

III. OVĚŘOVÁNÍ A OPTIMALIZACE DIDAKTICKÉHO TESTU¹

Relativně definitivní představu o vlastnostech testu můžeme získat až po důkladném vyzkoušení (ověření) testu na vzorku žáků. Toto ověřování se ovšem neprovádí pouze za účelem získání informací o kvalitě vytvořeného testu, nýbrž zejména proto, abychom mohli případné nevhodné vlastnosti testu odstranit, zmírnit nebo korigovat.

ANALÝZA VLASTNOSTÍ TESTOVÝCH ÚLOH

Při ověřování vlastností vytvořeného didaktického testu hledáme nejdříve odpověď na otázku, jaké vlastnosti mají jeho jednotlivé úlohy. Na kvalitě úloh je totiž závislá také kvalita testu jako celku. Analýza vlastností testových úloh se zaměřuje zejména na obtížnost úloh, na citlivost úloh a na tzv. nenormované odpovědi.

OBTÍŽNOST ÚLOHY

Obtížnost jednotlivých testových úloh můžeme posoudit podle toho, kolik žáků je dokáže správně vyřešit.

Při analýze obtížnosti se vypočítává buď **hodnota obtížnosti** Q a nebo **index obtížnosti** P . Hodnota obtížnosti udává procento žáků ve vzorku, kteří danou úlohu zodpověděli nesprávně anebo ji vynechali.

$$Q = 100 \frac{n_n}{n}$$

kde Q je hodnota obtížnosti, n_n je počet žáků ve skupině, kteří odpověděli nesprávně a nebo neodpověděli, a n je celkový počet žáků ve vzorku.

Index obtížnosti je procento žáků ve skupině, kteří danou úlohu zodpověděli správně.

$$P = 100 \frac{n_s}{n}$$

kde P je index obtížnosti, n_s počet žáků ve skupině, kteří odpověděli v dané úloze správně a n je celkový počet žáků ve skupině.

Lze snadno dokázat, že mezi hodnotou obtížnosti a indexem obtížnosti testové úlohy platí vztah

$$Q = 100 - P$$

¹ CHRÁSKA, M. *Didaktické testy*. Brno: Paido, 1999. s. 46-50.

O vysoké obtížnosti testové úlohy vypovídají vysoké hodnoty obtížnosti Q a naopak nízké hodnoty indexu obtížnosti P . Pro tyto významové nesrovnalosti se v posledních letech při popisu obtížnosti testových úloh dává přednost hodnotám obtížnosti Q .

Za velmi obtížné lze pokládat testové úlohy, u nichž hodnota obtížnosti Q je vyšší než 80. Velmi snadné jsou naopak ty úlohy, které vykazují hodnotu obtížnosti Q nižší než 20. Velmi obtížných (ale ani velmi snadných úloh) by nemělo být v testu příliš mnoho. Úlohy extrémně obtížné, u nichž se hodnota obtížnosti Q blíží ke 100, jsou nevyhovující a je nutno je z testu vyloučit. Úlohu extrémně snadnou, u níž se hodnota obtížnosti Q blíží k nule, je možno z psychologických důvodů doporučit jako úvodní úlohu v testu. Může totiž přispět k uklidnění žáků a k vytvoření potřebného pocitu jistoty. Zkušenosti ukazují, že nejvhodnější vlastnosti mají testové úlohy s hodnotou obtížnosti kolem $Q = 50$ (platí pro testy rozlišující).

CITLIVOST TESTOVÝCH ÚLOH

Citlivost úloh bývá často označována také jako rozlišovací hodnota, diskriminační hodnota, rozlišovací ostrost nebo jako rozlišovací schopnost úloh. Vysokou citlivost má taková úloha, kterou řeší s velkým úspěchem žáci, kteří mají celkově lepší vědomosti, zatímco žáci, kteří mají celkově horší vědomosti, v této úloze dosahují výsledků špatných. Citlivost úlohy tedy vyjadřuje, jak dalece daná úloha zvýhodňuje žáky, mající lepší vědomosti, před žáky, kteří mají vědomosti horší. K rozlišení žáků na žáky „s lepšími vědomostmi“ a na žáky „s horšími vědomostmi“ se většinou používá celkových výsledků ověřovaného didaktického testu.

Při posuzování citlivosti úloh se většinou nejdříve vzorek žáků rozdělí podle celkového počtu dosažených bodů (hrubého skóre) na dvě části: skupinu „lepší“ (s vyšším počtem dosažených bodů) a skupinu „horších“ (s nižším počtem dosažených bodů). Žáci se seřadí podle dosaženého celkového počtu bodů v testu, přičemž horní polovinu označíme jako „lepší“ (L) a spodní polovinu jako „horší“ (H). někdy je možné obě skupiny žáků vytvořit i z menšího počtu žáků, např. z 33 % nejlepších a 33 % nejhorších žáků apod.

Citlivost úlohy se dá exaktně posoudit pomocí výpočtu některého z koeficientů citlivosti, kterých byla navržena celá řada. Všechny tyto koeficienty mohou nabývat hodnot od - 1 přes nulu do + 1, přičemž platí, že čím vyšší hodnotu koeficient má, tím lépe úloha rozlišuje mezi žáky s lepšími vědomostmi a mezi žáky s horšími vědomostmi. Pokud koeficient citlivosti dosahuje hodnoty 0, znamená to, že úloha vůbec nerozlišuje mezi oběma skupinami žáků (žáci s lepšími i žáci s horšími vědomostmi jsou v této úloze stejně úspěšní). Záporné hodnoty koeficientu citlivosti vypovídají o tom, že úloha zvýhodňuje žáky, kteří mají v testu celkově horší výsledky. Kladné hodnoty koeficientu citlivosti naopak vypovídají o tom, že v úloze dosahují lepších výsledků žáci, kteří mají v testu lepší celkové výsledky.

METODY VÝPOČTU KOEFICIENTU CITLIVOSTI

a) Koeficient ULI

Nejjednodušším ukazatelem citlivosti testové úlohy je koeficient ULI (upper-lower-index). Jeho stanovení lze doporučit i v těch případech, kdy připravujeme test nestandardizovaný, protože jeho výpočet je velmi jednoduchý. Vychází totiž z rozdílu mezi obtížností úlohy ve skupině lepších a ve skupině horších žáků

$$d = \frac{n_L - n_H}{0,5 N}$$

kde d je koeficient citlivosti ULI, n_L je počet žáků z lepší skupiny, kteří danou úlohu zodpověděli správně, n_H je počet žáků ze skupiny horších, kteří úlohu řešili správně, a N je celkový počet žáků. Uvedený vztah platí pro případ, že obě skupiny byly vytvořeny na základě rozdělení všech žáků podle celkového dosaženého počtu bodů na polovinu.

U koeficientu ULI se požaduje, aby v případě úloh s hodnotou obtížnosti 30 – 70 bylo d alespoň 0,25 a u úloh s hodnotou obtížnosti 20 – 30 a 70 – 80 alespoň 0,15.

ANALÝZA NENORMOVANÝCH ODPOVĚDÍ

Vedle posuzování obtížnosti testových úloh a citlivosti testových úloh se v rámci analýzy vlastností úloh provádí také analýza tzv. **nenormovaných** odpovědí (Byčkovský, 1982), tj. rozbor odpovědí vynechaných a nesprávných.

Rozbor vynechaných odpovědí

Jestliže zjistíme, že některé odpovědi jsou vynechány, může to znamenat vedle neznalosti učiva také nepochopení formulace úlohy, nedostatek času k vypracování odpovědi atd. V literatuře se uvádí, že je třeba věnovat zvýšenou pozornost zejména těm otevřeným úlohám, ve kterých odpověď vynechalo více než 30 – 40 % žáků. U uzavřených úloh je však třeba věnovat zvýšenou pozornost i úlohám, kde neopovědělo více než 20 % žáků.

Rozbor nesprávných odpovědí

Rozbor nesprávných odpovědí je velmi jednoduchý u úloh s výběrem odpovědí.

U otevřených úloh je rozbor nesprávných odpovědí poněkud obtížnější. V těchto případech se doporučuje veškeré chyby žáků v určité testové úloze rozdělit do dvou kategorií, na tzv. **základní** a **vedlejší chyby**. Za základní chyby považujeme ty, které jsou způsobené skutečnou neznalostí učiva, jeho nepochopením nebo nezvládnutím. Vedlejší chyby jsou takové chyby, které jsou způsobené různými náhodnými vlivy, např. přehlédnutím, numerickou chybou ve výpočtu, nepřesností, špatnou čitelností textu atd. Jestliže v určité testové úloze převažují vedlejší chyby nad hlavními, může to znamenat, že v úloze úspěch řešení závisí více na jiných (náhodných) okolnostech než na stupni zvládnutí učiva. Takovou úlohu je třeba jako nevyhovující z didaktického testu vyloučit. V dobré testové úloze by počet hlavních chyb měl být vždy větší než počet chyb vedlejších.

ÚPRAVA VYTVOŘENÉHO DIDAKTICKÉHO TESTU

Z toho, co bylo uvedeno v předchozí kapitole, vyplývá, že *nevhodná testová úloha* se vyznačuje zejména následujícími vlastnostmi:

- Úloha je příliš obtížná anebo příliš snadná (hodnota obtížnosti Q je buď větší než 80 anebo menší než 20).
- Úloha málo rozlišuje mezi žáky s dobrými a špatnými vědomostmi (např. koeficient citlivosti d je u středně obtížných úloh menší než 0,25 apod.).

- V testové úloze je příliš mnoho vynechaných odpovědí (u otevřených úloh např. více než 30 – 40 %, u uzavřených úloh více než 20 %).
- Počet vedlejších chyb v úloze převažuje nad počtem hlavních chyb (u úloh otevřených).
- Žáci nevybírají ze všech nabídnutých distraktorů v úloze (u úloh s výběrem odpovědí).

Vytvoření definitivní podoby didaktického testu

Nevhodné (nebo „podezřelé“) úlohy je lépe z testu vyřadit a nahradit úlohami vhodnějšími. Proto je vhodné navrhnout (zvláště pro důležité části učiva) úloh více. Jestliže se určitá úloha jeví jako problematická a zkouší přitom důležitou část učiva, můžeme se pokusit o její úpravu (korekci). Někdy pomůže přesnější (srozumitelnější) formulace úlohy, jindy změna některého z distraktorů (u úloh s výběrem odpovědí) apod.

Pokud se v didaktickém testu užívá úloh více typů, doporučuje se úlohy stejného druhu soustředit do jedné části testu. Úlohy v takto vzniklých částech testu se řadí podle vzrůstající obtížnosti.

Jestliže chceme didakticky test používat jako seriózní prostředek měření, je třeba zabezpečit podmínky pro samostatnou práci žáků. Tento požadavek je zvláště naléhavý u testů, ve kterých jsou používány úlohy s výběrem odpovědí, protože kód správných odpovědí se velmi snadno napovídá. Nesamostatné práci žáků lze zabránit jednak důsledným dozorem, ale především tím, že vytvoříme dvě nebo více **ekvivalentních forem** testu. Ekvivalentní formu testu lze vytvořit několika způsoby.

První způsob spočívá ve změně pořadí úloh v testu. Při přeskupování úloh dbáme, aby i v tomto případě byla dodržena zásada, že úlohy v testu mají mít celkovou tendenci vzrůstající obtížnosti.

Jinou možností jak vytvořit ekvivalentní formu testu, je změna pořadí nabídek odpovědí (u úloh s výběrem odpovědí). V tomto případě je opisování sice znesnadněno, ovšem vyhodnocování odpovědí v testu je poněkud komplikovanější.

POUŽÍVÁNÍ DIDAKTICKÝCH TESTŮ VE ŠKOLNÍ PRAXI (STR. 72-78)

Z výsledků didaktického testu by měl učitel získat co nejvíce informací k hodnocení žáků, ale také (a to zejména) k optimalizaci svého dalšího pedagogického působení.

DIAGNOSTICKÝ ROZBOR VÝSLEDKŮ TESTU

Diagnostický rozbor výsledků by měl následovat prakticky po každém použití didaktického testu. Při tomto rozboru si učitel všimá především chyb, kterých se žáci dopustili, a hledá jejich pravděpodobné příčiny.

Příklad: Diagnostický rozbor výsledků u úloh s výběrem odpovědí

Test obsahoval celkem 18 uzavřených úloh s výběrem odpovědí a byl skórován tak, že za každou správnou odpověď získávali žáci vždy 1 bod. Jeden z vhodných způsobů analýzy výsledků uvádí následující tabulka.

V tabulce jsou žáci seřazeni podle dosaženého počtu bodů. V záhlaví tabulky jsou pod číslem úlohy uvedeny vždy příslušné správné odpovědi. U jednotlivých žáků jsou správné odpovědi vyznačeny čárkováním, nesprávné odpovědi písmenem, pod kterým žák tuto odpověď zvolil. Vynechané odpovědi jsou vyznačeny pomlčkou.

Pod tabulkou jsou uvedeny počty žáků, kteří v dané úloze odpověděli správně, a tyto počty jsou vyjádřeny také v procentech. Tabulka poskytuje v horizontálním směru informace o jednotlivých žácích a ve vertikálním směru informace o třídě. Lze snadno vyčíst, ve kterých úlohách jednotliví žáci i celá třída dosáhli dobrých výsledků a ve kterých se objevují největší nedostatky.

Nejvíce informací učitel získává rozbořením výsledků těch úloh, ve kterých žáci nejvíce chybovali. V uvedeném případě byl proveden důkladný rozbor zejména u těch úloh, které řešilo správně méně než 50 % žáků. Největší nedostatky byly zjištěny v úloze č. 14, v úloze č. 17 a v úloze č. 13.

Tabulka: Výsledky žáků v testu Mechanické vlastnosti kapalin a plynů

Jméno žáka	Počet bodů	1 D	2 C	3 B	4 A	5 D	6 B	7 C	8 A	9 C	10 D	11 B	12 B	13 D	14 B	15 A	16 D	17 B	18 D
1.....	18	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
2.....	17	/	/	/	/	/	/	/	/	/	/	/	/	C	/	/	/	/	/
3.....	15	/	/	/	/	/	/	/	/	/	/	/	/	A	C	/	B	/	/
4.....	15	/	/	/	/	/	/	/	/	/	/	/	/	-	C	/	/	A	/
5.....	14	/	A	C	/	/	/	/	/	/	/	/	/	A	C	/	/	/	/
6.....	14	/	/	/	/	C	/	/	/	B	/	/	/	C	/	/	/	/	B
7.....	13	/	/	/	/	/	/	/	/	/	B	/	/	C	/	B	/	A	B
8.....	13	/	/	/	/	/	/	/	/	A	/	/	/	C	C	/	/	D	B
9.....	13	/	/	/	/	/	/	/	B	/	/	/	/	A	C	/	B	D	/
10.....	12	B	/	/	/	/	/	/	/	A	/	C	/	/	-	/	/	D	B
11.....	11	A	/	/	/	/	/	/	/	/	A	/	C	/	C	/	B	B	A
12.....	10	/	/	/	/	/	/	B	B	/	C	C	/	A	C	C	/	D	/
13.....	9	/	/	/	/	/	/	-	B	D	-	A	/	/	-	B	B	/	A
14.....	9	/	/	/	/	/	D	-	/	/	-	/	/	A	A	-	-	-	A
15.....	9	/	/	/	/	C	/	/	/	D	/	D	/	B	C	B	B	A	B
16.....	9	/	/	/	/	/	D	/	/	B	/	D	C	A	C	B	-	D	/
17.....	8	/	A	/	/	/	/	D	/	A	B	C	/	C	/	B	B	A	B
18.....	8	/	/	/	/	/	C	/	/	B	C	A	A	B	C	/	A	D	C
19.....	8	/	/	/	B	/	/	/	/	A	C	D	A	B	D	B	B	D	/
20.....	7	/	/	/	D	/	C	B	B	/	C	A	/	B	A	/	A	A	B
21.....	7	/	/	/	/	/	D	/	B	A	C	A	C	/	A	B	B	D	A
22.....	7	/	/	/	/	C	A	/	C	/	C	A	C	B	C	D	B	A	/
23.....	6	/	/	C	/	/	/	B	C	/	A	A	C	A	C	B	A	C	A
24.....	6	/	/	A	D	/	/	D	D	D	B	A	/	/	C	B	A	C	A
25.....	5	A	/	A	C	/	C	/	C	A	B	A	/	A	D	C	B	A	/
26.....	5	/	A	C	/	C	C	/	D	D	C	C	/	A	D	B	C	-	/
Σ	268	23	23	21	22	22	18	19	16	13	12	10	20	6	5	11	9	6	12
%	100	88	88	81	85	85	69	73	62	50	46	38	77	23	19	42	35	23	46

POSOUZENÍ CELKOVÝCH VÝSLEDKŮ TŘÍDY

Dosažené výsledky třídy, případně školy, se obvykle posuzují podle průměrného počtu dosažených bodů.

Dosažené výsledky testování je výhodné znázornit graficky, protože z grafického znázornění je možno získat informace také o rozložení výsledků ve třídě. Nejčastěji se k tomuto účelu používá tzv. histogram četností. Histogram četností je v podstatě sloupcový diagram, u něhož se na vodorovnou osu nanáší dosažené výsledky testování (počty bodů) a na svislou osu počty žáků (četnosti).

KLASIFIKACE VÝSLEDKŮ TESTU

a) Intuitivní přístup ke klasifikaci

Někteří učitelé přistupují k převodu bodových výsledků na klasifikační stupně zcela subjektivně a sami víceméně intuitivně určují, kolik bodů je potřeba na dosažení určité známky. Pokud se jedná o učitele s velkou pedagogickou a odbornou zkušeností, většinou je toto hodnocení vcelku odpovídající.

Někteří odborníci doporučují jako optimální řešení otázky převodu bodového hodnocení na klasifikační stupně normativní přiřazování klasifikačních stupňů, které se provádí na základě posudku skupiny odborníků. Tato technika je ve školní praxi dobře použitelná. Vytvořený didaktický test necháme posoudit co možná největšímu počtu učitelů – odborníků

v daném předmětu – s tím, že je požádáme o návrh klasifikační stupnice. Z jednotlivých posudků možno určit průměr, čímž se eliminují extrémní názory.

b) Klasifikace na základě procenta správných odpovědí

Někdy se při převodu bodových výsledků na klasifikační stupně vychází z procenta správných odpovědí, kterého žák v testu dosáhl. Jeden z návrhů (Sedláčková, 1993), jak převádět dosažené počty bodů na klasifikační stupně touto metodou, uvádí tabulka.

Tabulka: Klasifikace podle procenta správně vyřešených úloh

Procento správně vyřešených úloh v testu - klasifikace běžná	Klasifikační stupeň
91 – 100	1
81 – 90	2
71 – 80	3
61 – 70	4
0 - 60	5

Učitelé touto metodou často postupují např. při hodnocení písemných prací. Zadávají např. celkem 4 rozsáhlejší úlohy a hodnotí takto: všechny vyřešené úlohy – výborně, tři vyřešené úlohy – chvalitebně, dvě vyřešené úlohy – dobře, jedna vyřešená úloha – dostatečně, žádná vyřešená úloha – nedostatečně. Užívání této metody převodu bodových výsledků na klasifikační stupně je v některých případech **diskutabilní**, protože nepřihlíží ke skutečnosti, že úlohy v testu mohou mít rozdílnou obtížnost. Odpoví-li např. žák 50 % úloh správně v testu s velmi obtížnými úlohami, může to znamenat dobrý výkon, zatímco stejný výsledek v testu velmi snadném může být výkonem velmi slabým. Tento způsob klasifikace může vcelku vyhovovat u testů, které neobsahují extrémně snadné ani extrémně obtížné úlohy.

Procentuální rozvržení musí stanovit pedagog citlivě na základě obtížnosti testu.

Hodnocení výkonu žáka podle procenta správně vyřešených úloh se důsledně používá u testů ověřujících (kriteriálních). Tyto testy jsou konstruovány tak, aby zkoušely jen nezbytné učivo, a výkon žáků se hodnotí většinou pouze dichotomicky, tj. *vyhověl* – *nevyhověl*. Aby žák v testu vyhověl, musí zodpovědět správně zpravidla 80 – 90 % všech úloh.

c) Klasifikace na základě normálního rozdělení

Bodové výsledky žáků v testu můžeme klasifikovat také na základě normálního rozdělení četností. U tohoto postupu se vychází z předpokladu, že výkonu dosažené v testu u dostatečně velké skupiny žáků odpovídají tzv. Gaussově křivce. To znamená, že nejvíce je vždy výkonů průměrných a na obě strany od průměru potom četností symetricky ubývá. Jestliže klasifikujeme ve shodě s touto zákonitostí, potom nejvíce žáků klasifikujeme stupněm 3, poněkud méně žáků stupněm 2 a 4 a nejméně žáků stupněm 1 a 5. Mezi jednotlivými autory není shoda v tom, kolik procent žáků by mělo být jednotlivými stupni klasifikováno. Všechny návrhy se však shodují v tom, že rozdělení jsou symetrická kolem stupně 3 a tento klasifikační stupeň má také největší četnost. Tabulka uvádí některá nejčastěji doporučovaná rozdělení.

Tabulka: Doporučovaná rozdělení klasifikačních stupňů

Klasifikační stupeň	Rozdělení (%)		
	a	b	c
Výborný (1)	7	10	15
chvalitebný (2)	24	20	20
dobry (3)	38	40	30
dostatečný (4)	24	20	20
nedostatečný (5)	7	10	15

Jestliže se rozhodneme klasifikovat žáky ve třídě podle normálního rozdělení, znamená to, že určitému procentu nejlepších žáků ve třídě přidělíme stupeň 1, určitému procentu žáků stupeň 2, určitému procentu stupeň 3, stupeň 4 a stupeň 5.