa vyberá účelovo, s prihliadnutím na ciele plánovaných úloh, ktoré má zastrešiť, pretože celý projekt tvorby testu je účelový proces a riadi sa vopred stanovenými cieľmi. V žiadnom prípade nemôže vzniknúť dobrý test (sektor) tak, že sa zvolí *akýkoľvek* text, a potom sa v ňom sekundárne hľadajú námety na úlohy. Východiskový text musí byť už vybraný tak, aby obsahoval všetky možnosti pre tvorbu úloh, ktoré vyplývajú zo špecifikácie pre daný didaktický test. To je hlavná podmienka úspechu, ale aj najväčšia ťažkosť či obmedzenie pri jeho hľadaní.

Texty z učebníc, cvičebníc a čítaniek používaných pri vyučovaní, s ktorými už žiaci pracovali alebo mohli pracovať, nie sú vhodné na použitie v didaktickom teste. Takéto texty je možné použiť iba v prípade, ak ide napríklad o inšpekčný test a pod.

Vytvorený východiskový text vytvára autor úlohy. Je zvyčajne kratší ako prevzatý text a nezvykne sa naň viazať viac, ako jedna úloha. Môže ho tvoriť jedno slovo, slovné spojenie, skupina slov, viacero slovných spojení, veta, súvetie, menší odsek alebo rozsahom menší nesúvislý text.

Zmiešaný východiskový text pozostáva z prevzatého textu doplneného o vytvorený text (napr. o vysvetlenie kontextu alebo dopĺňujúcu informáciu), pričom sú tieto texty graficky odčlenené.

4.3.2. Druhy textov podľa typu

Z *typologického* hľadiska rozoznávame nasledovné druhy textov, ktoré sa môžu uplatniť v teste:

- a) súvislé,
- b) nesúvislé,
- c) kombinované,
- d) zložené.

Súvislý text je vo svojej podstate naratívny text, vyjadrený prirodzeným jazykom a vyznačujúci sa tematickou, štylistickou a gramatickou jednotnosťou a ucelenosťou. Bývajú to texty vybrané obvykle z umeleckých, publicistických a odborných diel. Aj keď sú tieto texty skoro vždy úryvkami z oveľa väčšieho základného celku, prinášajú jednu ústrednú myšlienku alebo predkladajú jeden vecný problém, ktorý je v texte zreteľne viditeľný.

V *nesúvislých textoch* je zložka vyjadrená prirodzeným jazykom vo výraznej menšine (popisy obrázkov, máp, hlavičky tabuliek, označenie osí v grafe atď.), poprípade sa nevyskytuje vôbec.

V tejto kategórii sa vyskytujú aj nesúvislé texty, ktoré nepoužívajú symbolický či technický jazyk, ale ich nesúvislosť je daná ich špecifickou kompozíciou, resp. osobitným rozmiestnením textu na stránke.

Tento typ nesúvislých textov je príznačný najmä pre administratívnu oblasť (žiadosti, formuláre, objednávky, dotazníky, zápisnice).

Kombinované texty sú kombináciou oboch predchádzajúcich spôsobov. Časť textu je napísaná súvislým prirodzeným jazykom, časť textu tvorí symbolický jazyk alebo grafické vyobrazenia.

Zložené texty sú zoskupenia viacerých rovnorodých textov (súvislých alebo nesúvislých), napríklad niekoľko tabuliek alebo niekoľko krátkych správ.

Nie je potrebné podrobne vysvetľovať, že každý typ textu si vyžaduje inú čitateľskú kompetenciu, a to tak z hľadiska porozumenia jeho obsahu, ako aj z hľadiska techniky čítania. Spôsobilosť vedieť čítať príslušný typ textu tvorí vo väčšine vyučovacích predmetov súčasť obsahu vzdelávania, a preto sa jej ovládanie spravidla zahŕňa do cieľa skúšky. Na druhej strane sa však čítanie v škole považuje za nadpredmetovú kompetenciu, takže schopnosť čítať napríklad štatistickú tabuľku môže byť zahrnutá aj do takých predmetov, ako je slovenský jazyk a literatúra alebo dejepis, i keď primárne vychádza z učiva matematiky či prírodovedných predmetov.

4.3.3. Implicitné a explicitné informácie v texte

Čítanie s porozumením je v prvej fáze vždy *analýzou textu a extrahovaním informácií* z neho. Autor testových úloh si pri voľbe východiskového textu musí ujasniť, či považuje čítanie textu za inherentnú zložku skúšania, alebo mu text slúži len ako informačná základňa. V tomto druhom prípade volí text, v ktorom sú informácie *viditeľne zobrazené* a žiak ich má iba nájsť, vyhodnotiť a zahrnúť do odpovede. Často ide o fakty, pomenovania, číselné údaje a pod. Ak žiak vie, ktorý údaj či fakt potrebuje, vyjme ho z východiskového textu a potom s ním pracuje podľa pokynov v zadaní úlohy.

Ak však učiteľ chce čitateľskú analýzu textu *zrovnoprávniť* s konečným výsledkom a úspešnosť riešenia úlohy tak vlastne podmieniť kvalitatívnou úrovňou čítania východiskového textu, predostrie text, v ktorom potrebné informácie *nie sú priamo vyjadrené*. Napríklad v matematike a prírodovedných predmetoch musí žiak z uvedených údajov najprv vypočítať podľa priameho pokynu alebo na základe v škole osvojenej stratégie (matematický postup) správny výsledok a až potom ho podrobiť záverečnej interpretácii.

V testoch pre spoločenskovedné predmety sa často od žiaka očakáva *vyhľadanie vnútorných vzťahov* medzi explicitnými informáciami, vytvorenie deduktívnych záverov z predpokladov uvedených v texte a pod. V takýchto prípadoch až súhra správneho prečítania textu a jeho následného spracovanie podľa pokynov zadania privedie žiaka k úspešnému riešeniu úlohy.

4.3.4. Rozsah a obsahová náročnosť textu

Mechanicky uplatniteľné pravidlo, podľa ktorého treba zvoliť dĺžku textu, nejestvuje. Ak ide o súvislý naratívny text, mal by byť taký dlhý, aby čitateľ mohol pochopiť jeho zmysel, podstatu jeho výpovede, hlavnú ideu a pod. To sa spravidla nedá dosiahnuť pri texte, ktorý by obsahoval povedzme dve-tri krátke vety. Vo všeobecnosti však možno vysloviť pravidlo, že čím viac úloh bude od východiskového textu odvodených, tým dlhší by text mal byť.

Na druhej strane však musí autor testových úloh dozrieť na to, aby čítanie textu nevyžadovalo neprimerane dlhý čas. Jednak tým rastie čas potrebný na riešenie sektoru (testu), jednak rýchlo narastá únava a klesá pozornosť žiaka. A práve koncentrácia pozornosti na dôkladne prečítanie východiskového textu je kritickou podmienkou úspešnosti pri riešení úlohy.

Dĺžka textu nie je jediný parameter, ktorý usúvzťažňujeme s počtom úloh. Potenciálnou možnosťou skracovania východiskového textu je zvyšovanie jeho *informačnej hustoty*. Lenže tento spôsob informačného obohacovania textu nesmie dosiahnuť extrémnu mieru, pretože presýtený text vyvolá v žiakovi tzv. informačný stres. Východiskový text teda má byť vyvážený: nie je dobré, ak na využiteľné informácie pripadne neúmerne množstvo redundantného, nadbytočného textu, lebo to žiaka rozptyľuje a unavuje, ale nie je vhodný ani krajný opak, keď je text informačne preťažený, hoci sa to neskúsenému autorovi môže zdať výhodné pre tvorbu úloh.

Doteraz povedané sa vzťahuje prednostne na súvislé texty v prirodzenom jazyku. Základnou vlastnosťou nesúvislých textov – ako sme už uviedli – je radikálna redukcia komunikačných prvkov na faktické informačné jednotky. Texty vyjadrené v symbolickom jazyku prakticky neobsahujú žiadne alebo skoro žiadne sprostredkujúce znaky a skladajú sa takmer výlučne zo znakov-informácií. V tejto úspornosti spočíva veľká komunikačná prednosť takýchto textov, zároveň však môžu byť tieto texty práve z uvedeného dôvodu čitateľsky obťažné. Osobitne treba zdôrazniť, že v takýchto „čistých“ informačných textoch akákoľvek chyba či vynechaný znak môžu byť absolútnou prekážkou pre správne vyriešenie zadanej úlohy. Preto ich treba mnohokrát preskúmať a overiť v testovacej praxi, kým sa stanú pripravené na zaradenie do definitívneho testu.

4.3.5. Formatívne aspekty východiskového textu

Technike testovania vedomostí a zručností sa dosť často vyčíta určitá edukačná sterilita, akoby sa fáza kontroly učebných výsledkov žiaka, vykonaná touto formou písomnej skúšky, vymkynala z celkového charakteru vyučovania. Školské vzdelávanie má byť – a v mnohých pozitívnych prípadoch aj je – dynamickou jednotou *informačných a formatívnych procesov*, ktorých výsledkom má byť komplexne rozvinutý žiak. Ale pri tvorbe testu sa tento zreteľ často neberie do úvahy a testovanie sa vníma ako statický *rezultatívny akt*, inštalovaný už mimo vlastného vzdelávacieho procesu.

Treba poznamenať, že niekedy to tak naozaj je, niektoré formy testovaní sa dejú už mimo rámca vyučovania a neraz až po jeho úplnom ukončení. Koniec-koncov celá maturitná skúška má tento exterritoriálny ráz. Vo všeobecnosti sa to akceptuje ako prirodzená vec pri výstupných skúškach, teda počas sumatívnej kontroly na konci vzdelávacieho obdobia.

Pri testovaní, ktoré prebieha vnútri vzdelávacieho programu a ktoré je súčasťou tzv. formatívnej kontroly učebných výsledkov, by takéto videnie neobstálo a súčasná pedagogika to aj veľmi dôkladne skúma a pripomína. Samotný proces skúšania v takomto prípade má okrem výpovednej examinačnej funkcie aj funkciu rozvoja žiakovho poznávania; nejde teda iba o štatisticky zistiteľnú mieru žiakových aktuálnych schopností, ale aj o ich ďalšie rozvíjanie. Žiak odpovedá, ale zároveň sa aj učí.

Ak teda pri testovaní berieme do úvahy rozvíjanie vedomostnej úrovne žiaka, a to najmä jej kognitívno-operačnú stránku, niet dôvodu na to, aby zo zreteľa tvorcov testu vypadla aj zainteresovanosť na rozvíjaní ostatných zložiek osobnosti žiaka. Tento zámer sa môže premietnuť napríklad aj do výberu východiskových textov. Tieto texty by sa mali žiakovi prihovárať aj po stránke mravnosti, medziludskej otvorenosti a empatie, národného uvedomenia či sociálnej žičlivosti a pod.

Zatiaľ sme tento aspekt spomínali najmä v negatívnej rovine: východiskový text a následne ani odvozené testové úlohy nesmú obsahovať žiadne hodnotové výroky, ktoré naša spoločnosť, náš národ a štát považujú za nezlučiteľné so svojou občianskou podstatou, resp. sú v priamom rozpore s ústavou a zákonmi Slovenskej republiky. Ide najmä o rôzne otvorené či skryté prvky diskriminácie, o ktorých sme písali v kapitole o etike pedagogickej kontroly v škole (pozri kapitolu 1.5).

Ale výchovný aspekt pri realizácii vnútornej skúšky sa môže prejavíť aj v menej extrémnych súvislostiach a v pozitívnom zmysle. Pri voľbe východiskového textu by sa tvorca mal riadiť aj hľadiskom hodnotovej naplnenosti textu. Nejde o to, aby všetky texty obsahovali krikľavé školometské ponaučenia,

ale aby z nich k žiakovi vyžarovala pohoda, vyrovnanosť, „pozitívna energia“. Tieto aspekty sa prirodzene nebudú premietat do obsahu skúšky (s výnimkou prípadu, keď tieto otázky sú predmetom vzdelávania a skúšania, napríklad v dejepise alebo v občianskej náuke), ale môžu sa priaznivo podpísať na atmosfére skúšky a vnútornom naladení žiaka počas vypätej skúšobnej situácie.

4.3.6. Intaktnosť prevzatého textu a autorské práva

Každý prevzatý text má nejakého autora a každý uverejnený text je chránený podľa legislatívy o ochrane autorských práv. Iba pri denníkoch a rôznych drobných tlačovinách (reklamné plagáty) sa obyčajne neuvádza pôvodca a text alebo vôbec nie je autorsky chránený, alebo práva prechádzajú na majiteľa a vydavateľa produktu, napríklad denníka, rozhlasu či televízie.

Ostatné uverejnené texty knižného a časopiseckého charakteru, v súčasnosti už aj autorizované internetové súbory, sú chránené autorským zákonom. Keď tvorca testovej úlohy vyberie z takého prameňa určitý text alebo jeho časť (úryvok), musí rešpektovať jeho identitu, čo znamená:

- a) nemeniť znenie textu,
- b) uviesť autora textu, názov textu a prameň, z ktorého čerpal.

Nedotknuteľnosť použitého textu je daná autorským zákonom (§ 17, odsek 1, písmeno d zákona č. 618/2003 Z. z.). Zakázané sú akékoľvek nedovolené zmeny a zásahy. Práva na nedotknuteľnosť svojho diela sa autor podľa zákona nemôže vzdať.

Text teda nemožno pozmeňovať, možno ho ale napríklad doplniť o vysvetlivky. V pedagogickej praxi to môže znamenať napríklad vysvetlenie cudzieho slova slovenským ekvivalentom alebo zriedkavého či zastaraného slova v zátvorkách, ďalej napríklad doplnenie textu o konektor (časticu, odkazovacie zámeno), ak sa vo zvolenom texte prejavuje logická súvislosť či tematická odkázanosť na predchádzajúcu časť pôvodného textu, ktorá nebola do úryvku zahrnutá. Pri textoch v cudzích jazykoch je možné slová, ktoré nespádajú do danej úrovne podľa Spoločného európskeho referenčného rámca vysvetliť v slovníčku pod textom alebo v poznámke pod čiarou.

Pri textoch, ktoré pre účely testovej úlohy autor takto upraví, uvedie za bibliografickým odkazom poznámku: (Upravené).

Nijako však nemožno odporúčať vypustenie slova alebo celej pasáže v súvislosti s etickými ohľadmi, na ktoré sme upozornili vyššie. V takomto prípade by išlo o zásah do *zmyslu* textu. Ak autor testu usúdi, že niektorá časť textu nie je pre test vhodná, musí sa vzdať celého textu a nepokúšať sa ho zmeniť podľa vlastných predstáv.

4.3.7. Spôsob citovania pri prevzatých východiskových textoch

Pri úlohách, v ktorých zadaniach citujeme cudzí text (či už súvislý alebo nesúvislý), musí autor úlohy, ktorý ukážku vybral, podľa autorského zákona¹¹² uviesť povinné údaje o prameni, z ktorého čerpal. Ak údaje o prameni obsahujú informácie, ktoré by mohli byť pri riešení úloh spojených s citovaným textom nápovedné (napríklad, keď majú žiaci určiť žánre ukážky a názov diela ho prezradza), autor úlohy uvedie pri citovanom texte len údaje, ktoré možno uvádzať v teste, a úplný bibliografický odkaz pre potreby NÚCEM-u uvedie v poznámke.

¹¹²§ 25 zákona č. 618/2003 Z. z. (Autorský zákon): Bez súhlasu autora možno použiť krátku časť zverejneného diela vo forme citácie v inom diele len na účel recenzie alebo kritiky tohto zverejneného diela alebo na **vyučovacie účely, vedeckovýskumné účely** alebo umelecké účely. Takéto použitie musí byť v súlade so zvyklosťami a jeho rozsah nesmie presiahnuť rámec odôvodnený účelom citácie. **Pri citácii sa musí uviesť meno autora** alebo jeho pseudonym, ak nejde o anonymné dielo, alebo meno osoby, pod ktorej menom sa dielo uvádza na verejnosti, ako aj **názov diela a prameň**.

Bibliografický odkaz na prameň musí byť správny, úplný a podľa aktuálnej citačnej normy:¹¹³

- Pri tlačených zdrojoch nesmie chýbať **autor a názov publikácie, miesto vydania, vydavateľ, rok vydania, rozsah strán**, z ktorých citujeme, a **ISBN** (ak ho publikácia obsahuje).
- Pri časopisoch okrem autora a názvu príspevku musíme uviesť aj údaje o časopise: názov časopisu, ročník, rok vydania, číslo, rozsah strán, z ktorých citujeme a ISSN (ak ho publikácia obsahuje).
- Pri citovaní elektronických zdrojov (ktoré nemajú tlačенú verziu) tiež musíme uviesť autora a názov príspevku, úplnú internetovú adresu a za ňou v zátvorkách dátum, kedy bola adresa autorom úlohy navštívená a funkčná.
- Ak ide o elektronickú verziu tlačenej publikácie, knihy alebo časopisu, citujeme ako pri tlačných dielach, ibaže navyše uvedieme aj internetovú adresu a za ňou v zátvorkách dátum, kedy bola adresa autorom úlohy navštívená a funkčná.

Pri citovaní elektronických zdrojov treba zvlášť skontrolovať úplnosť, správnosť a funkčnosť skopírovanej internetovej adresy. Najčastejšie chyby pri uvádzaní internetových adries:

1. Uvádzanie neúplnej adresy.
 - **Nesprávne:** <http://www.snk.sk/> (27. 5. 2013)
 - **Správne:** <http://www.snk.sk/?otvaracie-hodiny> (27. 5. 2013)
2. Znakové chyby v adrese, ktoré sa prejavujú po vložení do textu (najčastejšie pri adresách, ktoré obsahujú znaky s diakritikou).
 - **Nesprávne:** http://www.nucem.sk/documents//27/medzinarodne_merania/pisa/publikacie_a_diseminacia/1_narodne_spravy/N%C3%A1rodn%C3%A1_spr%C3%A1va_PISA_2006.pdf (27. 5. 2013)
 - **Správne:** http://www.nucem.sk/documents//27/medzinarodne_merania/pisa/publikacie_a_diseminacia/1_narodne_spravy/Národná_správa_PISA_2006.pdf (27. 5. 2013)
3. Skopírovanie skrátenej adresy pri kopírovaní z vyhľadávачa (napr. z Googla) namiesto z riadka adresy internetového prehliadača.
 - **Nesprávne:** www.literacytrust.org.uk/assets/0001/7162/..._2012.pdf (27. 5. 2013)
 - **Správne:** http://www.literacytrust.org.uk/assets/0001/7162/PLRS_Report_2012.pdf (27. 5. 2013)

V prípade výskytu podobnej chyby v adrese je lepšie adresu ručne opísať alebo opraviť v texte pomocou klávesnice a následne opravenú adresu skontrolovať v internetovom prehliadači.

4.4. Tvorba, výber a počet distraktorov

Testy s úlohami s výberom jednej odpovede z viacerých ponúknutých možností majú v testovaní význačné postavenie, a to najmä v súvislosti s možnosťou automatického vyhodnotenia odpovedových hárkov. Preto sa tvorbe takýchto úloh venujeme samostatne.

Výhodou takýchto úloh je najmä:

- vysoká objektivita hodnotenia,
- veľmi efektívny proces vyhodnocovania (pri použití elektronického testu máme kvalitne pripravený test vyhodnotený okamžite, pri použití odpovedových hárkov môžeme test vyhodnotiť strojovo, pri ručnom hodnotení môže test hodnotiť laik – nepotrebujeme experta),
- proces testovania zaberie testovanému menej času,

¹¹³STN ISO 690 (01 0197). Informácie a dokumentácia. Návod na tvorbu bibliografických odkazov na informačné pramene a ich citovanie. SÚTN, 2012.

- jednoduchšie a rýchlejšie sa dá upriamiť pozornosť testovaného z jedného problému na iný problém (z jednej úlohovej situácie na inú úlohovú situáciu).

V niektorých situáciách je možné za výhodu pokladať aj skutočnosť, že kvalitu žiackej odpovede neovplyvňuje miera jeho vyjadrovacích schopností (ak ju práve nechceme merať).

Pre výber vhodných distraktorov sa v testovacej praxi používajú v skúšobných testovaniach aj otvorené úlohy s tvorbou krátkej odpovede. Potom autor otvorenú úlohu pretransformuje na uzavretú a znovu ju podrobí skúšobnému testovaniu. Analýza žiackych odpovedí z otvorených úloh získaných v prvom skúšobnom testovaní poskytuje autorom úloh dve hlavné výhody:

1. Medzi žiackymi odpoveďami sa môžu objaviť originálne, neočakávané ale správne odpovede. Autor úlohy môže potom buď rozšíriť kľúč správnych odpovedí, alebo doplniť či preformulovať zadanie úlohy tak, aby takéto odpovede vylúčil.
2. Zo získaných nesprávnych odpovedí môže autor úloh čerpať námety na tvorbu distraktorov.

Pri tvorbe distraktorov musí autor úlohy splniť dve požiadavky:

1. Musí zvoliť také distraktory, aby si každý z nich zvolilo aspoň minimálne percento žiakov. Napríklad, ak má mať úloha obťažnosť 40 %, tak správnu odpoveď si úspešne zvolí 60 % žiakov a zvyšných 40 % žiakov si zvolí niektorý z distraktorov. Ak si niektorý distraktor nezvolí nikto, alebo len mizivé percento žiakov, môžeme o takej úlohe povedať, že mala v podstate o jeden distraktor menej. Preto, ak má autor k dispozícii výsledky skúšobného testu, v ktorom bola úloha ešte otvorená, vyberá si spomedzi frekventovaných nesprávnych odpovedí. Podobne, ak je niektorý z distraktorov nefunkčný preto, že je **príliš zjavne nesprávny** (vyberá si ho iba minimum žiakov), napríklad nonsens distraktory, neplní funkciu distraktora a teda môžeme povedať, že takáto úloha má iba 3 možnosti (1 správnu a 2 distraktory). V takom prípade distraktor musíme nahradiť a úlohu znova skúšobne otestovať, alebo vyradiť.
2. Musí však zároveň prihliadať na to, aby zvolené distraktory odrážali meraný jav, aby umožňovali pri vyhodnocovaní didaktickú analýzu. Autor úlohy musí vedieť, čo sa dá usúdiť o vedomostiach a zručnostiach žiakov, ktorý si daný distraktor zvolili, a toto zistenie musí korešpondovať s tým, čo daná úloha testuje.

Na tomto mieste sa venujeme základom tvorby uzavretej testovej úlohy s jednou správnou odpoveďou a tromi distraktormi (ISCED 2). Pred prípravou testových úloh je potrebné znovu si preštudovať aktuálny Štátny vzdelávací program, všeobecný (k príslušnému stupňu vzdelávania) aj ten k svojmu predmetu a špecifikáciu testu, v ktorom majú byť úlohy z banky úloh použité.

Pri tvorbe úloh musí autor dodržiavať určité pravidlá. Musí dbať na to:

1. aby úloha nebola chytákom alebo hlavolamom,
2. aby úloha testovala to, čo má testovať (pri počúvaní v CJ netestovať gramatiku, pri testovaní porozumenia textu netestovať kontrolu a pod.)
3. aby distraktory a správna odpoveď boli rovnako dlhé a zložité,
4. aby distraktory typu: *c) žiadna z ponúknutých odpovedí nie je správna* alebo *d) všetky ponúknuté odpovede sú správne*, keď už ich použijeme, boli dobre uvážené,
5. aby distraktory v určitých situáciách existovali (nesmú to byť nonsens distraktory čiže neexistujúce slová alebo zjavne nesprávne, nezmyselné odpovede), napríklad zákaz vjazdu piatich traktorov (známa úloha v testoch v autoškole),
6. aby formulácia zadania na začiatku a v úlohe bola jednoduchá, aby to nebola úloha v úlohe,
7. aby existovala jednoznačne správna odpoveď (zvyčajne práve jedna),
8. aby hlavná myšlienka príp. dôležité informácie boli v úlohe, nie v možnostiach odpovede,

9. aby v zadaní úlohy bolo čo najmenej kategorických slov (*vždy, nikdy, všade* a pod.) alebo slov, ktoré môžu viesť k odpovedi založenej na žiakovom úsudku (*môže byť, malo by byť, niekedy* a pod.),
10. aby, ak je to možné, boli ponúknuté odpovede logicky zoradené (napr. podľa abecedy),
11. aby bolo zadanie zapamätateľné a zrozumiteľné, aby sa žiak k nemu nemusel zbytočne vracat' pri posudzovaní každej z ponúknutých odpovedí.

Vráťme sa k najčastejším chybám naznačeným v niektorých z predchádzajúcich príkladov.

Úloha 1 je chybné sformulovaná – z údajov z jedného dňa nie je možné odpovedať na otázku „ako dlho cesta **trvá**“. Na tomto príklade upozorňujeme na nenáležité zovšeobecnenie. Úlohu môžeme upraviť napríklad nasledujúcim spôsobom.

Jozef šiel v pondelok ráno do školy. Z domu odišiel o 7.23 h a do školy vošiel o 7.38 h. Určte, ako dlho Jozefovi trvala cesta do školy.

- (A) 10 min (B) 15 min (C) 20 min (D) zo zadania sa nedá určiť

Úlohu už pravdepodobne môžeme zaradiť do pilotného testovania, avšak mali by sme predtým skontrolovať distraktor D. Sme si naozaj istí, že neexistuje alternatívny spôsob prístupu k úlohe, ktorý by pripúšťal iné riešenia? S distraktormi typu „zo zadania sa nedá určiť“ pracujeme veľmi opatrne, resp. čo najmenej.

Úloha 2 sa viazala na ukážku 1, v ktorej bolo napísané, že Galileo Galilei sa narodil v roku 1564. Nemáme dôvod pochybovať o tom, že sa dožil viac než 36 rokov, a teda že žil aj začiatkom 17. storočia. Čo ale, ak žil ešte aj v roku 1651? To by mal 87 rokov a žiak toto nemusí vedieť, nevie to vylúčiť. Teda odpoveď, že Galilei žil v druhej polovici 17. storočia by mohla byť tiež správna. Pokúsme sa úlohu upraviť.

Na základe ukážky môžeme povedať, že Galileo Galilei prežil svoju mladosť v druhej polovici _____.

- (A) 15. storočia (B) 16. storočia (C) 17. storočia (D) 18. storočia

V takto sformulovanej úlohe však distraktor D. nezvolí takmer nik, teda je nevhodne zvolený. Zvyšuje pravdepodobnosť, že žiaci uhádnu správnu odpoveď tipovaním jednej z troch možností.

Pokúsme sa upraviť aj úlohu 3. V tejto úlohe sú distraktory B a D veľmi podobné.

Galileo Galilei a Aristoteles sa _____.

- (A) často stretávali
- (B) stretávali iba veľmi výnimočne lebo žili veľmi ďaleko v iných krajinách
- (C) nemohli stretnúť, avšak Galileo Galilei poznal mnohé z diel Aristotela
- (D) mnohé z nich uznával, avšak niektoré aj kritizoval
- (E) nemohli stretnúť, ale aj tak Aristoteles kritizoval Galileiho učenie

V tejto úprave sme zámerne ukázali aj pomerne častú tendenciu autorov formulovať správnu odpoveď oveľa dlhšiu než distraktory.

V nasledujúcom príklade si ukážeme, že niektoré priradovacie úlohy môžeme pomerne jednoducho preformulovať na úlohy s výberom odpovede.

Úloha 25. Ktoré z nasledujúcich trojíc majú elektrónovú konfiguráciu 2.8.8, 2.8 a 2.8.1, v tomto poradí?

- (A) Ne, F, Na
- (B) Ca²⁺, F, Na⁺
- (C) K⁺, F⁻, Mg²⁺
- (D) Cl⁻, F⁻, Na

Upravte úlohu do formy priradovacej úlohy, uvedomte si klady a zápory oboch foriem tejto úlohy. Diskutujte tiež o poradí ponúknutých odpovedí – teda o grafickom spracovaní (formátovaní) úlohy.

5. ŠTATISTICKÁ ANALÝZA TESTOVÝCH ÚLOH

Štatistická analýza testových úloh – položiek (položková analýza) – je nevyhnutnou súčasťou analýzy výsledkov testovania. NÚCEM ju realizuje po každom testovaní, vrátane meraní realizovaných v národnom projekte E-test. Hlavným cieľom položkovej analýzy je:

1. korekcia úlohy na základe dát zo skúšobného (pilotného) testovania,
2. korekcia konečného výsledku testu, ak je potrebné vylúčiť položku z hodnotenia.

Ďalším cieľom položkovej analýzy môže byť aj získanie dodatočných informácií o spôsobe, akým žiaci riešia konkrétne úlohy a využitie týchto informácií na formatívne účely (úpravu vyučovacích metód a postupov). Prostredníctvom analýzy uprednostňovania jednotlivých distraktorov v úlohách s výberom odpovede alebo kvalitatívnou analýzou postupov pri riešení otvorených úloh môžeme získať vhľad do postupov, ktoré uplatňujú žiaci pri riešení úloh. Predtým, ako sa úloha dostane do certifikačného testu, ako je napríklad Testovanie 5, Testovanie 9 alebo externá časť maturitnej skúšky (EČ MS), prechádza sériou pilotných testovaní. Cieľom týchto testovaní je overiť fungovanie úloh v testovej situácii. Na základe výsledkov pilotných testovaní sa úlohy v prípade potreby upravujú (zadanie, distraktory a pod.) a následne dochádza k ich opätovnému pilotovaniu. **Proces tvorby kvalitných úloh je mimoriadne časovo náročný a zároveň kladie veľké nároky na odbornosť a kompetencie tvorcov testových úloh.** V odbornej literatúre venujúcej sa testovaniu a tvorbe úloh sa často uvádza, že vytvorenie samotných distraktorov v úlohe s výberom odpovede je niekedy náročnejšie ako vytvorenie samotného zadania úlohy.

Výber vzorky na pilotné testovanie

Pilotné testovanie je určitá aproximácia toho, ako bude fungovať úloha v ostrom testovaní. Nevyhnutnou súčasťou kvalitného pilotného testovania je vhodný výber žiakov. Veľkosť vzorky pri pilotných testovaniach býva viac ako 200 žiakov ($N > 200$). Počet žiakov v ostrých testovaniach je podstatne vyšší, na Slovensku sa pohybuje v tisícoch, pri zriedkavejších niektorých maturitných predmetoch v stovkách. Zloženie pilotnej vzorky by malo kopírovať zloženie populácie, na ktorej chceme administrovať ostrý test. Vzorka má teda byť vzhľadom na populáciu reprezentatívna. Dôležitú úlohu tu zohrávajú premenné ako napríklad pohlavie, typ školy, kraj, známky žiakov a podobne. Ak vieme, že v populácii, ktorú chceme testovať, je pomer chlapcov a dievčat 52 : 48, tak je potrebné, aby aj vo vybranej vzorke bol podobný pomer. Ak by sa v pilotnom testovaní použila nereprezentatívna vzorka, mohlo by dochádzať k nepresným odhadom parametrov úloh. Napríklad, ak by boli úlohy pilotované len na prestížnom gymnáziu zameranom na matematiku, tak by sa z výsledkov mohlo zdať, že úlohy sú pre žiakov jednoduché. Ak by sme mali ale vo vzorke aj iné „bežné“ školy, priemerná úspešnosť by sa pravdepodobne znížila. Bez pilotného testovania nie je možné odhadnúť fungovanie úloh v ostrom testovaní. Mimoriadne originálne úlohy, ktoré sú na pohľad bezchybné, dopadnú niekedy v pilotných testovaniach z pohľadu psychometrických ukazovateľov veľmi zle. Tento efekt často nastáva napríklad pri pilotovaní úloh z matematiky, ktoré sú relatívne obťažné, ale vyžadujú si hlboké porozumenie testovaného učiva s následnou aplikáciou (implementáciou) týchto poznatkov do neštandardnej situácie. Ak po pilotnom testovaní dochádza k zmenám v úlohách (preformulovanie, výmena distraktora, atď.), tak je potrebné úlohu opätovne pilotovať. Aj keď sa niektoré zmeny v zadaní na prvý pohľad môžu zdať bezvýznamné, môžeme hovoriť viac menej o novej úlohe. Z výsledkov ďalšieho pilotovania je často zrejmé, že psychometrické parametre upravených úloh sa výrazne zmenili.

Položková analýza

Položkovú analýzu testov je možné vykonať na základe dvoch rôznych prístupov. Jedna metóda je založená na klasickej teórii testov (ďalej len CTT)¹¹⁴ a druhá na modernej testovej teórii, ktorá sa nazýva aj teória odpovede na položku (ďalej len IRT).¹¹⁵ Spomenuté prístupy sa výrazne líšia v predpokladoch, z ktorých vychádzajú, samotné výsledky obidvoch metód však spolu konvergujú a vzájomne sa dopĺňajú. Štatistické parametre sa využívajú na zhodnotenie kvality testovej položky. Celková kvalita testu je potom odvodená od kvality položiek, ktoré sú v ňom použité.

Položková analýza prostredníctvom CTT

Štatistické analýzy vychádzajúce z klasickej testovej teórie sú rutinne uplatňované pri analýze dát z meraní vo vzdelávaní a predstavujú základ pri hodnotení kvality jednotlivých testovacích nástrojov.

CTT pracuje s konceptom pravého a chybového skóre. Každé meranie je tvorené pravým skóre žiaka (skutočná úroveň schopnosti, ktorú meriame testom) a chybovým skóre. Podľa tohto prístupu má každý testovaný žiak skutočné skóre (T), ktoré by sme vedeli zistiť len v tom prípade, ak by v našich meraniach neboli žiadne chyby. V praxi sa ale spoliehame na hrubé skóre žiaka (X) a zároveň predpokladáme, že meriame s určitou nepresnosťou (e). Chyba musí byť náhodná (napr. únava, nedostatočná koncentrácia, a pod.). Napríklad, žiak dosiahne v teste skóre 15 bodov. Test pozostával z 20 úloh s výberom odpovede z viacerých možností. Správne vyriešil 13 úloh, k tomu dokázal uhádnuť správnu odpoveď na ďalšie tri úlohy, ale na jednu z nich odpovedal nesprávne, lebo zle označil odpoveď.¹¹⁶

$$15 = 13 + (3 - 1)$$

$$T = X + e$$

T – skutočné skóre žiaka

X – hrubé skóre žiaka

e – chyba

Medzi základné psychometrické ukazovatele úlohy podľa CTT zaraďujeme obťažnosť a rozlišovaciu schopnosť.¹¹⁷ V nasledujúcej časti stručne objasníme základné pojmy a na príklade si vysvetlíme ako využiť metódy a postupy CTT v bežnej pedagogickej praxi pri vyhodnocovaní písomných prác v školskej praxi.

Obťažnosť

Obťažnosť úlohy sa dá vypočítať pomerne jednoducho, pretože ju predstavuje percentuálne vyjadrený podiel žiakov, ktorí odpovedali na položku správne. Môže nadobúdať hodnoty od 0 do 100 %. Obťažnosť sa niekedy označuje aj ako P hodnota. Veľmi obťažné úlohy majú nízke hodnoty P a ľahké úlohy, naopak, vysoké.

Na základe obťažnosti môžeme úlohy rozdeliť na niekoľko skupín. Rozdelenie je ale arbitrárne, to znamená, že záleží na nás, aké intervaly pre ľahké, priemerne obťažné a mimoriadne obťažné úlohy si zvolíme. Ak na základe štatistických analýz odpovedí zistíme príliš vysokú alebo nízku hodnotu obťažnosti, je potrebná detailná didaktická analýza úlohy.

¹¹⁴Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart and Winston.

¹¹⁵Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). Fundamentals of Item Response Theory. Newbury Park, CA: SAGE.

¹¹⁶Halama, P. (2011). Princípy psychologickéj diagnostiky. Trnava: Filozofická fakulta TU.

¹¹⁷Kline, T.J.B. (2005). Psychological Testing: A practical Approach to Design and Evaluation. Thousand Oaks, CA: SAGE.

Ak ide o mimoriadne jednoduchú úlohu a má slúžiť ako „zahrievacia“, tak je možné úlohu v teste ponechať. Ak je cieľom testovania zisťovanie rozdielov medzi žiakmi, táto úloha prináša len veľmi málo informácií.

V prípade mimoriadne obťažnej úlohy môžeme uvažovať o viacerých vysvetleniach. Môže ísť o nejednoznačné zadanie, ktoré zmatlo študentov, alebo žiaci celkovo neporozumeli učivu, ktoré úloha testovala. Vo všeobecnosti sa neodporúča používať úlohy, ktorých obťažnosť P je nižšia ako 20 %. Podobné kritérium sa dá uplatniť aj pri mimoriadne ľahkých úlohách.

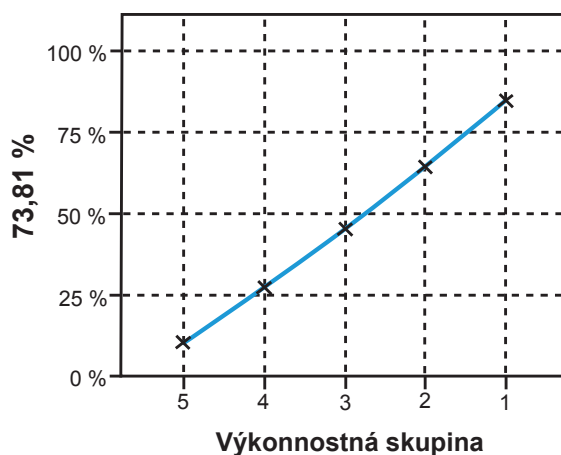
Výber úloh, ktoré sú zaradené do testu, na základe obťažnosti, výrazne závisí od cieľa a zámeru testovania. Ak je cieľom vybrať (napríklad na prestížnu školu) najlepších uchádzačov, test bude obsahovať väčší počet obťažnejších úloh.

Ukazovateľ obťažnosti je veľmi závislý od vzorky, na ktorej úlohu pilotujeme.

Diskriminačná schopnosť

Za predpokladu, že test aj úloha v teste merali ten istý konštrukt (napr. matematickú schopnosť), je prirodzené predpokladať, že žiaci s celkovým vyšším skóre v teste budú častejšie správne odpovedať na položku a naopak, žiaci s nižším skóre v teste budú odpovedať prevažne nesprávne. Schopnosť úlohy rozlišovať medzi najlepšími a najslabšími žiakmi (v teste) nazývame *diskriminačná schopnosť položky*. Existuje viacero postupov, ako môžeme ukazovateľ tejto schopnosti vypočítať.

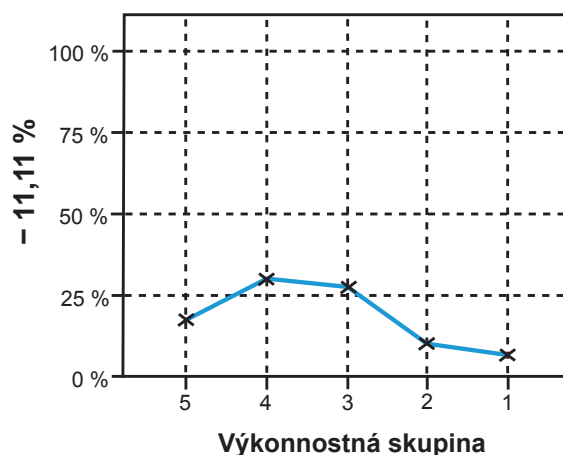
Najjednoduchší spôsob predstavuje metóda výkonnostných skupín. Všetkých žiakov rozdelíme na základe dosiahnutého skóre v teste do rovnako veľkých skupín. Počet skupín závisí od počtu testovaných žiakov. Pri menších testovaniach je vhodné pracovať napríklad s piatimi výkonnostnými skupinami. Pri väčších testovaniach (napr. T9 alebo EČ MS) je možné testový súbor rozdeliť aj na 10 skupín. V odbornej literatúre sa vedie polemika o tom, koľko percent žiakov majú obsahovať jednotlivé skupiny. Následne sa pre každú výkonnostnú skupinu určí priemerná úspešnosť a vnesie sa do grafu. Diskriminačná schopnosť tak predstavuje rozdiel v priemernej úspešnosti medzi skupinou najlepších a najslabších žiakov. Skupina označená číslom 1 predstavuje najlepších žiakov a skupina označená číslom 5 – najslabších žiakov. Číselné označenie výkonnostných skupín však nezodpovedaná známkam.



Obrázok 10. Položka s vynikajúcou rozlišovacou schopnosťou

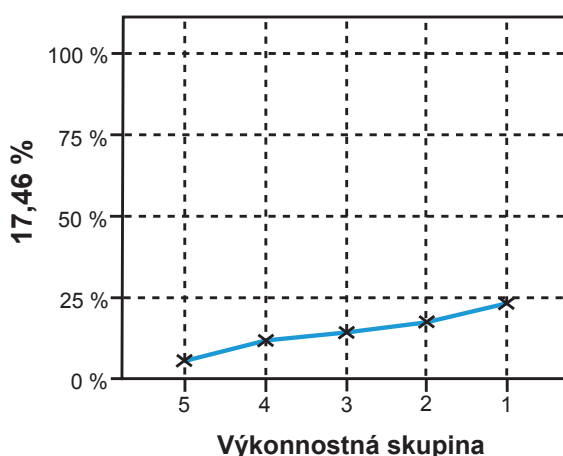
U položiek s dobrou rozlišovacou schopnosťou môžeme pozorovať veľký rozdiel v priemernej úspešnosti medzi prvou a poslednou výkonnostnou skupinou. Jedným z hlavných cieľov pilotného testovania je okrem iného aj identifikácia položiek s nedostatočnou rozlišovacou schopnosťou. Veľkosť rozdielu

v úspešnostiach medzi skupinami môže nadobúdať kladné aj negatívne hodnoty. Ideálna položka má vysoké kladné hodnoty.



Obrázok 11. Položka so zápornou rozlišovacou schopnosťou

Negatívnu diskriminačnú schopnosť majú často veľmi obťažné položky. Najslabší žiaci tak majú mierne vyššiu úspešnosť v porovnaní s najlepšimi žiakmi. Táto situácia môže nastať vtedy, ak je úloha nejasne naformulovaná alebo predstavuje určitý „chyták“. Položky s negatívnou rozlišovacou schopnosťou by sa nemali objaviť v národných meraniach ako Testovanie 5, Testovanie 9 alebo EČ MS. Rovnako problematická je aj situácia, keď je rozdiel medzi výkonnostnými skupinami veľmi malý. Priemerná úspešnosť sa medzi jednotlivými skupinami výrazne nemení. Môže to byť spôsobené extrémnou obťažnosťou položiek alebo môže ísť o problém s validitou položky.



Obrázok 12. Veľmi obťažná, slabo rozlišujúca položka

Existuje viacero indexov diskriminačnej schopnosti položiek, ktoré vychádzajú z priemerných úspešností najlepšej a najslabšej skupiny. Z didaktických dôvodov je vhodné používať aj grafické znázornenie, pretože priemerné hodnoty vstupujúce do výpočtov môžu byť niekedy mierne zavádzajúce (najmä pri menšom počte žiakov).

Point-Biserial

Jedným z najčastejšie používaných ukazovateľov rozlišovacej schopnosti položky je aj point-biserial korelácia (point-biserial correlation). Korelačný koeficient opisuje tesnosť vzťahu medzi celkovým

skóre a odpoveďou na položku; môže nadobúdať hodnoty od - 1 do 1. Vzorec na výpočet korelačného koeficientu je nasledovný:¹¹⁸

$$r_{pb} = \frac{M_s - M}{SD} \cdot \sqrt{pq}$$

r_{pb} – korelačný koeficient point-biserial

M_s – priemerné skóre žiakov, ktorí odpovedali správne

M – priemerné skóre všetkých žiakov

SD – štandardná odchýlka celého testu

P – percentuálny podiel žiakov, ktorí odpovedali správne

Q – percentuálny podiel žiakov, ktorí odpovedali nesprávne

Point-biserial korelácia jednotlivých položiek (r_{pb}) by mala mať hodnotu vyššiu ako 0,300. Pri výpočte point-biserial korelácie je možné postupovať aj tak, že počet bodov, ktoré žiak získal za analyzovanú položku, nie je započítaný do celkového skóre (táto procedúra prebieha automaticky pomocou štatistického softvéru). Takýto ukazovateľ sa potom nazýva korelácia položky so zvyškom testu (item-total correlation). Nízke hodnoty r_{pb} poukazujú na nízku validitu položky, resp. na jej slabú rozlišovaciu schopnosť.¹¹⁹

Vynechanosť

Mimoriadne dôležitým ukazovateľom kvality položky je vynechanosť (neodpovedanie na položku). Pri úlohách s výberom odpovede z viacerých možností tento ukazovateľ zvyčajne nie je problematický, pretože správnu odpoveď je možné uhádnuť (pri certifikačných testoch je pravdepodobnosť správnej odpovede 20 – 25 %, čo je pomerne vysoké číslo). Problém s vynechanosťou je skôr charakteristický pre úlohy s tvorbou krátkej odpovede. Položky môžu byť veľmi obťažné, ale môžu mať zároveň nízku vynechanosť. Ak sa príliš veľký počet žiakov nedokázal dopracovať k odpovedi (správnej alebo nesprávnej), tak je nevyhnutné podrobiť položku detailnej didaktickej analýze. V pilotných testovaní môže táto situácia nastať, ak ešte časť testovaných žiakov nepreberala učivo, ktoré položka testovala.

Analýza distraktorov

Výber a tvorba distraktorov zásadným spôsobom ovplyvňuje kvalitu úloh s výberom odpovede z viacerých možností. Tvorba vhodných distraktorov je mimoriadne dôležitá. Nevyhnutným predpokladom pri úlohách s výberom odpovede je, že existuje len jediná správna odpoveď a všetky distraktory sú skutočne nesprávne odpovede. Distraktory by si mali vyberať najmä žiaci, ktorí si nedokázali dostatočne osvojiť preberané učivo. S narastajúcou úrovňou schopnosti (zvládnutia učiva) by mala klesať preferencia jednotlivých distraktorov. Korelácia medzi výberom distraktora a celkovým skóre by mala byť negatívna. V prípade, že distraktor koreluje s celkovým skóre kladne, je nevyhnutná revízia položky. Krivka zobrazujúca výber distraktorov by mala mať v závislosti od narastajúcej úrovne schopnosti testovaných žiakov klesajúci priebeh. Pre kvalitu úlohy s výberom odpovede je dôležité, aby bol každý z distraktorov vybratý aspoň niekoľkými žiakmi. Ak sú napríklad dva distraktory zo štyroch zjavne nefunkčné (nevyberajú si ich ani neúspešní žiaci), pravdepodobnosť uhádnutia správnej odpovede je veľmi vysoká. Na základe položkovej analýzy je možné po pilotných testovaniach revidovať alebo odstrániť jednotlivé nefunkčné distraktory. V nasledujúcej časti uvedieme aj príklady štatistickej analýzy úloh s výberom odpovede z viacerých možností.

¹¹⁸Kline, T. J. B. (2005). *Psychological Testing: A practical Approach to Design and Evaluation*. Thousand Oaks, CA: SAGE.

¹¹⁹Furr, M. R. & Bacharach, V.R. (2008). *Psychometrics: An Introduction*. Thousand Oaks, CA: SAGE.

Ako urobiť jednoduchú položkovú analýzu v školskej praxi?

Detailné štatistické analýzy si vyžadujú veľkú vzorku testovaných žiakov, ktorá sa dá v bežnej pedagogickej praxi len ťažko dosiahnuť. Napriek tomu je možné základné ukazovatele kvality testu a položiek vyjadriť aj pri vzorke tridsiatich žiakov. Výsledky je potrebné brať s rezervou, môžu však slúžiť na orientačné zhodnotenie kvality položiek použitých v teste.

Predstavme si situáciu, že učiteľ rozdal test z matematiky svojim tridsiatim žiakom. Výsledky je možné veľmi jednoducho spracovať v programe MS EXCEL použitím funkcií AVERAGE, SUM a RANK (zoradenie).

V tabuľke 1 sú v stĺpcoch 01 až 10 body za jednotlivé úlohy (každá bola hodnotená nasledovne: 0 bodov – nesprávne; 1 bod – správne). Riadky predstavujú žiakov v triede. Pre každého žiaka najskôr vypočítame jeho hrubé skóre vyjadrené počtom dosiahnutých bodov (posledný stĺpec).

Žiakov si následne zoradíme podľa dosiahnutých bodov a rozdelíme ich na tri skupiny podľa počtu dosiahnutých bodov. V našich simulovaných výsledkoch sme rozdelili žiakov na tri rovnako veľké skupiny po 10 žiakov (skupina najslabších žiakov je označená oranžovou farbou, skupina najúspešnejších žiakov zelenou farbou).

Následne vypočítame úspešnosť žiakov v jednotlivých úlohách (percento žiakov, ktorí dokázali vyriešiť úlohu správne). Tento údaj vypočítame pre celý súbor žiakov (riadok Priemerná úspešnosť) a tiež osobitne pre skupinu najlepších a najslabších žiakov (riadok Priemerná úspešnosť – najlepší a riadok Priemerná úspešnosť – najslabší).

V dolnej časti tabuľky sú vypočítané hodnoty rozdielu priemerného skóre medzi hraničnými skupinami žiakov. Ukazovateľ rozlišovacej schopnosti použitých úloh tak predstavuje rozdiel v priemernej úspešnosti medzi týmito skupinami.

Z tabuľky je zrejmé, že najlepšie rozlišujúce sú úlohy č. 03, 06 a 09. Rozdiel medzi skupinami je až 80 %. Úloha číslo 04 má zápornú rozlišovaciu schopnosť. Znamená to, že skupina najlepších žiakov v teste bola pri riešení tejto úlohy menej úspešná ako najslabší žiaci. Ako sme spomínali vyššie, takéto úlohy sú veľmi problematické a vyžadujú si detailnú didaktickú analýzu.

Žiaci získali v teste priemerne 5,23 bodu, čo predstavuje úspešnosť 52,3 %. Z tabuľky je ale zrejmé, že úlohy sa medzi sebou líšia v obťažnosti (percentuálny podiel žiakov, ktorí vyriešili položku správne). Úloha č. 07 sa javí ako mimoriadne zložitá, nakoľko ju vyriešilo správne menej ako 20 % všetkých žiakov. Taktiež ale vidíme, že ak ju vyriešil žiak správne, tak to bol žiak zo skupiny priemerných alebo najlepších žiakov. Úloha si teda stále zachováva uspokojivú rozlišovaciu schopnosť, je však mimoriadne obťažná.

Tabuľka 11. Príklad analýzy výsledkov písomného testu z matematiky

Por. č. žiaka	Číslo úlohy										Počet bodov
	01	02	03	04	05	06	07	08	09	10	
8	0	0	0	1	0	0	0	0	0	0	1
7	0	0	0	0	1	0	0	0	0	0	1
11	1	0	0	1	0	0	0	0	0	0	2
2	1	1	0	1	0	0	0	0	0	0	3
5	0	1	0	1	0	0	0	0	0	1	3
14	1	1	0	0	1	0	0	0	0	0	3
18	0	0	0	0	1	0	0	1	1	0	3
3	1	1	0	1	1	0	0	0	0	0	4
13	0	0	1	0	1	1	0	1	0	0	4

Por. č. žiaka	Číslo úlohy										Počet bodov
	01	02	03	04	05	06	07	08	09	10	
20	1	0	0	0	1	0	0	0	1	1	4
21	0	1	0	1	1	0	0	1	1	0	5
27	1	1	1	0	0	1	0	1	0	0	5
28	1	1	0	0	1	1	0	1	0	0	5
19	1	1	0	1	0	1	0	0	0	1	5
23	1	1	0	1	1	0	0	0	1	0	5
24	1	0	1	0	0	1	0	1	1	0	5
29	0	1	1	0	0	1	1	1	0	0	5
4	1	1	0	0	0	1	1	1	1	0	6
6	0	1	1	1	1	0	0	1	1	0	6
12	0	1	1	1	1	1	0	1	0	0	6
15	1	1	1	0	1	1	0	1	1	0	7
1	1	1	1	0	1	1	0	1	1	0	7
25	1	1	1	0	1	1	0	1	1	0	7
26	1	1	0	1	1	1	0	0	1	1	7
30	0	1	1	0	0	1	1	1	1	1	7
16	1	1	1	0	1	1	0	1	1	0	7
22	1	1	1	1	1	0	0	1	1	1	8
9	1	1	1	0	1	1	1	1	1	0	8
10	1	1	1	1	1	1	0	1	1	1	9
17	1	1	1	0	1	1	1	1	1	1	9
Priemerná úspešnosť	66,7	76,7	50,0	43,3	66,7	56,7	16,7	63,3	56,7	26,7	5,23
Priemerná úspešnosť – najlepší	90	100	90	30	90	90	30	90	100	50	7,6
Priemerná úspešnosť – najslabší	50	40	10	50	60	10	0	20	20	20	2,8
Rozdiel	40	60	80	- 20	30	80	30	70	80	30	

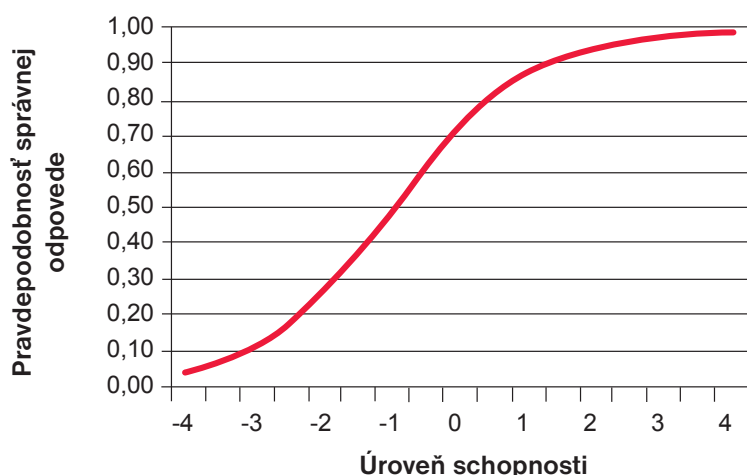
Podobným spôsobom je možné spracovať aj ďalšie základné ukazovatele, ako napríklad porovnanie priemernej úspešnosti chlapcov a dievčat v rámci celého testu. Vzťah medzi známkom na polročnom vysvedčení a výsledkom v T9 alebo EČ MS je pomerne silný, preto prichádza do úvahy aj analýza vzťahu medzi týmito dvoma ukazovateľmi. V ideálnom prípade by žiaci s lepšou známkom mali mať vyššiu úspešnosť. Na záver je potrebné povedať, že používanie percentuálneho vyjadrenia pri malých skupinách žiakov môže byť niekedy zavádzajúce a je potrebné na to prihliadať pri interpretácii výsledkov. Na druhej strane však aj takéto „hrubé“ odhady psychometrických charakteristík môžu dodať veľa cenných informácií, ktoré môže pedagóg vo svojej praxi využiť.

Položková analýza prostredníctvom teórie odpovede na položku (IRT)

IRT vychádza z dvoch predpokladov:¹²⁰

1. Výkon žiaka v teste môže byť predikovaný (aj vysvetlený) na základe určitých faktorov, ktoré nazývame latentný rys, resp. úroveň schopností (θ).
2. Vzťah medzi odpoveďou žiakov na položku a úrovňou ich schopnosti môžeme opísať prostredníctvom stúpajúcej krivky (*charakteristická krivka položky*). Krivka opisuje, aká je pravdepodobnosť správnej odpovede žiaka pri danej úrovni schopnosti.

¹²⁰Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: SAGE.



Obrázok 13. Charakteristická krivka položky

Pravdepodobnosť správnej odpovede žiaka na položku je vyjadrená nasledovne:¹²¹

$$P_{ij}(\theta_j) = c_i + (1 - c_i) \frac{\exp[Da_i(\theta_j - b_i)]}{1 + \exp[Da_i(\theta_j - b_i)]}$$

P_{ij} – pravdepodobnosť správnej odpovede na položku i žiakom j

θ_j – úroveň schopnosti žiaka j

a_i – rozlišovacia schopnosť položky i (sklon krivky)

b_i – obťažnosť položky

c_i – dolná asymptota alebo parameter pseudouhádnuteľnosti položky i

D – škálovacia konštanta (1 alebo 1,7)

Už na prvý pohľad je zrejmé, že metódy IRT sú podstatne náročnejšie a nie je ich možné uplatniť v bežnej pedagogickej praxi. Nakoľko v rámci projektu E-test pracujeme s veľkými počtami žiakov a máme potrebný softvér, IRT analýzy predstavujú rutinnú súčasť vyhodnotenia kvality testu a testových úloh. V nasledujúcej časti si priblížime len základné koncepty použité pri analýzách. Detailné vysvetlenie prístupu presahuje rozsah a zameranie tejto publikácie.

Na zistenie psychometrických vlastností úloh, pri použití IRT používame zhodnotenie charakteristických kriviek položiek. Prostredníctvom IRT môžeme pri jednotlivých položkách odhadnúť tri parametre (a , b , c – v závislosti od použitého modelu). Tvar krivky závisí od hodnoty odhadnutých parametrov.

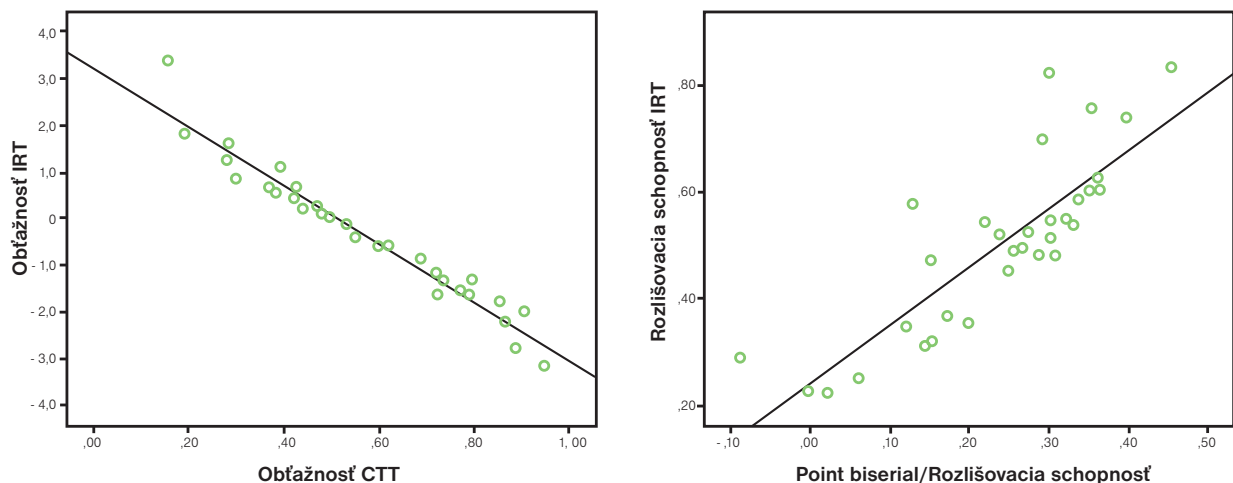
Parameter a predstavuje **rozlišovaciu schopnosť položky**. Vyššie hodnoty hovoria o lepšej rozlišovacej schopnosti. Čím je hodnota parametra a väčšia, tým je sklon charakteristickej krivky položky väčší.

Parameter b predstavuje **obťažnosť položky**. Položky so zápornou hodnotou b parametra sú ľahšie a položky s kladnou hodnotou b sú obťažnejšie. Hodnota parametra b zodpovedá úrovni schopnosti, pri ktorej majú žiaci 50 % pravdepodobnosť správnej odpovede. Hodnota parametra b je analógiou k indexu obťažnosti P (percentuálny podiel správnych odpovedí na danú položku) v CTT.

Parameter c predstavuje **faktor hádania** – pravdepodobnosť uhádnutia správnej odpovede. Zaoberať sa týmto parametrom má zmysel najmä pri úlohách s výberom odpovede z viacerých možností. Vyššia hodnota tohto parametra naznačuje, že správna odpoveď sa dá uhádnuť ľahšie, ako by sa dalo čakať (napr. viac ako 25 % pri 4 možnostiach odpovede).¹²²

¹²¹ De Ayala, R. J. (2009) The Theory and practice of Item Response Theory. New York: Guilford.

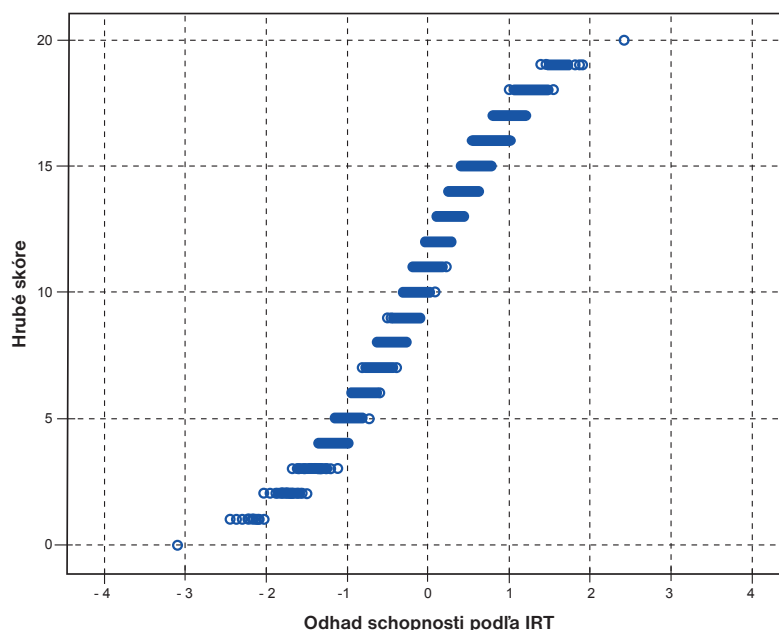
¹²² De Vellis, R. F. (2012) Scale development: Theory and applications: Thousand Oaks, CA: SAGE.



Obrázok 14. Vzťah medzi psychometrickými ukazovateľmi podľa CTT a IRT

Pri analýzach prostredníctvom IRT nepracujeme priamo s hrubým skóre jednotlivých žiakov, ale používame odhad latentnej schopnosti. Úroveň schopnosti je odhadnutá prostredníctvom komplikovaných matematických postupov a nie je v rozsahu tejto publikácie ich detailne priblížiť. Dôležité je ale objasniť vzťah medzi obťažnosťou položiek a úrovňou schopnosti. Obidve sú vyjadrené v rovnakých jednotkách, a tak je možné zhodnotiť primeranosť testu alebo položiek úrovni testovaných žiakov. Priemerná úroveň schopnosti je podľa konceptu IRT rovná nule a zároveň platí, že priemerne obťažná položka má nulovú hodnotu parametra b . Napríklad, ak má položka obťažnosť $b = 1$, tak žiak s úrovňou schopnosti $\theta = 1$ má 50 %-nú pravdepodobnosť, že vyrieši úlohu správne.

Na nasledujúcom obrázku je znázornený vzťah hrubého skóre podľa CTT a odhadu schopnosti prostredníctvom IRT. Môžeme vidieť, že priemerné skóre v teste (cca. 10 bodov) zodpovedá úrovni schopnosti $\theta = 0$. Z obrázku je tiež zrejmé, že pri rovnakom počte bodov môžu mať žiaci rôznu úroveň schopnosti. Pri použití dvoj- a troj-parametrového modelu je pri odhade schopnosti zohľadnené aj to, na ktoré konkrétne položky žiak odpovedal správne. Každá položka má tak svoju „váhu“.



Obrázok 15. Vzťahy medzi parametrami položiek v CTT a IRT

Položky je možné klasifikovať na základe hodnôt parametrov. Nasledujúce rozdelenie používa NÚCEM v rámci svojich analýz. Presné hranice je pomerne ťažké stanoviť, a preto je vždy potrebné zároveň brať do úvahy aj hodnotu rozlišovacej schopnosti položky.

Obťažnosť úloh podľa hodnoty parametra b :

- Veľmi ľahká úloha: menej ako $-1,5$
- Ľahká úloha: od $-1,5$ do $0,5$
- Stredne ťažká úloha: od $-0,5$ do $0,5$
- Ťažká úloha: od $0,5$ do $1,5$
- Veľmi ťažká úloha: viac ako $1,5$

Príklady analýzy úloh

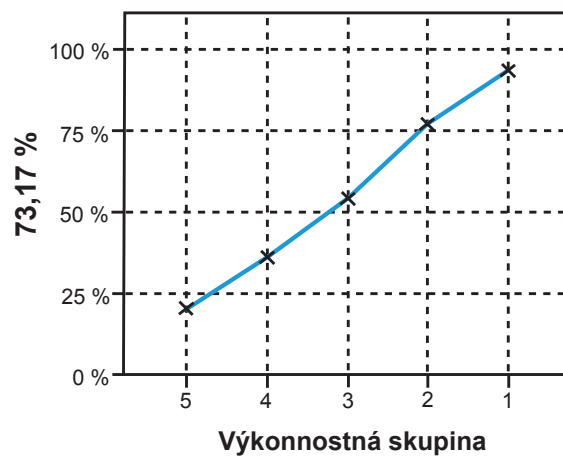
V nasledujúcej časti si stručne na príkladoch vysvetlíme ako prostredníctvom štatistických analýz môžeme zhodnotiť kvalitu položiek v teste. Súčasne pri každej úlohe uvedieme príslušné parametre podľa CTT aj podľa IRT.

Príklad 1.

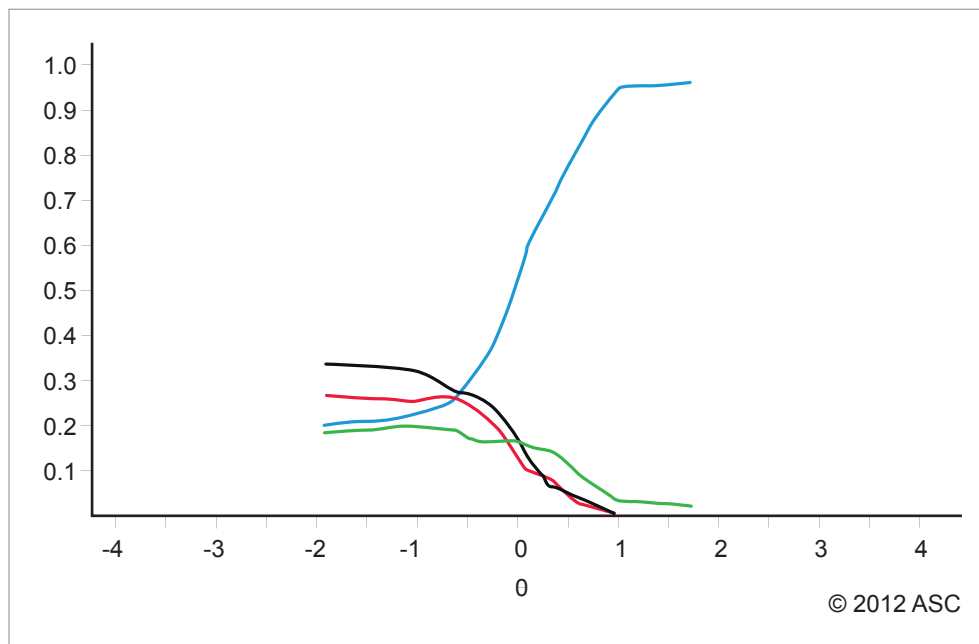
V ktorej vete sa nachádza sloveso v minulom čase a podmieňovacom spôsobe?

- 08**
- A** Nuž nemaj mi za zlé, že som ti to nedovravel.
 - B** Ak by sa naše potulky končili raz rozlúčkami.
 - C** A chladná rosa by sa ti bola vo vlasoch jagala.
 - D** Túlali sme sa tmavým mestom celkom sami.

Úloha bola súčasťou testu zo slovenského jazyka a literatúry v Testovaní 9-2013. Je to úloha s výberom odpovede zo štyroch možností. Položka má výborné rozlišovacie schopnosti. Rozdiel v priemernej úspešnosti medzi prvou a piatou výkonnostnou skupinou žiakov bol 73,17 % (obrázok 16). Na obrázku 17 je zobrazená analýza distraktorov tejto úlohy prostredníctvom IRT. Správna odpoveď C (modrá farba) prudko stúpa zároveň s narastajúcou úrovňou schopností testovaných žiakov. Distraktory A (červená farba) a D (čierna farba) si najslabší žiaci vybrali častejšie ako správnu odpoveď. V tabuľke 12 môžeme vidieť point-biserial koeficient položky (P. Bis), počet žiakov, ktorí si vybrali jednotlivé distraktory (N) a tiež percentuálny podiel týchto žiakov (p). Symbol X predstavuje žiakov, ktorí neodpovedali na položku. Na položku odpovedalo správne 56 % žiakov. Položka bola priemerne obťažná. Položka silne korelovala so zvyškom testu ($r_{pb} = 0,441$). Z pohľadu IRT šlo o mimoriadne dobre rozlišujúcu položku ($a = 1,417$), čo je zrejme z obrázku 17, na ktorom vidíme výber jednotlivých distraktorov. Pravdepodobnosť správnej odpovede prudko stúpa spolu s narastajúcou úrovňou schopnosti testovaných žiakov. Aj podľa IRT sa jednalo o priemerne obťažnú položku ($b = 0,142$). Napriek tomu, že išlo o úlohu s výberom odpovede, môžeme povedať, že distraktory boli pre najslabších žiakov pomerne atraktívne, a tak predpokladáme, že nedochádzalo k výraznému typovaniu správnej odpovede ($c = 0,213$).



Obrázok 16. Graf citlivosti



Možnosti = A - červená; B- zelená; C - modrá; D - čierna

Obrázok 17. Graf distraktorov

Tabuľka 12. Tabuľka distraktorov

	A	B	C	D	X
P. Bis.	- 0,22	- 0,10	0,44	- 0,29	- 0,04
p	14,0 %	13,0 %	56,0 %	17,0 %	0,0 %
N	2 716	2 571	11 102	3 421	91

Príklad 2.

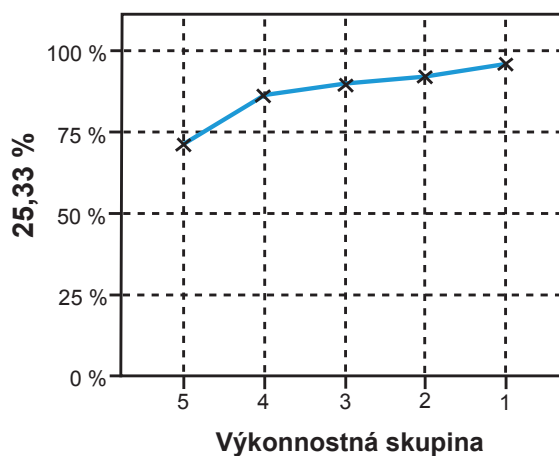
Aký význam má slovo *rozpumpovala* v krátkej správe?

Návštevníkov včerajšieho koncertu najviac rozpumpovala a primála vlniť sa v rytme samby najmä ohnivá hudba, za čo hudobníkom ďakujeme.

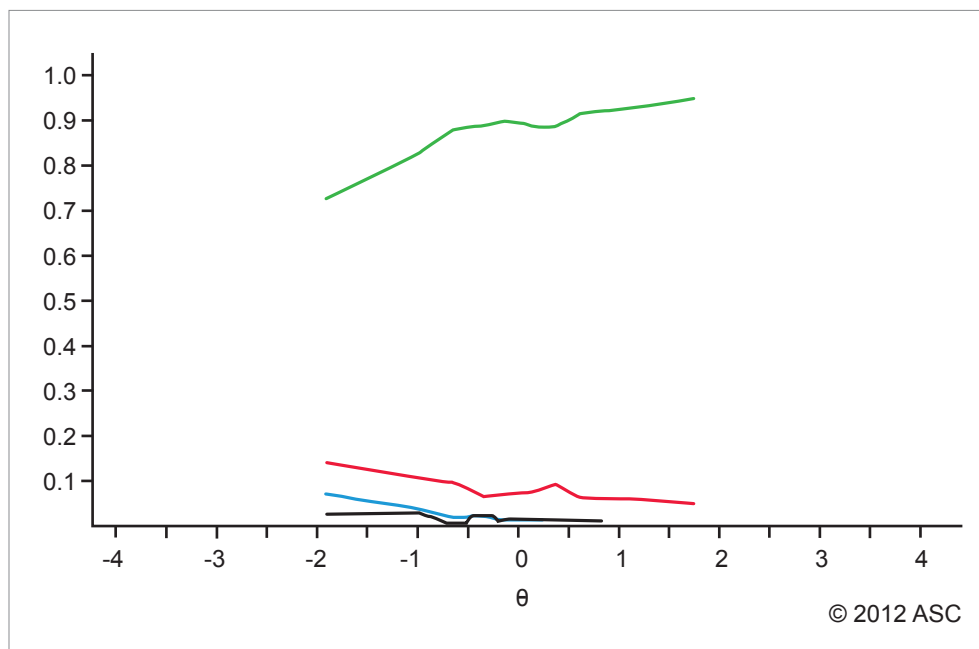
22

- A posilnila
 B rozprúdila
 C načerpala
 D udivila

Táto úloha bola súčasťou rovnakého testu ako predchádzajúca. Opäť to bola úloha s výberom odpovede zo štyroch možností. Z grafu citlivosti môžeme vidieť, že rozdiel v priemernej úspešnosti medzi prvou a piatou výkonnostnou skupinou bol pomerne malý (25,33 %). Úloha bola pre žiakov veľmi jednoduchá, správne ju dokázalo vyriešiť až 87 % žiakov. V tabuľke 13 sú uvedené počty jednotlivých žiakov, ktorí si vybrali príslušné distraktory. Tie boli pre žiakov málo atraktívne. Môžeme tu tiež vidieť, že aj point-biserial korelačný koeficient položky je pomerne nízky, čo je spôsobené tým, že aj najslabší žiaci mali takmer rovnakú priemernú úspešnosť ako žiaci s celkovo najvyšším skóre v teste. Aj z pohľadu IRT šlo o veľmi ľahkú úlohu ($b = -2,731$) s nízkou rozlišovacou schopnosťou (0,398). Úloha neprinesla o žiakoch veľa informácií, pretože ju takmer všetci žiaci vyriešili správne.



Obrázok 18. Graf citlivosti



Možnosti = A - červená; B- zelená; C - modrá; D - čierna

Obrázok 19. Graf distraktorov

Tabuľka 13. Tabuľka distraktorov

	A	B	C	D	X
P. Bis.	- 0,13	0,21	- 0,15	- 0,07	- 0,04
p	9,0 %	87,0 %	2,0 %	2,0 %	0,0 %
N	1 776	17 314	484	299	28

Príklad 3.

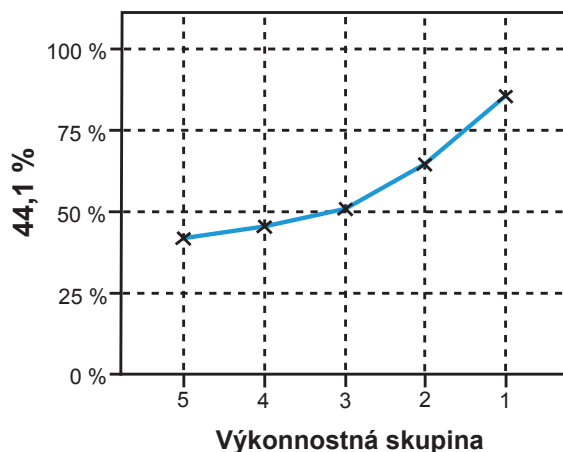
Ukážka 4

Sofokles – Antígona

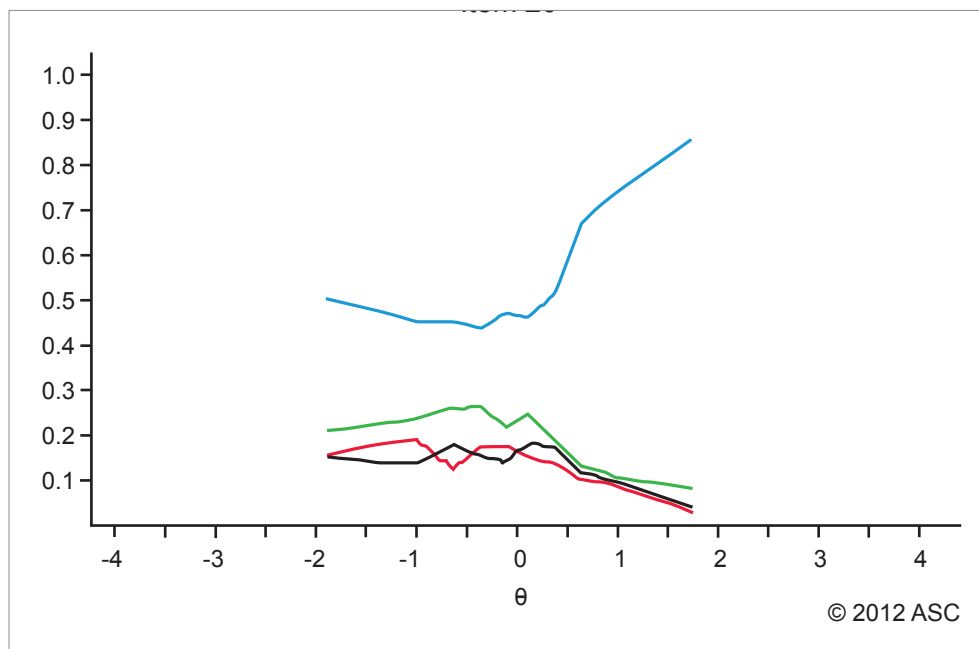
Ukážka je úryvkom

- 20 **A** z novely. **B** z balady. **C** z tragédie. **D** z románu.

Táto úloha tiež pochádza z testu zo slovenského jazyka a literatúry z Testovania 9-2013. 56,6 % všetkých žiakov dokázalo na ňu odpovedať správne. Úloha nemala dostatočnú rozlišovaciu schopnosť (44,1 %). Preferencia jednotlivých distraktorov síce klesala s narastajúcou úrovňou schopnosti žiakov. Zaujímavé však je, že najslabší a priemerní žiaci mali približne rovnakú úspešnosť pri riešení úlohy. Úlohy s výberom odpovede majú oproti úlohám s tvorbou krátkej odpovede niekoľko nevýhod. Často sú menej rozlišujúce a v niektorých prípadoch je možné aj uhádnuť správnu odpoveď. Ak má úloha 4 možnosti odpovede, tak vždy je minimálne 25 % šanca na uhádnutie správnej odpovede. Pri použití trojparametrového IRT modelu je možné zhodnotiť aj mieru „uhádnuteľnosti položky“. Táto miera je vyjadrená parametrom c . Pri úlohe s výberom zo štyroch možností by jeho hodnota nemala presiahnuť 0,250. V prípade tejto úlohy je hodnota parametra $c = 0,399$. Na základe toho môžeme usúdiť, že najslabší žiaci odpovedajú správne signifikantne častejšie, ako by sme predpokladali.



Obrázok 20. Graf citlivosti



Možnosti = A - červená; B- zelená; C - modrá; D - čierna

Obrázok 21. Graf distraktorov

Tabuľka 14. Tabuľka distraktorov

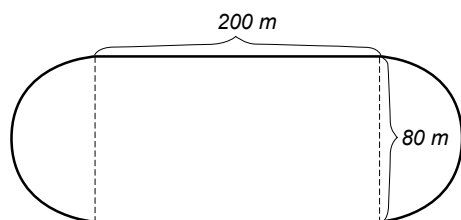
	A	B	C	D	X
P. Bis.	- 0,10	- 0,10	0,21	- 0,09	- 0,05
p	12,0 %	18,0 %	57,0 %	13,0 %	0,0 %
N	2 373	3 527	11 416	2 516	69

Príklad 4.

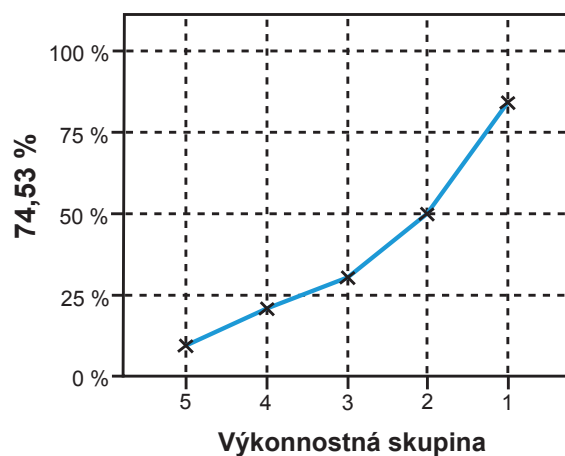
Námestie má tvar zložený z obdĺžnika a dvoch zhodných polkruhov (pozrite obrázok). Vypočítajte obvod tohto námestia a výsledok zaokrúhlite **na celé metre**.

- (A) 902 (B) 651 (C) 560 (D) 526

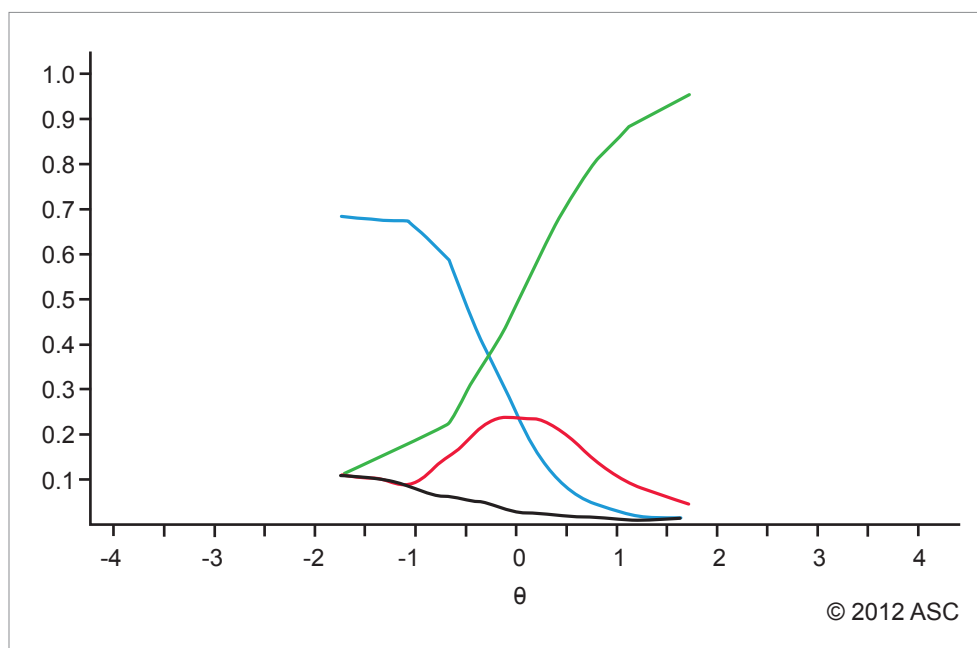
14



Posledná ilustračná úloha pochádza z testu z matematiky z Testovania 9-2013. Je to položka s výberom odpovede zo štyroch možností. V grafe citlivosti môžeme vidieť, že úloha veľmi dobre rozlišuje medzi najlepšimi a najslabšími žiakmi (74,53 %). Úlohu dokázalo vyriešiť správne 55,2 % žiakov. Celkovo teda šlo z pohľadu IRT o priemerne obťažnú položku s výbornou rozlišovacou schopnosťou ($\alpha = 0,921$). Pri detailnej analýze výberu jednotlivých distraktorov (tabuľka 14) sme ale zistili, že distraktor A kladne koreluje s celkovým skóre. Distraktor si vybralo približne 11 % žiakov. Pri analýze distraktorov z pohľadu IRT sme zistili stúpajúcu tendenciu voľby distraktora A v skupine priemerných žiakov. Najlepší žiaci dokázali distraktor A spoľahlivo vylúčiť.



Obrázok 22. Graf citlivosti



Možnosti = A - červená; B- zelená; C - modrá; D - čierna

Obrázok 23. Graf distraktorov

Tabuľka 14. Tabuľka distraktorov

	A	B	C	D	X
P. Bis.	0,04	0,47	- 0,41	- 0,13	- 0,06
p	11,0 %	39,0 %	43,0 %	6,0 %	1,0 %
N	142,00	480,00	528,00	80,00	9,00

Ako možno vidieť na konkrétnych príkladoch analyzovaných úloh, výsledky jednotlivých prístupov (CTT a IRT) sa navzájom dopĺňajú. Výhodou IRT metód pri analýze distraktorov je detailnejší pohľad na preferencie distraktorov a zároveň možnosť analyzovať vzťah medzi narastajúcou schopnosťou žiakov a zmenou preferencie distraktorov.

Rozdielne fungovanie položiek

Mimoriadne dôležitým aspektom kvality úloh je ich férovosť. Úlohy samotné by nemali nijakým spôsobom znevýhodňovať určitú skupinu žiakov. Zjednodušene povedané, ak majú dvaja žiaci rovnakú úroveň schopnosti, pravdepodobnosť úspešného vyriešenia úlohy by mala byť rovnaká. Prostredníctvom komplexných štatistických metód (napríklad Mantel-Haenszel štatistika, Logistická regresia a pod.) je možné identifikovať položky, u ktorých tento predpoklad neplatí. Hovoríme o tzv. rozdielnom fungovaní položiek (z anglického Differential Item Functioning/DIF).¹²³ Otázkou je, ktoré faktory môžu ovplyvňovať správne vyriešenie úlohy.

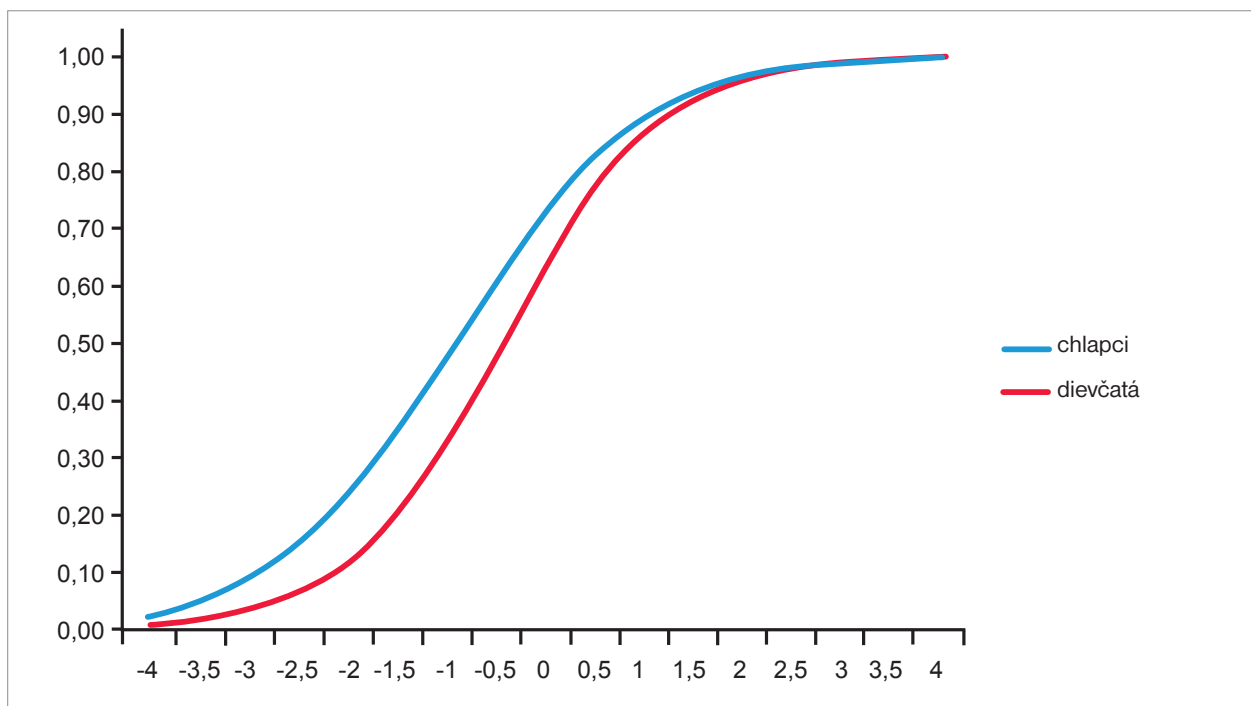
Najčastejšou analyzovanou premennou je pohlavie žiakov. Pri rozdielnom fungovaní úlohy môžu nastať dve situácie. Úloha medzi dvoma skupinami (chlapci a dievčatá) funguje odlišne z hľadiska obťažnosti (tzv. uniformný DIF) alebo funguje odlišne z hľadiska rozlišovacej schopnosti (tzv. neuniformný DIF). Na nasledujúcom obrázku sú znázornené obidve situácie. Porovnanie úspešnosti položiek medzi

¹²³Camilli, G. & Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. Thousand Oaks, CA: SAGE.

chlapcami a dievčatami často nemusí byť dostatočné na zhodnotenie férovosti položky. Existencia rozdielov v úspešnosti môže odzrkadľovať reálne rozdiely v meranej schopnosti, či rôznu úroveň skúseností s testovanou oblasťou medzi oboma skupinami. V takom prípade je férovosť položky zachovaná. Problém nastáva, ak pri rovnakej úrovni schopnosti majú skupiny rôznu pravdepodobnosť správnej odpovede, a tento rozdiel nevyplýva z meraného konštruktú.

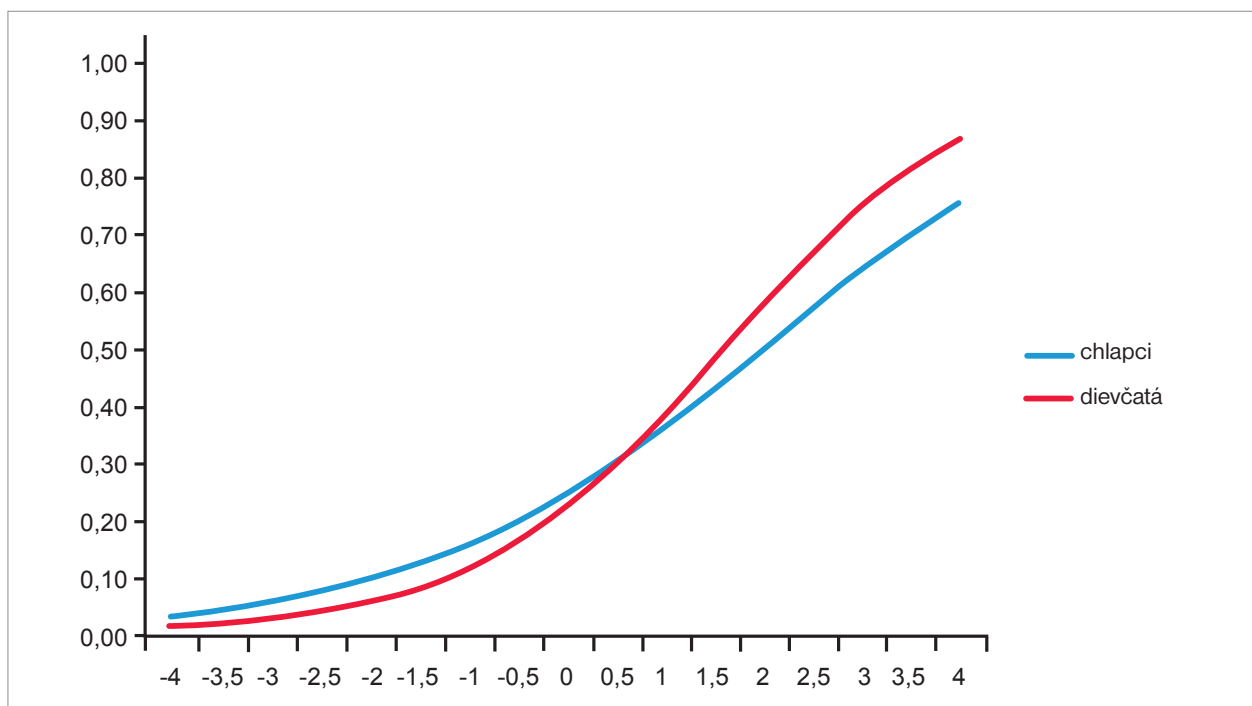
Ak úloha funguje odlišne medzi skupinou chlapcov a dievčat len v obťažnosti, hovoríme o uniformnom rozdielnom fungovaní položky (obrázok 24). V praxi ale nastáva aj situácia, keď je priemerná úspešnosť rovnaká, ale úloha medzi skupinami funguje odlišne v rozlišovacej schopnosti. Na obrázku 25 môžeme vidieť, že najslabší chlapci majú väčšiu pravdepodobnosť, že na položku odpovedia správne v porovnaní s najslabšími dievčatami. Situácia je však presne opačná v skupine najlepších žiakov. V tomto prípade možno pozorovať výrazne vyššiu pravdepodobnosť správnej odpovede u dievčat. Dôležitou súčasťou týchto analýz je aj zhodnotenie veľkosti rozdielneho fungovania úlohy (effect-size – vecná signifikancia rozdielu). Veľkosť rozdielu môže byť zanedbateľná až závažná. Na nasledujúcich obrázkoch môžete vidieť úlohy vykazujúce rozdielne fungovanie.

02. Počítačová zostava má veľkoobchodnú cenu 1200 eur. Maloobchodná cena je o 20 % vyššia ako veľkoobchodná cena. Vypočítajte maloobchodnú cenu počítačovej zostavy v eurách.



Obrázok 24. Ukážka odlišne fungujúcej úlohy a grafické zobrazenie uniformného DIF

- 19** Určte korene rovnice $\sin 2x = \sin x$ z intervalu $x \in (0^\circ; 360^\circ)$. Do odpoveďového hárka zapíšte v stupňoch súčet všetkých koreňov tejto rovnice z daného intervalu.



Obrázok 25. Ukážka odlišne fungujúcej úlohy a grafické zobrazenie neuniformného DIF

Prostredníctvom analýz položiek z Testovania 9 a EČ MS z rokov 2009 až 2013 sa nám podarilo identifikovať typické úlohy, ktoré vykazujú rozdielne fungovanie v skupine chlapcov a dievčat. Väčšinou šlo o úlohy s uniformným DIF (posun v obťažnosti položky). V nasledujúcich tabuľkách môžete vidieť, aké typy úloh spravidla zvýhodňujú chlapcov alebo dievčatá v testoch z matematiky a zo slovenského jazyka a literatúry.

Tabuľka 15.

Typické matematické úlohy vykazujúce rozdielne fungovanie medzi chlapcami a dievčatami

Merané vedomosti a zručnosti (obsahové zameranie položky)	Matematická operácia	Úlohy zvýhodňujúce chlapcov	Úlohy zvýhodňujúce dievčatá
Aplikácia rutinných matematických výpočtov v známych situáciách	výrazy, rovnice		x
Memorizácia	definície, vzorce		x
Riešenie viackrokových problémov (presnosť a systematický prístup)	slovné úlohy (známe situácie)		x
Viacnásobné riešenie problému	slovné úlohy („sedliacky rozum“)	x	
Nekonvenčné riešenie problémov	nezbierkové úlohy	x	
Kontextové úlohy	futbal, počítače, autá	x	

Tabuľka 16. Špecifiká položiek s rozdielnym fungovaním v testoch slovenského jazyka a literatúry

Merané vedomosti a zručnosti	Kontext	Úlohy zvýhodňujúce chlapcov	Úlohy zvýhodňujúce dievčatá
Jazyková zložka	Dialóg		
	Fakty		
	História		x
	Introspekcia		x
Literárna zložka	Dialóg		x
	Fakty		
	História	x	
	Introspekcia		x
Čítanie s porozumením	Dialóg		
	Fakty	x	
	História	x	
	Introspekcia		

* * *

Položková analýza je nevyhnutnou súčasťou procesu testovania v oblasti vzdelávania. Samotné zhodnotenie úspešnosti žiakov v danej úlohe je častokrát nedostačujúce. V našej práci rutinne aplikujeme obidva spomenuté postupy (CTT a IRT). Výsledky týchto analýz sa navzájom dopĺňajú a prinášajú detailnejší pohľad na samotné úlohy a test. Prostredníctvom rôznych štatistických metód je možné o jednotlivých úlohách získať podstatne detailnejšie informácie. Cieľom položkovej analýzy po pilotných testovaniach je predovšetkým identifikácia problémových úloh. V ideálnom prípade by sa do ostrých testovaní nemali takéto úlohy vôbec dostať. Popri analýzach pilotných testovaní je však potrebné realizovať položkovú analýzu aj po certifikačných testovaniach, nakoľko kontext testovania je úplne odlišný. Významným faktorom, ktorý ovplyvňuje výsledky testovania, je motivácia žiakov. Tá býva pochopiteľne výrazne vyššia v prípade certifikačných meraní (testy s veľkým vplyvom na študijnú kariéru žiaka). Detailná položková analýza, analýza výberu distraktorov, prípadne kvalitatívna analýza postupov pri riešení úloh prináša zároveň vzhľad do uvažovania a rozhodovania žiakov v procese riešenia úlohy.

Poznatky získané z týchto analýz je možné uplatniť pri tvorbe nových testov, cieľom ktorých je zredukovať počet potenciálne problematických úloh. Správy zo štatistických analýz slúžia ako podklad pre garantov a tvorcov úloh. Aby sme mohli výsledky certifikačných meraní pokladať za validné, reliabilné, objektívne a férové, musia byť tieto testy vytvorené z vysoko kvalitných úloh (či už z pohľadu štatistiky ale aj z pohľadu obsahového zamerania). Proces tvorby takýchto úloh patrí k najkomplexnejším a najzložitejším procesom v oblasti merania výsledkov vzdelávania.

6. BANKA ÚLOH PRO ADAPTIVNÍ TESTOVÁNÍ

Využití počítačů ve výuce

S tím, jak se nové ICT technologie rozvíjejí a adaptují, jsou stále více zapojeny do našich životů a nahrazují minulé aktivity a vzorce chování. Obvyklý názor, kterým je využití počítačů ve výuce posuzováno – „čím více tím lépe“ – by měl být podrobně zkoumán se zřetelem na sociální a osobnostní faktory rozvoje žáka. Počítačem podporované testování žáků přináší široké možnosti pro politiky, management škol, neomezené možnosti replikace testů, snadné řízení chování žáků, okamžitou realizaci výsledků a zpětné vazby. Tento typ testování, který je založen na využití počítačových modelů a algoritmů počítačem podporovaného testování, navíc přináší významný přínos ke zlepšení validity a reliability testu (Bartram, Hambleton, 2006). Nejrozšířenější a technologicky nejvyspělejší počítačem podporované (computer-based) testování je v současnosti Computerized Adaptive Testing (CAT).

V současnosti rozeznáváme několik typů podpory žákova učení počítači. Jsou to například tyto typy:

- **CAI** (computer-assisted instruction) – počítačem podporovaná výuka odpovídá počítačovým programům, které poskytují žákům možnosti opakování, cvičení, dril jednotlivých učebních aktivit.
- **CMI** (computer-managed instruction) – počítačem řízená výuka odpovídá programům, které nabízejí žákovi učební materiály, hodnotí jeho výkony a podle dosažených výsledků řídí další kroky žákova učení a zaznamenávají žákovy pokroky pro použití učitele. Jak CAI tak CMI nutně nevyžadují přítomnost učitele v průběhu učení žáka.
- **CEI** (computer-enhanced instruction) – počítačem rozšířená výuka, která poskytuje žákovi otevřenější možnosti, které podporují konkrétní výuku v určitém předmětu. V rámci tohoto typu počítačem podporované výuky je vždy přítomen učitel, který řídí učení se žákem a jeho využití počítačů ve výuce.

Adaptivní výukové systémy

Ve vzdělávání existuje několik typů technologií, které reprezentují adaptivní učení:

- **Diferencované učení** je ten případ, kdy je žákům k dispozici několik různých cest za poznáním, z nichž si každý (typicky na doporučení učitele) některou vybere.
- **Personalizované učení** znamená, že jsou výukové postupy nastavovány každému na míru na základě jeho schopností (znalostí) zjišťovaných před zahájením výukové aktivity (typicky formou pretestu). V průběhu výuky již k modifikacím nedochází.
- **Adaptivní učení** nastává tehdy, je-li výukový postup modifikován i v průběhu aktivity (např. testu) na základě bezprostředně získaných výsledků. [Jelínek, 2007].

Adaptivní diagnostika ve vzdělávání – modely

Pro adaptivní testování, ale také adaptivní učení se využívají v zásadě dva teoretické modely:

1. **Computer Adaptive Testing**, které je založeno na teorii odpovědi na položku - IRT (Item Response Theory). Teorie odpovědi na položku je sofistikovaná psychometrická teorie, která umožňuje přiřadit položky a žáky na jednu škálu a tak administrovat žákovi takové položky,

kteřé odpovídají jeho schopnostem. Tak je každému testovanému možno nabídnout kratší test se stejnou diagnostickou hodnotou (viz dále).

Nezbytnou součástí adaptivního testování založeném na modelu IRT, je banka úloh (někdy též položková banka). Dobrá banka úloh pokrývající všechny důležité aspekty sledovaného latentního rysu s jednorozměrnými úlohami zarovnanými na stejnou metrickou škálu je klíčová pro použití adaptivního testování. Wainer, Kiely (1987) uvádí, že slušný CAT algoritmus musí být postaven na bance úloh s minimálně 150 až 200 položkami, které jsou zarovnány na jednu škálu (viz dále).

2. **Knowledge spaces theory** – základní myšlenku modelu lze popsat příkladem: Představme si klasický didaktický test z matematiky sestávající ze sedmi lineárních rovnic. Je zřejmé, že prostá úroveň latentního faktoru „matematika“ získaná pomocí IRT nám příliš informací neposkytne. Ani kategorizace, závislá na předem určeném a relativně malém počtu kategorií, nemusí žákovi schopnosti popsat dokonale. Naopak detailní údaj, které položky v testu žák nezvládl, nám dá významně užitečnější informace (žák nezvládá dělení obou stran rovnice číslem). Cílem je tedy podávat takto obsažné informace, avšak v jednodušší (a tedy snáze interpretovatelné) podobě, než je seznam (ne)vyřešených úloh.

Informace týkající se znalostí určitého tématu je charakterizována jako množina otázek, problémů. Knowledge state (stav znalostí) jedince vzhledem k určitému tématu je formalizována jako podmnožina otázek (například v našem případě položek), které je testovaný schopen vyřešit. Množina stavů znalostí může tvořit znalostní prostor (knowledge space). Znalostní prostor příslušný dané relaci je množina stavů znalostí, která je uzavřená na sjednocení, obsahuje prázdný stav znalostí a úplný stav znalostí (množinu všech vlastností dané domény) a respektuje tuto relaci. Znalostní prostor je tedy částečně uspořádaný množina, která může být znázorněna Hassovými diagramy.

6.1. Obecný popis adaptivního testování

Teorie odpovědi na položku, IRT (Item Response Theory), poskytuje propracované techniky pro testování individuálních charakteristik probandů.¹²⁴ Tato nabízí vhodný aparát pro uplatnění obecných principů tzv. adaptivního testování. Základní idea adaptivního testování je jednoduchá a jasná: zadávejte testovanému žákovi k řešení pouze takové položky, které umožňují ideálně posoudit úroveň latentního rysu¹²⁵ (Linacre, 2000).

130

Ačkoli jsou **tradiční skupinová hodnocení**, zpravidla realizovaná na papíře, omezena rozsahem obtížnosti obsažených úkolů, bývají často považována za vhodnou metodu hodnocení široké škály schopností dětí. Zpravidla obsahují několik velmi obtížných a několik velmi snadných úloh pro většinu cílové skupiny. Pro dítě s nízkými schopnostmi se může jednat o zkušenost poškozující jeho sebeúctu; pro dítě výrazně nadané bude představovat skutečnou výzvu přinejlepším jen několik úloh. Takové testy **poskytují jen velmi málo informací o dětech, které se na škále nadání schopnostmi vyskytují na jejich krajích**, tyto testy mají rovněž vyšší toleranci chyb, což autorskými právy chráněné testy většinou opomíjejí (Tymms 1998).

Adaptivní hodnotící systémy obsahují větší počet úloh, které jsou přiměřené schopnostem žáků, a mnohem menší množství úloh, které jsou příliš snadné nebo příliš obtížné. Z tohoto důvodu mohou poskytnout v kratším čase přesnější a spolehlivější informace. Tato metoda hodnocení se obtížně realizuje na papíře, lze ji ale uskutečňovat formou rozhovoru vedeného na osobní bázi. Osobní rozhovory jsou ovšem pro hodnotitele časově náročné. Pokud je hodnotitelem učitel, který má třídu plnou žáků,

¹²⁴Jedinci, kteří jsou předmětem zkoumání.

¹²⁵Měřený pedagogický jev, konstrukt (např. vědomosti, dovednosti)

může být dokonce nemožné tuto metodu aplikovat, adaptivní hodnocení se proto ideálně uskutečňuje prostřednictvím počítače.

Hodnocení prostřednictvím počítače pracuje s předem určenými algoritmy – otázky jsou voleny na základě předešlých reakcí a odpovědí dítěte, které umožňují individualizovanější hodnocení (Viz např. Thissen a Mislevy 1990). Počítač se postupně soustřeďuje na úroveň schopností konkrétního dítěte – na zónu proximálního vývoje (Vygotskij, 2004) daného dítěte, v níž se pak může odehrávat další výuka.

Zkušenost tohoto typu testu je od tradičního testu velmi odlišná – test představuje výzvu, kterou je ale dítě schopné zdat, a je kratší.

Modely s latentní proměnnou

Pod označením modely¹²⁶ s latentními proměnnými rozumíme takovou skupinu statistických modelů, které popisují pozorované údaje pomocí jejich závislosti na nepozorované, matematicky zkonstruované charakteristice. Vzhledem k tomu, že řada jevů zkoumaných v pedagogice má teoretický charakter, je nutné tyto teoretické vlastnosti nějakým způsobem zobrazit. Provedeme to prostřednictvím formulování latentní proměnné, která bude formalizovaným obrazem teoretické vlastnosti mající věcný smysl. Teoretické vlastnosti existující pouze ideálně vytváříme zejména kvůli tomu, že potřebujeme nějakým způsobem vysvětlit vlastnosti empirické. Modely s latentními proměnnými představují poměrně univerzální metodu zkoumání struktury jevů a díky tomu se rozšířily v řadě vědních odvětví. Intelktové či osobnostní charakteristiky člověka se jen velmi těžko měří a tudíž musí být zpravidla modelovány na základě svých empirických projevů.

Základní modely IRT

Modely IRT jsou používány jako podklad pro statistický odhad parametrů, jež reprezentují polohu osob a položek na latentním kontinuu, neboli úroveň latentního rysu, již je možno přisoudit osobám a položkám. Jakmile jsou získány odhady relevantních parametrů, jsou obvykle provedeny statistické testy, jež mají určit, do jaké míry model s danými parametry umožňuje předpovídat odpovědi osob. Jinak řečeno jde o to, do jaké míry model a dané parametry mohou vysvětlit strukturu získaných dat.

Univerzálnost použití modelů s latentními proměnnými vyplývá především z jejich metodologické podstaty, kdy zajišťují konstrukci latentních proměnných na základě manifestních proměnných bez ohledu na to, jaký je jejich obsah. Aníž bychom brali v úvahu jednotlivé vědní obory, můžeme rozlišit zhruba následujících pět oblastí použití modelů s latentními proměnnými (Čepička, 2005):

1. *Odhad nepřímo měřitelných proměnných.* Tento typ použití lze označit jako základní, neboť právě k modelování latentních struktur byly modely s latentními proměnnými navrženy. Zprostředkovávají jejich měření a umožňují tak jejich kvantifikaci.
2. *Redukce počtu proměnných,* která slouží především ke zjednodušení popisu jevů. Tuto redukci můžeme provádět především na základě korelačních vztahů mezi redukovánými proměnnými. Existuje několik příčin, proč spolu proměnné korelují, ale pro oblast redukce přichází v úvahu především případy, kdy pozorované proměnné představují účinky vyvolané nějakou další proměnnou nebo kdy předpokládáme, že pozorované proměnné jsou indikátory téže vlastnosti, kterou měří a konečně jestliže existuje kauzální vztah mezi proměnnými.

¹²⁶Přes mnohoznačnost pojmu model jej můžeme charakterizovat jako zjednodušenou formu zobrazení zkoumaného úseku reality. Model je sestaven podle určitých pravidel, která dovolují napodobovat chování a vlastnosti zobrazované reality.

3. *Klasifikace proměnných.* Ta je spojena zejména s vytvářením nových, obecnějších druhových proměnných, které slouží k pojmenování shodných rysů původních proměnných. Ve faktorové analýze je míra této „společnosti“ vyjádřena prostřednictvím koeficientu komunality.¹²⁷

Modelů s latentními proměnnými, existuje celá řada. Vedle v minulosti nejvíce užívaných modelů klasické lineární faktorové analýzy jsou to například modely nelineární faktorové analýzy nebo modely latentní struktury. V tomto textu se budeme dále zabývat jen modely latentní struktury.

Modely latentní struktury se začínají objevovat v padesátých letech dvacátého století. Do té doby byly totiž používány k analýze binárních dat klasické modely faktorové analýzy, ale šířily se pochybnosti o jejich vhodnosti pro tyto případy.

Teorie odpovědi na položku (IRT) představuje skupinu matematických modelů, které popisují situaci interakce mezi jedincem a položkou. Modely, se kterými se v rámci IRT pracuje, lze na obecné úrovni rozdělit jednak podle počtu dimenzí, které jsou uvažovány, na jedno a vícedimenzionální, a také podle typu položek na modely pro dichotomní a polytomní položky. Více propracované jsou zatím modely pro dichotomně (resp. binárně) skórované položky, neboť IRT původně vzniklo pro účely zefektivnění a zkvalitnění výkonového testování. Vzhledem k tomu, že se IRT dobře osvědčilo v oblasti výkonových testů, byly posléze intenzivně vyvíjeny i modely pro složitější typy položek s více než dvěma skórovacími kategoriemi (např. posuzovací škály „Likertova“ typu typicky užívané v diagnostice osobnosti, postojů, apod.)

Základní typy IRT modelů pro dichotomní položky jsou tzv. jednoparametrový logistický (1PL) model (někdy zvaný též Raschův model), kdy jsou položky charakterizovány pouze jedním parametrem – obtížností, dvouparametrový logistický (2PL) model s parametry obtížnosti a rozlišovací účinnosti¹²⁸ a tříparametrový logistický (3PL) model s parametry obtížnosti, rozlišovací obtížnosti a pseudouhádnutelnosti.

6.2. Banka úloh

Při masových testováních se objevují desítky různých variant jednoho a téhož testu. Tyto varianty se snaží být podobné v tom smyslu, že se odlišují pouze konkrétním obsahem úloh a počtem zadání a jeho forma je stejná a základní statistické charakteristiky jsou podobné. Nicméně v praxi není možná dosáhnout úplné shody, a proto převažují takové varianty, které jsou ekvivalentní a navzájem záměnné. Proto nabývá mimořádné důležitosti otázka rovnocennosti výsledků získaných testováním v rámci vyplnění různých variant testu.

Současná teorie modelování a parametrizace testů IRT umožňuje realizovat proceduru srovnání ukazatelů různých variant a zajistit škálování na jediné metrické škále.

Základy teorie srovnání v rámci IRT jsou položeny v klasické práci B. D. Wrighta (1979) a N. N. Stonea a ve srovnatelné nedávné práci M. J. Kolena a R. L. Brennana (1995). S různými metodami srovnání je možné se seznámit v pracích autorů R. M. Smith, 1992 a W. Zhu, 1989 a dalších. V těchto pracích jsou představeny i postupy, které mohou být použity v klasické teorii testování (KTT). Nejznámější je metoda lineárního srovnání a metoda ekvipercilového srovnání. Nicméně uplatnění těchto metod, mimo známých principiálních ohraničení KTT, potřebuje ještě doplňující zjištění z ekvivalentnosti skupin respondentů, kteří vyplňují srovnávací varianty testů. V současnosti se zkoumají metody i procesy horizontálního srovnávání, které jsou v souladu s Raschovým modelem.

¹²⁷ komunalita – část rozptylu indikátora, kterou je možné vysvětlit působením skrytých faktorů

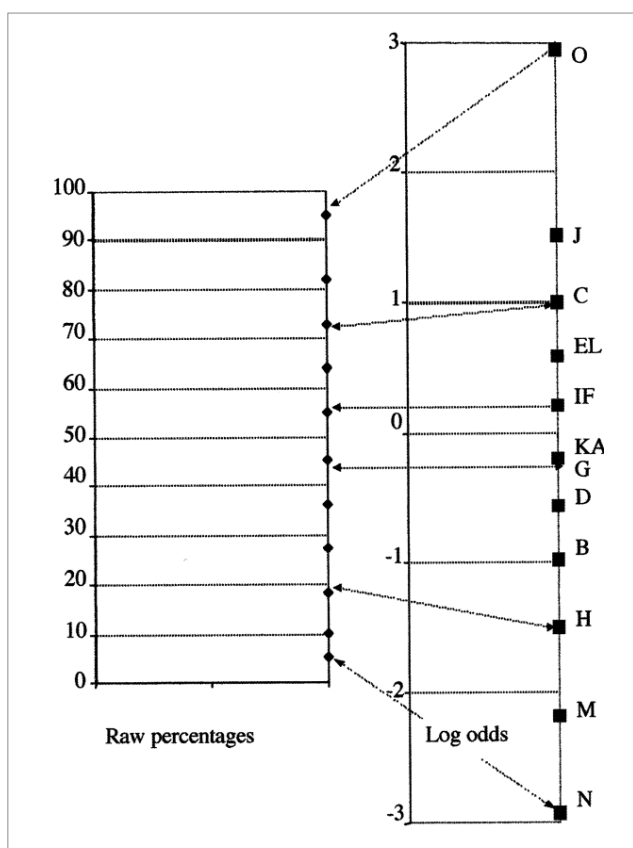
¹²⁸ rozlišovací schopnost

6.2.1. Tvorba banky úloh

Použití různých variant stejného testu

Základním požadavkem kladeným na banku úloh je dostatečný počet úloh, jejichž parametry jsou umístěny na společné škále a jsou jednodimenzionální. Aby mohla být tato podmínka splněna, musí být úlohy z různých testování na jednu škálu srovnány. Potřeba srovnání vzniká především při využívání různých variant jednoho a téhož testu. S pomocí metod matematické statistiky je možné ověřit hypotézu o stejné náročnosti souhlasných zadání různých variant stejného testu.

Při využití matematických modelů IRT úrovně schopností žáků a úrovně náročnosti úkolů testu mohou být zobrazeny na jedné metrické škále doprovázené charakteristikou přesnosti měření. Nicméně libovolná varianta testu bude mít svou vlastní škálu. Toto lze vysvětlit tím, že škála logitů¹²⁹ (viz obr. 1), na které jsou výsledky respondentů a parametry zadání je intervalová. Je známo, že intervalové škály nemají absolutní nulu. Tudíž na každé ze škál existuje neurčitost s výběrem nuly, která se odstraňuje procesem kalibrování. Existují různé způsoby provedení, často se při uplatnění Raschova modelu pro eliminaci uvedené neurčitosti každá škála centruje takovým způsobem, aby se suma obtížnosti zadání rovnala nule. To znamená, že škála se centruje ve střední hodnotě náročnosti zadání testu. Je zřejmé, že různé varianty testu mají různé střední hodnoty náročnosti zadání, v důsledku čehož vznikají problémy s porovnáním stupňů, získaných z různých variant testů.



Obrázek 26. Relativní schopnosti žáků A až O znázorněny ve výsledcích v podobě hrubého skóru (vlevo) a jako pravděpodobnost v logitech (zdroj: Bond, Fox, 2001).

Pokud chceme srovnat jednotlivé varianty na jednu škálu, použijeme speciální proceduru, která umožňuje nastavit spojení mezi testováním v různých variantách testů a umístit je na jednu širší škálu.

¹²⁹Logit je logaritmus podílu $(\pi/(1 - \pi))$; π je hrubý skór testovaného vyjadřujícího pravděpodobnost, že veličina Y (testovaný odpoví správně) nabývá hodnoty 1. Jde o logaritmickou transformaci hrubého skóru.

Ve výsledné míře založené na odpovědích stejné varianty testu, může být nastavena míra založená na odpovědích druhé varianty testu a závěry testování nejsou závislé na zvolené variantě testu. Jinak řečeno konečný stupeň zkoušení se základní úrovni nastavení bude stejný, nezávisle na variantách testu a nehledě na to, jak se mohly první stupně odlišovat. V literatuře se srovnání různých variant jednoho a téhož testu nazývá horizontálních srovnáním.

6.2.2. Metody pro vytvoření banky úloh

Banka úloh je množství kalibrovaných úloh (tj. úloh se známými parametry), ze kterých mohou být vytvořeny různé varianty stejného testu. Srovnání různých testů a vytvoření banky úloh jsou podobné úlohy, neboť obě procedury si kladou za cíl umístit měřené parametry z Raschova modelu na obecnou škálu. Rozdíl je pouze v nutnosti umístit na jednu škálu žáky vyplňující různé varianty testu. Cíl vytvoření banky úloh spočívá tedy v sestavení úloh z různých testů na jednu škálu. Takto banka úloh představuje množinu testových úloh, jejichž parametry jsou umístěny na jednotnou škálu tak, že různé podmnožiny těchto zadání poskytují pro testované žáky odpovídající stejný výsledek. Při správně sestavené bance úloh odpadá nutnost dalšího srovnávání zkoušených. Testové úlohy je však možné dát do banky úloh za předpokladu, že jsou srovnatelné, čehož dokážeme procedurou srovnání a umístěním na obecnou škálu.

6.2.3. Podmínky nutné ke srovnávání úloh

V literatuře se uvádí 4 podmínky nutné ke srovnávání úloh:

- 1. Testování stejné charakteristiky.** Testy u kterých je možné srovnání, musí měřit stejnou charakteristiku (hranice možností, schopnosti, úroveň přípravy v určité znalosti). Testy s různým obsahem nelze srovnávat, i když je lze škálovat. Jinými slovy nelze srovnat testy z matematiky a ruského jazyka.
- 2. Neměnnost hodnocení zkoušených** v souvislosti se zadáním testů. Hodnocení zkoušených musí být nezávislé na testu; z testu bychom měli získat odpovídající výsledky průběžného hodnocení.
- 3. Neměnnost charakteristik úloh** ve vztahu k výběru zkoušených. Výsledky kalibrace úloh a srovnávání nesmí záviset na vzorku zkoušených.
- 4. Nezávislost na metrice vybrané škály.** Škála, která je vybrána jako společná, nesmí ovlivňovat srovnání hodnot ve kterékoliv pozici. Jinak řečeno, interpretace testových bodů nesmí záviset na tom, co je vybráno jako základ při srovnávání a případné přetvoření škál musí být lineární (bez nelineárních deformací). Je zřejmé, že dodržení všech těchto podmínek je možné pouze u Raschova modelu, pouze s pomocí tohoto modelu je možný vědecký přístup k testování jako měření, čehož se týkají naše doporučení.

6.2.4. Metody srovnávání

Jak uvádíme výše, škála logitů, na které je promítnuto měření schopností zkoušených i obtížnost úkolů, je intervalová a nemá daný počátek. Z matematického hlediska to znamená, že parametry θ_n (úroveň schopností zkoušeného n) i b_i (úroveň obtížnosti i -tého testu) mají vzestupnou tendenci:

$$\theta_n = \theta'_n + \alpha, b_i = b'_i + \alpha \quad (1)$$

Pravdivost toho, že zkoušený s úrovní θ_n správně odpoví na úlohu b_i , v základním jednoparametřovém modelu Rasche se vyznačuje vztahem, který je neměnný ve vztahu k (1):

$$P(U_n = 1 | \theta_n, b_i) = \frac{1}{1 + e^{-(\theta_n - b_i)}} \quad (2)$$

Pro odstranění nepřesnosti s výběrem začátku sumarizace se obvykle používá centrování škály ve střední hodnotě obtížnosti všech úkolů dané varianty testu, tzn. při využití parametrů θ_n a b_i předpokládáme, že:

$$\bar{b} = \frac{\sum_{i=1}^k b_i}{k} = 0 \quad (3)$$

k – počet úloh v testu

Předpokládejme, že jsou nám známy výsledky testování dvou zkoušených se stejnou úrovní znalostí: první odpovídal na zadání první varianty a druhý na zadání druhé varianty. Necht' úroveň zkoušeného v první variantě je θ_1 a úroveň druhého zkoušeného je θ_2 . Protože tato hodnocení se vztahují ke stejné úrovni schopností, pak odlišnost mezi hodnotami θ_1 a θ_2 se shoduje s rozdíly mezi začátky dvou škál. Když určíme rozdíl $m = \theta_2 - \theta_1$, pak můžeme převést všechny významy parametrů zkoušených a úloh ze škály prvního testu na škálu testu druhého.

Analogická situace nastane, pokud známe odpovědi zkoušených na dvě zadání se stejnou úrovní obtížnosti. Identifikovat dvě zadání se stejnou úrovní nebo dva zkoušené se stejnou úrovní připravenosti je v praxi nemožné. Je však možné stejné zadání vložit do obou variant anebo dát stejnému žákovi obě varianty testu.

Na těchto tvrzeních jsou založeny dvě nejvíce používané metody srovnání:

1. metoda společných úloh (někdy též kotvících položek),
2. metoda společných zkoušených.

Metoda společných úloh

Metoda společných úloh je jednou z nejpoužívanějších metod srovnávání různých variant testů. Při tomto přístupu spojujeme dvě varianty, které mají část úkolů společných.

Jsou známy dva způsoby využití této metody:

1. Při přípravě úloh má každá varianta testu pár úkolů společných s druhými variantami. Tyto úkoly mohou být společné pro všechny varianty testu nebo jen pro některé z nich. Tímto způsobem každá varianta testu má podmnožinu úkolů, které se překrývají s druhými variantami a ostatní jeho úlohy jsou samostatné, tj. patří jen do jedné varianty. Tato metoda se jeví jako nejpoužívanější.
2. V případě, kdy použijeme test a jeho varianty nemají společné úlohy, pak použijeme jiný postup. Uživatel testu doplňuje každou jeho variantu tím samým množstvím úkolů (lépe již se známými parametry obtížnosti, tj. z kalibrované banky úkolů). Tento počet úkolů je svazujícím článkem mezi různými variantami testu. Tento přístup se doporučuje používat velmi obezřetně, protože je nutné garantovat stejnou validitu testu a doplnitelného množství úkolů.

Metoda společných testovaných

Metoda společných testovaných se používá méně díky své nepraktičnosti. Nicméně jestliže z různých důvodů nemůžeme využít metodu společných zadání, pak použijeme metodu společných testovaných, kdy jedné a téže skupině zkoušených předložíme vyplnění dvou variant testů. Je známo několik variant této metody:

1. Zapojení všech zkoušených do procesu srovnávání, tzn. dvě varianty podléhající srovnání se předloží stejné skupině zkoušených. Zde jsou možné dva přístupy:
 - a) stejná skupina/stejný čas – obě varianty se spojí v jeden text a ten se dá zkoušeným;
 - b) stejná skupina/jiný čas – zkoušení vyplňují zadání každé varianty v různé době.
2. Spojení obou variant se uskuteční s pomocí podmnožiny zkoušených, vyplňující obě varianty testu. Tento přístup je v podstatě analogický metodě společných zadání.

Další metody srovnávání

V literatuře se vyskytují i další metody srovnávání, které však mají specifitější charakter a jejich oblast použití je ohraničena.

U metody ekvivalentních skupin je vztah mezi dvěma variantami testu založen na předpokladu, že různé skupiny zkoumaných žáků plnících úkoly těchto variant jsou ekvivalentní, tzn., že schopnosti zkoumaných jsou podobné. Teoreticky mohou být ekvivalentní skupiny vybrány stratifikací výpisků z kterýchkoliv důležitých poznatků. Druhý přístup při utvoření takových skupin je náhodný výběr z obecného počtu zkoumaných žáků. Avšak prakticky je utvoření ekvivalentních skupin stejně obtížné, jako vytvoření paralelních variant testu. Rozdíly ve schopnostech zkoumaných žáků mají vliv na proceduru srovnávání a srovnatelnost hodnot zkoumaných žáků a zadání úloh. Při využití této metody srovnávání je nutné jako doplňující zkoumání problému smíšených hodnot pro zachování ekvivalentnosti skupin.

6.3. Adaptivní testování – Computer Adaptive Testing

Historie adaptivního testování

Kořeny adaptivního přístupu k testování osob sahají až na počátek 20. století. Alfred Binet při konstrukci svých inteligenčních testů zavedl jednoduché pravidlo – sdružoval položky do bloků dle obtížnosti. Trénovaný examinátor musel nejprve odhadnout individuální úroveň probandových schopností, aby určil adekvátní začátek testu. Poté, co byl zadán odpovídající blok položek, examinátor rozhodl, zda budou následovat položky snadnější nebo obtížnější. Na konci takového procesu byly odhadnuty dvě úrovně probandových schopností – úroveň bazální (*basal level*) a úroveň stropní (*ceiling level*). Protože bazální a stropní úroveň se u různých lidí lišily, výsledné sety administrovaných položek tak byly přizpůsobeny úrovni konkrétní osoby. V současnosti jsou adaptivní testy administrovány pomocí počítače, neboť jsou schopny v reálném čase provést řadu složitých matematických postupů v rámci naprogramovaných algoritmů, a tím nahradit rozhodovací proces trénovaného administrátora (Jelínek 2007).

Základní idea adaptivního testování je tedy jasná a přímočará – účelem je zadávat testované osobě k řešení pouze takové položky, které umožňují ideálně posoudit úroveň latentního rysu (Linacre, 2000). Praktickým důsledkem tohoto přístupu je efektivnější procedura testování ve srovnání s konvenčními testy s pevně stanoveným počtem a pořadím položek, neboť **položky s minimální informační hodnotou pro danou osobu nejsou vůbec administrovány**.

Adaptivity testu je možné dosáhnout různými způsoby, nicméně v zásadě existují dvě skupiny strategií – dvouúrovňové (*two-stage*) a víceúrovňové (*multi-stage*) strategie. Dvouúrovňové strategie patří mezi adaptivními přístupy k těm nejstarším a nejjednodušším. Tyto strategie zahrnují dvě etapy – administraci rozřazovacího (*routing*) testu a posléze samotného měřícího (*measurement*) testu. Rozřazovací test je většinou poměrně krátký a slouží k počátečnímu odhadu úrovně probandových schopností. Na základě výsledku tohoto testu je respondentovi předložena jedna z variant měřícího testu, která svou obtížností nejlépe odpovídá provedenému odhadu schopnosti.

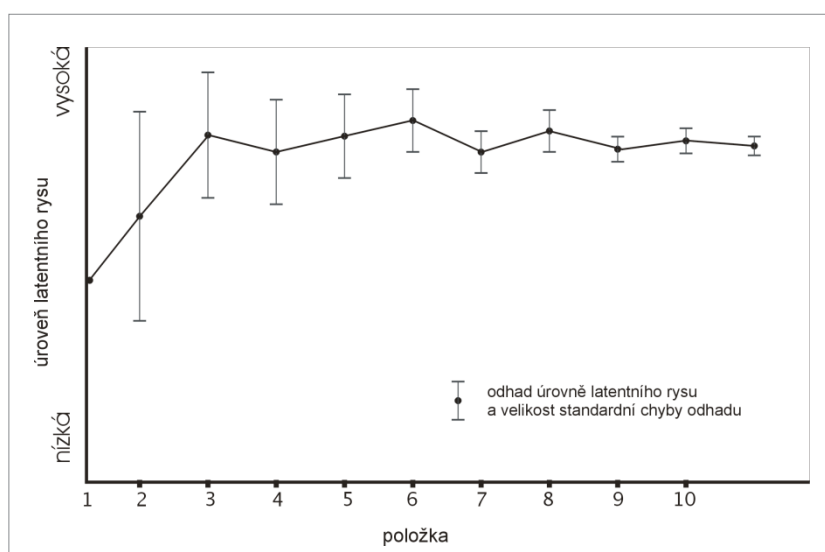
Algoritmus adaptivního testování

Jedním z nejdůležitějších předpokladů dobrého adaptivního testu je dostatečná velikost položkové banky, za jakési minimum je v tomto případě považováno 100 položek. Pro jednotlivé položky je třeba v procesu kalibrace odhadnout jednotlivé parametry (v závislosti na použitém IRT modelu). Toto odhadování parametrů je nutné provádět na dostatečně velkém souboru osob, nicméně charakteristiky položek nejsou na tomto souboru závislé, a měření vlastností probanda již není nutné interpretovat pouze v rámci populace, pro kterou byl test standardizován. Vzhledem k tomu, že uvažování v rámci IRT je většinou jednodimenzionální, bývá potřeba při budování položkové banky řešit problém multidimenzionality. Existují různé strategie, jak se s tímto problémem vypořádat, například vybalancováním obsahu (*content-balancing*) nebo izolací dimenzí do jednotlivých subtestů.

Po úvodním odhadu úrovně latentního rysu (např. na základě odpovědí na několik náhodně vybraných položek) jsou v rámci administrace adaptivního testu na základě předem zjištěných parametrů vybírány takové položky, které o daném jedinci s určitou odhadovanou úrovní latentního rysu podávají maximální množství informace. Vzhledem k faktu, že nejcitlivěji položka rozlišuje mezi osobami, u kterých se úroveň latentního rysu nachází v blízkosti hodnoty obtížnosti dané položky (a proband má tak přibližně 50 % šanci uspět, respektive odpovědět diagnosticky), počítač vybírá z položkové banky ty položky, které kritérium s určitou tolerancí splňují. Nicméně někdy bývá toto psychometrické kritérium v případě výkonových testů zmírněno až na úroveň 60, 70 nebo dokonce i 80 procentní pravděpodobnosti správné odpovědi, aby byl klient dostatečně motivován a měl po dokončení testu pocit optimálního výkonu; délka testu se takovou změnou prodlouží při zachované přesnosti měření přibližně o 10 – 20 % (Jelínek, 2007).

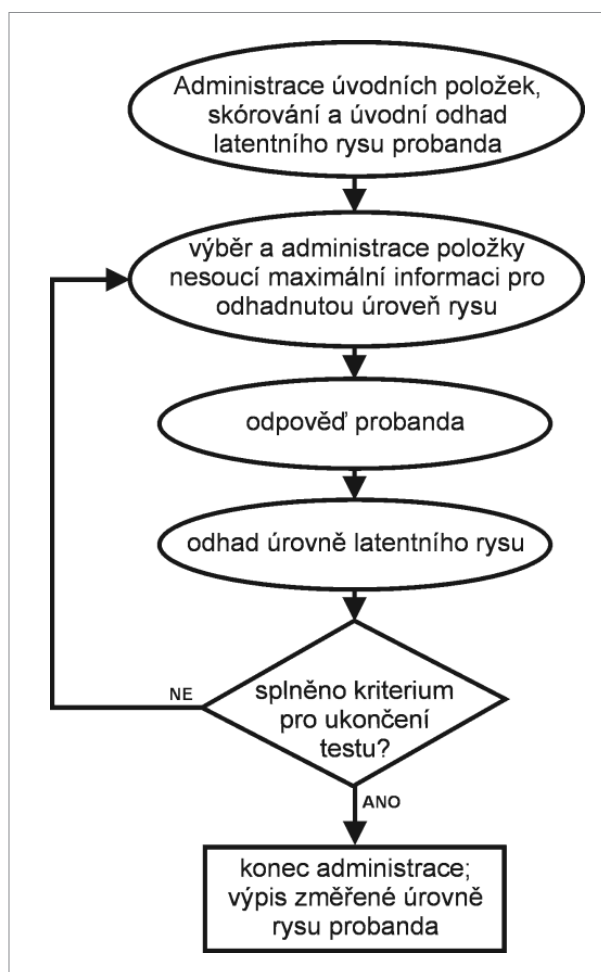
Administrace položek pokračuje do doby, než je splněno kritérium pro ukončení testu, kterým může být dosažení předem stanovené přesnosti odhadu rysu probanda nebo zařazení probanda s dostatečnou spolehlivostí vzhledem k předem definované úrovni latentního rysu, resp. je administrován maximální předem zvolený počet položek.

Jádro CAT tvoří algoritmus pro výpočet úrovně latentního rysu (z dosud administrovaných položek), a tedy i pro selekci následujících položek. Pokaždé, když respondent zodpoví novou položku, je přepočítána úroveň jeho schopnosti a dle její míry vybrána další.



Obrázek 27. Postupné zpřesňování odhadu úrovně latentního rysu (Zdroj: Jelínek, 2007)

Počítačem podporované adaptivní testování je realizováno podle následujícího algoritmu (viz obrázek 28). Postup realizace adaptivního testování je dále rozepsán v krocích 1 až 10.



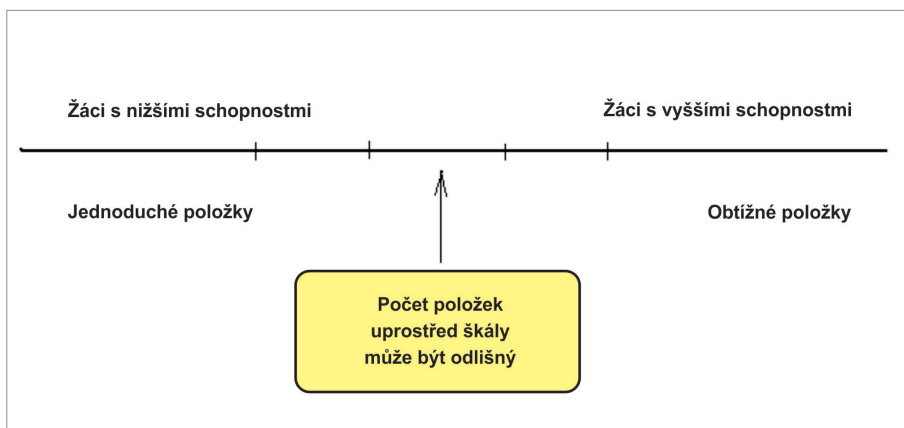
Obrázek 28. Algoritmus adaptivního testování (Zdroj: Jelínek, 2007)

1. Jak určit výchozí hodnotu Theta?¹³⁰

Jako výchozí odhady můžeme zvolit expertní hodnocení, například známky udělované učitelem. Tyto známky však nebudou zobrazeny na logitové stupnici. Možné řešení: požádat učitele aby kategorizoval studenty v jiném měřítku – například velmi schopný, průměrný a slabý. Úrovní může být 5 – v souladu se školními známkami. Poté bychom měli kategorizovat všechny položky, které máme k dispozici, do obtížnostní úrovně podle stejného počtu kategorií. Budeme mít tedy k dispozici několik skupin položek.

Pro každou skupinu položek spočítáme průměrnou obtížnost. Tuto průměrnou obtížnost můžeme považovat za výchozí hodnotu theta (Theta 1) pro zkoušeného v daném shluku. Představme si například zkoušeného, kterého jeho učitel klasifikoval jako velmi schopného studenta (studenti napravo v obrázku níže). V tomto případě bude jeho výchozí hodnota theta (Theta 1) odpovídat průměrné obtížnosti položek, které byly klasifikovány jako položky obtížné.

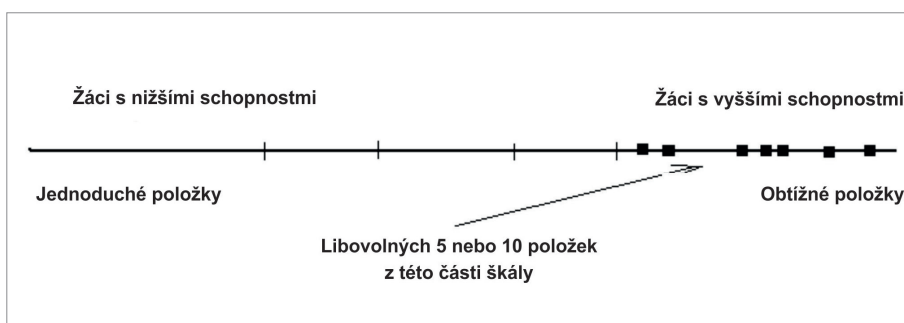
¹³⁰úvodní odhad latentního rysu (v našem případě schopností žáka)



Obrázek 29. Určení výchozí hodnoty theta

2. Krok 1: Jak zvolit první položku/položky?

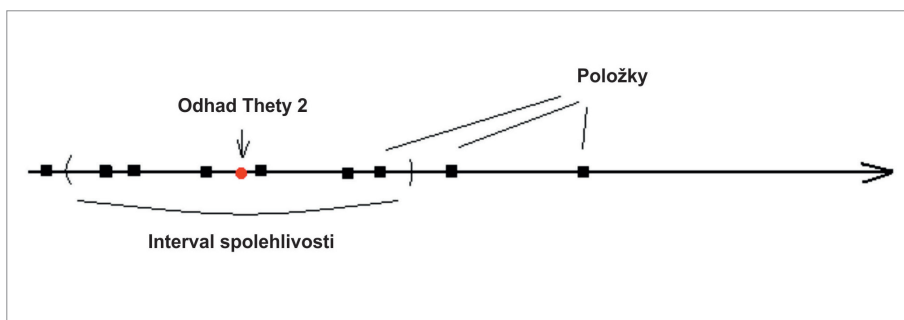
Na začátku zkušební procesu použijeme skupinu položek (5 či dokonce 10) namísto jediné. Všechny tyto položky by měly být zvoleny náhodně ze skupiny, pro kterou byl zkoušený klasifikován. Viz například obrázek níže s velmi schopným studentem.



Obrázek 30. Určení výchozí položky

3. Odhad Theta 2

Odhadujeme novou hodnotu theta (Theta 2) a SEM (standardní chybu měření). Bereme v potaz Thetu 1 jako výchozí hodnotu iteračního algoritmu (viz níže) a pracujeme se všemi položkami, které byly tomuto studentovi dány. Po odhadu vytvoříme interval spolehlivosti pro Thetu 2 (kde by se měla nacházet skutečná úroveň schopností žáka). Viz obrázek níže.



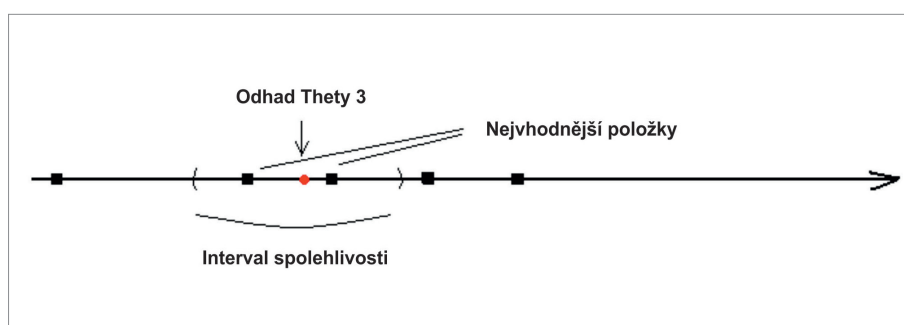
Obrázek 31. Druhý odhad parametru theta

4. Krok 2: Jak zvolit následující položku?

Je třeba poznamenat, že v této fázi bude SEM velmi velká, tedy i interval spolehlivosti bude velmi široký. V druhém kroku vezmeme jakýchkoliv 5 položek z daného intervalu spolehlivosti (zkráceně IS). Pokud se v intervalu nenachází 5 položek, můžeme zvolit položky, které jsou hranicím IC nejbližší.

5. Odhad Theta 3

Odhadujeme novou hodnotu theta (Theta 3) a SEM. Bereme v potaz Thetu 2 jako výchozí hodnotu iteračního algoritmu a pracujeme se všemi položkami, které byly tomuto studentovi dány (v kroku 1 a 2). Po odhadu vytvoříme interval spolehlivosti pro Thetu 3. V této fázi by měla být SEM i interval spolehlivosti menší než po kroku 1.



Obrázek 32. Další odhad parametru theta

6. Krok 3

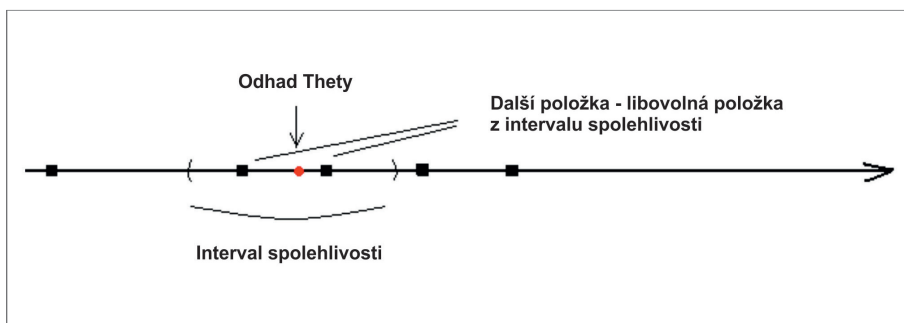
Nyní vybíráme pokaždé pouze jednu položku. Může to být jakákoliv položka z intervalu IS (viz obrázek výše). Pokud se v intervalu nenachází žádná položka, můžeme zvolit položku, která je hranicím IS nejbližší.

7. Odhad Theta 4

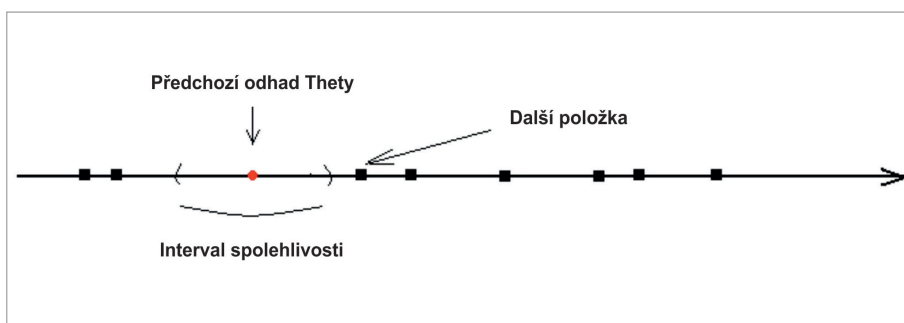
Odhadujeme novou hodnotu theta (Theta 4) a SEM. Bereme v potaz Thetu 3 jako výchozí theta hodnotu iteračního algoritmu a pracujeme se všemi položkami, které byly tomuto studentovi dány (v kroku 1, 2 a 3). Po odhadu vytvoříme interval spolehlivosti pro Thetu 4.

8. Další algoritmus bude cyklus

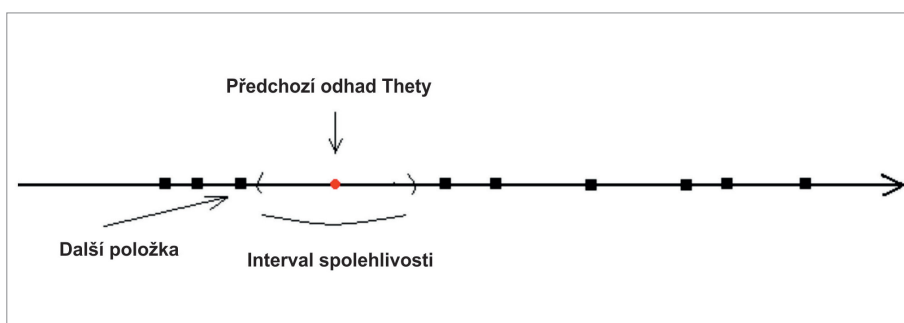
Další položka může být zvolena kterákoli z IS (viz obrázek 33 níže). Pokud v intervalu nejsou žádné položky, můžeme vybrat položku nejbližší intervalu: pokud byla předchozí položka dokončena správně, další položka by měla být napravo od IS (viz Obrázek 8); pokud byla předchozí položka dokončena chybně, další položka by měla být nalevo od IS (viz Obrázek 9).



Obrázek 33. Volba položky, pokud se nacházejí uvnitř IS



Obrázek 34. Volba další položky, pokud se uvnitř IS nenacházejí žádné položky a předchozí odpověď byla správná.



Obrázek 35. Volba další položky, pokud se uvnitř IS nenacházejí žádné položky a předchozí odpověď byla chybná.

9. Odhad Theta

Odhadujeme novou hodnotu theta a SEM. Bereme v potaz Thetu z předchozího kroku jako výchozí theta hodnotu iteračního algoritmu a pracujeme se všemi položkami, které byly tomuto studentovi dány (ve všech krocích). Po odhadu vytvoříme interval spolehlivosti pro novou Thetu.

10. Kontrola stop pravidla

Pokud ho není dosaženo, proces pokračuje opakováním kroků 8 a 9. Jakmile je ho dosaženo, zkoušení končí. Poslední odhad hodnoty Theta a její SEM se považují za studentův výsledek testu a SEM daného výsledku. Budou na logitové stupnici. Běžně jsou všechny odhady převedeny z logitové stupnici na jinou vhodnou stupnici (můžeme detailněji rozvést později).

Matematicky lze vyjádřit kroky 3 až 10 takto:

Hodnocení úrovně schopnosti žáka

Zadání

Máme dán test, který se skládá z n úloh, přičemž jejich obtížnost je stejná a předem známá $b_1 \dots \dots b_n$. V případě správného vyřešení úlohy dostává účastník testování jedničku, v opačném případě nulu. Nechtě jsou hodnocení odpovědí i -tého účastníka testování rovny:

$$x_{i1}, \dots, x_{in}; x_{ij} \in \{1, 0\}$$

V rámci jednoparametrického modelu Rasche je pravděpodobnost toho, že i -tý účastník testování správně vyřeší j -tou úlohu rovna:

$$[P_{ij}]_t = \frac{\exp([\theta_i]_t - b_j)}{1 + \exp([\theta_i]_t - b_j)} \quad (4)$$

Pravděpodobnost toho, že úloha nebude žákem vyřešena je rovna:

$$[Q_{ij}]_t = 1 - [P_{ij}]_t = \frac{1}{1 + \exp([\theta_i]_t - b_j)} \quad (5)$$

Pokud vezmeme v úvahu vzorce (4) a (5), můžeme vypočítat pravděpodobnost toho, že i -tý účastník získal jako výsledek j -té úlohy hodnotu x_{ij}

$$L_{ij} = P_{ij}^{x_{ij}} Q_{ij}^{i-x_{ij}} = \frac{\exp(x_{ij}(\theta_i - b_j))}{1 + \exp(\theta_i - b_j)} \quad (6)$$

kde b_j je obtížnost j -té úlohy a úroveň schopností i -tého žáka.

Zdůvodnění

Pokud použijeme vzorce (6) a za předpokladu o lokální nezávislosti můžeme vyjádřit pravděpodobnost, že za vyplnění všech úloh získá žák s úrovní schopností θ_i hodnocení x_{i1}, \dots, x_{in} .

$$L_i = \prod_{j=1}^n L_{ij} \prod_{j=1}^n \frac{\exp(x_{ij}(\theta_i - b_j))}{1 + \exp(\theta_i - b_j)} \quad (7)$$

Jestliže pohlížíme na L_i jako na funkci θ , získáme $L_i(\theta_i)$ jako funkci pravděpodobnosti. V souladu s metodou největší pravděpodobnosti jako hodnocení neznámého parametru θ_i vezmeme tu hodnotu, při které funkce pravděpodobnosti dosáhne maxima. Jelikož funkce L_i dosahují svého maxima v jednom bodě, budeme dále hledat maximum funkce $\ln(L_i)$. Hledání maxima funkce jedné proměnné budeme řešit pomocí iterační metody Newtona-Rafsona. Proto je třeba vyčíslit první a druhou odvozenou funkci pravděpodobnosti, přesněji, logaritmu funkce pravděpodobnosti:

$$\frac{d \ln(L_i)}{d \theta_i} = \sum_{j=1}^n (x_{ij} - P_{ij}) \quad (8)$$

$$\frac{d^2 \ln(L_i)}{d \theta_i^2} = - \sum_{j=1}^n P_{ij} Q_{ij} \quad (9)$$

Hodnotu θ_i nalezneme v rámci následujícího iteračního procesu, kde t je číslo předchozí iterace:

$$[\hat{\theta}_i]_{t+1} = [\hat{\theta}_i]_t - \left[\frac{d^2 L}{d\theta_i^2} \right]_t^{-1} \cdot \left[\frac{dL}{d\theta_i} \right]_t \quad (10)$$

K zastavení dochází, když hodnoty získané ve dvou sousedních iteracích se liší nevýznamně.

Pro hodnotu chyby úrovně schopností můžeme využít následující vzorec:

$$Err^2(\theta_i) = \frac{1}{-E\left(\frac{d^2 \ln(L_i)}{d\theta_i^2}\right)} \quad (11)$$

Popis algoritmu adaptivního testování

Jestliže i -tý účastník testování správně vyřešil všechny úlohy, tj. $\sum_{j=1}^n x_{ij} = n$, nelze získat jeho hodnocení úrovně schopností, je rovno $+\infty$.

Jestliže i -tý účastník testování nevyřešil ani jednu úlohu, tj. $\sum_{j=1}^n x_{ij} = 0$, nelze rovněž získat jeho hodnocení úrovně schopností, je rovno $-\infty$.

Jestliže i -tý účastník testování vyplnil úspěšně pouze část úloh, hodnocení jeho úrovně připravenosti můžeme získat s pomocí iteračního postupu následně:

Jako počáteční přiblížení hodnocení úrovně připravenosti vezmeme:

$$[\theta_i]_0 = \ln \left(\frac{\sum_{j=1}^n x_{ij}}{n - \sum_{j=1}^n x_{ij}} \right) \quad (12)$$

Dále, iteraci $(t+1)$ získáme z hodnoty získané při předcházející iteraci dle vzorce:

$$[\theta_i]_{t+1} = [\theta_i]_t + \frac{\sum_{j=1}^n (x_{ij} - [P_{ij}]_t)}{\sum_{j=1}^n [P_{ij}]_t [Q_{ij}]_t} \quad (13)$$

Pro vzorec (13) platí:

$$[P_{ij}]_t = \frac{\exp([\theta_i]_t - b_j)}{1 + \exp([\theta_i]_t - b_j)} \quad (14)$$

$$[Q_{ij}]_t = 1 - [P_{ij}]_t = \frac{1}{1 + \exp([\theta_i]_t - b_j)} \quad (15)$$

Iterační proces se zastaví, když druhá mocnina schopnosti žáka θ_i , získaná ve dvou sousedních iteracích bude menší než zadaná hranice, a to:

$$([\theta_i]_{t+1} - [\theta_i]_t)^2 < 0,00001 \quad (16)$$

Přitom přesnost získaného hodnocení úrovně připravenosti můžeme vyjádřit vzorcem, kde k je číslo poslední iterace:

$$Err(\theta_i) = \frac{1}{\sqrt{\sum_{j=1}^n [P_{ij}]_k [Q_{ij}]_k}} \quad (17)$$

Shrnutí výhod počítačového adaptivního testování

Za hlavní výhodu CAT je pokládána efektivnost, a to zejména ve smyslu nutnosti administrovat menší počet položek k dosažení stanovené míry přesnosti (Wainer et al., 2000). Mimo tuto hlavní výhodu jsou v literatuře uváděny i další klady počítačového adaptivního přístupu:

- zlepšená přesnost pro jedince s extrémními úrovněmi latentního rysu;
- vzhledem k tomu, že položková banka CAT testu v ideálním případě obsahuje poměrně velký počet položek, z nichž pouze malá část se stane součástí konkrétního testu, je tím zajištěna možnost opakované administrace testu bez nežádoucího efektu učení. Je tím také zamezeno kompromitaci testu tím, že se položky postupně stanou všeobecně známými, což je typické pro často používané a oblíbené;
- z povahy adaptivního principu vyplývá, že odpověď neodrážející úroveň probandova rysu – ať již z důvodu chybného úkonu, překlepu nebo hádání – nemá v konečném důsledku žádný vliv na finální odhad úrovně rysu probanda. Naopak v klasickém testu se taková odpověď projeví, a to snížením nebo zvýšením celkového skóru o váhu položky;
- v průběhu celého testování zůstává respondent dostatečně motivován, protože jsou mu předkládány pouze položky s obtížností odpovídající úrovni jeho rysu. CAT testování se tak stává lepším zážitkem, přičemž optimálně vytížení probandi mohou při určitých schopnostních testech zažít fenomén flow a zvýšit svou individuální výkonnost (Csíkszentmihályi, 1996);
- parametry položek mohou být neustále zpřesňovány na základě kontinuálně získávaných dat, zatímco v tradičních testech jsou parametry položek pevně dány;
- praktickou výhodou, kterou ocení zejména vydavatelé testů, je ta skutečnost, že test může být velmi snadno revidován. V okamžiku, kdy jakákoli položka z položkové banky přestává být funkční, je možné ji bez narušení integrity testu odstranit.

LITERATÚRA

- Albers, H. G. – Bolton, S.: *Testen und Prüfen in der Grundstufe*. München : Goethe-Institut. 1995, 200 s. ISBN 3-468-49673-7.
- Anderson, L. W. – Krathwohl, D. R.: *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York : Longman, 2001, 305 s. ISBN 80-200-0524-2.
- Bartram, D. – Hambleton, R. K. (Eds.): *Computer-based testing and the Internet: Issues and advances*. Chichester : John Wiley & Sons Ltd, 2006. 262 p. ISBN: 978-0-470-86192-9.
- Birnbaum, A.: *Some latent trait models and their use in inferring an examinee's ability*. In: Lord, F. M. – Novick, M. R. (Eds.): *Statistical theories of mental test scores*. Reading : MIT Press, 1968, p. 397 – 479.
- Bond, G. B. – Fox, Ch. M.: *Applying the Rasch model: fundamental measurement in the human sciences*. Mahwah : Lawrence Erlbaum Associates, Inc., 2001, 280 p. ISBN: 978-0-805-83476-5.
- Brown, G. – Bull, J. – Pendlebury, M.: *Assessing Student Learning in Higher Education*. London: Routledge, 1997, 316 s. ISBN 0-415-16226-2.
- Brožík, V.: *Hodnoty, normy a projekty sveta*. Bratislava : Nakladateľstvo Epocha, 1969, 155 s.
- Burjan, V.: *Evalvácia a hodnotenie výsledkov vyučovania matematiky*. Bratislava : Metodické centrum v Bratislave, 1993, 52 s.
- *Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*. Cambridge : Cambridge University Press, 2007, 260 s. ISBN 0521803136 (Hardback), 0521005310 (Paperback) http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (25. 10. 2008)
- Csikszentmihalyi, M.: *O štěstí a smyslu života. Můžeme ovládat své prožitky a ovlivňovat jejich kvalitu?* Praha : Nakladatelství Lidové noviny, 1996, 400 s.
- Čečetka, J.: *Příručný pedagogický lexikon*. Martin : Kompas, 1943, 447 s.
- Ebel, R. L.: *Essential of Education Measurement*. Englewood Cliffs : Prentice Hall, 1972, 580 s.
- Ebel, R. L.: *Content-Standard Test Scores*. In: *Educational and Psychological Measurement*. R. 13, 1962, č. 1, s. 15 – 20.
- Gondová, D.: *Taking first steps in teaching English: assessing learners*. Žilina : University of Žilina, 2010, 142 s. ISBN 978-80-554-0236-9.
- Harlen, W.: *The Role of Assessment in the Implementation of Science into Primary School*. In: *Science and Primary School*. Amsterdam : University of Amsterdam, 2005, s. 27 – 48.
- Hniličková, J. a kol.: *Didaktické testy a jejich statistické zpracování*. Praha : Státní pedagogické nakladatelství, 1972, 200 s.
- Hughes, A.: *Testing for Language Teachers*. Cambridge University Press. 1989, 172 s. ISBN 0-521-27260.
- Chráška, M.: *Didaktické testy. Příručka pro učitele a studenty*. Brno : Paido – edice pedagogické literatury, 1999, 95 s. ISBN 80-85931-68-0.

- Jelínek, M.: *Teorie odpovědi na položku a počítačové adaptivní testování*. (Disertační práce.) Brno : Masarykova Univerzita, 2007.
- Jelínek, M. – Květoň, P. – Denglerová, D.: *Adaptivní testování – základní pojmy a principy*. In: Československá psychologie, 2006, roč. 50, č. 2, s. 163 – 173. ISSN 0009-062X.
- Kardanova, Ju. E.: *Modelirovanie i parametrizacija testov: osnovy teorii i prilozhenia*. Moskva : Federalnyj centr testirovania, 2008, 296 s.
- Kolen M. J. – Brennan R. L.: *Test Equating: Methods and Practices*. New York : Springer, 1995, 333 p. ISBN: 978-0-387-94486-9.
- Kujal, P.: *Aplikace teorie odpovědi na položku: Odlišné fungování položek Eysenckova osobnostního dotazníku podle pohlaví*. (Diplomová práce.) Brno : Masarykova univerzita, 2008.
- Kulič, V.: *Chyba a učení*. Praha : Státní pedagogické nakladatelství, 1971, 244 s.
- Lapitka, M.: *Tvorba a použitie didaktických testov*. Bratislava : Štátny pedagogický ústav, 1996 (2. vyd.), 134 s. ISBN 80-85756-28-5.
- Lapitka, M.: *Tvorba a použitie didaktických testov*. Bratislava: SPN - ÚÚVU, 1990, 135 s. ISBN 80-08-00782-6.
- Lord, F. M. – Novick, M. R.: *Statistical theories of mental test scores*. Reading : Addison-Wesley, 1968.
- Lord, F. M.: *Some test theory for tailored testing*. In: W. H. Holtzman (Ed.): *Computer assisted instruction, testing, and guidance*. New York : Harper and Row, 1970, p. 139 – 183.
- Malinovská-Šalamonová, Z.: *Francúzska literatúra 40. a 50. rokov*. In: Svetová literatúra po roku 1945. Prešov : Krajský pedagogický ústav, 1999, 184 s.
- Maňák, J.: *Náčrt didaktiky pre ďalšie vzdelávanie pedagogických pracovníkov*. Bratislava : Slovenské pedagogické nakladateľstvo, 1989, 133 s.
- Niemierko, B.: *Rozwój teorii pomiaru dydaktycznego na świecie w latach 1960 – 1980*. In: Kwartalnik Pedagogiczny. R. 26, 1981, č. 1, s. 19 – 32.
- Onderčová, V.: *Ako si vytvoriť didaktický test? Prešov : MPC, 2003, 10 s.*
www.rocepo.sk/downloads/RocMetListy/RocML001.doc (21. 4. 2013)
- Papica, J.: *Psychometrické metody ke konstrukci testu a sestavování testových systémů*. Olomouc : Univerzita Palackého, 1987, 117 s.
- Popham, W. J.: *Criterion-Referenced Measurement*. Englewood Cliffs : Prentice Hall, 1978, 212 s.
- Rafajlovičová, R. – Štulrajterová, M.: *Skúšanie, testovanie a hodnotenie v edukačnom procese*. Bratislava : ŠPÚ, 2002, 53 s. ISBN 80-85756-67-6.
- Rasch, G.: *Probabilistic models for some intelligence and attainment tests*. Copenhagen : Danish Institute for Educational Research, 1960.
- Rosa, V.: *Metodika tvorby didaktických testov*. Bratislava : Štátny pedagogický ústav, 2007, 72 s. ISBN 978-80-89225-32-3.
- Rymešová, J.: *Sledování evaluačních systémů v různých zemích*. In: Metodológia hodnotenia výsledkov vzdelávania. Zborník príspevkov zo seminára v Budmericiach, 6. – 8. mája 1996. Bratislava : Metodické centrum v Bratislave, 1996, 60 s. ISBN 80-8052-008-9.
- Smith R. M.: *Applications of Rasch Measurement*. Chicago : Mesa Press, 1992.
- Švec, Š.: *Učiť sa byť a uspieť*. In: Pedagogická revue. R. 54, 2002, č. 1, s. 9 – 32. ISSN 1335-1982.
- Turek, I.: *Kapitoly z didaktiky. Didaktické testy*. Bratislava : Metodicko-pedagogické centrum, 1995, 102 s. ISBN 80-85185-96-2

- Valček, P.: *Slovník literárnej teórie. A – J*. Bratislava : Vydavateľstvo SSS, 2000, 280 s. ISBN 80-8061-121-1.
- Váross, M: *Úvod do axiológie*. Bratislava : Nakladateľstvo Epoque, 1970, 356 s.
- Vygotskij, L. S.: *Psychologie myšlení a řeči*. Výbor z díla. Praha : Portál, 2004, 136 s. ISBN: 80-7178-943-7.
- Wainer, H. – Kiely, G.: *Item clusters and computerized adaptive testing: A case for testlets*. Journal of Educational Measurement, 1987, Vol. 24, No. 3, 185 – 202.
- Wainer, H. – Mislevy, R. J.: *Item response theory, item calibration, and proficiency estimation*. In: Wainer, H. et al.: *Computerized adaptive testing: A primer*. Mahwah : Lawrence Erlbaum Associates, Inc., 2000, 335 p. ISBN: 978-0-805-83511-3.
- Weir, C. J.: *Understanding and Developing Language Tests*. New York : Phoenix ELT, 1993, 341 s. ISBN 978-0-13-947532-0.
- Linden, Wim J. (et al.): *Elements of Adaptive Testing*. New York : Springer, 2010, 438 p. ISBN: 978-0-387-85459-5.
- Wright B. D. – Stone M. N.: *Best Test Design*. Rasch Measurement. Chicago : Mesa Press, 1979, 240 p. ISBN: 978-0-941-93800-6.
- Zhu, W.: *Test Equating: What, why, how? In: Research Quarterly for Exercise and Sport*. 1998, Vol. 69, No 1, p. 11 – 23.

Vydal: © Národný ústav certifikovaných meraní vzdelávania
Miesto vydania: Bratislava
Rok vydania: 2015
Vydanie: prvé
Počet strán: 150
Náklad: 1500 kusov
Tlač: Róbert Jurových – Nikara, Krupina

ISBN 978-80-89638-28-4

ISBN 978-80-89638-28-4

