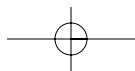
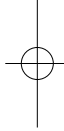
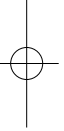
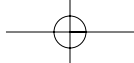


Rukověť autora testových úloh

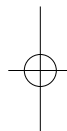
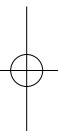
Radek Schindler a kol.





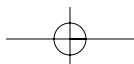
Na vzniku tohoto materiálu se podíleli: František Brož, Martina Hulešová, Martin Chvál, Štěpán Kovařík, Jan Matoušek, Marek Mičienka a Irena Smetáčková. Všem zmiňovaným spolupracovníkům děkujeme za jejich připomínky a cenná doporučení.

V textu jsou použity vybrané pasáže z publikací uvedených v seznamu použité literatury. Hlavním výchozím zdrojem byla Příručka pro tvorbu testových úloh, autorů Marka Mičienky a Ireny Moravcové-Smetáčkové, kterou vydalo Centrum pro reformu maturitní zkoušky v r. 2002.



Centrum pro zjišťování výsledků vzdělávání
Praha 2006

ISBN 80-239-7111-5



Obsah

	ÚVOD	...5
1.	TEST JAKO NÁSTROJ	...6
1.1	Vlastnosti testu	...6
2.	DIDAKTICKÉ TESTY	...7
2.1	Definice didaktického testu	...7
2.2	Didaktický test a ostatní formy zkoušek	...8
2.3	Vlastnosti didaktického testu	...10
2.4	Druhy didaktických testů	...19
2.5	Test rozlišující a test ověřující	...22
3.	ZAČÍNÁME TVOŘIT DIDAKTICKÝ TEST	...24
3.1	Od čeho začít aneb plánování testu	...24
3.2	Konstrukce testu	...30
3.3	Ověřování testu	...32
4.	ZAČÍNÁME S TESTOVÝMI ÚLOHAMI	...33
4.1	Jakou kvalifikaci by měl mít autor testových úloh?	...33
4.2	Testová úloha a její části	...34
5.	DĚLENÍ ÚLOH NA UZAVŘENÉ A OTEVŘENÉ	...36
5.1	Uzavřené úlohy	...36
5.2	Otevřené úlohy	...38
6.	TVORBA UZAVŘENÝCH TESTOVÝCH ÚLOH	...39
6.1	Úlohy s dvoučlennou volbou (dichotomické)	...39
6.2	Úlohy s výběrem z více odpovědí	...42
6.3	Úlohy přiřazovací	...48
6.4	Úlohy uspořádací	...49

7.	TVORBA OTEVŘENÝCH TESTOVÝCH ÚLOH	...52
7.1	Úlohy se stručnou odpovědí	...52
7.2	Úlohy se širokou odpovědí	...55
8.	OBEZNÁ DOPORUČENÍ PRO TVORBU TESTOVÝCH ÚLOH	...58
9.	HODNOCENÍ ODPOVĚDÍ	...62
10.	OVĚŘOVÁNÍ KVALITY TESTOVÝCH ÚLOH	...70
	ZÁVĚR	...81
	Použitá a doporučená literatura	...82
	Užitečné internetové odkazy	...83

Úvod

Objektivní měření výsledků vzdělávání již má v České republice určitou tradici, avšak počet odborníků, kteří se testováním profesionálně zabývají, je ve srovnání s vyspělým zahraničím poměrně nízký. Ze strany škol, jejich zřizovatelů, Ministerstva školství, mládeže a tělovýchovy (MŠMT) i veřejnosti jsou stále častěji slyšet hlasy volající po objektivních výsledcích vzdělávacího procesu. Lze konstatovat, že v kontextu probíhajících školských reforem a kurikulárních změn je problém pedagogického hodnocení poměrně živý. Hledají se způsoby, jak zlepšit stávající systém ukončování vzdělávání na jednotlivých stupních škol a přijímacího řízení na stupně následné. *Centrum pro zjišťování výsledků vzdělávání* je ve zmíněném procesu velmi angažováno a klade si za cíl vytvářet co nejkvalitnější didaktické testy jako nástroje pedagogického hodnocení.

Právě didaktickému testu, jak je používán v celostátních projektech Centra pro zjišťování výsledků vzdělávání, budou věnovány následující kapitoly. Cílem této příručky přitom není zevrubný popis tvorby a užití didaktického testu, ale soustředíme se zde hlavně na testové úlohy samotné, na to, co odlišuje dobré úlohy od špatných, zaměříme se na typická úskalí, která číhají na autory úloh, a uvedeme řadu doporučení pro tvorbu kvalitních testových úloh. Nevyhneme se přitom jistému množství teorie, ale pokusíme se ji doplnit názornými příklady.

Doufáme, že tato rukověť poslouží především autorům (nebo budoucím autorům) testových úloh – našim spolupracovníkům.

1. TEST JAKO NÁSTROJ

Ve světě, kde vládou informace, vzrůstá poptávka především po informacích ověřených, přesných, úplných a maximálně objektivních. Aniž si to uvědomujeme, testování pozvolna proniklo do našich životů a těší se velké pozornosti právě jako nástroj k získávání těch „správných“ informací. Některými testy získáme informace zcela zásadní, ba životně důležité povahy – např. u testu těhotenského nebo testů jaterních. Samotný fakt, že se něco nazývá testem, ještě není zárukou vysoké kvality získaných informací. Špatné testy mohou podávat informace liché, v horším případě dokonce zkreslené či zavádějící. Vždy záleží na účelu testování, na kvalitě testu, na způsobu jeho užití a na interpretaci zjištěných výsledků. Nedá se proto říci, že testování obecně je jednoznačně dobré nebo jednoznačně špatné.

1.1 VLASTNOSTI TESTU

Test představuje zkoušku, jejíž podmínky jsou pro všechny testované subjekty shodné a jejíž výsledky lze vyjádřit číselně. Test je objektivní tehdy, když se testování děje za dodržení předem stanovených podmínek a jsou limitovány subjektivní vlivy na straně testovaných i testujících. S oblibou se k němu využívají měřicí přístroje nebo hodnotící komise složené z nezávislých expertů daného oboru. Testovat lze kvalitu výrobků, přítomnost určitých chemických látek v prostředí, ale testovat lze samozřejmě i lidi a zvířata – je možné měřit a poměřovat jejich výkony, testovat jejich vlastnosti, dovednosti, znalosti apod. V humanitních vědách je tradiční doménou testování psychologie, která zkoumá lidské schopnosti, testuje určité znaky osobnosti nebo psychický výkon.

V pedagogice se nejčastěji mluví o **testu didaktickém** (*Achievement Test*), který slouží ke zjišťování úrovně zvládnutí učiva a užívá se k měření výsledků vzdělávání.

2. DIDAKTICKÉ TESTY

Vzdělávací proces není myslitelný bez zkoušek. Dokonce i samouk si musí čas od času „ověřit“, co již zvládl a co by si měl ještě zopakovat. Účelem zkoušek je průběžné nebo výstupní zjišťování osvojených znalostí a dovedností žáků v určité oblasti a za určité časové období. Didaktický test je jedním z mnoha nástrojů, které k tomuto účelu slouží.

Druhů didaktických testů a způsobů jejich užití je více. Jinak k didaktickému testu přistupuje učitel, který potřebuje uzavřít během dvou týdnů klasifikaci ve své třídě, jinak děkan fakulty, na niž má být přijato 300 z 2000 uchazečů a výsledky se musí zveřejnit do týdne, a ve zcela odlišné situaci je organizátor testování, které má ověřit úroveň znalostí a dovedností v matematice u všech maturantů z České republiky. Pro každou situaci je vhodný jiný typ didaktického testu.

V kontextu této brožury máme na mysli především didaktické testy, jaké používá Centrum pro zjišťování výsledků vzdělávání – tj. profesionálně vytvářené didaktické testy využitelné v celorepublikovém měřítku s vysokými nároky na kvalitu.

2.1 DEFINICE DIDAKTICKÉHO TESTU

Jedná se o nástroj určený k objektivnímu měření výsledků vzdělávání v předem vymezené konkrétní oblasti. Bývá tvořen souborem úloh, které se vztahují k vybraným částem obsahu vzdělávání a které jsou určeny k řešení během přesně vymezeného časového úseku. Součástí didaktického testu je dokumentace, která výslovně definuje, co test měří, jaké populaci žáků je určen, jak má probíhat jeho zadávání, jak mají být vyhodnocena žákovská řešení a jak lze interpretovat zjištěné výsledky.

2.2 DIDAKTICKÝ TEST A OSTATNÍ FORMY ZKOUŠEK

Zpětnou vazbu o úrovni zvládnutí probíraného učiva potřebuje žák i učitel. Formy zjišťování této úrovně mohou být různé v závislosti na tom, co je jejich hlavním cílem. Od ústního zkoušení, přes písemky, desetiminutovky, kompozice, praktické úkoly, projekty až po plošné (celostátní nebo i mezinárodní) didaktické testy.

POROVNÁNÍ ÚSTNÍHO ZKOUŠENÍ A DIDAKTICKÉHO TESTU

Každý z nás si jistě vzpomene na klasické (protože nejčastější) ústní zkoušení a mnohým se přitom vybaví i chvíle, kdy se u tabule cítili poškozeni nebo znevýhodněni. **Ústní zkoušení** má svůj význam při ověřování znalostí a dovedností žáka, a zejména při hodnocení žákova ústního projevu je nezastupitelné. Nemívá striktně daný časový limit (odpadá s tím spojený stres z nedostatku času) a umožňuje individuální přístup ke každému žákovi. V porovnání s písemným testem je však zatíženo řadou prvků, které mohou být zdrojem nerovných podmínek – různý obsah a různá obtížnost otázek, různá forma zadání, různá délka zkoušení, různé prostředí, nejasná kritéria hodnocení, žákův handicap v mluvené řeči, subjektivní vztah mezi učitelem a žákem, aktuální psychický stav učitele apod.

Didaktický test má obvykle písemnou formu a oproti ústní zkoušce sice neumožňuje individuální přístup ke každému žákovi,¹ nemůže ověřit některé žákovy kompetence v oblasti mluveného projevu, avšak má důležité přednosti, které z něj dělají spolehlivější nástroj plošného zjišťování výsledků vzdělávání. Jmenujme alespoň hlavní výhody didaktického testu:

- menší nebo žádný subjektivní vliv učitele na průběh řešení úloh a na hodnocení žákovských řešení;

¹ Individuální přístup k žákovi do jisté míry umožňuje tzv. computer-adaptive-testing (bližší viz kapitola 2.4).

- srovnatelné podmínky (úkol, čas, kritéria hodnocení) pro všechny žáky;
- snížená časová náročnost (za kratší čas lze vyzkoušet více žáků).

POROVNÁNÍ ŠKOLNÍ PÍSEMKY A DIDAKTICKÉHO TESTU

Co odlišuje běžnou školní písemku (prověrku, desetiminutovku apod.) od písemného didaktického testu jako nástroje plošného měření výsledků vzdělávání? **Písemka**, i když se v praxi také někdy nazývá testem a mnohé rysy didaktického testu bezesporu má, slouží víceméně k jednomu použití na malé skupině žáků (většinou jedna nebo dvě třídy). Písemka bývá menšího rozsahu, zpravidla má jednoho autora (učitele), nemá profesionální grafickou úpravu, nemá doprovodnou písemnou dokumentaci, pravidla hodnocení jsou jednoduchá, pokyny k řešení jsou žákům sdělovány ústně a možnost porovnání výsledků je omezena většinou jen na danou třídu.

Naproti tomu **didaktický test**, jak jsme jej vymezili výše, se rodí nepoměrně delší dobu (týdny až měsíce) a spolupracuje na něm tým odborníků. Bývá rozsáhlejší, rozmanitější co do typů použitých úloh, písemná dokumentace je bezpodmínečně nutná. K testu se pojí přesné písemné pokyny pro žáky, pro zadavatele a hodnotitele. Možnost srovnání výsledků je o mnoho širší, početný soubor žáků umožňuje matematicko-statistické zpracování výsledků a porovnání různých skupin žáků mezi sebou, porovnání škol podle typu, podle studijních oborů, regionů apod.

Didaktický test není alternativou ke zkouškám ústním, praktickým a jiným, ale spíše jejich vhodným doplňkem.²

² V současné době se vede odborná diskuse o tzv. osobním portfoliu či osobní složce žáka, která představuje nový nástroj pedagogického hodnocení. Jeho prostřednictvím se hodnocení stává komplexní a dlouhodobé, zahrnuje různé údaje o výsledcích (včetně výsledků v testech) a procesech žákova učení, které se shromažďují po celou dobu jeho školní docházky. Zavedení tohoto nového nástroje by současně podporovalo probíhající kurikulární reformu.

Bylo by chybou didaktické testy přeceňovat a činit z nich všemocné, neomylné a jediné správné nástroje pro měření výsledků vzdělávání, a na druhé straně je stejně nesprávné didaktické testy zcela zavrňovat.

Je třeba kriticky poznamenat, že didaktický test je jen nástroj (jak již bylo zmíněno výše) a jako takový může být použit správně nebo nesprávně, může být utvořen dobře i špatně, navíc i výsledky dobrého testu lze mylně interpretovat a činit z nich chybné závěry. V praxi ne každý didaktický test je vytvořen a použit optimálním způsobem a v takovém případě nejenže postrádá mnohé přednosti, ale jeho hodnota jako nástroje měření výsledků vzdělávání je pochybná. (Vzpomeňte na svá školní léta a na testy, které jste absolvovali. Pravděpodobně se vám vybaví i takové testy, které vás, mírně řečeno, rozčílily.) S určitou mírou nadsázky bychom mohli říci, že lepší žádný test, než test obsahově či konstrukčně špatný, protože výsledky, které jeho prostřednictvím získáme, mohou být nepřesné, neúplné, nesprávné... Některé testy mohou mít (ať už záměrně nebo bezděčně) diskriminační charakter – tj. předem zvýhodňují určitou skupinu žáků oproti jiné. Test je mocný nástroj a při jeho konstrukci i použití je třeba postupovat velmi obezřetně.

Lze shrnout, že didaktické testy mají v systému školního hodnocení své důležité a opodstatněné místo za předpokladu, že budou dobře utvořeny, náležitě použity a jejich výsledky budou správně a citlivě interpretovány.

2.3 VLASTNOSTI DIDAKTICKÉHO TESTU

Zmínili jsme významné přednosti didaktického testu a upozornili na možná nebezpečí související s jeho tvorbou a použitím. Zaměřme se nyní blíže na vlastnosti, které by kvalitní didaktický test měl splňovat. Jsou to především:

- objektivita
- validita
- spolehlivost (reliabilita)
- citlivost (diskriminace)

OBJEKTIVITA

Objektivita je klíčová, avšak vnitřně komplikovaná a jen zčásti dosažitelná charakteristika. V kontextu testování objektivitu chápeme zejména jako nepřítomnost výrazných subjektivních vlivů v průběhu testování. Ty přitom mohou mít původ jak v samotné tvorbě testu, tak v situaci řešení či následného vyhodnocování testu. Subjektivní vlivy na straně zadavatelů a hodnotitelů se snažíme minimalizovat tím, že stanovujeme přesná pravidla, kterými se řídí. Současně také snižujeme subjektivitu na straně žáků tím, že úlohy formulujeme naprosto jednoznačně, aby žáci neměli možnost rozumět úloze jiným, nepředpokládaným způsobem. V návaznosti na vyloučení subjektivních vlivů se objektivita spojuje se srovnatelností. Všichni testovaní žáci řeší shodný test a za shodných podmínek, čímž se zvyšuje srovnatelnost jejich výsledků.

Zkouška je tedy objektivní především tehdy, pokud výsledky nejsou ovlivněny subjektivními názory či postoji osoby, která je hodnotí. Taktéž samotný průběh zkoušky by neměl být ovlivněn prostředím konkrétní školy nebo třídy, v níž je prováděna. Výsledky některých druhů zkoušek mohou být významně ovlivněny působením celé řady vedlejších faktorů. Vzpomeňme na zmiňované vlastnosti ústní zkoušky.

Správně konstruovaný didaktický test je schopen poskytnout poměrně objektivní, a tedy srovnatelné výsledky, které závisí pouze na znalostech a dovednostech jednotlivých žáků. V testu jsou žákům předloženy stejné úlohy se stejným časovým limitem pro řešení a s předem určeným klíčem správných řešení. Žáci tak mají stejné výchozí podmínky, čímž je současně umožněno porovnávat výsledky jednotlivých žáků/tříd/škol mezi sebou. Objektivita a srovnatelnost nejsou samozřejmou vlastností každého písemného didaktického testu, ale pouze testu kvalitního a konstrukčně bezchybného. K zajištění objektivity přispívá **jednoznačná formulace úlohy testu, shodné podmínky při jeho zadávání a především přesná a pro všechny stejná pravidla hodnocení žakovských odpovědí.**

Dodejme, že nikdy nemůžeme dosáhnout objektivitu stoprocentní, jelikož testování je živá situace, v níž je zapojen velký počet lidí, a v souvislosti s tím se může objevit řada rušivých, nestandardních či „neobjektivních“ prvků (někomu dopsalo pero, někomu se chce na záchod, kolem školy jede hluchý kamion, učitel umožnil opisování apod.). I přesto lze u kvalitně sestaveného a správně použitého testu mluvit o vysoké míře objektivitu.

Objektivita může být snížena například nejednoznačně formulovanými úlohami, v nichž každý žák rozumí zadání jiným způsobem.

Příklad neobjektivní úlohy

Které zvíře žijící v ČR je největší?

KOMENTÁŘ: Úloha je nejednoznačná, jelikož umožňuje zahrnout i zvířata chovaná v zajetí (např. v ZOO) – i tato žijí v ČR. Další problém představuje slovo „největší“ (žák si může pod tímto pojmem představit zvíře nejhmotnější, nejvyšší, nejvyšší ve stoje na zadních apod.).

VALIDITA

Je test dostatečně validní? Měří to, co měřit má? Slouží k tomu účelu, za jakým byl konstruován? Tyto tři otázky se táží na jednu a tu samou věc – na validitu testu. Validita představuje shodu mezi výsledky testu a účelem, pro který byl test vytvářen. Nízká validita testu znemožňuje interpretaci výsledků, a tedy jejich další využití v pedagogické praxi. Zjednodušeně řečeno: výsledky testu nám mnoho neřeknou, pokud si nejsme jisti, co skutečně daný test měřil a co můžeme na základě jeho výsledků o žákovi zjistit.

Standardy pro pedagogické a psychologické testování (Praha, Testcentrum 2001) definují validitu jako míru, ve které empirické důkazy a teorie podporují interpretaci testových skóre při doporučeném způsobu užívání testu. Validita je proto nejdůležitějším hlediskem při tvorbě a hodnocení testu.

Provést validizaci znamená shromáždit výzkumné nálezy, a získat tak přesvědčivé vědecké důkazy pro navrhovaný způsob interpretace testových skóre získaných při jejich administraci. Navrhovanou interpretací je myšlen konstrukt nebo koncept, který má test měřit.

Validita testu je ovlivněna celou řadou faktorů jak v konstrukci testu, tak při jeho zadávání (možnost opisování apod.). Nízkou validitu mohou zapříčinit též nejasné pokyny pro řešení testu, příliš obtížný jazyk nebo složitá struktura testu.

Rozlišujeme několik druhů validity: **obsahovou, kriteriální a zjevnou.**³ Nás bude v první řadě zajímat **validita obsahová** – tj. míra, s jakou daný test opravdu měří tu konkrétní znalost nebo dovednost, kterou skutečně chtěli autoři testu měřit, resp. zda úlohy zvolené do testu dostatečně obsahově pokrývají uvažovanou oblast učiva a požadované úrovně znalostí a dovedností.

Obsahová validita je ohrožena nesprávnou konstrukcí úloh nebo chybným výběrem ověřovaného učiva. Konstrukční chyby vedoucí ke snížení validity spočívají nejčastěji v tom, že k vyřešení úlohy je nutné provést několik myšlenkových operací, z nichž některé nemusí být žákům známé. Autoři úloh mohou považovat určitou operaci za elementární či samozřejmou, a nepovažují ji proto za cíl úlohy. Pro žáky však právě tato operace může být překážkou k úspěšnému vyřešení úlohy. Příkladem by mohl být test z dějepisu, jehož součástí jsou úlohy s výchozími texty. Pokud by texty byly příliš dlouhé a jazykově komplikované, mohou vést k neúspěchu žáka, jehož historické znalosti jsou sice na vysoké úrovni, ale v textu se nedokázal správně zorientovat.

³ U různých autorů se můžete setkat s různým způsobem třídění validity na druhy. Uvádíme zde jen ty základní. V současných Standardech pro pedagogické a psychologické testování se mluví spíše o zdrojích důkazů, nebo lépe dokladů validity, které napomáhají správné interpretaci výsledků testů.

Příklad úlohy s nízkou validitou

(Úloha z matematiky)

Obec Kovač má 250 obyvatel. Každý rok se v obci narodí nebo se do ní přistěhují 3 obyvatelé. Každý rok 2 obyvatelé zemřou nebo se odstěhují. Vypočítejte, za kolik let se obec Kovač stane městem.

KOMENTÁŘ: Cílem úlohy je ověřit základní matematické operace. Při řešení úlohy žáci mají uvést postup výpočtu a závěr, který obsahuje počet let, za něž se daná obec stane městem. V zadání úlohy však není uvedena definice města. Předpokládá se, že žáci vědí, jaké podmínky musí být splněny, aby se obec změnila v město. Tato znalost však jistě nepatří do matematiky. Úspěšnost žáků při řešení úlohy tedy nezávisí pouze na znalostech a dovednostech matematických.

Zdrojem nízké validity může být také samotná formulace testové úlohy. V důsledku nejednoznačného vyjádření, použití cizích nebo neobvyklých slov či komplikované větné stavby mohou žáci pochopit úlohu odlišným způsobem, než jak ji autoři původně zamýšleli. V takovém případě má úloha pro žáky zcela jiný význam a jejich odpověď je reakcí na jiný podnět, než který v úloze spatřují autoři.

Výsledky též nejsou validní tehdy, pokud test ověřuje znalosti a dovednosti, které si žáci neměli možnost ve škole osvojit (buď proto, že je ještě neprobírali, nebo proto, že se jedná o znalosti a dovednosti získávané v mimoškolních aktivitách).

V jistém smyslu lze o snížené validitě mluvit i u takových úloh, které jsou obtížně řešitelné, nebo dokonce zcela neřešitelné pro žáky s určitým druhem postižení. Např. nevidomí žáci mohou jen velmi omezeně pracovat s informacemi v grafické podobě, neslyšící žáci jsou zase znevýhodněni v oblasti jazykové. Velmi zjednodušeně řečeno – dáte-li nevidomému test

napsaný v běžném černotiskovém písmu, změříte pouze to, že jej nemůže přečíst. Změříte žákův handicap, ale nikoli jeho znalosti a dovednosti.

Posouzení obsahové validity testu se v praxi ponechává na odborníkovi (nebo lépe skupině odborníků).

Z dalších druhů validity zmiňme **validitu kriteriální** (*Criterion-Related Validity*), kdy je dosažené skóre v testu porovnáváno s jinými kritérii. Kriteriální validitu lze odhadovat empiricky a numericky vyjádřit (např. korelačním koeficientem mezi naším testem a kritériem). Proto je kriteriální validita označována i jako **validita empirická**.

Kriteriální validitu lze dále dělit na **souběžnou** a **predikční**. V případě **souběžné** validity zjišťujeme vztah mezi naším testem a výsledkem v jiném testu (kritériem), o kterém je známo, že měří stejnou vlastnost. V případě, že není po ruce ověřený validní test, porovnávají se výsledky žáků dosažené v daném testu alespoň s jinými, všeobecně uznávanými údaji o žácích, které by měly s výsledky našeho testu souviset (např. školní klasifikací nebo výsledky v IQ testech).

Predikční validita vypovídá o tom, do jaké míry dokáže náš test předpovídat výsledky v budoucnu v oblastech, které nás primárně zajímají a kvůli kterým je test konstruován (např. souvislost mezi výsledkem testu studijních dovedností po 9. ročníku ZŠ a úspěšností v dalším studiu.)

Pro úplnost uvedme ještě **validitu zjevnou** (*Face Validity*), u níž se jedná o posouzení testu „na první pohled“, tj. nakolik je struktura a grafická podoba testu shodná u testů téže série (např. různé předměty maturitní zkoušky v daném roce) a nakolik je přiměřená schopnostem a dovednostem žáků.

SPOLEHLIVOST (Reliabilita)

Pokud je test, resp. celý proces testování, dostatečně reliabilní, znamená to, že se na jeho výsledky lze spolehnout. Reliabilita je komplikovaná veličina zahrnující v sobě **spolehlivost a přesnost**. Od každého měřícího nástroje (tedy i didaktického testu) chceme, aby byl spolehlivý a přesný. Pokud by například teploměr naměřil jednomu pacientovi v pětiminutových intervalech teploty 37,7; 35,9 a 39,2, asi bychom ho považovali za rozbitý a naměřeným hodnotám bychom nepřikládali žádnou váhu. Podobně i didaktický test poskytuje směrodatné informace o úrovni znalostí a dovedností žáků jedině tehdy, pokud dává přesné a spolehlivé informace – tj. je reliabilní.

Každý výsledek testování však v sobě přirozeně obsahuje jisté zkreslení. Toto zkreslení má původ v kvalitě testu, v porozumění úlohám, v podmínkách, za kterých je test zadáván, v momentální kondici žáka atd. Do hlavy žáka nevidíme a na dosaženou úroveň zvládnutí učiva usuzujeme z výsledku, kterého dosáhl v testu. Vzdálenost mezi skutečnými znalostmi a dovednostmi žáka a jeho dosaženými výsledky v testu může být různě velká. Jestliže je test dostatečně přesný, je tato vzdálenost malá a výsledek v testu je pak téměř totožný se žákovými skutečnými znalostmi a dovednostmi. Spolehlivé měření znamená, že test poskytuje stabilní, opakovatelné výsledky. V ideálním případě by měl stejný žák při opakovaném zadání testu dosáhnout shodného výsledku. Pokud test neměří spolehlivě, znamená to, že do výsledků se promítá řada vnějších, náhodných vlivů. Na základě výsledků málo reliabilních testů není možné činit významné závěry o žácích, protože nelze vyloučit příliš velký podíl náhody.

Vzorce pro výpočet reliability zde neuvádíme a případné zájemce odkazujeme na odbornou literaturu pojednávající o didaktických testech – např.: Chráska, M.: Didaktické testy. Brno, Paido 1999, Říčan, P.: Úvod do psychometrie. Bratislava, Psychodiagnostické a didaktické testy 1978, Hendl J.: Přehled statistických metod zpracování dat. Praha, Portál 2004 nebo Zvára, K., Štěpán, J.: Pravděpodobnost a matematická statistika. Praha, Matfyzpress 2002.

Běžnému pedagogovi pro praktické účely postačí, když dokáže údaje o reliabilitě testu správně interpretovat. Zmíňme alespoň některé důležité skutečnosti spojené s reliabilitou.

- K exaktnímu posouzení spolehlivosti testu slouží koeficienty reliability. Koeficient je reálné číslo nabývající hodnot od 0 pro případ naprosté nespolehlivosti a nepřesnosti (například kdybychom výsledky žáků losovali) až po hodnoty blízké 1 pro ideální, dokonale spolehlivý a přesný didaktický test, se kterým se však v praxi nesetkáme.
- Čím je koeficient reliability bližší hodnotě 1, tím je test přesnějším měřícím nástrojem a tím více se lze na jeho výsledky spolehnout. Pro individuální pedagogickou diagnostiku, kdy se podle výsledku v testu např. rozhoduje o přijetí nebo nepřijetí žáka na určitou školu, se požaduje koeficient reliability minimálně 0,8. Pro ostatní školskou praxi postačuje koeficient reliability pohybující se v blízkosti hodnot 0,6–0,7.
- Existuje více způsobů odhadu reliability. Tyto způsoby se mimo jiné liší tím, jakou chybu se příslušným způsobem daří odhalit. Jedním z odhadů je tzv. **vnitřní konzistence**. Zde předpokládáme, že každá úloha měří pouze tu znalost či dovednost jako celý test. Pro vnitřní konzistenci platí, že čím více úloh test obsahuje a čím jsou si úlohy podobnější, bližší (tj. ověřují co nejužší skupinu znalostí a dovedností), tím vyšší má reliabilitu.⁴ U testů s malým počtem úloh (10 a méně) zpravidla koeficient reliability dosahuje hodnoty okolo 0,6.
- Snaha o dosažení vnitřní konzistence jde však proti obsahové validitě testu, kdy se snažíme naopak pokrýt úlohami širší obsahovou oblast a předejeme tomu, že každá úloha měří něco trochu jiného, a na výsledek testu nahlížíme jako na kombinaci požadovaných znalostí a dovedností. V tako-

⁴ Testy zahrnující v sobě tzv. subtesty (např. poslechový subtest z cizího jazyka) mají obecně vnitřní konzistenci nižší. Naopak testy z předmětů, které jsou vnitřně obsahově homogenní a jejich poznatky jsou spolu úzce propojené, mají vnitřní konzistenci vyšší.

vém případě by bylo vhodnější volit jinou cestu odhadu reliability, např. testování pomocí paralelních forem testů. Blíže viz např. Řičan, P.: Úvod do psychometrie. Bratislava, Psychodiagnostické a didaktické testy 1978.

CITLIVOST (Diskriminace)

Citlivost vypovídá o schopnosti testu rozlišovat mezi žáky s různou úrovní skutečných znalostí a dovedností. Je-li test citlivý, měly by být výsledky žáků přiměřeně rozprostřeny po celé bodové škále. Pokud například všichni žáci dosáhnou v testu výborného výsledku nebo všichni špatného, test není citlivý – nerozlišil žáky mezi sebou. Optimální míra citlivosti se liší v závislosti na účelu testu. Test, jehož cílem je rozhodnout, kteří žáci mají být přijati na vysokou školu, musí být velmi citlivý. Naopak pro test, jehož prostřednictvím chceme ověřit, zda si každý jednotlivý žák osvojil určité učivo, není vysoká míra citlivosti nutnou podmínkou jeho úspěšného použití.

Příklad úloh s nízkou citlivostí

Úloha 1

Napište jméno prezidenta České republiky.

Úloha 2

Napište, jaké je národnostní složení obyvatel republiky Uganda.

KOMENTÁŘ: Obě úlohy mají velmi nízkou citlivost vůči maturitní populaci. Zatímco u první z nich můžeme předpokládat, že ji téměř všichni maturanti zodpoví správně, u druhé naopak jen málokdo bude znát správnou odpověď. Po těchto dvou úlohách na tom většina žáků bude stejně – tj. budou mít jednu odpověď správně a jednu špatně. V nejlepším případě odhalíme jedině (možná recesistu), který nebude znát jméno českého prezidenta, nebo žáka (pravděpodobně afrikanistu) se zcela mimořádnými znalostmi o Ugandě.

2.4 DRUHY DIDAKTICKÝCH TESTŮ

Způsobů, jakými lze zjišťovat školní znalosti a dovednosti, je celá řada a zkušený učitel je zná, používá, sám je též vytváří, kombinuje a jejich výsledky vyhodnocuje. Podobně i didaktických testů je mnoho druhů a liší se svým účelem, koncepcí, formou apod. Jinak bude vypadat test v autoškolce, jinak test u přijímacích zkoušek na vysokou školu. Neexistuje jeden univerzální (a nevhodnější) typ testu pro všechny příležitosti.

Následující tabulka ukazuje možné třídění testů podle řady kritérií, avšak toto třídění není jediné možné a není ani zcela vyčerpávající.

<i>Kritérium třídění</i>	<i>Testy</i>	<i>Popis</i>
<i>Forma zadání</i>	<i>Zadané na papíře</i>	<i>Test je žákům předložen v tištěné podobě. Úlohy mají textový, případně grafický charakter.</i>
	<i>Zadané ústně</i>	<i>Test zadává administrátor ústně nebo jej přehrává z audiozáznamu. Žák odpověď zapisuje nebo nahlas vyslovuje a zápis provádí administrátor.</i>
	<i>Zadané elektronicky</i>	<i>Test je zadáván prostřednictvím osobního počítače. Elektronické testování umožňuje tzv. computer-adaptive-testing, ve kterém není předem dáno přesné pořadí úloh v testu; počítač úlohy vybírá na základě předchozích odpovědí žáka.</i>
	<i>Speciální</i>	<i>Pro žáky se zdravotním postižením může být test zadáván prostřednictvím znakové řeči či bodového písma.</i>

Kritérium třídění	Testy	Popis
<i>Forma zadání</i>	<i>Kombinované</i>	<i>Test, z něhož žák získá jednotný výsledek, může být tvořen několika subtesty, z nichž každý je zadáván odlišnou formou. Příkladem je test z cizího jazyka, který obsahuje subtest na čtení s porozuměním (zadáván písemně) a subtest poslechový (zadáván ústně).</i>
<i>Povaha činnosti testovaného</i>	<i>Kognitivní</i>	<i>Test měří úroveň (kvalitu) znalostí a intelektových dovedností.</i>
	<i>Psychomotorické</i>	<i>Test ověřuje psychomotorické dovednosti, jako např. psaní na stroji.</i>
<i>Rozsah použití</i>	<i>Učitelské</i>	<i>Užívaný výlučně autorem testu (zpravidla učitelem).</i>
	<i>Širokého použití</i>	<i>Využívaný masově.</i>
<i>Tematický rozsah</i>	<i>Monotematické</i>	<i>Testují jedno téma učební látky.</i>
	<i>Polytematické</i>	<i>Testují učivo několika tematických celků.</i>
<i>Míra specifičnosti testu</i>	<i>Výsledků vzdělávání</i>	<i>Test měří to, co se žáci v dané oblasti naučili.</i>
	<i>Studijních předpokladů</i>	<i>Test měří úroveň obecnějších charakteristik - předpokladů, které jsou potřebné k dalšímu studiu.</i>
<i>Časové zařazení do výuky</i>	<i>Vstupní</i>	<i>Cílem je postihnout úroveň znalostí a dovedností na počátku výuky určitého celku.</i>
	<i>Průběžné</i>	<i>Cílem je poskytnout učitelům i žákům zpětnou vazbu k optimalizaci výuky. Učitel ověřuje, jak žáci učivo chápou a jak si je osvojují.</i>

Kritérium třídění	Testy	Popis
Časové zařazení do výuky	Výstupní	Slouží k hodnocení žáků na konci výukového období nebo po ukončení výuky určitého celku.
Měřená charakteristika výkonu	Rychlost	Test ověřuje, jak rychle žák dokáže řešit určené úkoly. Úkoly jsou obvykle méně komplexní a jejich počet výrazně překračuje možnosti dané časovým limitem.
	Úroveň	Test ověřuje, zda žák dokáže řešit specifické úkoly. Úkoly jsou náročnější a čas na jejich řešení je dostatečný.
Interpretace výsledků	Rozlišující	Cílem testu je vzájemně porovnat výsledky jednotlivých žáků. Na základě řešení testu jsou žáci seřazeni podle dosažené úspěšnosti. Zda je konkrétní žák hodnocen jako úspěšný nebo neúspěšný, závisí mimo jiné na výkonech ostatních žáků.
	Ověřující	Cílem testu je ověřit, zda si žák osvojil určité znalosti a dovednosti, které jsou předem stanoveny jako podstatné. Výsledek konkrétního žáka není porovnáván s výsledky jiných žáků, ale s předem stanovenými kritérii.

Didaktické testy v praxi vždy kombinují několik výše zmíněných charakteristik.

Zaměříme se nyní blíže na rozdělení testů podle toho, jaký druh informací jejich výsledky poskytují, jinými slovy – jakým způsobem lze testové výsledky interpretovat (viz poslední řádek tabulky).

2.5 TEST ROZLIŠUJÍCÍ A TEST OVĚŘUJÍCÍ

Oba hlavní druhy didaktických testů – test rozlišující a test ověřující – se od sebe v mnoha aspektech liší a **tvůrci testu musí mít od samého začátku jasno v tom, který z testů budou vytvářet**. Druh testu ovlivňuje jeho celkovou koncepci i charakteristiky úloh. Především je však každý z nich vhodný pro jiný typ zkoušky. Pokud tvůrci testu předem dostatečně neobjasní, o který z uvedených druhů se jedná, jaký má daný test cíl a jaký typ informací o žákovi poskytne, může se stát, že učitelé, žáci nebo jejich rodiče budou z výsledků rozpačití a budou mít dojem, že jim test nepřinesl to, co od něj očekávali.

Toto ostré rozlišení chápeme jako modelové, v praxi se testy spíše blíží jednomu či druhému typu.

TESTY ROZLIŠUJÍCÍ

(testy relativního výkonu, norm-referenced measurement, NR-testy)

Základním účelem rozlišujícího testu je **vzájemné porovnání žáků**.

- Výkon žáka je srovnáván s výkony ostatních žáků, resp. s průměrným výsledkem dané skupiny.
- Výsledky nám umožní stanovit pořadí žáků podle dosaženého skóre.
- Test musí být vysoce citlivý, aby dobře rozlišil jednotlivé žáky, proto se v rozlišujícím testu používají obtížnější úlohy s vysokou citlivostí.
- Úlohy zpravidla nepokrývají důsledně celou vymezenou oblast učiva.
- Test bývá obtížnější, na jednotlivou úlohu odpoví správně v průměru asi 50 % žáků.

- Výsledek testování může být uveden ve formě pořadí žáků od nejlepšího k nejhoršímu. Jinou formou je například vyjádření výsledků jednotlivých žáků v percentilech (v procentech vyjádřený podíl žáků, kteří dosáhli v testu stejného nebo horšího výsledku).
- Příkladem využití tohoto druhu testů jsou předmětové olympiády, kdy se stanovuje pořadí účastníků podle úspěšnosti, nebo přijímací zkoušky, u nichž je předem stanoven počet přijímaných uchazečů.

TESTY OVĚŘUJÍCÍ

(testy absolutního výkonu, criterion-referenced measurement, CR-testy)

Ověřující test **poměřuje výkon žáka s předem danými kritérii**. Výsledky pak vyjadřují míru splnění vytčených požadavků, informují nás o úrovni dosažených znalostí a dovedností žáka, ale neříkají nic o pořadí žáků.

- Výsledky žáka neporovnáváme s ostatními žáky ve skupině, ale pouze s dokonalým zvládnutím učiva (tzv. „ideálním žákem“).
- Pravidla hodnocení a klasifikace jsou stanovena předem. Součástí těchto pravidel je tzv. *minimální výkon v testu* nebo *hraniční skóre (cut-off score)* – minimální počet bodů, které žák musí získat, aby v testu uspěl.
- Ověřující test musí být validní vůči ověřovanému učivu.
- V testu se dbá na zastoupení všech předem vymezených požadavků, tematických celků a učebních cílů.
- Optimální test ověřující bývá spíše snazší, na jednotlivou úlohu odpoví správně v průměru cca 80 % žáků.
- Typickým příkladem použití ověřujícího testu je test v autošколе.
- Často se používá ověřujících testů v jednostupňové podobě (*pass-fail scaling, mastery tests*). Úlohy vztahené k jednomu okruhu požadavků jsou považovány za rovnocenné a kritériem je pouze „množství“ – čili dosažení

určitého procenta úspěšných odpovědí, aby mohl být daný okruh požadavků uznán za splněný.

- Jiný typ ověřujícího testu představuje test vícestupňový, který umožňuje odstupňovat požadavky podle různých úrovní odpovídajících příslušnému vzdělávacímu programu, přičemž vrstvy jsou uspořádány hierarchicky a každá nižší vrstva je částí požadavků vrstvy vyšší. Může se jednat o model dvouúrovňový (úroveň základní a rozšířená) nebo i tří či pětiúrovňový model (úroveň minimální, základní, rozšířená atd.).

V běžné školní praxi se často používají testy kombinované – nejsou ryze ověřující nebo rozlišující, ale mají charakteristiky obou druhů. Představte si situaci, kdy by učitel předem vyhlásil, že jedničku z písemky dostane jen 5 nejlepších žáků, dvojku 8 dalších atd. až po 2 nejhůřší, kteří dostanou nedostatečnou. Pojal tak písemku jako test ryze rozlišující. V takovém případě by se mohlo stát, že i žák, který se na písemku důkladně připravil a prokázal v ní výborné znalosti a dovednosti, by mohl dostat trojku, protože 13 žáků podalo ještě lepší výkon než on. Nebo naopak i velmi slabě napsaná písemka by byla ohodnocena jedničkou, protože většina ostatních ji napsala ještě hůře.

3. ZAČÍNÁME TVOŘIT DIDAKTICKÝ TEST

3.1 OD ČEHO ZAČÍT ANEB PLÁNOVÁNÍ TESTU

Vytvořit dobrý didaktický test určený pro široké použití je poměrně náročný proces, na kterém se podílí řada odborníků. Není možné, aby takový test vznikl za víkend, je to otázka týdnů, spíše měsíců. Tvorba celého testu probíhá v několika fázích a pokud některé z nich opomineme, zvyšuje se riziko, že vytvořený test nebude kvalitní.

Na tomto místě je třeba poznamenat, že různé didaktické testy se mezi sebou zcela pochopitelně liší i mírou náročnosti přípravy. Jinak vypadá příprava na desetiminutový test zadávaný ve vlastní třídě, kterým chceme ověřit, zda žáci ovládají v praxi pravidla pro psaní velkých písmen, a jinak se připravuje souhrnný (např. maturitní) celostátně zadávaný test z českého jazyka nebo test určený pro přijímací zkoušky na vysokou školu. Čím úžeji je vymezen obsah učiva a čím homogennější je cílová skupina (žáci jedné třídy, jednoho oboru), tím je plánování testu jednodušší.

Nejprve musíme vědět, **CO** přesně chceme testováním zjistit (účel testu, k čemu mají sloužit výsledky testu) a následně se ptáme, **JAK** toho dosáhnout, přičemž odpověď na účel testování často ovlivňuje i zvolený způsob a formu. V praxi plošného testování na krajské nebo celostátní úrovni zpravidla o první otázce rozhoduje příslušná instituce (krajský úřad nebo ministerstvo). Realizací pak pověří výkonnou složku (státní nebo soukromou organizaci zabývající se testováním) a jejím úkolem je, aby připravila konkrétní testy, zorganizovala jejich zadávání a následné zpracování výsledků.

Z výše uvedeného vyplývá, že dříve, než se pustíme do tvorby úloh, potřebujeme **znát účel testování, testované učivo a cílovou skupinu žáků**. Na základě těchto východisek pak určíme, jak dlouho by měl test trvat, jakou bude mít formu, jaký typ úloh by se v něm měl/neměl objevit, jakým způsobem se bude celý test vyhodnocovat apod. Ve fázi plánování nesmíme zapomenout také na to, kolik nám příprava a realizace testu zabere času, jakým způsobem se budou zpracovávat výsledky žáků, kolik budeme potřebovat spolupracovníků a jaké budou celkové finanční náklady.

V další fázi, ještě před samotnou tvorbou testových úloh, je třeba přesně vymežit, co bude obsahem testu – tj. očekávané znalosti a dovednosti žáka, tematické celky učiva a specifické cíle. Vznikne tzv. **specifikační tabulka**.

Ptáte se, k čemu je to dobré? Pokud bychom se jako autoři testu nebo autoři jednotlivých úloh nechali vést jen svým odhadem, zkušeností z výuky, učebnicemi nebo např. úlohami z předchozích testů, pravděpodobně bychom se více či méně minuli s požadavky testu. Ze samotného znění již

použitých úloh totiž nevyčteme zcela přesně a v plném rozsahu konkrétní specifické cíle (každý test pokrývá nutně jen část z celého testovaného učiva). Může se to zdát jako samozřejmost, ale překvapivě velké množství autorů se dívá spíše na předchozí testy, než do specifikačních tabulek. Možná je to dáno i tím, že mnoho testů postrádá řádnou obsahovou specifikaci v písemné podobě.

Specifikační tabulka stanovuje, jaká úroveň osvojení znalostí má být celým testem zkoušena. Podle účelu testu je ve specifikační tabulce upřesněno, zda se test zaměřuje na pamětní osvojování učiva, nebo je jeho cílem ověřit i vyšší úroveň dovedností, jako je porozumění poznatkům, aplikace poznatků, analýza a syntéza apod. V tomto ohledu jsou dobrou pomocí tzv. taxonomie výukových cílů (např. Bloomova taxonomie výukových cílů, Tollingerové taxonomie učebních úloh nebo Niemiervova taxonomie výukových cílů). Míru dosažení určité cílové kompetence vyjadřují aktivní slovesa definující žákovu dovednost, kterou lze testovat (viz např. tabulka níže). Při sestavování specifikační tabulky je učivo rozčleněno do několika tematických celků, kterým je přiřazena určitá váha, a tím i určitý počet úloh v testu.

Příklad specifikační tabulky

OBČANSKÝ ZÁKLAD

Písemný didaktický test pro maturitní ročník

Délka testu: 90 min.

Celkový počet úloh: 50

Tematické okruhy	Cílové dovednosti			Počet úloh	Úlohy v %
	A. Znalost s porozu- měním	B. Aplikace zna- lostí a práce s informacemi	C. Hodnotící posouzení		
	<i>žák dovede: popsat, defi- novat, vyjme- novat, určit, vymezit, vy- světlit, identifi- kovat, objas- nit, odhad- nout, přeložit, převést</i>	<i>žák dovede: porovnat, roz- lišit, rozhod- nout, rozpo- znat, aplikovat, řešit, demon- strovat, načrt- nout</i>	<i>žák dovede: posoudit, zhodnotit, zdůvodnit, vyvrátit, obhájit, rozhodnout</i>		
Člověk jako jedinec	1	2	1	4	8
Člověk a společnost	3	2	0	5	10
Občan a stát v dějinách a současnosti	4	8	3	15	30
Občan a právo	3	2	2	7	14
Občan a hospodářství	3	2	1	6	12
Občan, Evropa a soudobý svět	3	3	2	8	16
Člověk a praktická filozofie	2	2	1	5	10
Celkem úloh	19	21	10	50	100

Specifikační tabulka může obsahovat řadu dalších informací, které buď uve-
dený počet bodů nahrazují, nebo doplňují. Může se jednat například
o počet úloh či typ úloh. Čím podrobnější je specifikační tabulka, tím srov-
natelnější jsou testy, které jsou podle ní vytvořené, a současně tím přesněji
lze vymežit, co přesně test ověřuje. Přílišná podrobnost však ale může tvůr-
ce testu a tvůrce úloh svazovat. Proto specifikační tabulky obvykle vymezu-
jí pouze témata, úroveň osvojení a počet bodů či počet úloh.⁵

V praxi Centra pro zjišťování výsledků vzdělávání sestavuje specifikační
tabulku autor testu podle předem daných požadavků, které má test splňo-
vat, a autoři testových úloh se této specifikace musí držet. Specifikační
tabulka představuje jakýsi most mezi kurikulárními dokumenty a vlastním
testem. Slouží jako základní vodítko při konstrukci jednotlivých úloh i celého
testu.

Určitým doplňkem specifikační tabulky je tzv. **katalog požadavků** neboli
kompletní soupis toho, co bude předmětem zkoušky. Příkladem mohou být
zveřejněné katalogy požadavků k maturitní zkoušce, které obsahují kon-
krétně vymezené maturitní požadavky, event. jednotlivé specifické cíle.

5 Počet úloh je vhodným údajem při binárním hodnocení žákovských řešení. Pokud
se v testu používá vážené bodování (tj. za úlohy lze získat různý počet bodů), obsa-
huje specifikační tabulka počet bodů, který ukazuje váhu daného tématu a kognitiv-
ní operace v testu.

Ukázka z katalogu požadavků (specifických cílů) k maturitní zkoušce z českého jazyka zadávané MŠMT

1. PÍSEMNÝ DIDAKTICKÝ TEST

Žák

1.1 ovládá pravidla českého pravopisu

1.1.1 ovládá základní pravidla českého pravopisu (i/y; bě, pě, mě, vě a konkurující varianty; pravopis předložek a předpon; velká písmena; skupiny souhlásek a zdvojené souhlásky; hranice slov v písmu; interpunkce; psaní cizích slov);

1.2 provede slootovornou a morfologickou analýzu slovního tvaru

1.2.1 určí slovnědruhovou platnost slova;

1.2.2 rozliší funkčně/nefunkčně utvořené slovo;

1.2.3 identifikuje a opraví v daném kontextu morfologicky chybný tvar slova;

1.3 pochopí význam pojmenování a dovede ho popsat

1.3.1 posoudí vhodnost užití pojmenování v daném kontextu;

1.3.2 odhadne význam pojmenování v daném kontextu i mimo něj;

1.3.3 přiřadí k pojmenování v daném kontextu synonyma a antonyma;

1.3.4 rozezná obrazné a neobrazné pojmenování;

1.4 provede syntaktickou analýzu věty, souvětí a textu

...

Stejně jako u specifikační tabulky lze pak jednotlivým specifickým cílům přiřadit určitý počet úloh v závislosti na jejich výukovém významu. Úroveň osvojení poznatků již vyplývá z formulace jednotlivých specifických cílů.

PLÁNOVÁNÍ TESTU – SHRNUÍ:

- definování cíle testu, tj. k čemu mají sloužit získané výsledky;
- vymezení obsahu testu, tj. jaké znalosti a dovednosti mají být prostřednictvím testu ověřovány;
- určení žákovské populace, tj. pro jaký typ školy, studijní obor a ročník studia je test určen;
- zvolení účelu/druhu testu (podle kap. 2a);
- vypracování specifikační tabulky testu;
- stanovení výstižného názvu testu;
- odhad počtu spolupracovníků, časových a finančních nákladů pro konstrukci testu.

3.2 KONSTRUKCE TESTU

Ve fázi konstrukce testu jde především o vytvoření jednotlivých testových úloh a o jejich vhodné zakomponování do testu. Specifikace testu určuje, jaké testové úlohy se v testu použijí, v jaké míře apod. Praxe v Centru pro zjišťování výsledků vzdělávání je taková, že autor testu spolupracuje s týmem autorů úloh, kteří mu vytvářejí zásobu testových úloh. Následně pak autor testu spolu s expertní komisí posuzuje kvalitu jednotlivých úloh především s ohledem na:

- shodu úlohy se zadaným specifickým cílem a její validitu (že skutečně měří to, co má měřit);
- věcnou správnost, jednoznačnost a nezpochybnitelnost správné odpovědi (zda je „správné“ řešení opravdu správné, jednoznačné a nepřekrývá se s jinými variantami odpovědí);
- technickou kvalitu úlohy (zda je úloha vhodně formulována, je dostatečně srozumitelná, přiměřeně dlouhá, vztahuje se k zadání, nepředpokládá u žáka jiné, výrazně nadstavbové znalosti či dovednosti apod.);
- obtížnost úlohy (zda úloha svou obtížností odpovídá cílové skupině žáků a koncepci testu).

Má-li již autor testu k dispozici dostatečné množství kvalitních úloh odpovídajících svou povahou specifikační tabulce, může postoupit k dalším fázím konstrukce testu:

- výběr úloh do testu podle specifikační tabulky;
- řazení úloh do testu;
- provedení odhadu časové náročnosti testu a porovnání s časovým limitem uvedeným ve specifikační tabulce;
- volba způsobu hodnocení jednotlivých úloh a celého testu ve shodě se specifikační tabulkou;
- sestavení záznamového archu, do něhož žáci zapisují řešení úloh;
- popis organizace testování (vybavení učebny, povolené pomůcky, podrobný popis průběhu testování včetně zpracování písemných instrukcí pro zadavatele).

3.3 OVĚŘOVÁNÍ TESTU

Test určený pro široké použití musí projít fází ověřování, při níž je obvykle posouzen odbornými recenzenty po stránce obsahové, konstrukční a jazykové. Následně je zadán určité skupině žáků (vzorku z cílové skupiny) a na základě jejich výsledků se posuzuje, zda je test nebo jednotlivá úloha dostatečně kvalitní. Tento způsob **ověřování (tzv. pilotáž)** klade velké nároky na zabezpečení před únikem informací. Pokud je ověřován přímo celý test, je riziko prozrazení neúměrně vysoké. Proto se častěji než celé testy ověřují pouze jednotlivé úlohy nebo části budoucího testu. V takovém případě probíhá ověřování testu ještě před jeho konečným sestavením. U každé úlohy jsou statistickými metodami sledovány určité ukazatele, podle nichž je pak úloha schválena, upravena nebo vyřazena. (O pilotážích a analýze testových úloh pojednává podrobněji kapitola 10.) Test je sestavován z úloh, které samostatně prošly ověřením a jsou konstrukčně i obsahově vyhovující.

Detailní analýza charakteristik testové úlohy nám pomůže také při **sestavování paralelních verzí testu** (*parallel forms*), od nichž se očekává, že budou plně srovnatelné. Celá řada testů je zpracovávána v paralelních verzích, nejčastěji A a B. V praxi je dosažení naprosté srovnatelnosti a rovnocennosti obou verzí poměrně obtížné. Základním předpokladem je použití zcela identické specifikační tabulky při konstrukci obou verzí testu. Určitou možností představuje vytváření „paralelních“ úloh verze B podle vzoru verze A na principu změn některých údajů v zadání (jmen lidí, názvů předmětů, číselných hodnot a jiných atributů). Na základě zkušeností nelze tento způsob jednoznačně doporučit, protože ony druhé verze B bývají často konstruovány jako obtížnější. Určitým řešením je vytvářet zároveň dvojice rovnocenných úloh, které se pak rozlosují mezi verzi A i B. Paralelní verze testu lze vytvořit též změnou pořadí úloh.

4. ZAČÍNÁME S TESTOVÝMI ÚLOHAMI

4.1 JAKOU KVALIFIKACI BY MĚL MÍT AUTOR TESTOVÝCH ÚLOH?

Pro autora testových úloh je užitečné, má-li živé zkušenosti s učením daného předmětu u takové skupiny žáků, která je co nejvíce podobná skupině testované. Díky těmto zkušenostem autor snáze posoudí, co bude pro žáky snadné a co obtížné, co je zajímavé, jaké je jejich kulturní zázemí, úroveň a styl vyjadřování apod. Většina autorů úloh se rekrutuje z řad učitelů nebo bývalých učitelů, přičemž však platí, že být dobrým učitelem ještě neznamená být automaticky dobrým autorem testových úloh. Kromě zkušenosti je zapotřebí notná dávka tvořivosti, fantazie, citlivosti, intuice, logického uvažování, přesnosti a dalších vlastností, které se u budoucích autorů obtížně rozpoznávají předem. V cizině existují i profesionální autoři testů nebo testových úloh na plný úvazek, u nás se však jedná spíše o práci příležitostnou.

Vlastnosti dobrého autora testových úloh jsou podobné vlastnostem dobrého autora učebnic. Testové úlohy se do značné míry podobají otázkám v opakovacích cvičeních z učebnic. I při psaní učebnic je nutná velká míra tvořivosti a schopnost utvořit příklady a cvičení pro žáky vhodné a přitom zajímavé. Obecně se má za to, že napsat testovou úlohu je poměrně těžké. Proč to platí především o testových úlohách a už méně o cvičeních v učebnicích? Je zde jeden podstatný rozdíl – píše-li žáci test, nemohou při jeho řešení očekávat žádnou dodatečnou pomoc od učitele, od spolužáků nebo z jiných než povolených zdrojů. U cvičení to tak docela neplatí. Proto musí být testová úloha formulována naprosto jednoznačně a velmi srozumitelně. Testovou úlohu nelze žádným způsobem dovysvětlit, upřesnit, doplnit příkladem nebo prostě přeskočit. Dobrý autor testových úloh by tedy měl mít kromě přesného logického myšlení též vynikající formulační schopnost a v neposlední řadě by měl dokázat přijmout případnou kritiku své práce. Posuzování úloh se totiž v další etapě účastní předmětoví specialisté, jazykoví korektoři, analytici a autor celého testu, přičemž „úmrtnost“ úloh zvláště u začínajících autorů je poměrně vysoká.

4.2 TESTOVÁ ÚLOHA A JEJÍ ČÁSTI

Testová úloha je nejmenší, relativně samostatnou částí testu, na niž žák odpovídá. Relativně samostatná je proto, že na ni žák může odpovědět, aniž předtím vyřešil jiné úlohy, avšak výsledek konkrétní úlohy je (ba dokonce musí být) v konečném testovém výsledku provázán s výsledky ostatních úloh. Úloha má zpravidla podobu úkolu, otázky, neúplného textu, tvrzení, o jehož pravdivosti je třeba rozhodnout apod. Úloha může testovat jednu konkrétní znalost či dovednost, anebo je složená z více kroků a při hodnocení lze brát v úvahu též použitou strategii a správnost postupu řešení. Kromě termínu „testová úloha“ se můžeme v literatuře setkat ještě s označením „testová položka“ nebo „testový úkol“, v praxi se často užívá označení „otázka“, „příklad“ nebo „úkol“.

Obecně může mít testová úloha několik částí. Nazýváme je: **instrukce**, **výchozí text**, **kmen úlohy**, **alternativy**, **distraktory** a **správné řešení**. To, zda jsou všechny zmíněné části u konkrétní úlohy zastoupeny, záleží na jejím typu.

1. Instrukce – podává návod, co má žák dělat, jak má postupovat při řešení (týká se úloh s výchozím textem).

Přečtěte si následující text a odpovězte na úlohu.

2. Výchozí text – materiál (jde většinou o text, ale může to být též tabulka, graf nebo schéma), který je východiskem k řešení úlohy. Na základě výchozího textu žáci aplikují své znalosti. Výchozí text navozuje problém nebo situaci, kterou pak žáci posuzují či interpretují.

Otec začal na Jirkovi pozorovat určité změny. Jirka prochází bouřlivou proměnou po stránce tělesné a psychické. Jeho nálada se rychle mění, odmlouvá, kritizuje rodiče a jejich názory. Více mu záleží na názorech jeho vrstevníků. Odpoutává se od rodičů a patrná je snaha prosazovat vlastní názory a postoje.

(zdroj textu⁶)

3. Kmen úlohy – zadání ve formě otázky, pokynu nebo řidčeji neúplného tvrzení, které žák doplňuje vlastními slovy nebo výběrem z alternativ.

V které etapě života se Jirka nachází?

4. Alternativy – v úlohách s výběrem odpovědi jsou to všechna nabízená řešení, tj. správná i nesprávná.

A) adolescence

B) dospělost

C) mladší školní věk

D) puberta

5. Distraktory – v úlohách s výběrem odpovědi je to označení pro nesprávná řešení.

alternativy A, B a C.

6. Správné řešení

alternativa D.

6 Pokud byl výchozí text odněkud převzat (doslovně nebo v upravené podobě), uvádí se zdroj v závorce pod výchozím textem. Např.: (*Boleslavský deník 25. 3. 2002, upraveno*).

5. ÚLOHY UZAVŘENÉ A OTEVŘENÉ

Jeden a tentýž specifický cíl je zpravidla možné ověřit více typy úloh. Každý typ úlohy má určité vlastnosti, má své výhody a nevýhody a v jednom testu se obvykle kombinuje více druhů testových úloh. O tom, který typ úlohy použijeme, rozhoduje cíl testování, obsah učiva, požadované rozumové operace a také materiální a technické podmínky testování.

Úlohy je možné třídit podle mnoha různých hledisek, avšak tím základním je dělení podle způsobu, jakým žák úlohu řeší, na **úlohy uzavřené a otevřené**.

V úlohách uzavřených je žákovi nabízeno několik alternativ, z nichž žák vybírá jednu (v některých případech i více než jednu) správnou.

Otevřené úlohy vyžadují, aby žák sám odpověď vytvořil. Odpověď může být slovo, číslo, výpočet, jedna věta nebo i delší text.

5.1 UZAVŘENÉ ÚLOHY

Mezi uzavřené úlohy patří:

- **úlohy s dvoučlennou volbou (dichotomické)**
- **úlohy s výběrem z více odpovědí**
- **úlohy přiřazovací**
- **úlohy uspořádací**

VÝHODY UZAVŘENÝCH ÚLOH:

- Vyhodnocení je velmi objektivní, lze jednoznačně rozhodnout, zda je odpověď správná, nebo nesprávná.
- Vyhodnocení odpovědí je poměrně rychlé a při záznamu odpovědí do předtištěného formuláře může být do značné míry automatizováno použitím šablony nebo strojovým zpracováním.

- Uzavřené úlohy jsou vhodné pro žáky, kteří mají potíže s vlastním formulováním odpovědi nebo pomalu píší.
- Odpověď není závislá na žákově vyjadřovací či formulační schopnosti a zcela minimálně na jeho grafomotorické zdatnosti (jde nejčastěji o zakřížkování či zakroužkování správné odpovědi).

NEVÝHODY UZAVŘENÝCH ÚLOH

- Některé dovednosti produktivní povahy (jako zformulovat stanovisko, nakreslit obrázek apod.) se nedají uzavřenými úlohami testovat vůbec nebo jen velmi omezeně.
- Znevýhodňují nepozorné a roztržité žáky (mohou přehlédnout mezi alternativami správnou odpověď nebo se spletou při zaškrtnutí), ale také žáky přespříliš hloubavé, kteří nad alternativami znejistí a hledají v nich „chyták“.
- Existuje jistá pravděpodobnost uhodnutí správné odpovědi. Tato pravděpodobnost vzrůstá u úloh s malým počtem alternativ (2, 3) či při použití nevhodných distraktorů.
- Nelze vysledovat myšlenkový postup, kterým se žák dobral k řešení.
- Vytvoření dobré uzavřené úlohy není úplně snadná záležitost. Odhalení konstrukčních nedostatků nemusí být zjevné ze žakovských odpovědí a vyžaduje si hlubší rozbor a analýzu.
- Snadněji se opisuje.
- Učitelská veřejnost (alespoň u nás) je k nim dosud poněkud skeptická, ve výuce se používají jen málo.

5.2 OTEVŘENÉ ÚLOHY

Mezi otevřené úlohy řadíme:

- **úlohy se stručnou odpovědí**
- **úlohy se širokou odpovědí**

VÝHODY OTEVŘENÝCH ÚLOH:

- Některé dovednosti (zejména produktivní, kreativní povahy) se dají testovat téměř výhradně otevřenými úlohami.
- Při formulaci odpovědí jsou žáci nuceni používat odbornou terminologii, čímž zároveň prokáží, že ji znají a umějí správně používat.
- Autor úlohy či testu může z žakovských odpovědí snadno poznat, nakolik žáci porozuměli zadání, zda byla úloha obsahově nebo konstrukčně chybná, či nikoli.

NEVÝHODY OTEVŘENÝCH ÚLOH

- Otevřená úloha vyžaduje naprosto jasné a jednoznačné zadání, aby se předešlo všem (očima žáka) možným, avšak (očima autora testu) mylným interpretacím.
- Čím je úloha více otevřená, tím obtížnější je zajistit její objektivní hodnocení. Jeden hodnotitel se může od druhého v posouzení konkrétní odpovědi velmi lišit, byť se řídili stejnými instrukcemi. Tento problém narůstá s tím, jak narůstá počet testovaných žáků. Zpravidla se využívá více kritérií hodnocení a bývají odstupňovaná na širší škále než dvoustupňové „správně–nesprávně“.
- Žáci komunikačně slabší jsou v nevýhodě, jelikož formulace odpovědi zpravidla ovlivňuje její hodnocení.
- Náročné a pracné bývá stanovení jasných kritérií hodnocení a sestavení úplného kódového klíče pro všechna (možná i „nemožná“) řešení.

- Hodnocení je časově i personálně náročnější. Vyžaduje od hodnotitele soustředěné čtení žákova rukopisu. Případná digitalizace a automatické hodnocení jsou prakticky vyloučeny (jistou výjimku představují texty psané přímo na počítači).

6. TVORBA UZAVŘENÝCH TESTOVÝCH ÚLOH

V této kapitole pojednáme podrobněji o jednotlivých typech uzavřených úloh s důrazem na úskalí, která jsou spojena s jejich tvorbou.

6.1 ÚLOHY S DVOUČLENNOU VOLBOU (DICHOTOMICKÉ)

Úlohy s dvoučlennou volbou bývají označovány též jako úlohy dichotomické nebo z angličtiny přejatým termínem *true-false items*. Jejich vytvoření je rychlé a snadné (nebo se tak alespoň jeví). Skutečnost je taková, že svou strukturou jsou sice dichotomické úlohy jednodušší než jiné typy úloh, avšak vymyslet kvalitní dichotomickou úlohu úplně snadné není.

Řešení dichotomické úlohy spočívá ve volbě mezi dvěma alternativami (ANO–NE, PRAVDA–NEPRAVDA, SPRÁVNĚ–NESPRAVNĚ). Slabinou dichotomických úloh je právě počet alternativ, kdy má žák 50% šanci, že uhodne správné řešení.⁷ Proto bývají dichotomické úlohy řazeny do svazků složených

⁷ Hádání je regulérní metoda řešení v případě, že žák nezná správnou odpověď. Ostatně, většinou se nejedná o klasické hádání „naslepo“, ale o určitou žákovu strategii řešení, o úvahu vedoucí k volbě (tipu) mezi nejpravděpodobnějšími alternativami. Neodrazujeme žáky od hádání, už proto, že mu stejně nelze zabránit. Naopak, tím, že na hádání jako možnou strategii řešení upozorníme všechny, eliminujeme do jisté míry výhodu „testově zkušenějších“ žáků oproti ostatním, kteří by snad z obavy před špatnou odpovědí raději nezvolili žádnou. To platí tehdy, nejsou-li v testu nesprávné odpovědi penalizovány minusovými body.

z více jednotlivých dichotomických úloh (resp. podúloh). K dosažení příslušného bodového zisku je zapotřebí správně zodpovědět všechny, příp. jiný stanovený počet podúloh svazku. Podúlohy by se měly vztahovat k jednomu tematickému celku, měly by mít shodný nebo velmi blízký specifický cíl, k řešení by měly být použity stejné nebo podobné myšlenkové operace.

Svazky dichotomických úloh se někdy používají ve vazbě na výchozí text, v němž žák může nalézt potřebné informace.

Příklad svazku úloh s dvoučlennou volbou

Rozhodněte, zda jsou následující tvrzení pravdivá (ANO), nebo nepravdivá (NE). Správnou odpověď zakroužkujte.

10.1 Nádobu s objemem $V = 3 \text{ dm}^3$ obsahuje kapalinu o objemu $V_1 = 0,5 \text{ dm}^3$. Ani po přilítí další kapaliny o objemu $V_2 = 2 \text{ 500 ml}$ nádoba nepřeteče.

ANO NE

10.2 V cisterně je 25 hl tekutiny, což je stejné množství jako 25 m^3 .

ANO NE

10.3 Vzdálenost 1 cm na mapě s měřítkem 1 : 50 000 odpovídá vzdálenosti 0,5 km ve skutečnosti.

ANO NE

10.4 50 procent ze dvou hodin je totéž jako 2 procenta z 50 hodin.

ANO NE

DOPORUČENÍ PRO TVORBU DICHOTOMICKÝCH ÚLOH

- V instrukcích pro řešení testu musí být zcela jednoznačně určeno, jakým způsobem se označí odpověď. (Jistě jste už někdy narazili na test nebo formulář, kde jste váhali, zda máte zaškrtnout to, co platí, nebo škrtnout, co neplatí.) Zvolený způsob musí být jednotný pro celý test či skupinu testů a není třeba jej uvádět u každé úlohy znovu.
- Odpověď na tvrzení v úloze musí být jednoznačná (ANO, nebo NE) a nezávislá na dalších podmínkách.
- U daného tvrzení pokud možno vymezte kontext nebo konkrétní zdroj, k němuž se tvrzení vztahuje. (Zpravidla se neptáme na osobní názor či postoj žáka.)

Špatně

Každý může činit, co není zákonem zakázáno.

ANO NE

Správně

Rozhodněte, zda je následující tvrzení v souladu s Listinou základních práv a svobod, která je součástí ústavního pořádku ČR.

Každý může činit, co není zákonem zakázáno.

ANO NE

- Pravdivých i nepravdivých tvrzení by měl být v testu přibližně stejný počet a měla by být náhodně prostřídána, aby nevznikala schémata (např. ANO, ANO, ANO, NE, NE, NE).

- Pravdivá i nepravdivá tvrzení by měla mít přibližně stejnou délku, aby správná nebyla výrazně delší (např. proto, že jsou terminologicky přesnější).
- Uvážlivě nakládejte s používáním záporu. Odpověď „NE“ na negativně formulovanou otázku je vlastně „pastí“ dvojí negace. Odpověď na záporné tvrzení by měla být raději volena z alternativ PRAVDA–NEPRAVDA.

Špatně

Na Měsíci neplatí Newtonovy gravitační zákony.

ANO NE

Správně

Na Měsíci neplatí Newtonovy gravitační zákony.

PRAVDA NEPRAVDA

- V tvrzeních se vyhněte výrazům jako „vždy“, „často“, „téměř“, „nikdy“, „zřídka“ apod.
- Nepoužívejte vět vytržených z kontextu učebnice, ani je neobměňujte zařazením záporu.
- Tvrzení formulujte stručně.

6.2 ÚLOHY S VÝBĚREM Z VÍCE ODPOVĚDÍ

Široce používaným typem úloh jsou úlohy s výběrem odpovědi, přesněji s výběrem odpovědi z více než dvou alternativ (*multiple-choice items*). Úlohy

s výběrem z více odpovědí mohou velmi účinně testovat cíle jednoduché i velmi složité, např. porozumění či aplikaci poznatků. Jejich předností je snadné vyhodnocování a oproti dichotomickým úlohám jsou méně „náchylné“ k uhodnutí správného řešení.

Mají dvě základní podoby:

- **ukončená otázka**
- **nedokončená věta**

Ukončená otázka – kmen je běžnou otázkou a alternativy představují možné odpovědi. Žák vybírá jednu správnou. Někdy se používá též výběr jedné nejpřesnější odpovědi. (Méně často výběr více správných odpovědí nebo výběr jediné nesprávné odpovědi.)

Příklad ukončené otázky

Na trase Brno – Ostrava je úsek silnice dlouhý 10 km, kde je doporučená rychlost $70 \text{ km} \cdot \text{h}^{-1}$. Kolik času řidič automobilu ušetří, jestliže v tomto úseku překročí rychlost o $20 \text{ km} \cdot \text{h}^{-1}$?

- A) 8,6 min
- B) 3,0 min
- C) 1,9 min
- D) 6,7 min

Nedokončená věta – kmen úlohy je nedokončenou větou a jednotlivé alternativy představují různá možná dokončení. Žák má určit, která z alternativ doplní tvrzení tak, aby bylo správné a pravdivé.

Příklad nedokončené věty

Muž je ve vztahu k ženě jako chlapec k _____ .

- A) otci
- B) matce
- C) dívce
- D) chlapci

Tento typ úloh není optimální, protože nestimuluje konkrétní kognitivní operace, ale žák si musí domýšlet, co se po něm v úkolu vyžaduje.

DOPORUČENÍ PRO TVORBU ÚLOH S VÝBĚREM ODPOVĚDI

- Žáci musí být předem seznámeni s charakterem úloh – zda se v úlohách vyskytuje jedno (a právě jedno) správné řešení, nebo zda může být správných řešení více, resp. žádné. V testech, které připravuje Centrum pro zjišťování výsledků vzdělávání, používáme pouze jednu správnou odpověď a jen kladné hodnocení (za nesprávné odpovědi se body neodečítají).
- Počet alternativ volíme mezi 3 a 5, přičemž optimální a nejčastěji používaný je počet 4 (1 správné řešení a 3 distraktory). Nižší počet alternativ nahrává hádání a vyšší počet distraktorů má za následek jejich nižší kvalitu (dají se vyloučit na první pohled, a neplní tedy svou funkci).
- Otázku konstruuje až na základě přesného znění správného řešení.
- „Správné“ řešení musí být jednoznačně a nezpochybnitelně správné a musí být právě jedno (pokud není předem stanoveno jinak). Toto konstatování se zdá být zcela samozřejmé, avšak právě úlohy s výběrem odpovědi trpí mnohdy chybou spočívající v nejednoznačné správnosti řešení (např.

pochopení úlohy v jiném významovém kontextu, který autor nepředpokládal). Je třeba pečlivě posoudit všechny distraktory, zda jsou skutečně a zcela jednoznačně nesprávné.

Špatně

Které zvíře nepatří mezi ostatní?

- A) ovce
- B) koza
- C) tele
- D) zajíc

KOMENTÁŘ: *Autor měl zřejmě na mysli správné řešení D, avšak někteří žáci mohou za správné řešení považovat alternativu C, protože jako jediná označuje mládě.*

- V úloze se vyhýbáme všem slovům nebo údajům (resp. takovým jejich gramatickým tvarům), které by mohly sloužit jako nápověda.

Špatně

Someone who designs houses is an _____ .

- A) designer
- B) builder
- C) architect
- D) plumber

KOMENTÁŘ: *Tvar neurčitého členu „an“ předem vylučuje alternativy A, B a D.*

Správně

Someone who designs houses is a/an _____ .

- A) designer
- B) builder
- C) architect
- D) plumber

- Pro formulaci distraktorů využijte chyb, kterých se žáci nejčastěji dopouštějí, používejte v nich slova, která mají asociaci s výchozím textem nebo kmenem úlohy, využívejte formulací z učebnic nebo frází, které se zdají být pravdivé apod.

- Alternativy by si měly být podobné obsahem i formou. Správná odpověď by se neměla nijak významně lišit od distraktorů (např. stylem vyjádření, přesností formulace nebo délkou). Distraktory by měly být všechny přibližně stejně atraktivní – distraktor, který nebude využit, je zbytečný. Dbejme na to, aby nebylo možné dospět ke správnému řešení pouze vyloučením zcela nesmyslných či na první pohled banálních distraktorů.

Špatně**Co znamená, že obviněný bude „vzat do vazby“?**

- A) bude pevně svázán
- B) bude zadržen po nezbytně dlouhou dobu, aby nemařil vyšetřování případu
- C) bude přijat do společenství ostatních obviněných
- D) bude pozván na prohlídku vězení

KOMENTÁŘ: *Správná odpověď B je výrazně delší a distraktory A, C, D jsou natolik odtažitá od učiva, až absurdní, že je lze okamžitě vyloučit bez znalosti správné odpovědi.*

Správně**Co znamená, že obviněný bude „vzat do vazby“?**

- A) bude soudem odsouzen a uvězněn, dokud se neprokáže, že čin nespáchal
- B) bude po nezbytně dlouhou dobu uvězněn, aby nemařil vyšetřování případu
- C) bude uvězněn a nepřipuštěn k hlavnímu líčení, ve kterém nemůže vypovídat
- D) nebude uvězněn, ale bude pod neustálým dohledem policie a soudu

- Jednoduchý zápor používejte jen tam, kde je to opravdu nutné, a graficky ho zvýrazněte (tučné písmo nebo podtržení).

- Umístění správné odpovědi mezi distraktory volte zcela náhodně nebo třeba řaďte alternativy v abecedním pořádku podle začátečních písmen. Autoři tíhnou k tomu nedávat správnou odpověď na první místo. Žáci na to mohou přijít.
- Alternativy znějící „žádná z výše uvedených odpovědí“ nebo „všechny výše zmíněné odpovědi“ používejte zřídka.
- Alternativy by se neměly významově překrývat a taktéž formální podoba (příliš mnoho stejných slov) může být matoucí.

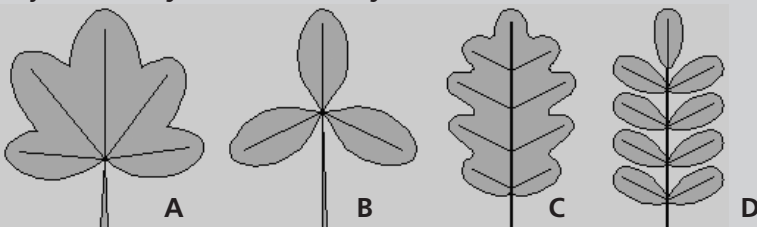
6.3 ÚLOHY PŘÍRAZOVACÍ

Přířazovací úlohy (*matching items*) jsou zpravidla tvořeny dvěma seznamy pojmů, výroků, čísel nebo symbolů, přičemž žák má v těchto seznamech nalézt vzájemně si odpovídající dvojice. Přířazovací úlohy jsou velmi vhodné pro testování konkrétních informací a jejich vzájemných vazeb. Jeden ze seznamů obsahuje jednu i více položek, ke kterým neexistuje protějšek z druhého seznamu (nemohou utvořit dvojici).

Příklad úlohy přiřazovací

K následujícím třem názvům typů listů (1–3) přiřadte jejich schematické zobrazení (A–D).

- | | | | | |
|------------------------------|---|---|---|---|
| 1 složený zpeřený | A | B | C | D |
| 2 dlanitě složený | A | B | C | D |
| 3 jednoduchý dlanitě členěný | A | B | C | D |



DOPORUČENÍ PRO TVORBU PŘÍŘAZOVACÍCH ÚLOH

- Seznamy by měly být poměrně krátké a stejnorodé – oba by se měly vejít na jednu stránku. Možnosti uspořádejte v určitém logickém pořádku. Pokud to nejde, úloha zřejmě obsahuje příliš rozmanité položky. Zkuste je rozdělit.
- Každý bod v prvním seznamu by měl být očíslován (každý z nich je vlastně úlohou) a každá položka v seznamu alternativ by měla být odlišena písmenem. Označení by mělo být zmíněno už v zadání úlohy.
- Uveďte více alternativ, než je položek výchozího seznamu, aby se eliminovala vylučovací metoda (co zbude).
- Vždy jasně definujte způsob vyznačení dvojic a pokud lze alternativy přiřadit více než jednou, výslovně na to upozorněte.

6.4 ÚLOHY USPOŘÁDACÍ

V uspořádacích úlohách (*ordering items*) se od žáka požaduje, aby předložené pojmy (tvrzení, čísla, symboly, části textu) uspořádal podle určitého pravidla. Instrukce musí vyjadřovat kritérium a způsob řazení. Prvky je možné řadit např. podle velikosti, chronologicky, podle míry obecnosti apod.

Příklad úlohy uspořádací

Výchozí text:

Čočková polévka

Tradiční čočková polévka je nejoblíbenější na jihu Indie.

Doba přípravy: 20 minut

Doba vaření: 15 minut

Porce: pro 4 osoby

Suroviny: 15 ml rostlinného oleje, 1 lžička hořčičných semínek, 5 celých červených chilli papriček, 10 lístků kari, 2 lžice rozetřeného česneku, 1 lžička kurkumy, 100 g omyté a přebrané červené čočky, 2–3 oloupaná a rozčtvrcená rajčata, 3–4 kuličky černého pepře, 1 zelená chilli paprička zbavená semínek a nakrájená na proužky, 100 g tamarindové dužiny, sůl.

Postup:

1. V hrnci rozpálíme olej. Vsypeme do něj hořčičná semínka a zahříváme je, dokud nezačnou prskat. Přisypeme celé chilli papričky, lístky kari a rozetřený česnek. Na pár sekund je za stálého míchání osmažíme.
2. Do hrnce přidáme kurkumu, čočku, rajčata, pepř, nakrájenou zelenou chilli papričku, tamarindovou dužinu a osolíme. Vlijeme asi litr vody.
3. Vodu přivedeme k varu a za občasného míchání vaříme, dokud čočka nezměkne.

Úloha:

Uspořádejte následující suroviny (A–F) v pořadí, v jakém se mají podle receptu přidávat do polévky (1. krok–6. krok).

- A) kurkuma
- B) voda
- C) sůl
- D) olej
- E) čočka
- F) rajčata

1.	krok	A	B	C	D	E	F
2.	krok	A	B	C	D	E	F
3.	krok	A	B	C	D	E	F
4.	krok	A	B	C	D	E	F
5.	krok	A	B	C	D	E	F
6.	krok	A	B	C	D	E	F

Určité problémy může působit bodové hodnocení uspořádacích úloh. Nesprávné řazení může být totiž provedeno mnoha způsoby, přičemž se jedná o různě závažné chyby. Nejjednodušším (avšak málo citlivým) způsobem hodnocení je 1 bod za zcela správné uspořádání a 0 bodů za všechna ostatní řešení. Jiný, složitější způsob skórování porovnává odchylky žákových odpovědí od správného pořadí s maximálními odchylkami vzniklými zcela opačným seřazením. Tento postup se doporučuje v případech, kdy je počet přiřazovaných položek větší než 5. (Blíže viz Byčkovský, 1982 nebo Chráska, 1999.)

Dalším postupem je hodnocení dvojic přiřazení. Například má-li správná odpověď znít „Jaro – Léto – Podzim – Zima“, hodnotíme, zda žáci správně uspořádali léto za jaro, podzim za léto, zimu za podzim. V uvedeném příkladu by se tedy hodnotily 3 uspořádané dvojice.

DOPORUČENÍ PRO TVORBU USPOŘÁDACÍCH ÚLOH

- U úloh uspořádacích, podobně jako u přiřazovacích, je nutné velmi přesně definovat způsob, jakým mají žáci vyznačit pořadí.
- V instrukci je nutné zmínit, který údaj má být uveden jako první, tj. směr uspořádání (od nejmenšího po největší, od nejmladšího po nejstarší apod.).

7. TVORBA OTEVŘENÝCH TESTOVÝCH ÚLOH

7.1 ÚLOHY SE STRUČNOU ODPOVĚDÍ

Otevřené úlohy se stručnou odpovědí (*short-answer items*) mohou být zodpovězeny jedním slovem, krátkým slovním spojením, symbolem, grafem, značkou nebo číslem. Na rozdíl od uzavřených úloh zde žáci prokazují schopnost vlastní formulace odpovědi, což značně snižuje možnost uhodnutí odpovědi.

Mezi velké výhody otevřených úloh lze počítat skutečnost, že se poměrně snadno navrhují. Učitelé i žáci jsou na tento typ úloh zvyklí, jelikož se podobá běžným otázkám v učebnicích a při zkoušení.

Úlohy se stručnou odpovědí mohou mít charakter **produkční**, kdy se od žáka očekává „vyprodukování“ odpovědi na přímou otázku, nebo **doplňovací**, kdy má úloha charakter doplnění neúplného tvrzení. (Doplňovací úlohy bývají v angličtině nazývány *cloze items* nebo *completion* či *supply items*.) Rozlišení na produkční a doplňovací se týká spíše formální podoby obou typů úloh, jinak mezi nimi nejsou z pedagogického hlediska významné rozdíly. Tam, kde si můžeme vybrat, dáváme přednost spíše typu produkčnímu před doplňovacím.

Příklad úlohy produkční

Napište jména dvou japonských měst, která byla na konci druhé světové války zničena atomovými bombami.

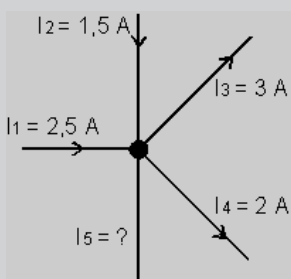
Příklad úlohy doplňovací

Na konci druhé světové války byla atomovými bombami zničena japonská města _____ a _____ .

Jiný příklad doplňovací úlohy

Na obrázku je uzel elektrické sítě.

Vyznačte do obrázku směr proudu procházejícího pátým vodičem.



DOPORUČENÍ PRO TVORBU OTEVŘENÝCH ÚLOH SE STRUČNOU ODPOVĚDÍ

- Volte takové úlohy, na něž se dá odpovědět velmi stručně. Odpověď formulujte zároveň se zněním úlohy.
- Nevyžadujte doslovné opakování textu z učebnice.
- Uvažte pečlivě všechny možné odpovědi a je-li jich mnoho, úlohu přeformulujte nebo vyřaďte.

Špatně

Co je jednotkou tlaku?

KOMENTÁŘ: *Na tuto otázku může být několik správných odpovědí – Pa, at, mb, torr apod. Odpověď může být zapsána i slovy, s gramatickými chybami nebo nepřesnostmi, navíc špatně čitelným rukopisem. V takových případech mohou vznikat pochybnosti, zda úloha byla zodpovězena správně.*

- Musíte si být jisti, že dokážete posoudit správnost žákovy odpovědi.

Špatně

Uveďte název alespoň jednoho německého filmu.

KOMENTÁŘ: *Žák uvedl odpověď „Vracím se z pekla“. Je prakticky nemožné posoudit, zda je jeho odpověď správná.*

- Na zápis odpovědi musí být ponecháno dostatečně velké místo.
- U doplňovacích úloh vynechávejte pouze podstatné údaje, ale jen tolik, aby smysl věty zůstal patrný.
- Vynechané údaje umísťujte nejlépe na konec věty.
- Z neúplné věty musí být jednoznačně zřejmé, co se má doplnit.

Špatně

Univerzita Karlova byla založena _____ .

KOMENTÁŘ: *Z formulace není jasné, zda má být doplněn rok, místo, jméno zakladatele apod. Nejednoznačné zadání v jistém smyslu znevýhodňuje žáky s širším než pouze „učebnicovým“ rozhledem, kterým se může formulace úlohy asociovat s jinou správnou odpovědí.*

Správně

V kterém roce byla založena Univerzita Karlova v Praze?

nebo:

Univerzita Karlova v Praze byla založena v roce _____ .

- Pokud má být odpověď číselný údaj spojený s nějakou jednotkou, vždy jednotku předem vypište.
- Předem stanovte, zda pravopis bude zahrnut do hodnocení, či nikoli. Jde o to rozhodnout, zda pravopisné chyby znamenají ztrátu bodů a pokud ano, pak je třeba definovat, jaké chyby (a jaké ne), v jaké míře výskytu, kolik bodů se za ně odečítá apod. Obvyklý přístup je takový, že, pokud pravopisná chyba nebrání pochopení podstaty odpovědi, nebere se při hodnocení v úvahu.

7.2 ÚLOHY SE ŠIROKOU ODPOVĚDÍ

Hranice mezi stručnou a širokou odpovědí není striktní. Obvykle se má za to, že pokud je odpověď v rozsahu jednoho slova až jedné věty, jedná se o stručnou odpověď. Dvě a více vět pak znamenají širokou odpověď. Zvláštní kategorii představuje otevřená úloha se širokou odpovědí – někdy označovaná jako strukturovaná písemná práce – kdy žák píše rozsáhlejší slohový útvar.

Určitým dělícím kritériem je také míra objektivity, s níž je možno hodnotit danou úlohu. Srovnajte následující příklady úloh:

Úloha 1

Přeložte do angličtiny větu: „Zítřka asi bude pršet.“

Úloha 2

Vystihněte jednou větou co nejlépe obsah následujícího textu.
(Následuje text.)

KOMENTÁŘ: Obě úlohy vyžadují odpověď jednou větou, ale zatímco u první úlohy je považována odpověď za stručnou, u druhé se kloníme spíše k označení široká právě pro menší míru objektivitu, s níž lze odpověď hodnotit.

Pokud má hodnotitel k ruce jednoznačný kódový klíč (pravidla pro hodnocení všech možných odpovědí), lze mluvit o objektivním hodnocení. To je případ úloh uzavřených a do určité míry i úloh otevřených se stručnou odpovědí. Úlohy se širokou odpovědí je prakticky nemožné objektivně hodnotit, i když se i zde vyvíjí nástroje, s kterými by se hodnocení široce otevřených úloh objektivitou alespoň blížilo hodnocení podle stanoveného kódového klíče. Nejzávažnějšími problémy v této oblasti jsou: přesné formulování kritérií hodnocení, ověřování vytyčených kritérií, proškolení hodnotitelů, kontrola jejich práce apod. V některých případech se jedna práce dává k ohodnocení více hodnotitelům a porovnává se jejich shoda. Ač se jedná o proces velmi zdoluhavý a finančně i organizačně náročný, úlohy se širokou odpovědí mají v testování své nezastupitelné místo.

Otevřené úlohy se širokou odpovědí lze doporučit zejména pro testování komplexních dovedností, jsou vhodné pro zkoušení vyšších úrovní osvojení učiva (např. řešení problémových situací, třídění poznatků, analýza, zhodnocení, obhajoba stanoviska apod.). Jejich velkou předností je schopnost testovat komplexně úroveň písemného projevu žáka. Nezanedbatelnou výhodou je i poměrně vysoká „odolnost“ vůči opisování. Nevýhody pak představuje především obtížné hodnocení (zejména čtení rukopisu), malá objektivita hodnocení (dokonce jeden a týž hodnotitel může stejnou práci ohodnotit různě po určitém časovém odstupu) a velká časová náročnost, kdy během jednoho testování lze použít pouze malé množství široce otevřených úloh a s dlouhým psaním výrazně vzrůstá únava žáků. Nesmíme zapomínat, že tento typ úloh více znevýhodňuje žáky, kteří mají v důsledku zdravotního postižení problémy s grafomotorikou, se souvislým písemným vyjadřováním nebo žáky, kteří trpí specifickými poruchami učení (dyslexie, dysgrafie apod.).

V otevřených úlohách se širokou odpovědí se od žáka očekává souvislejší písemný projev (např. pojednání na téma: *Význam díla K. H. Máchy pro českou poezii*), vyřešení zadaného problému (*Navrhněte postup, kterým je možno určit hustotu neznámé kapaliny*), popis konkrétního procesu (*Popište činnost čtyřdobého zážehového motoru*) apod. Očekávaný rozsah odpovědi je dobré předem vymezit rozsahem volného místa pro odpověď nebo přímo určeným počtem slov v jistém intervalu.

DOPORUČENÍ PRO TVORBU OTEVŘENÝCH ÚLOH SE ŠIROKOU ODPOVĚDÍ

- Promyslete si, jaký typ rozumové operace po žákovi požadujete. V instrukci buďte konkrétní (používejte slova jako „vysvětlete“, „zdůvodněte“, „popište“, „obhajte“, „porovnejte“, „zhodnoťte“, „ukážte na příkladu“, „odhadněte, co by se stalo, kdyby...“ apod.). Nezačínejte úlohu slovy „co“, „kdy“, „kdo“ (protože příliš svádějí k pouhému uvádění faktů) a vyhněte se též neurčitým výrazům jako „pokuste se“, „dokážete popsat?“ apod.

Špatně

Co vedlo k vypuknutí občanské války mezi Severem a Jihem?

Správně

Porovnejte rozdíly mezi Severem a Jihem při vypuknutí občanské války. Uveďte ekonomické podmínky, sociální podmínky a politické názory.

- Úlohu formulujte tak, aby bylo jednoznačně patrné, co se od žáka očekává, v jaké formě nebo stylistickém útvaru má být odpověď napsaná (popis pracovního postupu, úvaha, úřední dopis apod.), v jakém rozsahu, zda se hodnotí i pravopis a vymezte v případě potřeby i strukturu požadované odpovědi (co všechno má odpověď zahrnovat, resp. co zahrnovat nemá).

Správně

Popište postup při výrobě železa. Uveďte hlavní používané suroviny, nakreslete schéma vysoké pece a popište hlavní probíhající chemické reakce.

- Úlohy by neměly hodnotit postoje či názory zastávané žáky. Je oprávněné zjišťovat, do jaké míry žák rozumí argumentům pro tu či onu pozici, nikoli však „správnost“ jeho vlastního stanoviska.
- Používejte široce otevřené úlohy výhradně k testování takových specifických cílů, které nemohou být uspokojivě měřeny jiným typem úloh.

8. OBECNÁ DOPORUČENÍ PRO TVORBU TESTOVÝCH ÚLOH

- Při konstrukci testových úloh počítejte s dostatečnou časovou rezervou na revizi (obsahovou i jazykovou) a zapracování změn. Málokdy je úloha naprosto bezvadná hned na první pokus.
- Mějte na zřeteli, že nejprve je třeba znát co nejpřesnější vymezení rozsahu znalostí a dovedností, které chceme úlohou testovat, a teprve potom můžeme začít s tvořením úlohy. Pokud by tomu bylo naopak, může se stát, že úloha bude sice konstrukčně perfektní, ale mine se s cíli testování.
- Úlohy by neměly stavět na znalostech a dovednostech, které se ve škole neučí. Zároveň vždy pečlivě posuzujte, zda úloha pracuje s tím, co je v dané části učiva podstatné, a nezabývá se okrajovými jevy.
- Znění úlohy musí být jednoznačné, srozumitelné, gramaticky správné a pokud možno stručné.
- Vyhýbejte se úlohám kvízového charakteru. Úlohy nemusí být nutně zábavné, avšak je ku prospěchu, když jsou témata či výchozí texty úloh pro žáky zajímavé a staví na situacích, které žáci znají.
- Úlohy musí být na sobě nezávislé. Správná odpověď na jednu úlohu nesmí být podmínkou pro správné vyřešení úlohy další.

- Ujistěte se, že úloha má jedinou správnou a nezpochybnitelnou odpověď, na níž se shodnou i experti z daného oboru.
- Při psaní testových úloh se vyhýbejte reprodukování učebnic; necitujte přímo z učebnicových textů. Obvykle nezjišťujeme, nakolik se student naučil něco z učebnice nazpaměť. Přímé citace navíc bývají vytržené z kontextu, a proto nejednoznačné.
- Vyhýbejte se tzv. „chytákům“.

Příklad "chytáku"

Protože se chcete dobře vyspat, jdete si lehnout už ve 20:30 hod. a budík si nařídíte až na 9:15 hod. Kolik hodin budete spát?

Nebo z jiného soudku

Jak se jmenovala první manželka Rudolfa II.?

KOMENTÁŘ: *Rudolf II. byl svobodný.*

- Při využívání výchozího textu by všechny úlohy měly z tohoto textu vycházet. Výchozí texty nemají jen ilustrační charakter a neslouží pouze k „navození atmosféry“. Výchozí texty by měly být pokud možno stručné a srozumitelné (při testování běží čas a každý řádek navíc je zátěží a zdržením). Složitý text klade vyšší nároky na úroveň čtenářských dovedností žáka a výsledky tím mohou být ovlivněny.
- Nepoužívejte v úlohách dvojí zápor. Jedno negativní tvrzení použijte jen v odůvodněných případech a zápor graficky zvýrazněte (tučné písmo nebo podtržení).

Příklad zvýraznění záporu

Pan Novák si chce koupit pračku za 20 000 Kč. Má v hotovosti pouze 10 000 Kč. Pračku si hodlá koupit i přesto, že ji bude muset splácet. Pan Novák si nemůže pračku pořídit prostřednictvím:

- A) hypotečního úvěru
- B) kreditní karty
- C) leasingu
- D) spotřebitelského úvěru

- Věnujte pozornost genderové zatíženosti úloh. Genderová zatíženost představuje určitý rys úlohy, který zapříčiní, že jsou výsledky chlapců a dívek výrazně odlišné. Rozdíl může být způsoben tím, že úloha předpokládá zapojení takových znalostí a dovedností, které žáci získávají mimo školu a které jsou jednostranně přístupné pouze chlapcům či jenom dívkám. Např. v rámci testování z občanského a společenskovedního základu se ukázalo, že chlapci zpravidla lépe řeší úlohy se zaměřením na politické dění nebo techniku, dívky zas úlohy se zaměřením na rodinný život, psychiku jedince a sociální vztahy.
- Úloha by měla být neutrální též s ohledem na citlivá témata, která by se mohla některých žáků dotknout (zdravotní či sociální handicap, rasový původ, vyznání apod.). Bude-li např. výchozím textem báseň opěvující oči nebo zrak, může to u žáků s těžkým zrakovým postižením vzbudit silné emoce.
- Úlohy by měly být předem vyzkoušeny/ověřeny – např. pilotáží.
- Úloh vždy navrhujte o něco více, protože ve fázi odborné revize a následného ověřování se některé z nich mohou ukázat jako nevyhovující.

- Nestavte jeden typ úloh nad jiné, ale používejte typy úloh účelně podle druhu ověřované kognitivní operace.
- Pište jen několik úloh za den.
- Neopakujte bezmyšlenkovitě úlohy z cizích ani vlastních testů.
- Zkoušejte jednu úlohu vytvořit ve více variantách (např. jako uzavřenou dichotomickou, uzavřenou s výběrem odpovědi nebo otevřenou se stručnou odpovědí). Zkoumejte, který z typů nejlépe odpovídá vytčenému specifickému cíli.

Příklad variací jedné úlohy

Dichotomická úloha

Martin Luther King byl bojovník za lidská práva.

ANO NE

Úloha se stručnou odpovědí

Jak se jmenuje černošský bojovník za lidská práva, který byl v r. 1968 úkladně zavražděn?

Úloha s výběrem odpovědi

Který z následujících mužů byl bojovníkem za lidská práva?

- A) Jefferson Davis
- B) Martin Luther King
- C) John Quincy Adams
- D) John Wilkes Booth

9. HODNOCENÍ ODPOVĚDÍ

Při hodnocení žákovských odpovědí **lze užívat tzv. jednoduchého (binárního) skórování úloh**, kdy za správnou odpověď v každé úloze připisujeme vždy jen jeden bod. Ukazuje se, že tento způsob skórování u běžných testů s objektivně skórovatelnými úlohami plně vyhovuje. U svazků dichotomických úloh přidělujeme zpravidla body za celý svazek (viz kapitola 6.). Složitější a pracnější je tzv. **vážené skórování**, kdy přidělujeme různé počty bodů úlohám podle významu ověřovaného učiva. Toto skórování se dobře uplatní u úloh, které se výrazně liší časem, který je potřeba k jejich vyřešení (např. v matematice).

U **otevřených úloh** nastávají určité komplikace. Především – posouzení správnosti řešení a přidělení bodů i za částečně správné odpovědi musí provádět přesně instruovaný odborník, který zkoušenému učivu dokonale rozumí. I přesto je skórování otevřených úloh zatíženo určitou mírou subjektivity, kdy různí hodnotitelé mohou určitou odpověď hodnotit zcela rozdílně.

Příklad komplikací s hodnocením odpovědí

S kterým státem sousedí na jihu Slovenská republika?

KOMENTÁŘ: *Zdánlivě jde o jasnou otázku s jednoznačnou správnou odpovědí. Jak budou ale jednotliví hodnotitelé posuzovat například tyto žákovské odpovědi:*

„S maďarskou republikou.“

„S maďarským státem.“

„S Maďary.“

První odpověď obsahuje pouze pravopisnou chybu. Jak jsme zmínili výše, je třeba v pokynech popsat, jak k pravopisným chybám při skórování přihlížet. Druhou odpověď by někdo mohl posoudit jako správnou a někdo jako nesprávnou, čímž by vznikla nejednotnost v hodnocení. Podobně je tomu i s třetí odpovědí. Někdo by mohl usoudit, že žák zná to podstatné, co se měl naučit, jiní budou namítat, že nerozlišuje mezi národem a státem, což je závažný nedostatek.

Pokud klíč k hodnocení úlohy obsahuje veškeré zmíněné možnosti odpovědí a k nim závazný způsob hodnocení, je vše v pořádku. Kde ale vzít jistotu, že jsme dokázali předvídat všechny možné odpovědi a že nás nějaký žák nezaskočí svou „velmi originální“ odpovědí? Praxe ukazuje, že ani zkušení autoři testů nejsou vždy schopni předvídat všechny možnosti.

Tento problém lze do určité míry řešit:

- precizní formulací úlohy;
- vhodným typem bodování (správně–nesprávně versus vícestupňové bodování za částečně správnou odpověď);
- co nejpodrobnějším kódovým klíčem pro hodnocení;
- pečlivým výběrem a důkladným proškolením hodnotitelů;
- možností rychlé komunikace se všemi hodnotiteli (případný dodatek ke klíči se rozšíří přes internet všem hodnotitelům);
- používáním metody nezávislého hodnocení jedné úlohy dvěma (nebo více) hodnotiteli.

Kódový klíč by měl obsahovat:

- charakteristiku a příklad správné odpovědi, která je od žáků očekávána;
- charakteristiku a příklady odlišně formulovaných odpovědí, které jsou také zcela správné;

- charakteristiku a příklady částečně správných odpovědí (pokud je tento způsob hodnocení použit);
- charakteristiku a příklady nesprávných odpovědí (nejčastější či typické nesprávné odpovědi).

Příklad úlohy a příslušného kódového klíče

Úloha 13

Filip, narozen 2. 5. 1989, který 30. 6. 2004 dokončí 9. ročník základní školy, chce nastoupit do pracovního poměru první možný den, kdy to povolují zákony, jako pomocný dělník firmy Výkopy, s. r. o.

Napište datum (den, měsíc, rok), od kterého smí Filip začít pracovat pro firmu Výkopy: _____ .

KÓDOVÝ KLÍČ k úloze 13

	KÓD	ZNĚNÍ ODPOVĚDI
správná odpověď	11	1. 7. 2004 nebo 1. července 2004
chybná odpověď	51	2. 5. 2004
chybná odpověď	52	3. 5. 2004
chybná odpověď	53	2. 5. 2007
chybná odpověď	50	jakákoliv jiná odpověď, např. neúplné datum
chybějící odpověď	99	neuvezená odpověď

Pro hodnocení otevřených úloh se širokou odpovědí jsou nutná určitá kritéria hodnocení, která popisují způsob hodnocení žákovy odpovědi. Dodržení těchto kritérií částečně koriguje subjektivitu hodnocení.

Příklad vymezení kritérií hodnocení

Kritéria pro hodnocení strukturované písemné práce z českého jazyka

Strukturovaná písemná práce je tvořena několika otevřenými testovými úlohami, doba jejího trvání je 60 minut. V průběhu psaní mohou žáci používat Pravidla českého pravopisu. Vyhodnocení strukturované písemné práce probíhá podle následujících kritérií:

- **dodržení požadované formy odpovědi** (respektování zadání) – funkční styl, slohový postup, slohový útvar, ohled na slohotvorné činitele;
- **věcná správnost odpovědi** (bez obsahových nesmyslů, jednotlivé části si vzájemně neodporují /pokud ano, je to vždy vysvětleno/), **výstižnost, srozumitelnost a relevantnost textu** (text neobsahuje fráze, výplňkové výrazy, „bezobsažné pasáže“...);
- **volba jazykových prostředků, vytříbenost odpovědi** (posouzení volby jazykových prostředků – vzhledem k účelu a obsahu textu – v rovině hláskoslovné, morfologické, lexikálně-sémantické, syntaktické a textologické – prostředky textového navazování, aktuální členění výpovědi, slovosled...) **a její strukturovanost** (logická výstavba textu);
- **jazyková správnost odpovědi** (uplatnění znalosti pravidel českého pravopisu, zásad českého tvarosloví a syntaktických a slohotvorných principů češtiny).

Odpovědi žáků jsou hodnoceny v jednotlivých kritériích na škále:

- 1 – bez nedostatků a chyb
- 2 – s drobnými nedostatky a bez chyb
- 3 – se závažnějšími nedostatky a s méně podstatnými chybami
- 4 – s vážnými nedostatky a občasnými podstatnými chybami
- 5 – s nedostatky a chybami zásadního charakteru

Příklad použití kritérií hodnocení pro konkrétní úlohu

Jednotlivé široce otevřené úlohy budou hodnoceny z několika dílčích hledisek na škále 0–3 body:

Možný příklad záznamu hodnocení

oblasti hodnocení

	0	1	2	3
1. požadovaná forma odpovědi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. věcná správnost	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. výstižnost, relevantnost a originalita zpracování textu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. jazykové prostředky	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. pravopis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. tvarosloví	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. větná a textová syntax	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

U každé úlohy se budou posuzovat ty oblasti, které zahrnují dovednosti a znalosti, na jejichž ověření se úloha zaměřuje.

Zadání úlohy:

VÝCHOZÍ TEXT K ÚLOZE

Rychlý hokej z první třetiny se ve druhé třetině změnil v urputnou bitvu, množily se souboje a rostla nervozita domácích. Hráčů, kteří nemohli nalézt cestu k překonání Fraňka a defenzívy Slavie i diváků, kteří zasy-pávali led odpadky na znamení nevole s výroky rozhodčích.

(Hospodářské noviny 9. 4. 2004)

Úloha

Upravte výchozí text tak, aby byly odstraněny chyby a nedostatky v jeho výstavbě:

Text:

Tato úloha je zaměřena především na ověření dovednosti posoudit úryvek textu jako celek a upravit jeho formulační chyby a nedostatky. Hodnotit se tedy u odpovědi bude jednak to, zda žák dodržel požadovanou formu odpovědi a zda nezkreslil původní význam textu, a jednak zda dokáže k vyjádření požadované odpovědi volit vhodné a bezchybné jazykové prostředky.

PŘÍKLAD ODPOVĚDI 1:

Nudný hokej z první třetiny se ve druhé třetině změnil v urputnou bitvu. Množily se souboje a rostla nervozita domácích hráčů, kteří nemohli nalézt cestu k překonání Fraňka a defenzívy Slavie. Ke klidu hráčů nepomohli ani diváci, jenž zasypávali led odpadky na znamení nevole s výroky rozhodčích.

POSOUZENÍ ODPOVĚDI 1:

Z odpovědi je patrné, že žák pochopil zadání úlohy a rozpoznal ve výchozím textu problematiku, která je úlohou ověřována. Dokáže uplatnit znalosti z textologie na konkrétním příkladu, přičemž volí vhodné jazykové prostředky a neporušuje pravopis. Zbytečnou změnou prvního slova textu žák mírně obměňuje význam věty, resp. celého textu. Jednou žák použil nesprávný tvar vztažného zájmena *jenž*, což by při hodnocení mělo být v tabulce zachyceno, např.

oblasti hodnocení

0	1	2	3
---	---	---	---

7. větná a textová syntax

		X	
--	--	---	--

Obecně ale žák prokázal, že dokáže posuzovat úryvek textu jako celek a upravit nepřesné a chybné formulace tak, aby výrazně nezměnil význam textu. Žák dodržel požadovanou formu odpovědi, tzn.

oblasti hodnocení

0	1	2	3
---	---	---	---

1. požadovaná forma odpovědi

			X
--	--	--	---

PŘÍKLAD ODPOVĚDI 2:

Z první třetiny zápasu, kde se hrál velmi útočný hokej, se ve druhé třetině stala urputná bitva. Velká nervozita domácích vedla ke častým soubojům mezi hráči. Obránce Franěk se stal pro domácí hráče nepřekonatelným. Slávistická tvrdá defenzivní hra vzbuzovala u diváků velkou podporu, a tak když se rozhodčí rozhodli dát trest „tvrdým“ hráčům, diváci tento rozsudek vypískali a odměnili odpadky, které házeli na led.

POSOUZENÍ ODPOVĚDI 2:

Odpověď ukazuje, že i tento žák pochopil zadání úlohy, řeší ji ovšem komplikovaněji než žák předchozí. O nápravu formulačních nepřesností a chyb se pokouší rozsáhlejšími úpravami a formulacemi vlastními, čímž ale dochází k zásahům do přesného významu textu. Sám se dopouští formulačních neobratností nebo používá nevhodné lexikální prostředky, což by se do hodnocení určitě promítlo, např.

oblasti hodnocení

0	1	2	3
---	---	---	---

4. jazykové prostředky

	X		
--	---	--	--

I tento žák prokázal dovednost posuzovat úryvek textu jako celek, úpravy formulací však vykazují vyjadřovací neobratnost a význam textu byl v některých částech pozměněn. Tento žák dodržel požadovanou formu odpovědi, získal by ale méně bodů než žák předchozí.

10. OVĚŘOVÁNÍ KVALITY TESTOVÝCH ÚLOH

Před zařazením do ostrých testů musí každá úloha projít odborným posouzením, zda je obsahově, konstrukčně a jazykově bezchybná. Následně je ověřena na vybraném vzorku žáků v rámci tzv. pilotáží. **Pilotáž** je výběrové šetření na reprezentativním vzorku minimálně 200 žáků. Na základě výsledků řešení úloh v tomto souboru jsou odhadovány psychometrické⁸ charakteristiky úlohy. K odhadu jsou využívány standardní matematicko-statistické metody.

O procesu ověřování testových úloh a celých testů pojednávala kapitola 3. Na základě výsledků z pilotáží je pro každou úlohu vytvořena tzv. položková analýza, což je v podstatě rozbor žákovských odpovědí, který má za cíl:

- zjistit, zda úloha odpovídá probíranému učivu na jednotlivých typech škol;
- analyzovat pravděpodobné důvody zjištěné obtížnosti dané úlohy a definovat faktory, které její řešení ovlivňují;
- odhalit problematické úlohy a formulovat možné příčiny zjištěných problémů jako náměty pro další úpravy nebo tvorbu nových úloh.

Položková analýza ukáže, jak danou úlohu zvládali žáci z různých typů škol, jaké alternativy žáci vybírali, jak byla úloha obtížná a co případně mohlo činit žákům nečekané problémy. U testových úloh sledujeme především dvě základní charakteristiky – obtížnost a citlivost (diskriminační hodnotu).

⁸ Název „psychometrické“ je zde převzatý z oblasti psychologie. Na počátek 20. století jsou datovány první inteligenční testy včetně statistických parametrů vypovídajících o jejich kvalitě a kvalitě testových úloh. Uvnitř psychologie se začala rozvíjet samostatná disciplína – psychometrie. Charakter parametrů je v oblasti pedagogiky tentýž, proto je i zde použit převzatý termín „psychometrické“. U některých autorů se však můžeme setkat např. i s pojmem „edukometrické“ odvozeným od slova edukace – vzdělávání a výchova.

OBTÍŽNOST

Obtížnost úlohy je odvozována z počtu žáků, kteří úlohu úspěšně vyřešili. Obtížnost se vyjadřuje indexem obtížnosti P vycházejícím z procentuálního zastoupení žáků, kteří danou úlohu zodpověděli správně. Počítá se jako

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

kde P je **index obtížnosti**, n_s je počet žáků ve skupině, kteří na danou úlohu odpověděli správně, a n je celkový počet žáků.

Někdy se místo indexu obtížnosti používá hodnota obtížnosti Q . Ta je naopak odvozena od počtu žáků, kteří danou úlohu vyřešili chybně.

$$Q = 100 \cdot \frac{n_{ch}}{n} ,$$

kde Q je **hodnota obtížnosti**, n_{ch} je počet žáků ve skupině, kteří na danou úlohu odpověděli chybně, a n je celkový počet žáků.

Mezi indexem obtížnosti a hodnotou obtížnosti platí vztah

$$P + Q = 100, \text{ neboť } n_s + n_{ch} = n.$$

Čím je úloha obtížnější, tím je hodnota obtížnosti větší a index obtížnosti menší. Index obtížnosti i hodnota obtížnosti mohou být nejméně 0 a nejvíce 100. Pokud je například hodnota obtížnosti 30, znamená to, že ze všech získaných odpovědí bylo 30 % chybných a tedy 70 % správných. Za velmi obtížné lze považovat úlohy s hodnotou obtížnosti vyšší než 80. Naopak úlohy, jejichž hodnota obtížnosti je nižší než 20, jsou považovány za velmi snadné.

To, co zde bylo uvedeno, platí pro úlohy, které mají hodnocení pouze 0 (úloha řešena chybně) a 1 (úloha řešena správně). V případě, že hodnocení jedné úlohy může nabývat více hodnot, např. 0, 1, ..., MAX, lze výpočet indexu obtížnosti upravit.

$$P = 100 \cdot \frac{\bar{b}}{MAX},$$

kde \bar{b} je aritmetický průměr bodových hodnocení všech žáků v dané úloze a MAX je maximální počet bodů v úloze, který může získat jeden žák.

Hodnota obtížnosti Q se pak určí podle vztahu

$$Q = 100 - P.$$

V tomto případě nelze index obtížnosti či hodnotu obtížnosti interpretovat jako procento žáků, kteří danou úlohu řešili správně či chybně, ale jako „míru správnosti“ či „míru chybovosti“ při řešení úlohy.

Doposud jsme se dívali na obtížnost jako vlastnost úlohy. Snažili jsme se tuto obtížnost odhadnout na základě pilotáže na reprezentativním vzorku žáků. Přitom jsme si tedy vědomi toho, že procento žáků, kteří danou úlohu vyřeší správně, je dáno nejen „náročností“ (obtížností) úlohy, ale i „chytrostí“ (znalostmi a dovednostmi) žáků, kterým byla úloha předložena k řešení. Někdy však nemusí být naším cílem odhadnout obtížnost úlohy, ale pouhé porovnání obtížností jednotlivých úloh v daném testu nebo souboru testových úloh. Pracujeme zde tedy s nereprezentativním vzorkem a na výsledky se díváme jako více podmíněné znalostmi a dovednostmi v námi sledovaném vzorku. V takovém případě se počítá tzv. úspěšnost v úloze.

Úspěšnost v úloze se značí P a její výpočet je totožný s výpočtem indexu obtížnosti. Proto se oba parametry označují shodně P . Další interpretace úspěšnosti se řídí shodnými principy jako interpretace obtížnosti. Jejich rozdíl je pouze v optice, kterou se na výsledky díváme. Zatímco obtížnost pou-

žíváme primárně k tomu, abychom charakterizovali úlohu, úspěšností charakterizujeme žákovskou populaci, která úlohu řešila.

Vědomí závislosti obtížnosti úlohy na tom, pro jako populaci je úloha určena, je promítnuto do novější matematicko-statistické psychometrické teorie, tzv. item-response-theory (IRT). Zde již není obtížnost vyjádřena jedním číslem, ale grafickou závislostí obtížnosti úlohy na celkových znalostech a dovednostech testovaných žáků.

Velmi obtížných ani velmi snadných úloh by v testu nemělo být příliš mnoho. Úlohu extrémně snadnou lze z psychologických důvodů použít jako úvodní úlohu v testu. Může to přispět k uklidnění žáků, zvýšení motivace a pocitu jistoty. Pro testy rozlišující se uvádí optimální hodnota indexu obtížnosti kolem 50.

CITLIVOST

Další důležitou charakteristikou testových úloh je jejich **citlivost**, označovaná též jako diskriminační schopnost nebo rozlišovací ostrost. Citlivost úlohy vypovídá o její schopnosti rozlišovat mezi úspěšnými a slabými žáky, přičemž za úspěšného žáka je v tomto kontextu považován žák s vysokým počtem dosažených bodů v testu, zatímco slabý žák je ten, který v testu dosáhl slabých výsledků – nízkého bodového skóre. Citlivost se vyjadřuje tzv. diskriminačním indexem, který má hodnoty v pásmu od -1 přes 0 až do 1. Vysoké hodnoty diskriminačního indexu má taková úloha, kterou správně vyřešili převážně dobří žáci (úspěšní v celém testu). Nulovou hodnotu diskriminačního indexu má naopak taková úloha, kterou řešili slabší žáci stejně dobře jako úspěšní, a zápornou hodnotu diskriminačního indexu má úloha, v níž byli naopak slabší žáci úspěšnější (takové by se však v testu neměly objevit). Dá se říci, že citlivost vyjadřuje, jak dalece úloha zvýhodňuje „lepší“ (v testu úspěšnější) žáky. U pilotáží je diskriminace konkrétní úlohy závislá na ostatních úlohách v pilotážním souboru testových úloh.

Velice nízkých hodnot diskriminačního indexu mohou nabývat úlohy s komplikovanou formulací zadání, u nichž hrají roli odlišné strategie řešení. „Lepší“ žáci se snaží k řešení komplikované úlohy dojít složitými úvahami, v nichž mohou chybovat, zatímco „horší“ žáci se příliš nezaobírají složitými úvahami a odpověď zkusí s větším či menším štěstím uhodnout. Taková úloha je vlastně chytákem a v hotovém testu by se neměla objevit.

Jedním ze způsobů stanovení diskriminačního indexu je **metoda ULI** (upper-lower-index). Výpočet podle metody ULI je poměrně jednoduchý. Rozdělíme žáky podle celkové úspěšnosti na dvě poloviny či menší, stejně velké skupiny na opačných stranách podle celkové úspěšnosti. Dále pracujeme s výsledky v dané úloze zvlášť u skupiny, která je lepší či nejlepší (L) a horší či nejhorší (H).

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

kde d je koeficient citlivosti ULI, n_L je počet žáků z lepší skupiny, kteří danou úlohu řešili správně, n_H je počet žáků ze skupiny horších, kteří danou úlohu řešili správně, a N je celkový počet žáků (zde $0,5N$ je tedy velikost jedné poloviny žáků, v případě rozdělení na 2 stejně velké menší skupiny bude ve jmenovateli vzorce místo $0,5N$ velikost této jedné menší skupiny).

V praxi se však můžeme setkat se dvěma problémy:

- 1) žáky nelze rozdělit podle celkových výsledků na 2 stejně velké skupiny (jsou zde žáci na hranici rozdělení, kteří mají shodné bodové hodnocení);
- 2) úloha nebude mít pouze hodnocení 0 bodů (úloha řešena chybně) a 1 bod (úloha řešena správně).

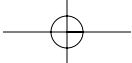
V takovém případě opět rozdělíme žáky na 2 skupiny podle celkových výsledků v testu, ale skupiny mohou být pouze přibližně stejně velké. Rozdělení učiníme tak, aby hranice mezi skupinami byla logickou hranicí mezi dvěma hodnotami celkového bodového hodnocení v testu. Dále spočítáme úspěšnost v dané úloze pro skupinu lepší i horší zvlášť. Zde můžeme použít i vzorce pro index úspěšnosti pro bodové hodnocení úlohy rozšířenější, než jen 0 a 1. Modifikovaný koeficient citlivosti ULI se pak spočítá podle vzorce

$$d = P_L - P_H,$$

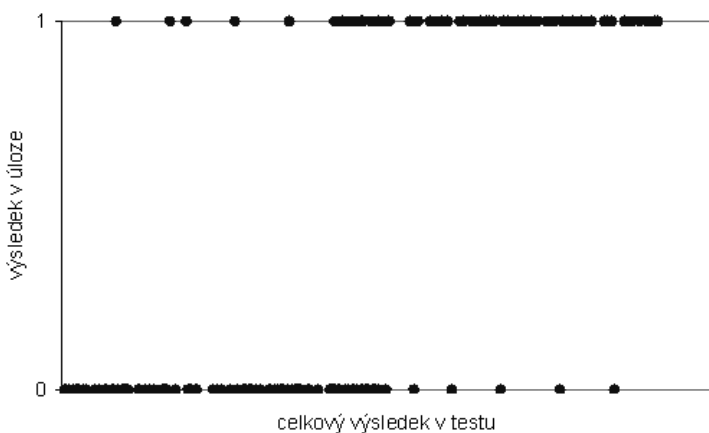
kde P_L je úspěšnost žáků v úloze z lepší skupiny a P_H úspěšnost žáků v úloze z horší skupiny.

Jiným v praxi používaným způsobem vyjádření diskriminace je stanovení tzv. **diskriminace RIT nebo diskriminace RIR**. Vypočítá se jako Pearsonův lineární korelační koeficient mezi skóre dosaženým v dané testové úloze a celkovým skóre (RIT) nebo mezi skóre dosaženým v dané testové úloze a celkovým skóre s vyloučením dané úlohy (RIR). Obzvláště koeficient RIT se v praxi používá velmi často vzhledem k obecnému rozšíření vzorce pro korelační koeficient ve většině běžných tabulkových editorů (včetně MS Excel) a nemusí být řešen problém rozdělení žáků na dvě stejně velké skupiny ani problém při hodnocení úlohy jinak než 0 nebo 1 bodem. Aby bylo možné označit úlohu za přiměřeně citlivou, měla by diskriminace RIT či RIR dosahovat nejméně hodnoty 0,25–0,30. Vzhledem k charakteru výpočtu bývá hodnota RIT o něco větší než hodnota RIR.

Diskriminace vypočtena indexem ULI i diskriminace RIR může mít své grafické znázornění.



Graf 1: Znáznornění výsledků v celém testu a v dané úloze. Jeden bod znázorňuje výsledek jednoho žáka.

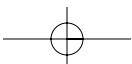
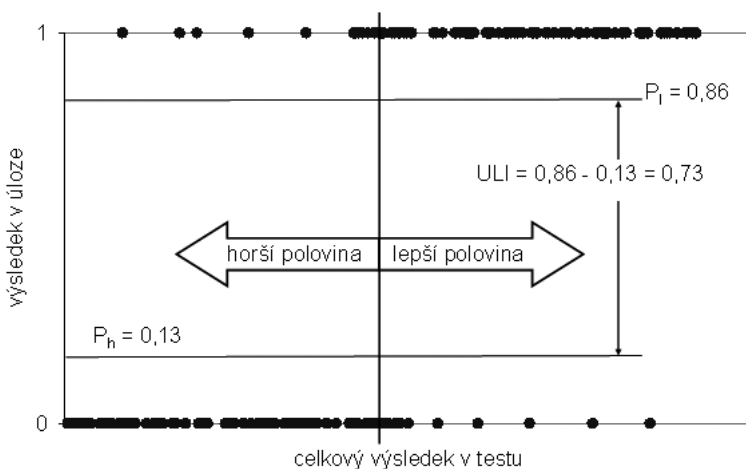


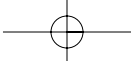
Na výsledcích, které jsou znázorněny na grafu vychází

$$ULI = 0,73$$

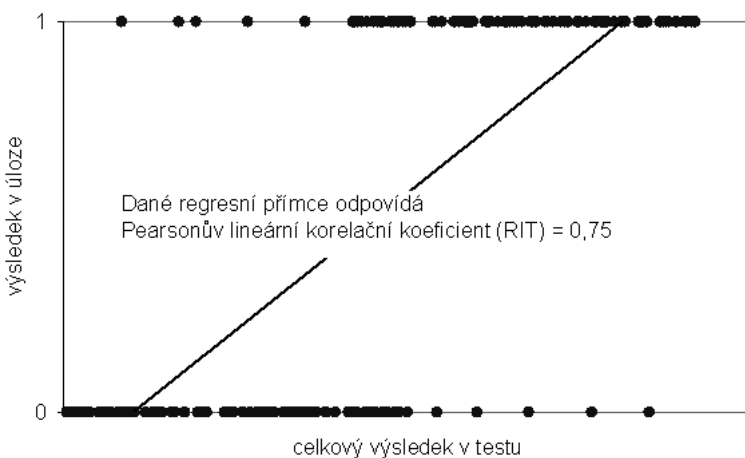
$$RIT = 0,75$$

Graf 2: Znáznornění výpočtu koeficientu citlivosti ULI.



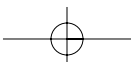
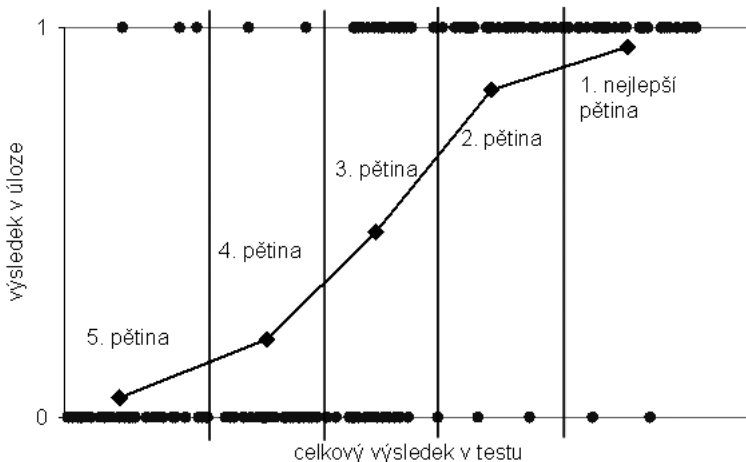


Graf 3: Výpočet koeficientu citlivosti RIT



Kromě těchto možností, kde je citlivost úlohy prezentována jedním inde-
xem, lze obtížnost prezentovat i graficky. Např. můžeme skupinu testova-
ných žáků rozdělit na 5 přibližně stejně velkých skupin seřazených podle
výsledků v celkové úspěšnosti v testu. V každé z těchto skupin se pak spo-
čítá průměrná úspěšnost v dané úloze a graficky se tato závislost znázorní
(viz Graf 4). Tento přístup je již blízky zmíněné item-response-theory.

Graf 4: Detailnější grafický náhled na obtížnost úlohy



POLOŽKOVÁ ANALÝZA

Položková analýza má především odhalit úlohy nevyhovující a upozornit na úlohy „podezřelé“, tj. takové, které mohou být zatíženy nějakými konstrukčními nedostatky. Pro ověřující testy se za podezřelé považují:

- úlohy velmi snadné (index obtížnosti nad 80) nebo úlohy velmi obtížné (pod 20). U těchto úloh vždy zkoumáme, není-li extrémní hodnota obtížnosti zapříčiněná nějakým nedostatkem v konstrukci úlohy.
- úlohy s diskriminačním indexem okolo 0. U průměrně obtížných úloh (index obtížnosti v rozmezí 30–70) by měla být hodnota diskriminačního indexu vyšší než 0,25. U úloh spíše snazších (index obtížnosti 70–80) a spíše obtížnějších (20–30) by měl diskriminační index dosahovat hodnoty minimálně 0,15.
- úlohy, u nichž více žáků neuvvedlo žádné odpovědi. Nezdopovězení úlohy může být kromě neznalosti způsobeno též neporozuměním zadání. Pokud se objevuje vyšší podíl chybějících odpovědí u úloh na konci testu, lze usuzovat na větší časovou náročnost testu. Žáci test prostě nestihli dokončit.
- uzavřené úlohy, u kterých nejsou dost atraktivní jednotlivé distraktory, tzn. některé nabídnuté alternativy odpovědí nikdo nezvolil. Takový distraktor neplní svou funkci a měl by být vyřazen či změněn. Při rozboru distraktorů rovněž posuzujeme jejich souvislost (korelaci) s celkovým výsledkem. Korelační koeficient mezi distraktorem a celkovým výsledkem v testu by měl být záporný (chybnou variantu volí celkově spíše slabší žáci).
- otevřené úlohy, u kterých je velký podíl nesprávných odpovědí. U každé takové úlohy zvažujeme možnost, zda žáci mohli úloze špatně porozumět např. kvůli nepřesně zformulovanému zadání.

Příklad interpretace výsledků položkové analýzy

číslo úlohy	index obtížnosti	diskriminační index	Volba jednotlivých alternativ v %				podíl vynechaných odpovědí
			A	B	C	D	
21	75	0,53	7	10	8	75*	0 %
22	26	0,13	23	27	26*	24	0 %
23	27	0,20	27*	0	65	8	0 %
24	11	-0,30	66	11*	10	9	4 %

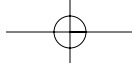
* správná řešení

Úloha 21

Na úlohu 21 odpovědělo správně celkem 75 % žáků. Distraktor A zvolilo 7 % žáků, distraktor B zvolilo 10 %, distraktor C zvolilo 8 % žáků. Úlohu 21 lze celkově považovat za kvalitní. Má dobrou hodnotu diskriminace (žáci celkově v testu úspěšní ji většinou řešili správně a žáci neúspěšní ji řešili spíše špatně), tudíž úloha dobře rozlišuje mezi dobrými a špatnými žáky. Všechny distraktory byly voleny obdobně často, což znamená, že je lze považovat za hodnověrné.

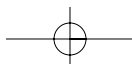
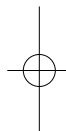
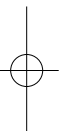
Úloha 22

Na úlohu 22 odpovědělo správně pouze 26 % žáků. Jednotlivé varianty volil přibližně stejný podíl žáků. Tato skutečnost oprávněně vede k domněnce, že většina žáků správné řešení prostě hádala. Úloha byla zřejmě příliš obtížná nebo nejednoznačná. Buďto žáci správnou odpověď neznali, a proto tipovali, nebo nedokázali od sebe odlišit málo jednoznačné alternativy.



Jak byste interpretovali výsledky úloh 23 a 24? Jsou tyto úlohy podle vašeho názoru bezproblémové? Splňují důležitá statistická kritéria? Případně jaké problémy u nich lze na základě položkové analýzy vysledovat?

Pracujeme-li s výsledky dostatečně početné skupiny žáků, můžeme při položkové analýze dále sledovat celou řadu statistických kritérií. Zajímá nás především úspěšnost žáků dělených podle typu školy (gymnázia, střední odborné školy, střední odborná učiliště) nebo též podle skupin oborů (školy umělecké, zdravotnické, hotelové apod.). Velmi užitečným údajem může být vzájemné porovnání úspěšnosti chlapců a dívek, dále můžeme sledovat úspěšnost žáků dělených do skupin podle poslední známky na vysvědčení (jak úlohu řešili jedničkáři, dvojkaři, trojkaři atd.), podle toho, kteří žáci se hlásí na VŠ, podle toho, zda mají nějaké zdravotní postižení (jak dopadli žáci se zrakovým postižením, žáci neslyšící atd.), nebo dokonce tato kritéria vzájemně kombinovat.



ZÁVĚR

Tato rukověť chce především „vést ruku“ začínajícím autorům úloh. Obrací pozornost k testové úloze jako základnímu prvku didaktických testů, jako k dílku stavebnice, na jehož kvalitě závisí úspěch celé stavby. Záměrně jsme se zde nevěnovali některým „pokročilejším“ tématům spojeným s testováním (sestavení celého testu, grafická úprava testového sešitu a záznamového archu, zadávání a vyhodnocování testu, matematicko-statistické metody pro analýzu výsledků apod.). Takové ambice tato brožura nemá, protože z hlediska autorů úloh se nejedná o podstatná témata. Pokud jste však po jejím přečtení dostali chuť do tvorby testových úloh, pak splnila svůj cíl. Jestli budete napříště k testovým úlohám (svým i cizím) o něco kritičtější, pak je to také v pořádku. Věříme, že se postupem času tento text zaplní rukopisnými poznámkami a na vašem psacím stole zaujme jeho místo vaše vlastní sbírka (sbírečka, štůsek, soubor či fascikl) výborných testových úloh.

Použitá a doporučená literatura

1. Alderson, J. Ch., Clapham, C., Wall, D.: Language Test Construction and Evaluation. (8. vydání) Cambridge 2004.
2. Burjan, V.: Tvorba a využívanie školských testov. in EXAM-info č. 1-7, Bratislava 1999-2005.
3. Byčkovský, P.: Základy měření výsledků výuky. Praha 1982.
4. Hendl J.: Přehled statistických metod zpracování dat. Praha 2004.
5. Chráska, M.: Didaktické testy. Brno 1999.
6. Hrabal, V., Lustigová, Z., Valentová, L.: Testy a testování ve škole. Praha 1992.
7. Katalogy požadavků k maturitní zkoušce. ÚIV – CERMAT Praha 2005.
8. Komenda, S., Mazuchová, J.: Tvorba a testování testu. Olomouc 1995.
9. Kubiszyn, T., Borich, G.: Educational Testing and Measurement. (6. vydání) New York 2000.
10. Mičienka, M., Moravcová-Smetáčková, I.: Příručka pro tvorbu testových úloh. Praha 2002.
11. Niemierko, B.: Pomiar wyników kształcenia. Warszawa 1999.
12. Říčan, P.: Úvod do psychometrie. Bratislava 1978.
13. Standardy pro pedagogické a psychologické testování. Praha 2001.
14. Van Krieken, R.: Writing and Administering Examinations. Arnhem 1995.
15. Zvára, K., Štěpán, J.: Pravděpodobnost a matematická statistika. Praha 2002.

Užitečné internetové odkazy

ACER – Australian Council for Educational Research (Austrálie)

<http://www.acer.edu.au/>

Organizace zabývající se pedagogickým výzkumem, tvorbou didaktických testů a problematikou pedagogických měření.

AERA – American Educational Research Association (USA)

<http://aera.net/>

Americká asociace pedagogického výzkumu věnuje hodně prostoru i problematice měření a evaluace vzdělávání.

Buros Center for Testing (USA)

<http://www.unl.edu/buros/>

Jedna z nejstarších institucí zabývající se problematikou měření a testování v pedagogice a psychologii.

Centrum pro zjišťování výsledků vzdělávání – CERMAT (ČR):

www.ceremat.cz

Státní organizace zabývající se evaluací výsledků vzdělávání, výzkumem a vývojem v této oblasti.

CITO – Národní institut pro vývoj testů (Holandsko)

<http://www.cito.nl/>

Největší evropská organizace specializovaná na tvorbu didaktických testů a problematiku pedagogických měření.

CKE – Centralna Komisja Ekzaminacyjna (PL)

<http://www.cke.edu.pl/>

Státní organizace, jejímž cílem je příprava a organizování systému pedagogické evaluace v Polsku. Spolupracuje s osmi regionálními útvary OKE (okręgowymi komisjami egzaminacyjnymi).

CRESST – National Center for Research on Evaluation Standards and Student Testing (USA)

<http://www.cse.ucla.edu/>

Velké vědecké pracoviště zabývající se výzkumem v oblasti pedagogických měření. Možnost bezplatně stáhnout množství výzkumných zpráv o různých aspektech testování.

ERIC – Clearinghouse on Assessment and Testing (USA)

<http://ericae.net/>

Na těchto stránkách lze najít rozsáhlý zdroj literatury o problematice testování.

ETS – Educational Testing Service (USA)

<http://www.ets.org/>

Významná organizace zabývající se tvorbou didaktických testů a problematikou pedagogických měření v USA.

EXAM testing, spol. s r. o. (SR)

<http://www.exam.sk/>

Slovenská soukromá firma zabývající se systematicky problematikou didaktických testů.

IEA – International Association for the Evaluation of Educational Achievement

<http://www.iea.nl/>

Mezinárodní instituce, která tradičně organizuje velké komparativní projekty v oblasti evaluace výsledků vzdělávání (např. TIMSS, SITES, PIRLS).

NCES – National Center for Education Statistics (USA)

<http://www.nces.ed.gov/>

Národní centrum pro statistiku vzdělávání, které organizuje projekt NAEP (National Assessment of Educational Progress) – periodické celonárodní testování amerických žáků ve 4., 8. a 12. ročníku.

NCME – National Council on Measurement in Education (USA)**<http://www.ncme.org/>**

Stavovská organizace sdružující odborníky z oblasti pedagogických měření. NCME vydává dva odborné časopisy věnované problematice testů (teoreticky laděný *Journal of Educational Measurement* a prakticky orientovaný *Educational Measurement: Issues and Practise*).

NFER – National Foundation for Educational Research (V. Británie)**<http://www.nfer.ac.uk/>**

Významná britská organizace věnující se pedagogickým výzkumům a tvorbě didaktických testů.

PISA – OECD Programme for International Student Assessment**<http://www.oecd.org/pages/>**

Organizace pověřená realizací periodických mezinárodních komparativních testování žáků v zemích OECD.

Practical Assessment, Research and Evaluation (USA)**<http://pareonline.net/>**

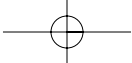
Elektronický časopis věnovaný problematice testování, hodnocení, měření výsledků vzdělávání. Prakticky orientované články.

SAT – Scholastic Aptitude Test (USA)**<http://www.collegeboard.com/>**

Zřejmě nejnámější didaktický test na světě. Každoročně ho absolvuje několik miliónů studentů.

SCIO (ČR)**<http://www.scio.cz/>**

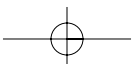
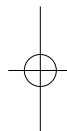
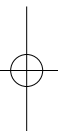
Soukromá společnost zabývající se problematikou didaktických testů.

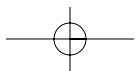
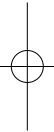
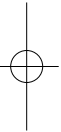
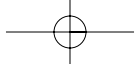
**ŠPÚ – Štátny pedagogický ústav (SR)****<http://www.statpedu.sk/>**

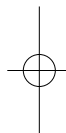
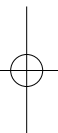
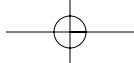
Státní organizace, která zabezpečuje účast SR v mezinárodním srovnávacím testování žáků (PISA, TIMSS, PIRLS...) a připravuje novou formu maturitních zkoušek na Slovensku.

ÚIV – Ústav pro informace ve vzdělávání (ČR)**<http://www.uiv.cz/>**

Příspěvková organizace, přímo řízená MŠMT ČR, zapojená do mezinárodních programů testování žáků (PISA).







Rukověť autora testových úloh

Zpracoval: Radek Schindler a kol.

Vydalo: Centrum pro zjišťování výsledků vzdělávání

Jeruzalémská 957/12

110 00 Praha 1

www.cermat.cz

Návrh obálky: Jan Jiskra

Sazba: TYPO+

Tisk: Nakladatelství Tauris, divize ÚIV

Praha 2006

Vydání první

ISBN 80-239-7111-5

